Assimilation variationnelle d'observations multi-échelles : Application à la fusion de données hétérogènes pour l'étude de la dynamique micro et macrophysique des systèmes précipitants

François Mercier

▶ **To cite this version:**

François Mercier. Assimilation variationnelle d'observations multi-échelles : Application à la fusion de données hétérogènes pour l'étude de la dynamique micro et macrophysique des systèmes précipitants. Climatologie. Université Paris Saclay (COmUE), 2016. Français. NNT : 2016SACLV046 . tel-01466267

HAL Id: tel-01466267
https://theses.hal.science/tel-01466267

Submitted on 13 Feb 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

NNT : 2016SACLV046

THESE DE DOCTORAT
DE
L'UNIVERSITE PARIS-SACLAY
PREPAREE A
L'UNIVERSITE VERSAILLES SAINT-QUENTIN EN YVELINES

ÉCOLE DOCTORALE N°579
Sciences mécaniques et énergétiques, matériaux et géosciences

Spécialité de doctorat : météorologie, océanographie, physique de l'environnement

Par

M. François Mercier

Assimilation variationnelle d'observations multi-échelles : Application à la fusion de données hétérogènes pour l'étude de la dynamique micro et macrophysique des systèmes précipitants.

Thèse présentée et soutenue à Guyancourt, le 05 juillet 2016 :

Composition du Jury :

Mme S. Thiria, Professeure, Université Versailles Saint-Quentin, Présidente
M. H. Andrieu, Directeur de Recherche, IFFSTAR, Rapporteur
M. R. Uijlenhoet, Professeur, Wageningen University, Rapporteur
M. E. Blayo, Professeur, INRIA / Université Joseph Fourier, Examinateur
M. P. Tabary, Ingénieur en chef des eaux, des ponts et des forêts, CNES, Examinateur
M. L. Barthès, Maître de conférences, Université Versailles Saint-Quentin, Directeur de thèse
M. A. Chazottes, Maître de conférences, Université Versailles Saint-Quentin, Co-encadrant de thèse

Résumé

L'étude des précipitations est généralement abordée d'un point de vue soit macrophysique soit microphysique. Dans le premier cas on se place à l'échelle des cellules de pluie, caractérisées par des quantités intégrées (taux de pluie) tandis que dans le second cas, la pluie étant caractérisée par ses distributions en tailles des gouttes (DSD – Drop Size Distribution), on travaille à l'échelle de la goutte et des phénomènes physiques qui l'affectent. Il existe une grande variété d'instruments permettant de caractériser les précipitations. Cependant, ils mesurent des grandeurs différentes à des échelles spatiotemporelles diverses. Cette thèse - en se basant sur l'assimilation de données – s'intéresse au développement d'algorithmes fusionnant des mesures hétérogènes en vue d'améliorer la restitution des précipitations à fine échelle et de faciliter ainsi l'étude de leur variabilité spatiotemporelle à échelles macro et microphysiques.

Dans un premier temps, nous nous intéressons à la restitution de cartes de pluie par assimilation de mesures d'atténuation d'ondes provenant de satellites géostationnaires TV-SAT en bande Ku ($10,7$-$12,7$ GHz). L'atténuation, due aux gouttes présentes sur la liaison, dépend du taux de pluie via une loi puissance. On dispose de séries temporelles d'atténuations provenant de 4 satellites permettant de couvrir au sol quelques dizaines de kilomètres carrés. On suppose que les champs de pluie sont transportés sans déformations par le vent (advection pure). On utilise ainsi un algorithme d'assimilation 4D-VAR pour reconstituer la carte de pluie à l'instant initial qui, propagée par le modèle, expliquera au mieux les observations. Cette méthode est appliquée à un évènement enregistré dans le sud de la France durant la campagne HyMeX. On restitue des cartes de précipitations sur une superficie de quelques dizaines de kilomètres carrés semblables à celles des radars de Météo France. Cette méthode pourrait être adaptée pour couvrir des zones où les radars sont absents (pays en développement, zones montagneuses), ou pour l'hydrologie urbaine et la gestion des crues éclairs.

L'étude précédente suppose des hypothèses fortes concernant la microphysique (DSD fixée). Sa compréhension et l'étude de sa variabilité sont pourtant d'une grande importance pour la paramétrisation de la pluie dans les modèles météorologiques ou pour l'estimation du taux précipitant à partir de la réflectivité radar (loi Z-R). Il convient donc d'étudier la DSD et sa variabilité. Dans cette optique, une nouvelle approche consistant à coupler des observations co-localisées (disdromètres et spectres Doppler de radars micro-ondes à visée verticale) dans un modèle de propagation via un algorithme d'assimilation 4D-VAR est présentée. Cette approche est dénommée VAMOS (Variational Assimilation of Multi-scale Observations for rain Studies). Il s'agit de restituer DSD et vents verticaux à la verticale des instruments. Le modèle de propagation représente la chute des gouttes modulée par le vent vertical, la turbulence et l'évaporation. L'impact du vent horizontal est étudié indépendamment de l'algorithme d'assimilation. En appliquant ce travail à 3 cas d'étude dans divers contextes instrumentaux et météorologiques, on montre que l'on restitue des profils de DSD et des vents physiquement cohérents. Ces restitutions sont validées grâce à une autre méthode utilisant 2 radars. On met aussi en évidence pour chaque cas l'importance des différents phénomènes physiques et leur signature sur les cartes radar. Ce travail pourrait être étendu pour prendre en compte, contrôler et paramétrer d'autres phénomènes physiques ou instrumentaux.

Abstract

The study of rainfall can be viewed at macro or microscopic scales. In the first case, we work at the rain cell level. The rain cells are characterized by integrated parameters, such as the rainfall rate. In the second case, we work at the raindrop level and have to study the physical phenomena which affect the drops. Rain is then characterized by its Drop Size Distribution – DSD. There are many different instruments able to monitor the rainfall, but they act at different scales and can measure the rainfall directly (f.i. rain gages) or indirectly (f.i. radars). In this thesis, we develop data assimilation algorithms for fusing heterogeneous measurements and for studying rainfall and its spatio-temporal variability at the macro and microphysical levels.

First, we rebuild rainfall maps with an assimilation algorithm using measurements of attenuations of satellites microwaves. These microwaves are in the Ku-band and come from TV-sats. They are attenuated by the raindrops, and this attenuation can be linked to the rain rate by a power law. In this study, we use one receiver able to measure the broadcasts from 4 satellites. Moreover, we assume that the rain cells are advected by the wind field without any distortion. Consequently, they only depend on the rain field at an initial time. Then, we use a 4D-VAR assimilation algorithm to retrieve this initial rain map which, once propagated by the model, appears able to reproduce the measured attenuations. This method is applied to a rain event located in the south of France. We retrieve rain maps covering a few dozens of square kilometers with features comparable with Météo France radar ones. This method could be adapted to monitor rainfall in areas not covered by operational radar networks (developing countries, mountainous areas), or applied to urban hydrology problems and flash flood monitoring.

So far, microphysical phenomena has been strongly constrained (we supposed a fixed DSD). But their understanding and spatio-temporal variability are crucial for rainfall parameterization in models, for the reflectivity/rainfall relations accuracy, or for calibration and intercomparison of heterogeneous instruments. And the study of rain microphysics has to go through the study of the DSD and its variability. For this purpose, we present a novel approach which consists in coupling colocalized observations (disdrometer data and Doppler spectra from vertically pointing microwave radars) in numerical models using 4D-VAR data assimilation. This approach is called VAMOS (Variational Assimilation of Multi-scale Observations for rain Studies). With this method, we retrieve DSD and vertical winds above the instruments. The numerical model simulates the drops fall modulated by vertical wind, turbulence and evaporation. Collisions between drops are not taken into account but could be added. The influence of the horizontal wind is studied apart from the assimilation algorithm. When applied to 3 case studies with different instruments and meteorological conditions, we show that we retrieve DSD profiles and winds physically coherent (thanks to the model) and coherent with the different observations, providing robust relationships. These retrievals are validated with another method using 2 radars. We also show for each event the relative importance of the different phenomena (evaporation, horizontal wind, turbulence) and their signatures on the radar maps. This work could be extended to take into account, control and parameterize other physical (evaporation, collisions) or instrumental (radar calibration) factors.

Remerciements

Allez, quelques remerciements pour commencer sans (presque) parler de science. Les lecteurs qui ne me connaîtraient pas – ils sont nombreux j'espère (plusieurs ? au moins un ?) – peuvent aller directement à la page suivante, ça ne va pas durer si longtemps.

Je commence par mes encadrants. Laurent, sa gentillesse, sa disponibilité, la confiance qu'il m'a toujours accordée, et tout ce qu'il m'a permis de faire durant cette thèse : le séjour en Australie, les collaborations qu'il m'a aidé à mettre en place. Aymeric, qui a été en permanence à l'écoute et prêt à m'aider et me soutenir (on peut noter aussi nos longs débats – quelque peu stériles quand on n'y connaissait rien ni l'un ni l'autre – sur les performances de la modélisation atmosphérique ou l'avenir de l'assimilation de données ou de la modélisation stochastique). Et Cécile, qui malgré ses autres occupations a toujours participé aux réunions, a relu mes papiers et m'a aidé pour l'enseignement et l'après-thèse.

Mon jury ensuite. Sylvie Thiria, que j'avais rencontrée dès mon arrivée au LATMOS et qui m'a toujours encouragé. Mes rapporteurs : Remko Uijlenhoet, qui avait déjà eu la gentillesse de participer à mon second comité de thèse, et Hervé Andrieu. Et mes examinateurs, Pierre Tabary, qui m'a ensuite accompagné de ses conseils dans ma recherche de post-doc, et Eric Blayo. D'autres gens m'ont aidé grandement durant cette thèse. Je remercie tout particulièrement Alain Protat, du Bureau of Meteorology de Melbourne, qui m'a accueilli là-bas 2 mois durant l'été 2015 (l'hiver chez lui...) et s'est toujours montré disponible pour me suivre durant cette période. Également Julien Delanoë, qui m'a permis de partir en Australie, et Julien Brajard, du LOCEAN à Jussieu, qui m'a formé à l'assimilation de données et à YAO, a participé à mon premier comité de thèse et m'a récemment conseillé pour des entretiens de post-doc.

Je remercie aussi Alessandro Battaglia, de l'université de Leicester, pour avoir participé à mes comités de thèse et m'avoir conseillé sur sa spécialité, les spectres Doppler. Et Jean-Marie Donier, de Météo France, avec qui j'ai pu échanger sur le Micro-Rain Radar. Et aussi les enseignants chercheurs avec qui j'ai travaillé dans le cadre de mon monitorat à l'UVSQ, en particulier Ronan Modolo, du LATMOS, et Bruno Lansard, du LSCE (Cécile et Aymeric aussi mais ils ont déjà été cités !).

Une pensée particulière bien entendu pour les 3 camarades doctorants qui auront commencé et fini en même temps que moi dans l'équipe : Meriem, Ruben, et Thomas. Les plus anciens, comme Jean-François et Nawal, et les doctorants et post-doctorants arrivés un peu plus tard, Nadia, Djallel, Sabrina, Quitterie, Kouadio, Renaud. On s'est bien amusé et c'est triste de partir. Et bien sûr les permanents de l'équipe SPACE, je ne les cite pas mais la bonne ambiance du quotidien (et du café du matin) leur doivent bien sûr beaucoup. Et aussi

un moment de recueillement pour les pauses tisane (mais pas pour les poissons).

Hors travail, je pense à tout ce que j'ai fait à Paris pendant ces 3 ans. A mes colocs successifs avec qui j'ai vécu durant la thèse, Evier, Démission, Cosette, Jean-Foutre, Hébinvoilà (si si, ils s'appellent comme ça). Et aux Fanfares Brass de Pneu et Zebra Rayures, longue vie à elles !

Et pour finir tous mes remerciements et mon amour à mes parents, Gérard et Maryse. Et à Sarah.

Bonne lecture à ceux ayant prévu d'aller au-delà de cette page !

Table des matières

Remerciements **I**

Nomenclature **VIII**

Introduction **1**

1 Contexte . . . 1
2 Objectifs de la thèse . . . 3
3 Plan . . . 4

Partie I Définir et mesurer la pluie **5**

1 Définir la pluie : de la goutte au taux précipitant . . . 6
1.1 Distribution volumique de gouttes de pluie : point de vue aléatoire . . . 6
1.2 Le taux de pluie . . . 7
1.3 Moments de la distribution . . . 8
2 Quels instruments pour mesurer la pluie ? . . . 10
2.1 Pluviomètres . . . 10
2.2 Disdromètres . . . 11
2.3 Radars . . . 14
2.4 Capteurs d'atténuations . . . 31

Partie II Dynamique des précipitations : phénomènes et modélisation **34**

1 À l'échelle de la cellule de pluie . . . 35
2 À l'échelle microphysique : évolution spatiotemporelle de la DSD . . . 38
2.1 Introduction / Revue des phénomènes . . . 38
2.2 EDP complète . . . 38
2.3 Transport . . . 39
2.4 Évaporation . . . 40
2.5 Collisions et désintégration de gouttes . . . 45
3 Relations vitesse-diamètre pour la pluie . . . 52
3.1 Introduction . . . 52
3.2 Paramétrisation utilisée dans cette thèse . . . 52
3.3 Écarts à la vitesse terminale de chute . . . 53
4 Discrétisation des EDP . . . 55
4.1 Introduction . . . 55

4.2 Schéma upstream . . . 56
4.3 Schéma de Smolarkiewicz . . . 57

Partie III Assimilation de données et mise en oeuvre informatique 59

1 Généralités sur l'assimilation de données . . . 61
1.1 Quelques notations . . . 61
1.2 Approche bayésienne, filtres particulaires . . . 61
1.3 Assimilation sans dépendance temporelle . . . 62
1.4 Ajout d'une dépendance temporelle : le filtre de Kalman . . . 64
1.5 4D-VAR . . . 66
2 Assimilation 4D-VAR . . . 67
3 Mise en oeuvre logicielle : YAO . . . 71
3.1 Un exemple simple : système proies / prédateurs . . . 71
3.2 Généralisation : le logiciel YAO . . . 74
4 Extension . . . 80
4.1 Contrôle de conditions initiales et aux limites . . . 80
4.2 Cas sans ébauche / Filtrage . . . 83

Partie IV Restitution de cartes de précipitations par assimilation d'atténuations TV-SAT 85

1 Présentation . . . 86
1.1 Objectifs généraux . . . 86
1.2 Autres études . . . 86
1.3 Spécificités de ce travail . . . 87
1.4 Plan . . . 87
2 Dispositif expérimental . . . 89
2.1 Capteur Ku . . . 89
2.2 Instruments pour validation . . . 91
3 Équations du 4D-VAR pour ce problème . . . 92
3.1 Rappel modèle direct . . . 93
3.2 Opérateur d'observation . . . 94
3.3 Filtrage . . . 95
4 Détermination de la vitesse d'advection . . . 96
5 Validation de l'approche sur données simulées . . . 97
5.1 Méthodes / Définitions . . . 97
5.2 Influence du nombre de capteurs et de la géométrie du problème : 15 simulations . . . 101
6 Résultats sur données réelles . . . 107
6.1 Description des évènements . . . 107
6.2 Vitesse d'advection . . . 108
6.3 Résultats / Efficience de l'algorithme . . . 110

6.4 Résultats / Restitutions intégrées : comparaison au radar 112
6.5 Résultats / Restitutions ponctuelles : comparaison aux pluviomètres . . . 118
7 Conclusions et perspectives 122

Partie V Restitution et étude des profils verticaux de DSD et de vents verticaux **124**
1 Présentation 125
1.1 Objectifs généraux / Approche VAMOS 125
1.2 Autres études sur l'inversion des spectres Doppler 127
1.3 Plan 129
2 Description de l'algorithme de restitution VAMOS complet 130
2.1 Modèle de propagation 130
2.2 Paramétrisation de la DSD en haut de la colonne 133
2.3 Opérateurs d'observation 134
2.4 Filtrage / Pénalisation 139
2.5 Conclusion 142
3 Données simulées 143
3.1 Introduction 143
3.2 Réglage relatif des coûts 144
3.3 Performances sur observations réalistes 147
4 Influence du vent horizontal 159
4.1 Présentation 159
4.2 Etude théorique : signature du vent horizontal sur les profils radar 159
4.3 Données simulées / Erreur modèle due au vent horizontal 166
4.4 Conclusions 178
5 Première cas d'étude : HyMex 179
5.1 Présentation et enjeux 179
5.2 Instruments disponibles / Présentation de l'évènement 179
5.3 Résultats : reproduction des observations 184
5.4 Résultats : vent vertical 187
5.5 Résultats : DSD / comparaison aux produits MRR 188
5.6 Conclusions et ouverture sur le deuxième cas d'étude 192
6 Deuxième cas d'étude : Finlande 194
6.1 Présentation et enjeux 194
6.2 Instruments disponibles / Présentation de l'évènement 194
6.3 Résultats 199
6.4 Conclusions, perspectives et ouverture sur le troisième cas d'étude 212
7 Troisième cas d'étude : Darwin 214
7.1 Présentation et enjeux 214
7.2 Instruments disponibles / Présentation de l'évènement 214
7.3 Résultats de l'assimilation (VAMOS) 223

7.4 Conclusions 243
8 Conclusions 244
8.1 Conclusions 244
8.2 Perspectives 245

Conclusion générale **247**

Bibliographie **251**

Table des figures **i**

Liste des tableaux **xviii**

Annexe A : Coalescence / break-up : paramétrisation **xxii**

Annexe B : Auto break-up **xxxix**

Annexe C : Discussion sur la paramétrisation de la vitesse terminale de chute des gouttes. **xli**
1 Un peu de physique xli
2 Relations empiriques xliv
3 Vitesse terminale de chute : dépendance à l'altitude xlv

Annexe D : Discrétisation des EDP **xlix**
1 Stabilité du schéma upstream xlix
2 Diffusion numérique li
3 Justification du schéma de Smolarkiewicz liii
4 Schémas pour l'équation d'advection 2D liii

Annexe E : Contrôle des paramètres d'un modèle dans un algorithme 4D-VAR. **lv**

Annexe F : Ecriture du terme de filtrage du chapitre IV sous forme de terme d'ébauche **lviii**

Annexe G : Détermination par triangulation de la vitesse d'advection des champs de pluie **lxii**
1 Théorie lxiii
2 Résultats lxxi

Annexe H : Évaluation rapide de l'impact de quelques erreurs modèle sur la restitution de cartes de précipitations **lxxviii**
1 Expérience jumelle parfaite lxxix
2 Erreur sur la longueur des liaisons lxxxi
3 Erreur sur la vitesse de déplacement lxxxii
4 Intérêt du filtrage lxxxiv

5 Conclusions lxxxv

Annexe I : Compléments sur la paramétrisation de la DSD en haut de la colonne dans VAMOS **lxxxvii**

1 Implémentation dans YAO lxxxvii

2 Lien expressions de Ulbrich (1983) et Testud et al. (2001) lxxxix

Annexe J : Surapprentissage dans VAMOS **xc**

Annexe K : Correction de l'atténuation due au radôme des radars à visée verticale **xciv**

Annexe L : Article "Estimation of Finescale Rainfall Fields Using Broadcast TV Satellite Links and a 4DVAR Assimilation Method", paru Journal of Atmospheric and Oceanic Technology. **xcvi**

Annexe M : Article "4-D-VAR assimilation of disdrometer data and radar spectral reflectivities for rain drop size distribution and vertical wind retrievals", révision Atmospheric Measurement Technique. **cxviii**

Nomenclature

Généralités

D : diamètre (mm)
v : vitesse (m/s)
V : volume (m^3)
S : surface (m^2)
m : masse (kg)
z : altitude (km)
x, y : coordonnées du plan (km)
c : vitesse de la lumière dans le vide (m/s)
λ : longueur d'onde (m)
$\mathcal{M}_f^m$: moment d'ordre m de la distribution f
$\mathcal{N}(a, b^2)$: loi normale, espérance a, écart-type b
$\mathcal{U}$: loi uniforme
$\mathcal{L}$: loi quelconque

Maillage des discrétisations

ΔD : pas de diamètre (mm)
Δt : pas de temps (s)
Δx, Δy, Δz : pas d'espace (m)
NT : nombre de pas de temps
ND : nombre de pas de diamètre
NX, NY, NZ : nombre de pas d'espace

Indices pour les grandeurs discrètes

p : classes de diamètres
g : goutte particulière
k : temps
i : espace (le plus souvent coordonnée verticale, associé à un j si horizontal)

Pluie

n : distribution en taille des gouttes de pluie $(m^{-3}mm^{-1})$

RR : taux de pluie (mm/h)
LWC : contenu en eau liquide (kg/m^3)
D_m : diamètre volumique moyen (mm)
N_{tot} : densité de gouttes (m^{-3})
Z : facteur de réflectivité radar $(mm^6m^{-3}$ ou $dBZ)$
v_{th} : vitesse terminale de chute des gouttes (m/s)

Météorologie générale

σ_{air} : intensité de la turbulence atmosphérique (m/s)
w : vent vertical (si vent horizontal négligé) (m/s)
v_x, v_y, v_z : trois composantes du vent (sinon) (m/s)
T : température $(^\circ C)$
P_{atm} : pression atmosphérique (Pa)
e_{sat} : pression de vapeur saturante (Pa)
h_r : humidité relative (%)
ν_a : viscosité cinématique de l'air (m^2/s)
D_v : diffusivité de la vapeur d'eau dans l'air (m^2/s)
R_d, R_v : constante des gaz pour l'air sec, pour la vapeur d'eau $(J/kg/K)$
ρ_a, ρ_e : masse volumique de l'air, de l'eau (kg/m^3)
K_T : conductivité de la chaleur dans l'air $(W/m/K)$
L_v : chaleur latente de vaporisation de l'eau liquide (J/kg)

Radars

r : distance radiale (m)
P_e, P : puissance (émise, reçue) (W)
τ_{on}, τ_{off} : temps de pulse, temps interpulse (s)
θ_0, ϕ_0 : ouverture de faisceau (rad)
σ : surface efficace de rétrodiffusion (m^2)
σ_{ext} : surface efficace d'extinction (m^2)
η : réflectivité (m^2/m^3)
W : vitesse Doppler moyenne (m/s)
LS : largeur spectrale (m/s)
Z_e : facteur de réflectivité radar équivalent $(mm^6m^{-3}$ ou $dBZ)$
η_v : réflectivité spectrale $(m^2/m^3/(m/s))$
κ : coefficient spécifique d'atténuation (m^{-1})
L : atténuation atmosphérique totale (dB)
$\mathcal{A}$: atténuation (dB)
N_{fft} : nombre de points d'un spectre Doppler
θ, ϕ : coordonnées sphériques

Assimilation de données (chapitre III)

x^t : état vrai du système ("true")

x^b : état d'ébauche ("background")

x^a : état analysé ("analyzed")

x^g : état d'initialisation ("first guess")

x : état quelconque (typiquement l'inconnue de la fonction de coût du 4D-VAR ou l'état courant à une itération quelconque du 4D-VAR)

y : observations

$M_{k,k+1}$: modèle (opérateur non linéaire) propageant un état de la date t_k à la date t_{k+1}

M_k : modèle (opérateur non linéaire) propageant un état de la date t_0 à la date t_k

$\boldsymbol{M_{k,k+1}}$, $\boldsymbol{M_k}$: Idem dans le cas linéaire (matrices)

$\boldsymbol{B}$: matrice de covariance d'erreur d'ébauche

$\boldsymbol{R}$: matrice de covariance d'erreur d'observations

$\boldsymbol{Q}$: matrice de covariance d'erreur modèle

H : Opérateur (non linéaire) d'observation

$\boldsymbol{H}$: opérateur d'observation linéaire ou linéaire tangent de H

J : fonction de coût

J^b, J^{obs}, J^f : parties ébauche, observations et filtrage de la fonction de coût

c^f : coefficient de pondération du terme de filtrage

Modèles (chapitre II)

κ_x, κ_y : diffusivité dans l'EDP de déplacement des champs des taux de pluie (m^2/s)

γ : taux d'augmentation du taux de pluie dans une cellule (s^{-1})

$\tilde{\epsilon}$: variable aléatoire caractérisant la variation aléatoire des taux de pluie dans une cellule

ξ : taux d'évaporation (kg/s)

Φ_{coll}, Φ_{coal}, Φ_{bu} : noyaux de collision, de coalescence, de break-up (m^3/s)

E_{coll}, E_{coal}, E_{bu} : efficacités de collision, de coalescence, de break-up

χ_{bu} : fonction de break-up (kg^{-1})

I_1, I_2, $I^{'}$: termes sources de l'EDP de dynamique de la DSD $(m^{-3}kg^{-1}s^{-1})$

ϕ, $\tilde{\phi}$, $\tilde{\phi}^*$: champ advecté continu, discret, état intermédiaire du schéma de Smolarkiewicz

ϕ_0 : condition initiale pour l'EDP d'advection

F : fonction utilisée pour simplifier l'écriture des schémas numériques

ϵ, S_c : coefficients du schéma de discrétisation de Smolarkiewicz

Capteur Ku (chapitre IV)

$\boldsymbol{v}$: vecteur vent (advection) $(\boldsymbol{v} = (v_x, v_y))$

δ : direction du vent (direction d'advection) ($^\circ$)

L : longueurs des liaisons Ku (km)

α : élévations (liaisons Ku) ($^\circ$)

γ : orientation d'une liaison Ku (azimut) ($^\circ$)
$abias$, $rmse$, $rbias$, $b95$: indicateurs statistiques pour évaluer la qualité des restitutions
NL : nombre de liaisons Ku
A : atténuation linéique (dB/km)

Profils de DSD (chapitre V)

α, θ, k : 3 coefficients d'une DSD gamma
Γ : fonction gamma d'Euler
N_{dis} : nombre de gouttes disdromètre (flux, sans unité)
η_{obs} : spectres Doppler en sortie des opérateur d'observation ($m^2/m^3/(m/s)$)
$MAPE$, $rbias$, MAE : indicateurs statistiques pour évaluer la qualité des restitutions

Introduction

1 Contexte

L'amélioration des connaissances sur les précipitations est un problème crucial aux implications scientifiques et sociétales considérables. La gestion des ressources en eau est en effet fondamentale pour l'agriculture et l'absence de pluie peut conduire à des sécheresses aux conséquences dramatiques. Au contraire, l'excès de précipitations provoque des inondations. Celles-ci peuvent être soudaines et violentes (crues éclair) et entraîner des dégâts humains et matériels considérables, notamment dans le sud de la France. La connaissance de la physique régissant la pluie et de sa variabilité spatiotemporelle est donc capitale.

Mais l'étude des précipitations peut être abordée à plusieurs niveaux. Si l'on se place à l'échelle des cellules de pluie (échelle macrophysique), un évènement pluvieux sera caractérisé par son étendue spatiotemporelle, et par les taux de pluie qu'il implique en son sein. Cependant, la pluie est avant tout composée de gouttes d'eau, et si on l'étudie à l'échelle de ces gouttes d'eau (échelle microphysique), l'évènement pluvieux sera caractérisé par ses distributions en tailles de gouttes (DSD - Drop Size Distribution). Caractériser ces DSD passera nécessairement par l'étude des phénomènes physiques qui les affectent (vent, collisions entre gouttes, évaporation, etc.). La connaissance des DSD permet ensuite d'estimer les grandeurs intégrées - taux de pluie, contenu en eau - définissant les évènements à l'échelle macrophysique.

Par ailleurs, et quelque soit l'échelle à laquelle on se place, on peut aborder l'étude des précipitations du point de vue de l'observateur ou du point de vue du modélisateur.

Les instruments permettant d'observer la pluie sont nombreux. Certains, généralement placés au sol, fournissent des mesures directes. Ils estiment des paramètres intégrés (pluviomètres, qui mesurent des taux de pluie), ou des DSD (disdromètres, qui enregistrent des flux de gouttes). D'autres sont des instruments de télédétection permettant des mesures indirectes et à distance. Il s'agit bien sûr en premier lieu des radars météorologiques. Dans leur utilisation opérationnelle (par exemple par les services météorologiques nationaux), ils scannent l'espace et mesurent des réflectivités (énergie rétrodiffusée par les gouttes de pluie), qui sont ensuite converties en taux de pluie via les célèbres relations Z-R. Celles-ci servent à déduire des taux précipitants - permettant une étude macro-physique - mais sont dépendantes

de la DSD, et donc de la microphysique. Les radars Doppler à visée verticale permettent eux d'obtenir des informations sur les DSD. Mais leurs mesures - indirectes - sont affectées par un certain nombre d'autres paramètres météorologiques (vent, turbulence, atténuation) qui rendent l'inversion délicate. Toujours dans le domaine de la télédétection, on dispose aussi de systèmes permettant de mesurer les atténuations d'ondes électromagnétiques provenant d'antennes relais de téléphonie mobile ou de satellites TV. En utilisant une relation atténuation-taux de pluie (fonction de la DSD), on obtient des informations sur les taux précipitants. Les satellites, enfin, ambitionnent aujourd'hui de mesurer la pluie depuis l'espace, mais il s'agit encore une fois de mesures indirectes (notamment via les températures de brillance), d'autant plus sujettes à des erreurs qu'elles sont intégrées sur des zones étendues alors que la pluie est un phénomène très variable.

Le fait que ces mesures sont de natures différentes et à différentes résolutions les rend difficilement comparables et parfois difficilement exploitables en l'état, notamment pour les QPE (Quantitative Precipitation Estimation - estimation quantitative des précipitations) (Berne and Krajewski (2013)), ou pour la validation des modèles (Mcfarquhar (2010)).

L'évolution spatiotemporelle de la pluie peut en effet aussi être modélisée. D'un point de vue macro-physique, on modélisera plutôt les déplacements et les déformations des cellules pluvieuses au gré des vents, et la variation de leurs taux de pluie moyens, sans tenir compte de la microphysique de la pluie. D'un point de vue microphysique, on modélisera les phénomènes physiques affectant les gouttes, conduisant typiquement à des systèmes d'équations aux dérivées partielles. Mais les premiers (modèles macro-physiques) se doivent d'être basés sur une bonne connaissance des processus microphysiques sous-jacents pour ne pas être déconnectés de la réalité. Et les seconds (modèles microphysiques) nécessitent la paramétrisation de nombreux phénomènes. Ils sont complexes et difficilement réglables. Les observations susceptibles de corroborer ces modèles microphysiques étant comme on l'a vu indirectes ou ne couvrant pas toute la colonne atmosphérique, ils souffrent généralement d'un manque de validation.

L'assimilation de données est une théorie mathématique permettant de coupler des observations dans des modèles numériques de propagation. Elle est par exemple utilisée dans les modèles de prévisions météorologiques pour déterminer le meilleur état initial sachant le modèle et les observations disponibles. Supposons par exemple que l'on dispose d'un modèle numérique (typiquement une discrétisation d'équations aux dérivées partielles) permettant de propager un champ initial d'un certain paramètre dans le temps, et que l'on dispose par ailleurs d'observations de ce paramètre réparties dans le temps (éventuellement indirectes et/ou intégrées). L'assimilation de données variationnelle va fournir un champ initial qui, une fois propagé par le modèle, sera le mieux à même d'expliquer les observations disponibles. Il s'agit donc d'une méthode fusionnant les différentes observations disponibles pour restituer des paramètres expliquant ces observations et dont la cohérence physique est assurée par le

modèle.

L'assimilation de données semble donc particulièrement bien adaptée aux précipitations, pour lesquelles on a vu que les approches "observations" et "modèles" seules pouvaient être insuffisantes. L'objectif de cette thèse est de développer des algorithmes d'assimilation de données pour fusionner des observations des précipitations dans des modèles numériques afin d'étudier la pluie et sa variabilité spatiotemporelle, à la fois à l'échelle microphysique et à l'échelle macro-physique. Nous allons à présent détailler quelque peu ces objectifs.

2 Objectifs de la thèse

Cette thèse s'intéresse au développement d'algorithmes fusionnant des mesures hétérogènes de la pluie en vue d'améliorer la restitution des précipitations et d'étudier leur variabilité spatiotemporelle, aux échelles macro et microphysiques. En pratique, nous travaillerons d'abord à l'échelle macro-physique. L'objectif de cette première application sera de restituer des cartes de précipitations à partir de mesures d'atténuations d'ondes électromagnétiques provenant de satellites géostationnaires TV-SAT en bande Ku. Cette restitution de cartes de pluie sera effectuée via un algorithme d'assimilation de données 4D-VAR utilisant un modèle d'advection des champs de pluie (déplacement sans déformations selon le vent). Nous évaluerons la qualité des cartes restituées en les comparant à des cartes fournies par les radars opérationnels de Météo France et à des mesures au sol de pluviomètres.

La relation entre atténuation de l'onde et taux de pluie est comme on l'a vu dépendante de la microphysique (de la DSD). Nous avons donc ensuite choisi d'utiliser la même technique d'assimilation de données 4D-VAR pour la restitution des profils verticaux de DSD et de paramètres météorologiques associés (vent vertical par exemple) pour l'étude de sa variabilité spatiotemporelle et des phénomènes physiques sous-jacents. Cet algorithme d'assimilation, baptisé VAMOS (Variational Assimilation of Multi-scale Observations for rain Studies), utilise des observations au sol (disdromètres) et en altitude (spectres Doppler provenant d'un radar à visée verticale). Il permet de fusionner ces observations hétérogènes dans un modèle de propagation verticale des DSD prenant en compte le vent vertical et l'évaporation. Les objectifs de ce travail, outre la fusion des observations en elle-même, sont d'abord l'étude des phénomènes physiques affectant la DSD durant la chute des gouttes (pour un évènement donné, est-ce l'évaporation qui est prépondérante, ou l'effet du vent horizontal, ou des mouvements verticaux de l'air ?). Enfin, nous développons une méthode qui pourrait par la suite être utilisée pour paramétrer ces phénomènes physiques directement à partir des observations. Nous validons par ailleurs la qualité de nos restitutions sur un cas d'étude en les comparant à celles obtenues indépendamment par une autre technique utilisant les spectres de deux radars Doppler à visée verticale.

3 Plan

Dans ce travail, nous couplons grâce à l'assimilation de données des observations et des modèles d'évolution des précipitations. Nous commencerons dans les 2 premiers chapitres de ce manuscrit par exposer les concepts et enjeux liés à la définition et à l'observation des précipitations, et à leur modélisation (aux échelles macro et microphysique). Le chapitre 3 s'intéressera plus particulièrement à la question de l'assimilation de données et à sa mise en œuvre dans le contexte de notre étude. Le lecteur qui serait familier avec ces concepts peut bien sûr s'abstenir de lire les chapitres correspondant et est alors encouragé à se reporter à la nomenclature (page vii) pour simplement se familiariser avec les notations utilisées. Les chapitres 4 et 5 du manuscrit présentent en détails les résultats obtenus pour les deux applications mentionnées plus haut (restitution de cartes de précipitations et étude des profils de DSD).

Dans le chapitre I, nous définissons tout d'abord les principales grandeurs caractérisant les précipitations (DSD, taux de pluie, contenu en eau liquide, etc.). Ensuite, nous étudions les différents instruments conçus pour mesurer la pluie et utilisés dans cette étude (pluviomètres, disdromètres, radars, capteurs d'atténuation) et précisons la nature de leur mesure.

Le chapitre II revient sur la modélisation de l'évolution spatiotemporelle de la pluie. Nous présentons d'abord rapidement les modèles d'advection-diffusion utilisés pour représenter l'évolution des cellules de pluie indépendamment de leur microphysique. Puis nous revenons plus en détails sur les modèles d'équations aux dérivées partielles développés pour représenter l'évolution des DSD et modélisant la physique affectant les gouttes. Nous nous attardons aussi quelque peu sur les relations vitesse de chute/diamètre des gouttes de pluie et sur la question de la discrétisation des équations aux dérivées partielles.

Le chapitre III concerne l'assimilation de données. Il présente rapidement les différents algorithmes d'assimilation existant et leurs spécificités et s'attarde davantage sur l'algorithme 4D-VAR, qui est utilisé dans ce travail. Nous introduisons ensuite YAO, un logiciel d'aide à la programmation de tels algorithmes.

Le chapitre IV présente la première application de ce travail concernant la restitution de cartes de pluie à partir de mesures d'atténuation d'ondes TV-SAT. Une étude bibliographique est fournie en début de chapitre afin de replacer cette étude dans son contexte, puis nous donnons les détails de l'algorithme de restitution utilisé et les résultats obtenus.

Le chapitre V présente la seconde application concernant l'étude des profils verticaux de DSD par la fusion de données radar et disdromètre. Comme précédemment, après avoir contextualisé ce travail, nous présentons notre algorithme (VAMOS) et les différents résultats obtenus sur données simulées et sur 3 cas d'étude réels.

Partie I

Définir et mesurer la pluie

Dans ce chapitre, nous allons introduire quelques grandeurs permettant de comprendre et de définir la pluie. Dans un premier temps, nous travaillerons avec une vision "volumique" de l'atmosphère, c'est-à-dire en figeant un volume d'atmosphère à une date donnée. Nous définirons ainsi la notion de distribution en taille des gouttes de pluie, puis utiliserons cette définition pour introduire les concepts de taux de pluie, contenu en eau liquide, etc.

Ensuite, nous présenterons le principe de fonctionnement de différents instruments utilisés pour mesurer les précipitations (pluviomètres, disdromètres, radars et capteurs d'atténuation) et la nature de leurs mesures.

1 Définir la pluie : de la goutte au taux précipitant

1.1 Distribution volumique de gouttes de pluie : point de vue aléatoire

Considérons un certain volume de l'espace, figé à une date donnée. Supposons qu'une goutte de pluie soit présente dans ce volume, et notons V_g son volume. Si c'est une sphère, on note D_g son diamètre, qui est défini par :

$$V_g = \frac{\pi}{6} D_g^3 \tag{I.1}$$

On peut inverser cette expression pour obtenir D_g à partir de ce volume : $D_g = \frac{6}{\pi} \sqrt[3]{V_g}$. Si le volume n'est pas une sphère, on obtient, dans ce cas-là, un "diamètre équivalent".

On suppose qu'un diamètre D suit une loi de probabilité $\mathcal{L}$, et que cette loi est homogène (ses paramètres ne changent pas) sur le volume considéré. On suppose qu'elle possède une densité de probabilité, notée f ($f(D)$ est une fonction du diamètre D, et est en mm^{-1}). On note$\int_{D_1}^{D_2} f(D)dD$ la probabilité qu'une goutte du volume ait un diamètre compris entre D_1 et D_2.

Soit N_{tot} le nombre total de gouttes dans le volume considéré (normalisé par le volume, donc en m^{-3}). N_{tot} peut également être vu comme une variable aléatoire, avec sa loi spécifique. On le suppose néanmoins connu. Il est traité ici comme un paramètre fixé. Les diamètres des N_{tot} gouttes suivent la loi $\mathcal{L}$. On définit la distribution en taille des gouttes de pluies (DSD) n, en m^{-4} (souvent exprimée dans la littérature ainsi que dans ce travail en $m^{-3}mm^{-1}$) par :

$$n(D) = N_{tot} f(D) \tag{I.2}$$

$\int_{D_1}^{D_2} n(D)dD = N_{tot} \int_{D_1}^{D_2} f(D)dD$ représente le nombre moyen de gouttes dans la gamme de diamètres $[D_1 : D_2]$.

Dans la réalité, les diamètres des N_{tot} gouttes effectivement présentes dans notre volume peuvent être vus comme N_{tot} réalisations de la loi $\mathcal{L}$ (la valeur N_{tot} pouvant elle-même être vue comme une réalisation de la distribution de probabilités sous-jacente au nombre de gouttes). En discrétisant l'espace des diamètres, on peut ranger les gouttes dans des classes de diamètre et évaluer le nombre de gouttes dans chaque classe. En notant ΔD le pas de cette discrétisation, on calcule l'histogramme des réalisations, en notant n_i le nombre de gouttes dans la i-ème classe de diamètres (i.e. dans un intervalle $[D_i : D_i + \Delta D]$). En divisant chacun des n_i par N_{tot}, on peut chercher à inférer $\mathcal{L}$, sa nature (gamma, log-normale, ...), et ses paramètres, c'est-à-dire à estimer f.

1.2 Le taux de pluie

Le taux de pluie mesure le flux d'eau à travers une surface pendant une durée donnée. Définissons le d'abord intuitivement. Notons S la surface considérée, Δt la durée (temps d'intégration). Supposons que durant le temps Δt, $N_{S,\Delta t}$ gouttes traversent la surface S. Notons $(V_p)_{p\in[|1:N_{S,\Delta t}|]}$ l'ensemble des volumes de ces gouttes. Le taux de pluie $RR_{S,\Delta t}$ est alors défini par :

$$RR_{S,\Delta t} = \frac{1}{S}\frac{1}{\Delta t}\sum_{p=1}^{N_{S,\Delta t}} V_p \tag{I.3}$$

On obtient ainsi une grandeur le plus communément exprimée en mm/h. Le cumul d'eau pendant la durée Δt, le plus souvent exprimé en mm, est naturellement la quantité $RR_{S,\Delta t}\Delta t$.

En définissant comme précédemment un diamètre ou un diamètre équivalent D_p pour la p-ième goutte, on a :

$$RR_{S,\Delta t} = \frac{\pi}{6}\frac{1}{S}\frac{1}{\Delta t}\sum_{p=1}^{N_{S,\Delta t}} D_p^3 \tag{I.4}$$

Comme précédemment, supposons à présent que l'on a, dans un volume situé immédiatement au-dessus de la surface S, des gouttes (réalisations de la loi $\mathcal{L}$) rangées par classes de diamètres. On a alors :

$$RR_{S,\Delta t} = \frac{\pi}{6}\frac{1}{S}\frac{1}{\Delta t}\sum_{p=1}^{N_D} K_p \tilde{D_p}^3 \tag{I.5}$$

Avec $\tilde{D_p}$ le diamètre cubique moyen des gouttes de la classe p, N_D le nombre de classes de diamètre, et K_p une grandeur sans unité correspondant au nombre de gouttes de la classe p traversant la surface S pendant la durée Δt.

Pour exprimer simplement K_p, nous supposons qu'il existe une relation entre le diamètre des gouttes et leur vitesse (il s'agit d'un flux). Ainsi nous considérons que toutes les gouttes de la classe p tombent avec une même vitesse notée v_p. Cette vitesse peut typiquement être une vitesse terminale de chute gravitationnelle (équilibre entre la gravité et la resistance de l'air), plus une correction par le vent vertical. On se rapportera au chapitre II, section 3 pour plus de détails sur les relations vitesse / diamètre. On la relie explicitement au diamètre par une fonction g telle que : $v_p = g(D_p)$. Si on note n_p le nombre de gouttes dans la *p-ième* classe de diamètres, on obtient alors :

$$K_p = n_p v_p \Delta D S \Delta t \tag{I.6}$$

Avec toujours ΔD la largeur de la *p-ième* classe de diamètres.

Dans les faits, la vitesse réelle d'une goutte est composée d'une partie déterministe à laquelle s'ajoute un aléa. En supposant cet aléa gaussien de moyenne nulle (soit $v_p = g(D_p) + \epsilon$, avec $\epsilon \sim \mathcal{N}(0, \sigma^2)$), on peut supposer que l'équation (I.6) reste vraie en moyenne.

En reportant (I.6) dans (I.5), on trouve pour le taux de pluie :

$$RR_{S,\Delta t} = \frac{\pi}{6} \sum_{p=1}^{N_D} n_p v_p \tilde{D}_p{}^3 \Delta D \tag{I.7}$$

On remarque que S et Δt n'apparaissent pas dans cette expression, ce qui repose sur deux hypothèses implicites :

- le volume sur lequel on a estimé la réalisation de notre DSD inclut la surface S.

- les gouttes traversant S pendant Δt viennent toutes d'un volume dans lequel les diamètres suivent la même loi.

Schématiquement, en notant S_{dsd} et h_{dsd} respectivement la surface de notre volume d'échantillonage parallèle à S et sa hauteur (dans la direction perpendiculaire à S), on doit avoir : $S_{dsd} > S$ et $h_{dsd} > v_p \Delta t$, pour tout p.

En travaillant avec des grands volumes sondés on peut passer des sommes aux intégrales. On peut donc réécrire (I.7) sous la forme suivante :

$$RR_{S,\Delta t} = \frac{\pi}{6} \int_D n(D) v(D) D^3 dD \tag{I.8}$$

On remarque que la DSD a remplacé le nombre de gouttes par bin.

Comme précédemment, on a supposé une relation vitesse/diamètre pour les gouttes, et que la DSD était homogène sur le volume sondé.

1.3 Moments de la distribution

Le moment d'ordre m de la variable aléatoire "diamètre des gouttes" définie par la distribution f, s'il existe, est noté $\mathcal{M}_f^m$ et s'exprime par :

$$\mathcal{M}_f^m = \int_D f(D) D^m dD \tag{I.9}$$

De la même manière, le moment d'ordre m de la DSD, noté $\mathcal{M}_n^m$, est défini par :

$$\mathcal{M}_n^m = \int_D n(D) D^m dD = N_{tot} \mathcal{M}_f^m \tag{I.10}$$

Nous allons rapidement présenter quelques moments de la DSD particulièrement importants dans l'étude des précipitations.

1.3.1 Ordre 0

Le moment d'ordre 0 est simplement le nombre volumiques de gouttes, N_{tot}, en m^{-3}. En effet, $\mathcal{M}_f^0 = \int_D f(D) dD = 1$ donc $\mathcal{M}_n^0 = N_{tot}$.

1.3.2 Ordre 3

Le moment d'ordre 3 de la DSD s'écrit $\int_D n(D)D^3dD$. Le cube qui apparaît au diamètre relie ce moment au volume d'eau présent. $\frac{\pi}{6}D^3$ étant le volume d'une goutte, $\frac{\pi}{6}\mathcal{M}_n^3$ estime le volume d'eau par unité de volume d'air (en m^3/m^3). En multipliant par la masse volumique de l'eau ρ_{eau}, on définit le contenu en eau liquide (liquid water content) par $LWC = \rho_{eau}\frac{\pi}{6}\mathcal{M}_n^3$, masse d'eau par unité de volume d'air, en $kg\ m^{-3}$.

1.3.3 Ordre 6

Le moment d'ordre 6 de la DSD, $\int_D n(D)D^6dD$, est appelé facteur de réflectivité radar (Z, voir l'Eq.(I.17) dans la section 2.3 pour son lien avec la mesure radar). Il s'exprime en $mm^6\ m^{-3}$.

On remarque enfin que le moment d'ordre 3.67 de la DSD est souvent assimilé au taux de pluie (voir par exemple Ulbrich (1983)), dans la mesure où la vitesse de chute des gouttes peut être approchée par une expression proportionnelle à $D^{0.67}$ (voir Eq.(I.8)). Dans ce travail, nous n'utiliserons cependant pas cette expression (voir chapitre II, section 3 pour une étude sur la vitesse terminale de chute des gouttes).

1.3.4 Cas discret / réalisation

Dans le cas où l'on ne travaille pas avec des distributions de probabilités ("théoriques" ou inférées) mais avec des nombres de gouttes par classe de diamètre (n_p), on peut bien entendu définir de la même manière les moments correspondant. En utilisant (abusivement) la même notation pour ce nouveau moment d'ordre m, on a :

$$\mathcal{M}_n^m = \sum_{p=1}^{N_D} n_p D_p^m \Delta D \tag{I.11}$$

Pour que ce moment estimé $\mathcal{M}_n^m$ reste fidèle au moment attendu, cela suppose que les classes de diamètres sont suffisamment petites pour que l'erreur due à la quantification reste négligeable. On voudra donc une erreur $\|n_p D_p^m - \sum_{o=1}^{n_p} D_{p,o}^m\|$ aussi petite que possible ($D_{p,o}$ correspond au *o-ième* diamètre de la classe p).

On pourra ensuite bien sûr utiliser (I.11) pour évaluer directement à partir des mesures le nombre total de gouttes, le contenu en eau liquide, ou le facteur de réfléctivité radar.

2 Quels instruments pour mesurer la pluie ?

Un certain nombre d'instruments sont utilisés pour mesurer, directement ou indirectement, la pluie. Nous allons dans cette section lister les principaux et donner leur principe de fonctionnement et leurs caractéristiques. On s'étendra davantage sur ceux utilisés dans la suite de ce travail.

2.1 Pluviomètres

Les pluviomètres sont des instruments utilisés pour mesurer des taux de précipitation (ou des cumuls de précipitations) au niveau du sol. La majorité des pluviomètres non manuels (notamment la totalité de ceux du réseau Météo France ainsi que ceux utilisés en validation dans le chapitre IV de ce travail) sont des pluviomètres à augets basculants. Leur principe de fonctionnement est relativement simple. L'eau est collectée sur une certaine surface, puis dirigée vers un petit réservoir, fixé sur un auget basculant. Lorsque l'eau contenue dans le réservoir dépasse une certaine masse, celui-ci bascule et l'eau pénètre alors dans un second auget. En notant les dates de basculement, on connaît la quantité d'eau tombée entre deux dates successives. En divisant par la durée entre ces deux dates et par la surface de collecte, on estime la masse d'eau tombée par unité de temps et de surface, et on remonte au taux de pluie en divisant par la masse volumique de l'eau.

Le prinicipe de fonctionnement de ces appareils et les erreurs de mesure associées sont détaillés dans un rapport de l'Organisation Météorologique Mondiale (Jarraud (2008), pages 144-145) ou dans une note de Météo France (Leroy (2000)).

Comme sources d'erreurs, on peut citer l'évaporation dans les augets entre deux basculements (en cas de pluie très faible), une perte d'eau pendant un basculement (pendant que l'auget bascule, l'eau qui tombe est perdue, et la quantité peut devenir significative en cas de pluie forte), ou encore le fait que de l'eau reste collée aux augets, et ne sort pas lors d'un basculement, ce qui surestime ensuite la quantité d'eau au basculement suivant. Ces erreurs sont estimées, selon le type de pluie, et éventuellement corrigées (Leroy (2000), Jarraud (2008)). Le vent est plus difficile à prendre en compte. Le fait que les gouttes ne tombent pas droit peut ainsi causer des erreurs significatives, renforcées par le fait que la forme du pluviomètre lui-même peut induire des courants verticaux écartant les petites gouttes du pluviomètre. Météo France estime les erreurs dans ces cas-là de 10% (vent à $5m/s$) à 25% (vent à $10m/s$) (Leroy (2000)).

On remarque également que les pluviomètres sont des instruments qui mesurent des flux d'eau (une masse d'eau traversant une surface durant un temps donné). Ces flux sont mesurés à travers des surfaces relativement petites (de l'ordre de 400 à $1000cm^2$), tant et si bien que l'on peut considérer, notamment par rapport aux mesures radar, qu'ils donnent des mesures ponctuelles de la pluie. Enfin, ils sont la plupart du temps installés au niveau du sol.

2.2 Disdromètres

Les disdromètres sont des instruments conçus pour estimer les distributions en taille des gouttes de pluie. Historiquement, les premières estimations de tailles de gouttes de pluie ont été réalisées en mesurant la taille de l'impact des gouttes sur du papier saupoudré de farine (Laws and Parsons (1943))) ou sur du papier absorbant (Marshall et al. (1947)). Ces mesures ont été réalisées parallèlement à l'essor des radars afin d'estimer la pertinence des mesures qu'ils fournissaient (Marshall et al. (1947), Marshall and Palmer (1948)), ou pour estimer l'impact de la pluie sur les sols (Laws and Parsons (1943)). Des instruments automatisés ont ensuite été développés. Nous en présentons 3 types ici, qui seront utilisés par la suite, à savoir le disdromètre Joss-Waldvogel (JW), le disdromètre 2D vidéo (2DVD), et le spectropluviomètre (DBS). La Figure I.1 montre des photographies de ces trois instruments.

FIGURE I.1 – Photographies des trois disdromètres utilisés dans ce travail. a) disdromètre Joss-Walvogel, de la société Distromet Ltd. (http://distromet.com/), image provenant du site de l'Atmospheric Measurement Research (https://www.arm.gov/instruments/disdrometer) ; b) disdromètre 2D vidéo, même source (http://www.arm.gov/instruments/vdis) ; c) spectropluviomètre du LATMOS.

2.2.1 Disdromètre Joss-Waldvogel (JW)

Le premier disdromètre développé l'a été par Joss and Waldvogel (1967). Il consiste en une surface de collecte de $50cm^2$ que les gouttes viennent frapper. L'impact produit par une goutte est transformé en un signal électrique, dont l'intensité est reliée à la taille et à la vitesse de la goutte. En supposant que les gouttes tombent à leur vitesse terminale de chute (et donc qu'il existe une relation non ambigüe entre diamètre et vitesse), on remonte à leurs diamètres. Dans sa forme commerciale, le disdromètre JW mesure des gouttes entre 0.3 et $5.5mm$, et fournit un histogramme de leur distribution avec une résolution en diamètre de $0.2mm$ jusqu'à $5mm$. Les gouttes de taille supérieure à $5mm$ sont rangées dans la dernière classe de diamètre, $5 - 5.5mm$ (Tokay et al. (2005)).

On peut citer quelques facteurs influant sur la qualité des mesures réalisées. Tout d'abord, on note que cet instrument suppose une relation vitesse/diamètre, or les gouttes ne tombent pas forcément à leur vitesse terminale de chute, du fait notamment de la présence de mouvements d'air verticaux (qui modulent cette vitesse), de turbulence, ou de collisions entre gouttes (il faut un temps de relaxation non nul après une collision pour que les gouttes résultantes reviennent à

leur vitesse terminale de chute). Par ailleurs, si deux gouttes tombent sur la surface de collecte dans un intervalle de temps très court, rien ne lui permettra de différencier leur impact. Il en résultera une erreur (éventuellement marginale) sur la taille de la plus grosse des deux gouttes. La plus petite ne sera elle pas comptabilisée. Ce disdromètre peut ainsi sous-estimer le nombre de petites gouttes, en particulier en cas de pluie forte (Tokay et al. (2005)) ou de bruit d'arrière-plan fort (pour des diamètres inférieurs à $0.8mm$, Tokay et al. (2003)).

2.2.2 Disdromètre 2D-vidéo (2DVD)

Une documentation pour le 2DVD (Two-Dimensionnal Video-Disdrometer) est disponible sur le site de son constructeur (`http://www.distrometer.at/en/`). On peut aussi citer l'étude Kruger and Krajewski (2002) qui estime ses performances pour la pluie (le 2DVD pouvant aussi servir pour la neige ou la grêle).

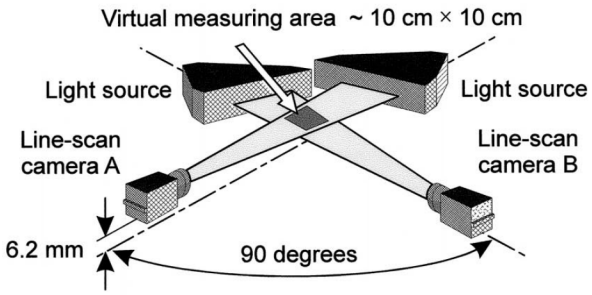

FIGURE I.2 – Figure extraite de Kruger and Krajewski (2002) illustrant le principe de fonctionnement du disdromètre 2D-vidéo.

Le 2DVD consiste en deux sources lumineuses, émettant dans deux directions orthogonales, et captées par deux caméras. La Figure I.2, extraite de Kruger and Krajewski (2002), explicite son fonctionnement. Les deux rayons perçus par les caméras se superposent (selon l'axe vertical) sur une surface de $100cm^2$. C'est la surface efficace de mesure de l'instrument. Les deux plans de $100cm^2$ ainsi définis sont séparés verticalement de $6.2mm$.

Lorsqu'une goutte vient traverser l'un des rayons, la caméra enregistre la forme de sa projection dans l'axe du rayon. En supposant la goutte symétrique autour d'un axe vertical, on obtient ainsi sa forme (son volume). L'algorithme de traitement restitue ensuite, pour chacun des deux couples rayon / caméra, un diamètre équivalent (si la goutte était sphérique) et un indice d'oblicité. Le temps mis par la goutte pour passer du premier rayon au second permet d'estimer sa vitesse. On obtient ainsi en sortie 6 informations pour chaque goutte : date d'arrivée, diamètre équivalent et oblicité pour chaque caméra, et vitesse de chute.

L'avantage d'un tel instrument est qu'il fournit des mesures à la fois de la vitesse et du diamètre des gouttes. On n'a ainsi plus besoin d'une forme analytique pour les relier, et on peut tester la cohérence entre ces deux informations. Kruger and Krajewski (2002) remarque qu'environ 20% des gouttes mesurées sur 2 ans ont été enregistrées avec une vitesse s'éloignant

de plus de 40% de la vitesse théorique de Atlas et al. (1973). Le vent horizontal est en revanche toujours une limitation importante pour cet appareil, d'autant plus que l'instrument est plus gros que le Joss-Waldvogel, et donc susceptible de créer autour de lui des courants ascendants plus forts. Le constructeur recommande de l'utiliser pour des gouttes avec une vitesse inférieure à $10m/s$, et donne alors une résolution en diamètre d'environ $0.2mm$. Kruger and Krajewski (2002) recommande de ne l'utiliser que dans des cas avec un vent horizontal faible près du sol.

Tokay et al. (2001) compare les résultats donnés par deux 2DVD et un JW, ainsi que par huit pluviomètres, tous répartis sur une surface totale de $5000m^2$. Il conclut (sur 8 évènements pluvieux) que les deux disdromètres sous-estiment les taux de pluie, en particulier le JW (de 18% pour le JW, de 12% pour le 2DVD). Tokay et al. (2001) indique que le JW sous-estime le nombre de petites gouttes ($< 1mm$), entraînant une sous-estimation du taux de pluie. Il explique la sous-estimation du 2DVD par le filtre appliqué pour écarter les gouttes dont la vitesse s'éloigne trop de la vitesse théorique terminale de chute.

2.2.3 Spectropluviomètre (DBS)

Le spectropluviomètre à deux rayons (Dual-Beam Spectropluviometer, DBS) est un instrument optique conçu notamment pour mesurer plus précisément les petites gouttes. Delahaye et al. (2006) détaille son fonctionnement. Son principe général est le suivant. Deux sources lumineuses, placées l'une au-dessus de l'autre, émettent deux nappes lumineuses superposées et séparées de $0.2mm$, de surface $100cm^2$ (la surface de collecte). Des photodiodes mesurent l'intensité de signal à l'autre extrémité des nappes. La Figure I.3a, extraite de Delahaye et al. (2006), présente ce dispositif.

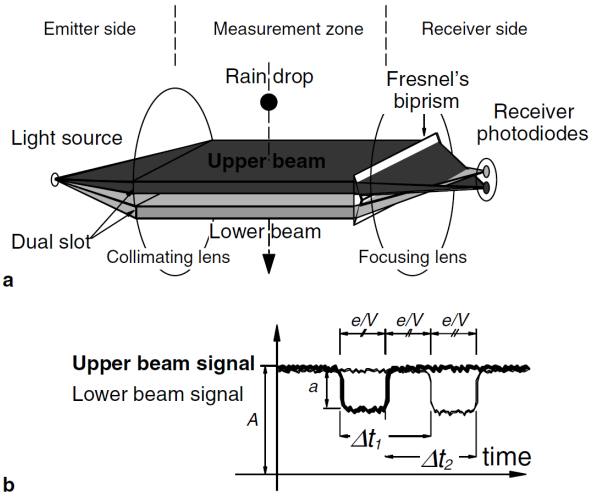

FIGURE I.3 – Figure extraite de Delahaye et al. (2006) illustrant le principe de fonctionnement du spectropluviomètre. Voir le texte pour les commentaires.

Lorsqu'une goutte vient couper un rayon, il en résulte un affaiblissement dans l'intensité captée par la photodiode. En superposant les signaux reçus par les deux diodes (Figure I.3b), on obtient des signaux décalés dans le temps (le temps que la goutte franchisse la distance entre les deux nappes) et dont la forme caractérise la forme de la goutte.

Un algorithme de traitement (voir Delahaye et al. (2006)) extrait ensuite de ces signaux trois paramètres pour chaque goutte détectée : la date d'arrivée, son diamètre équivalent, et sa vitesse. Comme dans le cas du 2DVD, on a ainsi accès aux données de chaque goutte, et on estime indépendemment leur diamètre et leur vitesse.

Krajewski et al. (2006) présente les résultats d'une campagne d'intercomparaison entre un JW, un 2DVD, un DBS et un Parsivel (autre disdromètre optique non présenté ici), pour 6 évènements pluvieux représentant environ $65mm$ d'eau. Il conclut à des différences significatives entre les différents instruments (supérieur à ce que l'on pourrait a priori attendre vue la proximité des instruments). Le JW sous-estime à nouveau très sensiblement le nombre de petites gouttes.

2.2.4 Nature des mesures des disdromètres

Dans tous les cas, les disdromètres (comme les pluviomètres) mesurent des flux de gouttes à travers leur surface de collecte, et généralement au niveau du sol. Considérer ces flux de gouttes comme représentatifs de la distribution en taille des gouttes dans un volume donné autour de l'instrument requiert donc un certain nombre d'hypothèses. Le volume sondé durant un intervalle de temps Δt par un disdromètre dépend en effet de la surface de collecte (S), de la direction et de la force du vent horizontal (si celui-ci n'est pas nul, on ne sonde pas un volume à la verticale de l'instrument), et, surtout, de la vitesse de chute des gouttes d'eau. Supposons ici que nous disposons d'un instrument mesurant à la fois diamètre et vitesse (2DVD ou DBS). Pour extrapoler la mesure de flux à un volume, il faut supposer que la vitesse des gouttes mesurée est restée constante pendant Δt, le temps d'intégration choisi. Et même sous cette hypothèse, le volume sondé sera naturellement beaucoup plus conséquent pour les gouttes rapides (grosses gouttes) que pour les gouttes lentes (petites gouttes). Le schéma de la Figure I.4 illustre cette situation.

Par exemple, pour des gouttes tombant à $10m/s$ (diamètre typiquement vers $6mm$), avec $\Delta t = 1min$, la hauteur sondée sera de $600m$, soit potentiellement une part significative de la colonne de pluie, tandis qu'elle ne sera que de $60m$ pour des gouttes tombant à $1m/s$ (diamètre typiquement vers $0.5mm$). Ajoutant à cela que les grosses gouttes sont généralement significativement plus rares que les petites, on peut supposer que la distribution de probabilités sous-jacente, quand bien même elle serait homogène sur la colonne, sera généralement moins bien échantillonnée pour les grosses gouttes. On peut essayer de réduire l'impact de ce problème. D'une part en groupant les grosses gouttes sur des classes de diamètre plus larges (mais on augmente l'erreur faite en supposant la même vitesse pour toutes les gouttes d'une classe). D'autre part en intégrant sur des durées différentes les petites et les grosses gouttes, pour tenter d'homogénéiser les volumes sondés (voir Figure I.4).

2.3 Radars

Dans cette section, nous allons présenter les mesures réalisées grâce aux radars météorologiques. Nous présenterons tout d'abord la nature de la mesure radar, sans nous

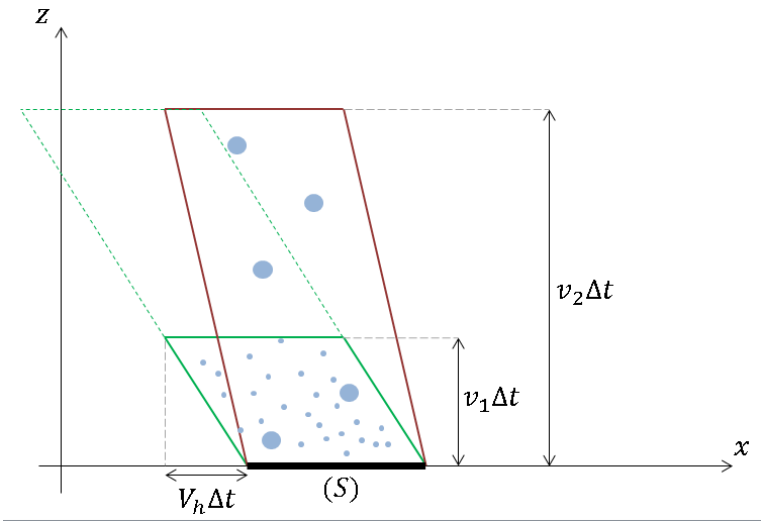

FIGURE I.4 – Volume sondé par un disdromètre. Schéma représentant un plan altitude/dimension horizontale. Le disdromètre est symbolisé par le trait épais noir au sol, sa surface de collecte est notée S. On considère deux classes de gouttes, des petites tombant à une vitesse constante v_1 et des grosses tombant à une vitesse constante $v_2 > v_1$. Toutes les gouttes ont aussi une même vitesse horizontale V_h parallèle à l'axe x (correspondant à un vent horizontal). Pendant la durée Δt suivant cette image, le disdromètre enregistre les gouttes provenant du cadre vert (gouttes lentes) ou rouge (gouttes rapides). On remarque qu'en augmentant le temps d'intégration pour les petites gouttes (cadre vert pointillé), on peut obtenir la même hauteur sondée ($v_2\Delta t$), mais pas les volumes (à moins d'avoir $V_h = 0$).

attarder trop longuement sur le fonctionnement de l'appareil. Puis nous détaillerons davantage l'intérêt de la mesure radar en visée verticale. Nous introduirons ainsi la notion de spectre Doppler dont nous soulignerons le lien avec les autres paramètres météorologiques associés, dont la distribution en taille des gouttes de pluie. Nous présenterons enfin rapidement les différents instruments utilisés dans ce travail.

On peut trouver la description détaillée de l'instrument et de ses mesures dans le livre Doviak and Zrnic (2014), ou, en français, dans le livre Sauvageot (1982). Ces notions sont aussi abordées plus rapidement dans les thèses de Tridon (2011) ou Rysman (2013). Ulaby et al. (1981) présente plus en détails les modèles de diffusion des ondes électromagnétiques par des particules atmosphériques.

2.3.1 Considérations générales : Équation radar

Un radar est un instrument constitué :

- d'un émetteur, qui produit une onde électromagnétique, soit intermittente (radars pulsés), soit continue et modulée en fréquence (radars FMCW, "Frequency Modulated Continuous Wave).

- d'une antenne concentrant, dans la mesure du possible, l'émission de cette onde dans une direction donnée.

- d'un récepteur, captant l'énergie électromagnétique renvoyée au radar par les différents objets présents dans l'atmosphère (gaz, eau liquide, solide, insectes, objets divers).

Volume sondé : porte / ouverture de faisceau

Dans ce paragraphe, nous allons introduire le principe de la mesure radar et définir le volume sondé par l'instrument, ainsi que les notions associées de porte et d'ouverture de faisceau.

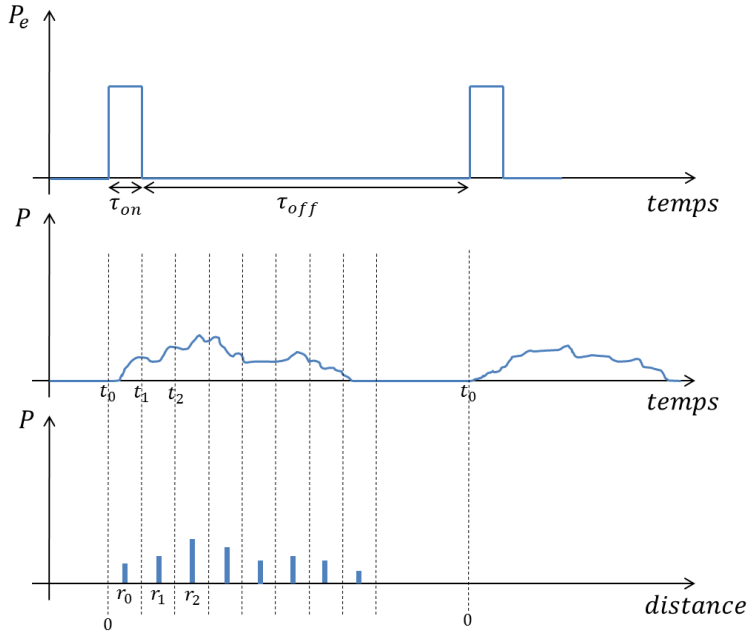

FIGURE I.5 – *Haut* : signal émis par un radar pulsé. On représente la puissance émise P_e en fonction du temps. Il s'agit d'une fonction créneau. Le temps d'impulsion est τ_{on}, le temps entre 2 impulsions τ_{off}. *Milieu* : Représentation schématique du signal renvoyé au radar par des cibles réparties dans la zone visée par le radar (typiquement des gouttes d'eau). Il s'agit de la puissance reçue P. On discrétise ensuite l'axe temporel (dates successives t_i entre 2 impulsions), et on estime la puissance moyenne reçue par le radar sur chaque période $[t_k : t_{k+1}]$. Chaque fenêtre temporelle correspond à une porte radar (zone de l'espace dans laquelle sont comprises les cibles provoquant l'écho enregistré). On définit ici la distance du milieu de chaque porte par $r_k = c\,(t_{k+1} + t_k)\,/2 - ct_0$ (c la vitesse de la lumière). *Bas* : Visualisation de la puissance reçue en fonction de la porte (distance au radar).

Supposons ici que l'on utilise un radar pulsé. Cela signifie qu'il émet une onde électromagnétique à une fréquence donnée λ, pendant une durée donnée τ_{on}, et avec une puissance émise P_e, suivie par une période durant laquelle il n'émet pas d'énergie (durée interpulse τ_{off}). Ce cycle est ensuite répété (Figure I.5, haut). Si une cible est présente à la distance r du radar, elle renvoie une partie de l'onde vers lui. Le temps mis par cette onde pour faire l'aller-retour du radar à la cible est $2r/c$, avec c la vitesse de la lumière. Si l'on a un seul écho, provenant d'une seule cible, on peut donc déduire la distance de cette cible à partir de la date à laquelle l'écho est reçu. Si l'on a des cibles distribuées dans l'espace, on ne reçoit pas en écho un tel Dirac (ou créneau) mais une distribution temporelle, comme celle représentée au milieu de la Figure I.5. En discrétisant ensuite l'axe du temps entre 2 impulsions et en calculant la puissance moyenne reçue sur chaque pas de discrétisation $[t_k : t_{k+1}]$, on obtient, via la même transformation temps/distance, la puissance renvoyée par les cibles présentes dans chaque "porte" radar, c'est-à-dire dans chaque volume situé dans la

même gamme de distance au radar (Figure I.5 et légende). En version (notation) continue, on obtient ainsi $P(r)$, la puissance reçue par le radar en fonction de la distance r.

Comme on l'a déjà mentionné, l'antenne du radar permet de focaliser l'énergie, c'est-à-dire de faire en sorte que le radar n'émette pas l'énergie de manière isotrope mais dans une direction préférentielle. Toute la puissance n'est cependant bien sûr pas émise dans une unique direction. On peut ainsi définir la densité de puissance émise dans chaque direction (formellement en W/sr). En coordonnées sphériques, une telle direction est définie par deux coordonnées θ et ϕ. En supposant (pour simplifier la représentation sur la Figure I.6) cette densité de puissance indépendante de ϕ, on peut la tracer, en $1D$, en fonction de la direction (qui se ramène à l'angle θ) (Figure I.6). On définit ensuite l'ouverture du faisceau (ouverture à $-3dB$) comme l'angle θ_0 englobant les directions dans lesquelles la densité de puissance émise est égale à au moins la moitié de sa valeur maximale. On considère généralement ensuite que les cibles renvoyant de l'énergie au radar sont situées dans l'intervalle correspondant. Plus l'intervalle est petit, plus la puissance est focalisée dans une direction et plus le volume sondé par le radar sera petit. En pratique (2 coordonnées de direction), on caractérise l'ouverture du faisceau par 2 angles θ_0 et ϕ_0, définissant un cône elliptique.

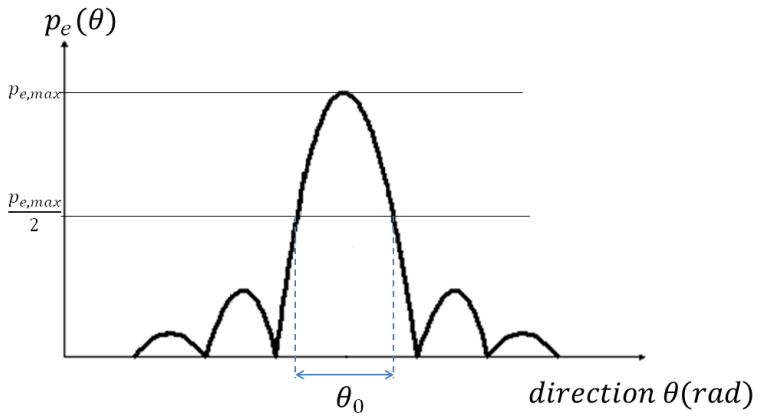

FIGURE I.6 – Définition de l'ouverture de faisceau radar θ_0. On considère un radar qui émet une onde. On représente ici la densité de puissance émise en fonction de la direction. La densité de puissance émise est maximale ($p_{e,max}$) dans une direction donnée. On définit ensuite l'angle θ_0 comme celui englobant toutes les directions dans lesquelles le radar émet une densité de puissance supérieure à la moitié de sa valeur maximale (d'où l'expression "ouverture à $-3dB$" pour désigner θ_0). On note que l'on suppose ici que l'émission est la même dans tous les plans contenant la direction d'émission maximale (on devrait formellement, en coordonnées sphériques, définir deux angles θ_0 et ϕ_0).

La Figure I.7 définit le volume sondé par le radar (V) pour une porte et une ouverture de faisceau données (voir légende).

Équation radar Une cible donnée va recevoir de l'énergie du radar vers lequel elle en ré-émet une partie. Cette cible est caractérisée par sa section efficace de rétrodiffusion, qui sera simplement notée σ (m^2) dans la suite (σ_p, par exemple, est la surface efficace de

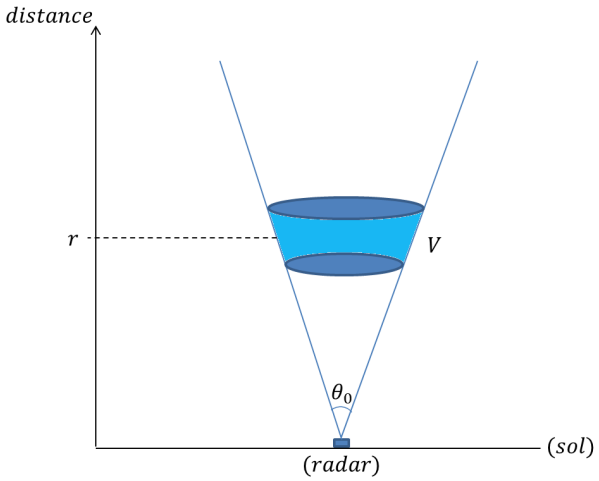

FIGURE I.7 – Visualisation du volume (noté V) sondé par un radar, pour une porte et une ouverture de faisceau données. On représente ici un radar à visée verticale (maximum de densité de puissance émis à la verticale). La distance r caractérisant la porte du radar considérée est définie sur la Figure I.5. L'ouverture du faisceau θ_0 est définie sur la Figure I.6.

rétrodiffusion d'une goutte d'indice p). Pour définir σ, on considère une cible "idéale", renvoyant toute la puissance interceptée, et de manière isotrope. On définit alors σ comme la section que devrait avoir cette cible idéale pour capter une puissance lui permettant de renvoyer au radar la puissance qu'il reçoit effectivement. En supposant les cibles réparties de manière homogène dans le volume sondé par le radar (pour une porte et une date données), on définit ensuite la réflectivité η (m^2/m^3) comme la somme des sections efficaces de rétrodiffusion des différentes cibles normalisée par le volume sondé. En référençant chaque cible présente par un indice p, et en notant V le volume sondé, on obtient :

$$\eta = \frac{1}{V} \sum_{p \in V} \sigma_p \tag{I.12}$$

L'équation radar permet ensuite de relier la puissance reçue (P, toujours pour une porte et une date données, en W), à cette réflectivité η. Elle s'écrit :

$$P = CL\frac{\eta}{r^2} \tag{I.13}$$

r est toujours la distance au radar de la porte considérée (en m). L est un coefficient caractérisant l'atténuation du signal due aux cibles présentes dans les portes inférieures (seule une partie du signal atteint la distance r). C est la constante radar, en Wm^3, englobant tous les termes affectant la relation puissance reçue - réflectivité et ne dépendant que des caractéristiques du radar utilisé. Cette constante dépendra de la puissance émise (P_e), de la longueur d'onde λ, du gain de l'antenne (caractérisant sa directionnalité), de l'ouverture du faisceau (θ_0, ϕ_0), du temps d'impulsion τ_{on} et des pertes dues à l'appareil lui-même. Cette constante est généralement déterminée empiriquement, pour chaque radar. On calibre le radar en visant des cibles proches et dont le comportement électromagnétique est connu (c'est-à-dire dont on connaît la réflectivité η).

TABLE I.1 – Bandes de fréquence micro-ondes classiquement utilisées pour la télédétection de la pluie ou des nuages. On indique, pour chaque bande, la lettre la désignant, les gammes de fréquence et de longueur d'onde λ correspondant, la limite haute de l'approximation Rayleigh pour la diffusion par des particules sphériques (voir texte) et les instruments utilisés dans ce travail émettant ou recevant dans cette bande.

bande	fréquence (GHz)	λ	limite Rayleigh	instrument
S	$2-4$	$7.5-15cm$	$\approx 10mm$	Bollène (opérationnel MF)
C	$4-8$	$3.75-7.5cm$	$\approx 5mm$	
X	$8-12$	$2.5-3.75cm$	$\approx 3mm$	
Ku	$12-18$	$1.6-2.5cm$	$\approx 1.8mm$	capteur d'atténuation
K	$18-27$	$1.13-1.66cm$	$\approx 1.2mm$	MRR
Ka	$27-40$	$0.75-1.13cm$	$\approx 0.9mm$	KAZR
W	$75-110$	$2.7-4mm$	$\approx 0.3mm$	WACR, BASTA

Diffusion des micro-ondes par des particules sphériques On suppose dans cette partie que les cibles sont des gouttes d'eau liquide (pluie ou nuage) sphériques. Dans ce cas-là, les interactions entre ces cibles et l'onde électromagnétique émise par le radar sont décrites par la théorie de Mie. Celle-ci permet de calculer numériquement les coefficients de diffusion et d'absorption d'une goutte en fonction de son diamètre, de la longueur de l'onde λ, et de l'indice de réfraction complexe de l'eau (lui-même fonction de la température et de la longueur d'onde). Pour une goutte de diamètre D, on peut ainsi notamment calculer la surface efficace de rétrodiffusion $\sigma(D)$.

Le Tableau I.1 présente les différentes bandes de fréquence utilisées dans ce travail (voir légende).

La Figure I.8 présente, pour quelques longueurs d'onde classiquement utilisées pour la télédétection pluie/nuages, l'évolution de la surface efficace de rétrodiffusion avec le diamètre des particules diffusantes. On remarque qu'aux hautes fréquences (petites longueurs d'onde, typiquement en bande W), la section efficace oscille en fonction du diamètre (à partir de $1mm$ environ). Au contraire, aux basses fréquences (bandes C et X), la relation est strictement croissante, et plutôt régulière, dans la gamme de diamètres considérée.

Il se trouve que les équations de la théorie de Mie peuvent être approximées et singulièrement simplifiées lorsque le diamètre des particules diffusantes est très petit par rapport à la longueur d'onde. On entre alors dans le cadre de l'approximation de Rayleigh, que l'on considère souvent valable pour des diamètres inférieurs à environ un dixième de la longueur d'onde. La $4^{\text{ème}}$ colonne du Tableau I.1 donne une limite haute (c'est-à-dire le diamètre maximal) approximative pour le domaine de validité de cette approximation pour différentes bandes de fréquence. On remarque qu'en bande S ou C, l'approximation est valable sur la majeure partie du spectre de diamètres couvert dans la pluie. Au-delà d'environ $8GHz$, l'approximation n'est plus adéquate pour la pluie.

Vus les instruments utilisés dans ce travail, nous ne travaillerons pas ici avec cette approximation. Nous allons néanmoins donner sa paramétrisation, car la plupart des radars

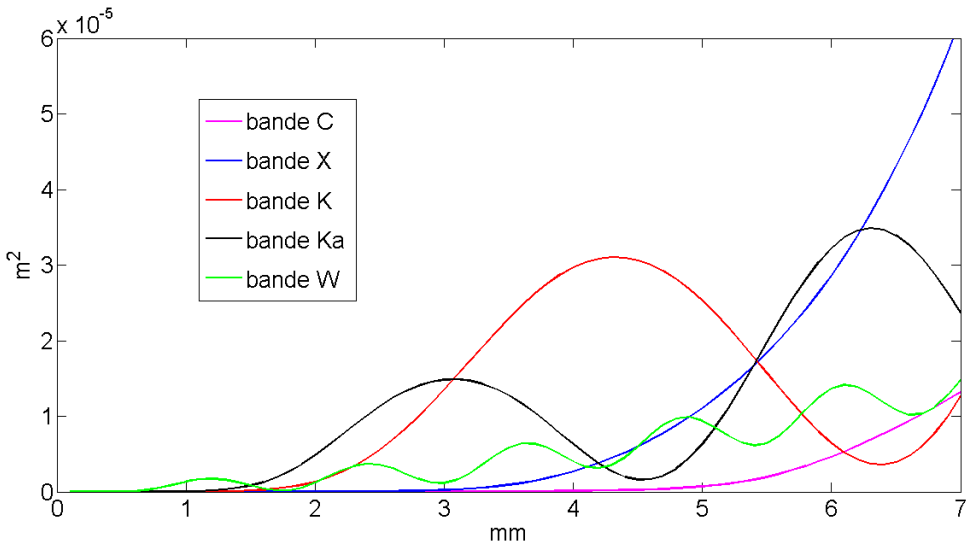

FIGURE I.8 – Pour différentes longueurs d'onde, représentation des surfaces efficaces de rétrodiffusion de particules sphériques en fonction de leur diamètre, calculées dans le cadre de la théorie de Mie. Les fréquences utilisées sont : $6GHz$ (bande C), $10GHz$ (bande X), $25GHz$ (bande K), $35GHz$ (bande Ka), $90GHz$ (bande W).

opérationnels pour la pluie sont en bande S ou C et utilisent l'approximation Rayleigh, si bien que certaines grandeurs classiquement utilisées en météorologie radar (Z_e, voir l'Eq.(I.18) plus loin) sont définies d'après sa paramétrisation.

Sous l'approximation Rayleigh, la surface efficace de rétrodiffusion $\sigma(D)$ d'une particule sphérique de diamètre D éclairée par une onde de longueur d'onde λ s'écrit :

$$\sigma(D) = \frac{\pi^5}{\lambda^4}|K|^2 D^6 \tag{I.14}$$

Où $|K|^2$ est le facteur diélectrique de l'eau, fonction de son indice de réfraction complexe, et généralement considéré constant (à 0.93) sur les gammes de fréquence et de températures considérées.

Lien réflectivité / DSD En notant D_p le diamètre de la p-ième goutte présente dans le volume V, on a donc :

$$\eta = \frac{1}{V}\sum_{p\in V}\sigma(D_p) \tag{I.15}$$

Pour un nombre de gouttes assez grand, on peut également utiliser la distribution en taille des gouttes de pluie $n(D)$ définie dans la section 1. En supposant cette distribution homogène dans le volume V, on a :

$$\eta = \int_D n(D)\sigma(D)dD \tag{I.16}$$

Pour rappel, η est en m^2/m^3 (surface par unité de volume).

On définit par ailleurs Z, le facteur de réflectivité radar, par :

$$Z = 10^{18} \int_D n(D) D^6 dD \tag{I.17}$$

Avec n en m^{-4}, et D et dD en m, on obtient un Z en mm^6/m^3. En échelle logarithmique, on définit le décibel Z (dBZ) comme l'unité de la grandeur $10log_{10}(Z/Z_{ref})$, avec $Z_{ref} = 1mm^6m^{-3}$.

On définit enfin Z_e, le facteur de réflectivité radar équivalent, par :

$$Z_e = 10^{18} \frac{\lambda^4}{\pi^5 |K|^2} \eta \tag{I.18}$$

On obtient là encore un Z_e en mm^6/m^3.

Si l'on se place dans le cadre de l'approximation Rayleigh (Eq. (I.14)) que l'on reporte dans l'Eq.(I.16), on obtient :

$$\eta = \int_D n(D) \frac{\pi^5}{\lambda^4} |K|^2 D^6 dD \tag{I.19}$$

En reportant cette expression dans Eq.(I.18), on a :

$$Z_e = Z \tag{I.20}$$

Dans le cadre de l'approximation Rayleigh, on a donc introduit une seule nouvelle grandeur, notée Z ou Z_e. Cette grandeur permet de comparer des mesures faites à différentes longueurs d'ondes : on s'affranchit de la dépendance de η à λ, et Z ne dépend plus que de la DSD (on note néanmoins que les volumes sondés, notamment, dépendent toujours de l'instrument, et donc de la longueur d'onde).

Si l'on est dans le cadre de la théorie de Mie, σ n'est plus inversement proportionnel à λ^4, Z et Z_e ne correspondent plus et on ne s'affranchit pas de la dépendance à λ. Dans ce travail, où nous ne serons pas dans le cadre de l'approximation de Rayleigh, nous utiliserons Z et Z_e comme des grandeurs différentes, la première permettant de caractériser la DSD, la seconde la mesure radar (au même tire qu'η qui lui est proportionnelle).

Atténuation On a vu (Eq.(I.13)) que relier réflectivité et puissance reçue par le radar implique d'estimer l'atténuation du signal durant sa traversée de l'atmosphère jusqu'à la porte considérée. Cette atténuation peut venir de différents corps : gaz (oxygène, vapeur d'eau), ou aérosols (nuages, pluie, ...). La Figure I.9, extraite de Brisseau (2005) présente l'atténuation (ou affaiblissement) linéaire produite par l'oxygène atmosphérique et par la vapeur d'eau (pour différentes valeurs d'humidité relative) en micro-ondes (au niveau de la mer et à $20\,^\circ C$, il s'agit donc plutôt de majorants de l'atténuation sur des colonnes atmosphériques).

Les radars météorologiques émettent à des fréquences plus ou moins éloignées des pics

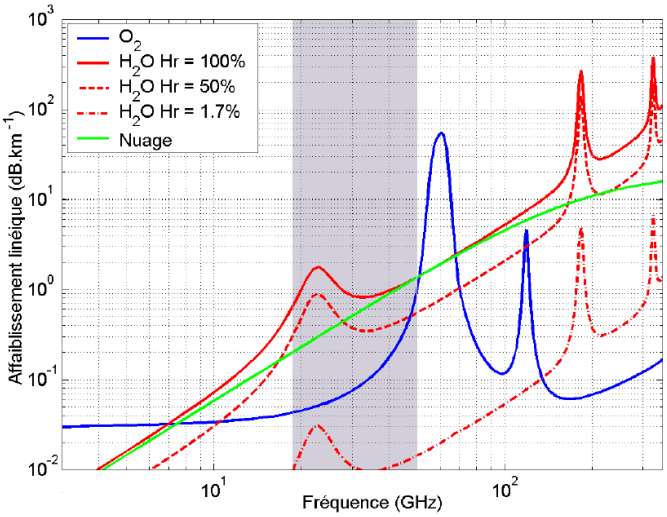

FIGURE I.9 – Selon la fréquence d'éclairage, atténuation spécifique des ondes due à l'oxygène atmosphérique (bleu) et à la vapeur d'eau pour différentes humidités relatives (rouge), à $1013HPa$ et $20\,^{\circ}C$. Figure adaptée de Liebe et al. (1993) et Brisseau (2005).

d'absorption de la vapeur d'eau et de l'oxygène (c'est-à-dire en-dessous de $20GHz$, ou, plus haut, entre 25 et $40GHz$ et autour de $90-100GHz$. On se concentre donc dans la suite sur l'atténuation du signal due aux gouttes d'eau liquide, prépondérante à ces fréquences dans la pluie.

De même que pour la rétrodiffusion des ondes par des particules, on définit une section efficace d'extinction, σ_{ext}, en m^2, qui est la somme des contributions des effets dûs à la diffusion et l'absorption. Mathématiquement, elle est calculée, par exemple pour des particules sphériques dans la théorie de Mie, comme la somme des sections efficaces d'absorption et de diffusion (Brisseau (2005)). Le coefficient spécifique d'atténuation κ (pour une date donnée et une porte radar données), en m^{-1}, est ensuite défini comme la somme des contributions de chaque particule, normalisée par le volume sondé. En utilisant les mêmes notations que pour l'Eq.(I.12), on a :

$$\kappa = \frac{1}{V}\sum_{p\in V}\sigma_{ext_p} \tag{I.21}$$

Et comme dans l'Eq.(I.16), en continu, et en utilisant la paramétrisation de Mie fournissant l'expression de σ_{ext} pour une particule sphérique de diamètre D, on peut utiliser la DSD et exprimer κ (pour l'altitude z) par :

$$\kappa(z) = \int_D n(z,D)\sigma_{ext}(D)dD \tag{I.22}$$

κ est toujours ici formellement en m^{-1}, et est souvent exprimé en dB/km.

Le facteur d'atténuation (contribution de la pluie au coefficient L de l'équation (I.13)), noté

$\mathcal{A}(r)$ (r indiquant la distance au radar), s'écrit alors :

$$\mathcal{A}(r) = e^{-2\int_0^r \kappa(z)dz} \tag{I.23}$$

On note que $\mathcal{A}$ est une grandeur sans unité et est généralement exprimé en décibels (dB).

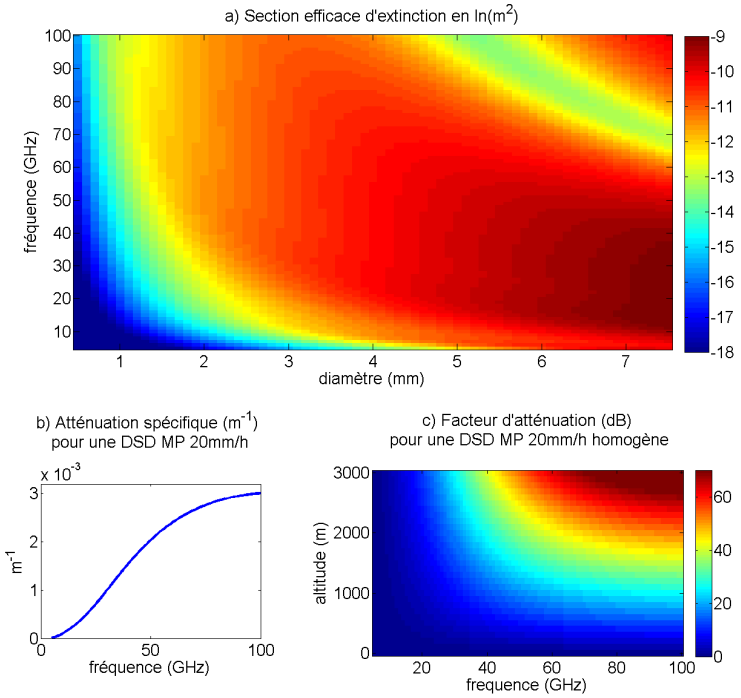

FIGURE I.10 – a) Influence du diamètre et de la fréquence sur les surfaces efficaces d'extinction. b) Atténuation spécifique produite par une DSD de Marshall-Palmer (avec un taux de pluie de $20mm/h$) selon la fréquence. c) Atténuation totale produite par cette même DSD selon l'altitude et la fréquence.

La Figure I.10a représente la surface efficace d'extinction σ_{ext}, découlant de la théorie de Mie, en fonction de la fréquence d'éclairage et du diamètre de la particule diffusante. On remarque que les grosses gouttes atténuent davantage le signal pour les fréquences basses et intermédiaires ($< 50GHz$). Les petites, en revanche, atténuent d'autant plus le signal que celui-ci est à haute fréquence. Dans la pluie, les petites gouttes sont généralement significativement plus nombreuses que les grosses (typiquement selon des distributions exponentielles). Au final, les hautes fréquences seront par conséquent davantage atténuées que les basses. Les Figures I.10b et c illustrent ce phénomène. On suppose dans la Figure I.10b que l'on a une DSD (n) exponentielle, paramétrée par les expressions de Marshall-Palmer (MP, Marshall and Palmer (1948)) pour un taux de pluie de $20mm/h$. On représente l'atténuation spécifique (κ) résultante en fonction de la fréquence. On voit que l'atténuation augmente continuellement avec la fréquence, et que la pente maximale se trouve aux alentours de $30GHz$. Sur la Figure I.10c, on suppose une colonne d'eau avec cette même DSD à toutes altitudes. On représente, toujours en fonction de la fréquence, le facteur d'atténuation ($\mathcal{A}$) en

dB. On voit qu'en dessous de $10-15GHz$, même à $3000m$ d'atitude, l'atténuation reste faible ($<3dB$). Au-dessus de $40-50GHz$ en revanche, le signal finit par être complètement atténué ($>40-50dB$ à $3km$). C'est pourquoi les radars opérationnels de pluie, qui doivent couvrir de vastes zones, travaillent généralement à basse fréquence (bandes C, S), et à d'autant plus basse fréquence qu'ils se trouvent dans des zones subissant des pluies intenses. Les radars nuages, opérant lorsqu'il ne pleut pas et sur de très petites gouttes, travaillent au contraire à plus haute fréquence (typiquement 35 ou $90GHz$). Ces radars peuvent aussi être utilisés pour des pluies pas trop fortes. Dans ces cas-là, ils sont utilisés en visée verticale pour atteindre les quelques kilomètres de nuages au-dessus du sol ou la couche de fonte, la distance couverte dans ces cas-là étant bien sur plus petite.

2.3.2 Visée verticale : le spectres Doppler

Principe Lorsqu'une onde électromagnétique est diffusée par une particule en mouvement, la fréquence de l'onde rétrodiffusée est modifiée par effet Doppler-Fizeau. Si la particule a une vitesse radiale (projection de sa vitesse sur l'axe de propagation de l'onde) v_{part}, la fréquence du signal reçu en retour par le radar sera affectée d'un facteur $1-2v_{part}/c$, avec c la vitesse de la lumière, soit une écart de $2v_{part}/\lambda$, avec λ la longueur de l'onde émise.

Considérons maintenant un radar à visée verticale. En supposant que notre radar a une ouverture de faisceau ($\theta_0 = \phi_0$) suffisamment petite, on peut supposer que les vitesses radiales des particules mentionnées plus haut sont des vitesses verticales, c'est-à-dire des vitesses de chute. En pratique, le radar reçoit, à chaque tir et pour chaque porte, un signal complexe (amplitude et phase). La phase de ce signal est affectée par l'effet Doppler induit par les gouttes qui tombent. En effectuant une transformée de Fourier d'un certain nombre d'échantillons (chacun résultant d'un tir), on obtient le spectre de puissance en fonction du décalage Doppler en fréquence, et donc en fonction de la vitesse de chute des particules diffusantes.

Puisque l'on mesure des phases, celles-ci sont forcément comprises dans $[0:2\pi]$ et il existe une vitesse limite de non ambiguïté (vitesse de Nyquist v_{Nyq}) au-delà de laquelle apparaît un repliement spectral. Cette vitesse de Nyquist dépend des caractéristiques du radar et s'écrit : $v_{Nyq} = \lambda/2\tau_{off}$, avec τ_{off} le temps interpulses. Le nombre d'impulsions N_{fft} utilisées pour calculer la FFT (64, 128, 256, ...) donne d'une part le nombre de bins (points de discrétisation) de la distribution d'énergie obtenue. v_{Nyq}/N_{fft} donne alors la résolution en vitesse du spectre. D'autre part, $N_{fft} * (\tau_{off} + \tau_{on})$ (τ_{on} la durée d'une impulsion) donne le temps nécessaire à l'obtention d'un spectre. On note qu'il est par ailleurs généralement nécessaire de moyenner quelques spectres successifs pour avoir des données plus régulières et moins bruitées.

Réflectivité spectrale On définit, pour une date et une porte radar donnée, la réflectivité spectrale $\eta_v(v)$, en $m^2/m^3/(m/s)$ (surface de rétrodiffusion par unité de volume et de vitesse), comme la réflectivité induite par les cibles se déplaçant à la vitesse v. On a par conséquent la relation :

$$\eta = \int_v \eta_v(v)dv \tag{I.24}$$

Dans la suite, on parlera indistinctement de réflectivité spectrale ou de spectre Doppler pour désigner la grandeur η_v.

On reprend ensuite l'équation radar (Eq (I.13)), en ne considérant que la puissance $P(z, v)$ (en $W/(m/s)$) rétrodiffusée par les cibles se déplaçant à la vitesse v (à l'altitude z), correspondant à la réflectivité spectrale η_v. En considérant un radar à visée verticale, principalement atténué par la pluie, on obtient :

$$P(z, v) = C\mathcal{A}(z)\frac{\eta_v(v)}{z^2} \tag{I.25}$$

Avec z l'altitude. C est toujours la constante radar, et $\mathcal{A}(z)$ l'atténuation du signal, définie par l'Eq.(I.23) (qui remplace le L de l'Eq.(I.13)).

Si l'on veut pouvoir déterminer la distribution en taille de gouttes, il nous faut établir un lien entre cette dernière, dépendant du diamètre, et la réflectivité spectrale, dépendant elle de la vitesse. Pour cela on définit d'abord une réflectivité spectrale "hydrologique" en diamètre, $\eta_{D,hy}$, en $m^2/m^3/m$ (surface par unité de volume et de diamètre), par :

$$\eta_{D,hy}(D) = n(D)\sigma(D) \tag{I.26}$$

Si l'on suppose qu'il existe une relation analytique entre le diamètre et la vitesse des gouttes s'écrivant $D(v)$, on peut définir une réflectivité spectrale "hydrologique" en vitesse, $\eta_{v,hy}$, en $m^2/m^3/(m/s)$ (surface par unité de volume et de vitesse), par :

$$\eta_{v,hy} = \eta_{D,hy}\frac{dD}{dv} \tag{I.27}$$

La question qui se pose ensuite est celle du lien entre cette grandeur, fonction de la DSD, de la longueur d'onde (via Mie dans l'expression du σ inclus dans $\eta_{D,hy}$) et de la relation vitesse/diamètre, et le η_v défini plus haut et plus directement lié à la mesure radar. La relation vitesse/diamètre étant le plus souvent basée sur une expression de la vitesse terminale de chute, il se trouve que l'expression de l'Eq.(I.27) est insuffisante pour caractériser la mesure. En effet, elle ne prend pas en compte certains phénomènes affectant directement les spectres Doppler, à savoir le vent vertical, la turbulence (et plus généralement les écarts résiduels de la vitesse de chute des gouttes à leur vitesse terminale théorique), et le bruit instrumental. On définit alors une nouvelle grandeur, $\eta_{v,simu}$, toujours en $m^2/m^3/(m/s)$, par (Tridon and Battaglia (2015)) :

$$\eta_{v,simu}(v + w) = \eta_{v,hy}(v) * \eta_{air}(v + w) + \eta_{noise} \tag{I.28}$$

Avec w le vent vertical et η_{air} et η_{noise} des gaussiennes représentant la turbulence atmosphérique et le bruit instrumental (voir ci-après).

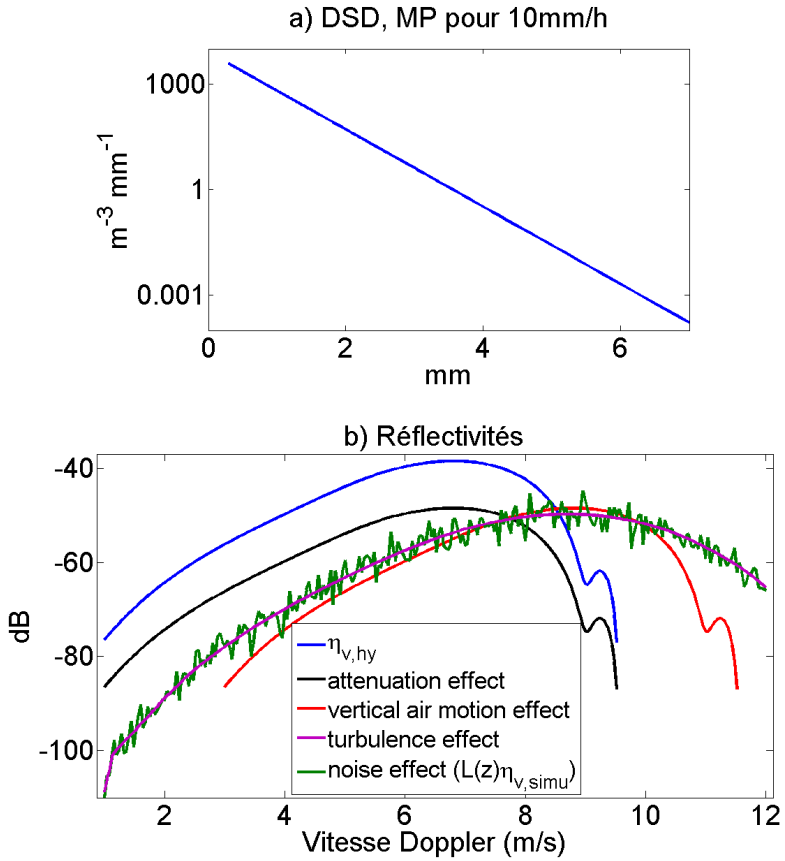

FIGURE I.11 – DSD simulée (haut) et réflectivités spectrales prenant en compte divers phénomènes. Voir texte pour la description détaillée.

La Figure I.11 présente les effets de chaque terme de cette équation. Sur la FigureI.11a, on présente une DSD exponentielle (respectant la paramétrisation de Marshall and Palmer (1948) pour un taux de pluie de $10mm/h$) à l'origine des différents spectres présentés en-dessous. En bleu sur la Figure I.11b, on représente un spectre hydrologique ($\eta_{v,hy}$) (pour un éclairage à $35GHz$), en utilisant la relation vitesse/diamètre de Atlas et al. (1973). On a ensuite pris en compte l'atténuation due aux couches inférieures de l'atmosphère, arbitrairement fixée à $10dB$. Il s'agit de la courbe noire représentant $\mathcal{A}(z)\eta_{v,hy}$.

En considérant l'influence du vent vertical (paramètre w de l'Eq.(I.28)), avec $w = 2m/s$, on obtient la courbe rouge de la FigureI.11b. On considère que w modifie de manière homogène la vitesse de toutes les cibles présentes dans le volume sondé. Une goutte ayant une vitesse théorique terminale de chute v (fonction de D) tombera en fait à la vitesse $v + w$. Le spectre se trouve donc translaté vers les grandes vitesses si le vent va vers le bas (positif), vers les faibles vitesses sinon.

On considère ensuite le fait que, au sein d'un volume et d'un temps d'intégration donné, toutes les gouttes d'un diamètre donné ne tombent pas à la même vitesse. On suppose plutôt ici que la variabilité est provoquée par la turbulence (variations de la vitesse verticale du vent sur de courtes périodes temporelles). Celle-ci est représentée analytiquement en convoluant le

spectre précédent avec une gaussienne centrée et possédant un écart-type σ_{air} caractérisant son intensité (représentation que l'on retrouve par exemple dans Gossard (1994), Williams (2002) ou Tridon and Battaglia (2015)). Dans l'Eq.(I.28), on utilise :

$$\eta_{air}(v) = \frac{1}{\sqrt{2\pi}\sigma_{air}} e^{-\frac{v^2}{2\sigma_{air}^2}} \tag{I.29}$$

Sur la Figure I.11b, on a aussi représenté en mauve le résultat de cette convolution avec $\sigma_{air} = 0.9m/s$. Cela provoque un lissage des spectres. On note :

- qu'on suppose ici que les variations locales du vent vertical se répercutent immédiatement sur la vitesse de chute des gouttes. On peut penser qu'en fait les grosses gouttes ont un temps d'inertie plus long et qu'elles sont moins sensibles à la turbulence. Cela se traduirait par un η_{air} non gaussien mais plus lourd du côté des petites vitesses de chute.

- que les gouttes, indépendamment du vent, ne tombent pas toutes à leur vitesse terminale de chute, notamment lorsqu'elles viennent de se "transformer" (suite à une auto-cassure ou un choc entre 2 gouttes). Là encore, il paraîtrait raisonnable de modéliser ce phénomène par une convolution, mais pas avec une gaussienne.

Finalement, on a ajouté sur la Figure I.11b (en vert), pour chaque bin de vitesse Doppler, un bruit blanc (gaussien) instrumental d'intensité $2dB$ (le η_{noise} de l'Eq.(I.28)). La courbe verte correspond donc au produit $\mathcal{A}(z)\eta_{v,simu}$, avec les notations des Eqs. (I.28) et (I.23).

On supposera dans la suite que $\mathcal{A}(z)\eta_{v,simu}$ est une bonne approximation du spectre mesuré par le radar $\mathcal{A}(z)\eta_v$ (Eq. (I.25)).

Moments des spectres Doppler On a vu que l'on pouvait exprimer la réflectivité η à partir de la réflectivité spectrale (Eq.(I.24), $\eta = \int_v \eta_v(v)dv$). La réflectivité correspond aussi au moment d'ordre 0 $\mathcal{M}^0_{\eta_v}$ du spectre Doppler. Le moment d'ordre 1 est la vitesse moyenne des cibles pondérée par leurs surfaces efficaces de rétrodiffusion. Ce moment, $\mathcal{M}^1_{\eta_v}$, aussi noté W, sera dans la suite désigné par "vitesse Doppler moyenne" (en m/s), que l'on écrit :

$$W = \frac{1}{\eta}\int_v \eta_v v dv \tag{I.30}$$

Le moment centré d'ordre 2 ($\mathcal{M}^2_{\eta_v}$) est la variance du spectre Doppler. Il caractérise donc la dispersion des vitesses. Dans la suite, on fera appel à l'écart-type associé LS (en m/s), désigné par "largeur spectrale", que l'on écrit :

$$LS = \sqrt{\frac{1}{\eta}\int_v \eta_v \left(v - W\right)^2 dv} \tag{I.31}$$

2.3.3 Quelques instruments utilisés dans cette thèse

Radars scannants Dans ce travail, nous n'utiliserons des radars scannants qu'en validation. Les données d'entrées du travail décrit dans le chapitre V ne proviennent en effet que de radars

à visée verticale.

Dans le chapitre IV, nous utiliserons le produit pluie du radar Météo France de Bollène. Ce radar opérationnel est un radar scannant en bande S faisant partie du réseau national de télédétection ARAMIS de Météo France (réseau décrit notamment dans Parent-du Châtelet (2003)). Il scanne l'espace avec des élévations (angles par rapport au sol) successives de 1.8 °, 1.2 ° et 0.8 °. Il répète un tel cycle toutes les 5 minutes. Son ouverture de faisceau est de 1.28 ° (Tabary (2007)). La Figure I.12 présente une photographie de l'appareil.

FIGURE I.12 – Le radar opérationnel en bande S de Météo France de Bollène. Photographie issue du site de l'institut (http://education.meteofrance.fr/la-mesure-des-precipitations-par-radar

À partir des mesures de réflectivité dans différents volumes (à différentes distances du radar et altitudes), Météo France fournit un produit pluie au sol à une résolution de 5 minutes et $1km^2$. L'algorithme permettant d'évaluer cette lame d'eau est décrit en détails dans Tabary (2007). Il est basé sur l'utilisation d'une relation $Z - R$, qui permet empiriquement de transformer les réflectivités en taux de pluie à partir d'un a priori sur la DSD (c'est-à-dire en supposant la DSD connue). Dans l'algorithme utilisé, on suppose une DSD exponentielle de Marshall and Palmer (1948), c'est-à-dire entièrement paramétrisée par le taux de pluie. On obtient ensuite la relation :

$$Z = 200R^{1.6} \tag{I.32}$$

avec Z le facteur de réflectivité radar en mm^6m^{-3} et R le taux de pluie en mm/h.

On note que de très nombreuses relations du type $Z = aR^b$ existent dans la littérature, avec a et b des coefficients dépendant du type de précipitation (donc de la saison, de la latitude, etc.). Les modifications de la DSD avec l'altitude lors de la chute des gouttes (voir chapitre II) sont aussi susceptibles d'affecter ces relations, d'autant que les mesures effectuées par les radars scannants sont réalisées d'autant plus haut que l'on s'éloigne du radar (ils ne tirent pas horizontalement et la Terre est ronde). Elles sont ensuite elles-même converties en taux de pluie selon des relations généralement étalonnées avec des mesures au sol, puis sont projetées au sol. Nous ne nous attarderons pas sur ces relations ici. Dans les faits, le perfectionnement (ou le dépassement) de telles relations serait une perspective naturelle du travail présenté dans le chapitre V.

Radars à visée verticale Nous allons à présent décrire rapidement les quatre radars à visée verticale utilisés dans le chapitre V. Les caractéristiques principales de ces instruments sont données dans le Tableau I.2. Leurs photographies sont présentées sur la Figure I.13.

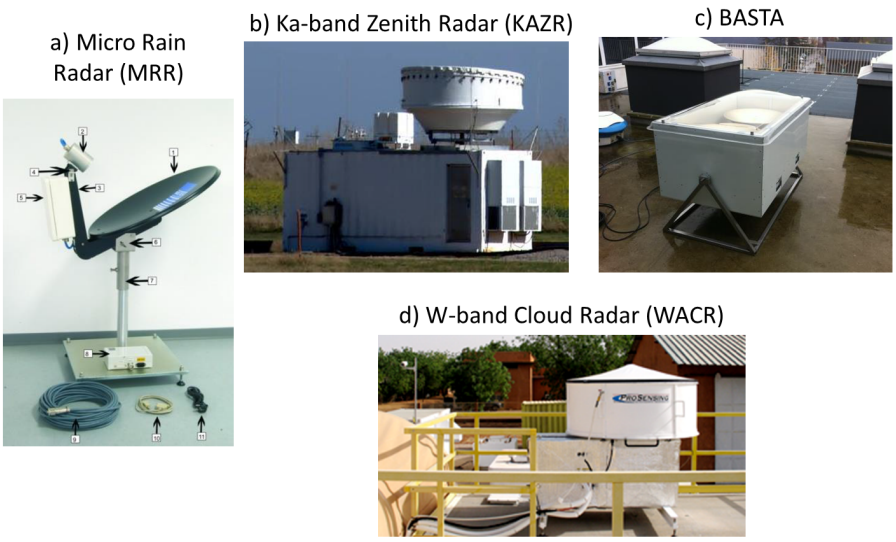

FIGURE I.13 – Photographies des 4 radars à visée verticale utilisés dans ce travail. a) le Micro-Rain Radar (extrait de METEK (1998)). b) le KAZR (extrait de Widener et al. (2012)). c) le radar BASTA (extrait du site du constructeur : http://www.meteomodem.com/basta.php). d) le WACR.

TABLE I.2 – Caractéristiques principales des 4 radars à visée verticale utilisés dans ce travail. Voir commentaires dans le texte.

radar		**MRR**	**KAZR**	**BASTA**	**WACR**
technique		FMCW	pulses	FMCW	pulses
puissance maximale		$50mW$	$200W$	$0.7W$	$119W$
fréquence (GHz)		24	35	94	95
longueur d'onde λ (mm)		12.5	8.6	3.2	3.2
ouverture de faisceau θ_0		2°	0.3°	0.4°	0.37°
v_{Nyq}(m/s)		11	11.9 ou 12.1	10	12.04
résolution Doppler (m/s)		0.19	0.047 ou 0.024	-	0.024
largeur des portes (m)		100	30	25, 100 ou 200	30
N_{fft}		64	256 ou 512	-	512
temps d'intégration		$10s$	$\approx 2s$	$1-10s$	$\approx 2s$
diamètre antenne		$60cm$	$2m$	$60cm$	$61cm$
$200m$	diamètre sondé (m)	7.6	3.1	2.0	1.9
	volume sondé (m^3)	4500	730	310	280
$2000m$	diamètre sondé (m^3)	70	13	15	14
	volume sondé (m^3)	390000	12000	17000	14000

Le premier de ces instruments est le Micro-Rain Radar (MRR), développé par la société METEK (http://metek.de/product-group/micro-rain-radar-24-ghz/). Il s'agit d'un

radar à $24GHz$ utilisant la technique Frequency Modulation Continuous Wave (FMCW). En pratique, il n'envoie pas des impulsions d'énergie mais une onde continue, dont la fréquence est modulée de $1.5MHz$ par des rampes linéaires avec une période de $0.512ms$. Cette technique permet en particulier d'utiliser des puissances plus faibles ($50mW$ en l'occurrence), et donc d'abaisser le coût de l'appareil. Un premier algorithme de traitement des données brutes transforme le signal brut en réflectivités spectrales (η_v, $m^2/m^3/(m/s)$), en appliquant l'équation radar et un algorithme d'estimation du bruit. Cet algorithme est décrit dans METEK (1998). Tridon (2011) l'a comparé à d'autres algorithmes existants et en a conclu qu'il n'est pas le meilleur algorithme pour ces mesures, dont il sous-estime généralement le niveau de bruit. Dans le chapitre V, nous utiliserons comme données d'entrées ces réflectivités spectrales. Nous appliquerons néanmoins un opérateur logarithmique qui, bien que conçu dans un autre but, devrait limiter les problèmes dûs au bruit sur les zones du spectre Doppler à faible énergie. Les spectres Doppler ont une résolution en vitesse de $0.19m/s$ et couvrent des vitesses de chute jusqu'à $11m/s$, ce qui sera généralement suffisant pour éviter les repliements spectraux sur les cas relativement calmes traités dans ce travail. Les spectres obtenus sont intégrés sur $10s$, et la largeur des portes est de $100m$ (voir Tableau I.2).

Le logiciel de traitement développé par le constructeur propose également un produit atténuation ainsi qu'un produit DSD, et par conséquent des produits pluie, contenu en eau liquide, et facteur de réflectivité radar (moment d'ordre 6 de la DSD). L'algorithme d'estimation de l'atténuation, préalable à toute estimation de DSD, est présenté dans METEK (1998). La DSD est ensuite calculée à partir de la réflectivité spectrale en inversant l'équation (I.26). Cela revient donc à négliger les effets du vent vertical et de la turbulence, et à assimiler la réflectivité hydrologique $\eta_{v,hy}$ à la réflectivité effectivement mesurée η_v. Cet algorithme utilise la relation vitesse-diamètre de Atlas et al. (1973) corrigée par l'altitude (Foote and Du Toit (1969)). Peters et al. (2005), notamment, évaluent ces produits et les conséquences induites par le fait de négliger le vent vertical.

Le travail décrit au chapitre V visant précisément à estimer des profils verticaux de DSD en tenant compte de modèles dynamiques de propagation, du vent vertical, et de la turbulence, nous utiliserons ces produits pour les comparer à nos restitutions.

Le second instrument utilisé est un radar en bande Ka, le KAZR (KA-band Zenith Radar). Il est développé et utilisé par l'Atmospheric Radiation Measurement (ARM) du département américain de l'énergie (`http://www.arm.gov/`) et construit par la société Millitech (`http://www.millitech.com/`). Ses caractéristiques sont fournies dans Widener et al. (2012) et le papier Kollias et al. (2014) revient sur ses performances dans la pluie.

Il s'agit d'un radar pulsé à $35GHz$, qui enregistre les spectres Doppler. Le développeur fournit les signaux bruts (issus de la FFT), c'est-à-dire, pour chaque date et chaque porte, la quantité d'énergie par bin de vitesse Doppler (en W), ainsi que la constante radar de calibration et un produit réflectivité (en dBZ), obtenu en sommant l'énergie sur toute la gamme de vitesses, en appliquant la constante radar et en appliquant un algorithme de correction du bruit. Pour

notre part, nous utiliserons les réflectivités spectrales, obtenues en appliquant la constante radar aux sorties brutes.

Par rapport au MRR, le KAZR a une ouverture de faisceau, une résolution en vitesse des spectres Doppler, et une résolution verticale sensiblement plus fines (voir Tableau I.2). Cela lui permet de sonder des volumes nettement plus petits. Par exemple, à $200m$ d'altitude, le KAZR sonde $730m^3$ contre $4500m^3$. A $2000m$ d'altitude, il sonde $12000m^3$, contre $390000m^3$ pour le MRR. Il couvre des vitesses jusqu'à $12m/s$, ce qui, là aussi, est généralement suffisant pour la pluie.

Le troisième instrument utilisé est le radar BASTA, développé au LATMOS et commercialisé par la société Meteomodem (http://www.meteomodem.com/fr/). Il s'agit d'un radar à $94GHz$, conçu principalement pour l'étude des nuages mais pouvant servir pour des pluies relativement faibles (lorsque l'atténuation n'est pas trop forte). Il n'enregistre pas les spectres Doppler mais seulement leurs deux premiers moments (réflectivité η et vitesse Doppler moyenne W). Il utilise la technique FMCW, et a pour le reste (ouverture de faisceau, largeur des portes, résolution temporelle) des caractéristiques assez similaires à celles du KAZR. Il servira en validation dans le chapitre V de ce travail. On peut se référer à Delanoë et al. (2014) pour plus de détails sur cet instrument.

Le dernier radar à visée verticale présenté est le WACR ("W-band Arm Cloud Radar", voir Widener and Mead (2004)). C'est également un radar en bande W ($95GHz$) possédant des caractéristiques très proches de celles du $KAZR$. Il enregistre également les spectres Doppler et est utilisé notamment par Tridon and Battaglia (2015), en combinaison du $KAZR$, pour inverser les spectres Doppler. Dans cette étude, nous l'utiliserons déployé aux côtés du $KAZR$ en Finlande pour validation.

2.4 Capteurs d'atténuations

De nouveaux systèmes d'observation des précipitations ont récemment vu le jour. Ces systèmes restent néanmoins actuellement en phase de recherche. Ils exploitent l'atténuation due à la pluie subie par des ondes électromagnétiques. Un certain nombre d'études ont été réalisées en utilisant des liaisons de télécommunications Terre-Terre (Leijnse et al. (2007), Leijnse et al. (2008), Schleiss and Berne (2010), Chwala et al. (2012), Overeem et al. (2013)). Maitra and Chakravarty (2005) et Barthès and Mallet (2013) travaillent avec des liaisons Terre-satellite.

C'est ce dernier système que nous présentons ici.

On a vu que les radars étaient des instruments actifs, émettant une onde électromagnétique et captant l'énergie rétrodiffusée par les gouttes de pluie. On a vu que plus on s'éloignait du radar, plus l'onde était atténuée par les gouttes de pluie (Eq.(I.22)), dans le cadre de la théorie de la diffusion de Mie. Pour un radar, cette atténuation est généralement considérée

comme un facteur limitant, à estimer d'abord pour pouvoir ensuite évaluer la réflectivité. Nous allons présenter ici le fonctionnement d'un appareil opportuniste utilisant l'atténuation d'ondes électromagnétiques pour estimer des taux de pluie.

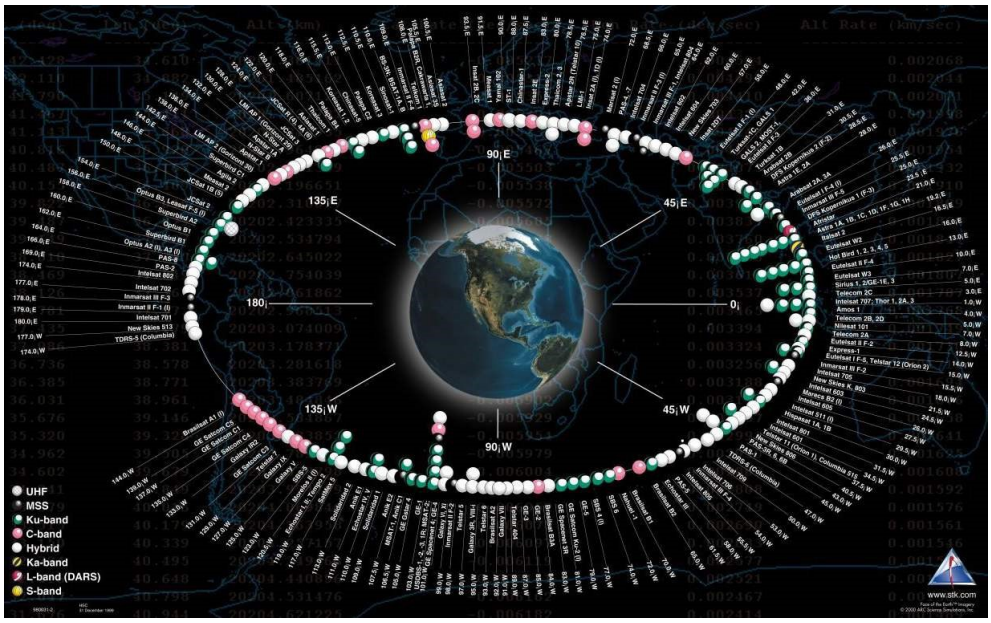

FIGURE I.14 – Représentation des différents satellites de télévision en orbite géostationnaire, colorés sur leur bande de fréquence d'émission. Figure extraite de http://fr.questmachine.org/Applications_des_satellites.

Le principe de fonctionnement de cet appareil (appelé capteur Ku dans la suite) est le suivant. Les satellites géostationnaires de télévision émettent des ondes électromagnétiques dans différentes bandes de fréquences (voir Figure I.14). Nous nous limitons ici à la bande Ku ($\approx 12GHz$, voir Tableau I.1 et satellites colorés en vert de la Figure I.14). Ces ondes, lors de leur traversée de l'atmosphère, sont atténuées (par les gaz, la glace, et majoritairement la pluie en bande Ku). On installe au sol un analyseur de spectre. Celui-ci permet de mesurer la puissance reçue. Connaissant la puissance émise par le satellite, on estime l'atténuation. La Figure I.15 résume ce fonctionnement.

Ce que l'on mesure ainsi est l'atténuation subie par l'onde sur un segment satellite-capteur. Si l'on suppose que l'atténuation provient uniquement des gouttes d'eau liquide d'un diamètre suffisant (ni de la glace, ni des nuages), on obtient une atténuation linéique sur un segment capteur-satellite en dessous de la bande brillante ou du nuage (voir Figure I.15), typiquement quelques kilomètres pour des conditions typiques en France en été. Pour transformer cette atténuation linéique (A, en m^{-1}) en taux de pluie (RR, en mm/h), on utilise ensuite une loi puissance (Barthès and Mallet (2013)) s'écrivant :

$$A = aRR^b \tag{I.33}$$

Avec a et b deux coefficients dépendant principalement de la fréquence, et dans une moindre mesure de la polarisation, de la DSD, et de l'élévation (angle β sur la Figure I.15). Barthès

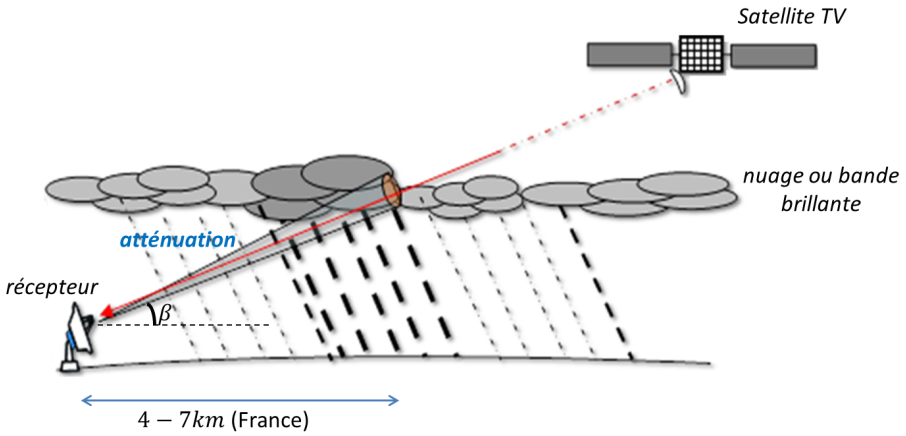

FIGURE I.15 – Schéma présentant le fonctionnement du capteur Ku. Une onde est émise en bande Ku par un satellite géostationnaire de télévision ("Satellite TV"). A cette fréquence, elle est atténuée essentiellement par les gouttes de pluie, donc en dessous de la bande brillante et du nuage. Un récepteur situé au sol mesure la puissance reçue, ce qui, connaissant la puissance émise, permet d'estimer l'atténuation. Projeté au sol et pour des conditions typiques en France en été, on obtient ainsi des mesures d'atténuations linéiques sur des segments de quelques kilomètres de longueur. β désigne l'élévation, angle selon lequel le satellite est vu depuis le sol.

and Mallet (2013) estime la précision de cette relation par rapport à des données de pluie au sol obtenues grâce au DBS (disdromètre optique, voir section 2.2.3). On note que dans la littérature (Barthès and Mallet (2013) par exemple), l'atténuation linéique, notée ici A, est souvent notée k. Dans la suite (chapitre IV), on notera $\mathcal{A}$ l'atténuation totale (intégration de A sur une longueur).

L'intérêt principal de ce système de mesure est qu'il est opportuniste (on utilise des ondes émises pour d'autres besoins), ce qui réduit très sensiblement son coût par rapport à un radar. Il donne par ailleurs des mesures intégrées, ce qui le rend très complémentaire aux pluviomètres, ponctuels (il est susceptible de mieux estimer la quantité d'eau tombée sur une zone mais est moins sensible aux extrêmes).

Ses inconvénients sont qu'il s'agit toujours d'une mesure indirecte de la pluie (comme pour le radar). La relation reliant l'atténation au taux de pluie implique de supposer une forme à la DSD.

On note enfin que, les satellites étant géostationnaires, pour des latitudes européennes, les liaisons seront toujours orientées vers le sud. Le nombre de satellites existant permet cependant d'avoir une ouverture d'environ 75 ° entre les liaisons les plus orientées vers l'est et les plus orientées vers l'est (voir plus en détails notre dispositif expérimental dans le chapitre IV).

Partie II

Dynamique des précipitations : phénomènes et modélisation

Dans ce chapitre, nous revenons sur la modélisation de la dynamique des précipitations. Dans un premier temps, nous étudierons les modèles d'advection-diffusion représentant l'évolution macroscopique des cellules de pluie. Puis nous détaillons les équations aux dérivées partielles représentant les phénomènes physiques agissant sur les gouttes au niveau microscopique. Enfin, nous revenons sur les relations liant diamètres et vitesses terminales de chute des gouttes de pluie, puis sur les schémas numériques de discrétisation permettant de résoudre de manière approchée les équations différentielles.

1 À l'échelle de la cellule de pluie

Dans cette partie, on souhaite décrire l'évolution spatiotemporelle de cellules de pluie telles que les verrait un observateur situé au sol, et indépendamment de considérations sur la microphysique de la pluie (à l'échelle des gouttes). Une cellule de pluie est ici définie comme un champ 3D (temps, 2 dimensions horizontales) de taux de pluie (RR) non nuls. On ne prend donc pas en compte la dimension verticale d'un tel champ. On cherche alors un moyen pratique, à l'échelle macroscopique, pour modéliser le devenir d'une telle cellule de pluie, en utilisant un nombre limité de paramètres, cette limite étant fonction de la quantité et de la diversité des observations disponibles pour estimer les paramètres.

Pour cela, on utilise souvent une équation aux dérivées partielles, potentiellement stochastique, d'advection-diffusion (voir par exemple Gupta and Waymire (1987), Jinno et al. (1993), Kawamura et al. (1997)) , qui peut s'écrire sous sa forme la plus générale :

$$\frac{\partial RR}{\partial t} = \frac{\partial}{\partial x}\left(v_x RR\right) + \frac{\partial}{\partial y}\left(v_y RR\right) + \frac{\partial}{\partial x}\left[\kappa_x \frac{\partial RR}{\partial x}\right] + \frac{\partial}{\partial y}\left[\kappa_y \frac{\partial RR}{\partial y}\right] + \gamma RR + \tilde{\epsilon} \qquad \text{(II.1)}$$

$RR(t, x, y)$ est le taux de pluie, fonction du temps et des deux coordonnées d'espace x et y. Les dérivées premières à droite du signe égal modélisent le transport de la cellule (advection), à la vitesse de déplacement (v_x, v_y), vitesse a priori fonction du temps et de l'espace. Les dérivées secondes correspondent au terme de diffusion ("étalement" de la cellule au cours du temps), avec κ_x et κ_y les coefficients de diffusion dans les deux directions de l'espace, eux aussi a priori fonction du temps et de la position. γRR est un terme source, modélisant la variation de la quantité d'eau dans la cellule (physiquement provoquée par l'évaporation, la condensation) au taux γ. ϵ est un terme aléatoire (typiquement un bruit blanc gaussien centré). Avec RR (en mm/h) réécrit en unités S.I. (m/s), on obtient formellement v_x et v_y en m/s, κ_x et κ_y en m^2/s, γ en s^{-1} et $\tilde{\epsilon}$ en m/s^2.

Jinno et al. (1993) se placent d'un point de vue Lagrangien (sa coordonnée spatiale x suit le déplacement de la cellule) et avec des coefficients constants (dans le temps) et homogènes (dans l'espace). Il propose alors deux solutions à cette équation. La première, sans le terme aléatoire, est analytique et gaussienne. La seconde, incluant le terme aléatoire, correspond à une décomposition de RR et ϵ en séries de Fourier. Cette méthode requiert cependant d'ajuster un certain nombre de paramètres (vitesse de déplacement, coefficients de diffusion, intensité du terme source). Cela n'est pas toujours possible (observations insuffisantes), ni nécessaire, (par exemple terme de diffusion inclus dans la diffusion numérique).

Une telle équation d'évolution d'un champ de pluie est souvent utilisée pour effectuer des prévisions de la pluie à (très) court terme ("nowcasting"), par exemple en extrapolant des images radar. Ce type de modèle semble particulièrement adapté pour certaines applications de l'hydrologie, par exemple la prévision des crues éclair ou l'hydraulique urbaine (voir Wilson

et al. (1998)). Les modèles de prévision numérique du temps n'utilisent en effet pas de telles équations pour déplacer les cellules de pluie mais plutôt pour déplacer des champs d'humidité ou des espèces chimiques (Allen et al. (1991)). Or on sait (Zawadzki (1973), par exemple) que sur de courtes périodes on peut supposer l'atmosphère "gelée", c'est-à-dire supposer que les cellules de pluie se déplacent au gré du vent sans déformation (diffusion) ni changement d'intensité (sources). L'équation (II.1) devient alors une pure équation d'advection :

$$\frac{\partial RR}{\partial t} = \frac{\partial}{\partial x}\left(v_x RR\right) + \frac{\partial}{\partial y}\left(v_y RR\right) \tag{II.2}$$

Si l'on suppose en plus les vitesses d'advection constantes, on se limite à la détermination de deux paramètres.

La Figure II.1 présente des cartes de lames d'eau à 5 minutes enregistrées par un radar en Ardèche en octobre 2013, sur une période de 25 minutes. On voit que bien que les taux de pluie en son sein changent légèrement au cours du temps, la cellule entourée en blanc est relativement bien conservée et semble se déplacer de manière assez constante du sud-ouest de la zone couverte vers le nord-est. Nous quantifierons bien sûr plus loin (voir chapitre IV section 4 et Annexe G) la direction et la vitesse de propagation, ainsi que la pertinence de l'hypothèse de l'atmosphère gelée.

La durée de validité d'une telle hypothèse dépend des conditions météorologiques et du terrain d'étude. Des pluies stratiformes, sur une zone sans relief, l'allongent, des pluies convectives, dans une région montagneuse (comme celles de la Figure II.1), la diminuent. Zawadzki (1973) trouve un temps de validité de 40 minutes en étudiant la corrélation entre des cartes issues de radars canadiens pour un épisode convectif. Météo France utilise cette hypothèse (avec un champ de vitesse constant dans le temps mais variable dans l'espace) pour extrapoler la pluie sur une heure à partir de données radars (Chèze et al. (2014)). Wilson et al. (1998) présentent un inventaire des modèles de nowcasting utilisant de telles modélisations, avec ou sans termes de sources et de diffusion.

Dans la suite, nous utiliserons des données issues d'un petit nombre de capteurs d'atténuation mesurées en Ardèche, région montagneuse (voir chapitre IV). Par conséquent, nous n'aurons pas suffisamment de données pour estimer des termes de diffusion et de sources. Nous limiterons donc à une advection simple. Elle ne sera cependant utilisée que pour déplacer des cellules de pluie sur de courtes périodes ($< 45min$). Tout cela sera réalisé sous la contrainte d'observations via un algorithme d'assimilation de données (voir chapitre III). De plus, la validité de cette hypothèse sera systématiquement vérifiée en étudiant la corrélation entre des cartes radars successives fournies par Météo France (comme celles de la Figure II.1).

On remarque que l'on se place ainsi dans le même contexte que l'étude réalisée par Zinevich et al. (2009), les différences principales venant de la nature du système d'observation et de l'algorithme d'assimilation utilisés. Bien que notre algorithme puisse techniquement opérer

avec un champ de vitesses variables, on se limitera ici à une vitesse d'advection constante et homogène. On se limite donc finalement à l'équation aux dérivées partielles :

$$\frac{\partial RR}{\partial t} = v_x \frac{\partial RR}{\partial x} + v_y \frac{\partial RR}{\partial y} \tag{II.3}$$

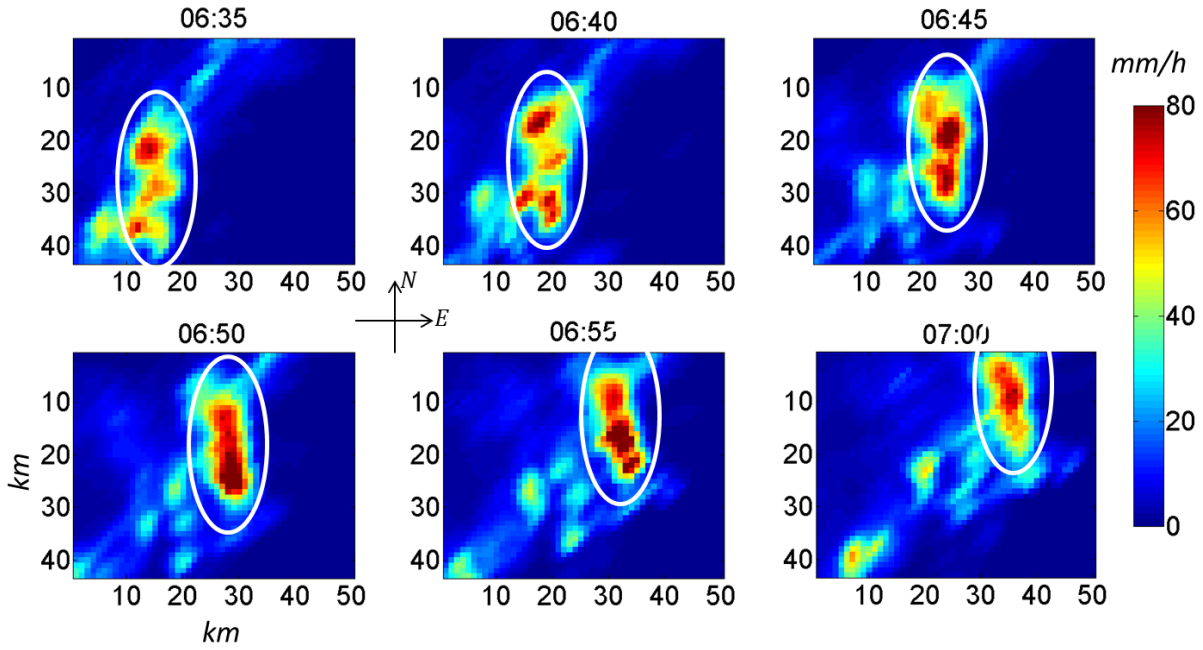

FIGURE II.1 – Produit lame d'eau (champ de taux de pluie à 5 minutes, en mm/h) issu du radar Météo France de Bollène, pour le 20 octobre 2013 entre 6:35 et 7:00 autour de Mirabel, Ardèche.

2 À l'échelle microphysique : évolution spatiotemporelle de la DSD

2.1 Introduction / Revue des phénomènes

Dans le chapitre I, nous avons défini la DSD $n(D)$, distribution des gouttes de pluie selon leur diamètre D. Dans cette section, nous allons considérer sa dynamique. Pour cela, nous allons considérer les différents phénomènes naturels susceptibles d'affecter cette distribution n. Nous introduirons différentes modélisations proposées dans la littérature pour traduire mathématiquement l'impact de ces phénomènes. Nous nous intéresserons aussi aux conditions sous lesquelles ces phénomènes jouent un rôle prépondérant ou non dans l'évolution de n.

Les phénomènes affectant n sont ceux impactant les gouttes d'eau lors de leur chute. Dans la suite, nous considérerons :

- la gravité (compensée par la résistance de l'air), qui fait tomber les gouttes (paramétrisation de la vitesse de chute étudiée rapidement dans la section 3).
- le champ de vent $3D$, qui déplace les gouttes horizontalement et module leur vitesse de chute (section 2.3).
- la condensation, susceptible d'augmenter le diamètre des gouttes (et donc le volume d'eau) (section 2.4).
- l'évaporation, susceptible de le diminuer (section 2.4).
- l'effet de collisions entre gouttes, qui peut résulter en la formation d'une seule goutte (coalescence) ou d'au moins 2 gouttes (break-up), ainsi que l'explosion des gouttes sous leur propre vitesse (auto break-up) (section 2.5 pour une description de ces phénomènes et des études sur leur impact et Annexes A et B pour une étude sur leur paramétrisation).

Nous considérerons ici la DSD $n(D)$ comme une grandeur continue, fonction du temps (coordonnée t) et de l'espace (coordonnées x, y, z). Le nombre de gouttes pour un diamètre donné (en m^{-4}), s'écrit donc :

$$n = n(t, x, y, z, D) \tag{II.4}$$

Dans cette section, nous présenterons tout d'abord l'équation aux dérivées partielles (EDP) représentant la dynamique de n dans sa forme la plus générale, susceptible de prendre en compte l'ensemble des phénomènes listés ci-dessus. Par la suite, nous reviendrons plus en détails sur chacun de ces termes.

2.2 EDP complète

L'EDP modélisant la dynamique de la DSD s'écrit, dans sa forme la plus générale (Hu and Srivastava (1995)) :

$$\frac{\partial n}{\partial t} = \left(\frac{\partial n}{\partial t}\right)_{collisions} + \left(\frac{\partial n}{\partial t}\right)_{evaporation} + \left(\frac{\partial n}{\partial t}\right)_{transport} \tag{II.5}$$

On a simplement ici décomposé la dynamique générale en termes modélisant chacun la part due à un phénomène précis. Les trois termes de droite représentent ainsi la variabilité de la DSD due respectivement aux collisions et explosions de gouttes ; à la condensation et à l'évaporation ; et au transport des gouttes dans l'espace.

On note que la paramétrisation des phénomènes physiques fait souvent intervenir la masse des gouttes m plutôt que leur diamètre D (typiquement pour des calculs énergétiques, voir section 2.5). Il peut alors être plus intéressant d'exprimer une DSD "massique", n_m, en $m^{-3}kg^{-1}$. On considère donc la distribution des gouttes par classe de masse et non plus par classe de taille, les deux grandeurs étant équivalentes pour des gouttes sphériques. Les deux grandeurs sont reliées par :

$$n_m = n\frac{dD}{dm} \tag{II.6}$$

Avec dD/dm la dérivée de la relation diamètre/masse. Pour une goutte sphérique (ρ_e la masse volumique de l'eau) :

$$m(D) = \frac{\rho_e \pi}{6} D^3 \tag{II.7}$$

Et on obtient :

$$\frac{dD}{dm} = \frac{2}{\rho_e \pi D^2} \tag{II.8}$$

On pourra ainsi utiliser indifféremment dans la suite une DSD en diamètre ou en masse.

Dans l'Eq.(II.5), les collisions et explosions de gouttes provoquent des échanges de masse, donc des variations du diamètre des gouttes (voir section 2.5) ; la condensation et l'évaporation entraînent elles des variations de masse en préservant l'intégrité des gouttes (voir section 2.4) ; et le transport des gouttes dans l'espace se fait par le vent et par la gravité (voir section 2.3).

2.3 Transport

En notant $\boldsymbol{v_{3D}}$ le vecteur vitesse des gouttes (fonction du temps, de l'espace et du diamètre), leur transport dans l'espace peut être représenté par l'équation d'advection :

$$\left(\frac{\partial n}{\partial t}\right)_{transport} + div\left[\boldsymbol{v_{3D}} n\right] = 0 \tag{II.9}$$

Dans toute la suite, on considérera que la vitesse des gouttes d'eau se réduit à la somme de la vitesse terminale de chute et de la vitesse du vent. La vitesse terminale de chute, notée v_{th}, est la vitesse correspondant à l'équilibre entre la gravité et la résistance de l'air. Elle est orientée selon z, dirigé vers le bas. On la suppose ici paramétrée en fonction du diamètre et éventuellement de l'altitude ($v_{th} = v_{th}(D, z)$). Sa paramétrisation et les écarts constatés à cette vitesse terminale sont l'objet d'une littérature abondante. Ces points sont évoqués plus en

détails dans la section 3 et dans l'Annexe C.

Les trois coordonnées du vent sont notées v_x, v_y et v_z, respectivement selon x, y et z. On suppose formellement ici que les variations de vitesse du vent sont communiquées instantanément à la vitesse de chute des gouttes d'eau. Compte tenu de la dynamique propre des gouttes due à leur masse, cela est bien sûr contestable, notamment lorsque le vent varie rapidement (turbulence). Dans la suite, on constatera néanmoins que le modèle utilisé pourra numériquement (c'est-à-dire indépendamment de la physique de l'EDP) générer des variations rapides de vitesse des gouttes, qui permettront indirectement une modélisation de la turbulence (voir chapitre V). Pour le vent vertical, dans certaines sections, quand on supposera la composante horizontale du vent nulle ($v_x = v_y = 0$), on utilisera la notation w plutôt que v_z .

La vitesse $3D$ des gouttes s'écrit donc finalement :

$$\boldsymbol{v_{3D}} = \begin{pmatrix} v_x \\ v_y \\ v_{th} + v_z \end{pmatrix} \tag{II.10}$$

L'Eq.(II.9) prend la forme suivante :

$$\left(\frac{\partial n}{\partial t}\right)_{transport} + \frac{\partial}{\partial x}\left(v_x n\right) + \frac{\partial}{\partial y}\left(v_y n\right) + \frac{\partial}{\partial z}\left(\left[v_{th} + v_z\right] n\right) = 0 \tag{II.11}$$

Ou encore, avec les dépendances aux coordonnées :

$$\left(\frac{\partial n}{\partial t}\right)_{transport}(t,x,y,z,D) + div\left[\begin{pmatrix} v_x(t,x,y,z) \\ v_y(t,x,y,z) \\ v_{th}(z,D) + v_z(t,x,y,z) \end{pmatrix} n(t,x,y,z,D)\right] = 0 \tag{II.12}$$

Dans la suite, nous travaillerons souvent dans une colonne d'atmosphère, c'est-à-dire en ne considérant qu'une seule coordonnée d'espace (z). On ne pourra donc pas prendre en compte les mouvements horizontaux. Nous discuterons bien sûr plus en détails de cette simplification dans la suite (voir chapitre V, section 4).

2.4 Évaporation

2.4.1 Introduction

Dans cette section, nous allons décrire le terme $\left(\frac{\partial n}{\partial t}\right)_{evaporation}$ de l'Eq.(II.5). Ce terme regroupe a priori les phénomènes affectant le diamètre des gouttes d'eau sans modifier leur intégrité. Cela signifie qu'il inclut les effets de la condensation et de l'évaporation, mais pas des collisions ou des explosions de gouttes. L'évaporation induit une perte de masse (et donc de diamètre), tandis que la condensation induit un gain de masse. La condensation est d'une importance cruciale dans les nuages (Hall (1980), Brenguier (1991)). Dans ce travail sur la

pluie, nous allons la négliger, comme indiqué par Hu and Srivastava (1995).

Mathématiquement, l'évaporation est représentée par une équation d'advection selon les diamètres (ou les masses des gouttes) à vitesse non constante. Cette vitesse d'advection est le taux d'évaporation, exprimé classiquement en kg/s (perte de masse par unité de temps), et noté ici ξ_m. On a alors :

$$\left(\frac{\partial n}{\partial t}\right)_{evaporation} = \frac{\partial}{\partial D}\left[\xi_m n_m\right] \tag{II.13}$$

On note que l'on travaille ici avec des grandeurs massiques. On aurait aussi pu définir un taux d'évaporation en diamètre (i.e. définir un ξ_D, en m/s). Il paraît néanmoins plus naturel de raisonner en perte de masse plutôt qu'en perte de diamètre. Le passage de l'un à l'autre se fait, toujours en supposant les gouttes sphériques, selon l'Eq.(II.8).

Le problème consiste ensuite à paramétrer ce taux d'évaporation. Nous allons détailler la paramétrisation proposée par Seifert (2008), qui sera ensuite utilisée dans ce travail.

2.4.2 Paramétrisation de Seifert (2008)

Quelques grandeurs physiques Définissons d'abord (indépendamment de la paramétrisation de Seifert (2008)) les équations utilisées pour évaluer certaines grandeurs physiques nécessaires au calcul du taux d'évaporation.

La pression atmosphérique P_{atm} (en Pa), tout d'abord, sera évaluée à partir de la pression au niveau de la mer P_0 (en Pa), de la température T (en K), du gradient de température α_T (en K/m), et de l'altitude z (en m) par la formule de nivellement barométrique (https://en.wikipedia.org/wiki/Barometric_formula) :

$$P_{atm} = P_0\left(1 - \frac{\alpha_T z}{T}\right)^{\frac{0.0342}{\alpha_T}} \tag{II.14}$$

Où 0.0342 est le coefficient (en K/m) correspondant au produit de la masse molaire de l'air par l'accélération de la pesanteur sur la constante universelle des gaz parfaits.

La pression de vapeur saturante e_{sat} (en Pa) sera calculée via la formule de Rankine, qui s'écrit (http://www.grasp.ulg.ac.be/nvdw/NVdw/Documents_files/thermo7.pdf) :

$$e_{sat} = 101325e^{13.7-\frac{5120}{T}} \tag{II.15}$$

La masse volumique de l'air, ρ_a (kg/m^3), sera évaluée à partir de la pression atmosphérique, de l'humidité relative h_r, de la température et de la pression de vapeur saturante, par (Brutsaert (2013)) :

$$\rho_a = \frac{P_{atm}}{R_d T}\left(1 - 0.378\frac{h_r e_{sat}(T)}{P_{atm}}\right) \tag{II.16}$$

Avec $R_d = 287.05 J/kg/K$ la constante des gaz pour l'air sec.

La diffusivité de la vapeur d'eau dans l'air, D_v (en m^2/s), sera calculée à partir de la température via (Gates (1980)) :

$$D_v(T) = 21.2.10^{-6}\left(1 + 7.1.10^{-3}T\right) \tag{II.17}$$

La viscosité cinématique de l'air, ν_a (en m^2/s), sera enfin évaluée à partir de la température et de la masse volumique de l'air par la formule de Sutherland (Sutherland (1893)) :

$$\nu_a(T, \rho_a) = \frac{1}{\rho_a}\frac{1.458.10^{-6}T^{\frac{3}{2}}}{T + 110.4} \tag{II.18}$$

Paramétrisation de l'évaporation Le taux d'évaporation des gouttes ξ_m, dans la paramétrisation de Seifert (2008), est une fonction du diamètre D des gouttes, de la température T, du taux d'humidité relative h_r et de la vitesse de chute des gouttes $v_{th} + v_z$. On suppose (pour pouvoir calculer des taux d'évaporation indépendamment du vent, plus difficile à évaluer en pratique que la température ou la pression car plus variable) que $v_z = 0$ et que les gouttes tombent à la vitesse v_{th}. On suppose toujours que v_{th} est elle-même fonction du diamètre et de l'altitude. Le modèle d'évaporation de Seifert (2008) s'écrit alors :

$$\xi_m(D, T, h_r, z) = 2\pi D G(T) f_{vent}(v_{th}(D, z), D) Su(h_r) \tag{II.19}$$

Avec T la température (K), h_r l'humidité relative, D le diamètre (m) et z l'altitude (m). Le coefficient G (en $kg/m/s$), est une fonction traduisant l'effet des conditions météorologiques, hors humidité, environnant la goutte. f_{vent} (sans unité) est le coefficient de ventilation, qui traduit l'effet du contact entre la goutte et l'air ambiant. Il dépend de la vitesse de chute et du diamètre. Su, le ratio de supersaturation, traduit l'influence de l'humidité relative de l'air ambiant.

Le ratio de supersaturation s'écrit :

$$Su(h_r) = h_r - 1 \tag{II.20}$$

On comprend que plus l'humidité s'approche de 1, plus l'air est proche d'être saturé, et plus le taux d'évaporation va être faible. Ici, on retrouve bien $Su(h_r = 1) = 0$, et donc $\xi_m(h_r = 1) = 0$: si l'air est déjà saturé, il n'y a pas d'évaporation. Comme h_r varie entre 0 et 1, on note également que $Su \leq 0$. On verra que G et f_{vent} sont des termes positifs, ce qui implique que l'évaporation ξ_m est négative ou nulle. Cela traduit bien la perte de masse inhérente à l'évaporation.

Le coefficient G s'écrit :

$$G(T) = \left(\frac{R_v T}{D_v(T) e_{sat}(T)} + \frac{L_v^2}{K_T R_v T^2} \right)^{-1} \tag{II.21}$$

R_v est la constante des gaz pour la vapeur d'eau. On prendra ici $R_v = 461.5 J/kg/K$ (http://www.lmd.ens.fr/jpduvel/Cours/Thermo_4_chgt_phase_eau.pdf).

K_T est la conductivité de la chaleur de l'air. Elle évolue lentement avec la température. Elle passe de $2.4 * 10^{-2} W/m/K$ à $0\,^\circ C$ à $2.65 * 10^{-2} W/m/K$ à $30\,^\circ C$ (http://www.electronics-cooling.com/2003/11/the-thermal-conductivity-of-moist-air/). Dans ce travail, on la considérera constante : $K_T = 2.5 * 10^{-2} W/m/K$.

L_v est la chaleur latente de vaporisation de l'eau. Elle dépend de la température et du point d'ébullition (donc de la pression atmosphérique). À $100\,^\circ C$ et $1 atm$, on a $L_v = 2.26 * 10^6 J/kg$ (https://en.wikipedia.org/wiki/Latent_heat). Cette valeur augmente jusqu'à environ $2.5 * 10^6 J/kg$ quand la température descend à $0\,^\circ C$. Là encore dans ce travail, on la considérera constante ($L_v = 2450 J/kg$) (https://fr.wikiversity.org/wiki/Changements_d'%C3%A9tats/Vaporisation_et_condensation). Cette valeur correspond à une température proche de $20\,^\circ C$, ce qui correspond assez bien aux conditions à Darwin où l'évaporation est significative.

Le facteur de ventilation f_{vent} s'écrit :

$$f_v(V_z, D) = 0.78 + 0.308 0.71^{1/3} N_{Rc}^{1/2}(v_{th}, D) \tag{II.22}$$

Avec

$$N_{Rc}(v_{th}, D) = \frac{v_{th} D}{\nu_a(T, \rho_a)} \tag{II.23}$$

Le taux d'évaporation augmente avec la vitesse de chute et le diamètre, et diminue avec la viscosité de l'air.

Dans ce travail, la température et l'humidité relative seront évaluées soit à partir de sorties de modèles, soit par des relevés au sol ou en altitude par des ballons-sondes, soit en combinant ces informations (voir dans le chapitre V).

2.4.3 Illustration

En nous basant sur une simulation simple, nous allons à présent illustrer les effets de l'évaporation, telle que paramétrée par Seifert (2008). Pour cela, nous résolvons numériquement l'Eq.(II.5) pour un vent nul ($v_x = v_y = v_z = 0$) et une vitesse terminale de chute suivant Atlas et al. (1973) (voir section 3). Nous ne prenons pas non plus les collisions en compte.

Sans vent, il n'y a aucun mouvement horizontal (les gouttes tombent verticalement). On peut ainsi se restreindre à une unique dimension d'espace (tout se passe dans une colonne). La

résolution numérique est faite via un simple schéma upstream (voir section 4). Pour une hauteur de chute de $3000m$, on considère un champ initial nul ($n(t_0) = 0$), puis on applique une condition limite en haut constante, correspondant à une DSD de Marshall and Palmer (1948) (forme exponentielle) correspondant à un taux de pluie de $6mm/h$. Pour le modèle d'évaporation de Seifert (2008), la température est fixée à $27\,^\circ C$ au sol, avec un gradient vertical de $-1\,^\circ/150m$, et une humidité constante et homogène dans la colonne. Aucun paramètre du modèle ne variant dans le temps, le système, une fois que les plus petites gouttes auront atteint le sol, va atteindre un régime continu. Tous les résultats présentés sont ceux de ce régime continu.

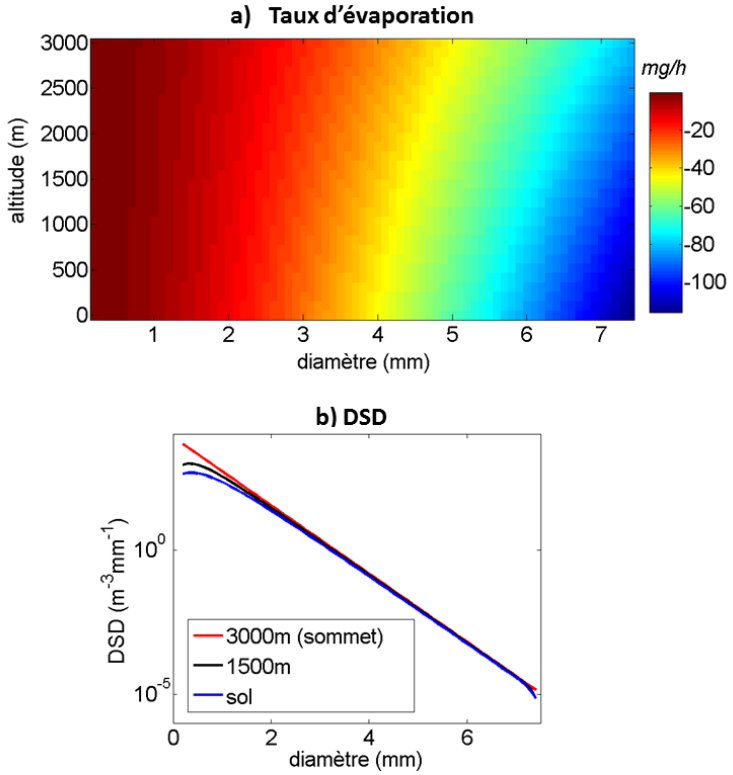

FIGURE II.2 – a) Taux d'évaporation massique selon l'altitude et le diamètre des gouttes sphériques. Calculé selon la paramétrisation proposée par Seifert (2008). Le taux d'humidité est homogène à 80%, la température est de $27\,^\circ C$ au sol, avec un gradient de $-1\,^\circ/150m$. L'unité est ici (pour des questions de lisibilité) le milligramme par heure. b) Impact selon l'altitude sur une DSD d'un tel taux d'évaporation. On propage verticalement, sans vent ni collisions, une DSD de Marshall and Palmer (1948) (pour un taux de pluie de $6mm/h$). La hauteur de chute est de $3000m$.

Sur la Figure II.2, on a représenté en haut le taux d'évaporation (massique), selon l'altitude et le diamètre, obtenu pour une humidité de 80%. On remarque, comme on s'y attendait, que les gouttes perdent, en valeur absolue, d'autant plus d'eau qu'elles sont grosses. On remarque également que le taux d'évaporation diminue avec l'altitude, du fait de la baisse de température. Le bas de la Figure II.2 représente les DSD à trois altitudes différentes, pour un taux d'évaporation constant sur la colonne. En haut de la colonne (rouge), on retrouve bien la DSD exponentielle fixée en condition limite. L'évaporation va ensuite diminuer plus fortement le nombre de petites gouttes.

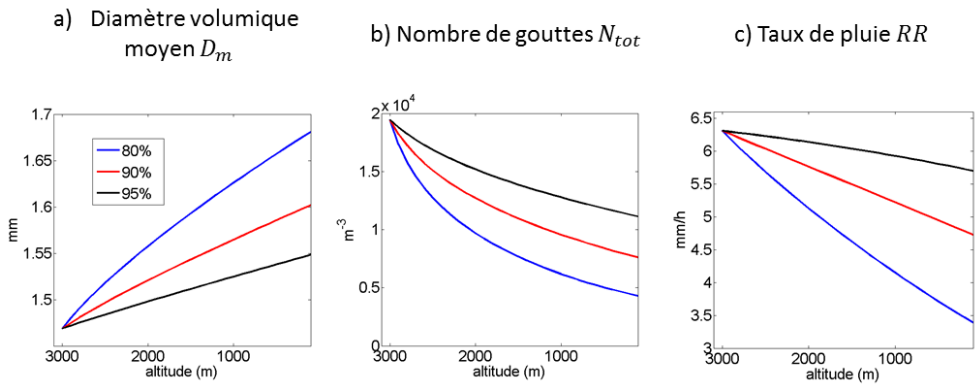

FIGURE II.3 – Impact de l'évaporation sur 3 paramètres issus de la DSD, selon l'altitude. On fait ici tourner un modèle de dynamique des DSD dans les mêmes conditions que sur la Figure II.2 (ni vent ni collisions, évaporation de Seifert (2008), DSD au sommet de la colonne de Marshall and Palmer (1948) pour un taux de pluie de $6mm/h$). Expérience réalisée pour 3 niveaux d'humidité : 80% (bleu), 90% (rouge) et 95% (noir). On représente : a) le diamètre volumique moyen D_m ($\mathcal{M}_4(n)/\mathcal{M}_3(n)$, en mm) ; b) le nombre de gouttes N_{tot} ($\mathcal{M}_0(n)$, en m^{-3}) ; c) le taux de pluie RR (proportionnel à $\int n(D)v_{atlas}(D)D^3dD$, en mm/h).

La Figure II.3 montre les effets de l'évaporation sur 3 paramètres issus de la DSD : le diamètre volumique moyen D_m, le nombre total de gouttes N_0, et le taux de pluie RR (voir légende), selon l'altitude, et pour trois taux d'humidité différents. On remarque que l'évaporation, en détruisant massivement les très petites gouttes, tend à faire augmenter le D_m (qui, en $3000m$, passe d'environ $1.46mm$ à $1.68mm$ à 80% d'humidité). Le nombre total de gouttes, piloté principalement par le nombre de très petites gouttes pour une DSD exponentielle, est bien sûr fortement affecté par l'évaporation. À 80% d'humidité, on détruit ainsi en $3000m$ plus des trois quarts des gouttes (courbe bleue). Même avec 95% d'humidité, on perd environ 40% des gouttes lors de la chute (courbe noire). Le taux de pluie est aussi fortement affecté, avec une baisse de près de 50% pour 80% d'humidité (courbe noire).

On conclut que l'évaporation semble jouer un rôle important, y compris pour des taux d'humidité assez forts, et d'autant plus significatif bien sûr que la hauteur de chute est grande. Il conviendra donc de la prendre en compte dans ce travail.

2.5 Collisions et désintégration de gouttes

Dans cette section, nous allons présenter les phénomènes susceptibles d'affecter l'intégrité des gouttes de pluie durant leur chute. Paramétrés, ces phénomènes sont ceux qui composeront le terme de collision de l'EDP (Eq.(II.5)), $(\partial n/\partial t)_{collisions}$.

De manière générale, ils regroupent deux catégories de phénomènes.

Tout d'abord les collisions entre gouttes à proprement parler. Il s'agit d'un phénomène complexe, pouvant déboucher sur différents scénarios (coalescence, c'est-à-dire fusion des deux gouttes impliquées dans la collision ; rebond sans modification des gouttes, ou explosion en une distribution de fragments). De nombreuses études existent pour déterminer sous quelles

conditions apparaissent ces différents phénomènes et pour les paramétrer de manière à ce qu'ils soient modélisables dans l'EDP (Eq.(II.5)). Nous présenterons rapidement ces phénomènes dans la section 2.5.1. La section 2.5.2 introduira la modélisation de ces phénomènes dans l'EDP à proprement parler. Pour une revue d'études proposant des paramétrisations explicites de ces phénomènes, on se reportera à l'annexe A. Enfin, la section 2.5.3 explorera rapidement l'influence de ces paramétrisations sur les DSD stationnaires obtenues via l'EDP (II.5) limitée au terme de collision et justifiera le choix fait dans ce travail de négliger ces phénomènes et les conditions sous lesquelles ce choix est envisageable.

On note que le terme de collision de l'EDP inclut aussi potentiellement l'explosion des gouttes sous leur propre poids. Ce phénomène, largement débattu dans la littérature, ne sera pas non plus pris en compte dans ce travail. L'annexe B revient néanmoins sur cette question.

2.5.1 Coalescence / Break-up : qu'est-ce qu'il se passe ?

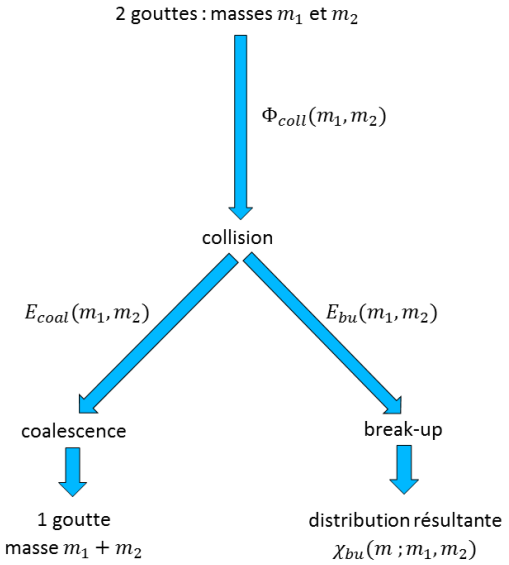

FIGURE II.4 – Évolutions possibles pour deux gouttes de masses m_1 et m_2 susceptibles de se percuter. Voir texte pour les définitions des noyaux de collision Φ_{coll}, des efficacités de coalescence et de break-up E_{coal} et E_{bu}, et de la fonction de break-up χ_{bu}.

Dans cette section, nous allons introduire et contextualiser les notations associées au devenir de deux gouttes susceptibles de se percuter. La section 2.5.2 introduira ces phénomènes dans l'EDP (Eq.(II.5)). Les paramétrisations des grandeurs présentées ici sont par ailleurs détaillées dans l'annexe A. La Figure II.4 introduit graphiquement ces notations. Nous allons la commenter ici. Supposons deux gouttes de masses m_1 et m_2 présentes dans un certain volume. On suppose qu'elles sont sphériques (donc que leur diamètre est déterminé par leur masse) et tombent à des vitesses connues, fonction éventuellement de leur masse, de l'altitude, du vent, etc. Dans un temps donné, elles peuvent se percuter ou ne pas se croiser. On définit ainsi à partir de cette probabilité de collision un noyau de collision, Φ_{coll}. (Dans la mesure où l'on travaille avec des grandeurs volumiques - les DSD - et que l'on cherche à

modéliser leur dérivée temporelle, Φ_{coll} n'est pas formellement une densité de probabilité mais une grandeur en m^3/s, associée à cette probabilité).

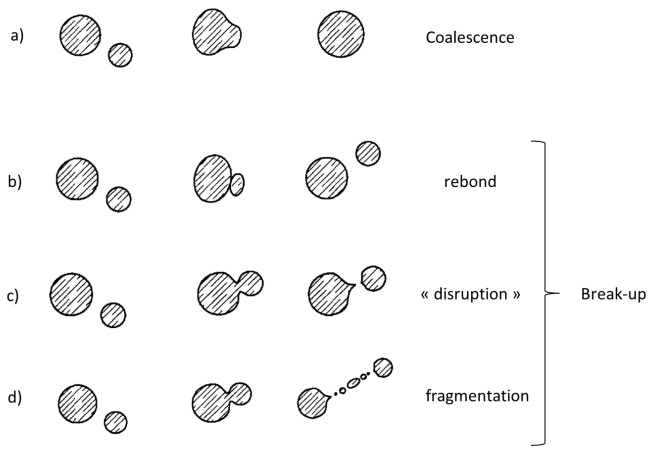

FIGURE II.5 – Croquis représentant les résultats possibles d'une collision entre deux gouttes, à savoir : a) coalescence ; b) rebond ; c) "disruption" (séparation à nouveau en deux gouttes) ; d) fragmentation. Figure adaptée de Orme (1997). Définition du break-up à partir de ces types d'impacts.

En cas de collision, plusieurs effets, illustrés sur la Figure II.5, peuvent en découler. Les gouttes peuvent rebondir l'une contre l'autre (sans échange de masse), fusionner pour n'en former plus qu'une, ou encore exploser en plusieurs fragments. Nous classons ces effets dans deux classes : la coalescence, correspondant à la formation d'une unique goutte, et le break-up, correspondant à la formation d'au moins deux gouttes. La probabilité d'avoir coalescence est noté E_{coal} ("efficacité de coalescence"), celle d'avoir break-up E_{bu} ("efficacité de break-up"). Les deux phénomènes étant exclusifs, on a naturellement :

$$E_{coal} + E_{bu} = 1 \tag{II.24}$$

On va encore définir le noyau de coalescence, Φ_{coal}, comme :

$$\Phi_{coal} = E_{coal}\Phi_{coll} \tag{II.25}$$

Il traduit la probabilité pour deux gouttes de masses m_1 et m_2 de n'en former in fine plus qu'une, de masse $m_1 + m_2$, via la coalescence résultant de la collision.

On définit de même le noyau de break-up, Φ_{bu}, comme :

$$\Phi_{bu} = E_{bu}\Phi_{coll} \tag{II.26}$$

Dans ce dernier cas, cependant, il ne se forme pas une unique goutte mais plusieurs dont la distribution des masses est supposée connue. Celle-ci, notée χ_{bu}, est appelée la fonction de break-up. La quantité $\Phi_{bu}(m_1, m_2)\chi_{bu}(m; m_1, m_2)$ correspond à la probabilité pour deux gouttes de masses m_1 et m_2 de former in fine des gouttes de masse m via une collision résultant en

break-up.

2.5.2 Coalescence / Break-up : modélisation pour l'EDP

Maintenant que les probabilités d'occurrence de la coalescence et du break-up ont été définies dans la section 2.5.1, nous allons, dans cette section, montrer comment intégrer les effets dans l'EDP modélisant la dynamique de la DSD (Eq.(II.5)). On peut retrouver notamment dans Hu and Srivastava (1995), Feingold et al. (1988), Prat and Barros (2007) ou Beheng (2010) les expressions données et rapidement justifiées ici.

On considère ici la variation du nombre de gouttes pour une masse donnée. On note ainsi respectivement I_1 et I_2 les variations positives (apports de gouttes) dues à la coalescence et au break-up, et $I_1^{'}$ et $I_2^{'}$ les variations négatives (pertes de gouttes) dues à la coalescence et au break-up. On a ainsi :

$$\left(\frac{\partial n_m}{\partial t}\right)_{collisions}(m) = I_1(m) + I_2(m) + I_1^{'}(m) + I_2^{'}(m) \tag{II.27}$$

On travaille ici avec des grandeurs massiques, et on a explicité seulement la dépendance à la masse m des quantités considérées (qui dans les faits dépendent aussi du temps et de l'espace). Toutes les grandeurs à droite du signe égal sont donc exprimées en $m^{-3}kg^{-1}s^{-1}$.

Apports par coalescence Par coalescence, deux gouttes de masses m_1 et m_2 ne peuvent produire qu'une goutte de masse $m = m_1 + m_2$. En conséquence, le nombre de gouttes produites dépendra du nombre de gouttes de chaque masse (m_1 et $m_2 = m - m_1$, avec $m_1 < m$) avant la collision, ainsi que du noyau de coalescence Φ_{coal}. On intègre ensuite cette expression sur toutes les valeurs de m_1 susceptibles de produire une goutte de masse m en s'associant à une autre goutte (c'est-à-dire pour $m_1 \in [0; m]$). On obtient finalement :

$$I_1(m) = \frac{1}{2}\int_0^m n_m(m_1) n_m(m - m_1) \Phi_{coal}(m_1, m - m_1) dm_1 \tag{II.28}$$

Le facteur 1/2 évite que chaque goutte ne soit comptée deux fois (une fois dans le calcul de $I_1(m_1)$ et une fois dans celui de $I_1(m - m_1)$). Avec n_m en $m^{-3}kg^{-1}$, m en kg et Φ_{coal} en m^3/s (avec comme on l'a vu $\Phi_{coal} = E_{coal}\Phi_{coll}$), on a bien I_1 en $m^{-3}kg^{-1}s^{-1}$, dérivée temporelle de la concentration massique de gouttes.

Pertes par coalescence La coalescence induit la perte des deux gouttes à l'origine de la collision. Si on considère la coalescence résultant de la collision d'une goutte de masse m percutée par une goutte de masse m_1 quelconque, la perte de gouttes par coalescence est proportionnelle au nombre de gouttes de masses m et m_1, et au noyau de coalescence Φ_{coal}, intégré pour toutes les valeurs de m_1. Soit :

$$I_1^{'}(m) = n_m(m)\int_0^{\infty} n_m(m_1) \Phi_{coal}(m, m_1) dm_1 \tag{II.29}$$

Apports par break-up Le nombre de gouttes de masse m créées par break-up à partir de gouttes de masses m_1 et m_2 est proportionnel au nombre de gouttes de chacune de ces masses, ainsi qu'au noyau de break-up Φ_{bu} et à la fonction de break-up χ_{bu}. On intègre ensuite pour toutes les masses m_1 et toutes les masses m_2. Soit :

$$I_2(m) = \frac{1}{2}\int_0^\infty \int_0^\infty \chi_{bu}(m; m_1, m_2) n_m(m_1) n_m(m_2) \Phi_{bu}(m_1, m_2) dm_1 dm_2 \tag{II.30}$$

χ_{bu}, la distribution des gouttes résultant du break-up selon leur masse, est en kg^{-1}.

Pertes par break-up La perte de gouttes par break-up s'écrit comme pour le cas de la coalescence (en utilisant le noyau de break-up), soit :

$$I_2^{'}(m) = n_m(m) \int_0^\infty n_m(m_1) \Phi_{bu}(m, m_1) dm_1 \tag{II.31}$$

On note que $\Phi_{coal} + \Phi_{bu} = \Phi_{coll}$ (voir Eqs.(II.24), (II.25) et (II.26)). On peut donc regrouper $I_1^{'}$ et $I_2^{'}$ dans un unique terme de pertes par collisions, $I^{'} = I_1^{'} + I_2^{'}$, qui s'écrit alors :

$$I^{'}(m) = n_m(m) \int_0^\infty n_m(m_1) \Phi_{coll}(m, m_1) dm_1 \tag{II.32}$$

2.5.3 Coalescence / Break-up : influence sur les distributions stationnaires

L'EDP (Eq.(II.5)) modélisant la dynamique de la DSD, sans terme d'évaporation et sans vent, peut se réduire à une dimension d'espace (colonne verticale) et s'écrit :

$$\left(\frac{\partial n}{\partial t}\right)(D, z, t) + \frac{\partial}{\partial z}\left[v_{th}(D, z) n(D, z, t)\right] = I_1(D) + I_2(D) + I^{'}(D) \tag{II.33}$$

Le deuxième terme à gauche du signe égal représente la chute des gouttes sous leur propre poids. Ceux de droite, correspondant aux effets de la coalescence et du break-up, sont écrits ici en fonction du diamètre et non de la masse des gouttes.

Avec des conditions aux limites (distributions en haut de la colonne) constantes et pour une hauteur de chute suffisante, la résolution numérique d'une telle EDP va conduire à un état stationnaire (ou état d'équilibre). C'est-à-dire qu'à partir d'une certaine hauteur de chute, la DSD n'évolue plus. Un grand nombre d'études cherchent cette distribution stationnaire, et comparent celles obtenues pour différentes conditions aux limites, ainsi que pour les différentes paramétrisations de l'efficacité de coalescence E_{coal} et de la fonction de break-up χ_{bu}. Parmi ces études , on peut citer Hu and Srivastava (1995), Seifert et al. (2005), Prat and Barros (2007), Prat and Barros (2007), Mcfarquhar (2010), Straub et al. (2010), Prat et al. (2012), Barthes and Mallet (2013).

Les Figures II.6, II.7 et II.8 présentent respectivement l'influence sur la distribution

stationnaire de la paramétrisation de la fonction de break-up, des conditions aux limites, et le temps mis pour atteindre une distribution stationnaire.

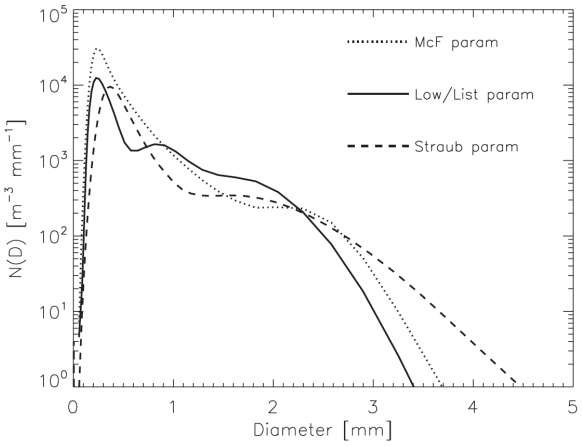

FIGURE II.6 – Figure extraite de Mcfarquhar (2010) présentant différentes distributions stationnaires. Distributions stationnaires obtenues pour un noyau de coalescence de Low and List (1982a) (traits continues et points) ou de Straub et al. (2010) (tirets) et avec une fonction de break-up de Low and List (1982b) (trait continu), de McFarquhar (2004) (points), de Straub et al. (2010) (tirets). Dans tous les cas, la condition aux limites en haut est une DSD de Marshall and Palmer (1948) pour un taux de pluie de $54mm/h$.

On voit sur la Figure II.6 que la coalescence et le break-up tendent à produire des distributions avec un ou plusieurs pics. Partant d'une distribution exponentielle, les très petites gouttes sont dans tous les cas quasiment toutes détruites. Un premier pic apparaît pour les 3 paramétrisations aux alentours de $0.2 - 0.3mm$. Les distributions sont ensuite assez différentes, que ce soit sur la présence et la localisation d'un second pic ou sur la queue, significativement plus lourde avec la paramétrisation de Straub et al. (2010). On peut noter que la difficulté existant pour obtenir des mesures, notamment sur les très petits diamètres, et la variabilité des DSD observées rendent difficiles la validation de telles simulations par l'observation. La difficulté à obtenir des mesures de la DSD à différentes altitudes complexifie le problème. Il est ainsi difficile de savoir si un équilibre est atteint dans la nature. Par ailleurs, toutes ces simulations sont réalisées dans des conditions forcément différentes de la nature (pas de vent ou en tous cas de turbulence, notamment).

On voit sur la Figure II.7 que la forme des distributions d'équilibre, dans le cadre de la paramétrisation de Low and List (1982a), ne dépend pas de l'intensité de la pluie, cette dernière ne jouant que sur la vitesse de convergence vers l'équilibre.

La Figure II.8 montre que le temps mis pour atteindre l'équilibre augmente sensiblement à mesure que le taux de pluie décroît. Ce comportement est assez naturel : plus le taux de pluie est faible, moins il y a de gouttes (surtout de grosses gouttes susceptibles de percuter plus de petites) et moins il y a de collisions.

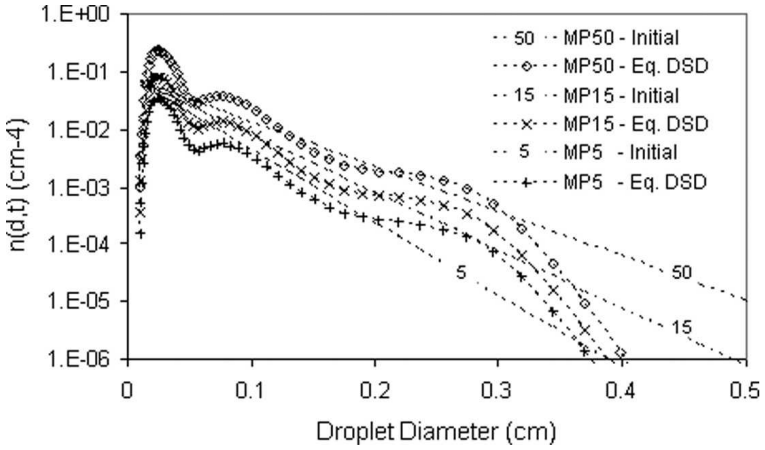

FIGURE II.7 – Figure extraite de Prat and Barros (2007). Distributions aux limites (marques) et stationnaires (pointillés) obtenues dans le cadre de la paramétrisation de Low and List (1982a) / Low and List (1982b) (efficacité de coalescence et fonction de break-up) pour différentes distributions aux limites. Celles-ci sont des distributions de Marshall and Palmer (1948) pour des taux de pluie de $50mm/h$ ("MP50"), $15mm/h$ ("MP15") et $5mm/h$ ("MP5").

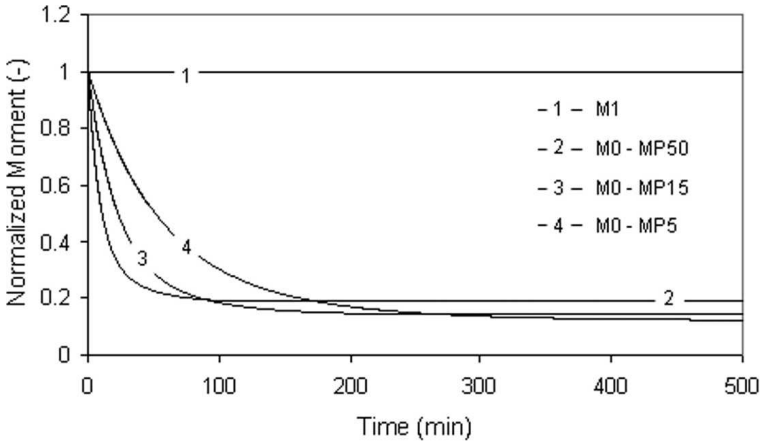

FIGURE II.8 – Figure extraite de Prat and Barros (2007). Evolution temporelle des moments d'ordre 0 et 1 de la DSD lors de la chute des gouttes, dans le cadre d'une efficacité de coalescence et d'une fonction de break-up de Low and List (1982a) et Low and List (1982b). Les conditions aux limites sont des distributions de Marshall and Palmer (1948) pour des taux de pluie de $50mm/h$ ("MP50"), $15mm/h$ ("MP15") et $5mm/h$ ("MP5").

Considérons des gouttes tombant assez lentement, à $2m/s$. Une hauteur de chute de $2km$ correspond pour elles à un temps de chute de $17min$. On voit sur la Figure II.8 qu'en un tel temps et pour des taux de pluie modérés ($15mm/h$), on est loin d'avoir atteint un état stationnaire. Pour des taux de pluie faibles ($5mm/h$), l'état initial n'a même été que peu affecté.

La coalescence/break-up est donc un processus assez lent, et pour des hauteurs de chute raisonnables, les DSD ne semblent pas avoir le temps d'atteindre un équilibre, d'autant moins que les taux de pluie sont faibles. Le fait que les paramétrisations proposées varient fortement et sont difficilement validables nous conduiront dans ce travail à négliger la coalescence et le break-up. Nous nous restreindrons bien entendu pour cela à des évènements calmes (taux de pluie sous les $10mm/h$).

3 Relations vitesse-diamètre pour la pluie

On a vu dans le chapitre I que la relation entre taux de pluie et DSD dépend de la vitesse de chute des gouttes. De même, la détermination de la DSD à partir d'un spectre Doppler (distribution d'énergie fonction de la vitesse Doppler mesurée par un radar) nécessite une relation liant vitesse et diamètre des gouttes. Dans cette section, nous allons présenter la relation théorique vitesse/diamètre utilisée dans ce travail et introduirons rapidement les discussions sur les écarts pouvant exister à cette vitesse. L'annexe C fournit une discussion plus détaillée sur les différentes relations existant dans la littérature et leur justification théorique et empirique.

3.1 Introduction

Notons v_g la vitesse d'une goutte de diamètre D_g. Nous allons dans cette partie écrire v_g sous la forme :

$$v_g = v_{th}(D_g) + w + \epsilon \tag{II.34}$$

$v_{th}(D_g)$ est la vitesse terminale de chute d'une goutte de diamètre D_g. La section 3.2 donne la paramétrisation utilisée dans ce travail pour v_{th}. On se reportera à l'annexe C pour une discussion de cette relation et sa comparaison relativement à d'autres relations existant dans la littérature. Par ailleurs, w est la vitesse du vent vertical, supposé constant (non aléatoire) sur la période considérée (par exemple la durée nécessaire pour enregistrer un spectre Doppler). w est supposée affecter toutes les gouttes identiquement. ϵ représente la partie stochastique de la relation vitesse/diamètre. On considère souvent dans les études sur les spectres Doppler que cet aléa est gaussien $\epsilon \sim \mathcal{N}(0, \sigma^2)$. Dans ce cas-là, la variabilité de la vitesse de chute autour de $v_{th} + w$ est due à la turbulence, qui est distribuée suivant une loi normale de moyenne nulle et d'intensité σ^2. Mais cette variabilité peut aussi être provoquée, entre autre, par des gouttes ne tombant pas à leur vitesse terminale de chute, indépendamment du vent et de la turbulence, notamment car elles viendraient d'exploser ou de percuter une autre goutte. La section 3.3 revient rapidement sur ces questions.

3.2 Paramétrisation utilisée dans cette thèse

Beaucoup de relations ont été proposées dans la littérature pour exprimer la vitesse terminale de chute de gouttes (voir Annexe C). Un grand nombre de ces relations prennent pour référence les mesures réalisées en laboratoire par Gunn and Kinzer (1949).

Dans cette étude, nous utiliserons pour v_{th} l'expression proposée par Atlas et al. (1973) (notée v_{atlas}), avec une correction pour l'altitude (voir ci-après). Cette relation conduit à des écarts inférieurs à $\pm 2\%$ par rapport aux mesures de Gunn and Kinzer (1949) sur la gamme de diamètres $(0.6 - 5.8mm)$. L'expression de Atlas et al. (1973) s'écrit :

$$v_{atlas}(D_g) = 9.65 - 10.3e^{-6D_g} \tag{II.35}$$

Avec toujours D_g le diamètre de la goutte considérée en cm. Parmi les études utilisant cette expression, on peut noter les restitutions de DSD à partir de mesures radars en UHF de Williams et al. (2000), ainsi que l'étude de Tridon and Battaglia (2015) qui l'utilise pour des diamètres au-dessus de $0.86mm$ et avec une correction pour l'altitude.

Par ailleurs, la force de traînée subie durant leur chute par les gouttes d'eau les freine. Et l'intensité de la force de traînée dépend de la masse volumique de l'air. Or celle-ci dépend des conditions de pression et de température, qui elles-même varient fortement avec l'altitude. On comprend donc que le freinage subi par les gouttes dépendra de l'altitude. La pression atmosphérique baissant lorsqu'on s'élève dans l'atmosphère, on peut supposer que la traînée baissera également, et que donc la vitesse terminale de chute des gouttes augmentera. L'Annexe C revient sur ces questions.

METEK (1998) propose une correction sous forme polynomiale, dérivée de la correction proposée par Foote and Du Toit (1969). Dans ce travail, nous utiliserons cette expression. Ainsi la vitesse terminale de chute des gouttes v_{th} s'écrit comme une fonction du diamètre D_g et de l'altitude z :

$$v_{th}(D_g, z) = \left[1 + 3.68.10^{-5} z + 1.71.10^{-9} z^2\right] v_{atlas}(D_g) \tag{II.36}$$

3.3 Écarts à la vitesse terminale de chute

Plusieurs études (Montero-Martínez et al. (2009), Villermaux and Eloi (2011), Larsen et al. (2014)) ont récemment noté qu'un nombre significatif de gouttes semblaient ne pas tomber à leur vitesse terminale de chute. Indépendamment des artefacts provoqués par les instruments de mesure eux-même (disdromètres), on peut penser que certaines gouttes ne suivent pas leur vitesse terminale car elles ont subi une modification de leur vitesse (collision avec une autre goutte, turbulence, explosion, etc.) et n'ont pas encore eu le temps de retrouver leur vitesse limite. Montero-Martínez et al. (2009), par exemple, remarque qu'un nombre significatif de gouttes tombent plus vite que leur vitesse terminale. En traçant l'histogramme des rapports des vitesses mesurées sur les vitesses terminales, il relève un maximum aux alentours de 2.8 (c'est-à-dire qu'un nombre important de gouttes tombe 2.8 fois plus vite qu'elles devraient le faire). Ce maximum est d'autant plus prononcé que la pluie est forte. Montero-Martínez et al. (2009) et Larsen et al. (2014) expliquent ce résultat par le break-up : un nombre significatif de gouttes (d'autant plus grand que les gouttes sont nombreuses, et donc qu'il pleut fort), lorsqu'elles sont vues par le disdromètre, viennent d'en percuter une autre et d'exploser, et les fragments vont plus vite que leur vitesse terminale car ils n'ont pas eu le temps de ralentir. Villermaux and Eloi (2011) expliquent différemment ces mesures. Il considère qu'un nombre significatif de gouttes explosent seules (car trop grosses). En se basant sur des rapports de forces, il justifie théoriquement qu'un maximum devrait alors apparaître aux alentours de 2.5 fois la vitesse terminale, ce qui semble relativement cohérent avec Montero-Martínez et al. (2009).

Dans les données utilisées dans le cadre de cette thèse, on n'a pas retrouvé de maximum

aux alentours de 2.5 ou 2.8, comme indiqué par Montero-Martínez et al. (2009) et Villermaux and Eloi (2011)). Nos cas d'étude correspondent à des pluies stratiformes, faibles, présentant très peu de grosses gouttes. Dans ce cas de figure, on se doit de souligner l'absence d'occurrence du break-up comme de l'auto-explosion des gouttes, considérées comme les principales causes de ce maximum par ces auteurs.

4 Discrétisation des EDP

4.1 Introduction

Toutes les équations aux dérivées partielles modélisant les phénomènes physiques décrits précédemment sont des équations d'advection, sur des espaces de dimensions diverses (coordonnées de temps, d'espace, de diamètre). Ces EDP peuvent, lorsque la vitesse d'advection est constante et homogène, être résolues explicitement (sous réserve de disposer d'une fonction donnant l'état initial du système). De manière générale, cependant, les vitesses d'advection ne sont pas constantes (le vent, par exemple, peut dépendre du temps), ni homogènes (la vitesse des chutes des gouttes dépend de l'altitude et de leur diamètre, le vent peut dépendre de la position dans l'espace, etc.). Dans ce cas-là, pour résoudre de manière approchée ces EDP, on définit un problème discret correspondant, problème discret qui lui pourra être résolu de proche en proche, en propageant l'information dans le temps et dans l'espace selon la direction et la vitesse d'advection. Pour définir un tel problème discret, il faut tout d'abord discrétiser (mailler) le temps et l'espace. Cela requiert d'une part de découper le temps et l'espace en "boîtes" (mailles), mais aussi de choisir où seront définies les variables du problèmes (typiquement au centre ou sur les bords des mailles). Ensuite, il s'agit d'approximer les dérivées sur cet espace discret.

Nous nous limiterons dans cette partie théorique à une description à une dimension d'espace. L'annexe D présente le cas à deux dimensions d'espace. Nous noterons t le temps et x la coordonnée d'espace. $\phi(t,x)$ sera notre variable d'état, c'est-à-dire le champ advecté (champ de taux de pluie, de DSD, etc.), $v(t,x)$ la vitesse d'advection. Vu qu'il n'y a qu'une seule dimension spatiale, l'advection se fait naturellement selon x. L'équation d'advection s'écrit alors :

$$\frac{\partial \phi}{\partial t} + \frac{\partial}{\partial x}(v\phi) = 0 \tag{II.37}$$

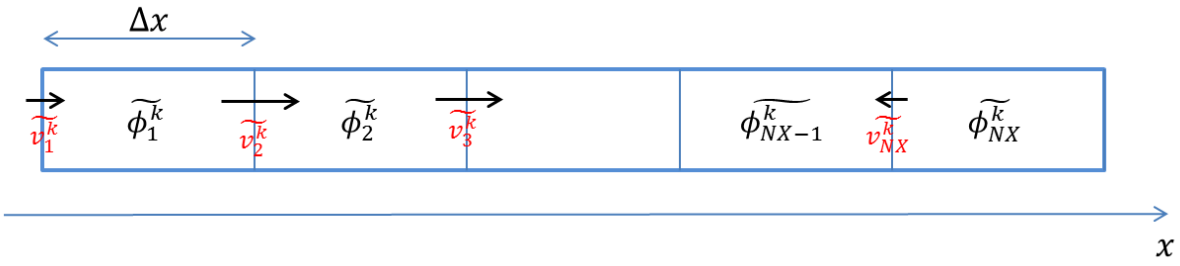

FIGURE II.9 – Représentation de la grille C d'Arakawa en une dimension. Le champ scalaire advecté ($\tilde{\phi}$) est évalué au centre des mailles (c'est-à-dire que l'on considère sa valeur homogène dans tout le point de maille) ; les vitesses d'advection ($\tilde{v}$) sont évaluées aux bords des mailles (permettant de définir naturellement les flux d'une maille à l'autre).

Dans tout ce travail, l'espace sera discrétisé selon une grille C d'Arakawa (Arakawa and

Lamb (1977)). Cela signifie que les vitesses seront évaluées sur les bords des mailles, et les scalaires (champs advectés) au centre des mailles. La Figure II.9 décrit cette grille. Les versions discrétisées des variables sont notées avec un tilde ($\tilde{v}$, $\tilde{\phi}$). L'exposant k désigne le temps, l'indice i la coordonnée d'espace. On note Δt et Δx les pas de temps et d'espace, et NT et NX le nombre de pas de temps et d'espace. Formellement, $\tilde{v}$ et $\tilde{\phi}$ sont donc des matrices de dimension $NT * NX$, dont les termes généraux s'écrivent $\tilde{v}_i^k$ et $\tilde{\phi}_i^k$, $k \in [|0 : NT - 1|]$, $i \in [|0 : NX - 1|]$.

Reste à présent à proposer un schéma numérique permettant d'approximer l'Eq.(II.37). Dans la suite, nous donnerons tout d'abord les équations du schéma upstream ainsi que ses conditions de stabilité. Par la suite, nous présenterons un autre schéma, celui de Smolarkiewicz (Smolarkiewicz (1983)). Finalement, nous comparerons rapidement les caractéristiques de ces deux schémas sur une application simple. L'Annexe D revient plus en détails sur la stabilité et la diffusivité de ces schémas.

4.2 Schéma upstream

Le schéma upstream permet de propager un champ discrétisé initial ($\tilde{\phi}_i^0$) à toutes les dates, d'itération temporelle en itération temporelle. Supposons connu le champ à la k-ième itération. L'équation permettant d'évaluer le champ à la $(k+1)$-ième itération s'écrit :

$$\tilde{\phi}_i^{k+1} = \tilde{\phi}_i^k - F\left(\tilde{\phi}_i^k, \tilde{\phi}_{i+1}^k, \tilde{v}_i^k, \Delta x\right) + F\left(\tilde{\phi}_{i-1}^k, \tilde{\phi}_i^k, \tilde{v}_{i-1}^k, \Delta x\right) \tag{II.38}$$

Avec la fonction F définie dans ce cadre par :

$$F(a, b, v, \Delta x) = \begin{cases} \frac{v\Delta t}{\Delta x} a & \text{si } v > 0 \\ \frac{v\Delta t}{\Delta x} b & \text{si } v < 0 \end{cases} \tag{II.39}$$

En se plaçant dans le cas où les vitesses sont positives aux deux bords de la maille considérée ($\tilde{v}_i^k > 0$ et $v_{i-1}^k > 0$), l'Eq.(II.38) se réécrit sous la forme plus simple :

$$\tilde{\phi}_i^{k+1} = \tilde{\phi}_i^k + \tilde{v}_{i-1}^k \frac{\Delta t}{\Delta x} \tilde{\phi}_{i-1}^k - \tilde{v}_i^k \frac{\Delta t}{\Delta x} \tilde{\phi}_i^k \tag{II.40}$$

Une condition nécessaire et suffisante pour que le schéma upstream soit stable est que la discrétisation vérifie la condition CFL (Courant-Friedrichs-Lewy), qui s'écrit dans ce cadre :

$$\max_{k,i}(\tilde{v}_i^k) \frac{\Delta t}{\Delta x} < 1 \tag{II.41}$$

Le schéma upstream présente néanmoins l'inconvénient d'être assez diffusif (voir plus bas le commentaire de la Figure II.10 ou une discussion théorique dans l'Annexe D). Dans la section suivante, nous introduisons donc le schéma de Smolarkiewicz.

4.3 Schéma de Smolarkiewicz

Le schéma de Smolarkiewicz a été développé pour limiter la diffusion numérique lors de la résolution de systèmes de grande taille (Smolarkiewicz (1983)). Nous conservons les notations discrètes définies plus haut, ainsi que l'expression de la fonction F de l'Eq.(II.39). Pour la k-ième itération temporelle, on définit un état intermédiaire de calcul, vecteur noté $\tilde{\phi}^*$. Les équations permettant d'évaluer le champ $\tilde{\phi}$ à la $(k+1)$-ième itération s'écrivent :

$$\begin{cases} \tilde{\phi}_i^* & = \tilde{\phi}_i^k - F\left(\tilde{\phi}_i^k, \tilde{\phi}_{i+1}^k, \tilde{v}_i^k, \Delta x\right) + F\left(\tilde{\phi}_{i-1}^k, \tilde{\phi}_i^k, \tilde{v}_{i-1}^k, \Delta x\right) \\ \tilde{\phi}_i^{k+1} & = \tilde{\phi}_i^* - F\left(\tilde{\phi}_i^*, \tilde{\phi}_{i+1}^*, \tilde{v}_i^*, \Delta x\right) + F\left(\tilde{\phi}_{i-1}^*, \tilde{\phi}_i^*, \tilde{v}_{i-1}^*, \Delta x\right) \\ \tilde{v}_i^* & = S_c \frac{\left(|\tilde{v}_i^k|\Delta x - \Delta t\left(\tilde{v}_i^k\right)^2\right)\left(\tilde{\phi}_{i+1}^* - \tilde{\phi}_i^*\right)}{\left(\tilde{\phi}_i^* + \tilde{\phi}_{i+1}^* + \epsilon\right)\Delta x} \end{cases} \tag{II.42}$$

La première ligne, permettant d'obtenir $\tilde{\phi}_i^*$, effectue une advection upstream classique. La deuxième ligne, aboutissant à $\tilde{\phi}_i^{k+1}$, représente une advection upstream à la vitesse d'antidiffusion non constante $\tilde{v}^*$. Cette deuxième équation limite la diffusion numérique induite par la première. La troisième ligne définit $\tilde{v}^*$. On peut trouver en Annexe D une justification de ces équations et de l'expression de $\tilde{v}^*$, ainsi que du coefficient Sc. Smolarkiewicz (1983) montre théoriquement que, pour la condition CFL définie pour le schéma upstream (Eq.(II.41)), le schéma de Smolarkiewicz est également stable, sous réserve d'avoir un coefficient $Sc < 4$. Smolarkiewicz (1983) préconise (empiriquement) de prendre Sc dans l'intervalle $[1 : 1.08]$. Le coefficient ϵ au dénominateur de l'expression de $\tilde{v}^*$ assure une valeur $\tilde{v}_i^* = 0$ dans le cas où $\tilde{\phi}_i^* = \tilde{\phi}_{i+1}^* = 0$.

La Figure II.10 présente les sorties des schémas upstream et de Smolarkiewicz pour un exemple de champ $2D$. Le champ initial (à gauche), ici une carte de taux de pluie, à la résolution spatiale $0.4 * 0.4km^2$, est advecté avec un pas de discrétisation de $1s$ vers le sud-est à la vitesse $(13.6; 10.4)m/s$, en haut via le schéma upstream, en bas via le schéma de Smolarkiewicz. On montre les résultats après $5min$ d'advection (centre), et $10min$ (droite). On constate immédiatement que le schéma upstream estompe rapidement les structures fines du champ initial et perd les valeurs extrêmes en lissant et étendant le champ. Le schéma de Smolarkiewicz, bien que présentant également une diffusion numérique manifeste, semble mieux préserver les propriétés du champ original.

Le schéma de Smolarkiewicz a été développé dans les deux buts conjoints d'être peu diffusif et bien adapté pour des problèmes en grandes dimensions, typiquement pour des applications en géophysique. Il convient donc bien pour les applications qui nous intéressent ici. On a en effet vu que dans le chapitre IV, nous utiliserons un modèle d'advection de champs de pluie sur des périodes courtes, sans sources ni déformations. L'un des objectifs d'un tel travail est d'être capable de restituer des valeurs extrêmes (notamment pour des applications en hydrologie). Un schéma trop diffusif empêcherait de produire de telles valeurs en quelques itérations. De même, dans le chapitre V, nous chercherons à restituer des DSD en haut de la colonne d'eau,

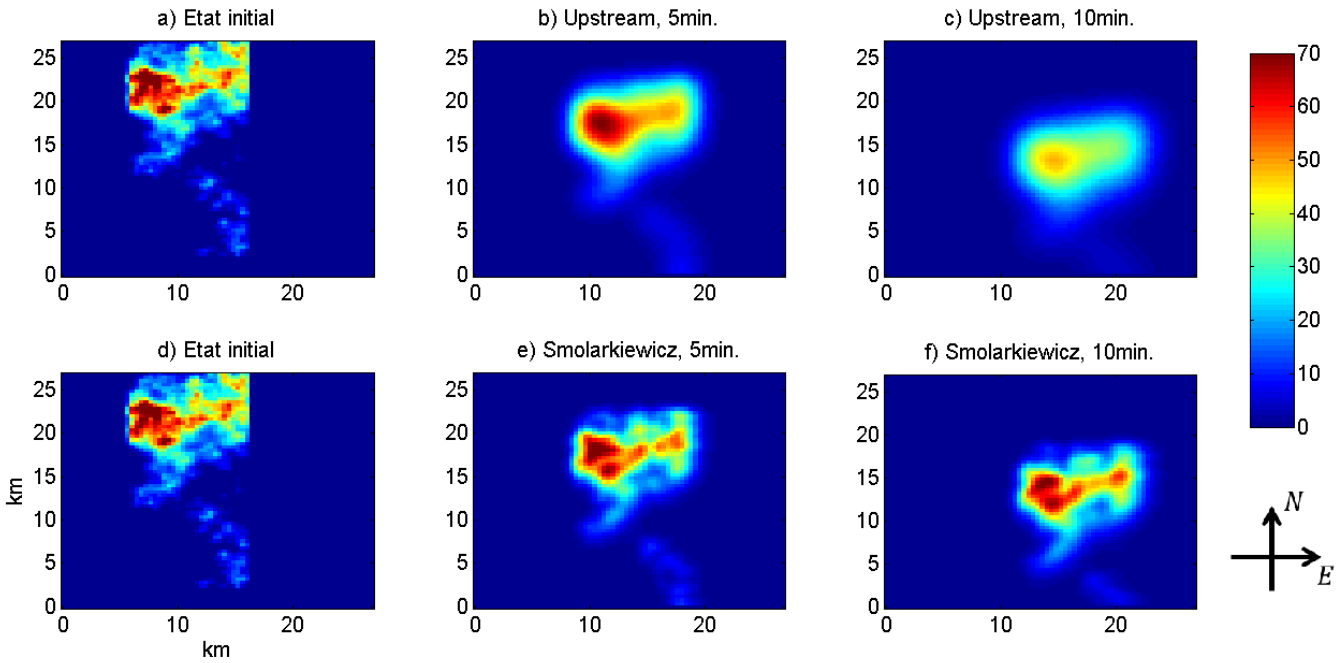

FIGURE II.10 – Expériences d'advection numérique avec les schémas upstream et de Smolarkiewicz. La résolution spatiale est de $0.4 * 0.4km^2$, la résolution temporelle de $1s$. La vitesse d'advection est de $13.6m/s$ vers le sud et $10.4m/s$ vers l'est. a) champ initial. b) champ initial advecté sur $5min$ par le schéma upstream. c) champ initial advecté sur $10min$ par le schéma upstream. d) champ initial. e) champ initial advecté sur $5min$ par le schéma de Smolarkiewicz. f) champ initial advecté sur $10min$ par le schéma de Smolarkiewicz.

donc potentiellement à quelques centaines/milliers de mètres d'altitude, DSD qui seront ensuite advectées jusqu'au sol où elles seront confrontées à des observations. Il faut que cette advection se fasse sans trop de diffusion parasite afin de pouvoir utiliser la plus grande part possible de l'information disponible au sol.

Partie III

Assimilation de données et mise en oeuvre informatique

La physique en général, et la météorologie en particulier, utilisent fréquemment des modèles dits déterministes. Ces modèles consistent généralement en la discrétisation de systèmes d'équations aux dérivées partielles visant à représenter la physique des phénomènes naturels sous-jacents. Ils sont déterministes dans la mesure où la connaissance des conditions initiales et des conditions aux limites les déterminent entièrement. Autrement dit, ces informations permettent de déterminer l'état du système en tout point de maille (temps et espace). De tels modèles ont fait la preuve de leur utilité, notamment dans le domaine des prévisions météorologiques. Ils sont néanmoins entachés d'erreurs multiples, qui peuvent être divisées en trois catégories. Tout d'abord celles quantifiant l'écart entre les lois de la nature et leur représentation sous forme d'équations différentielles. Les premières ne peuvent être approximées qu'imparfaitement, et de manière incomplète, par les secondes. Ensuite, on note que ces équations aux dérivées partielles ne sont résolues via une phase de discrétisation que de manière approximative. Une erreur est donc commise lors de cette phase. Enfin, les conditions initiales et conditions aux limites ne peuvent également être connues qu'imparfaitement, d'où une troisième source d'erreur.

Afin de déterminer l'état d'un système, on pourrait également ne se baser que sur les observations disponibles (sans l'aide du modèle). Là encore, nous nous devons de souligner les limites d'une telle approche. Tout d'abord, les observations ne peuvent être que parcellaires (on ne peut par exemple pas, pour des applications océanographiques ou météorologiques, avoir une bouée mesurant la température en chaque point à la surface de l'océan). Et même aux points où des mesures sont disponibles, celles-ci peuvent être indirectes et entachées d'erreurs. Elles ne représentent pas forcément la situation sur l'ensemble du point de maille considéré. Les observations peuvent également être de natures différentes et difficilement fusionnables sans informations a priori. Elles peuvent être distribuées dans le temps et donc là aussi difficilement comparables sans informations sur la dynamique du système sous-jacent.

L'assimilation de données va consister à utiliser à la fois la dynamique apportée par un modèle a priori et les observations disponibles pour estimer les états du système, ou les paramètres du modèle. L'objectif est bien sûr que cette estimation soit meilleure, dans un sens à définir, que celle qui aurait pu être réalisée en utilisant uniquement les aspects déterministes (a priori / modèle) ou uniquement les aspects stochastiques reposant sur les observations.

Dans ce chapitre, nous ferons tout d'abord un rapide tour d'horizon des différents algorithmes d'assimilation de données, et introduirons un certain nombre de notations. Ensuite, nous reviendrons plus en détails sur la théorie de l'algorithme 4D-VAR utilisé dans ce travail, dans un cadre classique. Puis nous donnerons des informations sur son implémentation numérique, en nous basant sur un exemple détaillé très simple, et nous présenterons le logiciel YAO utilisé à cet effet. Enfin, nous présenterons quelques extensions permettant d'adapter l'algorithme précédemment décrit à un cadre moins rigide.

1 Généralités sur l'assimilation de données

Les sources principales utilisées dans cette première partie présentant rapidement l'assimilation de données et ses principaux algorithmes sont les suivantes. Le livre Reich and Cotter (2015), et notamment son préambule, donne une vision d'ensemble de l'assimilation de données. Le cours de Marc Boquet (Bocquet (2005)) et le rapport de Nicolas Daget (Daget (2007)) présentent de manière détaillée les différents algorithmes d'assimilation de données classiquement utilisés en géophysique. La thèse Tran (2009) apporte des éclaircissements notamment sur l'assimilation bayésienne et les filtres particulaires, peu utilisés en géophysique. Le livre d'Eugenia Kalnay (Kalnay (2003)) revient sur la prévision météorologique numérique en général. Enfin, les notations utilisées dans cette partie sont celles définies par Ide et al. (1997).

1.1 Quelques notations

On considère un système physique, discrétisé, et caractérisé par un état vrai x^t. On cherche à estimer "au mieux" cet état vrai. On note x^a l'estimateur correspondant (état analysé).

Pour cela, on dispose d'observations, directes ou indirectes, de ce système physique. Ces observations, notées y, sont réparties dans le temps. Les observations disponibles à la date t_k sont notées y_k (l'état vrai étant alors x^t_k). Comme on l'a vu dans le préambule de ce chapitre, ces observations sont entachées d'erreurs.

On suppose par ailleurs que l'on dispose d'un modèle dynamique, a priori non linéaire (et entaché lui aussi d'erreurs). On note $M_{k,k+1}$ le modèle propageant un état de la date t_k à la date t_{k+1}. On note M_k le modèle propageant un état de la date t_0 à la date t_k. On a donc : $M_k = M_{k-1,k} \circ ... \circ M_{0,1}$.

L'assimilation de données va consister à estimer l'état x^a le plus vraisemblable (la notion de "plus vraisemblable" restant à déterminer) vues ces informations. Notons que les différents états définis ici (x^t, x^a et y) sont a priori des tableaux de dimensions quelconques. (Dans la suite, on supposera qu'ils sont systématiquement réécrits sous forme de vecteurs colonnes.)

1.2 Approche bayésienne, filtres particulaires

L'assimilation de données bayésienne consisterait à déterminer la distribution de probabilités complète de l'état vrai x^t sachant les observations y, ou, dans une vision séquentielle, à la date t_k, la loi de probabilité de x^t_k sachant $y_{p\leq k}$. Si l'on parvient à estimer cette distribution de probabilité, on peut ensuite choisir un estimateur (x^a) exprimant la notion subjective "d'état le plus vraisemblable". Cet estimateur peut alors être un estimateur de variance minimale, ou bien un estimateur de probabilité maximale (le maximum de la distribution trouvée), etc. Bocquet (2005) ou Tran (2009) reviennent plus en détails sur cette approche.

Il est délicat de déterminer ces distributions de probabilités quand le plus souvent les erreurs sont mal connues, qui plus est dans des systèmes de grande taille dont les modèles et

opérateurs sont non linéaires. Les filtres particulaires (voir Reich and Cotter (2015) ou Tran (2009)) utilisent des méthodes de Monte Carlo pour les approximer, mais impliquent un coût important pour générer un nombre satisfaisant d'échantillons. Nous n'y reviendrons pas plus en détails par la suite.

1.3 Assimilation sans dépendance temporelle

1.3.1 Interpolation statistique / BLUE

On se place à présent dans un cas plus simple, sans dépendance temporelle (tout se passe à une date fixée), et sans modèle dynamique.

On suppose que l'on dispose d'un état d'ébauche x^b (background). Celui-ci est une première estimation de notre état vrai, que l'on va chercher à améliorer grâce aux observations. Cette ébauche peut par exemple être une valeur climatologique, ou le résultat d'une estimation antérieure réalisée avec d'autres observations. L'erreur d'ébauche est la variable aléatoire mesurant l'écart entre l'état vrai et l'état d'ébauche (x^b - x^t). On suppose que les cumulants d'ordre 1 et 2 (espérance et variance) de cette variable aléatoire existent. L'erreur d'ébauche est de plus considérée sans biais (espérance connue et nulle). On note $\boldsymbol{B}$ sa matrice de variance/covariance (supposée définie positive), les termes diagonaux caractérisant la confiance à accorder aux termes de l'ébauche et les termes non diagonaux les corrélations entre eux.

On dispose par ailleurs d'observations (y). L'erreur de mesure qui leur est associée, appelée erreur d'observation (supposée sans biais), correspond à l'écart $y - \boldsymbol{H}x^t$ entre l'observation et la projection de l'état vrai dans l'espace des observations via un opérateur $\boldsymbol{H}$. Cet opérateur $\boldsymbol{H}$ est appelé opérateur d'observation. Il est ici supposé linéaire.

On note $\boldsymbol{R}$ la matrice (définie positive) de variance/covariance caractérisant l'erreur d'observations.

On note que si l'ébauche et les observations sont généralement bien définies, il peut être plus délicat de définir (i.e. de "remplir") les matrices de variance/covariance d'erreur d'ébauche et d'observations.

Dans le cadre de l'interpolation optimale, on définit ensuite l'état analysé (x^a) comme vérifiant les trois points suivants:

1. x^a est une combinaison linéaire des informations contenues dans l'ébauche et les observations.

La variable aléatoire caractérisant l'erreur d'analyse ($x^t - x^a$) est :

2. sans biais.

3. de variance minimale (c'est-à-dire que la trace de sa matrice de variance/covariance est minimale).

L'estimateur x^a vérifiant ces conditions est appelé $BLUE$, Best Linear Unbiased

Estimator. Sous les conditions définies plus haut (opérateur d'observation linéaire, erreurs sans biais et caractérisées par leurs matrices de variance/covariance), cet estimateur existe et est défini analytiquement à partir des données du problème. Il est bel et bien "BLUE", dans le sens où le point (1) ci-dessus assure le "linear", le (2) le "unbiased". Le "best" est défini ici comme l'estimateur minimisant la variance (3). On peut noter (voir Bocquet (2005)) que si toutes les distributions de probabilités des erreurs sont gaussiennes, il est également le meilleur estimateur au sens bayésien (les différents critères possibles d'estimation se recoupant).

Son calcul explicite nécessite néanmoins l'inversion de matrices potentiellement de très grandes tailles (quoique creuses), ce qui peut compliquer singulièrement sa mise en oeuvre effective.

1.3.2 Equivalence variationnelle : 3D-VAR

Le problème défini dans la section précédente peut également être vu sous un angle variationnel, c'est-à-dire comme un problème d'optimisation. Définissons la fonction de coût J suivante (caractérisant les critères que la solution devra minimiser) :

$$J(x) = \left(x - x^b\right) \boldsymbol{B}^{-1} \left(x - x^b\right)^T + (y - \boldsymbol{H}x)\, \boldsymbol{R}^{-1} (y - \boldsymbol{H}x)^T \tag{III.1}$$

On montre (voir Bocquet (2005)) que le x minimisant cette fonctionnelle est défini, unique et est le x^a défini plus haut, c'est-à-dire le $BLUE$. On peut donc exprimer le même problème sous cette forme variationnelle. C'est ce que l'on nomme l'assimilation 3D-VAR ($3D$ pour les trois dimensions potentielles d'espace). Cette vision présente deux avantages significatifs par rapport au $BLUE$. Tout d'abord elle ne nécessite pas d'inverser formellement des matrices, ce qui peut fournir un gain informatique conséquent. Par ailleurs, dans cette formulation, rien n'empêcherait l'opérateur d'observation $\boldsymbol{H}$ d'être non linéaire. La minimisation numérique d'une fonctionnelle passe généralement par le calcul de son gradient. Or, dans le cas linéaire, ce gradient s'écrit :

$$\nabla J(x) = 2\boldsymbol{B}^{-1} \left(x - x^b\right) - 2\boldsymbol{H}^T \boldsymbol{R}^{-1} (y - \boldsymbol{H}x) \tag{III.2}$$

Dans le cas non linéaire, l'opérateur d'observation n'est plus une matrice et s'écrit, appliqué à x, sous la forme $H(x)$. En notant le linéaire tangent $\boldsymbol{H(x)}$ (i.e. la jacobienne de l'opérateur H appliquée au point x) (c'est une matrice pouvant contenir une ou plusieurs composantes de x), on a :

$$\nabla J(x) = 2\boldsymbol{B}^{-1} \left(x - x^b\right) - 2\boldsymbol{H(x)}^T \boldsymbol{R}^{-1} (y - \boldsymbol{H}x) \tag{III.3}$$

Étant en mesure d'estimer le gradient de J, on peut ensuite utiliser un minimiseur numérique pour chercher à minimiser J. Néanmoins, dans ce cas-là, rien n'assure plus l'existence et l'unicité du minimum. Par ailleurs, dans la mesure où l'on utilise un algorithme

de minimisation numérique, il faut lui fournir un point de départ (noté x^g, "first guess"). Le résultat ensuite peut fortement dépendre de ce point de départ, notamment si la fonction J possède un grand nombre de minima locaux. Notons que généralement il est raisonnable de prendre $x^g = x^b$.

1.4 Ajout d'une dépendance temporelle : le filtre de Kalman

On revient à présent au contexte dans lequel les observations sont distribuées dans le temps et où l'on dispose d'un modèle numérique permettant de propager les états du système d'un pas de temps au suivant en simulant au mieux ce que l'on connaît de la physique sous-jacente.

Dans un premier temps, on suppose le modèle linéaire. L'opérateur $M_{k,k+1}$ défini plus haut peut donc s'écrire sous forme matricielle : $\boldsymbol{M_{k,k+1}}$. On définit l'erreur modèle à la date t_{k+1} comme $x^t_{k+1} - M_{k,k+1}(x^t_k)$. On note que cette erreur modèle englobe à la fois l'erreur commise en représentant les lois naturelles sous forme d'équations différentielles et celle liée à la discrétisation de ces dernières. Comme précédemment, on suppose que cette erreur possède une distribution de probabilités. On la suppose sans biais (espérance nulle) et on note sa matrice de variance/covariance $\boldsymbol{Q_k}$.

On suppose linéaire l'opérateur d'observation pour toute date t_k, noté $\boldsymbol{H_k}$. De même que précédemment, l'erreur d'observation $(y_k - \boldsymbol{H_k} x^t_k)$ est supposée, pour chaque date, sans biais et on note $\boldsymbol{R_k}$ sa matrice de variance/covariance. On suppose les différentes erreurs décorrélées entre elles et, surtout, décorrélées dans le temps (ce qui est souvent plus contestable).

Le filtre de Kalman est un filtre séquentiel, c'est-à-dire que l'on va chercher l'état optimal (dans un sens à définir) à une date donnée sachant les observations disponibles à cette date, puis propager cet état et l'erreur associée jusqu'à la date suivante grâce au modèle, puis recommencer. Il faut donc l'initialiser pour la date initiale (t_0), avec un état d'ébauche x^b_0 et une matrice de covariance d'erreur associée $\boldsymbol{B_0}$. Supposons à présent que pour la date t_k, on dispose d'un état d'ébauche x^b_k (qui n'est autre que l'état d'analyse à la date t_{k-1} propagé par le modèle jusqu'à t_k) et d'une matrice de covariance d'erreur associée $\boldsymbol{B_k}$. Le problème qui se pose alors à la date t_k est le même que précédemment, dans le cas du $BLUE$. En appliquant cette méthode, on calcule alors l'état analysé x^a_k, ainsi que la matrice de covariance d'erreur de l'analyse. On propage ensuite cet état d'analyse jusqu'à la date t_{k+1} par le modèle (on obtient ainsi l'ébauche pour la date suivante) et on utilise également le modèle et sa matrice de covariance d'erreur pour calculer la matrice de covariance associée à cette nouvelle ébauche à t_{k+1}. On recommence ensuite ce processus jusqu'à la fin de la période d'assimilation.

On montre que les états d'analyse obtenus dans ce cadre sont également optimaux, dans le sens d'une variance de l'erreur minimale. Comme pour le $BLUE$, il est également le meilleur estimateur au sens bayésien sous réserve que toutes les erreurs aient des lois de probabilité gaussiennes. C'est pourquoi il est généralement qualifié de filtre gaussien. Si ce n'est pas le cas, ses équations restent applicables mais la propriété ci-dessus n'est plus vérifiée.

Par ailleurs, l'inconvénient de ce formalisme est qu'il ne fonctionne que dans le cas où le modèle et les opérateurs d'observation sont linéaires, ce qui est rarement vérifié en pratique. Le

filtre de Kalman étendu contourne cette difficulté. On note également que le filtre de Kalman nécessite dans la phase d'analyse d'effectuer des produits de matrices de grande taille, ce qui peut poser problème au niveau du coût de calcul.

1.4.1 Filtre de Kalman étendu

La non linéarité du modèle ($M_{k,k+1}$) n'empêche bien sûr pas la propagation des états d'une date à l'autre mais empêche la propagation des matrices de covariance d'erreurs. Le filtre de Kalman étendu consiste simplement à linéariser les opérateurs autour des points d'application en utilisant leurs jacobiennes et en supposant les états calculés suffisamment proches des états vrais (ou les opérateurs "suffisamment réguliers") pour que cette hypothèse reste vraie. Par exemple, avec les notations précédentes, on écrira, au premier ordre :

$$\begin{aligned} M_{k,k+1}(x_k^a) &= M_{k,k+1}(x_k^t + (x_k^a - x_k^t)) \\ &\simeq M_{k,k+1}(x_k^t) + \boldsymbol{M_{k,k+1}(x_k^a)}.(x_k^a - x_k^t) \end{aligned} \tag{III.4}$$

Avec $\boldsymbol{M_{k,k+1}(x_k^a)}$ la jacobienne de $M_{k,k+1}$ évaluée au point x_k^a. Notons que pour des modèles complexes, l'écriture de cette jacobienne peut être assez lourde. Nous reviendrons sur cette question dans la section sur l'assimilation 4D-VAR.

Cette formulation permet d'estimer la matrice de covariance d'erreur d'ébauche à la date t_{k+1} à partir de celle à la date t_k et de $\boldsymbol{Q_k}$, matrice de covariance d'erreur modèle (voir Bocquet (2005) pour les calculs explicites). On note que dans ce cadre, même si les équations du filtre de Kalman restent applicables, rien ne garantit plus l'optimalité de la solution. Par ailleurs, la qualité de l'estimé dépend de la régularité des opérateurs. Si ceux-ci sont fortement non linéaires, l'erreur commise via la linéarisation peut être tout à fait significative.

1.4.2 Filtre de Kalman d'ensemble

Un algorithme souvent utilisé pour tenter de tenir compte de cette non linéarité est le filtre de Kalman d'ensemble. Comme on l'a vu, le principal problème posé par la non linéarité est qu'elle rend impossible la propagation des matrices de covariance d'erreur. Avec le filtre de Kalman d'ensemble, on ne va plus propager ces matrices mais les estimer à chaque date à partir d'un échantillon d'états supposés représentatifs de la distribution de probabilités sous-jacente.

Pour cela, on applique l'algorithme suivant :

- on choisit un nombre d'échantillons (L). On perturbe ensuite l'ébauche (à t_0) L fois. On crée ainsi L états d'ébauche.

- On se place à la date t_k. On suppose que l'on a L ébauches pour cette date, ainsi que la matrice de covariance d'erreur associée. On perturbe également L fois le vecteur contenant les observations disponibles (y_k). On calcule explicitement la matrice de covariance associée à ces L jeux d'observations (sans la propager, directement à partir des L échantillons).

- On applique les équations du filtre de Kalman pour calculer les états analysés correspondant.

- On calcule la matrice de covariance d'erreur de l'état analysé à partir des L états qu'on vient de calculer.

- On propage ces états jusqu'à la date t_{k+1} grâce au modèle (non linéaire).

- On calcule la matrice de covariance d'erreur d'ébauche à la date t_{k+1} à partir de ces L états propagés.

L'avantage de cet algorithme est qu'il ne requiert pas d'approximation linéaire du modèle. Par ailleurs, on espère que les "sauts" du modèle (fortes non linéarités) seront retranscrits dans l'échantillon. On note cependant (voir Bocquet (2005)) que bien que l'on ne propage plus explicitement les matrice de covariance d'erreurs, les équations utilisées pour calculer les états analysés sont toujours celles du $BLUE$, ce qui suppose toujours une hypothèse de gaussianité des erreurs. Par ailleurs, on peut noter que la justification de cet algorithme est semi-empirique (notamment le fait de bruiter également les observations). Enfin, il requiert un nombre assez conséquent d'échantillons et peut vite devenir coûteux.

1.5 4D-VAR

On a vu que dans un cadre non temporel, on pouvait faire correspondre à l'interpolation optimale une approche variationnelle (le $3D\text{-}VAR$, consistant à minimiser une fonction de coût) permettant de réduire le temps de calcul et fournissant un cadre plus propice au traitement de problèmes non linéaires. On peut faire de même ici avec le filtre de Kalman. En introduisant une approche variationnelle, le $4D\text{-}VAR$, on peut montrer que cette approche correspond exactement au filtre de Kalman dans le cas où les opérateurs d'observation et le modèle sont linéaires. La linéarité n'est cependant pas requise pour que la méthode soit applicable. Cet algorithme, que l'on utilise dans cette thèse, est décrit plus en détails dans la partie suivante.

2 Assimilation 4D-VAR

On se place ici dans le cadre classique de l'assimilation 4D-VAR en géophysique. Notons tout d'abord que les états du système à une date donnée t_k (par exemple l'état d'ébauche x_k^b ou d'analyse x_k^a) peuvent être un vecteur (si on travaille sur un espace à une dimension), une matrice, une matrice 3D, ou plus généralement un champ quelconque. Les variables x^a, x^b (toutes dates) peuvent donc être, par exemple et selon les cas, des matrices 2D, 3D ou 4D. C'est de ce dernier cas - classique pour les modèles météorologiques ou océanographiques - que vient le nom 4D-VAR. On l'utilise néanmoins systématiquement, dès lors que le temps est présent, afin de ne pas le confondre avec le 3D-VAR vu plus haut. Notons également que dans ce chapitre, seule la dépendance au temps sera explicitement mentionnée dans les noms de variables (généralement par l'indice "k"). Comme dans ce qui précède, la dépendance spatiale sera systématiquement sous-entendue. Les exposants ("a", "b", "g", "t") désignent uniquement ici les différents types d'états du système considérés (analyse, ébauche, initialisation, vrai).

On suppose par ailleurs que le modèle numérique de propagation ne requiert pas de conditions aux limites, ou que celles-ci sont fixées. On considère, de la même façon, sa paramétrisation fixée. Sous ces conditions, les états du système d'ébauche (x^b) et d'analyse (x^a) se déduisent, via le modèle, des états à l'instant initial (i.e. à t_0). On a ainsi : $x_k^b = M_k(x_0^b)$ et de la manière $x_k^a = M_k(x_0^a)$, où M_k est la composition des opérateurs propageant successivement les états de pas de temps en pas de temps jusqu'à t_k. L'assimilation va alors consister à déterminer l'état initial d'analyse (le x_0^a) qui, une fois propagé par le modèle, explique au mieux les observations disponibles, tout en restant proche de l'ébauche. Ce cadre n'est pas rigide. On pourrait par exemple fixer l'état initial et chercher à déterminer les conditions aux limites. La section 4 revient rapidement sur les implications qu'auraient un tel changement de cadre.

Comme dans le cas de l'assimilation 3D-VAR, on définit une fonction de coût, J (caractérisant les critères que la solution devra minimiser), comme :

$$J(x) = \left(x_0 - x_0^b\right) \boldsymbol{B}^{-1} \left(x_0 - x_0^b\right)^T + \sum_{k=0}^{NT} \left(y_k - H_k(x_k)\right) \boldsymbol{R_k}^{-1} (y_k - H_k(x_k))^T \qquad \text{(III.5)}$$

On s'est placé ici dans le cas général avec des opérateurs d'observation non linéaires.

Pour des commodités d'écriture, nous allons décomposer, comme Daget (2007), cette fonction de coût en deux composantes. La première, J^b, appelée le terme de rappel à l'ébauche, est composée de la première partie de l'équation (III.5) à droite du signe égal. La seconde, J^{obs}, appelée le terme de rappel aux observations, et composée du reste de la partie droite de l'équation (III.5). On a naturellement $J(x) = J^b(x) + J^{obs}(x)$. Comme leurs noms l'indiquent, J^b quantifie l'écart de l'état initial du système, x_0, avec le terme d'ébauche x_0^b, en tenant compte de la distribution des erreurs donnée par la matrice $\boldsymbol{B}$. Mathématiquement, ce

terme existait déjà tel quel dans les équations de l'assimilation 3D-VAR. Nous ne reviendrons donc plus dessus dans la suite de cette section. J^{obs} évalue l'écart entre les observations disponibles et les états du système projetés sur l'espace des observations. On note que cette formulation met en évidence le fait que le 4D-VAR, via la minimisation de la fonctionnelle J, va être une compétition entre un rapprochement à l'ébauche x_0^b ou aux observations y. Dans le cas où, par exemple, plusieurs ordres de grandeur séparent les valeurs des états du système et celles des observations, ou que l'on dispose d'un nombre d'observations très élevé, on pourrait être tenté d'ajuster les poids relatifs de J^b et J^{obs} dans la composition de J (en ajoutant par exemple un coefficient de poids α : $J(x) = \alpha J^b(x) + J^{obs}(x)$). On ne le fera jamais apparaître ici explicitement, dans la mesure où ce terme peut-être inclus dans les matrices $\boldsymbol{B}$ ou $\boldsymbol{R_k}$.

En réécrivant J^{obs} en utilisant le modèle, on peut obtenir une expression de J^{obs} ne dépendant plus que de l'état initial de l'inconnue (x_0) selon :

$$J^{obs}(x_0) = \sum_{k=0}^{NT} (y_k - H_k \circ M_k(x_0))\, \boldsymbol{R_k}^{-1} (y_k - H_k \circ M_k(x_0))^T \tag{III.6}$$

Plaçons-nous à présent dans le cas simplifié où tous les opérateurs sont linéaires. Il s'écrivent alors sous forme matricielle et on a donc :

$$J^{obs}(x_0) = \sum_{k=0}^{NT} (y_k - \boldsymbol{H_k M_k} x_0)\, \boldsymbol{R_k}^{-1} (y_k - \boldsymbol{H_k M_k} x_0)^T \tag{III.7}$$

Comme vu précédemment, l'opérateur M_k est la composition des opérateurs propageant successivement les états de pas de temps en pas de temps jusqu'à t_k. On note qu'ici, dans le cas linéaire, il s'agit simplement d'un produit matriciel : $\boldsymbol{M_k} = \boldsymbol{M_{k-1,k}} ... \boldsymbol{M_{0,1}}$.

Comme dans le cas du 3D-VAR, on peut montrer que l'estimateur x^a minimisant cette fonctionnelle dans le cas linéaire est le même que l'estimateur du filtre de Kalman.

Sa détermination, via un algorithme de minimisation numérique, passe généralement par le calcul de son gradient, qui s'écrit dans le cas linéaire :

$$\nabla J^{obs}(x_0) = -2 \sum_{k=0}^{NT} \boldsymbol{M_k}^T \boldsymbol{H_k}^T \boldsymbol{R_k}^{-1} (y_k - \boldsymbol{H_k M_k} x_0)\,. \tag{III.8}$$

Comme pour le 3D-VAR, le formalisme variationnel du 4D-VAR permet de passer au cas non linéaire. L'équation (III.6) définit la fonction de coût utilisée dans ce cas-là. Il s'agit à présent d'estimer son gradient. En notant comme précédemment $\boldsymbol{H_k(x_k)}$ le linéaire tangent (la jacobienne de l'opérateur H_k appliquée à l'état x_k) et $\boldsymbol{M_k(x)}$ l'opérateur M_k linéarisé au voisinage de l'état x, on a :

$$\nabla J^{obs}(x_0) = -2 \sum_{k=0}^{NT} \boldsymbol{M_k(x)}^T \boldsymbol{H_k(x)}^T \boldsymbol{R_k}^{-1} (y_k - H_k \circ M_k(x_0))\,. \tag{III.9}$$

On a vu que l'on avait $M_k = M_{k-1,k} \circ ... \circ M_{0,1}$. Si l'on parvient également à écrire chaque itération du modèle numérique sous la forme d'une composition de fonctions analytiques

(typiquement une fonction analytique par point de maille, répétée ensuite pour chaque point de maille), M_k se ramène à une composition de (très nombreuses) fonctions analytiques. $\boldsymbol{M_k(x)}$ se ramène alors au produit des matrices jacobiennes de chacune de ces fonctions évaluées à chaque date et chaque point de maille. Cette décomposition du linéaire tangent du modèle lors du calcul du gradient est appelée modèle adjoint. L'implémentation effective de son calcul constitue généralement la principale difficulté de la mise en oeuvre d'algorithmes 4D-VAR, notamment lorsque le modèle (et aussi éventuellement les opérateurs d'observation) sont lourds et complexes. De fortes non-linéarités dans ces opérateurs complexifient également la minimisation. La section 3 explicite son écriture sur un cas très simple.

Nous allons à présent illustrer le principe et le fonctionnement du 4D-VAR par deux schémas simples.

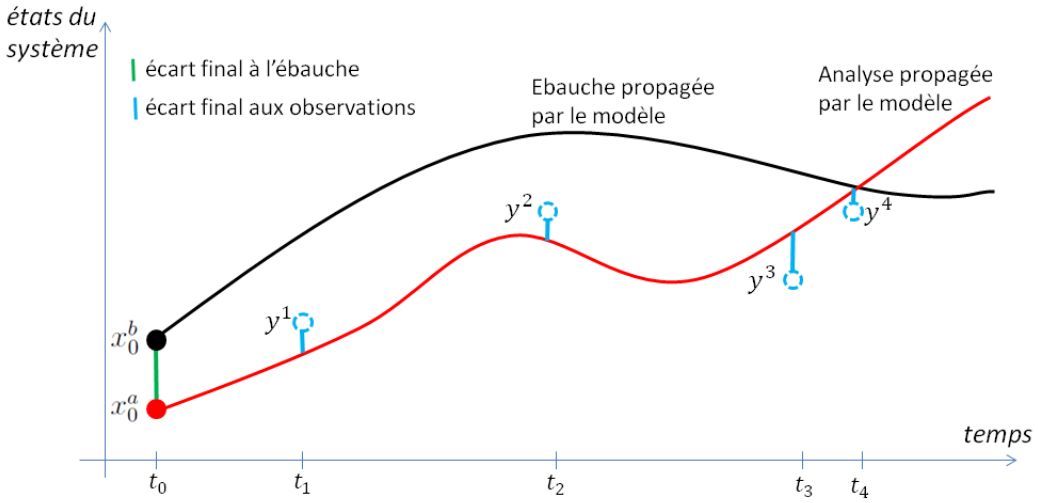

FIGURE III.1 – Schéma du résultat d'une expérience d'assimilation 4D-VAR.

La figure III.1 présente le résultat d'une expérience d'assimilation 4D-VAR pour un problème $1D$ (temps seul). L'espace des états est ici scalaire. Cela signifie que les états, à une date donnée t_k (x_k^t, x_k^a), sont des scalaires (et donc x^t ou x^a sont des vecteurs). On dispose ici d'une ébauche à l'instant initial x_0^b (point noir) qui peut être propagée dans le temps par le modèle (trait noir). 4 observations sont disponibles (cercles bleus). On note qu'ici ces observations sont situées dans le même espace que les états du système (on peut donc s'attendre à ce que les opérateurs d'observation $\boldsymbol{H_k}$ valent 1). Après avoir fait tourner un algorithme de minimisation numérique, on trouve l'état initial analysé x_0^a (point rouge) qui peut être propagé dans le temps par le modèle (trait rouge). Les traits bleus représentent l'écart final aux observations (qui, via les matrices de covariance d'erreur $\boldsymbol{R_k}$, déterminent $J^{obs}(x^a)$). Le trait vert à t_0 représente l'écart final à l'ébauche (qui, via la matrices de covariance d'erreur $\boldsymbol{B}$, détermine $J^b(x^a)$).

La figure III.2 schématise le fonctionnement de l'algorithme 4D-VAR et va nous permettre

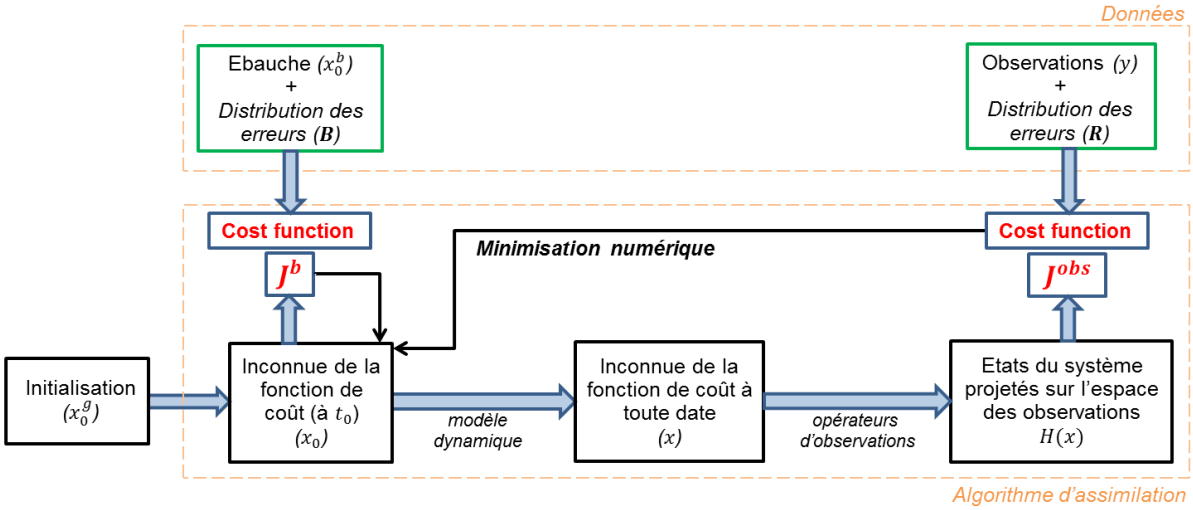

FIGURE III.2 – Schéma représentant le fonctionnement de l'algorithme d'assimilation 4D-VAR.

d'introduire la section suivante sur son implémentation effective. Les données du problème, à savoir l'ébauche x_0^b, les observations y, et leurs statistiques d'erreurs associées, $\boldsymbol{B}$ et $\boldsymbol{R}$, sont dans les cadres verts en haut du schéma. Au début du processus de minimisation (i.e. avant la 1ère itération du minimiseur numérique), on initialise l'état initial du système à la valeur x_0^g (first guess). Plaçons nous à présent à une itération quelconque du minimiseur et supposons qu'à cette itération de l'algorithme l'état initial du système vaille x_0. Cet état initial est ensuite propagé à toutes les dates par le modèle numérique (on obtient x, les différents états du système induits par x_0). À partir des états du système, les opérateurs d'observation ($\boldsymbol{H}$) permettent de déduire des grandeurs comparables aux observations (y). J^{obs}, la partie rappel aux observations de la fonction de coût, évalue l'écart entre les deux, pondéré par la confiance que l'on a en les observations ($\boldsymbol{R}^{-1}$). La partie rappel à l'ébauche de la fonction de coût (J^b) évalue elle l'écart entre l'ébauche (x_0^b) et l'état initial du système (x_0), pondéré là encore par la confiance que l'on a en l'ébauche ($\boldsymbol{B}^{-1}$). On itère ensuite le minimiseur numérique (méthode du gradient par exemple) pour actualiser l'état initial du système x_0. L'algorithme s'arrête lorsque le coût est suffisamment faible ou qu'il ne diminue plus de manière significative. Il peut aussi être arrêté lorsque l'on a effectué un nombre d'itérations donné.

3 Mise en oeuvre logicielle : YAO

Dans cette section, nous allons détailler le calcul et l'implémentation explicite d'un algorithme 4D-VAR sur un exemple très simple et concret. Nous appliquerons donc les principes vus dans la section précédente sur cet exemple et introduirons ensuite les concepts utilisés par le logiciel YAO pour faciliter l'implémentation du 4D-VAR sur des cas plus complexes. YAO a été développé au LOCEAN (Laboratoire d'Océanographie et du Climat : Expérimentations et approches numériques : https://www.locean-ipsl.upmc.fr/index.php?lang=en). Il s'agit d'un logiciel facilitant l'implémentation d'algorithmes d'assimilation variationnelle. L'article de Nardi et al. (2009) présente les concepts qu'il développe et utilise. Le rapport de Thiria et al. (2006) présente plus en détails le fonctionnement du logiciel et donne un exemple concret de mise en oeuvre. On peut aussi se référer au site internet du logiciel, qui propose notamment de télécharger ces différents papiers ainsi que des présentations : https://skyros.locean-ipsl.upmc.fr/~yao/index.html.

3.1 Un exemple simple : système proies / prédateurs

3.1.1 Présentation du problème

Nous allons détailler les étapes de mise en place d'un algorithme 4D-VAR sur un exemple très simple : le système proies/prédateurs. Considérons deux populations, l'une constituée de proies et l'autre de prédateurs. Supposons que nous cherchons à déterminer le nombre de proies et de prédateurs à une certaine date, prise comme origine des temps et notée t_0. Supposons que nous avons observé, par une méthode non précisée ici, le nombre de proies et de prédateurs à une certaine date postérieure, notée t_2. En utilisant les notations de la partie précédente, ces observations constituent le vecteur y_2. Ici, $y_2 = (y_2^L; y_2^R)$, où y_2^L et y_2^R désignent respectivement le nombre de proies et de prédateurs observés à t_2. Comme nous n'avons pas d'autres observations, on a $y = y_2$. Supposons par simplicité que les observations sont indépendantes et que nous n'avons pas de raison d'accorder plus de valeur à l'une qu'à l'autre. On a alors une matrice de variances/covariances correspondant à la variance des observations : $\boldsymbol{R_2} = \boldsymbol{Id}$.

Pour modéliser l'évolution des populations, nous allons utiliser le modèle de Lotka-Volterra (Lotka et al. (1925), Volterra (1926)). Celui-ci, en fixant arbitrairement tous ses paramètres à 1, est caractérisé par le système d'équations différentielles ordinaires suivant :

$$\begin{cases} (x^L)' = x^L(1 - x^R) \\ (x^R)' = x^R(-1 + x^L) \end{cases} \tag{III.10}$$

Où x^L et x^R désignent respectivement le nombre de proies et de prédateurs (fonctions du temps). On note comme précédemment que ce système n'est qu'une modélisation de la réalité et qu'une erreur (modèle) est déjà commise à ce niveau-là. Pour propager effectivement ses

états dans le temps, nous allons le discrétiser (i.e. résoudre l'équation différentielle de manière approchée). Une seconde erreur va être commise lors de cette discrétisation. Nous utilisons ici pour cela le schéma d'Euler (https://fr.wikipedia.org/wiki/M%C3%A9thode_d'Euler) , le plus simple / rapide à mettre en œuvre. Pour éviter d'augmenter le nombre de notations, nous allons utiliser les mêmes symboles pour les variables continues et discrétisées (c'est bien sûr abusif, l'écart entre les deux étant comme on l'a vu l'erreur de discrétisation). Sous ces conditions, on écrit, à la date t_{k+1} :

$$\begin{cases} x_{k+1}^L = x_k^L(1 + \Delta t - \Delta t x_k^R) \\ x_{k+1}^R = x_k^R(-1 + \Delta t + \Delta t x_k^L) \end{cases} \tag{III.11}$$

Avec Δt le pas de temps de la discrétisation.

En utilisant les notations de la section précédente, avec x_k le vecteur contenant les états du système à la date t_k :

$$x_{k+1} = \begin{pmatrix} x_{k+1}^L \\ x_{k+1}^R \end{pmatrix} = M_{k,k+1}(x_k) = \begin{pmatrix} x_k^L(1 + \Delta t - \Delta t x_k^R) \\ x_k^R(-1 + \Delta t + \Delta t x_k^L) \end{pmatrix} \tag{III.12}$$

Dans notre cas d'étude actuel on aura ainsi $x_2 = M_{1,2} \circ M_{0,1}(x_0) = M_{1,2}(x_1)$, avec $x_1 = M_{0,1}(x_0)$.

On rappelle qu'ici les observations disponibles portent directement sur les états du système (à la date t_2). L'opérateur d'observation H se réduit donc à $H(x) = H_2(x_2) = \boldsymbol{Id}$. On suppose également que l'on ne dispose d'aucune information sur l'ébauche, soit $\boldsymbol{B}^{-1} = 0$. Cela revient à dire dans la fonction de coût que $J^b(x) = 0$. Afin d'utiliser un algorithme de minimisation numérique, nous devons néanmoins l'initialiser à une certaine valeur, notée x_0^g (par exemple 0, ou y_2, ou une valeur quelconque).

3.1.2 Mise en œuvre

On a défini dans la partie précédente tous les éléments nécessaires à la mise en œuvre d'un algorithme d'assimilation 4D-VAR, à savoir les champs x^b, x_0^g, y ainsi que les opérateurs $M_{k,k+1}$, $\boldsymbol{B}$, $\boldsymbol{R}$ et H. On peut ainsi facilement exprimer la fonction de coût, qui s'écrit ici :

$$\begin{aligned} J(x_0) &= ||y_2 - x_2||_2^2 \\ &= \left(y_2^L - x_2^L\right)^2 + \left(y_2^R - x_2^R\right)^2 \end{aligned} \tag{III.13}$$

On suppose donc que l'on se place à une itération donnée de l'algorithme de minimisation et que celui-ci nous fournit une valeur pour x_0. On peut alors calculer explicitement $J(x_0)$, en utilisant la définition de la fonction de coût J (Eq. (III.13)) et le modèle : $x_2 = M_{1,2} \circ M_{0,1}(x_0)$

(Eq. (III.12)). Pour itérer l'algorithme de minimisation, nous allons avoir besoin du gradient de cette fonction de coût par rapport à l'état que l'on contrôle, c'est-à-dire x_0, soit :

$$\nabla_{x_0} J(x_0) = \begin{pmatrix} \frac{\partial J}{\partial x_0^L} \\ \frac{\partial J}{\partial x_0^R} \end{pmatrix} \tag{III.14}$$

Etant donnée la fonction de coût de notre modèle proies / prédateurs (Eq. (III.13)), on peut calculer explicitement et facilement $\nabla_{x_2} J$, le gradient de la fonction de coût J par rapport à x_2, qui s'écrit :

$$\nabla_{x_2} J(x_0) = \begin{pmatrix} \frac{\partial J}{\partial x_2^L} \\ \frac{\partial J}{\partial x_2^R} \end{pmatrix} = 2 \begin{pmatrix} y_2^L - x_2^L \\ y_2^R - x_2^R \end{pmatrix} \tag{III.15}$$

Pour "rétropropager" le gradient, dans un premier temps de t_2 à t_1, on va calculer l'adjoint du modèle. On sait que l'on a $x_2 = M_{1,2}(x_1)$. En utilisant le théorème des fonctions composées, on obtient l'expression suivante :

$$\nabla_{x_1} J = \boldsymbol{M_{1,2}^T(x_1)} \nabla_{x_2} J \tag{III.16}$$

Avec $\boldsymbol{M_{1,2}(x_1)}$ le linéaire tangent, c'est-à-dire la jacobienne de $M_{1,2}$, soit :

$$\boldsymbol{M_{1,2}(x_1)} = \begin{pmatrix} \frac{\partial x_2^L}{\partial x_1^L} & \frac{\partial x_2^L}{\partial x_1^R} \\ \\ \frac{\partial x_2^R}{\partial x_1^L} & \frac{\partial x_2^R}{\partial x_1^R} \end{pmatrix} \tag{III.17}$$

Si l'on se réfère à l'Eq. (III.12), on a vu que x_2 est défini analytiquement à partir de x_1. On peut donc calculer explicitement $M_{1,2}(x_1)$, qui s'écrit ici :

$$\boldsymbol{M_{1,2}(x_1)} = \begin{pmatrix} 1 + \Delta t - \Delta t x_1^R & -\Delta t x_1^L \\ \\ \Delta t x_1^R & -1 + \Delta t + \Delta t x_1^L \end{pmatrix} \tag{III.18}$$

Connaissant $\nabla_{x_2} J$ et $\boldsymbol{M_{1,2}(x_1)}$, on peut en déduire $\nabla_{x_1} J$ par l'Eq.(III.16). On procède de même pour calculer $\nabla_{x_0} J$ à partir de $\nabla_{x_1} J$ (rétropropagation du gradient de t_1 à t_0) :

$$\nabla_{x_0} J = \boldsymbol{M_{0,1}^T(x_0)} \nabla_{x_1} J \tag{III.19}$$

Avec :

$$\boldsymbol{M_{0,1}(x_0)} = \begin{pmatrix} 1 + \Delta t - \Delta t x_0^R & -\Delta t x_0^L \\ \\ \Delta t x_0^R & -1 + \Delta t + \Delta t x_0^L \end{pmatrix} \tag{III.20}$$

En combinant les équations (III.15), (III.16), (III.18), (III.19) et (III.20), on peut finalement en déduire $\nabla_{x_0} J(x_0)$, qui s'écrit :

$$\nabla_{x_0} J(x_0) = \boldsymbol{M_{0,1}^T(x_0) M_{1,2}^T(x_1) \nabla_{x_2} J} \tag{III.21}$$

Et donc explicitement :

$$\nabla_{x_0} J(x_0) = 2 \begin{pmatrix} 1+\Delta t - \Delta t x_0^R & \Delta t x_0^R \\ -\Delta t x_0^L & -1+\Delta t + \Delta t x_0^L \end{pmatrix} \begin{pmatrix} 1+\Delta t - \Delta t x_1^R & \Delta t x_1^R \\ -\Delta t x_1^L & -1+\Delta t + \Delta t x_1^L \end{pmatrix} \begin{pmatrix} y_2^L - x_2^L \\ y_2^R - x_2^R \end{pmatrix}$$

Connaissant x_0 et $\nabla_{x_0} J$, on peut finalement utiliser une méthode de minimisation numérique pour actualiser la valeur de x_0 et passer à l'itération suivante.

On pourrait par exemple faire appel au naïf algorithme du gradient à pas constant (https://fr.wikipedia.org/wiki/Algorithme_du_gradient). En notant α le pas de descente et x_0^{actu} le x_0 actualisé, on aura : $x_0^{actu} = x_0 - \alpha \nabla_{x_0} J(x_0)$. En imposant un critère d'arrêt (typiquement sur le coût ou le nombre d'itérations), notre algorithme termine.

3.2 Généralisation : le logiciel YAO

Nous allons à présent expliquer comment la partie précédente pourrait être prise en charge par le logiciel YAO. Pour plus d'informations sur le formalisme de YAO et le code en lui-même, on se raportera à Thiria et al. (2006). On expliquera ici la philosophie ainsi que les aspects spécifiques à l'implémentation d'un modèle dans YAO. Cette section a pour vocation d'introduire l'un des outils sur lequel repose cette thèse.

3.2.1 Mise en œuvre intuitive des concepts sur l'exemple précédent

La Figure III.3 présente les étapes successives de l'algorithme 4D-VAR lors d'une itération du minimiseur. Cela va nous permettre d'introduire de manière intuitive le concept de graphe modulaire utilisé par YAO. La partie gauche de la figure ("forward") représente la propagation de l'état initital (x_0), via le modèle, jusqu'à la date t_2, date à laquelle des observations sont disponibles. On calcule ensuite (analytiquement) la valeur de la fonction de coût et de son gradient par rapport à x_2. On rétropropage ensuite ce gradient jusqu'à t_0 ("backward") en utilisant les jacobiennes du modèle (définies elles aussi analytiquement pour chaque date). On actualise ensuite l'état initial grâce à ce gradient, en utilisant un algorithme de minimisation quelconque.

La remarque la plus importante pour nous permettre l'automatisation de ce processus est que le modèle qui propage les états de t_0 à t_1 ($M_{0,1}$) est le même que celui qui propage les états de t_1 à t_2 ($M_{1,2}$). Si l'on définit une "trajectoire" (c'est-à-dire une date initiale, un pas de temps, et une date finale), on comprend que la donnée du modèle direct ($M_{k,k+1}$) pour

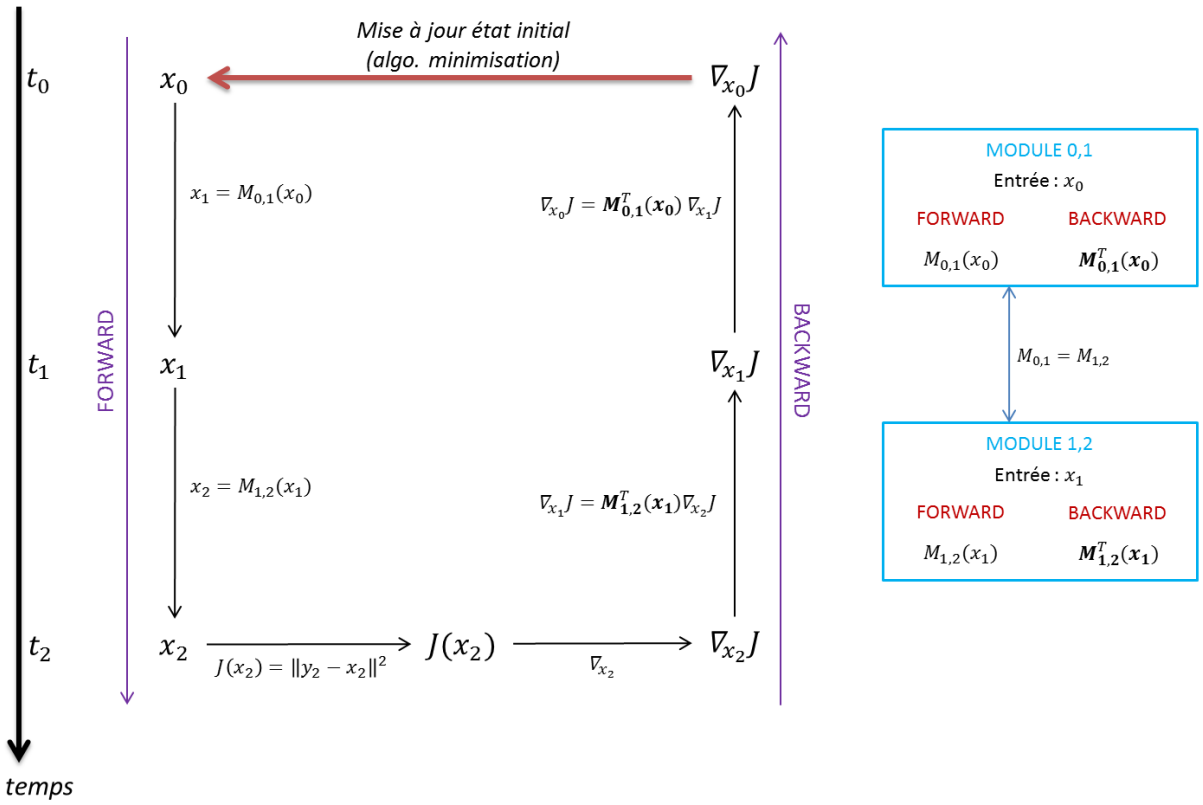

FIGURE III.3 – Détail des étapes d'une itération de l'algorithme 4D-VAR appliqué au problème proies/prédateurs.

n'importe quelle date (et de sa jacobienne ($\boldsymbol{M_{k,k+1}(x_k)}$) permettra à un programme de réaliser les étapes "forward" (passage de x_0 à x_k) et "backward" (passage de $\nabla_{x_k} J$ à $\nabla_{x_0} J$). On fournit par ailleurs à ce programme les expressions du coût J et de son gradient $\nabla_{x_k} J$. À partir d'un point de départ (x_0^g) et par l'intermédiaire d'un algorithme de minimisation, il pourra réaliser l'intégralité du processus.

YAO est un logiciel qui permet de réaliser relativement aisément l'écriture d'un tel code, y compris bien sûr pour des problèmes plus complexes.

3.2.2 Cas général

L'exemple précédent était simpliste sur plusieurs points : tout d'abord tous les états étaient des scalaires (il se basait d'ailleurs sur un modèle continu d'équations différentielles ordinaires et non d'équations aux dérivées partielles). On était donc sur un espace "0D" : à chaque date, tout se passait en un seul point de maille. Un modèle plus réaliste est généralement constitué d'un espace à plusieurs dimensions, typiquement une discrétisation de l'espace, avec donc un nombre important de points où évaluer les états du modèle à chaque date. Ensuite, il n'y avait qu'un espace. Notamment, les observations étaient définies sur le même espace que le modèle, il n'y avait pas de transformation à effectuer (calcul analytique, intégration spatiale / temporelle, projection, ...) pour transformer les sorties du modèle en données comparables aux

observations. Enfin, pour un point de maille donné, il n'y avait qu'un état à évaluer. On peut imaginer plusieurs variables (en météorologie : la température, la pression, l'humidité, le vent, ...) à propager.

Ces complexifications ne modifient pas fondamentalement le fonctionnement de l'algorithme d'assimilation. Sur l'exemple précédent (figure III.3), dans le modèle forward, la sortie du modèle à une date devient son entrée à la date suivante. S'il y a plusieurs espaces (par exemple un pour le modèle et un pour les observations), on va procéder entre ces espaces, à une date donnée, comme on procédait d'une date à l'autre. Les sorties du modèle (définies sur l'espace du modèle) deviennent ainsi les entrées de l'opérateur d'observation (défini sur l'espace des observations). Au sein d'un espace (par exemple celui du modèle), si l'on doit calculer des états en tous points de maille, on procédera là encore séquentiellement. On définira un module "générique", représentant le modèle pour un point de maille quelconque, prenant en entrée les états à la date précédente (sorties du modèle à la date précédente) et en un ou plusieurs points de maille. L'espace sera ensuite parcouru automatiquement, point de maille par point de maille, et notre module générique sera appliqué à chaque fois. Ce module pourra lui-même être subdivisé en plusieurs sous-modules (un pour la température, un pour la pression, ...), chacun prenant éventuellement en entrée les sorties d'un autre. On obtient ainsi intuitivement le concept de graphe modulaire formalisé dans YAO. En effet YAO est réalisé de telle façon qu'il permet de découper un modèle (et ainsi son adjoint) en briques élémentaires (les modules), chaque brique étant répétée en chaque point de maille et à chaque date dans un ordre défini par leur connectivité (induite par le graphe). On se reportera à Nardi et al. (2009) pour une définition plus formelle des graphes modulaires.

La figure III.4 représente schématiquement ce concept. En haut (Figure III.4a) est représenté un point de maille, sur lequel sont définis cinq modules et leur connectivité au sein de ce point de maille. Par exemple, le module F_1 prend en entrée une information venant de l'extérieur de ce point de maille (non encore définie), tandis que F_3 prend en entrée les sorties de F_1 et F_2 au même point de maille. Vues les connectivités, on imagine que pour chaque point de maille, on calculera, dans le forward, d'abord F_1 et F_2, puis F_3 et F_4, pour terminer par F_5.

En bas (Figure III.4b) est représentée l'intégralité d'un espace, avec la connectivité entre les 4 points de maille qui le composent. On voit par exemple que F_1, pour les points de maille (1.1) et (2.1), reçoit en entrée une information extérieure à la maille (par exemple une sortie à une date précédente), tandis qu'il reçoit en entrée sur les points de maille (1.2) et (2.2) la sortie de F_5 sur le point de maille immédiatement à sa gauche. Dans ces conditions, le calcul du forward sur la maille sera effectué en traitant les points de maille dans l'ordre suivant : (1.1) $\rightarrow$ (2.1) $\rightarrow$ (1.2) $\rightarrow$ (2.2).

YAO va nous permettre d'automatiser ces traitements. Pour cela, il requiert un certain nombre d'informations. Il faudra donc lui spécifier les éléments suivants :

1. une trajectoire (discrétisation du temps).
2. un ou des espaces (discrétisation de l'espace). Il peut y avoir comme on l'a vu plusieurs

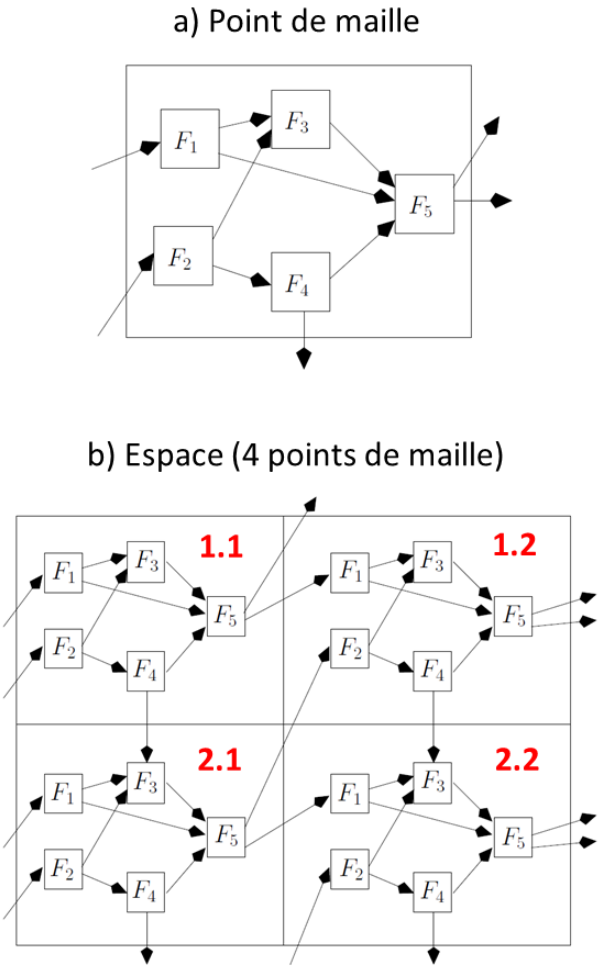

FIGURE III.4 – Illustration rapide du concept de graphe modulaire. Figure adaptée de Nardi et al. (2009), fig.3.

espaces, typiquement un pour le modèle et un pour les observations, chacun avec sa dimension et sa résolution spécifique. En météorologie, on peut par exemple imaginer un espace 3D avec un pas $dx * dy * dz$ pour le modèle et des observations qui sont des lames d'eau au sol, donc sur un espace 2D avec un pas $dx' * dy'$.

3. un ordre de calcul des espaces au sein de la trajectoire. Sur l'exemple précédent par exemple, à une date donnée, on calculera d'abord les variables météorologiques sur l'espace 3D du modèle avant de simuler les lames d'eau (via l'opérateur d'observation) sur l'espace 2D, pour les comparer aux observations.
4. un ou plusieurs modules (briques élémentaires pour le modèle comme pour les opérateurs d'observation), chacun rattaché à un espace. Ces modules seront répétés automatiquement en tout point de maille de l'espace. On aura, par exemple, dans un espace 3D, un premier module calculant l'humidité, un deuxième calculant la température, un troisième estimant les précipitations, etc. D'autres modules pourront eux résider dans un espace $2D$, par exemple, un opérateur d'observation transformant ces précipitations 3D en lames d'eau

au sol.

Ces modules sont caractérisés par plusieurs informations :

- Leur connectivité avec les autres modules (du même espace ou d'autres espaces), c'est-à-dire les entrées du module, ainsi que les paramètres dont il a besoin pour fonctionner. Par exemple, il serait cohérent qu'un module "précipitations" reçoive des informations venant de modules "température" et "humidité" au même point de maille, et de lui-même (module "précipitations") à une date précédente.

- leur code forward, c'est-à-dire l'expression de sa ou de ses sorties en fonction des entrées (entrées définies par la connectivité).

- leur code backward, c'est-à-dire l'expression de sa jacobienne (les dérivées des différentes sorties par rapport aux différentes entrées).

5. l'ordre de calcul des modules pour chaque point de maille de l'espace auquel ils sont rattachés. L'ordre de parcours des points de maille est automatique mais il est parfois nécessaire de faire plusieurs "passages" sur un espace, par exemple si l'on a besoin de connaître une variable en tout point de maille si l'on veut calculer une autre variable à la même date.
6. des observations (par exemple dans un fichier *.txt), avec leurs dates et leurs modules de rattachement (c'est-à-dire à quel module elles doivent être comparées, typiquement le module appliquant l'opérateur d'observation).
7. un module à contrôler (l'inconnue du problème). Par défaut, et comme cela est le cas dans cette section, il s'agit de l'état initial (à t_0) du modèle.
8. un état d'initialisation (x_0^g) pour le module à contrôler (par exemple dans un fichier *.txt).
9. tous les paramètres de la discrétisation (pas de temps, d'espace, nombre de pas, ...)

Avec ces informations, YAO va pouvoir exécuter l'algorithme. Il va d'abord propager l'état initial partout (modèle direct, forward), en procédant date par date, espace par espace, point de maille par point de maille, et module par module, en utilisant leur code forward (chaque module étant utilisé, sur son espace de rattachement, à chaque point de maille et à chaque date). YAO utilise ensuite les observations et le module "opérateur d'observation" et sa jacobienne pour évaluer la fonction de coût et son gradient par rapport aux sorties. Il procède ensuite "à rebours" pour évaluer le gradient par rapport aux entrées. Il utilise le code backward (associé aux jacobiennes) des modules et parcourt les modules, les espaces et la trajectoire dans l'ordre inverse de celui du forward défini plus haut.

On note que les matrices de covariance d'erreur ne sont pas définies explicitement (en tant que matrices). Cela serait mauvais en terme de performance et ce inutilement de par le fait qu'elles sont généralement très creuses. Elles sont donc traitées sous forme de module (typiquement dans le module opérateur d'observation) rattaché à un espace et une trajectoire.

3.2.3 Minimisation numérique sous YAO

On a vu que l'assimilation 4D-VAR était avant tout un problème d'optimisation numérique et qu'une fois gérées les difficultés de propagation des états et de rétropropagation du gradient, il fallait utiliser un minimiseur numérique (pour actualiser l'inconnue, le paramètre de contrôle). Le logiciel est prévu pour être couplé facilement avec deux algorithmes de minimisation développés à l'INRIA (Institut National de Recherche en Informatique et en Automatique : http://www.inria.fr/).

Le premier est l'algorithme M1QN3 (Gilbert and Lemaréchal (2006)). Il s'agit d'une méthode de minimisation numérique sans contraintes à direction de descente de type quasi-Newton. Contrairement à l'algorithme de Newton-Rhapson, qui nécessite le calcul de la jacobienne pour trouver le zéro d'une fonction, et donc le calcul de la hessienne (dérivée d'ordre 2), pour la minimisation d'une fonction, les algorithmes quasi-Newton utilisent des suites itératives pour approximer la hessienne à partir du gradient et réduire ainsi le coût de calcul. Cette méthode a été développée pour être utilisée pour des problèmes de grande dimension. Elle est donc pertinente pour les problèmes généralement traités par l'assimilation 4D-VAR.

Le second est l'algorithme M2QN1 (Lemaréchal and Panier (2000)). Il s'agit cette fois d'une méthode de minimisation numérique avec contraintes de bornes, c'est-à-dire qui permet de tenir la solution entre des bornes prédéfinies. C'est également un algorithme à direction de descente de type quasi-Newton.

On note que rien n'empêche l'utilisateur d'utiliser un autre algorithme de minimisation. Dans tout le travail présenté ici, nous utiliserons néanmoins l'algorithme M1QN3.

4 Extension

Jusqu'à présent, nous avons toujours supposé les paramètres du modèle connus et les conditions aux limites du système d'équations différentielles discrétisées inutiles ou parfaitement connues. L'inconnue de l'algorithme 4D-VAR (le paramètre de contrôle) était alors l'état initial du modèle numérique (x_0). La définition du problème dans ces termes n'est pas imposée par la théorie et nous pouvons imaginer élargir le cadre du 4D-VAR au contrôle des conditions limites, de paramètres du modèle, des conditions initiales ou de plusieurs de ces champs. Cela ne change en général pas fondamentalement le fonctionnement de l'algorithme mais modifie (et complexifie) quelque peu le formalisme employé. Dans la suite, on revient sur les implications qu'ont le contrôle des conditions limites. On abordera aussi l'abandon du terme de rappel à l'ébauche ou l'extension de sa définition. Le contrôle de paramètres du modèle, qui ne présente pas de difficultés majeures, est étudié sur l'exemple précédent (système proies / prédateurs) dans l'Annexe E.

4.1 Contrôle de conditions initiales et aux limites

Nous considérons dans cette section le cas où les paramètres du modèle sont supposés connus. On cherche à restituer un champ discrétisé possédant une dimension temporelle et une ou plusieurs dimensions spatiales. Pour cela, on devra contrôler l'état initial (comme précédemment), mais aussi les conditions aux limites du problème. Nous aurons, au moins dans certaines équations, besoin de faire apparaître explicitement les points de maille de la ou des dimensions spatiales impactées par ce changement. Afin de ne pas alourdir encore les notations, on se limite ici à une seule dimension d'espace. On peut imaginer par exemple que l'on est en train de modéliser la chute verticale de gouttes de pluie. (Ce sera le cas développé dans le chapitre V de ce travail). La dimension spatiale retenue est alors l'altitude. On discrétise donc l'axe vertical. On note $z_0 = 0$ l'altitude maximale (choisie comme origine pour cet axe) et dz le pas de discrétisation. La coordonnée du point de maille située à la distance $i.dz$ du sommet de la colonne est ainsi notée $z_i = i.dz$. La Figure III.5 représente la discrétisation du plan temps/altitude. On note $x_{k,i}$ la valeur de la variable d'état (x) discrétisée correspondant à la date t_k et à l'altitude z_i. Le vecteur x_k contient $NZ + 1$ éléments représentant les états du système (aux différentes altitudes, on considère la chute des gouttes de pluie) à la date t_k. On suppose que le modèle propage les états, dans le temps, du haut de la colonne (z_0) vers le bas (z_{NZ}), comme indiqué par la flèche sur la Figure III.5. Sous ces conditions, les paramètres de contrôle (qui coïncident alors avec les états nécessaires à l'exécution du modèle) sont x_0 (vecteur des états initiaux, en jaune sur la Figure III.5) et $(x_{k,0})_{k\in[|0:NT|]}$ (vecteur des conditions aux limites en haut). Dans la suite, ce dernier vecteur sera noté $x_L = (x_{k,0})_{k\in[|0:NT|]}$.

On reprend ici les notations précédentes. M_k correspond donc au modèle propageant les

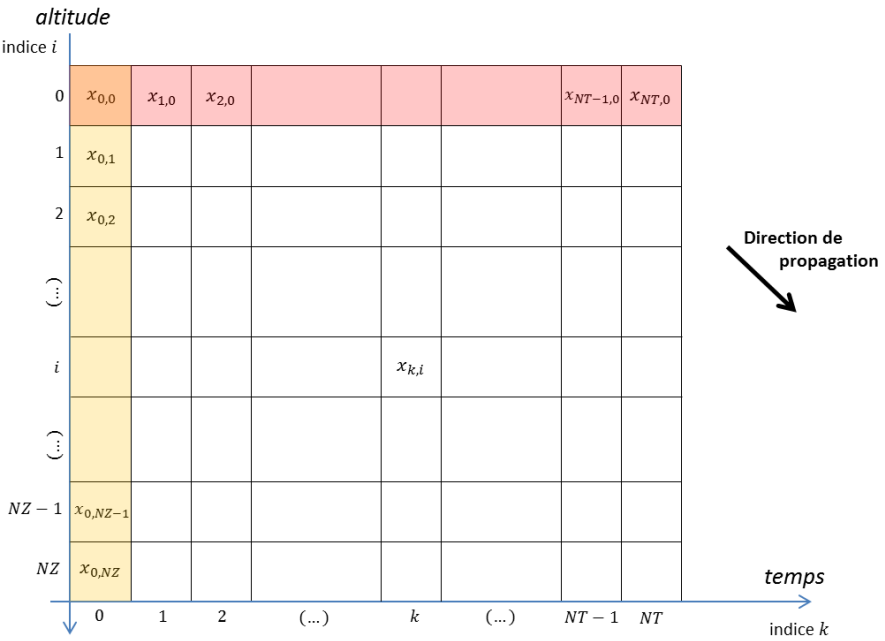

FIGURE III.5 – Schéma représentant la discrétisation du plan temps/altitude sur lequel sont définis des états du système dynamique. La direction de propagation de ce système est indiquée (propagation temporelle et du haut vers le bas), ainsi que les paramètres de contrôle de l'algorithme 4D-VAR associé, à savoir les conditions initiales (en jaune) et les conditions limites (en rouge).

états jusqu'à la date t_k. On a alors :

$$x_k = M_k\left(x_0, (x_{p,0})_{p\leq k}\right) \tag{III.22}$$

Il est important de remarquer que maintenant, le modèle, pour propager les états jusqu'à la date t_k, a besoin de connaître l'état initial x_0 mais aussi les conditions aux limites en haut jusqu'à la date t_k $(x_{p,0})_{p\leq k}$.

L'écriture formelle de la fonction de coût du 4D-VAR dans ce cadre ne pose alors pas de soucis. L'équation générale (Eq.(III.5)) peut parfois être conservée telle qu'elle. On rappelle que x_0^b désigne l'ébauche sur la condition initiale. On pourrait également faire ressortir un terme d'ébauche correspondant aux conditions aux limites, noté x_L^b. La partie rappel à l'ébauche de la fonction de coût, constituée des rappels aux conditions initiales (x_0^b) et aux conditions aux limites (x_L^b) s'écrirait alors :

$$J^b(x_0, x_L) = \left(x_0 - x_0^b\right)\boldsymbol{B_0}^{-1}\left(x_0 - x_0^b\right)^T + \left(x_L - x_L^b\right)\boldsymbol{B_L}^{-1}\left(x_L - x_L^b\right)^T \tag{III.23}$$

Avec $\boldsymbol{B_0}$ et $\boldsymbol{B_L}$ les matrices de covariance d'erreur d'ébauche sur la condition initiale et sur la condition aux limites.

Sur le même principe, on ajoute la dépendance aux conditions aux limites dans l'Eq.(III.6)

(rappel aux observations). Elle s'écrit maintenant :

$$J^{obs}(x_0, x_L) = \sum_{k=0}^{NT} \left(y_k - H_k \circ M_k \left(x_0, (x_{p,0})_{p\leq k}\right)\right) \boldsymbol{R_k}^{-1} (y_k - H_k \circ M_k \left(x_0, (x_{p,0})_{p\leq k}\right))^T \quad \text{(III.24)}$$

Ou pour alléger la notation :

$$J^{obs}(x_0, x_L) = \sum_{k=0}^{NT} \left(y_k - H_k \circ M_k \left(x_0, x_L\right)\right) \boldsymbol{R_k}^{-1} (y_k - H_k \circ M_k \left(x_0, x_L\right))^T \quad \text{(III.25)}$$

Il en découle l'expression du gradient suivante :

$$\nabla J^{obs}(x_0, x_L) = -2 \sum_{k=0}^{NT} \boldsymbol{M_k}(\boldsymbol{x})^T \boldsymbol{H_k}(\boldsymbol{x})^T \boldsymbol{R_k}^{-1} \left(y_k - H_k \circ M_k(x_0, x_L)\right). \quad \text{(III.26)}$$

On remarque que cette expression est une version modifiée du gradient présenté dans l'Eq.(III.9). Là où on ne contrôlait que les conditions initiales, on controle maintenant à la fois les conditions initiales x_0 et les conditions aux limites x_L. La fonction de coût permet de minimiser une distance entre les observations et les états du modèle projetés dans l'espace des observations. Ici, cette distance est mesurée en propageant les conditions initiales mais aussi les conditions aux limites. Ces conditions aux limites proviennent d'autres dates que t_0. Puisqu'on les contrôle, on ne cherche plus le gradient de la fonction de coût uniquement par rapport à x_0 $(\partial J/\partial x_{0,i})_{i\in[|0:NZ|]}$, mais aussi par rapport aux conditions aux limites $(\partial J/\partial x_{k,0})_{k\in[|0:NT|]}$. Dans la mesure où ces gradients étaient déjà évalués auparavant dans le parcours à rebours de l'algorithme backward, cela ne change cependant que peu la réalisation effective de l'algorithme.

De manière plus générale, on note que tous les $(\partial J/\partial x_{k,i})_{k\in[|0:NT|],i\in[|0:NZ|]}$ sont évalués dans l'algorithme backward. Les seuls à servir explicitement dans la minimisation sont les dérivées de la fonction de coût par rapport aux paramètres de contrôle. Tout état peut donc facilement être contrôlé. Il n'est cependant "raisonnable", en pratique, que de contrôler les inconnues du modèle (conditions initiales x_0 et aux limites x_L), les autres états en découlant explicitement.

On note enfin qu'il est parfois possible de se débarrasser du contrôle de l'état initial d'un système. Typiquement, dans le cas de processus intermittents (i.e. avec des plages de valeurs nulles), on peut considérer un état initial connu et nul en se plaçant avant le début de l'occurrence du phénomène modélisé. On peut ainsi forcer $x_0 = 0$. Comme on le verra au chapitre V, si l'on modélise la chute de gouttes de pluie, on peut constater qu'il ne pleut pas avant une certaine date (aucune goutte quelque soit l'altitude). En fixant l'origine des dates (t_0) suffisamment longtemps avant cette date, il est raisonnable de supposer des conditions initiales nulles $(x_0 = 0)$. Il ne reste plus alors qu'à déterminer les conditions aux limites.

4.2 Cas sans ébauche / Filtrage

Rien ne nous oblige dans le formalisme du 4D-VAR à avoir un terme de rappel à l'ébauche. Si l'on dispose de suffisamment d'observations, on peut tout simplement se passer de l'ébauche et du coût associé. Cela implique que le modèle soit assez stable et que les observations permettent de lever les ambiguïtés. Il arrive cependant - ce sera le cas dans les deux applications développées dans cette thèse - que l'on n'ait pas suffisamment d'observations pour couvrir tous les états. Pour autant, on peut ne pas disposer d'assez d'informations pour pouvoir déterminer un champ d'ébauche. Rien ne nous empêche alors d'ajouter dans la fonction de coût d'autres termes : de filtrage, de pénalisation, de bilan de masse, etc. Nous n'avons pas vocation ici à être exhaustif mais donnons simplement quelques exemples et contraintes sur ces termes.

4.2.1 Pénalisation

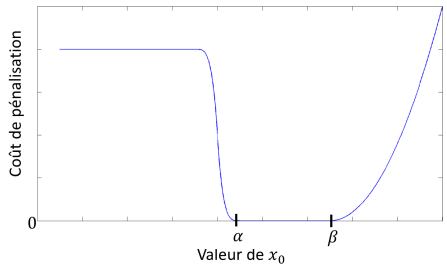

FIGURE III.6 – Exemple arbitraire d'une fonction de pénalisation d'un paramètre x_0. Ici, on veut contraindre x_0 dans l'intervalle $[\alpha : \beta]$, et le coût associé est $\mathcal{C}^1$.

Supposons que l'on contrôle l'état initial x_0 et que celui-ci soit un scalaire. Sans avoir a priori d'information sur sa valeur, on peut vouloir le contraindre entre deux bornes (dans un intervalle de valeurs "raisonnables"). Dans ce cas-là, on peut remplacer le terme de rappel à l'ébauche de la fonction de coût par un terme de pénalisation comme celui tracé sur la figure III.6. Sur cet exemple, x_0 est contraint entre les valeurs α et β.

On note que pour la minimisation de la fonction de coût, il va falloir calculer la dérivée de ce terme de pénalisation par rapport à x_0. On a donc tout intérêt à choisir une expression dérivable, et à la dérivée régulière (au moins $\mathcal{C}^1$ donc). Par exemple, la fonction tracée sur la figure III.6 est $\mathcal{C}^1$ et polynomiale par morceaux (donc $\mathcal{C}^\infty$ sur les morceaux).

4.2.2 Filtrage / Lissage

Il arrive que pour des raisons de stabilité numérique, on choisisse pour le modèle un pas de temps plus petit que la résolution la plus fine des observations disponibles. On peut par exemple imaginer un cas, toujours sans dimension temporelle, où les observations seraient intégrées sur $4dt$, avec dt le pas de temps du modèle. Une solution avec des oscillations d'ampleur quelconque sur chacune des fenêtres d'observation successives aurait alors le même coût qu'une solution constante sur les mêmes fenêtres. La Figure III.7 présente rapidement ce

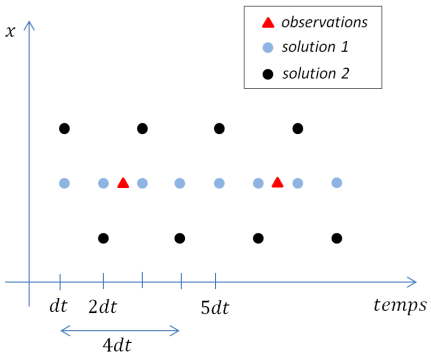

FIGURE III.7 – Exemple de problème justifiant l'utilisation d'un filtre. Voir texte pour les commentaires.

phénomène. On représente des observations intégrées à $4dt$ (triangles rouges), et deux solutions à résolution dt (ronds bleus et noirs), une constante, une oscillante. Dégradées à résolution $4dt$, ces deux solutions donnent les mêmes valeurs que les observations.

Si l'on veut, comme c'est le cas généralement, privilégier les solutions les plus lisses en l'absence d'observations suffisamment bien échantillonnées (c'est-à-dire sur la Figure III.7 privilégier la solution "bleue" face à la solution "noire"), on a intérêt à forcer l'algorithme en ajoutant un coût du type :

$$\sum_k \left[x_k - \frac{x_{k-1} + x_k + x_{k+1}}{3} \right]^2 \tag{III.27}$$

Ce genre de filtre pénalisant les hautes fréquence permet de tenir chaque point près de sa moyenne locale. On note par ailleurs, dans ce cas précis, qu'il s'agit d'un coût quadratique et qu'il peut s'écrire sous la forme d'un terme de rappel à l'ébauche $x^T \boldsymbol{B} x$, avec $\boldsymbol{B}$ une matrice. L'Annexe F fournit une preuve de ce résultat et l'expression de la matrice $\boldsymbol{B}$ correspondante pour un filtre à deux dimensions (dans le cadre de l'application présentée au chapitre IV).

4.2.3 Généralisation : ajout d'autres informations dans le coût

Plus généralement, on peut formellement ajouter dans la fonction de coût n'importe quelle information a priori sur les contraintes que doivent respecter les paramètres de contrôle, dès lors que ces contraintes s'expriment analytiquement à partir de ces paramètres de contrôle (ce qui permet le calcul des gradients). Il pourrait également s'agir de relations de conservation de masse ($\forall k, \sum_k x_{k,i} = cstte$, si elle n'est pas assurée par le modèle), de longueurs/temps de corrélation (sans avoir d'informations sur les valeurs absolues d'un champ, on peut savoir qu'il possède certaines propriétés statistiques), etc.

Tous ces termes s'apparentent à des termes de rappel à l'ébauche (dans la mesure où il s'agit d'ajouter des informations a priori sur les paramètres de contrôle), mais ils ne suivent pas forcément le formalisme développé plus haut, en particulier en ne s'appliquant pas seulement à l'état initial / aux conditions aux limites, et en n'étant pas forcément linéaires.

Partie IV

Restitution de cartes de précipitations par assimilation d'atténuations TV-SAT

1 Présentation : État de l'art et objectifs

1.1 Objectifs généraux

L'objectif du travail présenté dans ce chapitre est de construire des cartes de précipitations (taux de pluie) à fine échelle (typiquement $500 * 500m^2$ et $1min$) à partir de mesures d'atténuation d'ondes électromagnétiques opportunistes en bande Ku issues de satellites de télévision.

La production de telles cartes se fait généralement à partir de mesures réalisées par pluviomètres, par radar, ou par combinaison de ces deux types d'instruments. Or ces instruments présentent un certain nombre de défauts. On a vu (chapitre I) que les pluviomètres ne fournissaient que des données ponctuelles. Saisir la variabilité des précipitations nécessite donc de déployer un réseau dense, augmentant significativement les coûts et les difficultés d'entretien. Les radars sont quant à eux bloqués par les reliefs, et coûtent cher. Les zones montagneuses et les pays en développement ne sont ainsi pas toujours couverts par des réseaux opérationnels. Leur utilisation en hydrologie pose également encore beaucoup de questions (Berne and Krajewski (2013)).

La technique présentée dans ce chapitre, en utilisant un système de mesure et une technique de reconstruction de champs de pluie originaux, est une version prototype d'une méthode qui pourrait permettre d'avancer sur ces questions. Le travail présenté ici, basé sur le système de mesure détaillé par Barthès and Mallet (2013), a été publié (Mercier et al. (2015a), fourni en Annexe L) et ensuite étendu à une application à l'hydrologie urbaine (sur données simulées : Akrour (2015), Akrour et al. (2015)).

1.2 Autres études

Nous allons dans un premier temps présenter rapidement les différentes études réalisées dans un objectif similaire à celui de ce travail, à savoir reconstruire des cartes de pluie à partir d'affaiblissements d'ondes électromagnétiques opportunistes.

Giuli et al. (1991) et Giuli et al. (1999) ont proposé une méthode restituant par tomographie des cartes de pluie à partir de données simulées d'atténuation d'ondes à $30GHz$. Ils utilisent 12 liaisons à la géométrie choisie et couvrant une zone de $400km^2$ (voir Giuli et al. (1991), Fig.1). Zinevich et al. (2008) utilise également une technique de tomographie à partir de données simulées, mais avec des liaisons réalistes correspondant à un réseau de communication cellulaire couvrant une zone de $3200km^2$ en Israël. Il utilise un maillage plus fin et obtient de meilleurs résultats en zone urbaine, où les liaisons sont plus denses et plus courtes. D'autres études appliquent également des techniques de tomographie pour la restitution de cartes sur des zones urbaines à partir de données simulées (Cuccoli et al. (2009) à Florence, Zohidov et al. (2014) à Nantes). Zinevich et al. (2009) travaille lui sur des données réelles enregistrées autour de Tel-Aviv et avec une technique de reconstruction basée sur un algorithme d'assimilation de données par filtre de Kalman étendu (voir chapitre III) avec un modèle d'advection pure pour propager

les champs de pluie (voir chapitre II). Il dispose de données de 23 liaisons de téléphonie mobile (à 18, 22 et $23GHz$) et restitue des cartes de pluie à $1.5 * 1.5km^2$ et $1min$ sur $800km^2$. Plus récemment, Overeem et al. (2013) et Overeem et al. (2015) ont reconstruit par krigeage des cartes de pluie couvrant tous les Pays-Bas ($\approx 35000km^2$) à partir des données de 2400 liaisons télécom, la plupart à des fréquences entre 37 et $40GHz$, et d'une longueur moyenne de $3.1km$. Ils produisent des cartes à $15min$ et $0.9km^2$.

Une des difficultés à surmonter pour toutes ces études travaillant avec des données télécom est précisément l'obtention des données auprès des opérateurs. Zinevich et al. (2009) travaille ainsi avec un réseau relativement peu dense. Overeem et al. (2013) a pour la première fois eu accès à des données couvrant un pays entier, mais ne dispose pour chaque liaison que du maximum et du minimum d'énergie transmise toutes les 15 minutes. Chwala et al. (2012), Chwala et al. (2015) disposent eux de données à $1min$ dans le sud de l'Allemagne, tandis que cette technique a récemment été testée pour la première fois dans un pays non couvert par un réseau opérationnel de radars, le Burkina Faso (Doumounia et al. (2014)), avec une liaison de $29km$ à $7GHz$. D'autres études proposent de reconstruire des cartes de pluie en fusionnant ce type de données d'atténuation avec des pluviomètres et des radars (Bianchi et al. (2013) autour de Zurich par exemple).

1.3 Spécificités de ce travail

Ce travail diffère de toutes ces études car il utilise non pas des données télécom (sol-sol) mais des microondes issues de satellites de télévision (espace-sol). Cette caractéristique oriente toutes les liaisons disponibles vers le sud dans l'hémisphère nord (il s'agit de satellites géostationnaires, positionnés au dessus de l'équateur) et rend la longueur et l'orientation des liaisons fortement dépendantes de la latitude (en France, on obtient des liaisons projetées au sol de $3 - 7km$ et orientées à $\pm 40^\circ$ par rapport au sud).

Cependant, cela permet d'éviter les contraintes posées par la géométrie, la longueur et la fréquence des liaisons existantes. Par exemple les liaisons télécom ont tendance à être plus basses en fréquence en Afrique pour diminuer l'atténuation (et "affronter" ainsi de plus fortes pluies) et pour pouvoir être plus longues. La diminution de la fréquence comme l'allongement des liaisons rend d'autant plus imprécises et plus moyennées les restitutions de pluie.

Nous utiliserons par ailleurs pour reconstruire les cartes de pluie en elles-même un algorithme d'assimilation de données basé sur un modèle direct d'advection pure (comme Zinevich et al. (2009)), mais utilisant la théorie du 4D-VAR et non du filtre de Kalman étendu (voir chapitre III).

1.4 Plan

Dans la suite de ce chapitre, nous présenterons tout d'abord (section 2) le dispositif expérimental (données d'entrée comme données de validation). Puis nous donnerons les équations formelles de l'algorithme d'assimilation 4D-VAR utilisé (section 3), en rappelant le

modèle direct de propagation des champs de pluie, et en détaillant les opérateurs d'observations et de filtrage développés. Nous donnerons ensuite rapidement (section 4) la méthode employée pour déterminer la vitesse d'advection nécessaire à l'utilisation du modèle direct de propagation. La section 5 vérifiera sur données simulées la fonctionnalité de l'algorithme décrit. Nous détaillerons dans la section 6 les résultats sur les cas d'études issus de la campagne HyMeX (Hydrological Cycle in Mediterranean Experiment). La section 7 concluera et présentera les perspectives de ce travail.

2 Dispositif expérimental

2.1 Capteur Ku

Le capteur Ku utilisé dans cette étude est un instrument muni de 4 têtes de réception pouvant capter simultanément le signal provenant de 4 satellites géostationnaires en bande Ku différents. Il a été installé durant l'année 2013 dans le village de Mirabel, en Ardèche (voir Figure IV.1). Mirabel est situé dans la région Cévennes-Vivarais. Il s'agit d'une région montagneuse sujette, notamment à l'automne, à des épisodes de pluie intense (épisodes cévenols, ou plus généralement Méditerranéens), avec une forte variabilité spatiale et temporelle, amplifiée par la topographie. La géographie de cette région est décrite notamment dans Zwiebel (2015). Ses caractéristiques météorologiques, et notamment les épisodes cévenols, ont été abondemment étudiés (Molinié et al. (2012), Zwiebel (2015)). Les épisodes méditerranéens sont en effet susceptibles de produire des crues éclair dévastatrices, et un nombre conséquent d'inondations ont eu lieu par le passé (Vaison la Romaine en 1992, Hérault en septembre 2014, région de Cannes en octobre 2015, etc.). Les enjeux humains et financiers associés à ces inondations ont fait que de nombreux instruments ont été déployés, notamment par l'OHMCV (Observatoire Hydrométéorologique Cévennes-Vvarais, http://www.ohmcv.fr/) et dans le cadre de la campagne européenne HyMeX (Hydrological Cycle in Mediterranean Experiment : http://www.hymex.org/). (Bousquet et al. (2015)).

TABLE IV.1 – Caractéristiques des 4 liaisons microondes Terre-satellites géostationnaires enregistrées par le capteur Ku installé à Mirabel en 2013. Les indices les référençant sont ceux mentionnés sur la Figure IV.1. Pour les azimuth, 180 ° = vers le sud ; 270 ° = vers l'est.

indice (ouest vers est)	Position orbitale	Fréquence (GHz)	Polarisation	Azimuth	Élévation (α)
1	23.5 ° W	12.130	H	223.8 °	28.5 °
2	12.5 ° W	12.718	H	203.5 °	35.9 °
3	5 ° W	12.543	H	193.4 °	37.7 °
4	28 ° E	12.032	H	148.2 °	33.5 °

Le capteur Ku installé à Mirabel peut donc mesurer la puissance reçue de 4 satellites géostationnaires de télévision. Les caractéristiques de ces satellites sont résumées dans le Tableau IV.1. Leurs positions orbitales sont répaties entre 28 ° est et 23.5 ° ouest, et ils émettent tous entre 12 et 13GHz, en polarisation horizontale (notons que l'impact de la polarisation n'est pas étudié dans ce travail). Depuis Mirabel, situé à 44.6 ° nord, les satellites sont vus avec des azimuths entre 148 et 224 °, et des élévations (angles au sol) entre 28.5 et 37.7 °. On a vu que le signal était atténué à ces fréquences très majoritairement par la pluie (et non la glace ou les nuages). La longueur atmosphérique sur laquelle le signal est susceptible d'être atténué dépend donc de l'altitude de l'isotherme 0 ° C ("iso-0"). Les évènements étudiés ici (voir section 6) ayant lieu à l'été ou à l'automne 2013, cet isotherme

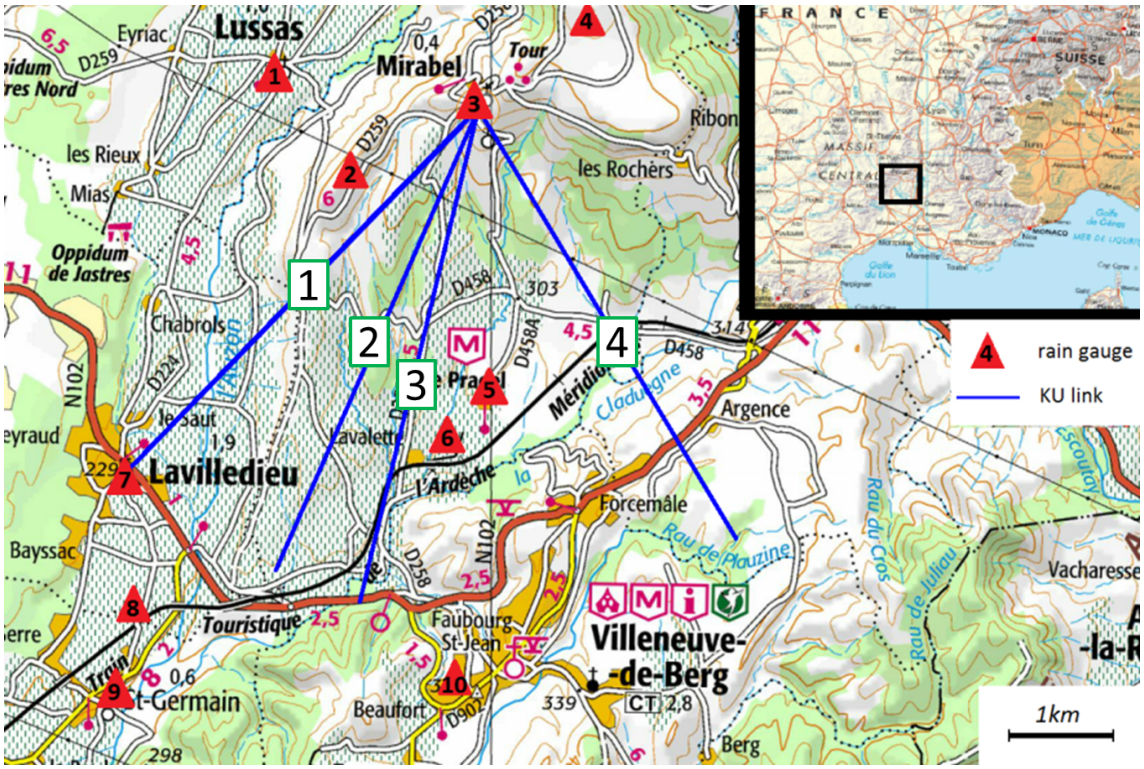

FIGURE IV.1 – Dispositif expérimental autour de Mirabel, Ardèche. Le capteur Ku est positionné dans le village de Mirabel, et capte 4 liaisons satellites (caractéristiques dans le Tableau IV.1). On représente ici ces liaisons en-dessous de l'isotherm 0 ° (là où la pluie peut atténuer le signal) avec un iso-0 vers $2800m$, et projetées au sol (traits bleus). On représente aussi la position des 10 pluviomètres du réseau HPiconet (triangles rouges).

sera compris entre 2.5 et $4km$, donnant des liaisons de longueurs entre 4 et $8.5km$. Nous projeterons ensuite ces liaisons au sol (en supposant les taux de pluie homogènes verticalement). La longueur des liaisons ainsi projetées au sol s'écrit : $z_0/tan(\alpha)$, avec α l'élévation décrite dans le Tableau IV.1 et z_0 l'altitude de l'iso-0. La Figure IV.1 présente les 4 liaisons Ku projetées au sol avec une longueur d'environ $5km$ correspondant à un iso-0 à environ $2800m$ d'altitude.

Dans la suite, nous ne reproduirons pas directement des champs de taux de pluie mais des champs d'atténuations linéiques à $12GHz$ (les deux étant reliés par l'Eq.(I.33) du chapitre I, section 2.4). En supposant sans erreur cette relation, l'atténuation à $12GHz$, A_{12}, est alors reliée à A_f, l'atténuation à la fréquence f, par :

$$A_{12} = a_f a_{12}^{-\frac{b_{12}}{b_f}} A_f^{\frac{b_{12}}{b_f}} L^{1-\frac{b_{12}}{b_f}} \tag{IV.1}$$

Avec L la longueur de la liaison et a_{12}, a_f, b_{12}, b_f les coefficients a et b de la loi puissance (I.33) pour les deux fréquences.

C'est ces atténuations qui seront utilisées comme observations. On note que toutes les fréquences des mesures utilisées ici (Tableau IV.1) sont proches de $12GHz$, ce qui limite les erreurs commises en appliquant l'Eq.(IV.1) (erreurs néanmoins non quantifiées ici).

2.2 Instruments pour validation

L'objectif de ce travail est de reproduire des champs de taux de pluie. La validation de sa pertinence peut donc être faite sur sa capacité à reproduire d'une part les quantités d'eau sur des zones entières, d'autre part la structure des champs de pluie. Ceci peut être fait en confrontant nos résultats à des observations radar. On peut aussi chercher à estimer la capacité de cette méthode à reproduire des taux de pluie ponctuels. Ceci sera fait en confrontant nos résultats à des mesures par pluviomètres, et à comparer ses performances sur ce point à celles d'un radar. Nous décrivons ici rapidement le radar et les pluviomètres disponibles.

2.2.1 Radar Météo France Bollène

Les données radar disponibles ici sont issues du radar opérationnel de Météo France de Bollène, dont les caractéristiques sont données dans le chapitre I, section 2.3.3 (radar scannant en bande S, photographie Figure I.12). Il est situé à $40km$ à l'est de Mirabel, en bordure est de la vallée du Rhône, à $312m$ d'altitude.

2.2.2 Pluviomètres HPiconet

TABLE IV.2 – Position des pluviomètres du réseau HPiconet. Les indices sont ceux définis sur la Figure IV.1.

indice	Nom	altitude (m)	latitude ($^\circ$N)	longitude ($^\circ$E)
1	Lussas	290	44.612	4.471
2	Mirabel Blaches	440	44.601	4.481
3	Mirabel mairie	500	44.607	4.499
4	Mirabel Mt Redon	630	44.614	4.515
5	Mirabel Pradel, fermes	280	44.583	4.499
6	Mirabel Pradel, vignes	270	44.580	4.450
7	Lavilledieu	230	44.577	4.453
8	St Germain, école	200	44.560	4.450
9	St Germain, station	190	44.550	4.449
10	Villeneuve de Berg	300	44.555	4.495

Les pluviomètres utilisés seront ceux du réseau HPiconet (http://www.ohmcv.fr/hpiconet/hpiconet-fr.html). Il s'agit d'un réseau de 22 pluviomètres à augets (caractéristiques au chapitre I, section 2.1) placés en 10 points dans la zone étudiée. La Figure IV.1 montre (triangles rouges) les positions de ces 10 points. Le Tableau IV.2 donne leur altitude et leur emplacement précis. Ils couvrent ainsi une zone d'environ $60km^2$. Leur résolution temporelle est de $5min$.

3 Équations du 4D-VAR pour ce problème

Nous allons dans cette section donner toutes les caractéristiques de l'algorithme d'assimilation 4D-VAR utilisé pour restituer des cartes de pluie à partir des atténuations enregistrées par le capteur Ku. On rappelle que la théorie du 4D-VAR a été donnée au chapitre III, section 2. Définir un tel algorithme 4D-VAR requiert de fournir tous les éléments non contrôlés des Eqs.(III.5) et (III.6), plus un minimiseur numérique et un état d'initialisation ("first guess" A^g).

L'inconnue de notre problème (champ x du 4D-VAR) ne sera pas ici un champ de taux de pluie mais un champ d'atténuations linéiques (en dB/km) à $12GHz$. Ces champs seront ensuite convertis (après la fin du processus d'assimilation) en taux de pluie en utilisant la loi puissance (Eq.(I.33)). Les erreurs associées à cette loi puissance ne seront ainsi pas incluses dans le processus d'assimilation (à part dans l'Eq.(IV.1), utilisée comme on l'a vu pour "ramener" les atténuations à $12GHz$) mais seulement dans la conversion finale en taux de pluie. Nous n'évoquerons donc que des champs d'atténuation dans cette partie théorique mais présenterons indifféremment des cartes de pluie ou d'atténuations dans les sections suivantes.

Notations

L'inconnue du problème (champ d'atténuations linéiques) sera notée A (au lieu de x dans le chapitre III sur l'assimilation de données). L'indice k désignera le temps (date t_k) et les indices i et j la position dans l'espace, selon les axes x (ouest-est) et y (nord-sud). On note NT, NX, NY, Δt, Δx et Δy les nombre de points mailles et les pas de discrétisation (en temps et en espace). Toutes les grandeurs considérées ici étant discrètes, on n'utilise pas de tildes pour les grandeurs discrètes (comme dans le chapitre III et au contraire du chapitre II). v_x et v_y désignent les vitesses de déplacement des champs de pluie (c'est-à-dire la force du vent) vers l'est et vers le sud. Elles sont supposées constantes et homogènes. Les autres notations sont celles des chapitres II ou III.

On note que dans toutes les équations formelles du 4D-VAR impliquant des calculs matriciels, A_k (atténuations à la date t_k) doit être un vecteur. Le champ (matrice) A_k défini ici sera donc dans ce cas-là implicitement (sans changer de notation) réécrit sous forme vectorielle. Soit :

$$A_k = \begin{pmatrix} A_{1,1} \\ A_{1,2} \\ \dots \\ A_{1,NY} \\ A_{2,1} \\ \dots \\ A_{NX,NY} \end{pmatrix}. \tag{IV.2}$$

Grandeurs à définir

Pour poser correctement notre algorithme 4D-VAR, il nous faut définir :

- une ébauche A_0^b et la matrice de covariance d'erreur associée $\boldsymbol{B}$.
- une initialisation A^g (first guess).
- des observations (vecteur y_k à la date t_k) et leurs matrices de covariance d'erreurs associées ($\boldsymbol{R_k}$).
- un modèle numérique de propagation spatiotemporelle des champs d'atténuation (noté M_k). On rappelle que $A_k = M_k(A_0)$ avec A_0 et A_k les champs d'atténuation (à reconstruire via l'algorithme d'assimilation) à t_0 et t_k. Voir section 3.1.
- des opérateurs d'observation H_k pour projeter les états du modèle sur l'espace des observations. Voir section 3.2.
- éventuellement des termes de pénalisation / filtrage (évoqués dans le chapitre III, section 4.2). Voir section 3.3.
- un algorithme de minimisation numérique.

Premiers choix et simplifications

Le minimiseur numérique sera l'algorithme $m1qn3$, décrit au chapitre III, section 3.2.3. N'ayant pas d'informations a priori sur la structure ou sur l'intensité des champs de pluie, nous n'utiliserons pas d'ébauche (soit $\boldsymbol{B}^{-1} = 0$). Les observations à la date t_k seront les atténuations totales (en dB) enregistrées sur chacune des 4 liaisons par le capteur Ku. L'espace des observations est donc, quelque soit la date, un espace de dimension 4. Des observations sont disponibles toutes les $10s$. On prendra par ailleurs : $\forall k$, $\boldsymbol{R_k} = \boldsymbol{Id}$. Au début du processus de minimisation, on prendra pour A^g un champ homogène d'une valeur faible et non nulle (typiquement $5.10^{-3} dB/km$).

3.1 Rappel modèle direct

On utilise pour propager les champs de pluie une EDP d'advection pure à vitesse constante, décrite dans le chapitre II section 1 par l'Eq.(II.3).

Cette équation est discrétisée par le schéma de Smolarkiewicz à deux dimensions, décrit en une dimension dans le chapitre II section 4 par l'Eq.(D.23) (et en deux dimensions en Annexe D). Les pas de discrétisation dans le temps et selon les deux directions d'espace seront, dans l'étude sur données réelles :

$$\begin{cases} \Delta t = 1s \\ \Delta x = 500m \\ \Delta y = 500m \end{cases} \tag{IV.3}$$

Dans l'étude sur données simulées, on utilise les valeurs voisines $\Delta x = \Delta y = 400m$. Dans le schéma de Smolarkiewicz, on prend $Sc = 1.04$ et $\epsilon = 10^{-14}$ (valeur non critiques, voir Smolarkiewicz (1983)).

3.2 Opérateur d'observation

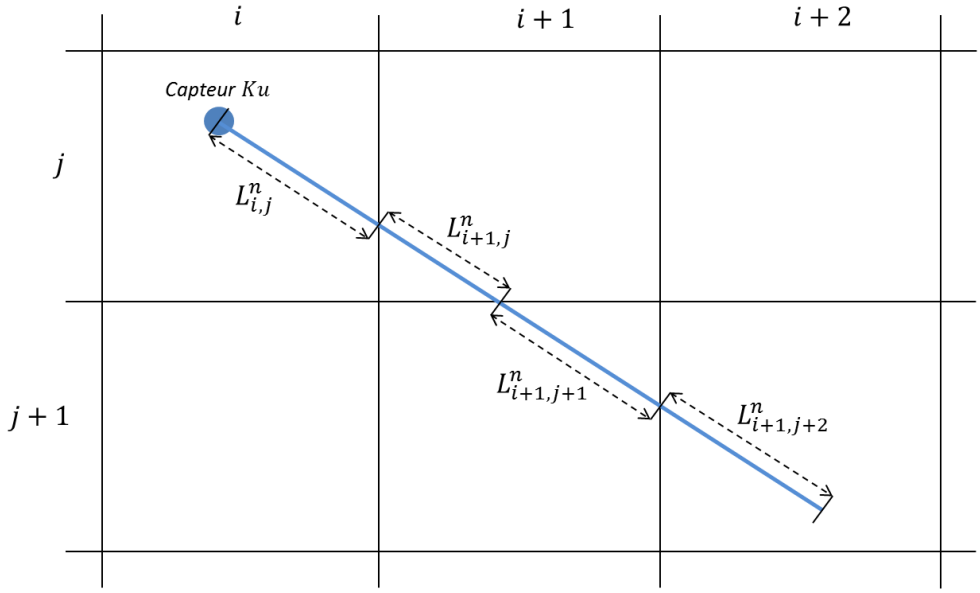

FIGURE IV.2 – Pour la n-ième liaison Ku projetée au sol (définie arbitrairement pour cet exemple), définition de la longueur traversée par la liaison sur chaque point de maille d'une grille de discrétisation. ($L^n_{i,j}$ est la longueur traversée dans le pixel (i,j).) $L^n = \sum_{i,j} L^n_{i,j}$ est la longueur totale de la liaison (km). Appliquée, à la date t_k, au champ d'atténuations linéiques A_k (dB/km), la n-ième composante de l'opérateur d'observation fournira ici la quantité $L^n_{i,j}A_{k,i,j} + L^n_{i+1,j}A_{k,i+1,j} + L^n_{i+1,j+1}A_{k,i+1,j+1} + L^n_{i+1,j+2}A_{k,i+1,j+2}$.

L'opérateur d'observation H_k à la date t_k projette l'inconnue du problème à la date t_k (champ d'atténuations linéiques A_k) sur l'espace des observations à cette même date (4 valeurs d'atténuations le long de liaisons Terre-satellite projetées au sol). La Figure IV.2 représente le problème pour une portion de grille à la date t_k contenant la n-ième liaison projetée au sol ($n \in [|1;4|]$). Les observations étant ici de la même nature à toutes dates, on a, $\forall k,\ H_k = H$.

Pour évaluer l'atténuation que provoquerait le champ A_k le long de la n-ième liaison, on effectue la somme des atténuation linéiques dans chaque point de maille traversé par la liaison, pondérée par la longueur de liaison la traversant. Formellement, on a ainsi :

$$\left(H(A_k)\right)_n = \sum_i \sum_j L^n_{i,j} A_{k,i,j} \tag{IV.4}$$

Avec $L^n_{i,j}$ la longueur de la n-ième liaison projetée au sol traversant le point de maille (i,j). Avec $n \in [|1;4|]$.

La partie écart aux observations de la fonction de coût est à présent pleinement définie. On rappelle qu'elle est notée J^{obs}. Elle s'écrit finalement :

$$J^{obs}(A_0) = \sum_{k=0}^{NT} ||y_k - H \circ M_k(A_0)||^2 \tag{IV.5}$$

3.3 Filtrage

Description

On a vu que nous n'introduirons pas formellement de terme d'ébauche, au sens habituel de l'assimilation, car nous ne disposons pas d'informations a priori sur la structure ou l'intensité des champs d'atténuations que l'on cherche à reconstruire. Dans une approche similaire à celle de Giuli et al. (1991) (mais avec un filtre différent), nous allons néanmoins introduire un terme de filtrage. Celui-ci est supposé introduire de la corrélation entre les points de maille voisins du champ d'atténuations. Ce terme de filtrage, partie de la fonction de coût, sera noté J^f. Il s'écrit :

$$J^f(A) = c^f \sum_{k=0}^{NT} J_k^f(A_k) \tag{IV.6}$$

Avec c^f un coefficient permettant de pondérer les deux termes de la fonction de coût, et pour la date t_k :

$$J_k^f(A_k) = ||A_k - m_k||^2 \tag{IV.7}$$

Formellement, A_k désigne toujours ici le champ d'atténuations à la date t_k réécrit sous forme vectorielle (Eq.(IV.2)). m_k désigne la matrice (elle aussi réécrite sous la même forme vectorielle) contenant les atténuations moyennées localement. Son terme général, $(m_k)_{i,j}$ s'écrit :

$$(m_k)_{i,j} = \frac{1}{9} \sum_{\substack{p=-1,0,1 \\ q=-1,0,1}} A_{k,i+p,j+q} \tag{IV.8}$$

On montre néanmoins dans l'Annexe F que ce terme de filtrage peut s'écrire sous la forme d'un terme d'ébauche.

Toutes les grandeurs nécessaires à l'exécution de l'algorithme d'assimilation 4D-VAR, et donc à la restitution de cartes d'atténuations à partir des mesures du capteur Ku, sont maintenant définies.

4 Détermination de la vitesse d'advection

Dans cette partie, nous donnons la méthode simple employée dans ce travail pour déterminer la vitesse de déplacement des champs de pluie à partir de cartes radars. On remarque que dans le cadre de cette thèse, nous avons développé une méthode originale de triangulation pour déterminer la vitesse de déplacement de champs de pluie directement à partir de mesures d'atténuations par capteur Ku. Cette méthode ne fonctionnant qu'avec au minimum les données de deux capteurs Ku voisins disponibles, elle n'a pas été employée dans les applications présentées ici (où l'on ne dispose que d'un seul capteur comme vu dans la section 2). Elle est présentée en détails, et évaluée sur données simulées, dans l'Annexe G.

Nous nous en tenons donc ici à une méthode simple pour déterminer une vitesse d'advection à partir de l'examen de la corrélation entre des cartes radar couvrant la cellule pluvieuse considérée à des dates successives. Notons t_k les dates successives où sont disponibles les cartes radar, Δt leur résolution temporelle et NT leur nombre. Notons RR_k chacune de ces cartes. Il s'agit d'un champ de taux de pluie, à résolution Δx (axe ouest-est) et Δy (axe sud-nord). Supposons ce champ de taille $NX * NY$. On note $Co_{k,k+1}$ la fonction d'autocorrélation entre les champs aux dates t_k et t_{k+1}. On a, pour $m \in [|1 : NX|]$ et $n \in [|1 : NY|]$:

$$Co_{k,k+1}(m,n) = \frac{1}{(NX-m)}\frac{1}{(NY-n)}\sum_{i=1}^{NX-m}\sum_{j=1}^{NY-n} RR_k(i,j)RR_{k+1}(i+m,j+n) \qquad \text{(IV.9)}$$

Notons $(m_k; n_k)$ le couple maximisant la fonction $Co_{k,k+1}$. On définit ensuite $v_{x_k} = m_k\Delta x/\Delta t$ et $v_{y_k} = n_k\Delta y/\Delta t$. On définit alors les paramètres d'advection comme la moyenne de ces paramètres (sur tout l'évènement), soit simplement :

$$v_x = \frac{1}{NT}\sum_{k=1}^{NT} v_{x_k} \qquad \text{(IV.10)}$$

Et de même pour v_y.

On peut noter que la valeur des maximum de corrélation (les $Co_{k,k+1}(m_k, n_k)$) fournissent une bonne information sur le caractère advectif ou non de l'évènement (plus la corrélation est forte plus on est proche d'une advection pure). L'écart-type de la série temporelle des $(v_x; v_y)$ fournit lui une bonne information sur le caractère constant ou non de la vitesse d'advection.

5 Validation de l'approche sur données simulées

Dans cette section, nous allons utiliser des données simulées pour évaluer la qualité des restitutions de notre algorithme sur 15 situations diverses selon si 1 ou 2 capteurs Ku sont disponibles.

Travaillant sur données simulées, nous travaillerons sans distinction avec des champs d'atténuations linéiques (en m^{-1}) ou de taux de pluie (en mm/h). Dans la mesure où nous supposons parfaite la relation puissance qui les relie, les deux approches sont strictement équivalentes. Les observations seront donc indifféremment des atténuations (sans unité) ou des taux de pluie intégrés sur les liaisons (en $m.mm/h$).

Dans cette section, le modèle de propagation (advection pure, voir section 3.1) sera à la résolution $\Delta x = \Delta y = 400m$ et $\Delta t = 1s$. On note toujours NT, NX et NY le nombre de pas de temps et d'espace. On a ici $NT \approx 1800$ (durée d'une simulation d'environ $30min$) et $NX = NY = 68$ (soit une carte d'environ $27 * 27km^2$). Les observations ont une résolution de $10s$ (on dispose donc de $NT/10$ observations pour chaque liaison Ku). On note NL le nombre de liaisons disponibles (on aura ici $NL = 4$ - un capteur Ku captant 4 satellites, ou 8 - 2 capteurs Ku captant chacun 4 satellites).

Dans la section 5.1, nous définirons tout d'abord la notion d'expérience jumelle, permettant d'évaluer l'algorithme d'assimilation sur données simulées. Puis nous donnerons la méthode utilisée pour simuler des cartes de pluie (ou d'atténuations), puis les critères mathématiques retenus pour évaluer les performances du modèle. Dans la section 5.2, nous étudierons les performances de l'algorithme sur 15 simulations avec différentes géométries et cartes de pluie pour le problème. On note que dans ces simulations, nous n'étudierons pas l'impact de potentielles erreurs modèles (par exemple mauvaise connaissance de la vitesse d'advection). Une étude à ce sujet sur quelques exemple particuliers peut être trouvée en Annexe H.

5.1 Méthodes / Définitions

5.1.1 Principe des expériences jumelles

Nous allons donc dans cette section évaluer notre algorithme d'assimilation sur des données simulées. C'est-à-dire que nous allons tester sa capacité à restituer des cartes de pluie parfaitement connues. Ce procédé, représenté schématiquement sur la Figure IV.3, est nommé expérience jumelle. La procédure pour effectuer une telle expérience jumelle est la suivante.

Tout d'abord, nous nous donnons une carte de pluie correspondant à l'état initial (à t_0) de notre système (voir paragraphe suivant pour la méthode employée pour produire une telle carte). C'est l'inconnue de notre problème, ce que nous chercherons ensuite à restituer. Cette carte d'atténuations (matrice de taille $NX * NY$) est notée A_0^{vrai}. Nous nous donnons par ailleurs une vitesse d'advection (vecteur $\boldsymbol{v}$) et choisissons le nombre et la géométrie des liaisons Ku.

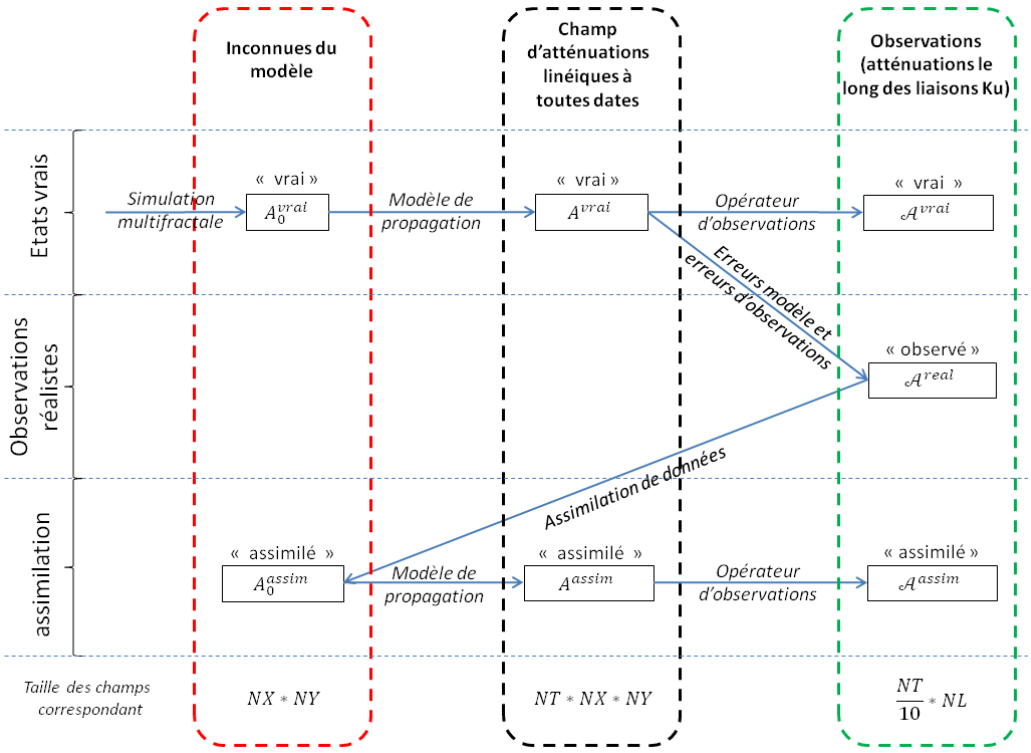

FIGURE IV.3 – Description du processus de test de l'algorithme de restitution des champs de pluie sur données simulées (expérience jumelle). Voir texte pour la description.

Nous utilisons ensuite notre modèle numérique de propagation (advection pure) pour propager A_0^{vrai} à toutes les dates. On obtient ainsi A^{vrai} (tableau $3D$ de taille $NT * NX * NY$). En appliquant ensuite l'opérateur d'observation le long des différentes liaisons Ku (voir section 3.2), on obtient les observations "vraies", $\mathcal{A}^{vrai}$ (matrice $NT/10 * NL$, on rappelle qu'on dispose d'une observation toutes les $10s$). Formellement $\mathcal{A}^{vrai} = H\left(A^{vrai}\right)$. Toutes ces grandeurs sont appelées les états "vrais" de notre problème. Ils sont représentés sur la ligne 1 de la Figure IV.3.

On peut ensuite ajouter des erreurs à ces observations ou les rendre plus réalistes. On peut typiquement ajouter un bruit. On obtient ainsi des observations dites "réalistes", notées $\mathcal{A}^{real}$.

Ces observations $\mathcal{A}^{real}$ constituent les données d'entrée de l'algorithme d'assimilation (les observations). On va ensuite faire tourner l'algorithme pour tenter de restituer A_0^{vrai}. En sortie d'assimilation, on obtient l'état initial "assimilé" du système, A_0^{assim}, comparable à l'état initial "vrai", A_0^{vrai} (Figure IV.3, ligne 3). Pour tester la capacité de l'algorithme à produire des champs cohérents non pas seulement à l'état initial mais à toute date, on peut aussi relancer le modèle de propagation pour avoir l'état assimilé à toutes dates, A^{assim}. En appliquant l'opérateur d'observations, on obtient les observations assimilées, $\mathcal{A}^{assim}$, dont l'écart à $\mathcal{A}^{real}$ fourni la valeur de J^{obs} en sortie d'assimilation.

5.1.2 Simulation de cartes de pluie

La première chose à faire pour effectuer une expérience jumelle est de simuler une carte de pluie pour l'instant initial du problème (état vrai A_0^{vrai} de la Figure IV.3).

Pour cela, on utilisera un générateur $2D$ multifractal, développé par Lovejoy and Schertzer (1990) et utilisant les paramètres déterminés par Verrier et al. (2010). Ce générateur est capable de reproduire la variabilité spatiale naturelle de la pluie. Dans ce travail, nous travaillerons sur des pluies assez fortes. Les cartes simulées présenteront ainsi des taux de pluie maximaux entre 80 et $100mm/h$, avec un pourcentage de pluie sur la carte fixé à 60%.

5.1.3 Caractéristiques des liaisons simulées

Dans cette section 5, nous utiliserons des liaisons simulées dont les caractéristiques sont données dans le Tableau IV.3. Ainsi, nous supposerons que chaque capteur Ku enregistre les émissions de 4 satellites. Les azimuts des satellites ainsi visés seront espacés de 20 $^\circ$, et centrés sur le sud. Les liaisons projetées au sol ont des longueurs entre 6 et $7km$.

TABLE IV.3 – Description des 4 liaisons Terre-satellite simulées et projetées au sol utilisées dans cette section.

	liaison 1	liaison 2	liaison 3	liaison 4
azimut ($^\circ$)	210	190	170	150
longueur (km)	6.5	6.1	6.1	6.0

Selon les cas, nous supposerons disposer d'1 ou 2 capteurs Ku. Dans ce dernier cas, le deuxième capteur sera placé à $2km$ au sud et $0.8km$ à l'est du premier. Cette configuration n'est pas arbitraire. Il s'agit de celle qui permettrait une détermination efficace des conditions d'advection directement à partir des mesures d'atténuation par triangulation dans le cas où 2 capteurs seraient disponibles, tout en conservant l'intérêt pour l'assimilation de données d'avoir des liaisons suffisamment proches (qui se croisent). Cette question, non primordiale dans le cadre de cette thèse, est abordée en détails dans l'Annexe G.

5.1.4 Zones de validation

Nous allons à présent définir des zones (partie des cartes de pluie) sur lesquelles nous évaluerons les performances de notre algorithme.

Supposons le modèle de propagation parfait (sans erreur). Considérons à présent une carte (i.e. un champ, un maillage) rectangulaire, et un capteur Ku disponible quelque part sur cette carte, captant 4 liaisons Ku. Seule une partie du champ de pluie rectangulaire va survoler à un moment ou à un autre les liaisons Ku. Et donc notre algorithme ne sera susceptible de restituer la pluie que sur une certaine zone de la carte. Nous n'avons donc bien sûr aucune prétention à restituer la pluie hors de cette zone, étant donné que nous n'avons aucun moyen de la mesurer.

La Figure IV.4 illustre cette situation. On a placé un capteur Ku au nord-est de la carte rectangulaire (cadre noir), et on suppose que la vitesse d'advection est donnée par le vecteur $\boldsymbol{v}$ (direction nord-nord-est). La partie de la carte en gris foncé ("grande zone d'assimilation") correspond à la partie de la carte initiale qui sera advectée au-dessus des liaisons Ku durant la durée T (durée de la simulation). C'est sur cette zone que nous prétendons restituer la pluie.

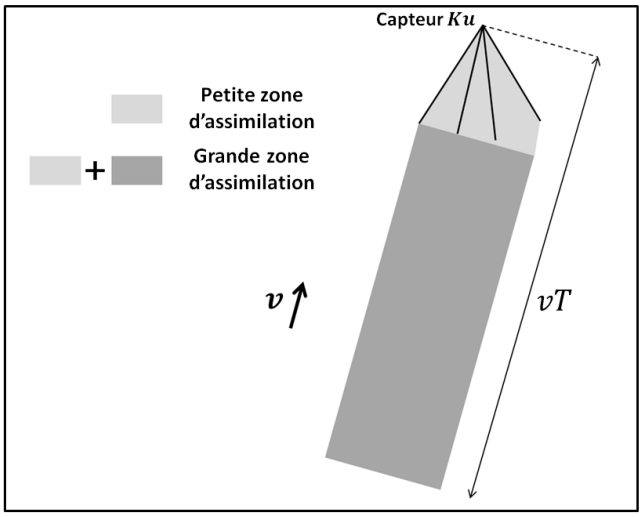

FIGURE IV.4 – Définition des zones d'assimilation pour l'évaluation des performances de l'algorithme de restitution des cartes de pluie. Voir texte pour les commentaires.

Nous introduisons également une "petite zone d'assimilation" (en gris clair sur la Figure IV.4), correspondant à la zone directement couverte par les liaisons Ku (traits noirs), pour deux raisons principales. Tout d'abord parce que la surface de la grande zone d'assimilation dépend des conditions d'advection ($\boldsymbol{v}$) et de la durée de la simulation (T). La petite zone d'assimilation a une surface fixe, ne dépendant que de la géométrie des liaisons. Ensuite parce que le modèle d'advection pure n'est bien sûr pas parfait et est toujours soumis à des erreurs plus ou moins fortes sur les cas réels. La différence de performance de l'algorithme entre les deux zones devrait permettre d'estimer la pertinence du modèle.

5.1.5 Indicateurs statistiques

Supposons que l'on veuille comparer deux champs, typiquement un champ vrai, par exemple A_0^{vrai}, et un champ assimilé, noté A_0^{assim}. Nous allons définir 4 indicateurs permettant d'avoir une bonne idée de l'écart entre ces deux champs, et donc de la qualité de la restitution. On introduit d'abord le biais absolu, noté $abias$, et défini par :

$$abias = \frac{1}{NX * NY} \left| \sum_{i=1}^{NX} \sum_{j=1}^{NY} \left(A_{0,i,j}^{vrai} - A_{0,i,j}^{assim} \right) \right| \tag{IV.11}$$

Avec NX et NY les dimensions des champs considérés.

On introduit ensuite le biais relatif (en %), noté $rbias$, et défini par :

$$rbias(\%) = 100. \frac{\sum_{i=1}^{NX} \sum_{j=1}^{NY} \left(A_{0,i,j}^{vrai} - A_{0,i,j}^{assim} \right)}{\sum_{i=1}^{NX} \sum_{j=1}^{NY} A_{0,i,j}^{vrai}} \tag{IV.12}$$

Puis le biais au quantile à 95%, noté $b95$ et défini par :

$$b95 = A95^{vrai} - A95^{assim} \tag{IV.13}$$

Avec $A95^{vrai}$ et $A95^{assim}$ les quantiles à 95% des deux champs.

On définit enfin la rmse (root mean square error) par :

$$rmse = \sqrt{\frac{1}{NX * NY} \sum_{i=1}^{NX} \sum_{j=1}^{NY} \left(A_{0,i,j}^{vrai} - A_{0,i,j}^{assim} \right)^2} \tag{IV.14}$$

On note bien que dans cette section, on présentera non seulement les valeurs de ces indicateurs pour différents couples de champs, mais aussi des statistiques de ces indicateurs eux-mêmes pour des séries de couples de champs (par exemple si l'on effectue 15 simulations en calculant à chaque fois un biais absolu, on peut estimer ensuite la moyenne des biais absolu, ou leur écart-type, etc.).

5.2 Influence du nombre de capteurs et de la géométrie du problème : 15 simulations

Dans cette section, nous étudions les restitutions de notre algorithme d'assimilation pour 15 simulations. D'une simulation à l'autre, la carte initiale de pluie vraie (A_0^{vrai}), la vitesse et la direction d'advection changent. La géométrie des liaisons ne varie en revanche pas. Nous utilisons les liaisons Ku décrites précédemment (Tableau IV.3) avec un ou deux capteurs, le second lorsqu'il existe étant donc disposé à $2km$ au sud et $0.8km$ à l'est du premier.

Les vitesses d'avection (module v et direction δ) utilisées sont fournies dans le Tableau IV.4. Elles varient de 9 à $21m/s$ (ce qui correspond bien à la gamme de vitesses de déplacement mesurées sur les cas réels, voir section 6), et de 180 ° (advection vers le nord) à 270 ° (advection vers l'est), directions correspondant là aussi aux données réelles. Les cartes initiales (A_0^{vrai}) sont toujours générées avec le simulateur multifractal (section 5.1.2). Elles sont présentées, restreintes à la grande zone d'assimilation, sur les Figures IV.5, IV.6 et IV.7, 1ère colonne. Dans tous les cas, la durée des simulations (temps T de la Figure IV.4) est d'environ $30min$.

TABLE IV.4 – Conditions d'advections pour les 15 simulations présentées dans la section 5.2. On rappelle $\delta = 180$ ° = déplacement vers le nord ; $\delta = 270$ ° = déplacement vers l'est.

simu	1	2	3	4	5	6	7	8
v (m/s)	16.7	17.1	17.9	16.8	13.0	8.9	21.0	21.2
δ	215 °	201 °	253 °	180 °	215 °	190 °	220 °	228 °

simu	9	10	11	12	13	14	15
v (m/s)	14.1	17.1	17.4	17.0	12.7	14.1	18.1
δ	226 °	217 °	227 °	229 °	251 °	216 °	244 °

Les observations vraies sont ici bruitées avant d'être insérées dans l'algorithme d'assimilation. Formellement, on a :

$$\mathcal{A}^{real} = \mathcal{A}^{vrai} + 2\mathcal{U}([-1;1]) \tag{IV.15}$$

Avec $\mathcal{U}$ la loi uniforme.

Ajouter une telle erreur correspond bien à l'incertitude sur une liaison Ku (de l'ordre de $0.5dB$, voir Barthes and Mallet (2013)).

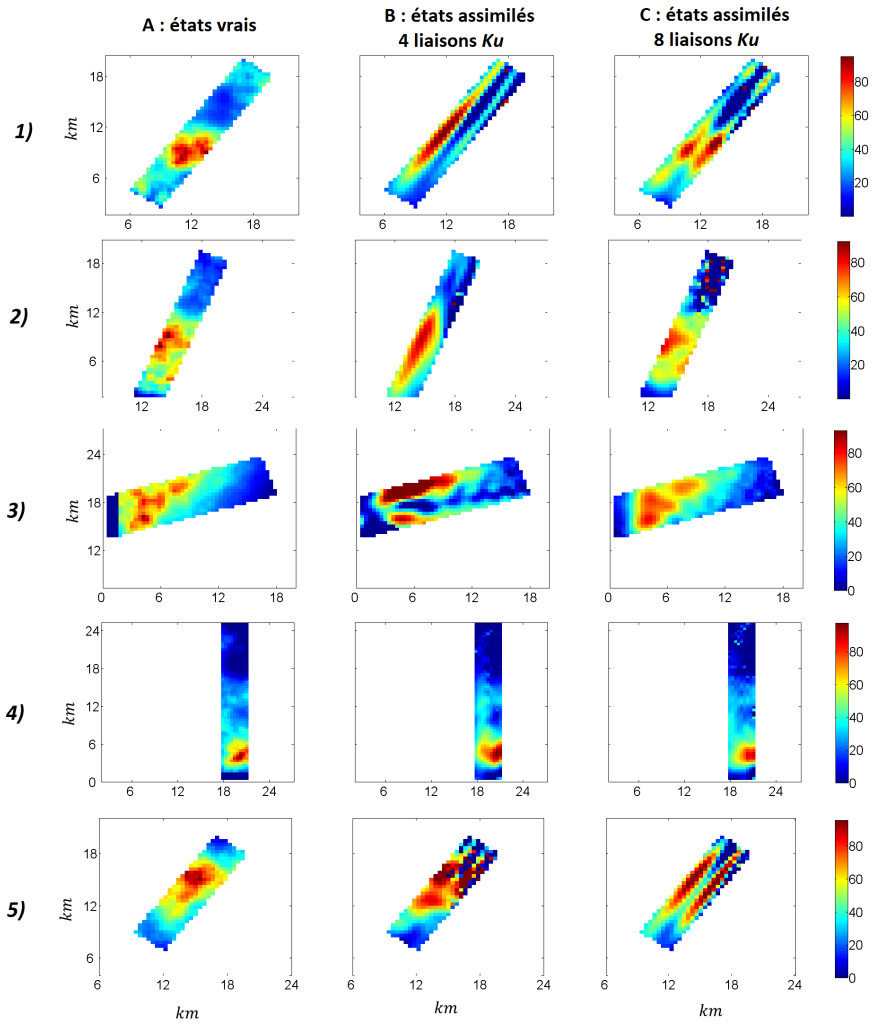

FIGURE IV.5 – Pour chacune des 15 simulations de la section 5.2. A) Champs initiaux vrais (A_0^{vrai}) restreints à la grande zone d'assimilation. B) Champs initiaux assimilés (A_0^{assim}) dans le cadre d'une expérience d'assimilation avec 1 capteur Ku. C) Idem avec 2 capteurs. Dans tous les cas, le(s) capteur(s) Ku est (sont) situé(s) au nord-est de la grande zone d'assimilation. Les vitesses d'advection correspondant à chaque simulation sont données dans le Tableau IV.4. Simulations 1 à 5 sur cette Figure.

Les Figures IV.5, IV.6 et IV.7 présentent, pour les 15 simulations, les champs initiaux vrais (A_0^{vrai}, colonne 1) et assimilés (A_0^{assim}) à partir des données de 4 liaisons Ku (1 capteur) (colonne 2), et assimilés à partir des données de 8 liaisons Ku (2 capteurs) (colonne 3). Les cartes sont

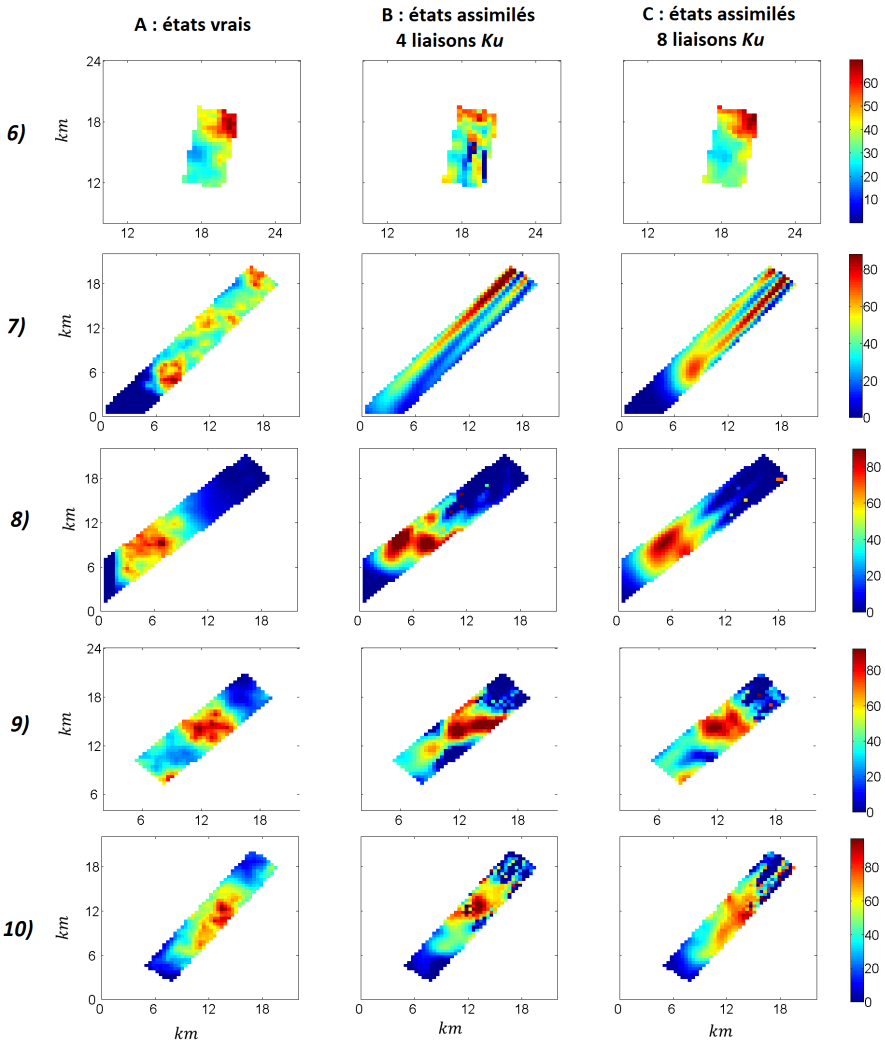

FIGURE IV.6 – Idem Figure IV.5 pour les simulations 6 à 10.

restreintes à la grande zone d'assimilation (définie dans la section 5.1) puisqu'on a vu que notre algorithme ne pouvait restituer la pluie hors de cette zone.

On constate de manière générale que la forme comme la position des cellules de pluie sont mieux déterminées avec 2 capteurs Ku qu'avec un seul (c'est particulièrement visible par exemple sur les simulations 2, 3, 9 ou 11), même s'il arrive que la restitution soit très difficile même avec 2 capteurs (simulation 7 par exemple), voire qu'elle semble parfois meilleure avec 1 capteur qu'avec 2 (simulation 5). Il semblerait que les simulations avec advection vers le nord-est soient également relativement plus difficile à restituer (simulations 1, 5, 7, 12) et celles vers l'est plus faciles (simulations 3, 13, 15). On peut enfin noter que les champs restitués, malgré la présence dans la fonction de coût de l'algorithme d'assimilation du terme de filtrage, semblent parfois assez bruités (simulation 10 par exemple, en particulier sa partie nord-est). Ce phénomène ne pose pas de problème majeur ensuite dans le coût car les valeurs très fortes dans les champs initiaux sont très rapidement lissées par la diffusion numérique lors de l'advection. Dans la suite, c'est bien néanmoins ces champs initiaux que l'on compare, car c'est tout de même eux qu'il s'agit de restituer et que le terme de filtrage a été introduit

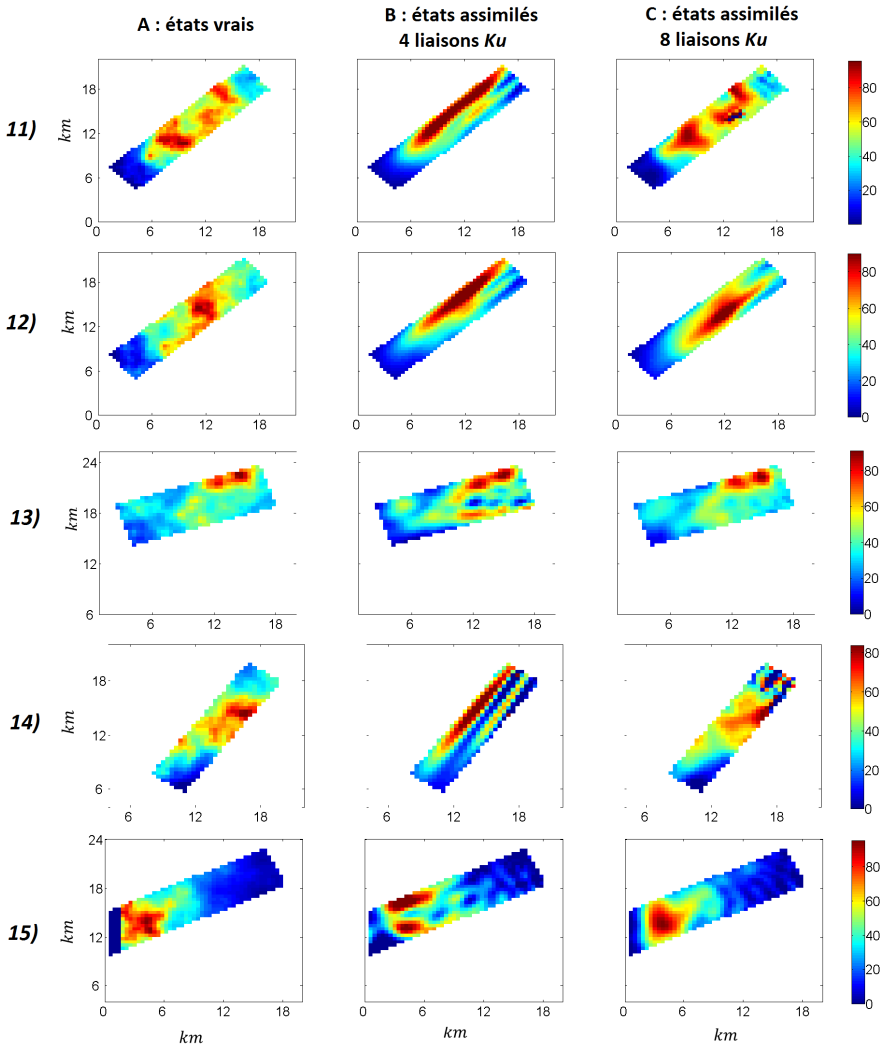

FIGURE IV.7 – Idem Figure IV.5 pour les simulations 11 à 15.

pour limiter ce phénomène.

Le Tableau IV.5 donne, pour chaque simulation, les valeurs des différents indicateurs statistiques définis plus haut (section 5.1). Le biais absolu (*abias*) donne l'erreur moyenne en mm/h, le biais relatif (*rbias*) rapporte cette valeur au taux précipitant moyen, la *rmse* indique la capacité de l'algorithme à restituer la structure fine des champs et le biais au quantile 95% ($b95$) la capacité de l'algorithme à reproduire les valeurs fortes. Le Tableau IV.6 fournit les valeurs moyennes et les écarts-types de ces indicateurs sur les 15 simulations, et en séparant les évènements donc l'advection se fait plus vers l'est que vers le nord (soit $\delta > 225^\circ$) et ceux où elle se fait au contraire plus vers le nord que vers l'est (soit $\delta < 225^\circ$).

Ces données confirment que les performances de l'algorithme sont sensiblement meilleures avec 2 capteurs qu'avec un seul. Cela est vérifié sur tous les critères en moyenne (Tableau IV.6). Cela est aussi vérifié pour 12 simulations sur 15 sur le biais, pour 14 simulations sur 15 pour la *rmse*, et également pour 14 simulations sur 15 pour le $b95$ (Tableau IV.5). On remarque de

TABLE IV.5 – Statistiques d'erreurs entre états initiaux vrais (A_0^{vrai}) et assimilés (A_0^{assim}) pour chacune des 15 simulations réalisées dans le section 5.2, pour des expériences d'assimilation avec 1 ou 2 capteurs Ku. Les indicateurs utilisés sont définis dans la section 5.1

simu	moyenne	*abias* (mm/h)		*rbias* (%)		*rmse* (mm/h)		*b*95 (mm/h)	
		$4Ku$	$8Ku$	$4Ku$	$8Ku$	$4Ku$	$8Ku$	$4Ku$	$8Ku$
1	38.5	2.7	0.5	+7.0	−1.3	29.0	16.9	−8.8	0.41
2	38.3	0.2	1.9	−1.0	−5.0	16.4	20.9	−7.4	−4.2
3	39.6	3.4	0.2	+9.0	−0.5	26.4	6.3	−16.4	−0.68
4	27.6	0.8	0.7	−3.0	−2.5	8.6	6.2	−9.7	−3.3
5	46.3	1.0	1.1	+2.0	+2.4	30.3	25.1	−17.4	−19.4
6	38.0	3.7	1.0	+10.0	+2.6	15.6	4.9	+8.4	−0.82
7	43.9	5.5	1.6	+13.0	+3.6	29.5	19.5	−14.5	−12.2
8	31.8	1.3	0.6	−4.0	+1.9	18.0	9.4	−20.3	−8.7
9	40.4	1.0	0.4	+2.0	+1.0	21.6	10.5	−13.3	−2.5
10	42.0	5.6	0.8	+13.0	−1.9	22.7	17.5	−7.2	−1.9
11	48.4	2.2	0.9	+5.0	+1.9	22.8	10.8	−20.2	−6.8
12	43.9	2.0	0.7	+5.0	+1.6	20.5	12.1	−16.1	−7.6
13	38.0	1.5	1.0	−4.0	+2.6	11.6	4.6	−4.4	−1.9
14	40.3	0.0	0.3	0.0	−0.7	28.5	14.8	−15.4	−3.2
15	32.5	0.3	0.2	+1.0	−0.6	22.1	7.4	−6.4	−1.8

plus qu'au niveau du biais, les résultats paraissent satisfaisants pour toutes les simulations avec 2 capteurs (erreur maximale de 5%). Avec 1 seul capteur, seules trois simulations présentent des erreurs de plus de 10%, et aucune de plus de 15% (Tableau IV.5). Le biais moyen sur les 15 simulations est faible dans tous les cas ($0.8mm/h$ avec 2 capteurs, $2.1mm/h$ avec un seul, Tableau IV.6). Avec 2 capteurs, il ne semble pas y avoir de décalage systématique, avec en moyenne sur les 15 simulations une surestimation de la pluviométrie de seulement 0.3%. Avec un seul capteur, on trouve une légère surestimation, de 3.6% (Tableau IV.6).

On remarque en revanche que dans les 2 cas, l'algorithme a tendance à surestimer le quantile à 95% (décalage moyen de 4.7% avec 2 capteurs et de 10.6% avec un seul). Cela signifie que l'algorithme produit des valeurs fortes plus fortes que l'état vrai. Cette caractéristique semble assez systématique, puisqu'elle est vérifiée pour 14 simulations sur 15, que l'on dispose d'1 ou de 2 capteurs. Cela vient encore une fois probablement d'un problème de diffusion numérique : ces valeurs extrêmes étant lissées rapidement par le schéma d'advection, elles ne pèsent ensuite rien dans la fonction de coût.

Au niveau de la structure fine des champs, caractérisée par la *rmse*, on constate (Tableau IV.5) des différences importantes d'une simulation à l'autre, avec par exemple avec 2 capteurs des erreurs allant de $4.6mm/h$ (simulation 13) à $25.1mm/h$ (simulation 5). Ces résultats sont cohérents avec ce qu'on avait pu voir sur les images des champs restitués (voir Figure IV.7 pour la simulation 13 et IV.5 pour la simulation 5). De telles comparaisons point à point à si fine résolution doivent cependant être utilisées avec prudence : dans les cas avec données réelles, où existeront d'autres sources d'erreur (sur $\boldsymbol{v}$ en particulier), il ne sera pas forcément

TABLE IV.6 – Statistiques d'erreurs entre états initiaux vrais (A_0^{vrai}) et assimilés (A_0^{assim}) sur les 15 simulations réalisées dans le section 5.2, pour des expériences d'assimilation avec 1 ou 2 capteurs Ku. Les indicateurs utilisés sont définis dans la section 5.1. On distingue également les évènements avec une advection majoritairement orientée vers l'est ($\delta > 225^\circ$) et vers le nord ($\delta < 225^\circ$).

tous évènements			
		$4Ku$	$8Ku$
$abias$ (mm/h)	moyenne	2.1	0.8
	std	1.8	0.5
$rbias$	moyenne	+3.6%	+0.3%
$b_{95\%}$ (mm/h)	moyenne	−10.6	−4.7
rmse (mm/h)	moyenne	21.6	12.4
	std	6.6	6.3
évènements nord-sud			
		$4Ku$	$8Ku$
$abias$(mm/h)	moyenne	2.4	1.0
$b_{95\%}$ (mm/h)	moyenne	−9.0	−5.6
rmse (mm/h)	moyenne	21.8	14.8
évènements est-ouest			
		$4Ku$	$8Ku$
$abias$(mm/h)	moyenne	1.7	0.6
$b_{95\%}$ (mm/h)	moyenne	−13.4	−4.3
rmse (mm/h)	moyenne	19.9	10.3

pertinent de regarder cet indicateur (voir section 6).

Commentons enfin les résultats du Tableau IV.6 sur la distinction des évènements avec advection majoritairement vers l'est ou vers le nord. Mis à part $b95$ avec 1 seul capteur Ku, tous les indicateurs donnent de meilleurs résultats dans le cas d'une advection orientée plutôt vers l'est. Ceci s'explique par la géométrie des liaisons. Lors d'une advection strictement vers l'est, les liaisons sont franchies l'une après l'autre par la cellule, et toutes les parties de celle-ci survolent successivement toutes les liaisons. Lors d'une advection vers le nord, certaines parties des cellules ne touchent qu'une seule liaison, et parfois immédiatement avant d'atteindre la latitude du capteur Ku, rendant d'autant plus difficile la restitution des cartes de pluie.

6 Résultats sur données réelles

Nous allons maintenant présenter nos résultats obtenus sur des données réelles enregistrées en Ardèche en 2013.

On note que dans cette section, les états vrais (A_0^{vrai} et A^{vrai}, voir Figure IV.3) ne sont pas connus. Nous disposons d'observations du capteur Ku, qui sont des mesures intégrées, indirectes et entachées d'erreurs de cet état vrai. Comme sur la Figure IV.3, ces observations seront notées $\mathcal{A}^{real}$. L'algorithme produit ensuite le meilleur estimateur de l'état vrai (sachant toutes les hypothèses du problème), noté toujours A_0^{assim} (carte initiale) et A^{assim} (cartes à toutes dates). En appliquant l'opérateur d'observations à A^{assim}, on obtient les observations assimilées (ou la reproduction des observations par l'algorithme d'assimilation), $\mathcal{A}^{assim}$. L'écart entre $\mathcal{A}^{real}$ et $\mathcal{A}^{assim}$ caractérise toujours la partie écart aux observations (J^{obs}) de la fonction de coût.

Dans cette section, nous décrivons d'abord les évènements étudiés (6.1), puis nous revenons sur la détermination de leurs vitesses d'advection associées (6.2). Les sections 6.3, 6.4 et 6.5 donnent les résultats, respectivement sur la capacité de l'algorithme à reproduire les observations, et sur sa confrontation aux mesures radar et pluviomètres.

6.1 Description des évènements

On a vu (section 2) que notre capteur Ku avait été installé dans le village de Mirabel, en Ardèche, en 2013. On a relevé durant cette période 8 évènements pluvieux pour lesquels étaient simultanément disponibles les mesures d'atténuations du capteur Ku, les mesures de réflectivité du radar Météo France de Bollène et les mesures des pluviomètres du réseau HPiconet.

Quelques caractéristiques de ces évènements (durée, atténuations moyennes et maximales le long des liaisons Ku, altitude de l'iso-0, quantité d'eau mesurée par les pluviomètres) sont présentées dans le Tableau IV.7.

L'altitude de l'iso-0 est déterminée à partir de la mesure de la température réalisée à la station Météo France d'Aubenas, située à $9km$ à l'ouest de Mirabel à $289m$ d'altitude, et en supposant un gradient de $-1\,^{\circ}C/150m$. On vérifie la pertinence de ce gradient en comparant avec la température relevée à la station Météo France de Berzème, située à $8km$ au nord-est de Mirabel et à $740m$ d'altitude.

Les évènements étudiés sont assez diversifiés en intensité, avec des cumuls d'eau moyens enregistrés par les pluviomètres variant d'à peine $1mm$ à plus de $20mm$. Tous durent en revanche entre 30 et $45min$. Les évènements longs (le 20/10 par exemple) ont été divisés en plusieurs parties. On a en effet vu que l'on devait se restreindre à des durées assez courtes pour d'une part que le modèle d'advection pure reste raisonnable et que d'autre part on évite une trop grande diffusion numérique. On peut déjà voir que ces évènements semblent assez variables dans l'espace, les pluviomètres voyant des cumuls d'eau variant parfois du simple au

TABLE IV.7 – Caractéristiques des évènements étudiés de l'année 2013. Dates (colonnes 2, 3, 4), atténuations (dB) maximales et moyennes relevées le long des 4 liaisons Ku (colonnes 5 et 6), estimation de l'iso-0 (colonne 7), et cumuls d'eau (mm) enregistrés par les 10 pluviomètres du réseau HPiconet : pluviomètre ayant vu le moins d'eau (colonne 8), le plus d'eau (colonne 9) et moyenne sur les 10 (colonne 10).

N°	Date	Début	Fin	attenuation (dB)		iso-0	cumul pluviomètres (mm)		
				moyenne	max		min	max	moyenne
1	07/09	14:00	14:35	1.64	5.29	3.5 km	0.1	2.3	0.7
2	07/09	17:35	18:10	1.19	5.00	3.0 km	5.0	12.4	8.4
3	08/09	14:40	15:25	2.51	12.78	3.9 km	0.1	12.0	4.8
4	28/09	18:05	18:40	1.20	9.89	3.5 km	0.8	4.3	1.9
5	29/09	02:05	02:50	4.37	13.97	2.8 km	16.2	31.5	21.7
6	04/10	17:45	18:25	1.95	7.49	3.0 km	8.3	12.8	10.0
7	20/10	06:35	07:05	2.56	11.28	2.5 km	4.2	9.4	6.6
8	20/10	08:20	08:50	5.32	12.48	2.5 km	15.6	21.5	18.0

double (évènement 5) voire plus (évènement 3).

Les Figures IV.8 et IV.9 (traits continus) présentent les relevés d'atténuations (en dB) du capteur Ku pour chacun des évènements et chacune des 4 liaisons. On remarque là aussi des différences assez sensibles entre les enregistrements réalisés le long des différentes liaisons, tant en intensité (par exemple pour l'évènement 1, maximum autour de $5dB$ pour la liaison 1 et autour de $2dB$ pour la liaison 4) que dans la forme des séries temporelles (voir évènement 3 par exemple).

6.2 Vitesse d'advection

Nous avons vu (section 3) que le modèle numérique utilisé dans l'algorithme d'assimilation était basé sur une EDP d'advection pure, et que dans le cadre de ce travail nous allions estimer la vitesse d'advection en maximisant la corrélation croisée entre les cartes radars successives fournies par le radar de Bollène (voir description de la technique dans la section 4). La Figure IV.10 présente pour deux des évènements étudiés les cartes radars successives (toutes les 10 minutes) sur des zones de 1000 à $2000km^2$ centrées sur le capteur Ku.

On voit que pour l'évènement 3 (08/09, Figure IV.10A), la propagation du champ de pluie ne semble pas du tout obéir à un modèle d'advection. La première petite cellule présente à 14:50 vers la coordonnée ($10km$ nord / $22km$ est) semble s'être développée très fortement depuis la carte précédente et se résorbe tout aussi rapidement ensuite. La cellule principale se déplaçant vers le nord-est entre 15:00 et 15:20 change elle significativement d'intensité et de forme. On peut donc d'ors et déjà s'attendre à des résultats insatisfaisants si l'on traite cet évènement avec notre algorithme, puisque le modèle direct de propagation sera incapable de restituer la dynamique de l'évènement perçue sur les cartes radar.

Au contraire, l'évènement 7 (20/10, Figure IV.10B) semble correspondre beaucoup mieux

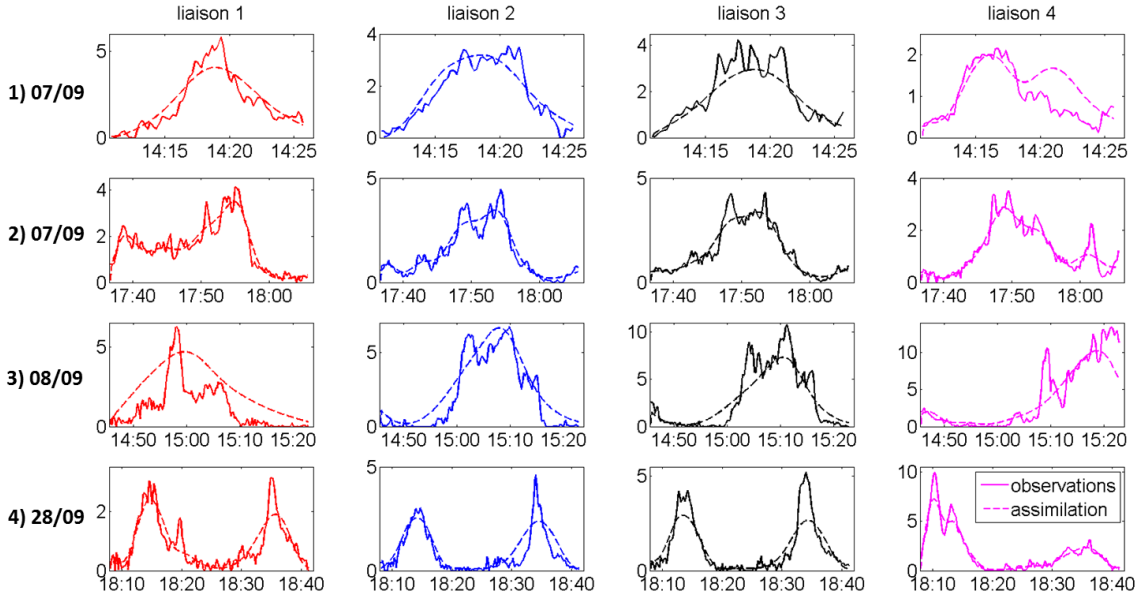

FIGURE IV.8 – Pour chaque évènement (1 à 8, lignes) et chaque liaison Ku (1 à 4, colonne) : observations du capteur Ku ($\mathcal{A}^{real}$, en dB, traits continus) et reproduction des observations par l'algorithme d'assimilation ($\mathcal{A}^{assim}$, traits pointillés, en dB). Évènements 1 à 4 sur cette Figure.

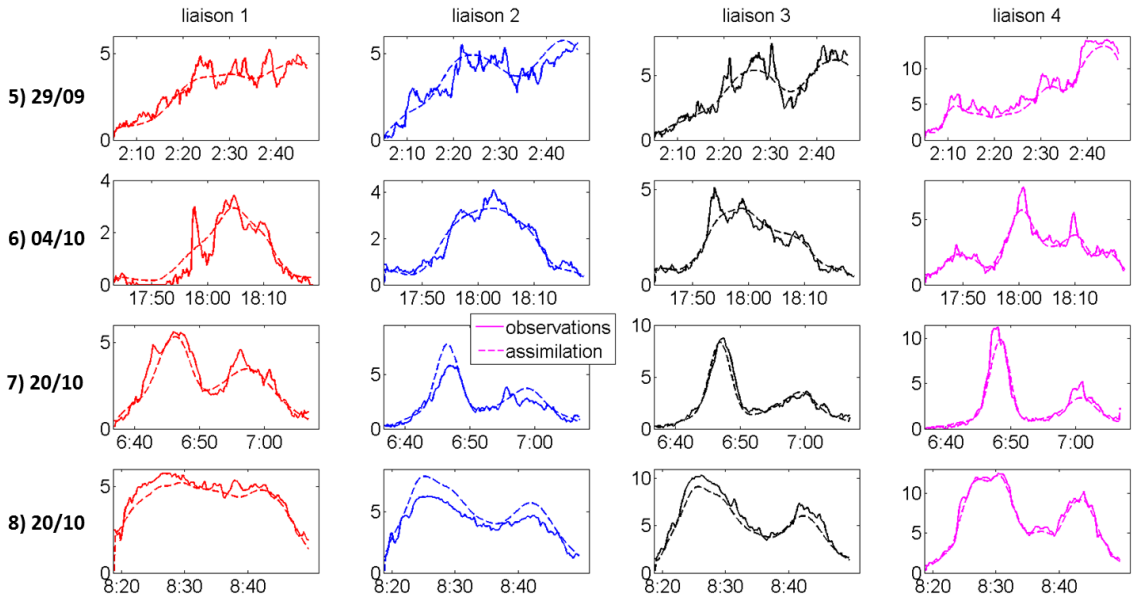

FIGURE IV.9 – Idem Figure IV.8, évènements 5 à 8. Unité : dB.

à un modèle d'advection pure. La quantité d'eau n'est bien sûr pas strictement conservée mais sur les $30min$ de l'évènement, la cellule reste globalement intacte et se déplace régulièrement vers le nord-est. On a donc de bien plus grands espoirs pour un tel évènement.

Afin d'évaluer un peu plus formellement ce caractère advectif ou non des évènements, et pour pouvoir estimer avant même l'application de l'algorithme d'assimilation si celui-ci sera ou non susceptible de restituer une carte de pluie réaliste, nous allons introduire 2 indicateurs.

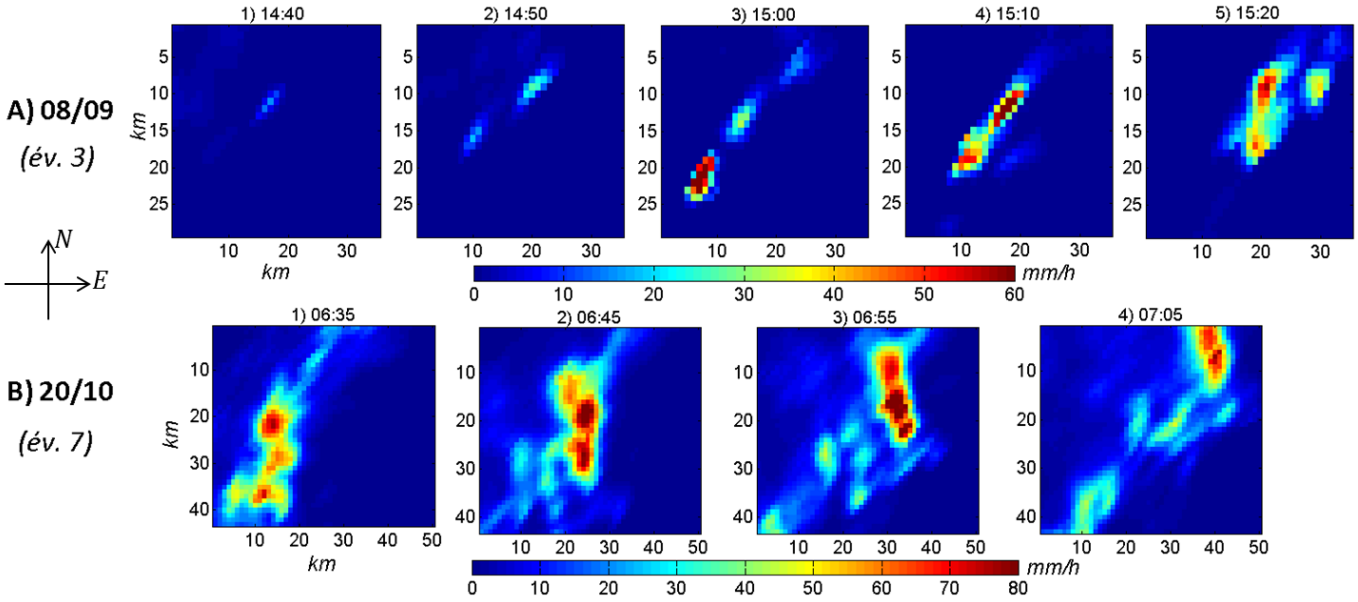

FIGURE IV.10 – Cartes radar successives pour 2 évènements de l'année 2013. A) Évènement peu advectif : 28/09, 14:40 à 15:20. A) Évènement davantage advectif : 20/10, 6:35 à 7:05. Les cartes présentées sont centrées sur la position du capteur Ku.

Pour chaque couple de cartes radar (séparées de 5 minutes), nous enregistrons la valeur du maximum de corrélation croisée (le $Co_{k,k+1}(m_k, n_k)$ de l'Eq.(IV.9)). Pour un évènement donné, nous obtenons ainsi une série temporelle de valeurs de maximums de corrélation. Nous calculons alors la moyenne et l'écart-type de cette série temporelle. Ces valeurs sont reportées pour chaque évènement dans le Tableau IV.8 (moyenne, colonne 3 ; écart-type, colonne 4). La première valeur sera d'autant plus forte que l'on a globalement des caractéristiques advectives fortes pour l'évènement considéré. La deuxième valeur sera d'autant plus basse que ces caractéristiques advectives persistent durant tout l'évènement. On constate que 3 évènements (2, 3, 6) présentent des coefficients de corrélation moyens (0.81) plus faibles que les autres (0.89 à 0.92). L'évènement 3 présente aussi une grande dispersion dans ses coefficients de corrélation (écart-type à 0.082). Les évènements 1, 2 et 6 présentent des écarts-types entre 0.025 et 0.030. Les 4 autres évènements (4, 5, 7, 8) ont des écarts-type en dessous de 0.02. Nous reviendrons sur ces données dans la section 6.4.1 lorsque nous comparerons les restitutions de notre algorithme aux cartes radar.

Le Tableau IV.8, colonnes 5 et 6, donne aussi les vitesses d'advection obtenues par l'analyse des cartes radar. On constate que dans tous les cas, l'advection se fait vers le nord-est, allant d'une advection quasiment sud-nord (évènement 4) à majoritairement ouest-est (évènement 3).

6.3 Résultats / Efficience de l'algorithme

Nous allons dans un premier temps rapidement évaluer la faculté de l'algorithme d'assimilation à restituer les observations du capteur Ku (c'est-à-dire à minimiser la fonction

TABLE IV.8 – Restitution des paramètres d'advection par analyse de la corrélation entre cartes radar successives pour les évènements de l'année 2013. On calcule entre chaque couple de cartes successives une valeur maximale de corrélation croisée, puis la moyenne (colonne 3) et l'écart-type (colonne 4) de ces valeurs sur tous les couples. On définit la vitesse d'advection (colonnes 5 et 6) comme la moyenne sur tous les couples des décalages permettant de maximiser les corrélations croisées. Les évènements présentant les caractères advectifs les plus marqués sont en gras.

N°	Date	Qualité advection		Paramètres d'advection	
		coeff. corr. moy.	std. coeff. corr.	vitesse	direction
1	07/09	0.90	0.025	$11.5m/s$	210°
2	07/09	0.81	0.026	$12.6m/s$	199°
3	08/09	0.81	0.082	$6.2m/s$	238°
4	**28/09**	**0.90**	**0.015**	$\mathbf{13.6m/s}$	$\mathbf{191^\circ}$
5	**29/09**	**0.89**	**0.019**	$\mathbf{7.9m/s}$	$\mathbf{208^\circ}$
6	04/10	0.81	0.030	$12.1m/s$	194°
7	**20/10**	**0.92**	**0.018**	$\mathbf{16.2m/s}$	$\mathbf{220^\circ}$
8	**20/10**	**0.89**	**0.015**	$\mathbf{16.2m/s}$	$\mathbf{220^\circ}$

de coût du 4D-VAR). Nous allons donc comparer, pour chaque évènement, $\mathcal{A}^{real}$ et $\mathcal{A}^{assim}$. On rappelle qu'une bonne restitution des observations est une condition nécessaire pour que l'algorithme soit performant mais ne garantit pas qu'il restitue correctement les champs de pluie, dans la mesure d'une part où il peut y avoir plusieurs solutions au problème et d'autre part qu'il peut exister des erreurs modèle.

Les Figures IV.8 et IV.9 déjà évoquées présentent, pour chacun des 8 évènements et des 4 liaisons Ku, les mesures ($\mathcal{A}^{real}$, traits continus) et la sortie de l'algorithme d'assimilation ($\mathcal{A}^{assim}$, traits pointillés).

On constate que :

- globalement, la dynamique enregistrée par les liaisons Ku est bien reproduite par l'algorithme. À la fois pour les évènements ne présentant qu'un pic principal (une cellule) (évènement 1 ou 6), ou ceux présentant deux pics (4, 7, 8). Les différences de timing entre liaisons sont bien restituées. Par exemple sur l'évènement 3 le fort décalage temporel entre les liaisons est bien reproduit (les maxima d'atténuations sont restitués aux bons moments pour chaque liaison, à savoir vers 14:55 sur la liaison 1, vers 15:20 sur la liaison 4). Les différences d'intensité sont bien reproduites (voir par exemple le premier pic de l'évènement 4, simultané sur les 4 liaisons (vers 18:15) mais à $2.5dB$ sur la liaison 1 et à $10dB$ sur la liaison 5).

- quelques fausses détection existent (évènement 1, liaison 4 : l'algorithme restitue un second pic vers 14:20 qui n'existe pas), ainsi que quelques erreurs significatives (évènement 3, liaison 1 : le maximum est relativement bien placé mais la cellule est beaucoup trop lisse, la dynamique est mal restituée). Ces erreurs proviennent vraisemblablement d'erreurs modèle. Par exemple pour l'évènement 3, l'advection pure dans le modèle se fait quasiment d'ouest en est. La cellule est donc forcée de traverser successivement toutes les liaisons (ce qui explique que $\mathcal{A}^{assim}$, pointillés

Figure IV.8, ligne 3, ait sensiblement la même forme pour les 4 liaisons). Si dans les faits la cellule se déforme durant son passage d'une cellule à l'autre, l'algorithme ne peut produire qu'une moyenne entre les différents signaux perçus.

- le signal restitué est de manière générale lissé par rapport aux mesures. Les valeurs extrêmes sont souvent sous-estimées. Prenons l'exemple de l'évènement 4 du 28/09 (Figure IV.8, ligne 4). La seconde cellule de pluie (passant sur les liaisons Ku vers 18:35) est largement lissée lors de sa restitution par l'algorithme d'assimilation. Ceci est la conséquence d'une autre erreur modèle. Comme nous l'avons déjà mentionné, il existe une diffusion numérique qui tend à rapidement détruire les forts gradients éventuellement présents dans les cartes de pluie à l'instant initial d'un évènement. Les séries temporelles d'atténuations le long des liaisons sont donc nécessairement lissées par ce phénomène, et le sont d'autant plus que la cellule considérée a été advectée sur une longue distance avant de toucher la liaison. C'est une des raisons faisant qu'on limite la durée des évènements considérés à 30 minutes (quitte à les diviser en plusieurs sous-évènements).

6.4 Résultats / Restitutions intégrées : comparaison au radar

Nous allons à présent comparer les champs restitués par l'algorithme d'assimilation à ceux restitués par les mesures du radar Météo France de Bollène. Dans un premier temps (section 6.4.1), nous constaterons à partir de différents critères statistiques que notre algorithme d'assimilation ne prétend restituer, comme nous le pressentions, les cartes de pluie que pour des évènements suffisamment advectifs. Nous établirons ensuite, à partir de ces résultats, un critère sur les corrélations entre cartes radar du Tableau IV.8 pour définir cette notion de "suffisamment advectif". Nous nous restreindrons ensuite (section 6.4.2) aux 4 évènements vérifiant ce critère et présenterons différents résultats.

6.4.1 Résultats pour 8 évènements selon leur caractère advectif

Le radar de Bollène fournit des cartes de pluie à $5min$ et $1km^2$. Notre algorithme produit lui des cartes à $1s$ et $0.5*0.5km^2$. Afin de comparer les deux produits, nous dégradons la résolution de nos cartes jusqu'à celle du radar. Nous dégradons également, dans un second temps, les deux produits à une résolution de $2*2km^2$ (et toujours $5min$). Le Tableau IV.9 présente, pour chacun des 8 évènements, divers indicateurs caractérisant l'écart entre les deux produits. Les résultats sont présentés pour des cartes restreintes à la petite ou à la grande zone d'assimilation. On calcule ainsi la droite de régression linéaire ainsi que le coefficient de corrélation associé, pour des cartes à 1 et $4km^2$, ainsi que les biais absolu et relatif (rapporté au taux de pluie radar moyen).

Si la propagation effective des champs de pluie n'est pas régie par une advection pure, on peut s'attendre notamment à ce qu'il y ait création ou destruction de pluie durant l'évènement. Et puisque notre modèle d'advection conserve lui strictement la quantité totale d'eau, on peut dans ces cas-là s'attendre à avoir des biais lorsque l'on compare les résultats dans la grande zone d'assimilation (non directement couverte par les liaisons Ku mais indirectement via le modèle de propagation). Les biais relatifs dans cette zone, pour les différentes évènements, sont donnés

TABLE IV.9 – Comparaison entre cartes radar et restituées par notre algorithme d'assimilation pour chaque évènement de l'année 2013, pour 2 résolutions différentes ($1*1$ et $2*2km^2$), et pour les deux zones d'assimilation (voir Figure IV.4). Coefficients de corrélation, droites de régression, biais (biais positif = surestimation de la pluie par notre algorithme par rapport au radar), taux de pluie radar moyen.

petite zone d'assimilation							
Évènement	Coeff. corrélation		droite de régression		biais	tx pluie moy.	biais
	$1\ km^2$	$4\ km^2$	$1\ km^2$	$4\ km^2$	$mm\ h^{-1}$	$mm\ h^{-1}$	relatif
4 (28/09)	0.85	0.92	$0.98x + 1.3$	$1.08x + 0.5$	+1.1	5.2	+21%
5 (29/09)	0.77	0.84	$0.95x + 0.5$	$0.94x + 1.1$	-0.7	24.9	−2.8%
7 (20/10)	0.81	0.88	$0.44x + 6.1$	$0.55x + 2.5$	-7.9	25.1	−31%
8 (20/10)	0.39	0.82	$0.40x + 26$	$0.83x + 3.5$	-2.7	47.1	−5.7%
1 (07/09)	0.69	0.67	$0.19x + 4.4$	$0.24x + 2.6$	-12.8	21.3	−60%
2 (07/09)	0.17	0.32	$0.13x + 0.3$	$0.19x + 1.5$	-5.5	9.0	−61%
3 (08/09)	0.62	0.66	$0.47x + 9.9$	$0.54x + 6.9$	+2.7	13.8	+20%
6 (04/10)	0.42	0.63	$0.36x + 8.9$	$0.54x + 4.8$	-0.3	14.4	−2.1%
grande zone d'assimilation							
Évènement	Coeff. corrélation		droite de régression		biais	tx pluie moy.	biais
	$1*1\ km^2$	$2*2\ km^2$	$1*1\ km^2$	$2*2\ km^2$	$mm\ h^{-1}$	$mm\ h^{-1}$	relatif
4 (28/09)	0.48	0.52	$0.68x + 1.3$	$0.74x + 0.97$	-0.1	3.6	−2.7%
5 (29/09)	0.51	0.53	$0.53x + 3.8$	$0.51x + 4.7$	-6.9	34.7	−20%
7 (20/10)	0.81	0.87	$0.58x + 3.5$	$0.68x + 2.0$	-2.1	12.2	−17%
8 (20/10)	0.64	0.78	$0.72x + 4.7$	$0.85x + 1.3$	-6.5	35.4	−18%
1 (07/09)	0.72	0.70	$0.21x + 2.0$	$0.23x + 1.3$	-10.8	16.0	−68%
2 (07/09)	0.41	0.49	$0.26x + 0.34$	$0.30x + 0.14$	-6.6	9.0	−73%
3 (08/09)	0.52	0.67	$0.45x + 14$	$0.61x + 8.1$	+3.5	19.5	+18%
6 (04/10)	0.43	0.61	$0.33x + 5.6$	$0.45x + 3.5$	-7.9	19.9	−40%

dans la colonne 8 du Tableau IV.9. On constate que trois évènements présentent des biais largement plus forts que les autres, les évènements 1 (sous-estimation de 68% par l'algorithme d'assimilation comparé au radar), 2 (sous-estimation de 73%) et 6 (sous-estimation de 40%). Ces trois évènements présentent également des équations de régression très plates (pentes à $1km^2$ de 0.21, 0.26, 0.33 dans la grande zone d'assimilation et 0.19, 0.13, 0.36 dans la petite). Cela implique que les champs restitués sont fortement lissés par rapport aux champs radar. Les coefficients de corrélation sont assez bas également (voir Tableau IV.9). Ces deux dernières remarques (sur la pente de régression et le coefficient de corrélation) sont également vraies, dans une moindre mesure, pour l'évènement 3, et pour l'évènement 5 uniquement dans la grande zone d'assimilation. Au contraire, on peut noter que dans la petite zone d'assimilation et à $4km^2$ de résolution, les évènements 4, 5, 7 et 8 présentent tous des coefficients de corrélation au-dessus de 0.8 tandis que les 4 autres présentent des coefficients de corrélation en-dessous de 0.7. On note que l'évènement 8 présente un coefficient de 0.39 à $1km^2$ mais bien de 0.82 à $4km^2$, suggérant

un problème d'échelle sur lequel nous reviendrons.

Revenons à présent sur les coefficients de corrélation (moyenne et écart-type) entre cartes radar successives du Tableau IV.8. On remarque que les évènements 4, 5, 7 et 8 sont les 4 présentant à la fois une corrélation moyenne au-dessus de 0.85 et un écart-type en-dessous de 0.02. Or ils semblent bien être les mieux restitués par l'algorithme d'assimilation. L'utilisation des valeurs et de la dispersion des maximums de corrélation entre cartes radar successives semble donc être un bon moyen de déterminer l'opportunité d'utiliser notre algorithme de restitution. Même s'il parait raisonnable, le critère numérique mentionné ici (corrélation moyenne au dessus de 0.85 et écart-type en-dessous de 0.02), ayant été réglé a posteriori sur les évènements disponibles, demande bien sûr à être vérifié.

Dans la suite, nous restreindrons notre analyse aux évènements ainsi définis comme "suffisamment advectifs", à savoir les évènements 4, 5, 7 et 8.

6.4.2 Restriction à 4 évènements

Nous allons donc à présent restreindre l'analyse aux 4 évènements advectifs (évènements 4, 5, 7, 8). Les Figures IV.11 à IV.14 présentent, pour chacun de ces évènements, les champs de pluie, restreints à la grande zone d'assimilation, mesurés par le radar de Bollène (A) et restitués par l'algorithme d'assimilation (B). Tous ces champs sont dégradés à résolution $5min$, $1km^2$. Dans tous les cas, le capteur Ku est au nord-est de la zone d'assimilation (sur le pixel $26km$ nord, $21km$ est). La Figure IV.15 présente, pour chaque évènement, des boîtes à moustaches de la différence entre les champs radar et assimilés restreints à la grande zone d'assimilation, pour des résolution $1km^2$ et $4km^2$.

Concentrons nous tout d'abord sur l'évènement 4 du 28 septembre. On peut voir sur les séries temporelles d'atténuations le long des liaisons Ku (Figure IV.8, ligne 4) qu'il y a deux pics distincts dans l'évènement, le premier vers 18:10, le second vers 18:35. La carte radar pour cet évènement (Figure IV.11A) montre que le second pic correspond à une cellule de pluie dont l'intensité augmente graduellement de 18:25 à 18:40 alors qu'elle se dirige vers le nord en s'approchant du capteur Ku. En étudiant les résultats de l'algorithme d'assimilation pour cet évènement, on peut noter que :

- le taux de pluie moyen près des liaisons Ku (c'est-à-dire dans la petite zone d'assimilation) est correctement reproduit (biais moyen de $1.1mm/h$, voir Tableau IV.9). Le taux de pluie moyen est également bien reproduit dans la grande zone d'assimilation (biais moyen de $-0.1mm/h$). Cela montre une fois encore qu'un avantage de ce système de mesure et de cette méthode de restitution est que toute la pluie tombant sur une zone donnée est vue par les liaisons Ku (grâce à la très bonne résolution temporelle du système : $10s$) et restituée par l'algorithme. Ensuite, grâce au caractère intégré des mesures et au modèle de propagation dont l'effet est de relier les mesures effectuées aux dates successives, la quantité totale d'eau tombée durant un évènement est bien reproduite.

- La cellule la plus au nord (correspondant au premier pic de la Figure IV.8, ligne 4) est correctement reproduite (Figure IV.11.1), tandis que la cellule la plus au sud est surestimée

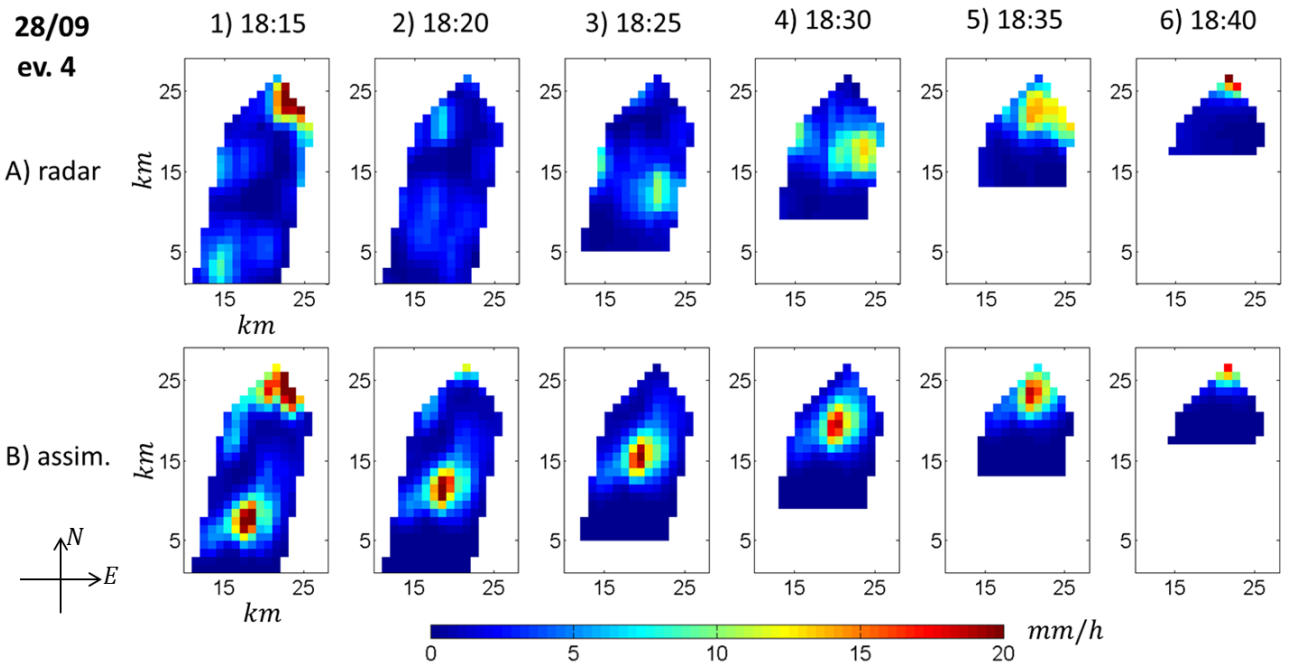

FIGURE IV.11 – Evènement 4 (28/09/2013). Champs de pluie à résolution $1km^2$ tel que restitués : A) par le radar Météo France de Bollène. B) par notre algorithme d'assimilation des mesures d'atténuation du capteur Ku. On restreint les champs à la grande zone d'assimilation définie par la Figure IV.4. Cartes toutes les 5 minutes sur tout l'évènement.

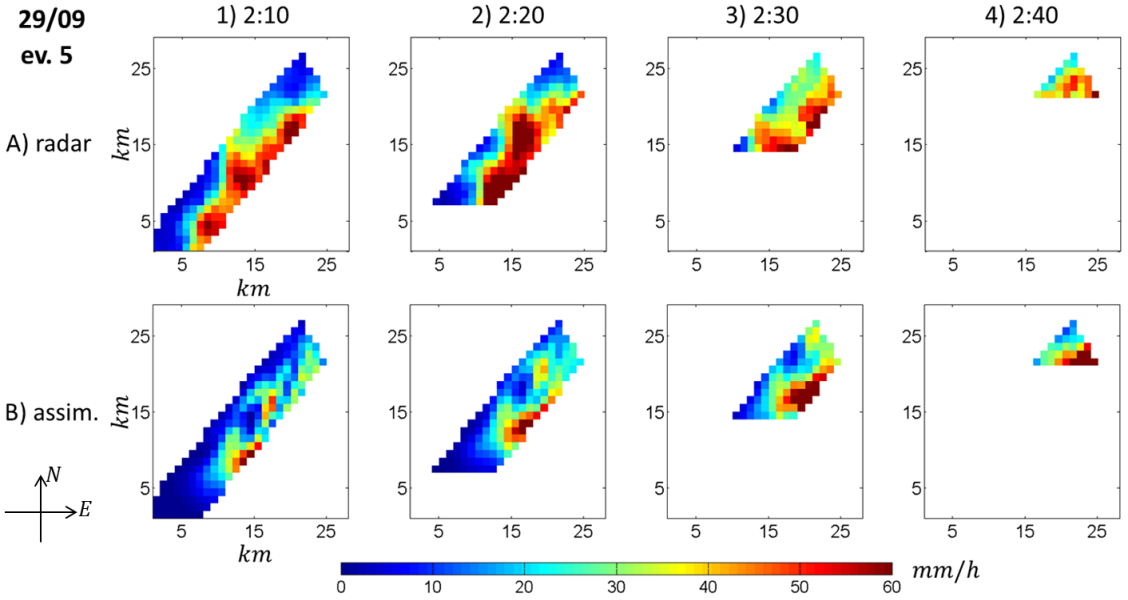

FIGURE IV.12 – Idem Figure IV.11 pour l'évènement 5 (29/09/2013). Cartes toutes les 10 minutes sur tout l'évènement.

sur la carte de pluie initiale (Figure IV.11.1). C'est une conséquence de l'erreur modèle. On suppose en effet que la cellule évolue sans déformations et sans variations de la quantité totale d'eau. Or on a vu qu'ici l'intensité de la cellule augmentait lors de sa progression. En conséquence, l'algorithme d'assimilation surestime la quantité d'eau à l'instant initial de manière à ce qu'advectée au-dessus des liaisons Ku, elle provoque une atténuation

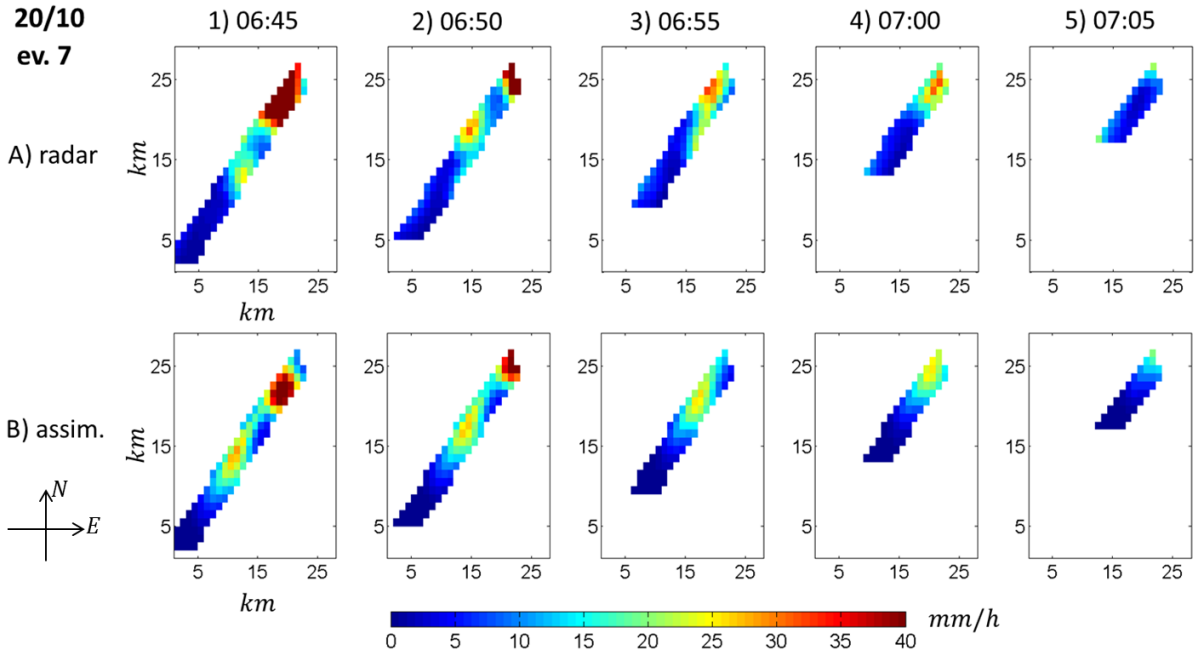

FIGURE IV.13 – Idem Figure IV.11 pour l'évènement 7 (20/10/2013). Cartes toutes les 5 minutes sur tout l'évènement.

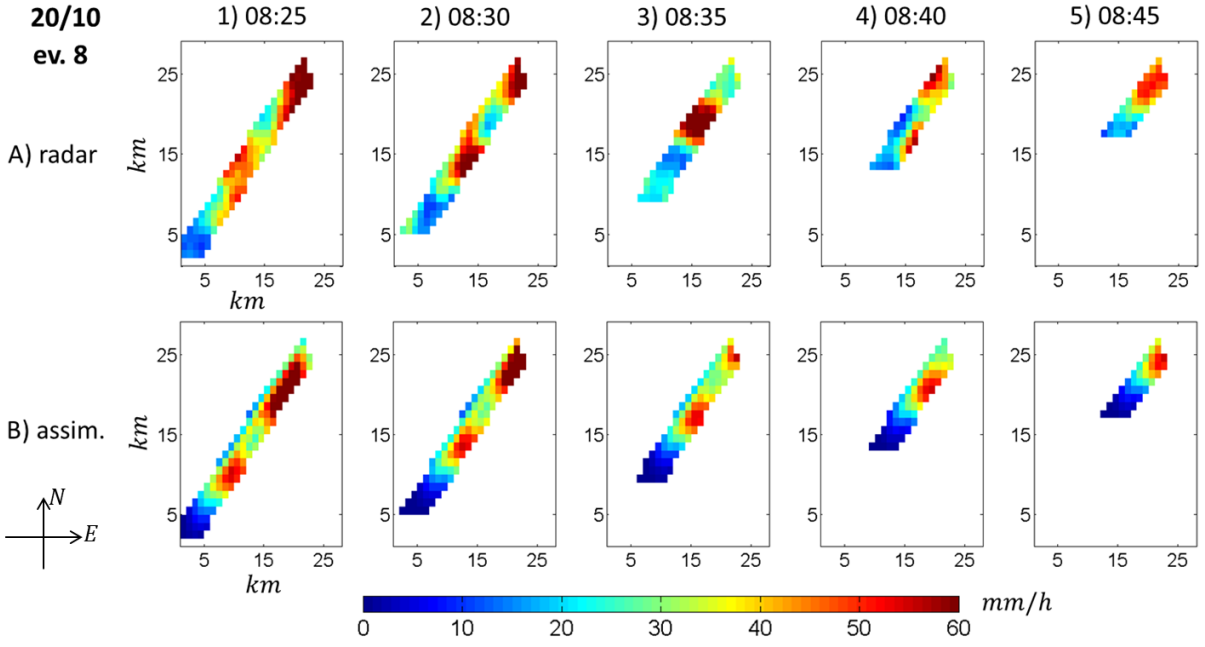

FIGURE IV.14 – Idem Figure IV.11 pour l'évènement 8 (20/10/2013). Cartes toutes les 5 minutes sur tout l'évènement.

correspondant aux mesures. Cette erreur modèle peut être vue sur cet évènement en comparant le coefficient de corrélation dans la zone directement couverte par les liaisons (0.85) avec celui dans la grande zone d'assimilation (seulement 0.48).

Les trois autres évènements étudiés confirment le premier point. La quantité totale d'eau

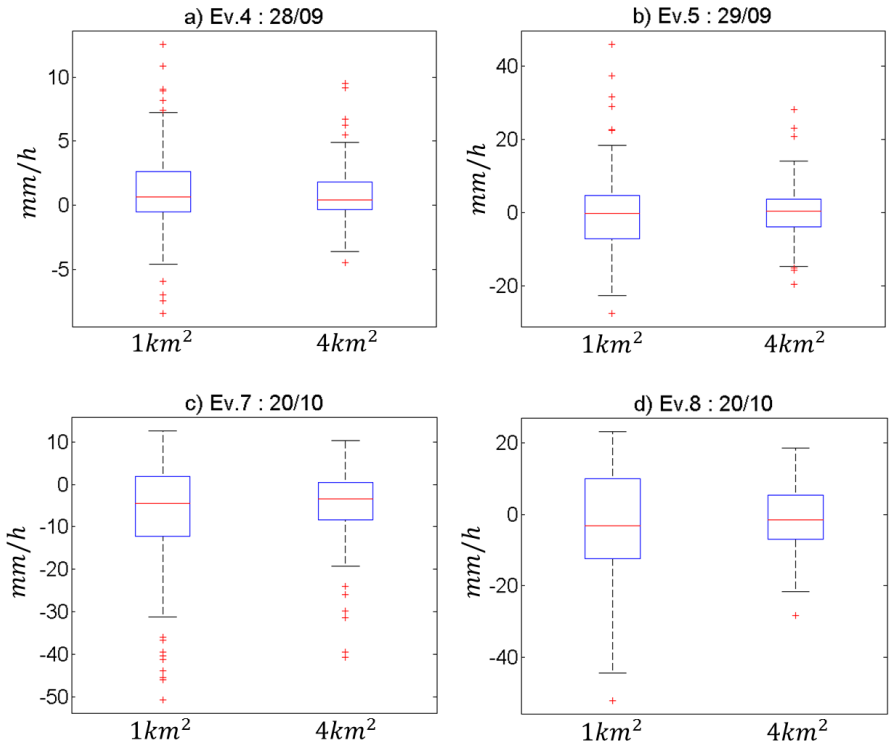

FIGURE IV.15 – Boîtes à moustache des différences entre cartes radar et assimilées pour les 4 évènements advectifs à 1 ou $4km^2$ de résolution. On représente dans chaque cas la médiane (traits rouges), les quartiles (boîtes bleues) et les déciles (traits noirs).

restituée par l'algorithme est compatible avec celle vue par le radar (voir Tableau IV.9 et Figure IV.15), que ce soit dans la petite ou dans la grande zone d'assimilation.

Concernant la localisation géographique précise des cellules à l'instant initial et durant l'évènement (Figures IV.12, IV.13 et IV.14), les meilleurs résultats sont obtenus pour les deux évènements du 20/10 (évènements 7 et 8), où les cellules sont relativement bien placées dès la carte initiale et leurs intensités correctement évaluées. C'est d'ailleurs pour ces deux évènements que l'on obtient les meilleurs coefficients de corrélation avec le radar dans la grande zone d'assimilation (Tableau IV.9). Quoique relativement bien localisée, la cellule de pluie initiale pour l'évènement 5 (29/09, Figure IV.12) est assez sensiblement sous-estimée (on a d'ailleurs un biais de -20% dans la grande zone d'assimilation). Il semblerait que bien qu'à plus grande échelle on ait de bonnes caractéristiques advectives pour cet évènement (Tabelau IV.8), l'intensité des cellules tende à diminuer en arrivant au niveau des liaisons Ku.

La tendance déjà mentionnée de l'algorithme à lisser les champs de pluie est assez claire, que ce soit comme on l'a déjà vu sur les reproductions des séries temporelles d'atténuation (Figure IV.8 et IV.9) ou lors du calcul de la droite de régression entre cartes radar et assimilées. Pour ce dernier critère, c'est vrai en particulier pour les évènements 5 (29/09) et 7 (20/10), avec des pentes de 0.44 et 0.40 dans la petite zone d'assimilation à $1km^2$ de résolution.

On peut aussi noter que les comparaisons point à point entre cartes radar et assimilées à la plus fine résolution ($1km^2$) donnent de bons résultats dans la petite zone d'assimilation

pour trois évènements (4, 5 et 7, avec des coefficients de corrélation respectifs de 0.85, 0.77 et 0.81), mais échouent pour l'évènement 8, avec un coefficient de corrélation de 0.39 et une ordonnée à l'origine de $26mm/h$, une valeur particulièrement forte. On voit très bien cet échec également sur la boîte à moustaches correspondant sur la Figure IV.15, qui montre des quartiles de l'écart radar-assimilation très élevés (autour de $10mm/h$). Visuellement (Figure IV.14), les résultats semblent pourtant assez bons. Lorsqu'on dégrade la résolution des champs à $2 * 2km^2$, les coefficients de corrélation augmentent sensiblement pour tous les évènements, à 0.92, 0.84 et 0.88 pour les évènements 4, 5 et 7, et surtout de 0.39 à 0.82 pour l'évènement 8 (Tableau IV.9), et les largeurs des boîtes à moustaches (quartiles des erreurs) diminuent sensiblement (Figure IV.15). Cela signifie que même si les champs assimilés et radar peuvent être très différents à la plus fine résolution (parfois à cause de légers décalages spatiaux), l'algorithme semble donner de bien meilleurs résultats dès lors que la résolution est un peu dégradée. Néanmoins, la petite taille de la zone considérée comme le petit nombre d'évènements traités font qu'il faudrait plus d'expériences pour appuyer statistiquement ces conclusions.

Nous concluons en remarquant que les comparaisons point à point sont bien entendu nettement moins satisfaisantes dans la grande zone d'assimilation (à cause des erreurs modèle) mais que l'on peut noter que :

- la zone couverte de manière satisfaisante peut être largement plus grande si le modèle d'advection est réaliste sur une plus grande durée. Par exemple, les résultats sont très bons sur les deux zones d'assimilation pour les évènements du 20/10 (évènements 7 et 8). La région considérée ici (Ardèche), montagneuse et sujette aux pluies orographiques, est particulièrement défavorable pour un modèle d'advection pure. On peut s'attendre à obtenir de meilleurs résultats en plaine.

- les performances comme la surface couverte pourraient aussi être améliorées en augmentant le nombre de capteurs utilisés.

6.5 Résultats / Restitutions ponctuelles : comparaison aux pluviomètres

Nous allons à présent estimer la capacité de l'algorithme à restituer des pluviométries ponctuelles. Nous comparons donc ses performances aux mesures des pluviomètres HPiconet. Afin de comparer ses performances dans ce domaine à celle d'un radar, nous effectuons systématiquement les mêmes comparaisons entre pluviomètres et cartes radar. Dans tous les cas, nous considérons que la valeur correspondant à un pluviomètre donné (soit pour les cartes radar, soit pour celles assimilées) est celle restituée pour le pixel dans lequel est situé le pluviomètre. En particulier, nous n'interpolons pas avec les valeurs des pixels voisins si le pluviomètre se situe au bord d'un pixel. Les écarts obtenus dans tous les cas (radar ou assimilation) et la taille des pixels ($1km^2$ pour le radar, $500 * 500m^2$ pour l'assimilation) comparés à la surface sondée par un pluviomètre, justifient cette façon de procéder.

La Figure IV.16 présente, pour chacun des 4 évènements advectifs, les cumuls d'eau enregistrés par le pluviomètre de Mirabel Pradel fermes (situé dans la petite zone d'assimilation, voir Tableau IV.2) (bleu) et reproduits par le radar (vert) et l'algorithme d'assimilation (rouge) au même point. La Figure IV.17 présente les mêmes résultats pour les pluviomètres de St Germain école (traits continus) et de Villeneuve de Berg (pointillés) (deux pluviomètres situés dans la grande zone d'assimilation, voir Tableau IV.2). Le Tableau IV.10 mélange les données des 7 pluviomètres du réseau situés dans la grande zone d'assimilation (on exclut les pluviomètres 1, Lussas ; 2, Mirabel Blaches et 4, Mirabel Mt Redon) et des 4 évènements advectifs. Il présente alors, pour différentes résolutions (5 à 25 minutes), la droite de régression linéaire et le coefficient de corrélation correspondant entre les données pluviomètres et les sorties de l'algorithme d'assimilation (colonne 3) et entre les données pluviomètres et les cartes radar (colonne 4). Il présente également le biais global (toujours sur les 4 évènements) et le taux de pluie moyen enregistré par les pluviomètres.

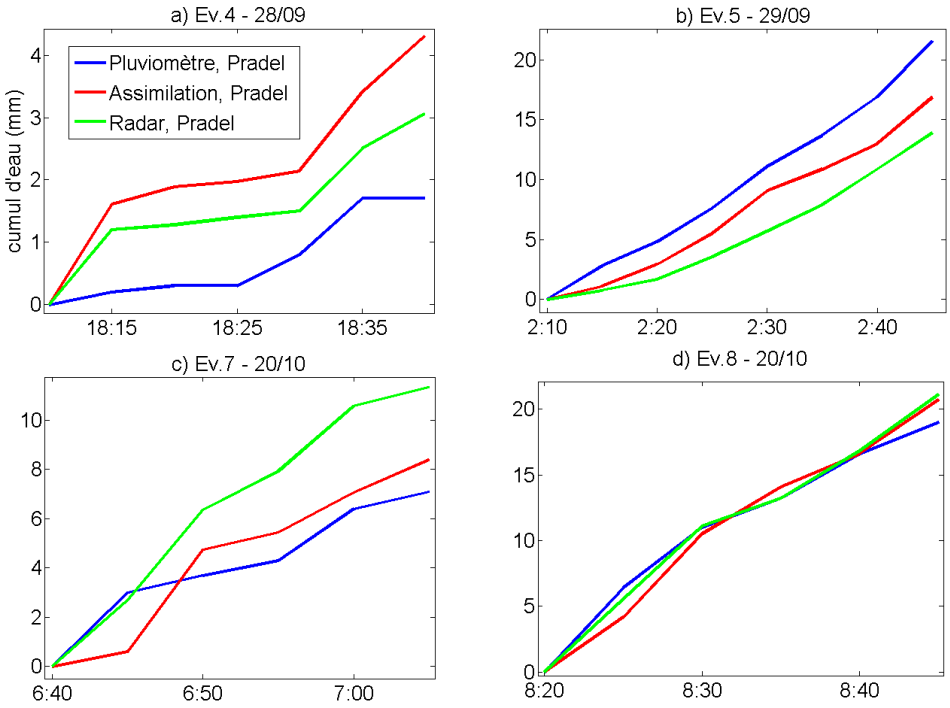

FIGURE IV.16 – Cumuls d'eau pour les 4 évènements advectifs considérés pour le pluviomètre de Mirabel Pradel fermes (pluviomètre 5 du Tableau IV.2). Pour chaque évènement, on présente les mesures effectuées par le pluviomètre (bleu), et les données restituées sur le pixel lui correspondant par le radar de Bollène (vert) et par l'algorithme d'assimilation (rouge).

Sur la Figure IV.16, représentant les cumuls d'eau pour un pluviomètre situé dans la petite zone d'assimilation, directement couverte par les liaisons Ku et donc moins soumise aux erreurs modèle, on note que l'algorithme fournit des résultats comparables à ceux du radar, avec des cumuls d'eau restitués sur l'évènement plus proches des mesures du pluviomètre pour 3 évènements (5, 7, 8) et plus éloignés pour 1 évènement (4). Sur la Figure IV.17, représentant les cumuls d'eau pour 2 pluviomètres situés dans la grande zone d'assimilation, plus sujette aux erreurs modèle, les résultats sont un peu moins bons, avec des

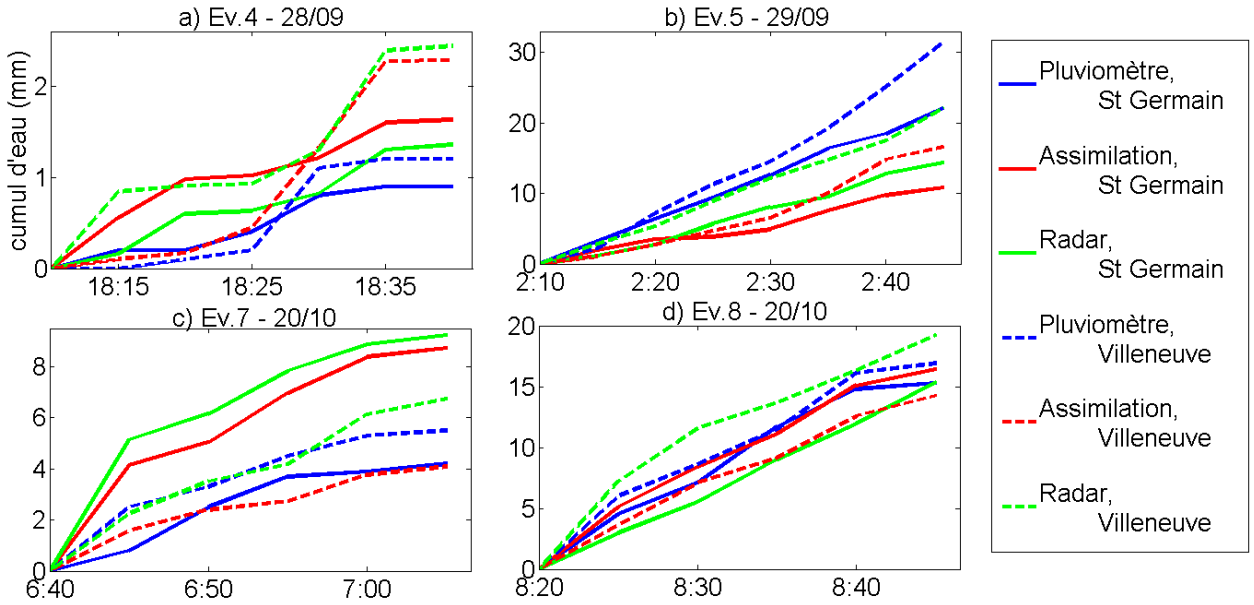

FIGURE IV.17 – Idem Figure IV.16 pour les pluviomètres de St Germain école (traits pleins) et de Villeneuve de Berg (pointillés) (respectivement pluviomètres 8 et 10 du Tableau IV.2).

écarts plus importants entre pluviomètres et algorithme d'assimilation. Les résultats (cumul d'eau en fin d'évènement) sont dans ce cas comparables à ceux du radar pour 3 cas (évènement 7, les deux pluviomètres, évènement 4, le pluviomètre de Villeneuve), meilleurs pour 2 cas (évènement 5, les deux pluviomètres) et moins bons pour 3 cas (évènement 8, les deux pluviomètres et évènement 4, le pluviomètre de St Germain).

On note sur les résultats du Tableau IV.10 que l'algorithme d'assimilation sous-estime légèrement le taux de pluie moyen (biais moyen de $-3.8mm/h$, à comparer au taux de pluie moyen de $24.5mm/h$). On remarque que le radar sous-estime également les taux de pluie aux points où sont placés les pluviomètres, mais dans une moindre mesure ($-1.9mm/h$). Les raisons conduisant à cette sous-estimation sont difficiles à exhiber (avec seulement 4 évènements), mais ces résultats sont dans tous les cas cohérents avec ceux du Tableau IV.9, qui montraient que l'algorithme sous-estimait légèrement les taux de pluie moyens par rapport à ceux mesurés par le radar pour 3 des 4 évènements considérés sur la petite zone d'assimilation (biais positif seulement pour l'évènement 4), et pour tous les évènements sur la grande zone d'assimilation.

Pour un temps d'intégration de 5 minutes, le coefficient de corrélation entre pluviomètres et algorithme d'assimilation est assez faible (0.63, voir Tableau IV.10), et la pente de la droite de régression également (0.53). Ce dernier résultat confirme la tendance de notre algorithme à lisser les champs et à gommer les valeurs extrêmes, tendance d'autant plus prononcée ici que nous confrontons des données intégrées à des données ponctuelles. Nous recherchons ensuite un temps d'intégration pour lequel les résultats se correspondraient mieux. Les coefficients de corrélation entre assimilation et pluviomètres sont de 0.76 et 0.77, respectivement à 10 et 15

TABLE IV.10 – Estimation ponctuelle de la pluie. Coefficients de corrélation, équation des droites de régression à différentes résolutions temporelles (5 à 25min), biais moyen entre pluviomètres et sorties d'assimilation (colonne 3) et entre pluviomètres et cartes radar (colonne 4). Taux de pluie moyen enregistré par les pluviomètres. On mélange ici les données des 4 évènements advectifs (évènements 4, 5, 7, 8) et des 7 pluviomètres situés dans la grande zone d'assimilation.

		KU / pluvios	radar / pluvios
coefficient de corrélation	$5-min$	0.63	0.69
	$10-min$	0.76	0.79
	$15-min$	0.77	0.83
	$20-min$	0.78	0.83
	$25-min$	0.82	0.84
équation droite de régression linéaire	$5-min$	$0.53x+0.64$	$0.61x+0.64$
	$10-min$	$0.63x+0.79$	$0.69x+0.85$
	$15-min$	$0.62x+1.08$	$0.69x+1.16$
	$20-min$	$0.61x+1.52$	$0.68x+1.58$
	$25-min$	$0.57x+1.58$	$0.64x+1.56$
biais (mm/h)		-3.8	-1.9
tx pluie moyen pluvios (mm/h)		24.5	

minutes, tandis que les pentes de la droite de régression sont de 0.63 et 0.62. L'algorithme semble donc produire de bien meilleurs résultats ponctuels pour un temps d'intégration supérieur ou égal à 10 minutes. Pour des temps d'intégration plus longs, le coefficient de corrélation n'augmente plus significativement (0.82 à 25 minutes). On note ici que les coefficients de corrélation entre pluviomètres et radars sont du même ordre de grandeur, allant de 0.69 à 5min à 0.79 à 10min et 0.84 à 25min. Bien qu'il faille toujours rester prudents du fait du petit nombre d'évènements considérés, il semblerait que l'ordre de grandeur des résultats obtenus ici (coefficients de corrélation plafonnant vers 0.8-0.85 et pentes des droites de régression sensiblement inférieures à 1) vienne du fait que l'on compare des données de natures différentes, intégrées (radar et assimilation) et ponctuelles (pluviomètres). Le fait que les pentes soient légèrement plus faibles dans le cas de l'algorithme d'assimilation pourrait s'expliquer par le lissage provoqué par le modèle de propagation, et les coefficients de corrélation légèrement plus faibles par l'ajout d'erreurs modèles.

7 Conclusion et perspectives

Conclusions

Nous avons présenté ici une nouvelle technique permettant de reconstruire des champs de pluie à partir de mesures d'atténuations le long de liaisons microondes Terre-satellites TV en bande Ku. Cette technique utilise un algorithme d'assimilation 4D-VAR, qui suppose que la dynamique des cellules de pluie est gouvernée par une équation d'advection pure (déplacement sans déformations). Sous réserve que l'évènement de pluie considéré soit effectivement "raisonnablement" advectif (sous des critères définis précisément), cet algorithme restitue des cartes de pluie à résolution $1s$ et $0.5 * 0.5km^2$, avec des caractéristiques comparables à celles fournies par un radar à résolution $5min$ et $1km^2$. Les caractéristiques de l'advection (vitesse de déplacement) sont déduites de l'analyse de cartes radar successives lorsque, comme ici, un seul capteur Ku est disponible. Elles pourraient être déduites directement des mesures d'atténuations avec au moins 2 capteurs (voir Annexe G).

Plus en détails, nous avons d'abord évalué, sur données simulées, la faisabilité de notre méthode et sa robustesse. Nous avons noté la plus grande difficulté de l'algorithme à localiser les cellules selon l'axe nord-sud, due à la géométrie des liaisons (qui pointent vers le sud). Nous avons ensuite travaillé avec des données réelles enregistrées dans le sud de la France en 2013. En comparant les restitutions avec des cartes radar et des mesures de pluviomètres, nous avons montré que l'algorithme fournissait de bons résultats dans la zone directement couverte par les liaisons Ku, avec en particulier une très bonne restitution de la quantité totale d'eau tombée. Cependant, nous avons noté que les performances de l'algorithme, notamment dans la zone couverte seulement indirectement via le modèle d'advection, dépendaient fortement de la faculté du modèle en question à restituer la dynamique réelle des champs. Nous avons ainsi introduit un critère (sur la corrélation entre cartes radars successives) limitant l'étude aux évènements "suffisamment advectifs" (4 évènements sur les 8 enregistrés). Même si les performances de l'algorithme, notamment sur la restitution de la localisation et de la géométrie des cellules, étaient bien meilleures pour ces évènements, nous avons vu que les erreurs modèles restaient la principale source d'erreur, limitant notamment la durée des évènements traités à $30 - 45min$.

Perspectives

Nous avons ensuite noté que la région d'étude (Cévennes-Vivarais, région montagneuse) ne facilitait pas le travail. La zone est en effet sujette à des pluies orographiques, limitant davantage l'utilisation d'un modèle d'advection pure. Cependant, les performances de l'algorithme pourraient être sensiblement améliorées par l'utilisation de plusieurs capteurs Ku. Cela permettrait en effet :

- d'élargir la zone couverte (que ce soit directement ou via le modèle de propagation).

- de déterminer les caractéristiques d'advection directement via la méthode de triangulation présentée en Annexe G, et donc de se passer complètement de cartes radar.

- éventuellement de complexifier le modèle de propagation (par exemple en ajoutant un coefficient faisant varier l'intensité des cellules lors du déplacement).

L'algorithme a commencé à être adapté dans ce sens. Akrour (2015) et Akrour et al. (2015) ont, sur données simulées uniquement, étudié un cas d'application à l'hydrologie urbaine. Ils se placent dans une ville et supposent qu'un nombre significatif de capteurs Ku est disponible (quelques dizaines sur une zone de $30 * 30km^2$, en testant plusieurs configurations avec différentes géométrie et différents nombres de liaisons), et montrent ainsi que l'algorithme peut être performant sur une zone significativement agrandie.

Une autre extension possible de ce travail serait d'adapter l'algorithme d'assimilation pour fusionner des mesures de différentes natures, pluviomètres et capteurs d'atténuation en particulier. En effet, les mesures ponctuelles des pluviomètres pourraient permettre de lever certaines ambiguïtés venant des caractères intégré et non isotrope des liaisons Ku (mauvaise discrimination des cellules selon l'axe nord-sud ou difficulté à reproduire les valeurs extrêmes). Des études de faisabilité à ce sujet sont en cours.

Au niveau de l'algorithme en lui-même, outre le perfectionnement du modèle mentionné plus haut, deux autres améliorations semblent envisageables à plus ou moins long terme. Tout d'abord l'amélioration du terme de filtrage. Celui-ci est pour l'instant un opérateur mathématique simple visant à lisser les champs de pluie. Ses inconvénients sont qu'il est nul (dans le sens où le coût qu'il ajoute est nul) pour des champs de pluie homogènes, et qu'il apporte un coût proportionnel au carré de l'intensité des champs (une atténuation moyenne deux fois plus faible donne un coût quatre fois plus faible). Il pourrait être plus judicieux, quoique plus complexe mathématiquement, de le remplacer par un ou des opérateurs prenant en compte la physique des champs de pluie ou leurs caractéristiques statistiques. On pourrait par exemple imaginer évaluer au sein de l'algorithme une fonction d'autocorrélation du champ de pluie et la comparer à une fonction d'autocorrélation prédéfinie. Ensuite, nous supposons que les gouttes tombent verticalement (les liaisons Ku sont projetées orthogonalement au sol et le problème est restreint à un plan). Une amélioration possible serait de prendre en compte le déplacement horizontal des gouttes durant leur chute dû au vent horizontal, c'est-à-dire à la vitesse d'advection (déterminée par ailleurs).

Nous concluons en remarquant que la méthode décrite ici pourrait être une alternative crédible et peu coûteuse aux radars pour l'observation de zones non couvertes (en particulier régions montagneuses ou pays en développement). Les mesures intégrées et l'algorithme d'assimilation permettant de bien restituer le total de précipitations, cette méthode semble bien adaptée pour l'hydrologie, et pourrait également être utilisée dans les zones urbaines, où un grand nombre de mesures seraient potentiellement disponibles en ajoutant un petit système de réception dédié aux antennes de télévision par satellites personnelles existantes.

Partie V

Restitution et étude des profils verticaux de DSD et de vents verticaux

1 Présentation

1.1 Objectifs généraux / Approche VAMOS

L'étude du chapitre IV sur la restitution de cartes de précipitations à partir d'atténuations d'ondes en bande Ku se situait à une échelle macrophysique, dans le sens où l'on modélisait le déplacement des cellules de pluie indépendamment des phénomènes physiques affectant les gouttes d'eau. Par ailleurs, la restitution de cartes de taux précipitants était basée sur l'utilisation d'une loi puissance liant atténuation et taux précipitant. Or les coefficients de cette loi puissance (en particulier en bande Ku et au-delà) dépendent de la distribution en taille des gouttes de pluie (DSD). De même, toujours dans le domaine de la télédétection, les paramètres de la relation $Z - R$, permettant de convertir la réflectivité radar en taux de pluie, dépendent de la DSD. Ces paramètres sont généralement basés sur une distribution particulière et supposés constant sur la colonne atmosphérique. Or en pratique, les taux de pluie estimés à partir de réflectivités par cette relation $Z - R$ peuvent varier d'un facteur 2 selon la DSD (List (1988)). De plus, les radars opérationnels (tout comme les systèmes comme celui du chapitre IV) mesurant la pluie à différentes altitudes, il apparaît crucial d'étudier la variabilité verticale des DSD. Or cette étude doit nécessairement se faire à l'échelle microphysique, c'est-à-dire à l'échelle de la goutte de pluie et des phénomènes physiques l'affectant durant sa chute.

On a vu au chapitre II que ces phénomènes étaient nombreux (vent $3D$, évaporation, collisions entre gouttes) et que leurs effets sur les DSD s'entremêlaient, si bien qu'il était difficile de déterminer leur influence relative. Certaines études sur des évènements spécifiques montrent l'influence que peuvent avoir des phénomènes particuliers. Par exemple Kollias et al. (2001) montrent que les courants ascendants peuvent causer un déficit de grosses gouttes ($> 3mm$) dans le cœur de l'ascendance et un surplus dans la périphérie. VanZanten et al. (2005) montrent que sous l'influence de l'évaporation, la bruine peut ne pas atteindre le sol sous des conditions météorologiques spécifiques. L'influence des collisions entre gouttes a elle fait l'objet d'une littérature abondante (voir Annexe A) et est aujourd'hui encore débattue (Villermaux and Bossa (2009), Barros et al. (2010), voir Annexe B). L'influence de ces phénomènes sur les DSD peut être représentée via des modèles de propagation d'équations aux dérivées partielles (EDP, voir chapitre II) mais ces modèles nécessitent une paramétrisation complexe, et, bâtis à partir de simulations numériques ou d'expériences en laboratoire, ils souffrent généralement d'un manque de validation dans des conditions réelles (Mcfarquhar (2010)).

On a vu également au chapitre I que différents appareils de mesure permettaient d'obtenir, directement ou indirectement, des informations sur la DSD (disdromètres, radars Doppler à visée verticale). Mais la possibilité d'utiliser directement ces observations pour l'étude des profils de DSD doit être nuancée. Les disdromètres effectuent une mesure directe d'un flux de

gouttes, mais cette mesure n'est réalisée qu'au sol (on n'obtient pas d'informations sur les profils verticaux des distributions) et est très ponctuelle (surfaces de collecte de $5 - 10cm^2$). Les radars effectuent des mesures intégrées sur des volumes plus conséquents et à différentes altitudes, mais ces mesures sont indirectes et affectées par différents phénomènes (vent vertical, atténuation, turbulence). Les deux types d'observations sont donc difficilement comparables directement, et difficilement utilisables directement pour la validation d'un modèle de propagation.

L'idée principale de la méthode développée ici et baptisée VAMOS (Variational Assimilation of Multi-scale Observations for rain Studies) consiste à fusionner les observations disponibles via un algorithme d'assimilation de données 4D-VAR (voir la théorie chapire III) contraint par un modèle de propagation des DSD. L'objectif de VAMOS est de restituer des profils verticaux de DSD et de paramètres atmosphériques associés comme le vent vertical basés sur toutes les observations disponibles (spectres Doppler et disdromètres) en tirant le meilleur parti de chacune d'elles, et cohérentes avec la physique incluse dans le modèle de propagation.

Une telle démarche exige de développer le modèle en fonction des observations disponibles. Utiliser un modèle trop complexe, nécessitant la paramétrisation d'un grand nombre de phénomènes et possédant un grand nombre de degrés de liberté alors que l'on disposerait de peu d'observations serait utopique et inutile : plusieurs jeux DSD / paramètres du modèle seraient susceptibles d'expliquer de manière satisfaisante les observations. Au contraire, un modèle trop simpliste ne parviendrait pas à expliquer la variabilité des observations disponibles.

Il convient donc d'étudier systématiquement la capacité du modèle (et donc de la physique qu'il inclut) à expliquer ou non les observations et à le complexifier en fonction de la quantité et de l'indépendance des observations disponibles.

Une telle approche devrait permettre, pour chaque évènement pluvieux de rendre comparables (de fusionner) les observations disponibles, de restituer les profils verticaux de DSD et de vents verticaux et d'étudier la prépondérance du rôle des différents phénomènes physiques impliqués. Cela se fera via l'étude de la capacité de la physique du modèle de propagation à expliquer les observations et de l'étude - complémentaire - de la variabilité non expliquée des observations.

Les objectifs à plus long terme (au-delà du cadre de cette thèse) de ce travail peuvent être de proposer des paramétrisations des phénomènes physiques impliqués basées directement sur les différentes observations disponibles, ou encore la simulation d'observations d'autres natures ou à d'autres échelles (satellite par exemple) pour calibration ou intercomparaison.

En conclusion, on note que dans le travail présenté ici, les objectifs seront principalement :

- la fusion de données hétérogènes de mesure des DSD, données provenant d'instruments (radars et disdromètres) très différents de par la nature, la localisation et l'échelle spatiotemporelle de leurs mesures, et donc a priori difficilement comparables.

- l'étude des profils verticaux de DSD et de paramètres atmosphériques liés (vent, turbulence atmosphérique, humidité relative) restitués via VAMOS, un algorithme d'assimilation de données 4D-VAR intégrant les observations disponibles dans un modèle de propagation.

- l'étude des phénomènes physiques prépondérants dans chaque épisode pluvieux étudié.

1.2 Autres études sur l'inversion des spectres Doppler

L'approche VAMOS proposée ici pour restituer des profils de DSD, par le fait qu'elle lie les observations réalisées à différentes dates et échelles par un modèle de propagation dans un algorithme d'assimilation de données, est nouvelle. Néanmoins, la question de l'inversion des spectres Doppler a déjà été largement traitée dans la littérature. Nous présentons ici quelques unes de ces études et leurs principes théoriques. On note bien néanmoins que nos objectifs vont au-delà de l'inversion de spectres Doppler, puisqu'il s'agit principalement de fusionner des observations hétérogènes et d'étudier la dynamique des précipitations à échelle microphysique. La restitution du vent vertical à fine échelle, en particulier, passe néanmoins par l'utilisation d'informations contenues dans les spectres Doppler et notre travail à ce niveau est bien entendu fortement lié à celui réalisé dans ces différentes études.

Généralement, les travaux visant à inverser les spectres Doppler travaillent spectre par spectre (en ne prenant pas en compte leur corrélation). Les premiers travaux sur ce sujet (Rogers (1964), Hauser and Amayenc (1981)) utilisaient des spectres Doppler à longueurs d'ondes assez élevées (bande C ou X, autour de $10 - 15GHz$) et paramétraient les DSD sous forme exponentielle. En ajoutant le vent vertical, dont l'influence sur les spectres Doppler est, comme on l'a vu, prépondérante, ils obtenaient 2 ou 3 inconnues (1 ou 2 paramètres d'une loi exponentielle selon si l'un d'entre eux est supposé fixe ou non et vent vertical) à déterminer à partir de chaque spectre. La non prise en compte de la turbulence (qui lisse les spectres, Gossard (1994)) et l'ambiguïté entre influence sur les spectres du vent et de la forme des DSD complique néanmoins l'application de telles techniques.

D'autres auteurs ont alors proposé l'utilisation de profileurs de vent (Gossard (1988), Williams (2002)). Radars à basse fréquence (UHF ou VHF), leur particularité est d'être affecté non seulement par la diffusion Rayleigh due aux gouttes de pluie mais aussi par la diffusion de Bragg due aux variations de la vitesse de l'air clair. L'idée générale est alors de déterminer la vitesse verticale de l'air grâce au pic de diffusion induit par la diffusion de Bragg. On corrige ensuite l'effet de ce vent vertical sur la partie du spectre due à la pluie (diffusion Rayleigh) pour déterminer la DSD. Cette technique fonctionne assez bien sur des pluies faibles et avec peu de turbulence mais peut être plus délicate à mettre en oeuvre si les deux pics de diffusion sont mêlés (Williams (2002)). Williams (2012) propose de remédier à ce problème en utilisant deux radars, un en UHF (sensible aux diffusions de Bragg et de Rayleigh) et un en VHF (principalement sensible à la diffusion Rayleigh). Kollias et al. (2002)

note néanmoins diverses limitations pour cette méthode : la taille des profileurs les rend peu mobiles, leur résolution (temporelle, verticale et en vitesses Doppler) rend moins précises ces restitutions, leur ouverture de faisceau implique des volumes sondés importants, augmentant l'élargissement spectral, et les premières mesures disponibles en UHF se situent généralement assez haut (typiquement $1.5km$ d'altitude, voir Williams (2012)).

Parallèlement ont été développées des méthodes basées sur l'utilisation de spectres Doppler à haute fréquence, typiquement en bande W, dont la particularité est d'être sensibles à la diffusion de Mie (Lhermitte (1988), Kollias et al. (2002), Giangrande et al. (2010), Giangrande et al. (2012)). Les radars à ces fréquences sont généralement plus petits et possèdent des résolutions plus fines. Ces méthodes sont basées sur la résonance existant dans la courbe donnant selon la théorie de Mie la relation entre surfaces efficaces de rétrodiffusion et diamètre des gouttes (voir par exemple Fig.4 de Kollias et al. (2007) ou Figure I.8 au chapitre I de ce manuscrit). Ce minimum de section efficace de rétrodiffusion se situe par exemple vers $5mm$ à $35GHz$ et vers $1.7mm$ à $95GHz$. Il se traduit par un minimum sur les spectres Doppler localisé à vitesse Doppler théorique fixe (sous réserve que la DSD soit suffisamment régulière et que les gouttes tombent, indépendamment de l'effet du vent, à leur vitesse terminale de chute). Ces techniques consistent alors à rechercher, pour chaque spectre Doppler, le minimum d'énergie effectivement mesuré, et à supposer que la différence avec la valeur théorique correspond à la vitesse verticale du vent. On inverse ensuite la DSD à partir du spectre corrigé de cette vitesse du vent. La turbulence atmosphérique et le petit nombre de grosses gouttes rendent généralement invisible ce minimum en bande Ka mais ces méthodes ont donné de bons résultats en bande W (Giangrande et al. (2012) par exemple). Elles souffrent néanmoins encore de certains désavantages. Le premier est leur limitation à la bande W, sujette à de fortes atténuations, limitant la méthode à des pluies faibles. Le second est la difficulté qu'il y a à partir des données d'un seul radar à s'affranchir de problèmes de calibration et d'estimation de l'atténuation par la pluie ou par le radôme. Il convient donc généralement soit de se limiter à la détermination de la forme de la DSD, soit d'ajouter d'autres informations (typiquement provenant d'un disdromètre au sol).

Récemment, une autre technique a été proposée combinant des spectres Doppler en bande Ka et W (utilisation des spectres Doppler différentiels) : Tridon et al. (2013), Tridon and Battaglia (2015). Cette méthode propose ainsi de s'affranchir des problèmes de calibration des radars et de restituer DSD, vents verticaux et turbulence à partir des spectres différentiels. On note qu'elle est cependant toujours limitée par la forte atténuation en bande W.

Nous concluons en notant que la méthode de restitution des profils verticaux de DSD que nous présentons ici (VAMOS) peut a priori être appliquée aux différentes fréquences mentionnées, tout en subissant bien entendu les limitations mentionnées pour chacune d'elles. On note néanmoins d'ors et déjà que l'on compte sur la cohérence physique apportée par le

modèle de propagation (une goutte enregistrée à une altitude à une date donnée le sera à nouveau un peu plus bas un peu plus tard) pour pouvoir obtenir des informations sur le vent vertical ou la turbulence à partir de moins de données que ces méthodes (en particulier en utilisant seulement un radar en bande Ka, sans bande W).

1.3 Plan

Dans ce chapitre, nous commencerons (section 2) par présenter VAMOS, notre algorithme de restitution des DSD et vents verticaux en détails, à la lumière des résultats théoriques présentés précédemment sur les observations disdromètres et radar (chapitre I), sur les modèles de propagation (chapitre II) et sur l'assimilation de données (chapitre III). Ensuite, dans la section 3, nous évaluerons les performances de notre algorithme sur des données simulées, afin de mettre en évidence sa réaction à différentes sources d'erreurs, et notamment sa gestion de la turbulence atmosphérique. La section 4 reviendra en détails sur la question de l'influence du vent horizontal sur nos résultats. Celui-ci n'est en effet pas pris en compte dans l'algorithme de restitution (on suppose que les gouttes tombent verticalement et le problème est réduit à une colonne atmosphérique). Dans cette section, nous présenterons une petite étude théorique permettant de simuler la signature de différents profils de vents horizontaux sur les mesures radars. Puis nous testerons VAMOS en présence de vent en simulant des mesures disdromètres et spectres Doppler dans un modèle avec vent horizontal puis en les assimilant dans notre algorithme sans vent horizontal. Nous étudierons ensuite les résultats sur trois cas d'étude. Le premier, présenté en section 5, concerne un évènement de pluie stratiforme ayant eu lieu en Ardèche. Nous étudierons la capacité de notre algorithme à reproduire les observations et la structure des champs de DSD et de vents verticaux restitués, mais nous ne disposerons pas de données permettant une validation indépendante des résultats. Ce travail de validation sera fait sur le deuxième cas d'étude (section 6). Celui-ci concerne un évènement similaire ayant eu lieu en Finlande. Pour ce cas, enregistré par deux radars Doppler à visée verticale en bande Ka et W, nous comparerons nos résultats, produits uniquement avec les données des disdromètres et du radar en bande Ka, avec les restitutions fournies par l'algorithme développé dans Tridon and Battaglia (2015) et basée sur les données des deux radars. Une fois cette étape de validation accomplie, nous appliquerons notre algorithme à une troisième et dernier cas d'étude (section 7) plus complexe, enregistré à Darwin, au nord de l'Australie, une zone à climat tropical. Nous nous attarderons en particulier sur l'influence sur les résultats du vent horizontal, de l'évaporation et sur la gestion de la turbulence par notre algorithme.

2 Description de l'algorithme de restitution VAMOS complet

Dans cette section, nous allons décrire l'algorithme d'assimilation 4D-VAR utilisé dans ce chapitre (VAMOS). Nous présentons ici l'algorithme dans sa forme la plus extensive effectivement implémentée et testée dans le cadre de ce travail. Il sera parfois simplifié dans la suite (par exemple non prise en compte de l'évaporation dans les section 5 et 6).

On rappelle que la théorie de l'assimilation 4D-VAR a été décrite au chapitre III. Dans ce travail, on cherche à restituer des DSD en haut d'une colonne atmosphérique et un champ temps/altitude de vents verticaux. Le modèle de propagation utilisé propage ces DSD à toutes dates et altitudes. Les observations sont composées de mesures de flux de gouttes au sol par disdromètres (voir chapitre I, section 2.2) et de spectres Doppler (voir chapitre I, section 2.3.2).

Dans la section 2.1, nous détaillerons le modèle de propagation des DSD (EDP continue, équations discrétisées et paramètres et inconnues de la discrétisation). Dans la section 2.2, nous détaillerons le principe de la paramétrisation d'inconnues du problème (DSD en haut de la colonne atmosphérique) sous forme de lois gamma. Puis nous détaillerons dans la section 2.3 les opérateurs d'observation, qui, nous le verrons, sont aussi utilisés pour simuler les données des disdromètres et des radars Doppler. La section 2.4 détaillera les termes de régularisation ajoutés dans l'algorithme 4D-VAR. Nous conclurons avec la section 2.5.

Rappelons d'ores et déjà que la plupart des hypothèses et choix présentés dans cette section ont été justifiés (par exemple chapitre II, section 2.5.3 sur les collisions entre gouttes) ou le seront dans la suite (section 4 sur le vent horizontal ou section 3 pour une étude sur données simulées). L'objet de cette section n'est donc pas de justifier ces choix mais de décrire formellement l'algorithme utilisé.

2.1 Modèle de propagation

Dans cette section, nous allons présenter le modèle direct de propagation des DSD utilisé par l'algorithme d'assimilation 4D-VAR, ainsi que les inconnues de celui-ci, qui seront les inconnues (les paramètres de contrôle) de l'algorithme d'assimilation. Nous présentons d'abord l'EDP (continue) utilisée, puis le schéma de discrétisation choisi et détaillons les inconnues et paramètres d'un tel modèle.

Équation aux dérivées partielles

Nous rappelons que l'équation aux dérivées partielles (EDP), régissant l'évolution spatiotemporelle des DSD, dans sa forme la plus générale, a été fournie et justifiée au chapitre II, section 2.

Nous avons mentionné (section 1) que, dans tout ce travail, nous négligerons les effets du vent horizontal. La justification et les limites de cette hypothèse seront présentées dans la section 4. Sans vent horizontal, l'EDP n'induit aucune variabilité dans le plan (x, y) (les gouttes

tombent verticalement). Le problème se ramène donc à l'évolution temporelle d'une colonne atmosphérique, c'est-à-dire à une coordonnée d'espace, la hauteur au-dessus du sol, z. La DSD n est alors une fonction du diamètre D, de l'altitude z et du temps t :

$$n = n(t, z, D) \tag{V.1}$$

Nous avons par ailleurs vu dans le chapitre II, section 2.5.3, que nous négligerons dans ce travail la coalescence et le break-up (en nous limitant à des évènements de pluie ne présentant pas de taux de pluie supérieurs à $10mm/h$ pour limiter les erreurs commises par cette simplification). Sous ces conditions, l'EDP d'évolution des DSD (équation (II.5)) s'écrit (voir chapitre II section 2) :

$$\left(\frac{\partial n}{\partial t}\right) + \frac{\partial}{\partial z}\left(\left[v_{th} + w\right] n\right) = \frac{\partial}{\partial D}\left[\xi_m n \frac{dD}{dm}\right] \tag{V.2}$$

On rappelle que :

- v_{th} est la vitesse terminale de chute. C'est ici une fonction du diamètre et de l'altitude, selon la paramétrisation de Atlas et al. (1973) corrigée pour l'altitude par Foote and Du Toit (1969). On utilise l'équation (C.18) du chapitre II, section 3.

- w est la vitesse du vent vertical, fonction du temps et de l'altitude.

- ξ_m est le taux d'évaporation (kg/s), fonction du diamètre des gouttes et de paramètres météorologiques (température, pression, humidité relative). Paramétrisé selon Seifert (2008), voir les détails chapitre II section 2.4 (Eq. (II.19)).

- $\frac{dD}{dm}$ est la dérivée de la relation diamètre-masse pour des gouttes sphériques (voir Eq. (II.8)).

Discrétisation

La programmation effective de l'algorithme 4D-VAR nécessite bien sûr de discrétiser ce modèle sur un maillage du temps et de l'espace. L'espace (axes verticaux des z et des diamètres D) est maillé selon une grille C d'Arakawa (voir chapitre II section 4). Les pas de discrétisation sont les suivants :

$$\begin{cases} \Delta t = 5s \\ \Delta z = 100m \\ \Delta D = 0.1mm \end{cases} \tag{V.3}$$

On rappelle que NT, NZ et ND correspondent respectivement aux nombres de pas de temps, d'altitude et de diamètre. Les deux premiers dépendent de l'évènement considéré (durée et altitude du nuage ou de la bande brillante). ND lui est fixé à 74 pour une gamme de diamètres linéairement répartis entre les valeurs minimale de $0.2mm$ et maximale de $7.5mm$.

L'EDP (Eq. (V.2)) est discrétisée en altitude (terme en $\partial/\partial z$) selon le schéma de

Smolarkiewicz (chapitre II, section 4.3) et en diamètre (terme en $\partial/\partial D$) selon le schéma upstream (chapitre II, section 4.2). Les pas définis (cf. Eq.(V.3)) permettent d'assurer la stabilité des schémas pour des vitesses de chute allant jusqu'à $20m/s$ (soit un vent vertical d'environ $10m/s$ pour les plus grosses gouttes, a priori bien supérieur aux valeurs caractéristiques pour les évènements stratiformes traités ici) et pour tous les taux d'évaporation réalistes. Pour le schéma de Smolarkiewicz, on prendra $Sc = 1.04$ et $\epsilon = 10^{-4}$ (cf. chapitre II section 4 ou Smolarkiewicz (1983)).

Notations

Comme dans le chapitre IV, nous ne travaillerons désormais qu'avec des grandeurs discrétisées. Par conséquent, contrairement au chapitre II, nous omettrons les tildes sur les grandeurs discrètes. On rappelle que l'indice k renvoie au temps (i.e. à la date t_k), l'indice i renvoie à l'altitude (i.e. à la couche z_i) et l'indice p au diamètre (i.e. à la classe de diamètre D_p). La variable n_k fait, par exemple, référence à un champ dépendant de l'altitude et du diamètre (éventuellement réécrit sous forme vectorielle). Il s'agit du nombre de gouttes à la date t_k . Ainsi, $n_{k,i}$ renvoie à la DSD au point de grille correspondant à la date t_k et à l'altitude z_i.

Par ailleurs, les gouttes tombant (généralement) du haut de la colonne atmosphérique vers le sol, nous fixerons le zéro des hauteurs (l'altitude z_0) en haut de la colonne, c'est-à-dire à l'altitude maximale couverte par notre modèle (déterminée par l'altitude de la bande brillante). $n_{k,0}$ fait donc référence à la DSD en haut de la colonne (condition limite supérieure du modèle). Sur le même principe, $n_{k,NZ-1}$ fait référence à la DSD au niveau du sol. Pour finir, $n_{0,i}$ représente la condition initiale du modèle (DSD à t_0).

Les autres grandeurs de l'Eq. (V.2) sont également discrétisées. w est dans cette partie le champ des vent verticaux et $w_{k,i}$ le vent vertical à la date t_k et l'altitude z_i. $v_{th,i,p}$ la vitesse terminale des gouttes de diamètre D_p à l'altitude z_i, et $\xi_{m,k,i,p}$ le taux d'évaporation pour les gouttes de diamètre D_p à l'altitude z_i et à la date t_k.

Paramètres/Inconnues

Dans la configuration actuelle, pour utiliser le modèle précédemment défini, il faut fournir :

- une condition initiale $n_{0,i}$ $(\forall i)$
- une condition limite supérieure $n_{k,0}$ $(\forall k)$
- les valeurs des vitesses terminales de chute $v_{th,i,p}$ $(\forall i, p)$, taux d'évaporation $\xi_{m,k,i,p}$ $(\forall k, i, p)$ et vents verticaux $w_{k,i}$ $(\forall k, i)$.

On définit alors comme "paramètre du modèle" les grandeurs calculées indépendamment de l'algorithme d'assimilation, et qui lui sont fournies dans un fichier séparé. On définit comme "inconnues du modèle" ou "inconnues de l'algorithme d'assimilation" les grandeurs contrôlées par l'algorithme d'assimilation 4D-VAR, c'est-à-dire que l'on cherche à restituer à partir du modèle et des observations disponibles.

Dans ce travail :

- Les vitesses terminales de chute ($v_{th,i,p}$), les taux d'évaporation ($\xi_{m,k,i,p}$) et les conditions initiales de la DSD ($n_{0,i}$) sont des paramètres du modèle. Les deux premiers sont calculés via les paramétrisations présentées précédemment. La condition initiale sur la DSD est fixée à 0 ($n_{0,i} = 0, \forall i$). Pour que cette hypothèse soit vraie, il suffit de définir le début des évènements de pluie considérés à un instant où l'on constate (grâce au radar et au disdromètre) qu'il ne pleut à aucune altitude.

- les conditions limites supérieures (DSD en haut de la colonne à toutes dates $n_{k,0}$) et les vents verticaux (à toutes dates et altitudes : $w_{k,i}$) sont les inconnues de l'algorithme d'assimilation. On remarque que l'on ne cherche pas ici (comme c'est souvent le cas dans le domaine de l'assimilation de données) à contrôler l'état initial d'un système mais plutôt à contrôler ses conditions limites et le vent vertical (qui est généralement considéré comme un paramètre du modèle). Ces aspects ont été évoqués plus en détails dans le chapitre Assimilation de données (cf. chapitre III, section 4).

On note qu'au tout début du processus d'assimilation, (première itération de la minimisation numérique), il faut fournir à l'algorithme un point de départ (first guess, exposant g), c'est à dire des valeurs pour les inconnues $n^g_{k,0}$ et w^g. Ces valeurs ne semblent pas critiques dans la mesure où nos restitutions sont fortement contraintes par les observations. Nous utiliserons donc $w^g = 0$ et pour $n^g_{k,0}$ des valeurs faibles et arbitraires, correspondant à un taux de pluie d'environ $0.1mm/h$.

On note que le champ des vents verticaux est de taille $NT * NZ$ et celui des DSD en haut de la colonne de taille $NT * ND$. On a donc $NT\,(NZ + ND)$ inconnues. Afin de limiter ce nombre d'inconnues, nous allons dans la suite paramétriser les DSD en haut de la colonne. La section suivante présente la gestion de cette contrainte.

2.2 Paramétrisation de la DSD en haut de la colonne

Nous allons donc exprimer, dans l'algorithme 4D-VAR, les DSD en haut de la colonne ($n_{k,0}$) sous forme de DSD gamma. Plaçons nous à une date t_k quelconque. La DSD correspondant est $(n_{k,0,p})_{p\in[|0:ND-1|]}$. En supposant les diamètres des gouttes distribués selon une loi gamma, on obtient cette nouvelle expression pour la *p-ième* classe de diamètres de la DSD :

$$n_{k,0,p} = \frac{\alpha}{\Delta D}\int_{D_p}^{D_p+\Delta D} pdf_\gamma(x;k,\theta)dx \tag{V.4}$$

Avec α un coefficient traduisant le nombre total de gouttes par unité de volume (unité : m^{-3}) et pdf_γ la densité de probabilités de la loi gamma, qui dépend de deux coefficients k et θ (respectivement paramètres de forme et d'échelle) :

$$pdf_\gamma(x;k,\theta) = \frac{x^{k-1}}{\Gamma(k)\theta^k}e^{-\frac{x}{\theta}} \tag{V.5}$$

Avec Γ la fonction gamma :

$$\Gamma(k) = \int_0^{+\infty} t^{k-1} e^{-t} dt \tag{V.6}$$

On se reportera à l'Annexe I pour l'implémentation de cette paramétrisation dans le logiciel YAO. Cette Annexe I fournit aussi une discussion liant les coefficients introduits ici à ceux utilisés par Ulbrich (1983) ou ceux utilisés par Testud et al. (2001).

En forçant ainsi les DSD en haut de la colonne sous forme gamma, nous réduisons, à chaque date (t_k) le nombre d'inconnues de ND (les $(n_{k,0,p})_{p\in[|0:ND-1|]}$) à 3 (les 3 coefficients α_k, k_k et θ_k). Nous réduisons donc le nombre total d'inconnues (en comptant le vent vertical) de $NT\,(NZ+ND)$ à $NT\,(NZ+3)$, soit pour un NZ variant entre 15 ($1500m$ de chute) et 30 ($3000m$), une division du nombre d'inconnues d'un facteur 3 à un facteur 5.

L'algorithme restituera donc, pour chaque date, les trois coefficients d'une DSD gamma en haut de la colonne atmosphérique. Ceux-ci sont néanmoins immédiatement convertis en histogrammes $(n_{k,0,p})$. Ce sont ces histogrammes, et non les coefficients α, k et θ, qui sont ensuite propagés par le modèle numérique. La DSD restituée sera ainsi gamma en haut de la colonne mais pas forcément aux autres altitudes (où il s'agira plutôt d'une mixture de gammas).

Par ailleurs, il peut exister une certaine ambiguïté entre ces trois coefficients, surtout quand, comme ici, on ne dispose pas d'observations sur tous les diamètres (très petites gouttes difficiles à observer). **Cela signifie que deux triplets a priori très différents de coefficients peuvent résulter en des DSD** (histogramme $n_{k,0,p}$) **très proches** dans la gamme de diamètres considérés. **Par conséquent, dans la suite, nous montrerons généralement des figures représentant la DSD ou ses moments plutôt que des séries temporelles des coefficients gamma restitués.**

2.3 Opérateurs d'observation

2.3.1 Disdromètres

Nous allons tout d'abord décrire l'opérateur d'observation pour les données d'un disdromètre. Celui-ci permet de simuler, pour une date quelconque, le nombre de gouttes enregistré par l'instrument à partir des sorties du modèle de propagation (DSD). À la date t_k, la sortie de cet opérateur sera notée $N_{dis,k}$. Elle sera fonction de la DSD au niveau du sol $(n_{k,NZ-1})$ via l'opérateur d'observation H_{dis} (soit $N_{dis,k} = H_{dis}\,(n_{k,NZ-1})$).

L'instrument et la nature de sa mesure ont été décrits dans le chapitre I, section 2.2. Nous avons ainsi vu que le disdromètre ne mesure pas directement la DSD, mais un flux de gouttes. Par ailleurs, le temps d'intégration de ses mesures ne sera pas forcément celui du modèle de propagation. En effet, nous avons vu que pour des raisons de stabilité numérique, le modèle a un pas de temps de $\Delta t = 5s$. Les données disdrométriques que nous utiliserons seront généralement dégradées à une résolution plus large (typiquement 1 ou $2min$, voir sections 5, 6 et 7 sur les études de cas). Enfin, nous avons vu que notre modèle de propagation travaillait

avec des classes de diamètre régulièrement espacées de largeur $\Delta D = 0.1mm$. Il peut être utile de modifier ces classes de diamètre pour utiliser les données des disdromètres, notamment en élargissant leur largeur pour les grosses gouttes, nettement plus rares que les petites. On note $NDdis$ le nombre de pas de diamètre des observations des disdromètres.

L'opérateur d'observation pour le disdromètre devra donc :

- calculer un flux de gouttes à partir de la DSD en bas de la colonne ($n_{k,NZ-1}$).
- intégrer ce flux de gouttes du pas de temps du modèle (Δt) à celui de l'instrument (noté T_{dis}). Pour ne pas avoir à interpoler, on suppose que T_{dis} est un multiple de Δt, et on note $i_{T_{dis}}$ le rapport des deux pas ($T_{dis}/\Delta t$).
- transposer ce flux de gouttes de la gamme de diamètres du modèle à celle de l'instrument. Pour ne pas avoir à interpoler, on choisit des pas de diamètres pour l'instrument compris dans ceux du modèle (par exemple $[0.5; 0.6; 0.7; 0.8; 0.9; 1.0]mm$ pour le modèle et $[0.5; 0.7; 1.0]mm$ pour l'instrument).

Ces trois étapes successives, les formules numériques les exécutant, ainsi que les tailles des espaces associés, sont résumés sur la Figure V.1. L'ensemble des opérations permettant de passer de la DSD en sortie de modèle de propagation ($n = (n_{k,i,p})_{\forall k,i,p}$) à la simulation de la mesure disdromètre ($(N_{dis,k,pdis})_{\forall k,pdis}$) constitue l'opérateur d'observation du disdromètre H_{dis}.

2.3.2 Spectres Doppler

Nous allons à présent détailler l'opérateur d'observation pour les spectres Doppler. Celui-ci permet de simuler, pour une altitude et une date quelconque, un spectre Doppler à partir des sorties du modèle de propagation (vent vertical et DSD). Pour la date t_k et l'altitude z_i, la sortie de cet opérateur sera notée $\eta_{obs,k,i}$. Elle sera fonction de la DSD à toutes altitudes (n) et du vent vertical à l'altitude z_i (w_i) via l'opérateur d'observation H_{dop} (soit $\eta_{obs,k,i} = H_{dop}(n, w_i)$).

Il est important de noter qu'en procédant ainsi nous ne négligeons pas l'effet du vent vertical sur la mesure radar, au contraire d'autres algorithmes (METEK (1998) par exemple).

La notion de spectre Doppler et son lien avec la DSD et les autres variables atmosphériques (vent vertical, turbulence) ont été décrits dans le chapitre I, section 2.3.2. Le spectre Doppler était alors noté η_v (unité : $m^{-1}(m/s)^{-1}$). La grandeur η_{obs} définie ici correspond (à quelques hypothèses près listées ci-dessous) à ce η_v intégré à la résolution des observations disponibles et auquel on a appliqué un opérateur logarithmique.

L'opérateur d'observation des spectres Doppler va être composé, comme celui des disdromètres, d'une composition d'opérations successives permettant, à partir des sorties du modèle de propagation (vent vertical, DSD) de simuler l'observation radar. Ces opérations

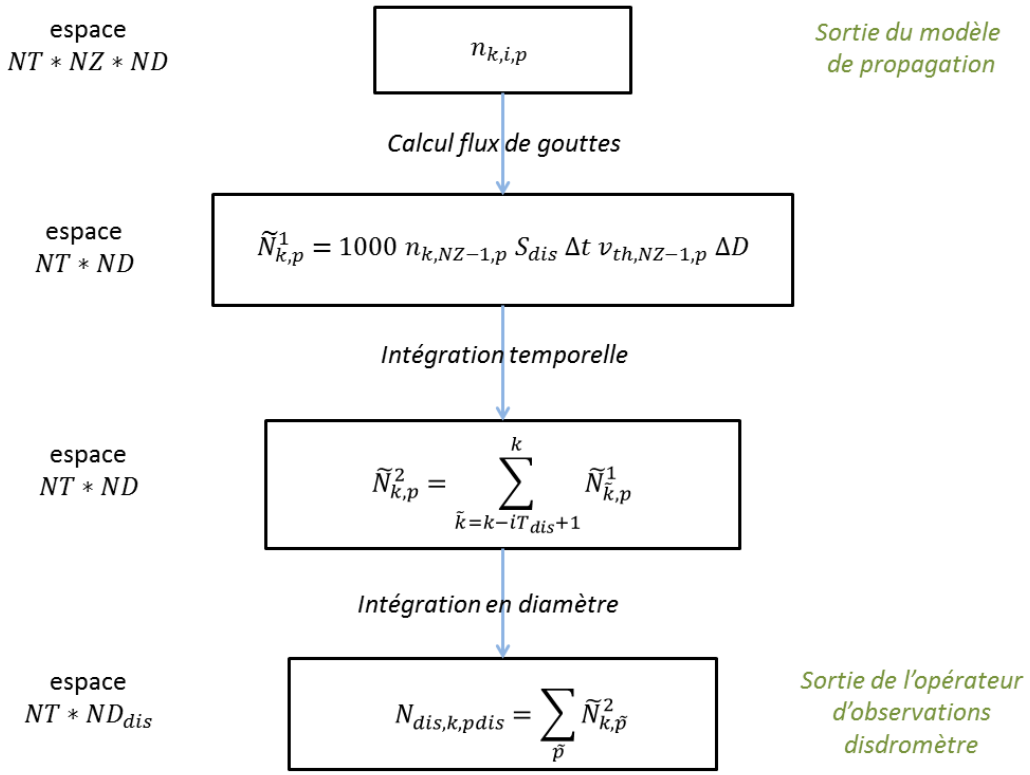

FIGURE V.1 – Schéma représentant l'opérateur d'observation pour un disdromètre, décomposé en trois étapes. L'entrée de l'opérateur d'observation est la DSD en sortie du modèle de propagation (n). On transpose ensuite la DSD au niveau du sol ($i = NZ - 1$) en flux de gouttes (noté ici $\tilde{N}^1$), via la surface de collecte du disdromètre S_{dis}, le pas de temps du modèle Δt, le pas de diamètre ΔD et la vitesse terminale de chute des gouttes de diamètre D_p. On obtient ensuite $\tilde{N}^2$ (deuxième intermédiaire de calcul) en intégrant $\tilde{N}^1$ sur la durée T_{dis}, temps d'intégration du disdromètre. Pour obtenir la sortie N_{dis} de l'opérateur d'observation, on intègre $\tilde{N}^2$ sur les classes de diamètre du disdromètre.

peuvent être classées en trois catégories :

- "physique" : il s'agit de mettre en oeuvre les diverses équations présentées dans le chapitre I, section 2.3.2 et reliant théoriquement DSD, atténuation, vent vertical et turbulence aux spectres Doppler.

- "résolution" : il s'agit d'interpoler à la résolution des observations les grandeurs calculées avec les pas de temps et d'espace du modèle de propagation.

- "numériques" : il s'agit d'appliquer un opérateur (typiquement un logarithme) pour équilibrer les contributions des différents ordres de grandeur présents dans les observations pour faciliter la minimisation numérique.

La Figure V.2 présente tout le schéma (les 8 opérations successives) permettant de calculer $H_{dop}(n, w_i)$. Les opérations "physiques" sont colorées en rouge, celles ayant trait aux changements de "résolution" en bleu et celles "numériques" en vert. On revient rapidement ci-dessous sur ces différentes étapes. Toutes les équations présentées ici sont expliquées (en version continue) dans le chapitre I, section 2.3 et illustrées dans ce même chapitre par la

Figure I.11.

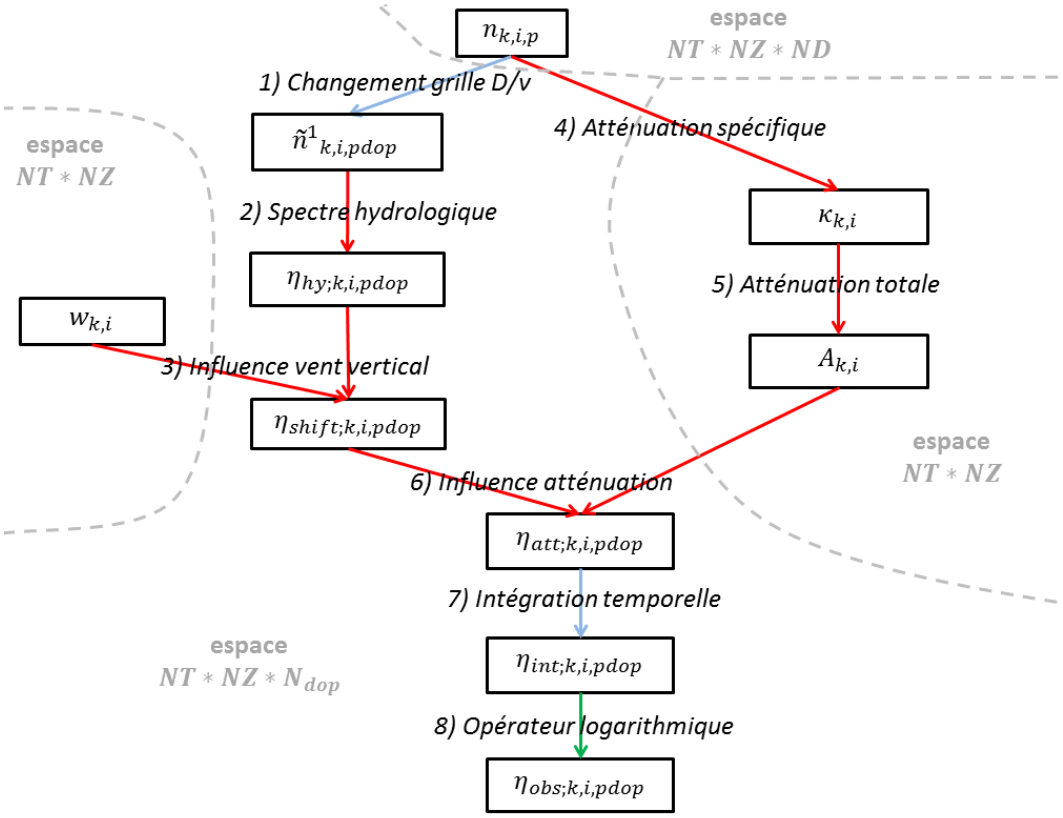

FIGURE V.2 – Schéma résumant les opérations successives effectuées par l'opérateur d'observation des spectres Doppler. Ses entrées sont les DSD en sortie de modèle de propagation ($n_{k,i,p}$) et le vent vertical ($w_{k,i}$). Ses sorties sont les spectres Doppler simulés (logarithmiques) $\eta_{obs,k,i,pdop}$. Voir description détaillée dans le texte.

1. La première opération consiste à projeter les DSD de l'espace du modèle de propagation, composé de ND classes de diamètres à pas fixe ΔD sur celui de l'espace des vitesses Doppler, composé de $Ndop$ classes de vitesses à pas fixe (en vitesse) ΔV_{dop} ($Ndop$ et ΔV_{dop} dépendant du radar utilisé). Diamètre de gouttes et vitesses de chute étant reliés comme on l'a vu (chapitre II, section 3) par les relations de Atlas et al. (1973) et Foote and Du Toit (1969), cette opération consiste à interpoler les valeurs de n sur une nouvelle grille de diamètres. La quantité obtenue est notée $\tilde{n}^1$ (intermédiaire de calcul). *pdop* représente une position quelconque sur la gamme de vitesses (ou de diamètre) Doppler. Le diamètre correspondant est noté $Ddop_{pdop}$.
2. On calcule ensuite le spectre hydrologique (η_{hy}) à partir de la DSD de l'étape précédente ($\tilde{n}^1$). Celui-ci est fourni par l'application successive des Eq. (I.26) et (I.27), qui s'écrivent ici :

$$\eta_{hy;k,i,pdop} = \tilde{n}^1_{k,i,pdop} \sigma(Ddop_{pdop}) \frac{dD}{dv}(Ddop_{pdop}) \tag{V.7}$$

Avec $\sigma(Ddop_{pdop})$ la section efficace de rétrodiffusion pour une goutte de diamètre $Ddop_{pdop}$ et dD/dv la dérivée de la relation diamètre-vitesse.

3. L'influence du vent vertical (décalage en vitesse des spectres) est ajoutée au spectre hydrologique η_{hy} via une version discrétisée de l'Eq (I.28), sans turbulence ($\eta_{air} = 0$) ni bruit ($\eta_{noise} = 0$). On obtient ainsi le spectre "shifté" η_{shift}.
4. On estime ensuite l'atténuation. Pour cela, on calcule d'abord l'atténuation spécifique κ de la couche d'air considérée à partir de la DSD n. Il s'agit ici d'appliquer l'Eq. (I.22), qui s'écrit pour la couche z_i :

$$\kappa_{k,i} = \sum_{p=0}^{ND-1} n_{k,i,p}\sigma_{ext}(D_p)\Delta D \tag{V.8}$$

Avec $\sigma_{ext}(D_p)$ la section efficace d'extinction pour une goutte de diamètre D_p.

5. L'atténuation totale $\mathcal{A}_{k,i}$ pour la couche z_i (due aux gouttes présentent dans cette couche et dans les couches inférieures) est calculée à partir de l'atténuation spécifique $\kappa_{k,i}$ et de l'atténuation totale sur la couche inférieure $\mathcal{A}_{k,i+1}$ via l'équation suivante (équivalent discret à l'Eq. (I.23)) :

$$A_{k,i} = A_{k,i+1}e^{-2\kappa_{k,i}} \tag{V.9}$$

6. On ajoute ensuite l'effet de cette atténuation $\mathcal{A}$ sur le spectre "shifté" η_{shift}. On obtient ainsi le spectre atténué η_{att} en multipliant η_{shift} par $\mathcal{A}$:

$$\eta_{att;k,i,pdop} = \mathcal{A}_{k,i}\eta_{shift;k,i,pdop} \tag{V.10}$$

7. Pour finir, on intègre ce spectre atténué de la résolution du modèle ($\Delta t = 5s$) jusqu'à celle des observations ($30s$ à $2min$ selon le radar utilisé). On obtient ainsi le spectre "intégré" η_{int}.
8. Et on applique un opérateur de régularisation logarithmique. On obtient le spectre "observé" η_{obs}, calculé à partir du précédent par l'opérateur :

$$\eta_{obs} = ln\left(10^{10}\eta_{int} + 1\right) \tag{V.11}$$

Cet opérateur a pour but d'équilibrer les contributions des différentes parties du spectre au coût. En effet, plusieurs ordres de grandeur séparent les énergies rétrodiffusées selon les vitesses de chute.

On peut noter quelques remarques importantes sur cet opérateur d'observation :

- la dernière étape (8) ne fait pas à proprement parler partie de l'opérateur d'observation (puisque dans les faits elle est appliquée à la fois aux sorties du modèle et aux observations réelles, qui ne sont bien sûr pas "naturellement" sous cette forme logarithmique). Algorithmiquement en revanche (i.e. pour YAO), il s'agit bien d'une partie de l'opérateur

d'observation, les observations qui lui sont fournies en entrée ayant été pré-traitées pour apparaître sous cette forme là.

- Le calcul de l'atténuation totale pour une couche donnée (étape 5) nécessite de connaître l'atténuation spécifique (et donc la DSD) pour chacune des couches inférieures. En pratique, l'algorithme commencera donc par calculer (modèle de propagation) les DSD à toutes les altitudes, en procédant dans le sens de chute (du haut de la colonne vers le sol). Il calculera ensuite à rebours (du sol vers le haut de la colonne) les atténuations. Ensuite seulement il évaluera les spectres dans le sens de chute (du haut de la colonne vers le sol).

- on a considéré ici (étape 3) qu'il n'y avait pas de turbulence ($\eta_{air} = 0$). Cela signifie que la turbulence n'est pas représentée explicitement dans l'opérateur d'observation (ni comme paramètre, ni comme inconnue du problème). A priori, cela devrait générer une erreur modèle potentiellement tout à fait significative. Nous verrons cependant plus loin et en détails (voir section 6) que la turbulence est tout de même modélisée, bien qu'indirectement, par l'algorithme (dans le sens où l'élargissement spectral qu'elle produit, voir Figure I.11, est modélisé).

2.4 Filtrage / Pénalisation

Nous allons dans cette section présenter les termes de filtrage et de pénalisation ajoutés à la fonction de coût de notre algorithme VAMOS. Celle-ci sera donc la somme du terme d'écart aux observations (J^{obs}) et des termes décrits ici. De tels termes ont déjà été évoqués dans le chapitre III, section 4.2. Les termes de filtrage ont pour objectif d'éviter qu'une résolution temporelle du modèle plus fine que celle des observations entraîne des oscillations non voulues, en forçant l'algorithme à choisir la solution la plus régulière. Les termes de pénalisation ont pour but de contenir certaines inconnues dans des intervalles prédéfinis ou proches de 0 (minimisation sous contraintes). On note que le poids relatif de ces différentes termes (entre eux et vis-à-vis de J^{obs}) peut et doit être ajusté avant toute expérience d'assimilation. Nous reviendrons dans la section 3 sur la procédure utilisée pour effectuer cet ajustement a priori.

2.4.1 Filtrage

Dans un premier temps, on ajoute donc à la fonction de coût un terme de filtrage, noté J_r. Un tel terme a déjà été évoqué dans le chapitre III, section 4.2.2. Ce terme prend ici en entrée toutes les inconnues du problème, soit les trois séries temporelles de coefficients d'une DSD gamma en haut de la colonne, α, k et θ, ainsi que le champ temps/altitude de vents verticaux w. Pour exprimer J_r, on introduit les deux fonctions $J_{r\gamma}$ et J_{rw}, qui prennent en entrée une série temporelle x à NT éléments. La première s'écrit comme dans l'Eq.(III.27) :

$$J_{r\gamma}(x) = \sum_{k=0}^{NT-1} \left[x_k - \frac{x_{k-1} + x_k + x_{k+1}}{3} \right]^2 \tag{V.12}$$

La seconde s'écrit sous une forme légèrement différente :

$$J_{rw}(x) = \sum_{k=0}^{NT-1} \left[x_k - \frac{x_k + x_{k+1}}{2} \right]^2 \tag{V.13}$$

On note que ces deux termes ont le même objectif, la différence de formulation n'étant qu'accessoire. Il s'agit des équations de récurrence de filtres passe-haut d'ordre 2 qui coupent à 0.25 fois la fréquence d'échantillonage.

Le terme de filtrage s'écrit finalement :

$$J_r(\alpha, k, \theta, w) = \left[J_{r\gamma}(\alpha) + J_{r\gamma}(k) + J_{r\gamma}(\theta) \right]^2 + \left[\sum_{i=0}^{NZ-1} J_{rw}(w_i) \right]^2 \tag{V.14}$$

On a vu précédemment que le pas de temps du modèle de propagation était de $\Delta t = 5s$. Or les observations disponibles (données disdromètres, spectres Doppler) ont généralement une résolution minimale supérieure à cette valeur (typiquement moyenne mobile sur 1 ou $2min$ d'observations à $10s$). Rien ne permet donc à l'algorithme 4D-VAR d'évaluer la variabilité de ses inconnues en-dessous de cette résolution. Le terme de filtrage a pour objectif de faire en sorte qu'il privilégie les solutions les plus régulières. Il est ainsi forcé via son coût à produire les champs les plus lisses possibles (tout en restant compatibles avec les observations).

2.4.2 Pénalisation

Nous allons à présent introduire 2 termes de pénalisation. Le premier, $J_{p\gamma}$ portera sur les coefficients de la DSD gamma en haut de la colonne. Les instruments ne fournissant pas d'informations sur le spectre entier de tailles de gouttes (typiquement les petites gouttes, en-dessous de $0.4mm$, ne sont pas observées), ce terme empêchera l'algorithme de produire des DSD non physiques, en contraignant les coefficients dans un intervalle prédéfini. Le second, J_{pw}, portera sur le vent en bas de la colonne (w_{NZ-1} et éventuellement w_{NZ-2}). Le vent étant essentiellement restitué grâce aux spectres Doppler et ceux-cis n'étant généralement pas disponibles près du sol, on force dans ces cas-là l'algorithme à restituer des vents proches de 0. De tels termes de pénalisation avaient déjà été évoqués dans le chapitre III, section 4.2.1.

Terme $J_{p\gamma}$ Le terme de pénalisation $J_{p\gamma}$ a pour objectif de garder les coefficients de la DSD gamma en haut de la colonne (séries temporelles α, k et θ) dans un intervalle prédéfini. Notons x un scalaire quelconque, et x_l et x_L les bornes supérieures et inférieures dans lesquelles on veut le contraindre. Notons c_x un coefficient qui lui est associé. On définit alors la fonction f_{pen} par :

$$f_{pen}(x) = \begin{cases} 2c_x \text{ si } x < x_l - 2 \\ c_x \left[2 - (x + 2 - x_l)^4\right] \text{ si } x_l - 2 < x < x_l - 1 \\ c_x \left(x - x_l\right)^4 \text{ si } x_l - 1 < x < x_l \\ 0 \text{ si } x_l < x < x_L \\ \left(x - x_L\right)^2 \text{ si } x > x_L \end{cases} \tag{V.15}$$

Cette fonction est celle que l'on avait représentée en exemple dans le chapitre III, section 4.2.1 (Figure III.6). On note que :

- la fonction n'est pas symétrique : le coût ($f_{pen}(x)$) est nul dans l'intervalle $[x_l : x_L]$ mais augmente plus vite pour les valeurs en-dessous de x_l. Cela est justifié par le fait que l'on veut généralement absolument éviter des valeurs négatives, alors que des valeurs fortes sont moins gênantes, notamment pour la stabilité numérique.

- le coefficient c_x n'est pas critique mais permet de régler le poids limite pour les valeurs très basses.

- la fonction est polynomiale par morceaux, ce qui simplifie la programmation de sa dérivée pour YAO.

- la fonction et sa dérivée sont continues, ce qui évite des problèmes d'instabilité numérique.

On applique ensuite cette fonction, pour toute date, à chacun des trois coefficients de la loi gamma. On définit ainsi formellement l'opérateur de pénalisation $J_{p\gamma}$ par :

$$J_{p\gamma}(\alpha, k, \theta) = \sum_{k=0}^{NT-1} f_{pen}(\alpha_k) + f_{pen}(k_k) + f_{pen}(\theta_k) \tag{V.16}$$

Les intervalles ($[x_l : x_L]$) dans lesquels on veut garder les coefficients α, k et θ sont respectivement $[0 : 8000]m^{-3}$, $[0 : 3]$ et $[0 : 10^{-3}]m$. Ces valeurs permettent de décrire une vaste gamme de DSD. En les transcrivant sous la formes des coefficients introduits pas Ulbrich (1983), on peut ainsi faire varier μ_{ulb} dans $[-1 : 2]$ et No_{ulb} et Λ_{ulb} dans $[0 : +\infty]$ (voir l'Annexe I et/ou ll'article Ulbrich (1983)).

Terme J_{pw} Le terme de pénalisation J_{pw} a simplement pour objectif de forcer le vent vertical à 0 lorsqu'il n'y a pas de spectres Doppler disponibles pour le restituer correctement (comme c'est généralement le cas près du sol). On utilise donc, pour un scalaire x quelconque, la fonction f_{pen2} :

$$f_{pen2}(x) = x^2 \tag{V.17}$$

2.5 Conclusion

Nous avons dans cette section décrit en détails l'algorithme d'assimilation 4D-VAR qui sera utilisé dans ce chapitre. Nous rappelons qu'il a pour but de restituer les DSD en haut de la colonne atmosphérique et le champ temps/altitude de vent vertical. Il utilise pour cela un modèle numérique de propagation des DSD prenant en compte la gravité, le vent vertical et l'évaporation, ainsi que des observations provenant de disdromètres situés au sol et mesurant des flux de gouttes et de radars mesurant des spectres Doppler.

La Figure V.3 schématise tout ce processus.

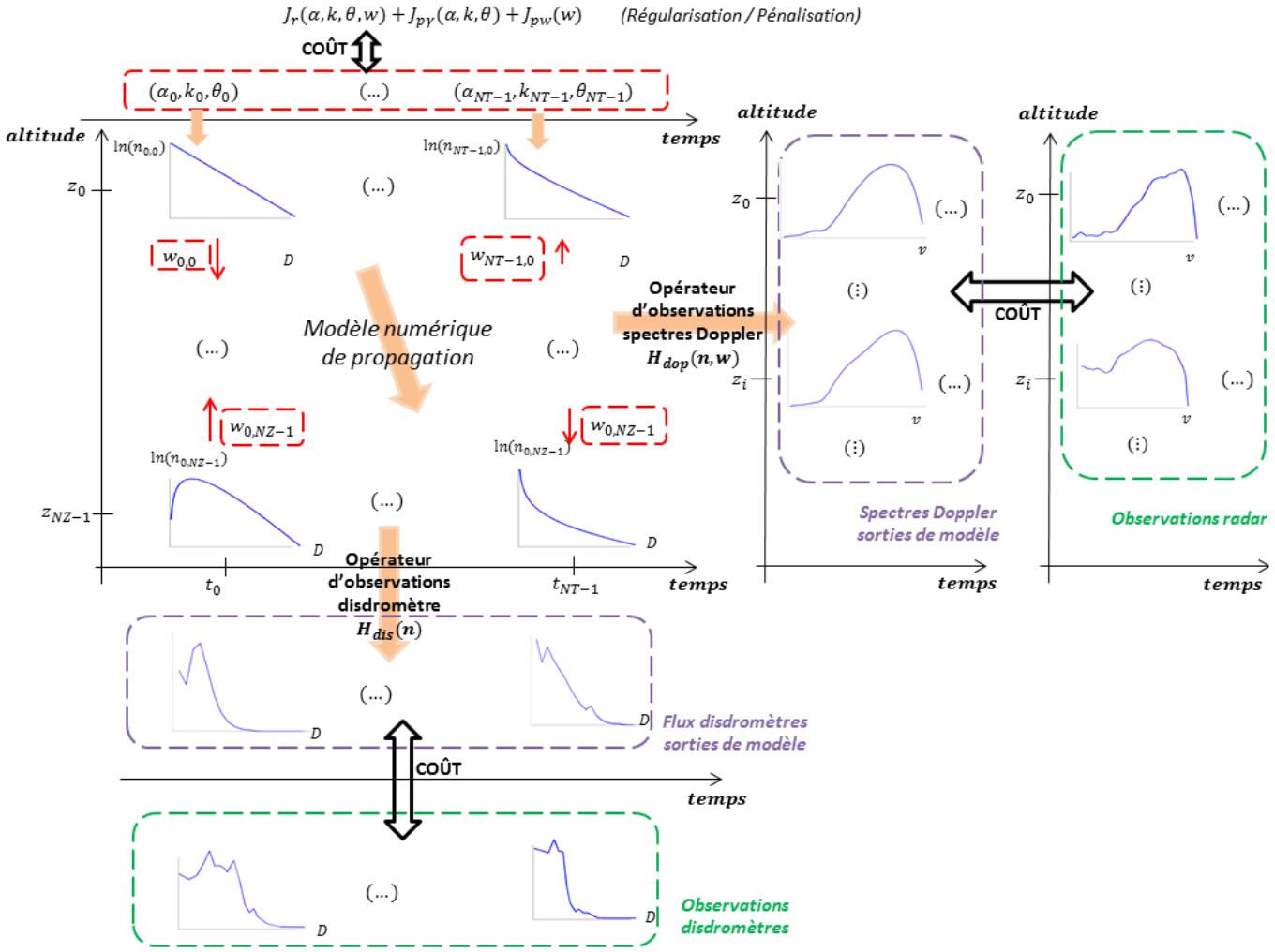

FIGURE V.3 – Schéma résumant l'algorithme d'assimilation 4D-VAR utilisé dans ce travail. Les inconnues de l'algorithme (ce que l'on veut restituer) sont encadrées en rouge. Il s'agit du vent vertical (w) à toutes dates et toutes altitudes et des trois coefficients d'une DSD gamma (α, k, θ) caractérisant les DSD en haut de la colonne atmosphérique à toutes dates. Ces DSD sont ensuite propagées à toutes dates et altitudes par le modèle numérique de propagation. Elles sont transformées en grandeurs comparables aux observations (cadres violets) par les opérateurs d'observation, que ce soit pour les mesures disdromètres (en bas) ou radars (à droite). Le coût est calculé en évaluant l'écart entre ces grandeurs et les observations effectivement disponibles (cadres verts). On ajoute à ce coût des termes de pénalisation et de régularisation des inconnues (haut de la figure).

3 Validation sur données simulées

3.1 Introduction

Dans cette section, nous allons évaluer les réponses de VAMOS dans diverses situations sur données simulées. Les objectifs de cette partie sont de tester la capacité de l'algorithme à restituer les DSD et les vents verticaux en présence notamment de bruit et de turbulence atmosphérique, et d'évaluer ainsi sa robustesse. Nous verrons également comment VAMOS parvient à restituer les effets de la turbulence sur les spectres Doppler sans dégrader la restitution des autres paramètres.

Dans un premier temps (section 3.2), nous verrons la procédure utilisée pour choisir a priori les poids relatifs des différentes termes composant la fonction de coût du 4D-VAR (équilibrage des termes d'observation des spectres Doppler, des disdromètres, et des termes de filtrage et de pénalisation). Puis dans la section 3.3, nous étudierons en détails une expérience d'assimilation sur des données simulées réalistes, et verrons notamment que bien qu'elle ne soit pas paramétrée explicitement dans l'algorithme, la turbulence peut néanmoins être gérée par le modèle utilisé.

Avant cela, nous allons revenir sur le concept d'expérience jumelle en assimilation de données.

Retour rapide sur le concept d'expérience jumelle / Notations

Nous utiliserons toujours ici les notations introduites dans la section 2.1. n désigne le champ temps/altitude de DSD ($n_{k,i}$ désigne la DSD à la date t_k et à l'altitude z_i). w désigne le champ temps/altitude des vents verticaux. On rappelle que les inconnues de l'algorithme d'assimilation (ce que l'on cherche à restituer) sont les DSD en haut de la colonne atmosphérique, paramétrisées sous forme de DSD gamma, et les vents verticaux à toutes dates et toutes altitudes. η_{obs} désigne les spectres Doppler (Eq. (V.11)) et N_{dis} les observations disdromètres.

Dans cette section, nous allons effectuer une expérience d'assimilation sur données simulées, procédé appelé expérience jumelle. Le principe général d'une expérience jumelle consiste à tester l'algorithme d'assimilation en travaillant sur des données simulées en utilisant le modèle direct et les opérateurs d'observations. La Figure V.4 représente schématiquement tout ce processus. La procédure pour effectuer une telle expérience jumelle est la suivante.

Tout d'abord (Figure V.4, première ligne), nous fixons (simulons) les inconnues de notre modèle de propagation (DSD en haut $n_{k,0}$ et vent vertical $w_{k,i}$). Avec le modèle, on propage ensuite ces DSD à toutes dates et toutes altitudes. On obtient ainsi n. On applique alors les opérateurs d'observations (décrits dans la section 2.3) pour simuler des observations de disdromètres (N_{dis}) et de spectres Doppler (η_v). Tous les états décrits jusqu'à présent sont les

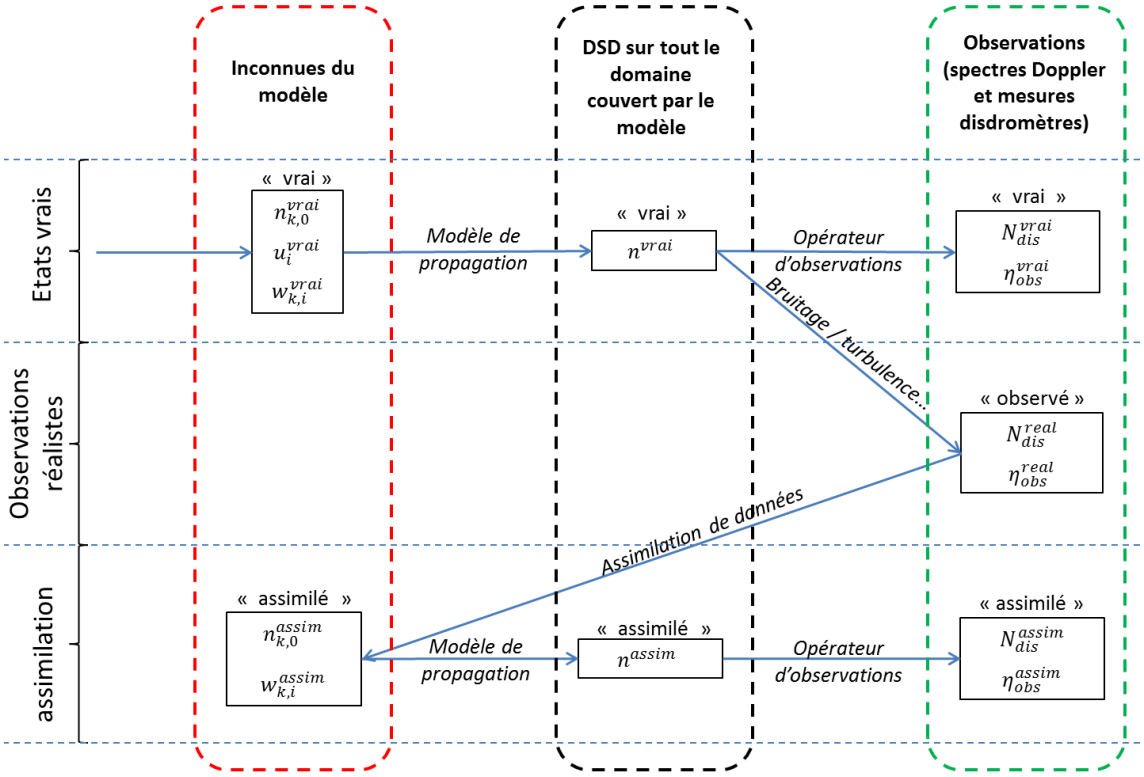

FIGURE V.4 – Description du processus d'expériences jumelles pour la section 3. Voir texte pour la description.

états vrais de notre problème (vérité simulée). Il sont désignés par l'exposant *vrai*. L'objectif de l'algorithme d'assimilation sera ensuite de restituer au mieux ces états vrais (soit restituer n^{vrai} et w^{vrai}).

Les observations vraies (N_{dis}^{vrai}, η_{obs}^{vrai}) sont ensuite perturbées (turbulence, bruit instrumental, Figure V.4 ligne 2) pour être rendues plus réalistes. On obtient ainsi les états réalistes (exposant *real*). Ces grandeurs constituent les données d'entrée de l'algorithme d'assimilation.

En appliquant l'algorithme d'assimilation 4D-VAR, on obtient les inconnues du problème assimilées (Figure V.4 ligne 3). L'assimilation fournit une DSD en haut de la colonne ($n_{k,0}$) et un vent vertical à toutes dates et tous instants ($w_{k,i}$). Propagée verticalement par le modèle de propagation, on obtient la DSD à toutes dates n. En appliquant les opérateurs d'observations, on obtient les observations assimilées. On désigne ces grandeurs assimilées par l'exposant *assim* ($n_{k,0}^{assim}$, w^{assim} par exemple).

3.2 Réglage relatif des coûts

Nous décrivons rapidement dans cette section la procédure employée pour équilibrer les différents termes de la fonction de coût. On rappelle que celle-ci comprend :

- un terme d'écart aux observations radar (spectres Doppler).

- un terme d'écart aux observations disdromètres. Ces deux premiers termes constituent l'écart aux observations (J^{obs}).

- un terme de filtrage, agissant sur chaque inconnue (les trois séries temporelles de

paramètres gamma en haut de la colonne et le champ de vents verticaux).

- un terme de pénalisation pour les paramètres gamma en haut de la colonne.

- un terme de pénalisation pour le vent vertical aux altitudes où il n'y a pas de spectres Doppler disponibles.

Avant d'effectuer une expérience d'assimilation, il convient de donner un poids relatif à chacun de ces termes dans le calcul final de la fonction de coût. Montrons d'abord que ce problème ne se pose pas pour le dernier terme mentionné (pénalisation du vent vertical si pas de spectre Doppler disponible). Ce terme agit (quasiment) indépendamment des autres. En effet, il n'agit que quand il n'y a pas de spectres Doppler disponibles, donc justement quand le terme d'écart aux observations est absent. De plus, le vent vertical n'est pas propagé par le modèle (il est simplement utilisé pour propager les gouttes). Une modification du vent vertical sur une couche ne modifiera pas le vent vertical ailleurs. Enfin, cette absence de spectres Doppler touche généralement au plus les 2 couches les plus près du sol. Or le vent vertical sur une couche affecte la propagation des gouttes vers la couche inférieure. Donc on n'affectera ici qu'au plus une couche (celle immédiatement au-dessus du sol), et seulement les très petites gouttes, seules susceptibles d'être affectées rapidement par des vents assez faibles.

Ce terme de pénalisation agissant indépendamment des autres, on pourra lui affecter un poids quelconque.

3.2.1 Équilibrage des observations radar / disdromètre

Nous allons tout d'abord voir quels poids relatifs affecter aux termes d'écart aux observations radar et disdromètre. Au tout début de la minimisation (first guess), nous avons vu que l'on initialisait les vents verticaux à 0 et les DSD avec des coefficients gamma correspondant à des taux de pluie presque nuls. Lors de la première itération du minimiseur numérique, les sorties de modèle vaudront donc quasiment 0. Par conséquent, le terme d'écart aux observations radar vaudra quasiment la somme des carrés de ces observations, soit formellement $\sum_k \sum_i \sum_{pdop} \left(\eta^{real}_{obs,k,i,pdop}\right)^2$. De la même manière le terme d'écart aux observations disdromètre vaudra alors quasiment $\sum_k \sum_{pdis} \left(N^{real}_{dis,k,pdis}\right)^2$.

Nous allons alors pondérer ce terme par le coefficient $\sum_k \sum_i \sum_{pdop} \left(\eta^{real}_{obs,k,i,pdop}\right)^2 / \sum_k \sum_{pdis} \left(N^{real}_{dis,k,pdis}\right)^2$ (ou pondérer le premier terme par l'inverse de cette quantité).

Ainsi, les deux parties de J^{obs} seront équilibrées lors de la première itération du minimiseur. Cela ne signifie pas qu'elles resteront équilibrées par la suite, mais il s'agit du réglage a priori le plus raisonnable, dans la mesure où l'on ne veut pas affecter une plus grande confiance à l'un des systèmes de mesure plutôt qu'à l'autre.

3.2.2 Équilibrage termes de filtrage et de pénalisation

Supposons que l'on veuille estimer les poids relatifs à donner dans la fonction de coût aux termes de filtrage / pénalisation pour un évènement réel. On travaille alors en deux temps.

1) On règle d'abord ces poids de manière empirique sur une expérience jumelle "référence" effectuée dans une configuration similaire (mêmes instruments disponibles, mêmes résolutions). Pour cela, on procède encore en deux temps.

- On effectue une expérience d'assimilation sans filtrage ni pénalisation. On calcule alors sur les états assimilés obtenus les valeurs que prendraient ces termes de filtrage / pénalisation s'ils étaient appliqués avec un poids 1 (on calcule par exemple $J_{p\gamma}(\alpha^{assim}, k^{assim}, \theta^{assim})$). Puis on ajuste leur poids pour faire en sorte qu'ils ajoutent le même coût que le terme d'écart aux observations à la fin de l'assimilation (ou 0.1, ou 10 fois le coût, on ajustera ensuite).

- Les poids ainsi obtenus sont généralement assez insatisfaisants, dans la mesure où les résultats sans aucun filtre sont généralement extrêmement bruités. Calculer le coût de filtrage / pénalisation qu'auraient ces sorties d'assimilation donne donc généralement des valeurs fortes, parfois sans commune mesure avec ce qu'elles auraient values en présence d'un filtre même très faiblement pondéré. Il convient donc, partant de cette base naïve, d'effectuer plusieurs expériences d'assimilation avec différents poids pour le filtrage / pénalisation. On ajuste ainsi les poids de manière totalement empirique, en comparant les résultats de l'assimilation aux états vrais. On obtient des poids pour chaque terme dans cet expérience jumelle "référence".

2) Lorsqu'on passe ensuite à notre expérience sur données réelles, on réajuste simplement ces poids selon la nouvelle masse d'observations disponibles. Par exemple, si l'on travaille sur une expérience où la somme des observations des spectres Doppler et des disdromètres vaut 2 fois ce qu'elle valait sur notre expérience jumelle "référence", on multiplie simplement par 2 tous les poids des termes de filtrage / pénalisation.

On peut noter que la méthode présentée ici est bien sûr empirique et assez grossière. Il ne faut cependant pas trop s'en formaliser. Les termes de pénalisation / filtrage agissent généralement (par définition) sur des zones du problème où la partie "observations" de la fonction de coût n'agit pas (par exemple J^{obs} agit à la résolution temporelle des observations - $30s$ à $2min$ - tandis que le filtrage agit à celle du modèle : $5s$). Ces termes ne sont donc pas vraiment en compétition mais plutôt complémentaires avec J^{obs}. Il s'agit donc simplement de leur donner des poids suffisamment forts pour qu'ils ne soient pas totalement négligeables et suffisamment faibles pour qu'ils n'écrasent pas le problème en ramenant tous les champs à 0, mais on constate généralement qu'ils produisent des résultats très similaires sur plusieurs ordres de grandeurs. La procédure ci-dessus suffit donc largement dans la plupart des cas.

3.3 Performances sur observations réalistes

3.3.1 Présentation

Nous allons à présent réaliser une expérience jumelle sur observations réalistes (voir le principe Figure V.4). Pour cela, nous simulerons des inconnues vraies (DSD en haut de la colonne et vents verticaux, voir section 3.3.2) puis des observations vraies, via le modèle de propagation et les opérateurs d'observations. Nous ajouterons ensuite aux spectres Doppler vrais ainsi obtenus l'influence d'une turbulence atmosphérique d'intensité $0.7m/s$ (on a vu que la turbulence était représentée mathématiquement par une convolution avec une gaussienne centrée et qu'elle avait pour effet de lisser les spectres Doppler). Nous ajouterons aussi un bruit instrumental. Les données disdromètres seront également rendues plus réalistes (voir section 3.3.3). La section 3.3.5 présentera les résultats de l'assimilation dans VAMOS de ces observations réalistes.

3.3.2 Paramétrisation du modèle et inconnues vraies de l'algorithme

Dans cette section, nous allons présenter le cadre de cette expérience jumelle, c'est-à-dire :

- la paramétrisation du modèle de propagation.
- les inconnues vraies simulées (coefficients d'une DSD gamma en haut de la colonne, vent vertical) que l'on cherchera à restituer.
- les caractéristiques des instruments de mesure simulés.

Paramétrisation du modèle Nous utiliserons ici le modèle de propagation décrit en section 2.1 avec un vent vertical variable avec le temps et l'altitude mais sans évaporation. On justifie cela par le fait que l'évaporation n'est pas contrôlée par l'algorithme mais est simplement un paramètre du modèle. Il n'y a donc pas lieu d'"évaluer" le terme d'évaporation du modèle : celui-ci agira dans tous les cas de manière fixe selon le taux d'humidité et la température donnés. Le code de l'évaporation en lui-même (sens direct et adjoint) a été validé séparément sur une expérience jumelle pure (observations réalistes = observations vraies).

La hauteur de chute des gouttes (formellement $z_0 - z_{NZ-1}$) est de $1400m$, caractéristique des évènements enregistrés dans l'hémisphère nord et présentés dans les section 5 et 6. La durée de l'évènement simulé est de $6h50$, comparable là-aussi aux longs évènements stratiformes qui seront étudiés par la suite.

Inconnues vraies (simulées) et caractéristiques des instruments simulés Nous allons à présent décrire les inconnues vraies utilisées dans cette expérience jumelle. Il s'agit donc de simuler d'une part des séries temporelles de coefficients gamma pour les DSD en haut de la colonne, d'autre part un champ temps/altitude de vent vertical à la résolution $5s/100m$.

Pour la DSD, nous utilisons une série temporelle de taux de pluie enregistrés au sol par disdromètre dans l'Ardèche pour le cas présenté dans la section 5. Il s'agit d'un long évènement de pluie stratiforme, avec des taux de pluie à $2min$ constamment entre 0 et $6mm/h$. Ces taux

de pluie sont transcrits en DSD exponentielles selon la paramétrisation de Marshall and Palmer (1948). Ces DSD exponentielles (donc a fortiori gamma) sont nos DSD vraies en haut de la colonne atmosphérique.

Pour le vent vertical, on simule d'abord un champ de variables gaussiennes centrées réduites indépendantes. Puis on le moyenne localement pour avoir des longueurs et temps de corrélation réalistes (c'est-à-dire correspondant aux restitutions de vent obtenues par Williams (2002) ou Giangrande et al. (2010) pour des cas similaires), soit environ $4-5min$ et $500m$ verticalement. On multiplie les champs obtenus pour avoir des valeurs réalistes également, soit dans la gamme $[-1.5; 1.5]m/s$. Un extrait de ce champ de vents verticaux est présenté sur la Figure V.8A (en section 3.3.5).

Pour les instruments, on utilise les caractéristiques de ceux disponibles pour le cas présenté en section 5. Soit pour le radar : résolution verticale de $100m$ et résolution en vitesse Doppler de $0.19m/s$, couvrant la gamme $[0.56-9.5]m/s$ en vitesse et $[100-1400]m$ en altitudes. Pour le disdromètre, on utilise 19 classes de diamètre (de largeurs croissantes) couvrant la gamme $[0.4-7.5]mm$,avec une surface de collecte $0.01m^2$. Nous effectuerons en fait deux expériences jumelles, avec des observations intégrées à $10s$ (résolution minimale du radar de la section 5) ou sur $2min$. Par défaut, nous présenterons les résultats avec observations à $2min$.

Nous disposons donc à présent de toutes les données (paramètres et inconnues) nécessaires à l'exécution du modèle de propagation et des opérateurs d'observations. On obtient alors toutes les valeurs vraies du problème : DSD (n^{vrai}), vent vertical (w^{vrai}), ainsi que les observations de DSD au sol par disdromètre (N_{dis}^{vrai}) et spectres Doppler (η_{obs}^{vrai}). Pour évaluer les performances de l'algorithme d'assimilation dans des conditions réalistes (mais sans erreurs modèle, ce cas sera traité en section 4.3 sur l'influence du vent horizontal), nous allons perturber ces observations en incluant les phénomènes, physiques et instrumentaux, non pris en compte dans les opérateurs d'observations.

3.3.3 Simulation d'observations réalistes

Dans cette section, nous présentons la procédure utilisée pour rendre plus réalistes les observations vraies obtenues précédemment en appliquant les opérateurs d'observations aux sorties du modèle de propagation.

Spectres Doppler

Le lien théorique entre DSD, paramètres atmosphériques (vent, turbulence) et spectres Doppler a été décrit en détails dans le chapitre I, section 2.3.2. L'opérateur d'observation correspondant, transformant dans notre algorithme d'assimilation (VAMOS) les sorties de modèle (DSD, vent vertical) en spectres Doppler, a lui été décrit dans ce chapitre en section 2.3.2. Il s'agit d'une chaîne d'opérateurs codant une partie de la théorie du chapitre I, section 2.3.2. Il prend en compte la diffusion de Mie, l'atténuation par les couches atmosphériques

inférieures, ainsi que l'effet du vent vertical. Deux phénomènes ne sont cependant pas inclus dans l'opérateur d'observation : la turbulence atmosphérique et le bruit instrumental.

Pour rendre plus réalistes les observations vraies, nous nous proposons ici d'ajouter ces deux effets. Pour cela, nous reprenons les spectres hydrologiques et vents verticaux en sortie de modèle (vrais) et appliquons une version discrétisée de l'Eq.(I.28). Nous reprenons ensuite la chaîne de traitement de l'opérateur d'observation du modèle (voir Figure V.2 en section 2.3.2), ajoutant l'atténuation, l'intégration temporelle et l'opérateur logarithmique.

La turbulence est représentée théoriquement par l'opérateur de convolution de l'Eq.(I.28) avec une gaussienne centrée et présentant un certain écart-type (σ_{air}, voir Eq.(I.29)) caractérisant son intensité. On rappelle que la turbulence lisse les spectres Doppler en modulant à haute fréquence la vitesse de chute des gouttes de pluie et donc leur signature Doppler (voir chapitre I, section 2.3.2 et Figure I.11).

On utilise une turbulence d'intensité $\sigma_{air} = 0.7m/s$, valeur assez élevée mais réaliste comparée à d'autres études restituant σ_{air} (Peters et al. (2005) ou Tridon et al. (2013)). On ajoute également un bruit instrumental d'intensité $\eta_{noise} = 1dB$.

Données disdromètres

Là encore, nous allons essayer de représenter au mieux la réalité pour simuler des observations disdromètres réalistes. Pour cela, on considère, pour une date t_k quelconque, la DSD vraie dans la couche verticale immédiatement au-dessus du sol, c'est-à-dire d'indice $NZ-1$, correspondant aux altitudes $[0 : \Delta z]$ (les 100 premiers mètres de la colonne avec $\Delta z = 100m$). Cette DSD s'écrit $n^{vrai}_{k,NZ-1}$. Elle est naturellement en $m^{-3}mm^{-1}$. Le nombre de gouttes de la classe de diamètre p présentes dans la maille à la date t_k à la verticale d'un disdromètre de surface de collecte S est $n^{vrai}_{k,NZ-1,p} S\Delta D\Delta z$ (c'est déjà comme cela que l'on a bâti l'opérateur d'observation des disdromètres, voir section 2.3.1). Ici, au lieu de procéder "théoriquement", nous allons tirer des gouttes et observer leur devenir.

Considérons l'une de ces gouttes. Notons D_g, v_g et z_g respectivement son diamètre, sa vitesse de chute et sa position verticale (son altitude) à la date t_k. Nous supposons que :

- son diamètre est uniforme dans la classe de diamètre à laquelle elle appartient, soit : $D_g \sim \mathcal{U}([D_p : D_{p+1}])$.

- sa vitesse est sa vitesse terminale modulée par la turbulence, soit : $v_g \sim \mathcal{N}(v_{th}(D_g), \sigma^2_{air})$, avec σ_{air} l'intensité de la turbulence définie au paragraphe précédent ($0.7m/s$ ici).

- son altitude est supposée uniformément distribuée dans la maille : $z_g \sim \mathcal{U}([0 : \Delta z])$.

Pour chaque goutte, on tire ainsi un diamètre, une vitesse et une position. On en déduit son altitude à la fin du pas de temps (à la date $t_k + \Delta t$), par $z_g - v_g\Delta t$. Si cette valeur est négative ou nulle, cela signifie que la goutte a atteint le sol, et donc qu'elle a été enregistrée observée par le disdromètre.

C'est ainsi qu'on obtient des observations disdromètres "réalistes", en incluant notamment l'effet de la turbulence atmosphérique.

Nous disposons maintenant d'observations réalistes, N_{dis}^{real} et η_{obs}^{real}. Nous pouvons donc passer à leur assimilation. Pour évaluer la qualité des champs restitués nous aurons besoin d'indicateurs introduits dans la section suivante.

3.3.4 Indicateurs statistiques

Comme nous l'avions fait dans le chapitre IV section 5.1.5 sur le capteur Ku, nous allons introduire quelques indicateurs statistiques qui nous permettront d'évaluer les performances de notre algorithme de restitution. Supposons que l'on souhaite comparer un champ A^{vrai} et un champ générique A^{assim}, de taille quelconque (ici nous considérons des champs de taille $NX * NY$).

Pour des champs présentant des valeurs loin de 0 et couvrant des gammes de valeurs limitées (par exemple des largeurs spectrales ou des vitesses Doppler moyennes), nous utiliserons des indicateurs relatifs sans risques de donner trop d'importance aux valeurs proches de 0. Les deux indicateurs correspondant introduits ici sont le $MAPE$ (Moyenne (en %) des valeurs absolues des erreurs relatives - Mean Absolute Percentage Error), qui s'écrit formellement :

$$MAPE(\%) = 100.\frac{1}{NX.NY}\sum_{i=1}^{NX}\sum_{j=1}^{NY}\frac{\left|A_{i,j}^{vrai} - A_{i,j}^{assim}\right|}{\left|A_{i,j}^{vrai}\right|} \tag{V.18}$$

Et le biais relatif $rbias$, déjà décrit au chapitre IV section 5.1.5 :

$$rbias(\%) = 100.\frac{\sum_{i=1}^{NX}\sum_{j=1}^{NY}\left(A_{i,j}^{vrai} - A_{i,j}^{assim}\right)}{\sum_{i=1}^{NX}\sum_{j=1}^{NY}A_{i,j}^{vrai}} \tag{V.19}$$

Pour des champs ayant des valeurs proches de 0 (par exemple le vent vertical ou les taux de pluie), on utilise le MAE (Moyenne des erreurs absolues - Mean Absolute Error), qui s'écrit formellement :

$$MAE = \frac{1}{NX.NY}\sum_{i=1}^{NX}\sum_{j=1}^{NY}\left|A_{i,j}^{vrai} - A_{i,j}^{assim}\right| \tag{V.20}$$

3.3.5 Résultats : Reproduction des observations / Résultats généraux

Dans cette section, nous étudierons tout d'abord la capacité de l'algorithme d'assimilation à expliquer les observations. Ensuite, nous étudierons les DSD restituées. Enfin, nous nous attarderons sur les caractéristiques du champ de vent vertical restitué par notre algorithme VAMOS.

Par ailleurs, nous avons vu que nous effectuions notre expérience jumelle pour deux temps d'intégration des observations : $10s$ ou $2min$. Puisque nous conclurons que le deuxième cas est le plus satisfaisant (voir section 3.3.6), les résultats présentés par défaut (lorsque rien d'autre n'est spécifié) sont ceux correspondant à $2min$.

Nous évaluerons également (voir Tableau V.2 en particulier) les indicateurs statistiques précédents pour différentes résolutions des champs considérés (l'algorithme nous fournit des

TABLE V.1 – Comparaisons entre champs vrais, réalistes et assimilés, pour les deux premiers moments des spectres Doppler (vitesse Doppler moyenne W et largeur spectrale LS) et pour le nombre de gouttes enregistrées par le disdromètre. On présente pour les deux premiers cas la $MAPE$ (Eq.(V.18)) et le biais relatif (Eq.(V.19)) et pour le troisième cas la MAE ((V.20)). La dernière colonne donne la moyenne des champs vrais.

			réal-vrai	assim-réal	assim-vrai	mean(vrai)
radar	vitesse Doppler moyenne (W)	$MAPE$ (%) rbias (%)	0.40 −0.36	1.6 +0.48	1.6 +0.12	$6.2m\ s^{-1}$
	largeur spectrale (LS)	$MAPE$ (%) rbias (%)	12 +12	3.3 −2.3	9.7 +9.5	$1.3m\ s^{-1}$
disdro	nbre de gouttes (mm^{-1})	MAE	42	38	27	689

champs au pas de temps du modèle, $5s$, mais on peut dégrader cette résolution). Notre objectif étant de restituer des paramètres à la résolution des observations, nous présenterons par défaut des résultats (cartes ou indicateurs) correspondant à une résolution de $2min$.

Reproduction des observations

Le Tableau V.1 (dernière ligne) présente la MAE (moyenne des erreurs absolues) entre les observations disdromètres réalistes et vraies (colonne 4), assimilées et réalistes (colonne 5) et assimilées et vraies (colonne 6), ainsi que le nombre moyen de gouttes vrai. On constate que l'assimilation reproduit très bien les observations vraies, avec une erreur moyenne de 27 gouttes/mm pour un nombre moyen de 689 gouttes/mm. On constate également que l'écart entre champs assimilé et réaliste est plus fort ($38mm^{-1}$) qu'entre assimilé et vrai ($27mm^{-1}$), et sensiblement égal à l'écart entre vrai et réaliste ($42mm^{-1}$). Cela signifie que l'assimilation explique bien les structures du champ vrai mais pas le bruit ajouté aux observations vraies au sol pour les rendre plus réalistes.

On note que l'on compare ici des flux de gouttes et non des DSD. Les premiers sont proportionnels aux produit des secondes par la vitesse de chute des gouttes. Cela rééquilibre le poids relatif des différentes gammes de diamètre (une DSD exponentielle présente un nombre prépondérant de petites gouttes mais cela est moins vrai lorsqu'on pondère par la vitesse de chute).

Les données disdromètres semblent donc bien reproduites.

Intéressons nous à présent aux spectres Doppler. On a vu que la différence entre spectres Doppler vrais et réalistes était l'ajout de turbulence et de bruit instrumental. Comme on l'a vu au chapitre I, section 2.3.2, la turbulence a pour effet de lisser les spectres Doppler. Elle augmente donc leur largeur spectrale LS (écart-type, moment centré d'ordre 2) sans modifier la vitesse Doppler moyenne W (espérance, moment centré d'ordre 1). Cet effet se voit bien sur le Tableau V.1 (lignes 2 à 5, colonne 4), qui présente, pour W et LS, les $MAPE$ et biais relatifs entre spectres Doppler réalistes et vrais. On voit que le passage des spectres vrais à

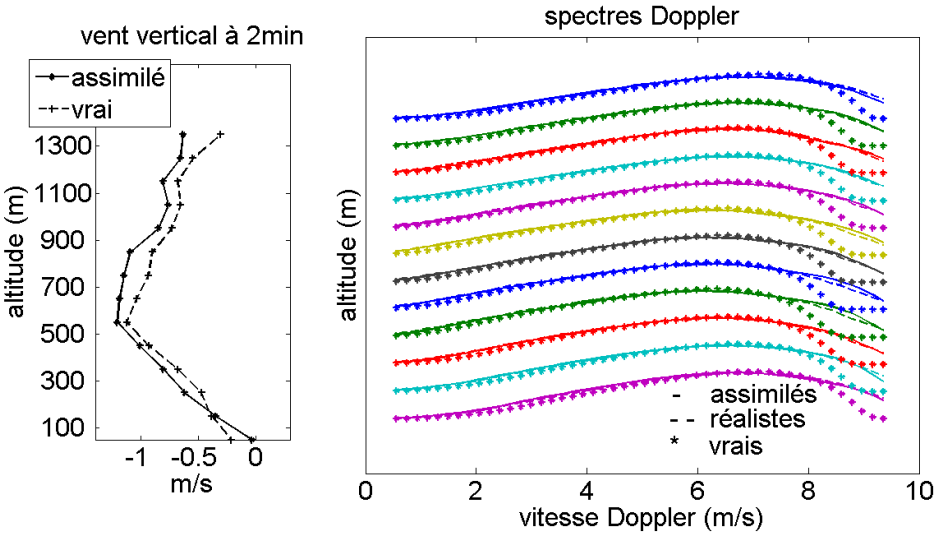

FIGURE V.5 – Exemple de profil vertical de vent vertical (gauche) et de spectres Doppler (logarithmiques, droite). Pour le vent, on présente les données vraies (pointillés) et assimilées (continu). Pour les spectres, on présente les données vraies (étoiles), réalistes (incluant la turbulence, tirets) et assimilées (continus). Le vent est positif vers le bas.

réalistes ne modifie que marginalement la vitesse Doppler moyenne (-0.36%), mais augmente sensiblement la largeur spectrale ($+12\%$). L'examen des résultats de l'assimilation (colonnes 5 et 6 du Tableau V.1) nous montre :

- que l'algorithme reproduit bien la vitesse moyenne (vraie ou réaliste puisque ce sont sensiblement les mêmes). Les erreurs sont faibles, $MAPE$ de 1.6% et biais inférieur à 0.5%.

- que l'algorithme se rapproche davantage des spectres réalistes que des spectres vrais (l'algorithme sous-estime la largeur spectrale de 2.3% par rapport à la valeur réaliste mais surestime de 9.5% la valeur vraie).

Ceci est confirmé par l'affichage des spectres à une date particulière. C'est ce qui est fait sur la Figure V.5 (droite). Pour une date donnée de l'évènement, on a représenté les spectres Doppler vrais, réalistes et assimilés à toutes les altitudes (échelle logarithmique). On constate que les spectres réalistes sont plus lisses que ceux vrais (phénomène particulièrement visible vers $8.5 - 9m/s$), et que les spectres assimilés sont plus proches des spectres réalistes que de ceux vrais.

On conclut donc que l'algorithme a reproduit correctement les spectres Doppler incluant la turbulence (spectres réalistes). A ce stade, on ne sait pas encore comment il a pu reproduire l'effet de la turbulence. Cela s'est-il fait au détriment de la DSD, ou d'une autre manière ? En effet, ajouter des petites (tombant entre 0.5 et $5m/s$) et des grosses gouttes ($> 8m/s$) tout en retirant des intermédiaires ($5 - 8m/s$) pourrait permettre d'expliquer les résultats de la Figure V.5. Nous verrons néanmoins dans le paragraphe suivant que ce n'est pas le cas car la DSD est bien restituée. Nous reviendrons en détail sur la gestion de la turbulence dans la section 3.3.6.

Restitution des DSD

Nous avons vu que notre algorithme d'assimilation reproduisait de manière satisfaisante les observations qu'on lui fournissait. Nous allons à présent étudier sa capacité à restituer les DSD et les paramètres associés (taux de pluie, contenu en eau liquide, etc.).

La Figure V.6 présente (A, C) les DSD vraies en haut et en bas de la colonne atmosphérique (soit $n_{k,0}^{vrai}$ et $n_{k,NZ-1}^{vrai}$). On peut notamment remarquer que les petites gouttes (bas des images) arrivent au sol (Figure V.6C) plus tard que les grosses (haut des images), puisqu'elles tombent plus lentement. Les images B et D présentent les DSD assimilées correspondant ($n_{k,0}^{assim}$ et $n_{k,NZ-1}^{assim}$). On constate que pour les deux altitudes, les DSD semblent très bien restituées par notre algorithme. La Figure V.6E enfin présente les taux de pluie vrai et assimilé au sol. Là encore, les résultats sont satisfaisants.

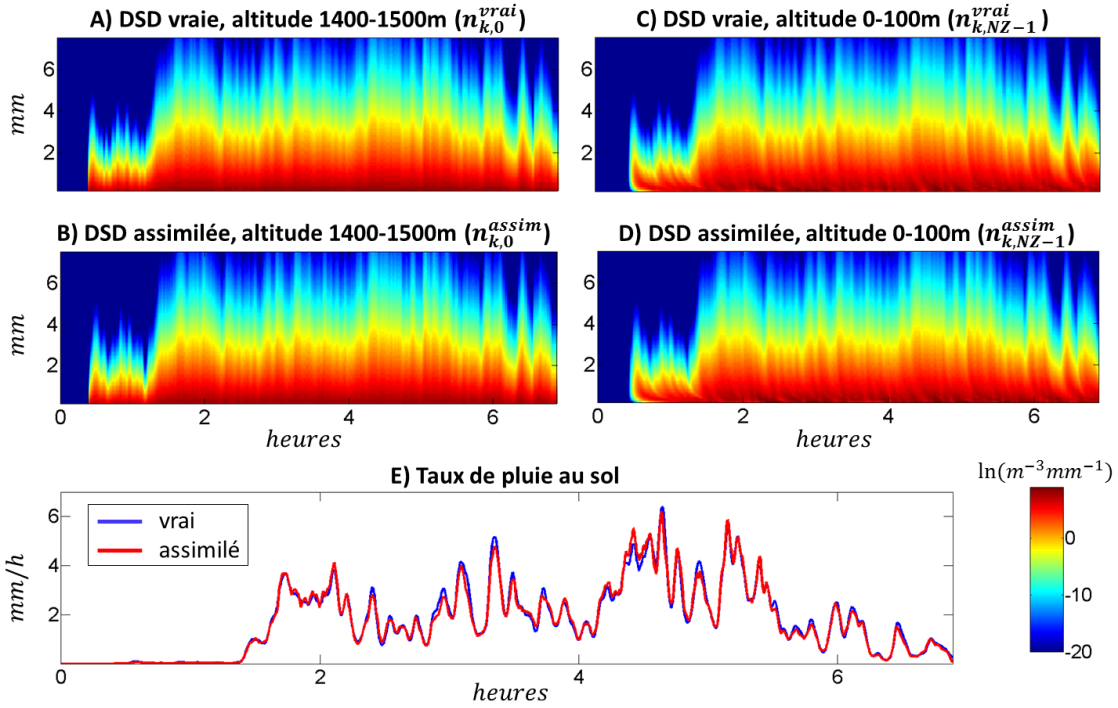

FIGURE V.6 – DSD et taux de pluie vrais et assimilés. A) C) DSD vraies en fonction du temps en haut et en bas de la colonne atmosphérique. B) D) Idem assimilées. Unité : $ln(m^{-3}mm^{-1})$. E) Taux de pluie au sol vrais (bleu) et assimilés (rouge) selon la date.

Nous allons maintenant regarder plus spécifiquement quelques paramètres issus de ces DSD. Le Tableau V.2 présente les MAE (moyenne des erreurs absolues) entre différents paramètres vrais et assimilés, pour différentes résolutions ($10s$ à $8min$) et pour deux temps d'intégration des observations ($10s$ ou $2min$). On s'intéresse pour l'instant seulement au cas où les observations ont une résolution de $2min$. Pour le taux de pluie (RR), le facteur de réflectivité radar équivalent (Z), le diamètre volumique moyen (D_m) et, dans une moindre mesure, le contenu en eau liquide (LWC), on obtient des MAE entre paramètres vrais et assimilés faibles comparées à la valeur absolue moyenne (dernière ligne). Ceci est particulièrement vrai pour des résolution supérieures ou égales à la résolution des observations

TABLE V.2 – Comparaison (MAE, voir Eq.(V.20)) entre champs vrais et assimilés pour le vent vertical (colonne 3) et différents paramètres tirés de la DSD : taux de pluie (RR), facteur de réflectivité radar (Z), diamètre volumique moyen (D_m), contenu en eau liquide (LWC) et nombre de gouttes (N_{tot}). On présente les résultats pour des observations intégrées à $10s$ ou $2min$ (colonne 2) et pour des champs dégradés à différentes résolution ($10s$ à $8min$). On présente aussi le biais absolu et la valeur absolue moyenne. Pour chaque situation, la meilleure performance selon la résolution des observations est en gras.

résolution	obs. window	vent ($m\ s^{-1}$)	RR ($mm\ h^{-1}$)	Z (dBZ)	D_m (μm)	LWC ($\mu g\ m^{-3}$)	N_{tot} (m^{-3})
$10s$	$10s$	*0.21*	*0.27*	*0.96*	*28.5*	*14.5*	*337*
	$2min$	0.46	0.51	1.05	43.0	19.2	**257**
$30s$	$10s$	*0.17*	*0.24*	*0.94*	*27.4*	*14.2*	*334*
	$2min$	0.40	0.34	**0.92**	33.4	14.8	**251**
$1min$	$10s$	*0.17*	*0.24*	*0.91*	*26.4*	*13.8*	*330*
	$2min$	0.29	**0.17**	**0.83**	**26.2**	**10.7**	**246**
$2min$	$10s$	*0.16*	*0.23*	*0.87*	*25.0*	*13.5*	*323*
	$2min$	**0.14**	**0.14**	**0.80**	**24.7**	**10.0**	**241**
$4min$	$10s$	*0.15*	*0.22*	*0.81*	*23.3*	*13.1*	*311*
	$2min$	**0.12**	**0.13**	**0.77**	23.5	**9.54**	**234**
$8min$	$10s$	*0.14*	*0.22*	*0.74*	*21.8*	*12.8*	*296*
	$2min$	**0.098**	**0.12**	**0.72**	22.4	**9.13**	**225**
abs. bias	$10s$	*0.10*	*0.20*	*0.69*	*22.6*	*13.0*	*305*
	$2min$	**0.054**	**0.087**	0.71	22.7	**8.67**	**212**
true absolute mean		0.34	2.18	27.6	1120	142	1070

(i.e. $\geq 2min$). Par exemple, à résolution $2min$, on a une MAE sur RR de $0.14mm/h$ contre un taux de pluie moyen de $2.18mm/h$, ou une MAE sur D_m de $0.025mm$ contre une valeur moyenne vraie de $1.12mm$. Tous ces paramètres sont donc très bien restitués par notre algorithme, ce qui est cohérent avec les résultats visuels de la Figure V.6.

Les résultats pour le nombre de gouttes (N_{tot}, dernière colonne du Tableau V.2) sont cependant nettement moins satisfaisant, avec par exemple (toujours à résolution $2min$) une MAE de 241 gouttes/m^3 pour une valeur moyenne de 1070 gouttes/m^3.

Ce résultat est lié au fait que, les DSD étant globalement exponentielles (exactement exponentielles en haut de la colonne, et mélange d'exponentielles aux niveaux inférieurs), le nombre total de gouttes N_{tot} (dans la gamme de diamètres considérée, soit $[0.2 : 7.4mm]$) est très majoritairement déterminé par le nombre de petites gouttes (typiquement $< 0.4 - 0.5mm$). Au contraire, les autres paramètres du Tableau V.2 sont déterminés par des moments assez élevés de la DSD (6 pour Z, 3 pour LWC, autour de 3.7 pour RR). Les petites gouttes les influencent donc peu.

Or les très petites gouttes ne sont vues par aucun appareil de mesure. Le disdromètre est supposé ne voir que les gouttes de diamètre $> 0.4mm$, et les énergies sur les spectres Doppler sont très faibles pour les petites gouttes (cf. par exemple Figure V.5 : pour des vitesses

Doppler vers $2m/s$, on a environ 65 fois moins d'énergie que vers $7m/s$). L'algorithme est donc relativement libre à ce niveau-là.

La Figure V.7 confirme visuellement ce résultat. Elle montre, pour une altitude donnée $(900-1000m)$, les séries temporelles des nombres de gouttes entre 0.2 et $0.4mm$ (haut) et entre 3 et $4mm$ (bas) vrais (bleu) et assimilés (rouge). On voit bien que les résultats sont très bon dans le second cas, et assez médiocres dans le premier.

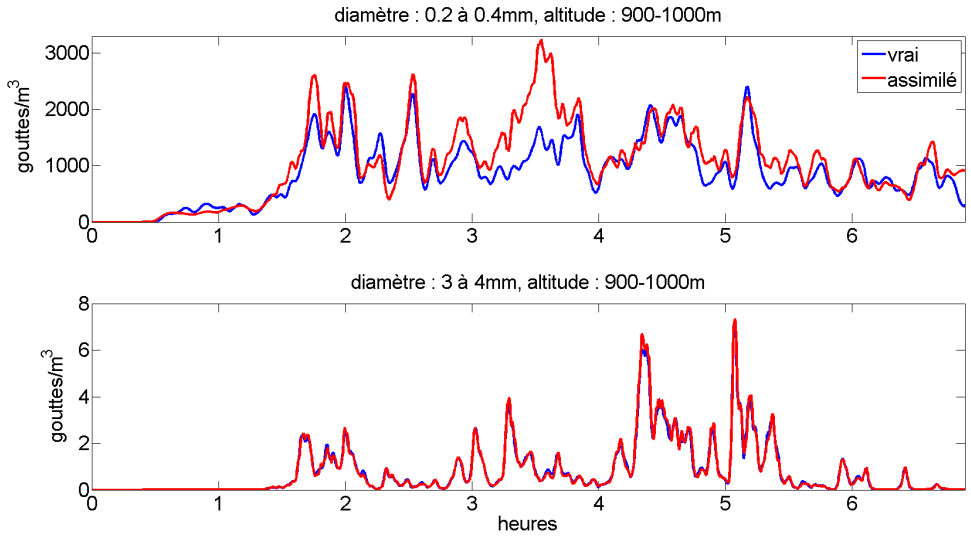

FIGURE V.7 – Nombre de gouttes dans la gamme $0.2-0.4mm$ (haut) et $3-4mm$ (bas) selon la date dans la couche $900-1000m$. Données vraies (bleu) et assimilées (rouge).

Nous concluons donc ce paragraphe par les résultats suivants :

- notre algorithme restitue de manière très satisfaisante les DSD pour les diamètres contrôlés par au moins l'un des instruments de mesure ($> 0.4mm$).

- les résultats pour les petites gouttes ($< 0.4mm$) doivent être utilisés avec beaucoup de prudence, et il faut en particulier se méfier du paramètre N_{tot} (moment d'ordre 0 de la DSD).

- cela ne remet pas en cause la restitution des paramètres classiquement tirés de la DSD (RR, Z, LWC, D_m).

Restitution du vent vertical

Intéressons nous à présent au second champ restitué par l'algorithme : le vent vertical.

La Figure V.8 (A et B) présente les champs des vents verticaux vrais et assimilés pour toute la colonne atmosphérique entre $2h45$ et $4h15$ d'évènement (résolution $2min$, vent positif vers le bas). On constate visuellement que le vent vrai semble très bien reproduit par l'algorithme d'assimilation, même si celui-ci semble produire un champ (légèrement) plus perturbé.

On représente ensuite (Figure V.8C, D, E), selon l'altitude, les moyennes (C) et écart-type (D) des vents vrais (bleu) et assimilés (rouge), ainsi que les autocorrélations temporelles des champs vrai et assimilé (E). On constate qu'il existe un très léger biais sur la moyenne, avec

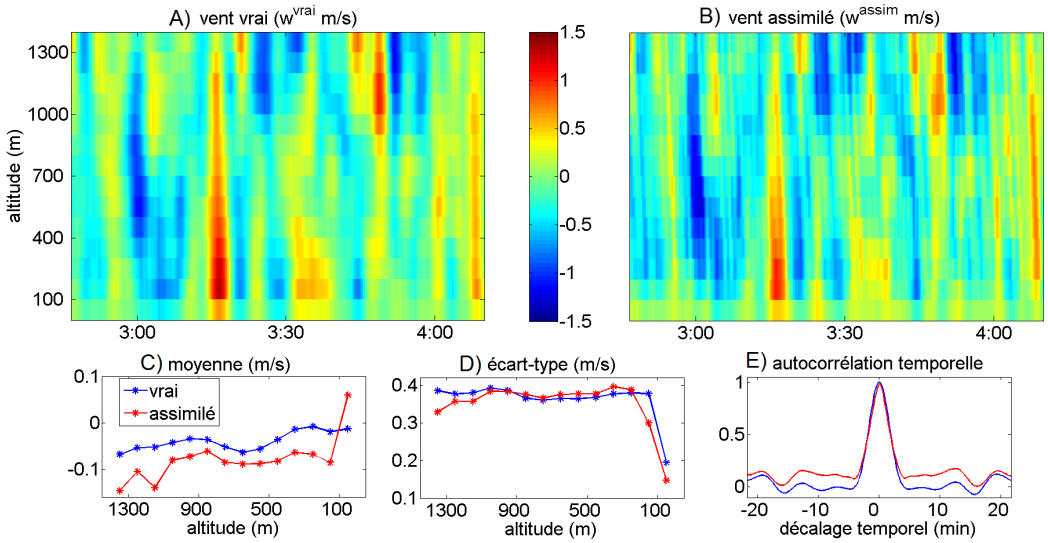

FIGURE V.8 – Restitution du vent vertical (résolution : $2min$). A) Vent vertical vrai. B) Assimilé. C) Selon l'altitude, moyenne des vents verticaux vrais (bleu) et assimilés (rouge). D) Idem écart-type. E) Fonction temporelle d'autocorrélation pour les vents verticaux vrais (bleu) et assimilés (rouge). Vents positifs vers le bas.

un vent assimilé environ $0.04m/s$ en-dessous du vent moyen. Son ampleur est tout de même très limitée. Les écarts-types et temps de corrélation sont eux très bien reproduits.

Reprenons à présent le Tableau V.2. La colonne 3 donne la MAE entre vents verticaux vrais et assimilés pour différentes résolutions du champ et temps d'intégration des observations. Ici, on considère exclusivement le temps d'intégration de $2min$. On constate là encore que les erreurs sont faibles pour des résolutions supérieures ou égales à $2min$ (MAE de $0.14m/s$ à résolution $2min$). Les erreurs sont en revanche sensiblement plus fortes pour les résolutions inférieures. Nous allons y revenir dans la section suivante (section 3.3.6 sur la turbulence).

3.3.6 Gestion de la turbulence

Dans cette partie, nous montrons que notre algorithme d'assimilation parvient à reproduire les effets de la turbulence atmosphérique sur les données radar (élargissement des spectres Doppler) sans altérer la restitution des paramètres physiques sous-jacents (DSD, vent vertical à $2min$). Pour cela, il utilise la fluctuation des vents verticaux à la résolution du modèle ($5s$) sur des fenêtres temporelles correspondant au temps d'intégration des observations ($2min$ ici).

On a vu (Tableau V.1 et Figure V.5) que l'algorithme d'assimilation reproduisait très précisément les spectres Doppler réalistes (qui sont lissés par la prise en compte d'une turbulence). Il parvient donc à restituer l'effet de la turbulence sur les spectres Doppler. On a vu également (Figure V.6 et V.8) que la DSD (et ses moments) et le vent vertical à résolution $2min$ étaient également bien restitués. L'algorithme n'utilise donc pas la DSD pour expliquer

la turbulence (notamment parce que la DSD est directement contrôlée au sol par le disdromètre). Nous allons donc expliquer comment l'algorithme procède pour restituer la turbulence qui n'était pas a priori prise en compte dans le modèle.

Nous avons vu que dans le cas par défaut (résultats présentés jusqu'à présent), les observations (spectres Doppler notamment) étaient intégrées à $2min$, tandis que le vent est toujours restitué par l'algorithme à la résolution du modèle : $5s$. Supposons que, sur une période temporelle de $2min$, le vent à $5s$ oscille autour d'une certaine valeur moyenne (bruit blanc ou autre). Sur les spectres Doppler à $5s$, ces oscillations se traduiront par des décalages vers les petites vitesses Doppler (vent à $5s$ vers le haut freinant les gouttes) ou vers les grandes (vent vers le bas qui les accélère). Si l'on moyenne sur $2min$ les 24 spectres Doppler à $5s$ correspondant, on obtiendra bien un spectre Doppler moyen plus lisse que s'il n'y avait pas eu ces oscillations.

La turbulence est donc gérée par le modèle en faisant fluctuer le vent vertical sur des intervalles de temps correspondant au temps d'intégration des observations.

Le Tableau V.2 met en lumière ce phénomène. Toujours pour des observations à résolution $2min$, on obtient sur le vent (colonne 3) des erreurs fortes pour des résolutions en-dessous des $2min$ et nettement plus faibles ensuite. Cela signifie que le vent à $10s$ ou $30s$ est déconnecté de la réalité (bruit haute fréquence de moyenne nulle) puisque ces résolutions sont utilisées pour modéliser la turbulence. Moyenné à $2min$ et au-delà, les fluctuations haute fréquence se compensent, et le vent vertical correspond à nouveau à l'état vrai.

Observons à présent ce qu'il se passe si l'on utilise des observations intégrées à seulement $10s$, soit quasiment la résolution du modèle ($5s$). Le Tableau V.2 présente aussi ces résultats. Intéressons nous d'abord au vent (colonne 3). Cette fois-ci, le vent est forcé à être cohérent avec les observations dès la résolution $10s$. La MAE du vent est donc sensiblement plus faible à cette résolution que dans le cas précédent (0.21 contre 0.46). Mais cela signifie également que l'algorithme ne peut plus utiliser les fluctuations haute fréquence du vent pour expliquer la turbulence (ces oscillations créeraient de trop gros coûts sur les observations radar à $10s$). Par conséquent, les observations sont moins bien expliquées. Les MAE à résolution moins fines ($2min$ et au-dessus) sont aussi légèrement plus fortes (par exemple pour le vent, erreur à $8min$ de 0.14 contre 0.098, pour le LWC, erreur à 13.1 contre 9.54).

Il convient donc plutôt d'utiliser des observations intégrées sur des périodes suffisamment longues (typiquement plus de $30s$) afin de permettre à l'algorithme d'utiliser les fluctuations du vent vertical à $5s$ au sein de ces périodes pour gérer la turbulence.

3.3.7 Conclusions

On a réalisé dans cette section une expérience jumelle prenant en compte le vent vertical et introduisant l'effet de la turbulence, non modélisée explicitement dans l'algorithme d'assimilation VAMOS, sur les observations réalistes. On tire de cette expérience les

conclusions suivantes :

- VAMOS retrouve les états vrais (DSD, vent vertical) de manière très satisfaisante à la résolution des observations ($2min$ ici).

- il est capable de reproduire les effets de la turbulence (lissage des spectres Doppler) sans altérer la restitution de la DSD et du vent vertical à $2min$. Tout cela est réalisé en faisant fluctuer le vent vertical à $5s$ autour d'une valeur moyenne obtenue sur une fenêtre d'intégration de $2min$, correspondant à la résolution des observations.

- cette manière de représenter la turbulence n'est pas explicitement une discrétisation de l'opérateur de convolution classiquement utilisé pour la représenter (l'algorithme ne comprend pas une inconnue "intensité de la turbulence"). Il est donc a priori difficile d'estimer une "intensité" de la turbulence directement à partir de ces fluctuations haute fréquence du vent. Nous reviendrons sur ces questions sur les données réelles (section 7).

- il est important que le temps d'intégration des observations vaille (au moins) quelques pas de temps du modèle (soit au moins $30s$) pour justement permettre à l'algorithme d'utiliser les fluctuations du vent pour représenter la turbulence.

- l'algorithme échoue à reproduire le nombre de gouttes de très petits diamètres ($< 0.4mm$), gouttes qui ne sont pas observées par les instruments disponibles.

- tous ces résultats demandent bien sûr à présent à être confirmés sur données réelles, où d'autres erreurs viendront s'ajouter (en particulier erreurs modèles dues à la non prise en compte du vent horizontal et de la coalescence/break-up ou erreurs faites en supposant les inconnues et paramètres du modèle constants sur chaque point de maille et en supposant que les instruments voient un point de maille).

4 Influence du vent horizontal

4.1 Présentation

Dans cette section, nous allons étudier l'influence du vent horizontal sur les mesures radar et disdromètres et évaluer en présence de vent horizontal la qualité des restitutions de notre algorithme 4D-VAR VAMOS (qui comme on l'a vu néglige le vent horizontal).

Pour cela, nous adopterons plusieurs point de vue. Dans la section 4.2, nous présenterons une étude théorique simple sur la forme des profils verticaux radar de réflectivité selon le vent horizontal (la signature du vent horizontal sur les mesures de radars à visée verticale). Dans la section 4.3, nous vérifierons numériquement ces résultats théoriques en simulant des observations radar via un modèle numérique de propagation des DSD avec vent horizontal. Puis nous testerons les performances de l'algorithme d'assimilation (sans prise en compte du vent horizontal) sur des observations simulées avec vent horizontal. La section 4.4 conclura cette étude.

4.2 Etude théorique : signature du vent horizontal sur les profils radar

4.2.1 Introduction

Dans cette section, nous allons développer un modèle théorique d'équations différentielles ordinaires très simple visant à simuler la signature de quelques profils de vents horizontaux sur des cartes temps/altitude de mesures de radars à visée verticale.

Plus précisément, nous allons exprimer l'évolution dans le temps de la position dans le plan (x, z) de gouttes de diamètres fixés tombant à leur vitesse terminale de chute et soumises à un vent horizontal représenté par une fonction simple de l'altitude (nulle, linéaire, logarithmique). En particulier, nous supposons qu'il n'y a pas de vent vertical, ni de turbulence, qu'il n'y a pas d'évaporation, de collisions entre gouttes, et que le vent horizontal ne dépend ni du temps, ni de x, la position sur l'axe horizontal. On suppose par ailleurs une cellule ponctuelle de pluie (représentant par exemple l'avancée d'un front), se déplaçant à une vitesse donnée en haut de la colonne atmosphérique et lâchant continument des gouttes. Ayant calculé, pour chaque goutte lâchée, sa position à toutes dates, nous pourrons ensuite :

- à une date donnée, déterminer la position de toutes les gouttes lâchées dans un plan (x, z), puis tracer la courbe correspondant à la position des gouttes lâchées par la cellule.

- Pour une abscisse (un x) donnée, simuler (de manière très simpliste bien sûr) la vision d'un radar placé à cette abscisse en représentant la position selon z (l'altitude) des gouttes présentes à cette abscisse en fonction du temps. Nous comparerons alors rapidement ces résultats théoriques avec des données radar réelles. L'idée est de se donner une base théorique simple permettant ensuite d'estimer visuellement pour une carte de réflectivité donnée l'influence du

vent horizontal sur cet évènement pluvieux et la pertinence qu'il y a à le traiter ou non avec notre algorithme 4D-VAR négligeant le vent horizontal.

4.2.2 Notations / Position du problème

Notons d'abord que vu le nombre assez important de symboles nécessaire pour écrire les équations de cette section, nous avons introduit des grandeurs qui ne seront utilisées nulle part ailleurs. Ces symboles sont donc indépendants du reste du manuscrit.

On suppose que le nuage (ou la bande brillante) est situé à l'altitude z_0. On considère une cellule ponctuelle, située à z_0, qui lâche des gouttes. Elle se déplace à la vitesse horizontale u_{cell}. On suppose que les gouttes sont lâchées à des dates t_{i_0} successives et on fixe l'origine des abscisses ($x = 0$) de manière à ce qu'à $t = 0$, la cellule se trouve en $x = 0$. Considérons à présent une goutte quelconque i. On note x_i et z_i ses positions horizontales et verticales (altitude). Ce sont des fonctions du temps t. On suppose que durant sa chute, la goutte tombe à la vitesse v_{th}, fonction de son diamètre D et de l'altitude selon les expressions d'Atlas et al. (1973) et de Foote and Du Toit (1969). Le diamètre étant fixé (ce n'est pas une variable du problème), on peut exprimer la vitesse v_{th} comme une fonction affine de z de la forme :

$$v_{th}(z) = \beta + \gamma z \tag{V.21}$$

Avec β la vitesse terminale de Atlas et al. (1973) (qui correspond à la vitesse d'une goutte au niveau du sol), et $\gamma = 3.68.10^{-5}\beta$ la correction selon l'altitude de Foote and Du Toit (1969).

Une fois lâchée par la cellule, la goutte a une vitesse horizontale qui est celle du vent horizontal, noté u, et fonction de l'altitude ($u = u(z)$). La goutte i étant par définition lâchée par la cellule à la date t_{i_0}, on a la position initiale de la goutte :

$$\begin{cases} x_i(t_{i_0}) = t_{i_0} u_{cell} \\ z_i(t_{i_0}) = z_0 \end{cases} \tag{V.22}$$

Les vitesses horizontale et verticale de la goutte correspondant bien sûr aux dérivées de sa position, on a les équations différentielles caractérisant la trajectoire de la goutte i (pour $t > t_{i_0}$) :

$$\begin{cases} x_i' = u(z_i) \\ z_i' = -v_{th}(z_i) \end{cases} \tag{V.23}$$

En utilisant l'Eq. (V.21), la deuxième équation différentielle se réécrit :

$$z_i' = -\beta - \gamma z_i \tag{V.24}$$

4.2.3 Résolution

À partir de l'Eq.(V.24), on peut exprimer explicitement $z_i(t)$ par :

$$z_i(t) = Ae^{-\gamma t} + B \tag{V.25}$$

Avec les constantes : $A = (z_0 + \beta/\gamma)\, exp\,(-\gamma t_{i_0})$ et $B = -\beta/\gamma$.

L'équation différentielle pour la position horizontale (cf Eq. (V.23)) devient alors :

$$x_i' = u(Ae^{-\gamma t} + B) \tag{V.26}$$

Exprimer $x_i(t)$ requiert donc, selon l'expression de u, de trouver une primitive de la fonction du temps $u(Ae^{-\gamma t}+B)$. Nous allons à présent examiner trois cas. D'abord un vent horizontal nul, puis un vent horizontal augmentant linéairement avec l'altitude, et enfin un vent augmentant logarithmiquement avec z.

Dans les différents cas de figure suivants, on considère la trajectoire d'une goutte une fois seulement qu'elle a été lâchée (à partir de t_{i_0}). Avant cela, elle se déplace avec la cellule pluvieuse à la vitesse u_{cell}.

Vent horizontal nul On prend ici $u(z) = 0$, qui n'a rien à voir avec la vitesse de la goutte au moment de son lâché (u_{cell}).

Sans vent horizontal, l'Eq.(V.26) se ramène bien sûr à $x_i' = 0$. La goutte ne bouge pas selon x et on obtient une valeur de x_i constante (dépendant uniquement de la date du lâché) :

$$x_i(t) = t_{i_0} u_{cell} \tag{V.27}$$

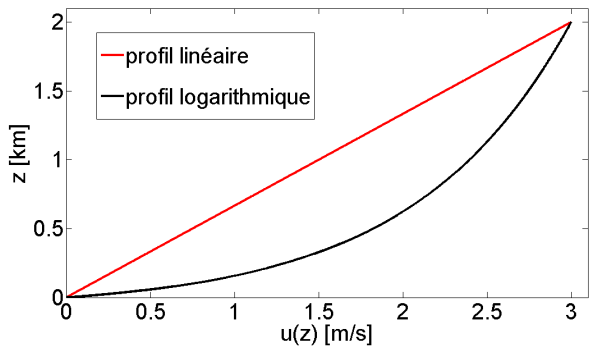

FIGURE V.9 – Profils verticaux de vent horizontaux utilisés dans la section 4.2. Vent nul au sol, de $3m/s$ à l'altitude maximale ($z_0 = 2km$). Rouge : vent horizontal linéaire : $u(z) = 1.5z$. Noir : vent horizontal logarithmique : $u(z) = -4.03 + 0.92ln(80 + 1000z)$. ($z$ en km, u en m/s).

Vent horizontal linéaire On suppose à présent un vent horizontal d'expression $u(z) = \alpha z$. Ce vent augmente linéairement avec z, et vaut 0 au sol. La Figure V.9, courbe rouge, représente ce profil de vent pour $\alpha = 1.5s^{-1}$. L'Eq. (V.26) devient alors :

$$x_i' = \alpha \left(Ae^{-\gamma t} + B \right) \tag{V.28}$$

dont la solution a pour expression :

$$x_i(t) = \frac{-\alpha}{\gamma} Ae^{-\gamma t} + \alpha Bt + cstte_{lin} \tag{V.29}$$

Avec $cstte_{lin}$ une constante (dépendant de la date du lâché t_{i_0}) assurant la condition $x_i(t_{i_0}) = t_{i_0} u_{cell}$.

Vent horizontal logarithmique On suppose enfin un profil logarithmique de vent de la forme :

$$u(z) = a_{log} + b_{log} ln(c_{log} + z) \tag{V.30}$$

La Figure V.9, noir, représente ce profil de vent pour $a_{log} = -4.03 m/s$, $b_{log} = 0.92 S.I.$, $c_{log} = 80m$. On a alors (Eq. (V.26)) :

$$x_i' = a_{log} + b_{log} ln(c_{log} + Ae^{-\gamma t} + B) \tag{V.31}$$

En réorganisant les constantes, on peut réécrire cette expression :

$$x_i' = \tilde{a}_{log} + b_{log} ln(1 + \tilde{c}_{log} e^{-\gamma t}) \tag{V.32}$$

Avec $\tilde{a}_{log} = a_{log} + b_{log} ln\left(B + c_{log}\right)$ et $\tilde{c}_{log} = A / \left(B + c_{log}\right)$.

Cette expression s'intègre en :

$$x_i(t) = \tilde{a}_{log} t + \frac{b_{log}}{\gamma} DL\left(-\tilde{c}_{log} e^{-\gamma t}\right) + cstte_{log} \tag{V.33}$$

Avec $cstte_{log}$ une constante (dépendant de la date du lâché t_{i_0}) assurant la condition $x_i(t_{i_0}) = t_{i_0} u_{cell}$, et DL la fonction (dilogarithme) telle que $DL(t) = \sum_{k=1}^{\infty} t^k / k^2$, $\forall |t| < 1$.

4.2.4 Plan (x, z) à t fixé

Pour une goutte donnée (c'est-à-dire une date de lâché t_{i_0} donnée), nous obtenons donc, selon le profil de vent horizontal, l'expression analytique de sa trajectoire (i.e. sa position au cours du temps). Nous allons à présent fixer la date. (On considère $t = T$). Nous allons alors représenter sur le plan (x, z) la position de toutes les gouttes présentes à cette date T. Nous avons pour chaque goutte lâchée à la date t_{i_0} l'expression analytique selon t de $(x_i(t), z_i(t))$. Nous pouvons également, pour $t = T$, considérer l'expression de $(x_i(T), y_i(T))$ comme une fonction de la date de lâché t_{i_0}. En faisant varier t_{i_0}, nous avons donc une représentation paramétrique de la position des gouttes présentes à T dans le plan (x, z). La Figure V.10 représente les courbes

ainsi obtenues, pour les 3 profils de vent (paramétrisations de la Figure V.9), et pour différents diamètres de gouttes (c'est-à-dire différentes vitesses de chute).

Ici, on considère les gouttes lâchées par une cellule positionnée à $2km$ d'altitude et se déplaçant à la vitesse u_{cell} constante. On voit, pour les différentes configurations, les courbes correspondant aux positions des gouttes lâchées par cette cellule, qui dépendent de la date du lâché et de la vitesse horizontale du vent.

Cette Figure représente donc ce que l'on verrait si l'on photographiait le plan (x, z) dans ces conditions. Elle est donc comparable par exemple aux photographies faites de fronts de pluie.

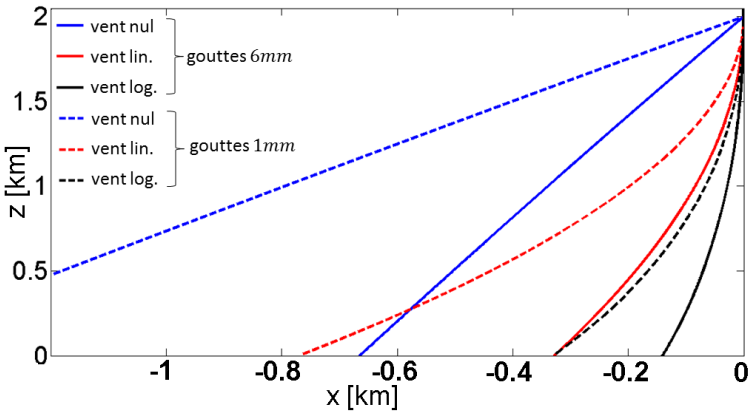

FIGURE V.10 – Pour deux diamètres de gouttes et trois profils de vent horizontal, position dans un plan (x, z) à une date donnée (T) de gouttes lâchées en continu par une cellule de pluie ponctuelle située à l'altitude $z_0 = 2km$. Traits continus : gouttes de diamètre $1mm$; pointillés : $6mm$. Bleu : vent horizontal nul (mais la cellule se déplace tout de même à la vitesse horizontale $u_{cell} = 3m/s$). Rouge : Vent horizontal linéaire (cf Fig. V.9). Noir : Vent horizontal logarithmique (cf Fig. V.9). On place le 0 de l'axe horizontal à la position de la cellule à la date $t = T$.

On remarque par exemple que pour un vent logarithmique et des gouttes de $6mm$ (courbe noire pleine), les gouttes ont pris environ $0.15km$ de retard sur la cellule qui les a lâchées durant leur chute. S'il n'y a pas de vent (courbe pleine bleue), les gouttes durant leur chute n'avancent plus selon x (contrairement à la cellule qui continue d'avancer) et prennent donc plus de retard : environ $0.7km$. Les petites gouttes (pointillés), mettant plus de temps à tomber, prennent plus de retard que les grosses sur la cellule.

4.2.5 Plan (t, z) à x fixé

L'objectif de ce paragraphe est de simuler schématiquement ce que verrait sous ces conditions un radar à visée verticale placé au sol en fonction du temps et de l'altitude. Pour cela, nous commençons par nous donner une abscisse (un X tel que $x = X$) correspondant à la position au sol du radar. Ensuite, nous cherchons quelles gouttes sont présentes au-dessus du radar à une date $t = T$ (c'est-à-dire quelles dates de lâché t_{i_0} permettent d'obtenir des gouttes au-dessus du radar à $t = T$). Cela revient en pratique à chercher l'ensemble des dates de lâché $t_{i_0}*$ se trouvant au-dessus du radar à l'instant T. Ensuite, et si cet ensemble est non

vide, on note z_{i*} l'ensemble des positions correspondant. On fait ensuite varier la date T pour obtenir la série temporelle des gouttes situées à la verticale du radar.

Dans le cas sans vent horizontal, on a $x_i(T) = t_{i_0} u_{cell}$, ce qui fournit naturellement $t_{i*} = X/u_{cell}$. En reportant cette expression dans l'expression de z_i, on obtient une expression analytique pour la courbe recherchée. Dans les autres cas, ne pouvant résoudre explicitement l'équation $x_i(T) = X$, on la résout numériquement et on obtient une suite de points (t, z_{i*}) définissant la courbe recherchée.

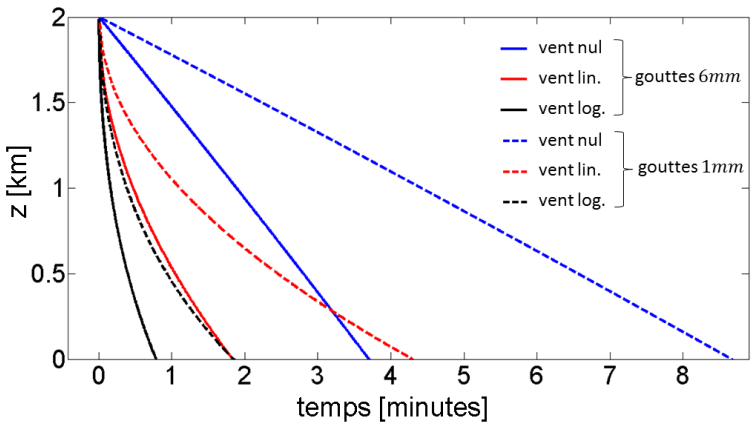

FIGURE V.11 – Simulation théorique simple de la vision qu'un radar en visée verticale aurait de l'arrivée de fronts de pluie, selon la taille des gouttes et le profil de vent horizontal. On représente, pour des dates successives, la position verticale des gouttes situées au-dessus d'une abscisse donnée, pour une cellule de pluie ponctuelle située à l'altitude $z_0 = 2km$ et lâchant continument des gouttes de diamètre fixe. Traits continus : gouttes de diamètre $1mm$; pointillés : $6mm$. Bleu : vent horizontal nul (mais la cellule se déplace tout de même à la vitesse horizontale $u_{cell} = 3m/s$). Rouge : Vent horizontal linéaire (cf Fig. V.9). Noir : Vent horizontal logarithmique (cf Fig. V.9).

La Figure V.11 représente les courbes obtenues dans le plan (t, z) pour les trois profils de vent (paramétrisations de la Figure V.9) et pour différents diamètres de gouttes. Elle représente les points du plan (t, z) desquels un radar recevrait de l'énergie pour une cellule ponctuelle lâchant des gouttes d'un diamètre unique. Elle peut aussi représenter le front d'une cellule non ponctuelle.

On note :

- que dans tous les cas, le radar voit d'abord la pluie à $2km$ d'altitude (à $t = 0$), puis la voit s'abaisser progressivement jusqu'à atteindre le sol. (Les gouttes ont pris du retard durant leur chute sur la cellule qui les a lâchées donc elles arrivent plus tard à la verticale du radar)

- Sans vent horizontal (bleu), les gouttes partent de z_0 puis descendent "diagonalement". Le temps mis pour atteindre le sol (sur la Figure le temps nécessaire à une courbe pour passer de $z = 2km$ à $z = 0$) correspond alors au temps de chute des gouttes : $8min30$ pour $D = 1mm$, $3min40$ pour $D = 6mm$.

À l'autre extrémité (non représentée ici), si le vent horizontal était constant sur toute la colonne, les gouttes se déplaceraient horizontalement à la même vitesse que la cellule de pluie (quelque soit l'altitude), et la pluie arriverait simultanément à toutes les altitudes (on aurait des traits verticaux).

Pour les situations intermédiaires, avec un vent horizontal croissant avec l'altitude, la cellule se déplace tout de même plus vite horizontalement que les gouttes aux altitudes inférieures. On voit donc bien la pluie d'abord en haut de la colonne, puis de plus en plus bas. Le temps mis par la courbe pour atteindre le sol est d'autant plus court que le vent horizontal est homogène sur la colonne.

- On a donc remarqué que sans vent horizontal, le front présente une signature diagonale pour le radar (Figure V.11, bleu). On verra dans la suite que notre algorithme 4D-VAR, qui travaille sans vent horizontal, reproduit effectivement des profils de réflectivités diagonaux. Au contraire, avec des vents linéaires ou logarithmiques, on remarque une signature davantage exponentielle (Figure V.11, rouge et noir).

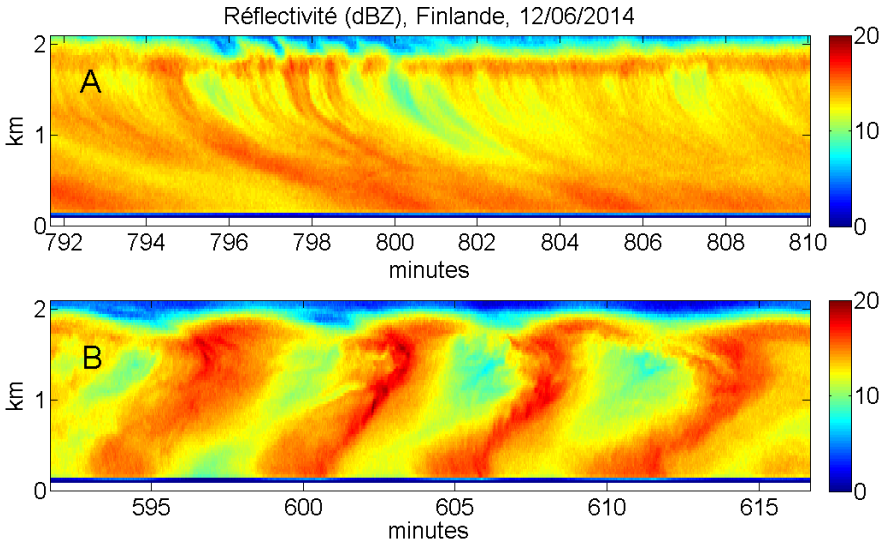

FIGURE V.12 – Réflectivités (dBZ) sur des périodes d'environ $20min$ mesurées par le radar Doppler à visée verticale KAZR (voir chapitre I, section 2.3.3, Tableau I.2) en Finlande le 12/06/2014 (voir section 6 la description de l'évènement de pluie complet correspondant).

La Figure V.12 représente des mesures réelles de réflectivité enregistrées par un radar en bande K en Finlande le 12/06/2014 (qui sera étudié dans la suite en section 6). Dans tous les cas, la bande brillante se trouve vers $z = 1.8km$. On s'intéresse donc seulement aux mesures sous cette altitude.

La première image (Figure V.12A), qui couvre une période temporelle d'environ $20min$, montre une période où l'on voit des signatures se rapprochant assez de celles obtenues théoriquement pour des profils de vent linéaires ou logarithmiques. Le temps mis par les cellules identifiables pour parcourir les $1.8km$ verticaux est d'environ $5min$ (voir par exemple

la cellule présente à $1.8km$ d'altitude vers $795min$ et qui touche le sol vers $800min$). Il s'agit ici de profils de réflectivités, donc influencés par toutes les tailles de gouttes (selon la DSD) mais majoritairement par les plus grosses gouttes. S'agissant néanmoins d'un évènement stratiforme faible (voir section 6), les gouttes $> 2.5mm$ sont quasiment absentes et on peut s'attendre (grossièrement) à avoir une influence majoritaire des gouttes autour de $1.5 - 2mm$. Pour de telles gouttes, l'étude théorique précédente (Figure V.11) indiquerait des temps de "chute" de $1min30$ (profil logarithmique, noir) à $3 - 4min$ (profil linéaire, rouge). On pourrait donc penser pour un tel évènement que l'on a un vent horizontal plus faible en moyenne sur la colonne ou devenant significatif seulement près de la bande brillante. Dans tous les cas, notre algorithme sans vent horizontal, pour lequel, dans un tel cas, on s'attend à avoir des courbes diagonales mettant environ $6min$ à passer de z_0 à 0, semble relativement adapté. On note bien que ces considérations n'ont pas pour ambition de démontrer l'applicabilité de l'algorithme mais seulement de chercher a priori et très qualitativement à anticiper quels évènements seront "traitables" ou non.

La Figure V.12B montre un cas où les échos radar viennent d'abord du sol avant de remonter vers les couches supérieures de l'atmosphère. De tels cas peuvent a priori être expliqués par des gradients négatifs de vent horizontaux (le vent horizontal étant plus fort dans les basses couches, les gouttes prennent de l'avance sur la cellule qui les a lâchées). On remarque néanmoins que près du sol ($< 500m$), les courbes se redressent à nouveau, signe que le vent tend tout de même vers 0 en approchant le sol. Bien entendu, notre algorithme produira dans ce cas-là une erreur tout à fait significative. La section 4.3 suivante reviendra sur ce problème en simulant numériquement de telles observations.

4.3 Données simulées / Erreur modèle due au vent horizontal

4.3.1 Introduction

Dans cette section, nous allons utiliser un modèle numérique de propagation des DSD tenant compte du vent horizontal pour simuler des mesures radars et disdromètres en présence d'un vent horizontal (section 4.3.4). Nous vérifierons ainsi les considérations théoriques de la section 4.2 précédente dans un cadre plus réaliste (vent vertical, propagation de DSD et pas de gouttes individuelles, simulations de réflectivités). Ensuite (section 4.3.5), nous assimilerons ces observations simulées dans VAMOS, notre algorithme 4D-VAR ne prenant pas en compte le vent horizontal. Cela revient mathématiquement à ajouter une erreur modèle (le modèle de propagation de l'algorithme 4D-VAR ne permet pas d'expliquer les observations). On teste ainsi les performances de notre algorithme en présence d'un vent horizontal.

Comme dans la section 3 précédente, nous travaillerons donc ici sur des expériences jumelles (expériences d'assimilation sur données simulées). Le cadre est néanmoins légèrement différent, puisque les observations sont ici simulées avec un modèle différent de celui utilisé dans la phase d'assimilation. La section 4.3.3 reviendra donc rapidement sur le concept d'expériences jumelles

appliqué à ce contexte. Avant cela (section 4.3.2), nous présentons le modèle direct avec vent horizontal et ses paramètres et inconnues.

4.3.2 Modèle direct avec vent horizontal

Nous allons dans cette section utiliser un modèle numérique de propagation des DSD avec vent horizontal. Ce modèle numérique est toujours basé sur l'EDP présentée dans Hu and Srivastava (1995) (Eq. (II.5)). Nous négligeons toujours les effets des collisions entre gouttes. L'objectif ici étant spécifiquement l'étude du vent horizontal et de son influence sur la mesure radar et les restitutions de notre algorithme, nous supposons également nulle l'évaporation (soit un taux d'évaporation $\xi_m = 0$). Le vent horizontal est supposé dirigé dans une unique direction. Nous notons x cette direction horizontale et u la vitesse du vent horizontal. On suppose u constant dans le temps et selon x, et ne dépendant donc que de altitude z. Le problème posé est donc à 4 dimensions (temps t, direction x, altitude z et diamètre D). On parlera de modèle $4D$ dans la suite (par opposition au modèle $3D$ - t, z, D - utilisé par l'algorithme d'assimilation). La DSD n s'écrit donc :

$$n = n(t, x, z, D) \tag{V.34}$$

Et l'EDP régissant son évolution est :

$$\left(\frac{\partial n}{\partial t}\right) = \frac{\partial}{\partial z}\left(\left[v_{th}(z, D) + w(t, x, z)\right] n\right) + u(z)\frac{\partial n}{\partial x} \tag{V.35}$$

Avec toujours v_{th} la vitesse terminale de chute, paramétrisée selon Atlas et al. (1973) et Foote and Du Toit (1969), et w le vent vertical.

Cette EDP est ensuite discrétisée, nous fournissant le modèle numérique de propagation des DSD. On a vu (section 2) que dans le cas sans vent horizontal (et donc dans l'algorithme d'assimilation), le terme en $\partial/\partial z$ était discrétisé selon le schéma de Smolarkiewicz (1983). Ici, les deux termes de l'EDP (en $\partial/\partial x$ et en $\partial/\partial z$) sont discrétisés selon un schéma upstream (pour des questions de simplicité et de temps de calcul). Les pas de discrétisation sont :

$$\begin{cases} \Delta t = 5s \\ \Delta x = 100m \\ \Delta z = 100m \\ \Delta D = 0.1mm \end{cases} \tag{V.36}$$

Comme dans la section 2, nous n'utiliserons dans la suite que des grandeurs discrétisées. L'indice k renvoie toujours au temps (date t_k), l'indice i à l'altitude (couche z_i), et l'indice p au diamètre (classe de diamètre D_p). L'indice m renverra à la position selon l'axe horizontal (position x_m). La vitesse terminale de chute des gouttes est supposée paramétrisée et pleinement connue (paramètre du modèle). Le modèle numérique, pour être exécuté, nécessite

alors :

- une condition initiale $n_{0,i,m}$ $(\forall i, m)$. Comme dans la section 2, on la prendra nulle ($n_{0,i,m} = 0$) en plaçant t_0 avant le début de l'évènement de pluie.
- une condition limite supérieure $n_{k,0,m}$ $(\forall k, m)$
- un champ $3D$ de vents verticaux $w_{k,i,m}$ $(\forall k, i, m)$.
- les valeurs des vents horizontaux selon l'altitude u_i $(\forall i)$.

On note que le temps de calcul qu'induirait potentiellement un tel modèle inclus dans VAMOS ainsi que la difficulté à le contraindre avec les observations disponibles et sans a priori nous ont convaincus pour l'instant de travailler sans dimension horizontale. Cette question reste néanmoins ouverte pour la suite.

4.3.3 Méthodologie utilisée : expériences jumelles

Dans cette section, nous allons utiliser le modèle qui vient d'être décrit (section 4.3.2) pour simuler des observations radar et disdromètres en présence de vent horizontal, puis nous assimilerons ces observations simulées dans notre algorithme d'assimilation 4D-VAR sans vent horizontal. Nous effectuons ainsi une expérience d'assimilation sur données simulées, procédé appelé expérience jumelle. La Figure V.13 reprend la Figure V.4 de la section 3 en l'adaptant au contexte présent. Nous décrivons la procédure utilisée ici pour réaliser une expérience jumelle.

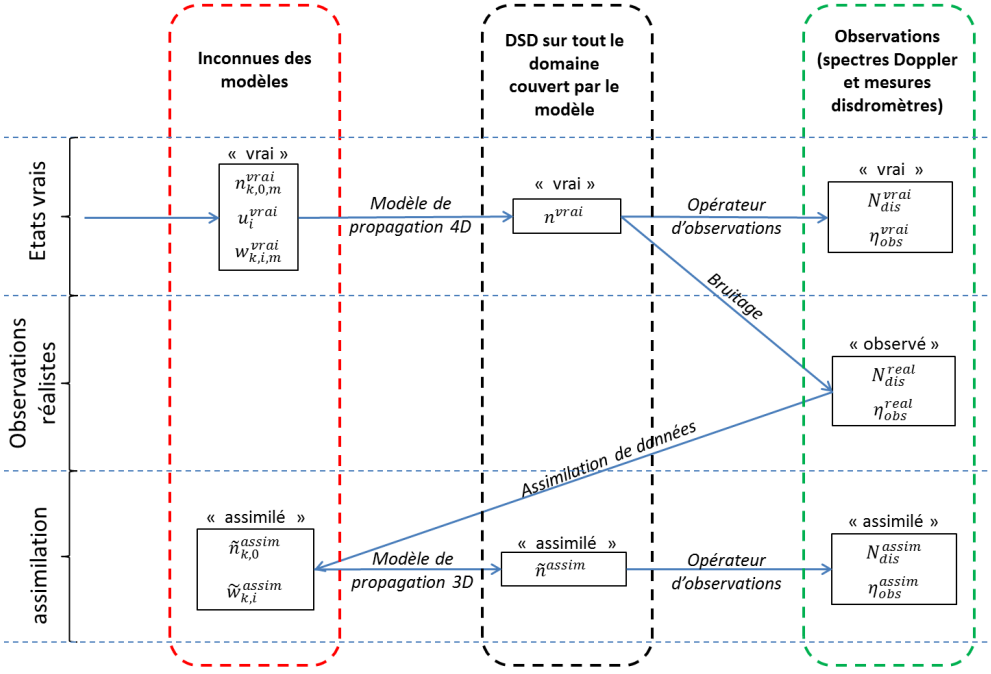

FIGURE V.13 – Description du processus d'expériences jumelles pour la section 4.3. Voir texte pour la description.

Tout d'abord, nous fixons (simulons) les inconnues de notre modèle de propagation avec vent horizontal ($n_{k,0,m}$, $w_{k,i,m}$ et u_i, voir section 4.3.4). Avec ce modèle, on propage ensuite les

DSD à toutes dates et toutes altitudes. On obtient ainsi $n_{k,i,m}$. On suppose ensuite que l'on a placé un radar et un disdromètre à la position x_{mr} selon l'axe horizontal. On utilise alors $n_{k,i,mr}$, $w_{k,i,mr}$ et les opérateurs d'observation (décrits dans la section 2.3) pour simuler des observations de disdromètres (N_{dis}) et de spectres Doppler (η_v). Ces états sont les états vrais du problème, désignés par l'exposant $vrai$. L'objectif de l'algorithme d'assimilation sera ensuite de restituer au mieux ces états vrais à la verticale du radar (soit restituer $n^{vrai}_{k,i,mr}$ et $w^{vrai}_{k,i,mr}$).

Ces observations vraies (N^{vrai}_{dis}, η^{vrai}_{obs}) sont ensuite bruitées pour être rendues plus réalistes. On obtient ainsi les grandeurs réalistes (exposant $real$). Elles constituent l'information qui sera utilisée par l'algorithme d'assimilation.

En appliquant VAMOS, on obtient les inconnues du modèle assimilées. On note bien que dans le cas présent, le modèle de propagation utilisé pour créer les états vrais ($4D$, avec vent horizontal) n'est pas le même que celui utilisé dans la phase d'assimilation ($3D$, sans vent horizontal). Les dimensions étant différentes entre états vrais et assimilés, nous ajoutons un tilde pour désigner les grandeurs assimilées. L'assimilation fournit une DSD en haut de la colonne ($\tilde{n}_{k,0}$) et un vent vertical à toutes dates et tous instants ($\tilde{w}_{k,i}$). Propagée verticalement par le modèle de propagation, on obtient la DSD à toutes dates. En appliquant les opérateurs d'observation, on obtient les observations assimilées. On désigne ces grandeurs assimilées par l'exposant $assim$. Dans ce travail, nous comparerons les DSD à la verticale des instruments ($n^{vrai}_{k,i,mr}$ et $\tilde{n}^{assim}_{k,0}$, mr la position du radar), les vents verticaux ($w^{vrai}_{k,i,mr}$ et $\tilde{w}^{assim}_{k,i}$) et les observations (N^{vrai}_{dis} et N^{assim}_{dis} face à η^{vrai}_{v} et η^{assim}_{v}).

4.3.4 Etats vrais simulés

Dans cette section, nous présentons les états vrais simulés. Ceux-ci comprennent d'une part les inconnues vraies du modèle de propagation avec vent horizontal (DSD en haut de la colonne $n^{vrai}_{k,0,m}$, vent vertical $w^{vrai}_{k,i,m}$ et profil de vent horizontal u^{vrai}_{i}, pré-requis du modèle de propagation), et d'autre part les états vrais qui en sont déduits par propagation dans le modèle (DSD partout n^{vrai}) ainsi que les observations radar et disdromètres simulées via les opérateurs d'observation (N^{vrai}_{dis} et η^{vrai}_{v}).

Inconnues simulées du modèle 4D Nous allons d'abord présenter les choix effectués pour fixer les inconnues du modèle de propagation avec vent horizontal. Il s'agit ici d'expliquer le procédé suivi pour obtenir des inconnues vraies réalistes. La Figure V.14 représente ces inconnues.

Pour simuler des profils de vent horizontal réalistes par rapport aux cas de pluies traités dans ce travail, nous utilisons les données de deux radiosondages réalisés en Finlande le 12/06/2014 (voir section 6 les détails sur l'évènement de pluie correspondant), ainsi qu'un produit fusionnant des données obtenues par radiosondages et des sorties de modèle numérique de prévision du temps pour Darwin, Australie, le 04/11/2014 (voir section 7). Nous obtenons ainsi trois profils temps/altitude de vents horizontaux. Deux d'entre eux, présentant des fluctuations non monotones sur les deux premiers kilomètres de la colonne,

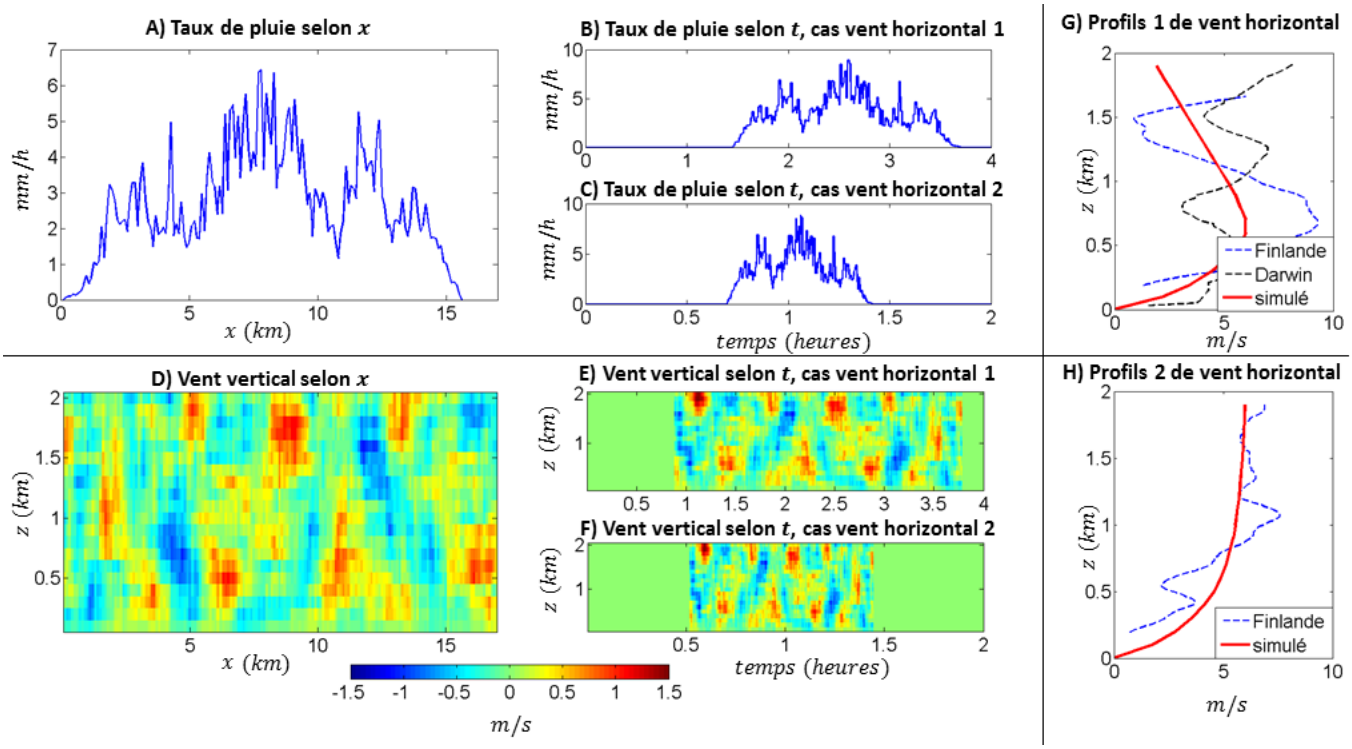

FIGURE V.14 – États vrais simulés nécessaires à l'exécution du modèle 4D de propagation des DSD. A) Taux de pluie en haut de la colonne atmosphérique à l'instant initial selon la position horizontale. B) Taux de pluie selon la date à la verticale des instruments de mesure, correspondant à la cellule du A propagée par le profil 1 de vent horizontal (G, rouge). C) Idem, profil 2 de vent horizontal (H, rouge). D) Cellules de vent vertical à l'instant initial selon l'altitude et la position horizontale. E) Vent vertical selon la date et l'altitude à la verticale des instruments de mesure, correspondant à la cellule du D propagée par le profil 1 de vent horizontal (G, rouge). Positif vers le bas. F) Idem, propagation par le profil 2 de vent horizontal (H, rouge). G) Vent horizontal selon l'altitude tel que : pointillés bleus : mesuré par radiosondage en Finlande le 12/06/2014, 11:36 ; pointillés noirs : estimé par fusion de radiosondages et de sorties de modèle à Darwin, Australie, le 04/11/2014, moyenné sur $17h$-$21h$; continu rouge : simulé dans le cadre de ce travail ("profil 1"). H) Idem G tel que : pointillés bleus : mesuré par radiosondage en Finlande le 12/06/2014, 17:23 ; continu rouge : simulé dans le cadre de ce travail ("profil 2"). Voir aussi le texte pour la description de ces champs.

sont présentés sur la Figure V.14G (pointillés). Le troisième, présentant un vent horizontal augmentant régulièrement avec l'altitude, est présenté sur la Figure V.14H (pointillés). Nous simulons ensuite deux profils standards simples correspondant plus ou moins aux profils mesurés (Figure V.14G et H, rouge), le premier monotone (H, profil 1), le second augmentant rapidement sur les 500 premiers mètres de la colonne puis décroissant ensuite (G, profil 2). Dans tous les cas, le vent horizontal est nul au niveau du sol et ne dépasse pas $6m/s$.

Pour la DSD en haut de la colonne, nous utilisons une série temporelle de taux de pluie à $1min$ mesurée par disdromètre en Finlande en 2014 (voir section 6). Cette série temporelle est ici transposée en taux de pluie en fonction de la position sur un axe horizontal couvrant $15km$ (cellule de pluie de $15km$). Ce taux est représenté sur la Figure V.14A. Ce taux de pluie est

ensuite transporté dans le temps et selon x à la vitesse horizontale en haut de la colonne correspondant au profil de vent 1 (V.14G) ou 2 (V.14H). On obtient ainsi, pour chaque profil de vent horizontal, un taux de pluie en haut de la colonne à toutes dates et toutes positions selon l'axe horizontal x. En utilisant la relation de Marshall and Palmer (1948), on en déduit une DSD en haut de la colonne $n^{vrai}_{k,0,m}$ (exponentielle). En supposant le radar et le disdromètre positionnés en un certain point de l'axe horizontal, on peut représenter le taux de pluie en haut de la colonne à la verticale de ce point en fonction du temps. C'est ce qui est présenté sur la Figure V.14B, pour le profil de vent horizontal 1 et sur la Figure V.14C pour le profil de vent horizontal 2. Le vent en haut de la colonne étant plus fort pour le profil 2 ($6m/s$) que pour le profil 1 ($2m/s$), la cellule de pluie est nettement plus contractée avec ce profil (la même cellule passe au-dessus d'un point plus rapidement).

Pour le vent vertical, on procède de manière similaire. On commence par simuler une cellule (x, z) de vent vertical. Celle-ci est représentée sur la Figure V.14D. Pour cela, on simule un champ de variables gaussiennes centrées réduites indépendantes. Puis on lisse (itérations de moyennes locales) ce champ jusqu'à obtenir des longueurs de corrélations d'environ $400m$ en altitude et $700m$ en horizontal. Finalement, on recale ce champ pour obtenir des valeurs dans la gamme $[-1.5 : 1.5]m/s$. Cette cellule de vent vertical est ensuite propagée via le modèle pour les profils de vent horizontal 1 et 2. On obtient bien ainsi des valeurs pour w à toutes dates, altitudes, et positions horizontales. On a représenté les champs temps/altitudes de vents horizontaux à la verticale du radar et du disdromètre sur les Figure V.14E (profil 1 pour le vent horizontal) et Figure V.14F (profil 2 pour le vent horizontal). On obtient des champs temps/altitudes de vent verticaux de temps de corrélation de 3 à $15min$ (selon le vent horizontal), correspondant bien, par exemple, aux restitutions de Giangrande et al. (2010) ou de Williams (2002).

Observations simulées Les DSD exponentielles en haut de la colonne sont ensuite propagées à toutes dates et altitudes par le modèle numérique grâce au champ de vent vertical et aux profils de vents horizontaux. On peut alors simuler ce que voient le radar et le disdromètre. Dans ce paragraphe, nous allons étudier rapidement ces grandeurs "observées".

La Figure V.15 représente, pour le profil 1 de vent horizontal (Figure V.14G), le nombre de gouttes au-dessus des instruments de mesure en fonction du temps pour trois classes de diamètres (voir légende). Le profil 1 de vent horizontal présente un vent nul au sol, puis un maximum à $600m$ d'altitude (à $6m/s$), puis une décroissance linéaire jusqu'au sommet de la colonne (environ $2m/s$ à $2km$ d'altitude). Vue la discussion de la section 4.2.5 sur un modèle théorique simple, on s'attendait dans ce cas de figure là à voir d'abord, à la verticale du radar, des gouttes là où le vent horizontal est maximal (vers $600m$) et seulement plus tard en haut de la colonne (un peu comme sur la Figure réelle V.12, bas). C'est bien ce que l'on constate ici : on voit d'abord des gouttes un peu au-dessus du sol, et seulement plus tard en haut de la colonne. Cet effet est d'autant plus marqué que la vitesse terminale de chute (et donc le

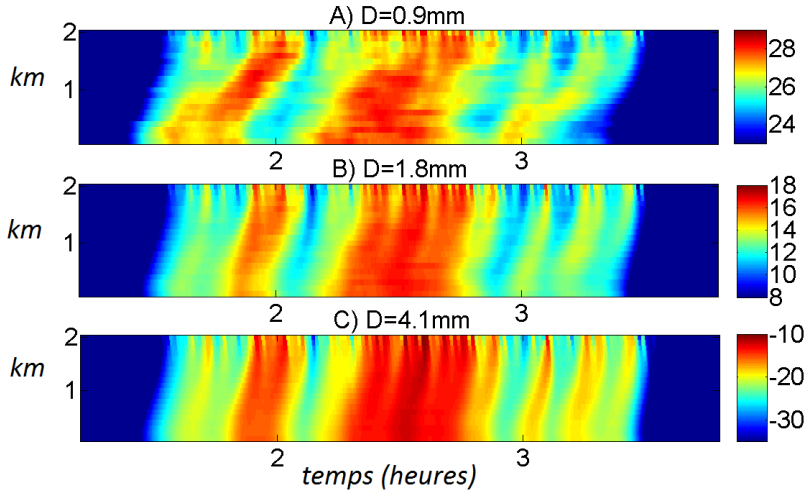

FIGURE V.15 – Pour le profil 1 de vent horizontal : Nombre (logarithmique) de gouttes au-dessus du radar simulé, pour l'état vrai (simulé) et pour trois classes de diamètres : A) $0.9-1.0mm$; B) $1.8-1.9mm$; C) $4.1-4.2mm$. Formellement, on représente ici $n^{vrai}_{k,i,mr,p}$, $\forall k$ (temps), $\forall i$ (altitude), pour un mr arbitrairement fixé (position du radar) et pour trois p (classes de diamètres) différents. Unité : $10log_{10}\,(m^{-1}mm^{-1})$.

diamètre des gouttes) est faible (Figure V.15A contre V.15C). On note également que si les structures verticales sont très claires, bien régulières et délimitées pour les grosses gouttes (C) et dans une moindre mesure pour les moyennes (B), le champ présentant les résultats pour les petites gouttes (A) semble nettement plus perturbé. Cela est dû à l'effet du vent vertical, qui module la vitesse de chute des gouttes (vent vertical au-dessus du radar représenté sur la Figure V.14E). Celles-ci sont d'autant plus affectées, proportionnellement, qu'elles tombent lentement, et donc qu'elles sont petites.

Nous allons maintenant voir ce que l'on obtient lorsque l'on simule des observations radar (η^{vrai}_{obs}) à partir des sorties du modèle de propagation (n^{vrai} et w^{vrai}). On rappelle que l'opérateur d'observation pour les spectres Doppler a été décrit en section 2.3.2. La Figure V.16A présente le facteur de réflectivité radar équivalent (en dBZ) vu par notre radar simulé en fonction du temps et de l'altitude. La Figure V.17A présente elle la vitesse Doppler moyenne (m/s) (voir légendes). (Les champs assimilés (B) seront commentés dans le paragraphe suivant).

Sur le champ de réflectivités (Figure V.16A), on remarque bien également les structures (du bas vers le haut) vues sur la Figure V.15 (DSD). En prenant un profil de vent horizontal décroissant avec l'altitude, on a donc bien réussi à simuler des observations comparables à ce que l'on avait observé en Finlande (Figure V.12A). Sur le champ de vitesses Doppler moyennes, en revanche (Figure V.17A), on remarque que l'on ne retrouve pas vraiment ces structures. Les structures de ce champ semblent en revanche davantage correspondre à celles du champ de champ des vents verticaux au-dessus du radar (Figure V.14E) (par exemple : maximum de W à environ $8m/s$ vers $t = 2.5h$ / $z = 1.5 - 2km$ correspondant à un maximum de vent vertical à environ $+1.5m/s$; et minimum à $5m/s$ vers $t = 3.2h$ / $z = 0 - 1km$ correspondant à un minimum de vent vertical à $-1m/s$. Ce résultat était assez attendu dans la mesure où le vent

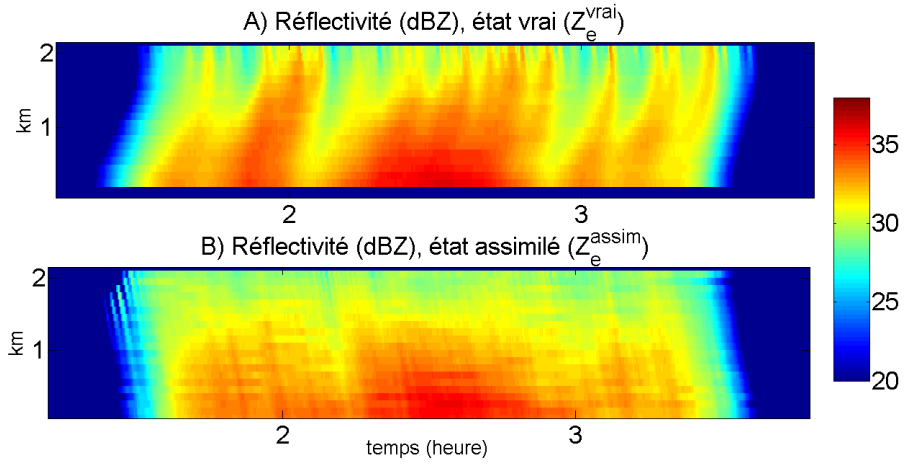

FIGURE V.16 – Pour le profil 1 de vent horizontal : facteur de réflectivité radar équivalent (proportionnel à la réflectivité η, moment d'ordre 0 de la réflectivité spectrale η_v, voir Eq.(I.18)) au-dessus du radar simulé. A) État vrai. B) État assimilé sans vent horizontal. Unité : dBZ.

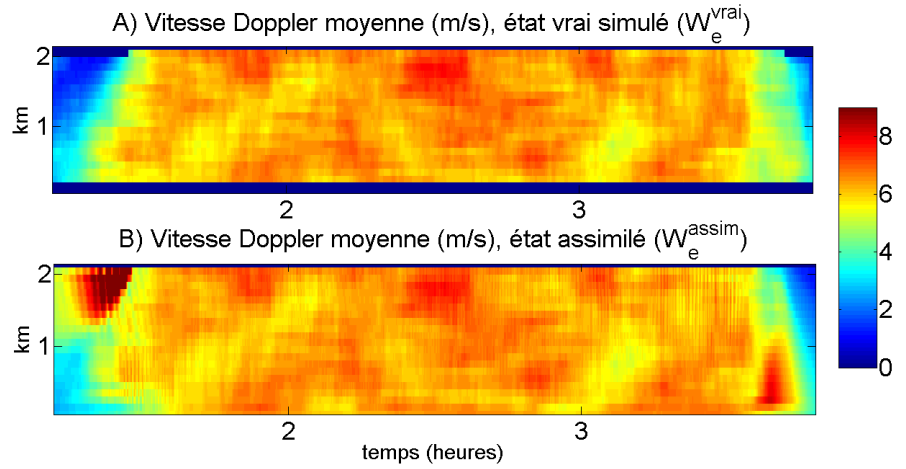

FIGURE V.17 – Pour le profil 1 de vent horizontal : Vitesse Doppler moyenne (moment centré d'ordre 1 de la réflectivité spectrale η_v, voir Eq.(I.31)) au-dessus du radar simulé. A) État vrai. B) État assimilé sans vent horizontal. Unité : m/s, positif vers le bas.

vertical agit directement sur les spectres Doppler en les translatant selon la vitesse et donc module directement W également (W étant l'espérance du spectre Doppler). W est cependant également affecté par l'altitude (la vitesse terminale des gouttes augmente avec l'altitude) et par la forme de la DSD. On peut donc s'attendre à ce que W soit majoritairement déterminé par le vent vertical sous réserve que celui-ci soit suffisamment fort.

Nous pouvons également regarder les simulations d'observations disdromètres (N_{dis}^{vrai}) à partir des sorties du modèle de propagation (n^{vrai}) (opérateur d'observation décrit en section 2.3.1). La Figure V.18A présente le nombre de gouttes (flux de gouttes) enregistré par le disdromètre en fonction de leur diamètre et du temps. On note en particulier qu'avec le profil 1 de vent vertical, les petites gouttes arrivent au sol avant les grosses. Comme elles tombent moins vite que les grosses, elles ont parcouru une distance horizontale plus grande durant leur chute (elles ont plus de temps pour prendre de l'avance sur la cellule qui les a lâchées). Il sera

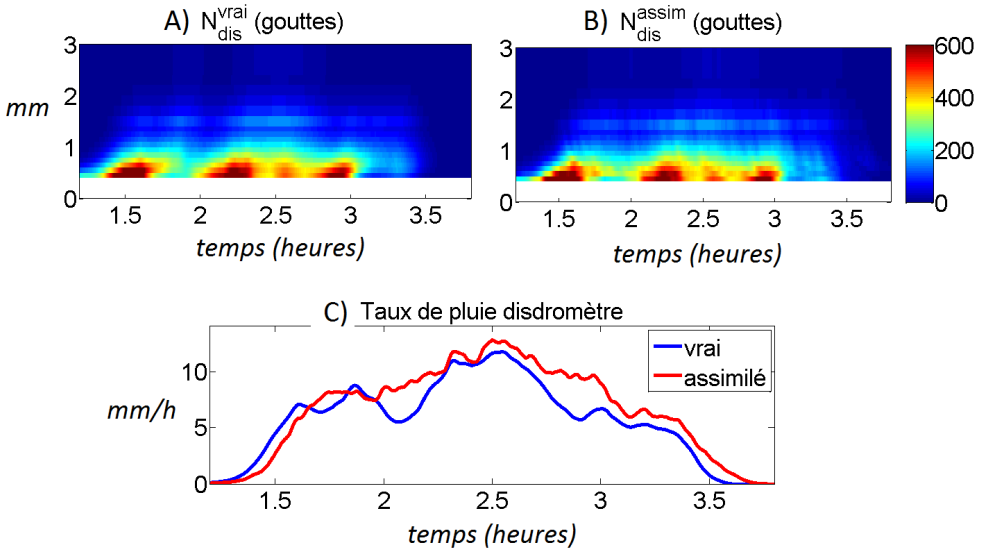

FIGURE V.18 – Pour le profil 1 de vent horizontal : A-B) nombre de gouttes en fonction du temps et du diamètre vues par le disdromètre simulé : A) État vrai. B) État assimilé. C) Série temporelle des taux de pluie correspondant : bleu : état vrai ; rouge : état assimilé.

intéressant de regarder sur les mesures disdromètres réelles si on retrouve ces structures dans les cas où l'on suspecte un gradient négatif de vent horizontal (typiquement le cas de la Figure V.12B).

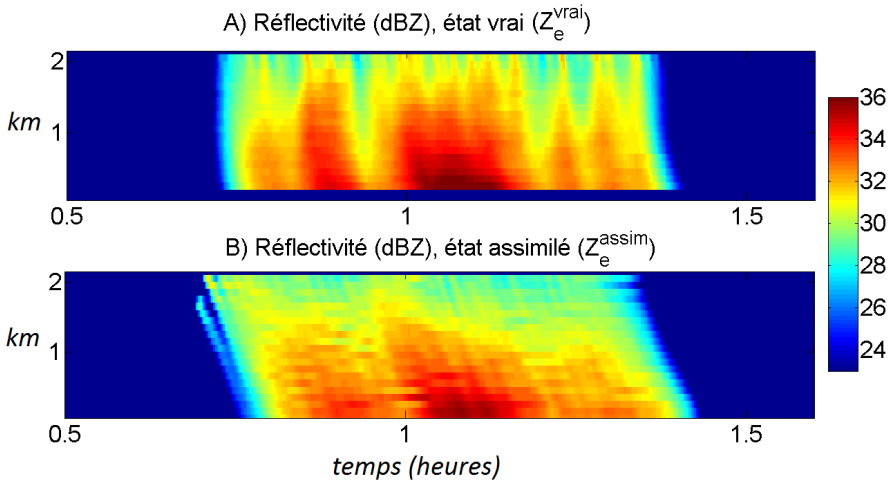

FIGURE V.19 – Pour le profil 2 de vent horizontal : idem Figure V.16.

Les Figures V.19A et V.20A présentent également Z_e^{vrai} et W^{vrai}, cette fois-ci pour le profil 2 de vent horizontal (vent croissant avec l'altitude). On retrouve cette fois-ci bien sur le champ de réflectivité les structures simulées précédemment dans un cas simple et observées en Finlande (section 4.2.5, Figures V.11 et V.12, haut). Comme dans le cas précédent, le champ de vitesses Doppler moyennes (Figure V.20) est principalement déterminé par le vent vertical (représenté Figure V.14F).

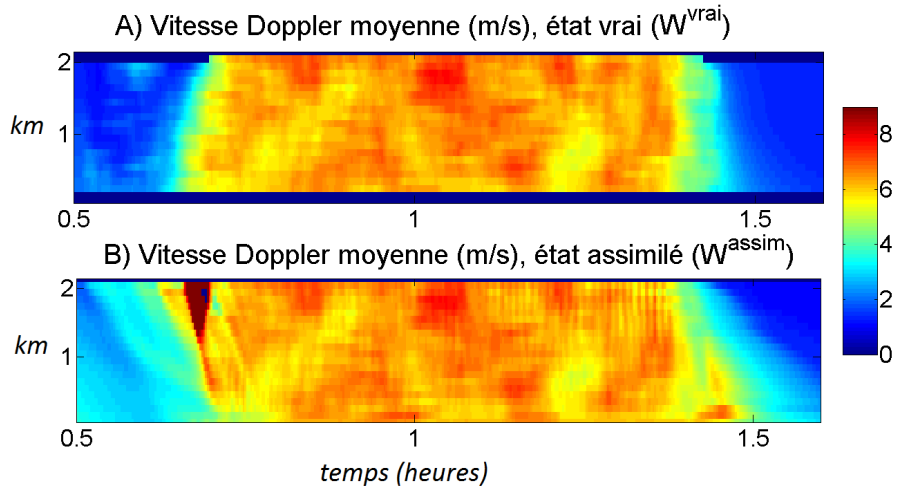

FIGURE V.20 – Pour le profil 2 de vent horizontal : idem Figure V.17.

4.3.5 Etats assimilés / 4D-VAR avec modèle sans vent horizontal

Nous allons à présent assimiler les observations précédemment simulées dans notre algorithme VAMOS ne prenant pas en compte le vent horizontal. L'objectif est d'étudier sa réponse à cette erreur modèle, et de voir s'il peut tout de même nous fournir des restitutions pertinentes de la DSD et du vent vertical.

On va discuter plus en détails les Figures introduites dans la section précédente. Les Figures V.16B et V.19B présentent les facteurs de réflectivité radar équivalents (sortie de l'algorithme 4D-VAR), pour les profils de vent horizontal 1 et 2. Elles sont bien sûr à comparer aux Figures V.16A et V.19A qui présentaient la grandeur vraie correspondante.

La Figure V.21A présente, selon l'altitude, et pour les deux profils de vent horizontal, les Z_e moyens, pour les états vrais et assimilés. Les Figures V.21D et E présentent les DSD moyennes en haut et en bas de la colonne, là encore pour les états vrais et assimilés. Enfin, la Figure V.18B présente le nombre de gouttes observées par le disdromètre, pour l'état assimilé, et la Figure V.18C, rouge, le taux de pluie correspondant.

On constate sur les Figures V.16B et V.19B qu'en sortie d'assimilation, les champs de réflectivité présentent bien des structures diagonales. Cela est dû à l'absence de vent horizontal dans le modèle (voir les résultats théoriques de la section 4.2.5). Les champs vrais présentant des structures paraboliques (profil 2, Figure V.19A) ou inversées dans le temps (profil 1, Figure V.16A), l'algorithme d'assimilation ajuste au mieux dans le temps ces structures diagonales pour minimiser l'erreur en moyenne, en restant au plus près des états vrais.

On remarque par ailleurs qu'en moyenne, les DSD sont bien restituées sur les gouttes $< 2mm$ (très majoritaires pour un évènement de plus assez faible comme celui-ci), mais semblent souffrir d'un léger biais positif sur les plus gros diamètres (que ce soit en haut (Figure V.21 D) ou en bas (Figure V.21 E) de la colonne). Ce phénomène n'est pas dû à un ajustement sur les données des

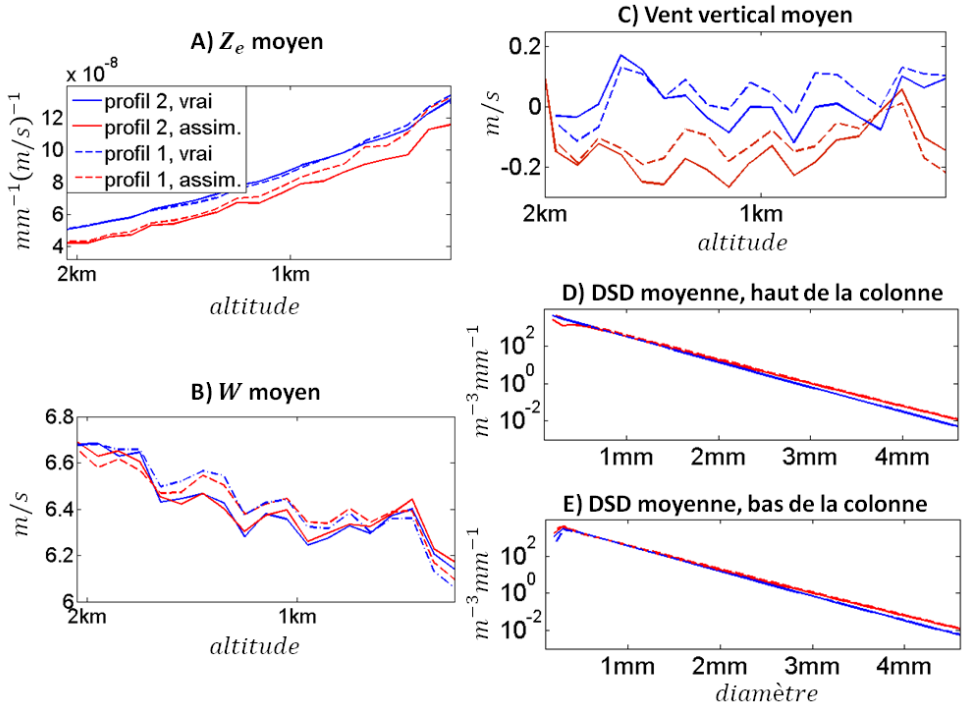

FIGURE V.21 – Comportement moyen durant l'évènement de plusieurs grandeurs vraies (bleu) et assimilées (rouge), pour les simulations avec profil de vent horizontal 1 (traits pointillés) ou 2 (traits pleins). On représente : A) le facteur de réflectivité moyen selon l'altitude. B) Le Doppler moyen selon l'altitude. C) Le vent vertical moyen selon l'altitude. D-E) La DSD moyenne respectivement en haut ($z = 2km$) et en bas de la colonne ($z = 0$). Les moyennes temporelles sont calculées sur le coeur des évènements, soit entre les heures 0.83 et 1.28 pour le profil de vent horizontal 1, et entre les heures 2.08 et 3.19 pour le 2.

disdromètres (Figure V.18C), dans la mesure où l'on constate là aussi une légère surestimation du taux de pluie.

Une explication possible à ce biais est que la présence de vent horizontal entraîne une arrivée des gouttes à la verticale du radar plus rapprochée dans le temps entre les différentes altitudes que ce que peut reproduire l'algorithme d'assimilation sans vent horizontal. On voit cela par exemple sur la Figure V.19 : les pentes moyennes des structures sur le champ vrai (A) sont plus fortes (en valeur absolue) que les pentes des diagonales de l'état assimilé (B). Or, les pentes de ces diagonales, sur un champ de réflectivité, seront d'autant plus fortes que la DSD est composée de grosses gouttes (qui tombent plus vite). On peut donc penser que l'algorithme 4D-VAR augmente artificiellement ces pentes en surestimant le nombre de grosses gouttes (mais pas trop, sinon cela ne correspondrait plus notamment aux observations au sol).

On note néanmoins que ce biais positif n'est pas critique, la surestimation des taux de pluie au sol, par exemple, étant généralement autour de +10%.

On constate également sur la Figure V.18C (taux de pluie au sol), que les séries temporelles de pluie assimilées sont en avance par rapport à l'état vrai. Là encore, cela est dû au fait que sans vent horizontal, les gouttes arrivent plus tard au sol qu'avec un vent

horizontal. L'algorithme ajuste en moyenne les temps d'arrivée sur la colonne. Les séries temporelles assimilées sont donc en avance en haut de la colonne, en retard en bas (assez visible également sur la Figure V.16 par exemple).

Intéressons nous à présent aux vitesses Doppler moyennes et au vent vertical. Les Figures V.17B et V.20B présentent la vitesse Doppler moyenne assimilée (sortie de l'algorithme 4D-VAR), pour les profils de vent horizontal 1 et 2. Elle sont bien sûr à comparer aux Figures V.17A et V.20A qui présentent la grandeur vraie correspondant. La Figure V.21B présente, selon l'altitude, et pour les deux profils de vent horizontal, les W moyens, pour les états vrais et assimilés. La Figure V.21C présente les vents verticaux moyens vrais et assimilés, toujours selon l'altitude. La Figure V.22, enfin, présente les champs de vent verticaux à la verticale du radar, vrais et assimilés, pour les deux profils de vents horizontaux.

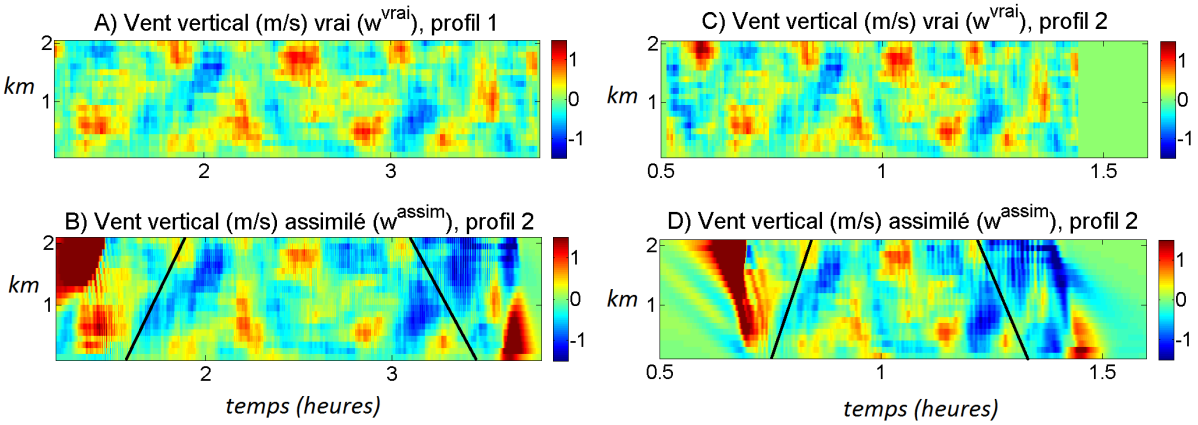

FIGURE V.22 – Vent vertical au-dessus du radar simulé, états vrais et assimilés, pour les deux profils de vent. Profil de vent horizontal 1, vent vertical vrai (A) et assimilé (B). Profil de vent horizontal 2, vent vertical vrai (C) et assimilé (D). Sur les représentation des vents verticaux assimilés, les deux traits noirs délimitent la zone où la structure des champs vrais est retrouvée. Le vent est positif vers le bas.

On constate (Figures V.17, V.20 et V.21B) que les champs de vitesses Doppler moyennes semblent très bien restitués, tant en moyenne temporelle qu'en variabilité. On ne constate pas de biais à ce niveau-là. Le vent vertical moyen est en revanche sous-estimé quelque soit l'altitude (Figure V.21C), là encore assez faiblement (de l'ordre de $0.2m/s$). Ceci est vraisemblablement dû aux mêmes causes que précédemment. La DSD étant légèrement trop lourde sur les grosses gouttes, les spectres "hydrologiques" (sans prise en compte du vent vertical) correspondant sont également légèrement trop lourds sur les grandes vitesses (les gros diamètres). La vitesse Doppler moyenne de ces spectres hydrologiques (leur espérance) est donc également légèrement trop forte. Le vent vertical est alors utilisé par l'algorithme pour compenser ce léger biais.

On constate (Figure V.22) néanmoins que les structures des champs vrais de vent verticaux (pour les 2 profils de vents horizontaux) sont bien reproduites dans le coeur de l'évènement (zone comprise entre les cadres noirs sur la Figure V.22). Le biais évoqué ($0.2m/s$) est en effet faible comparé aux valeurs caractéristiques de vent ($1.5m/s$). La capacité de l'algorithme à retrouver

les structures de vent vertical malgré des erreurs très conséquentes sur le vent horizontal est un résultat très positif sur sa robustesse. On note néanmoins également :

- que l'algorithme produit des résultats faux sur les périodes de transition (début et fin d'évènement, extérieur des cadres noirs de la Figure V.22).

- a tendance à restituer des champs de vents verticaux davantage bruités que les champs vrais. Ce dernier point n'est pas directement dû à l'influence du vent horizontal. Il s'agit d'un problème de minimisation numérique qui est évoqué rapidement dans l'annexe J.

4.4 Conclusions

Dans cette section, nous avons tout d'abord bâti un modèle théorique très simple permettant d'estimer la signature de quelques profils de vents horizontaux sur les cartes temps/altitude de réflectivités fournies par les radars Doppler à visée verticale. L'objectif était de pouvoir évaluer grossièrement et rapidement sur de telles cartes si l'influence du vent horizontal était ou non prépondérante pour l'évènement de pluie correspondant. Nous avons ainsi vu que des structures apparaissant sur les cartes de réflectivités présentant des pentes fortement négatives et non diagonales pouvaient être le signe d'un vent horizontal fort. Au contraire, des structures diagonales peuvent correspondre à un vent horizontal faible ou nul.

Nous avons ensuite utilisé un modèle de propagation des DSD à une dimension horizontale pour simuler des spectres Doppler et des mesures disdromètres pour 2 profils de vents horizontaux. Nous avons ainsi pu confirmer les résultats théoriques précédents. Puis nous avons assimilé ces observations simulées dans notre algorithme d'assimilation 4D-VAR VAMOS sans vent horizontal. Ceci nous a permis de montrer sa robustesse et sa capacité à retrouver les principales structures de vent vertical et de DSD malgré l'erreur commise en négligeant le vent horizontal. Nous avons néanmoins constaté l'inaptitude du modèle à reproduire les phases de transitions (début / fin d'évènement). Les valeurs obtenues présentent, de plus, un léger biais, tendant à surestimer les queues des DSD (et donc les taux de pluie, typiquement de 10%) et à dévier légèrement les vents verticaux vers les valeurs négatives (c'est-à-dire vers le haut) (typiquement de $-0.2m/s$).

5 Première cas d'étude : HyMex

5.1 Présentation et enjeux

Nous allons dans cette section nous concentrer sur un premier cas d'étude. Celui-ci a eu lieu en France, en Ardèche, à l'automne 2013. Il s'agit d'un évènement de pluie faible, avec a priori assez peu d'évaporation et de mouvements horizontaux ainsi que des taux de pluie au sol assez réguliers (voir section 5.2.2). Il semble donc particulièrement approprié pour tester VAMOS, notre algorithme de restitution, dans un contexte assez simple. Nous étudierons ainsi son comportement sur un premier cas réel et les structures des champs de DSD et de vents verticaux restitués. Cette section reprend globalement les résultats présentés dans l'article Mercier et al. (2015b) (fourni en Annexe M). Les deux instruments utilisés pour fournir les données d'entrée (observations) sont le MRR (Micro-Rain Radar) et le DBS (Dual Beam Spectropluviometer) (voir chapitre I ou section 5.2.1 ci-dessous). Le premier est un radar dédié à l'étude des précipitations. Il est fourni avec un logiciel inversant rapidement les spectres Doppler pour fournir des estimations de DSD (notamment sans prendre en compte le vent vertical). Il sera donc intéressant de comparer nos restitutions de DSD avec ce produit.

On note cependant d'ores et déjà que l'on ne dispose pas d'autres instruments (radars, disdromètres) susceptibles de permettre une validation indépendante des résultats (toutes les données disponibles sont nécessaires au bon fonctionnement de notre algorithme d'assimilation). Cette étape de validation sera effectuée ultérieurement, sur un autre cas d'étude, en section 6. Nous nous contentons ici d'évaluer la capacité de VAMOS à reproduire les observations disponibles et la cohérence des champs restitués comparativement à d'autres études effectuées dans des contextes météorologiques similaires.

Dans un premier temps (section 5.2), nous présenterons les instruments et produits disponibles. Nous introduisons aussi l'évènement considéré, le site où il a été enregistré, ses caractéristiques, ainsi que les choix effectués pour l'assimilation (par exemple de négliger l'évaporation). Ensuite, nous évaluerons la capacité de notre algorithme ainsi paramétré à reproduire les observations (section 5.3). Par la suite, nous nous concentrerons plus spécifiquement sur le vent vertical restitué (section 5.4). Enfin, dans la section 5.5, nous comparerons les DSD restituées par notre algorithme VAMOS à celles fournies par le produit MRR. Nous étudierons plus en détails la structure de nos champs de DSD. La section 5.6 conclura et introduira l'utilisation de jeux de données plus étoffés faite dans la suite.

5.2 Instruments disponibles / Présentation de l'évènement

Dans cette section, nous traitons donc un premier cas d'étude ayant eu lieu dans le sud-est de la France, en Ardèche, à l'automne 2013, dans le cadre de la campagne HyMeX (HYdrological cycle in the Mediterranean EXperiment, voir par exemple Bousquet et al. (2015) ou le site internet http://www.hymex.org/). La Figure V.23 présente une carte

topographique de la région étudiée et montre l'emplacement du Pradel, localité où étaient situés les instruments de mesure utilisés dans cette étude. La région et ses caractéristiques géographiques et météorologiques ont déjà été évoquées dans le chapitre IV (section 2.1).

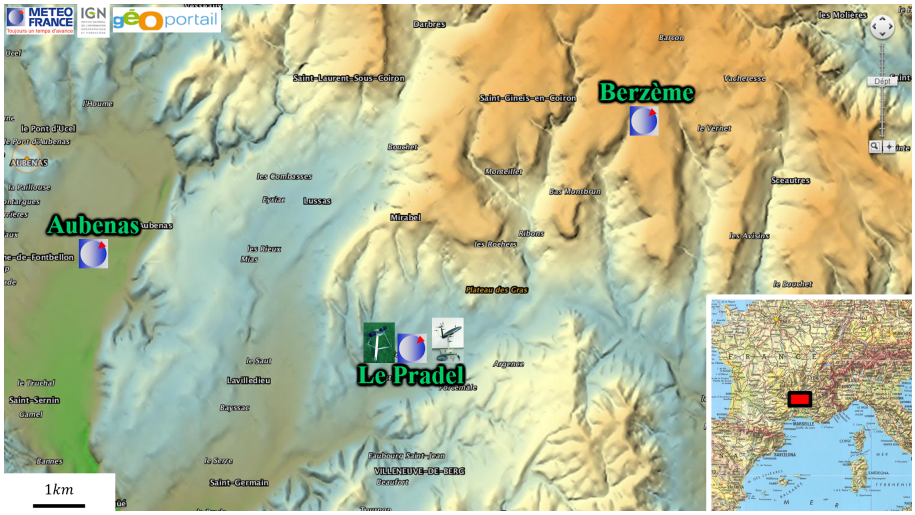

FIGURE V.23 – Localisation de la zone d'étude dans le sud-est de la France sur une carte topographique (source IGN). Les deux instruments utilisés dans l'algorithme d'assimilation sont au Pradel. Les données de stations Météo France au sol sont disponibles au Pradel, à Aubenas et à Berzème.

La section 5.2.1 présente les instruments et produits disponibles dans le contexte de ce travail. La section 5.2.2 présente l'évènement pluvieux étudié.

5.2.1 Instruments disponibles / Produits MRR

Données d'entrée de VAMOS Nous avons vu que notre algorithme d'assimilation réclamait comme données d'entrée (observations) des spectres Doppler et des mesures de flux de gouttes au sol. Dans cette étude, le radar utilisé est un MRR (Micro-Rain Radar) et le disdromètre un DBS (Dual Beam Spectropluviometer). Ces deux instruments ont été présentés dans le chapitre I (voir Tableau I.2 en section 2.3.3 pour le radar et section 2.2.3 pour le disdromètre).

Rappelons en quelques mots leurs caractéristiques principales. Le MRR est un radar Doppler à visée verticale à $24GHz$. Ses résolutions sont de $100m$ en vertical, $10s$ en temps et $0.19m/s$ en vitesse Doppler. Il fournit des mesures à partir de sa $3^{ème}$ porte ($[200-300]m$). (Quand nous parlerons de mesures radar au sol, il s'agira donc de mesures à $200/300m$ d'altitude.) Dans cette étude, les spectres Doppler seront intégrés à $2min$. Le DBS est un disdromètre optique ayant une surface de collecte de $0.01m^2$. On utilisera ses mesures pour les gouttes $> 0.4mm$. On travaillera également sur des données intégrées à $2min$.

Ces deux instruments sont localisés au village du Pradel, à $278m$ d'altitude (voir Figure V.23).

Produits MRR En plus des mesures des spectres Doppler, on a vu (chapitre I section 2.3.3) qu'il existait également un logiciel accolé au MRR fournissant, pour chaque spectre, une estimation de la DSD associée. Bien que celle-ci prenne en compte l'atténuation, elle néglige les effets du vent vertical et de la turbulence atmosphérique. À partir de cette DSD sont calculés divers moments, notamment le taux de pluie (RR), le contenu en eau liquide (LWC), le facteur de réflectivité radar (Z) ou le diamètre volumique moyen (D_m). Toutes ces quantités seront qualifiées dans la suite de produits MRR. Elles seront comparées aux restitutions de notre algorithme d'assimilation VAMOS.

Autres données : stations météorologiques On utilisera par ailleurs les données de trois stations Météo France au sol réparties dans la zone d'étude (voir Figure V.23). La première est également située au Pradel. La seconde est à Aubenas ($7.5km$ plus à l'ouest, $180m$ d'altitude). La troisième est à Berzème ($7.5km$ plus au nord-est, $650m$ d'altitude). Ces stations nous fourniront en particulier des informations sur la température, l'humidité relative et le vent horizontal au sol.

Limites On peut d'ors et déjà noter que l'on ne dispose pour cette étude que d'un nombre limité d'informations. On ne dispose notamment pas d'un autre radar Doppler à visée verticale. Nous ne pourrons pas valider nos restitutions. N'ayant pas de radiosondages, pas plus que d'autres mesures en altitude, nous n'aurons pas d'information sur les profils verticaux de vent ou l'humidité. Ce manque d'information nous contraint à un modèle simple : nous ne pouvons en choisir un plus complexe dont il faudrait pouvoir régler les paramètres supplémentaires. Nous n'avons pas non plus la possibilité de valider pleinement nos résultats. Ce sont les principales raisons qui nous conduiront par la suite à chercher d'autres études de cas plus documentées.

5.2.2 Présentation de l'évènement / Configuration de l'assimilation

L'évènement de pluie considéré ici a eu lieu dans l'après-midi du 12 octobre 2013, de 16:30 à 21:30. Il s'agit d'un épisode long de pluie faible et régulière. Ainsi, les mesures du disdromètre avec une résolution de $2min$ présentent des taux de pluie toujours entre 1 et $6mm/h$, présentés sur la Figure V.24 (rouge).

L'évènement semble également assez homogène dans l'espace. Ainsi, les trois stations Météo France situées dans la zone relèvent des taux de pluie très proches : $11.5mm$ au Pradel, $12mm$ à Aubenas et $12.9mm$ à Berzème. La Figure V.25 (partie gauche) présente les trois premiers moments des spectres Doppler observés par le MRR (réflectivité Z_e, vitesse Doppler moyenne W et largeur spectrale LS). On peut remarquer deux phases assez distinctes dans cet évènement. Jusqu'à 18:15 environ, on a des réflectivités relativement basses ($< 25dBZ$) correspondant surtout à des vitesses Doppler moyennes faibles ($< 5m/s$), signe a priori de la présence principalement de petites gouttes (tombant lentement). Après 18:15, on a au contraire des passages assez courts (quelques minutes) avec des réflectivités plus fortes (jusqu'à $35dBZ$) et des vitesses Doppler moyennes sensiblement plus fortes (jusqu'à $8m/s$), a

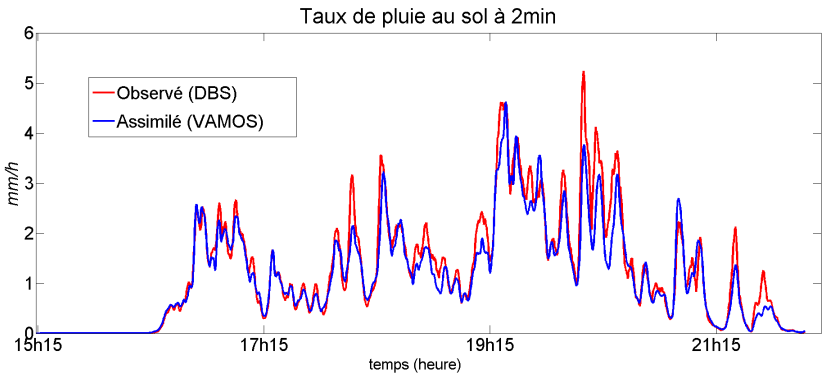

FIGURE V.24 – Taux de pluie au sol tels qu'observés par le DBS (rouge) et restitué par l'algorithme d'assimilation VAMOS (bleu).

priori signe de la présence de plus grosses gouttes. Les mesures au sol (Figure V.24) présentent également des taux de pluie faibles dans la première phase ($< 3mm/h$) et des pics un peu plus forts ensuite (jusqu'à un peu plus de $5mm/h$).

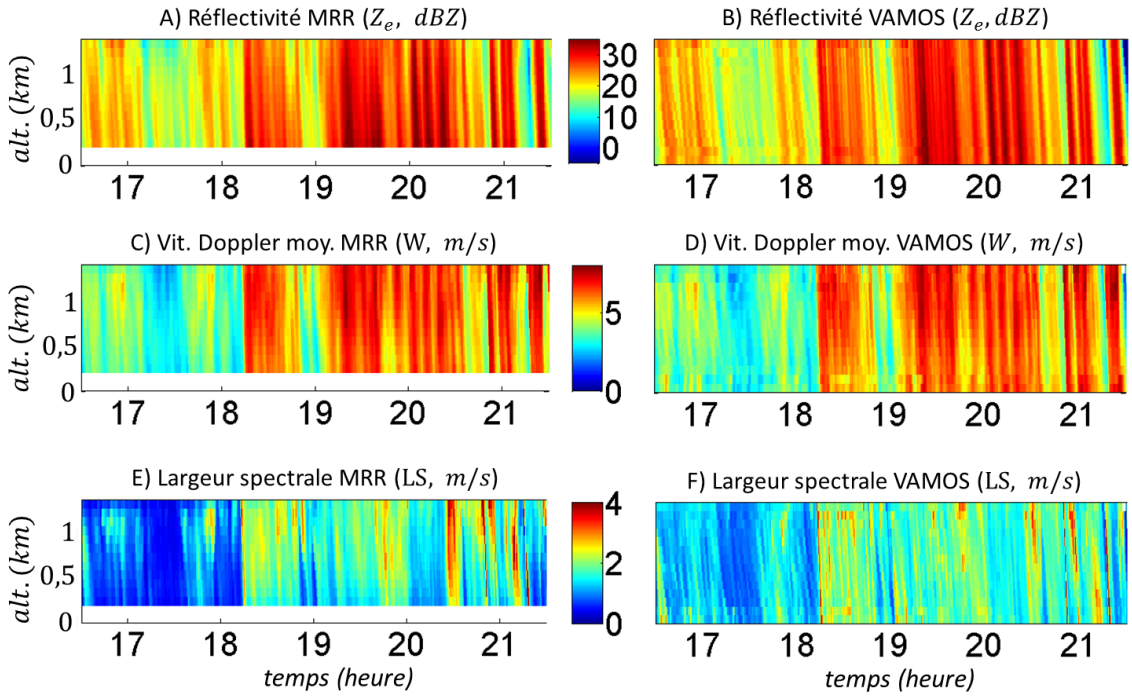

FIGURE V.25 – 3 premiers moments observés et assimilés des spectres Doppler. Gauche : Observations MRR ; Droite : Assimilation (VAMOS). Haut : Facteur de réflectivité radar équivalent Z_e en dBZ (proportionnel à la réflectivité η, moment d'ordre 0 du spectre Doppler). Milieu : Vitesse Doppler moyenne W (espérance). Bas : Largeur spectrale LS (écart-type).

La température au sol (au Pradel, altitude $278m$) est de $8.5\,^\circ C$, relativement constante durant l'évènement. En supposant un gradient thermique de $-1\,^\circ/150m$, on trouve un isotherme $0\,^\circ$ à environ $1550m$, cohérent avec les mesures du MRR (voir par exemple les produits MRR sur la Figure V.28, sur lesquels la bande brillante apparaît clairement vers $1500m$).

Notre algorithme d'assimilation ne traitant pas la couche de fonte, nous nous limiterons à la gamme $[0-1400]m$. Avec comme on l'a vu (section 2) un pas vertical de $\Delta z = 100m$, on aura ainsi $NZ = 14$ couches verticales. On rappelle que l'on utilise un pas de temps de $\Delta t = 5s$. On ne prendra pas en compte l'évaporation (cela équivaut formellement à un taux d'humidité de 100%). Cette hypothèse est justifiée d'abord par le fait que les stations au sol montrent une quasi-saturation de l'atmosphère (humidité relative de 97% à Berzème et 99% à Aubenas). Ces relevés ne prouvent bien sûr pas la saturation sur toute la colonne atmosphérique. Cependant, deux raisons nous poussent à supposer l'atmosphère saturée partout :

(1) - Sur les profils verticaux de réflectivité (par exemple Figure V.25), on ne voit pas de signature qui nous conduirait à envisager un effet significatif de l'évaporation. Une atmosphère saturée près du sol mais pas en altitude entraînerait a priori une réflectivité relativement constante sur les premières portes puis légèrement croissante ensuite, lorsque l'on monte. En effet, à la fréquence considérée, l'atténuation - susceptible de "compenser" l'évaporation sur les profils de réflectivité - est limitée pour les taux de pluie observés. Sur les deux études de cas suivantes (sections 6 et 7), on verra au contraire assez clairement de telles signatures.

(2) - Pour des évènements comparables (pluie stratiforme faible avec températures assez basses), l'étude des sorties du modèle AROME de Météo France indique généralement des taux d'humidité assez homogènes sur la colonne.

En l'absence d'informations quantitatives sur les taux d'humidité et notre algorithme ne contrôlant pas ce paramètre, il semble raisonnable de supposer l'atmosphère saturée à toutes altitudes (comme elle l'est effectivement au sol).

Les vents horizontaux près du sol à Berzème (altitude $650m$) sont faibles, entre 1.0 et $2.3m/s$ sur l'évènement. Là encore, cela ne prouve pas qu'ils sont également faibles sur toute la colonne. Cependant, les profils de réflectivité observés (Figure V.25A) semblent présenter des structures relativement diagonales. Les résultats de la section 4 suggèrent que cette signature correspond plutôt à un vent horizontal faible. Nous allons donc appliquer notre algorithme à cet évènement pour évaluer sa capacité à reproduire ces structures.

Notations / Vocabulaire Contrairement aux expériences jumelles (sur données simulées) précédemment réalisées (3.3 ou section 4.3), on ne connait pas ici l'état vrai de notre système. On dispose des états observés (spectres Doppler η_{obs}^{obs} et mesures disdromètres N_{dis}^{obs}), qui sont les données d'entrée de notre algorithme, et des états assimilés, composés : (1) - des inconnues de l'algorithme : DSD en haut de la colonne paramétrées sous forme gamma et vents verticaux. (2) - de ces inconnues propagées par le modèle : DSD partout n^{assim}. (3) - de ces états auxquels on a appliqué les opérateurs d'observations : spectres Doppler η_{obs}^{assim} et mesures disdromètres N_{dis}^{assim}.

Tous les résultats présentés dans les sections suivantes seront issus de champs dégradés à la résolution des observations, c'est-à-dire $2min$. On a vu (section 3.3.6) que pour le vent, l'algorithme utilise la résolution du modèle ($5s$) pour représenter la turbulence. Nous ne

TABLE V.3 – Comparaison entre états observés et assimilés pour les trois premiers moments des spectres Doppler (réflectivité η, vitesse moyenne W, largeur spectrale LS) et le nombre de gouttes disdromètre. On distingue le cas où l'on considère toutes les altitudes, et où l'on sépare les données en trois gammes d'altitudes. Les indicateurs présentés (et introduits en section 3.3.4) sont la MAE (moyenne des erreurs absolues, Eq.(V.20)), la MAPE (moyenne - en % - des valeurs absolues des erreurs relatives, Eq.(V.18)) et le biais relatif *rbias* (Eq.(V.19)). Biais positif = surestimation par l'assimilation.

		toutes altitudes	sommet $(1100-1400m)$	milieu $(700-1000m)$	bas $(200-500m)$
η	MAE (dB)	1.2	1.8	0.86	1.1
	bias (dB)	-0.017	$+0.60$	-0.12	-0.40
W	MAPE (%)	4.4	5.7	3.5	4.4
	rbias (%)	-0.39	-1.5	$+0.16$	-0.15
	moyenne(obs) (m/s)	5.11	5.28	5.18	4.87
LS	MAPE (%)	14	14	12	17
	rbias (%)	$+5.9$	$+3.1$	$+3.5$	$+12$
	moyenne(obs) (m/s)	1.20	1.25	1.22	1.11
DBS	MAE (nbre de gouttes : mm^{-1})	102			
	mean(obs) (nbre de gouttes : mm^{-1})	1080			

reviendrons pas sur ce phénomène dans cette section où l'on ne dispose pas d'autres instruments pour diagnostiquer plus en détails les performances de notre algorithme sur ce point. Nous y reviendrons cependant dans les sections 6 et 7 suivantes.

5.3 Résultats : reproduction des observations

Dans cette section, nous étudions la capacité de notre algorithme à restituer les observations qu'on lui a fournies (spectres Doppler et mesures de gouttes au sol). Dans cette optique, la Figure V.24 présente les taux de pluie au sol observés (par le DBS, rouge) et restitués par l'algorithme (bleu). Le Tableau V.3 présente quelques indicateurs d'erreurs entre états observés (par le DBS ou le MRR) et assimilés (VAMOS) pour les trois premiers moments des spectres Doppler (réflectivité η, vitesse moyenne W et largeur spectrale LS) ainsi que pour le nombre de gouttes disdromètre (flux de gouttes au sol). La Figure V.25 présente les champs temps/altitudes observés et assimilés pour ces trois premiers moments des spectres Doppler. La Figure V.26 enfin présente (A, à gauche), pour chaque porte verticale, les spectres observés et assimilés moyennés sur une vingtaine de minutes et (B, à droite) la part du coût total final de l'assimilation 4D-VAR après minimisation due à chacune de ces portes.

On constate tout d'abord (Figure V.24) que les taux de pluie observés au sol sont bien reproduits par notre algorithme, tant en valeur qu'en variabilité temporelle. On peut tout de même noter que les pics observés (par exemple le maximum global de la série vers $20h$) ont

tendance à être un peu sous-estimés par VAMOS. Nous reviendrons sur ce point plus bas. On note également (Tableau V.3, bas) que l'erreur sur le nombre de gouttes au sol est plus fort que dans l'expérience jumelle de la section 3.3. Cela reste tout de même assez faible (102 gouttes/mm) ramené à la valeur moyenne observée (1080 gouttes /mm).

Intéressons nous à présent aux observations radar. Tout d'abord, on note que les structures diagonales observées sur la carte de réflectivité MRR (Figure V.25A) sont correctement reproduites par notre algorithme (Figure V.25B). Cela laisse supposer une faible influence du vent horizontal sur cet évènement (voir section 4.3).

Ensuite, on constate que la vitesse Doppler moyenne W est très bien restituée au-delà des 200 premiers mètres (non vus par le radar). Visuellement (Figure V.25, C et D), les champs semblent très bien se correspondre. Statistiquement (Tableau V.3, colonne 3), on constate en effet une erreur moyenne (MAPE) faible (4.4%) et aucun biais significatif. Concernant la largeur spectrale LS, les restitutions semblent un peu moins bonnes mais les tendances et les structures du champ sont également assez correctement restituées. La MAPE n'est pas trop forte (14%), mais il existe cette fois-ci un biais assez significatif (+5.9%). En regardant la structure verticale de ces erreurs (Tableau V.3, colonnes 4 à 6), on constate que le biais ($rbias$) est sensiblement plus fort en bas de la colonne (+12%). Ce phénomène peut être dû à la diffusion numérique induite par notre modèle de propagation. Celle-ci "étale" les DSD, et donc les spectres, et ce d'autant plus que la hauteur de chute (temps d'advection) est grande, et donc que l'on s'approche du sol.

Concernant la réflectivité, on constate qu'en moyenne (sur toute la colonne, Tableau V.3 colonne 3), on commet une erreur absolue moyenne (MAE) relativement forte ($1.2dB$) mais sans biais. Cependant, en analysant la structure verticale de cette réflectivité (Tableau V.3 colonnes 4 à 6), on constate que notre algorithme la surestime ($+0.6dB$) sur le haut de la colonne et la sous-estime ($-0.4dB$) sur le bas. Ce phénomène est particulièrement visible sur la Figure V.25 (haut). Sur la figure observée (A), on voit bien que la réflectivité Z_e augmente sensiblement lorsque l'on se rapproche du sol. VAMOS (B) ne reproduit pas cette tendance. On voit également très clairement ce phénomène sur la Figure V.26A. L'algorithme d'assimilation surestime les spectres Doppler sur les portes les plus élevées, les sous-estime près du sol, et donne les meilleurs résultats vers le milieu de la colonne. Par conséquent, l'erreur ainsi créée (Figure V.26B) est forte sur le haut de la colonne et minimale au milieu. On comprend que l'algorithme d'assimilation, ne pouvant reproduire la structure verticale observée par le MRR, effectue un compromis sur la colonne. Il restitue ainsi bien la réflectivité moyenne sur l'ensemble de la colonne. On peut encore noter que ce phénomène explique la sous-estimation des pics de taux de pluie au sol par l'algorithme. Cependant, ces mesures supplémentaires au sol contraignent notre algorithme. On a ainsi des spectres assimilés qui ont tendance à être plus proches des mesures MRR sur le bas de la colonne que sur le haut (voir l'erreur sur la Figure V.26B ou dans la Tableau V.3, biais de $+0.6dB$ sur le haut et seulement $-0.4dB$ sur le bas).

La question de l'origine de cette tendance (réflectivité qui augmente quand on s'approche du sol) sur les mesures MRR n'est pas claire. Il ne s'agit pas d'un problème d'atténuation (qui induit bien le fait que moins d'énergie revient au radar depuis les portes les plus élevées), celle-ci étant prise en compte dans notre algorithme. Il est clair que rien dans notre algorithme ne permet de provoquer une telle baisse de réflectivité au cours de la chute des gouttes. La baisse constatée ici est très régulière et systématique. Elle est difficile à élucider dans mesures supplémentaires. Elle pourrait être due à un effet de la microphysique (coalescence / break-up, non pris en compte dans notre modèle), à des effets du vent horizontal (bien qu'il ne semble pas affecter la structure de la carte de réflectivité), aussi bien qu'à un problème de calibration du radar (fonction de transfert TF de l'appareil, voir MRR (1998)), conjugué au fait que l'on néglige l'atténuation due aux gaz (qui ne peut à elle seule expliquer cette tendance mais y contribue probablement).

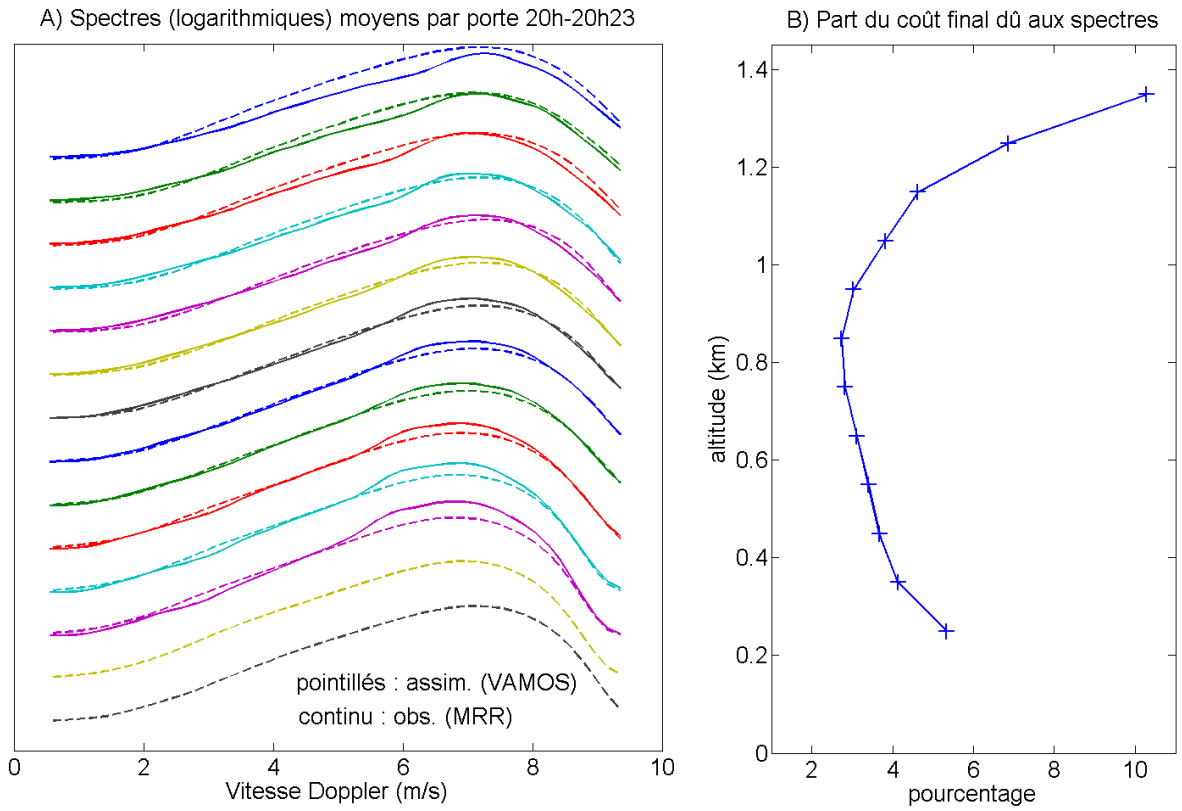

FIGURE V.26 – A) Pour toutes les altitudes, spectres Doppler observés (MRR, traits continus) et assimilés (VAMOS, traits pointillés) moyennés sur la période 20:00 - 20:23. On remarque que pour les deux portes les plus basses on ne dispose pas d'observations MRR. B) Part de la fonction de coût J après minimisation due à l'erreur commise sur les spectres Doppler pour chaque altitude (sur l'évènement entier). On rappelle que la fonction de coût comprend aussi un terme d'écart aux mesures disdromètres et des termes de filtrage / pénalisation. La somme des parts présentées ici ne fait donc pas 100% (mais 51%, voir texte).

On peut encore noter qu'à la fin du processus de minimisation, le coût (J) était composé à 51% par l'écart aux spectres Doppler (avec la répartition verticale de la Figure V.26B), à 42%

par l'écart aux mesures disdromètres et à 7% par les termes de pénalisation / filtrage. Cette répartition est très satisfaisante : elle montre que l'algorithme de minimisation ne s'est pas focalisé sur l'un des instruments mais a bien effectué un compromis entre les mesures disponibles. Le terme de filtrage n'a pas non plus écrasé les autres.

Globalement, on conclut que notre algorithme semble proposer un compromis entre le forçage apporté par le modèle de propagation (conservation des gouttes durant la chute) et les deux types de mesures disponibles. Il n'explique pas toutes les structures observées mais donne une solution dont la stricte cohérence physique est assurée par le modèle et qui est compatible avec les données des deux instruments de mesure.

Nous allons à présent examiner le vent vertical assimilé.

5.4 Résultats : vent vertical

Nous nous intéressons donc maintenant aux sorties de VAMOS, et commençons avec le champ temps/altitude des vents verticaux. Ce champ est présenté sur la Figure V.27A pour tout l'évènement (découpé en deux parties). La Figure V.27 présente également le vent moyen sur l'évènement selon l'altitude (B) et l'écart-type des séries temporelles de vent, pour chaque altitude également (C).

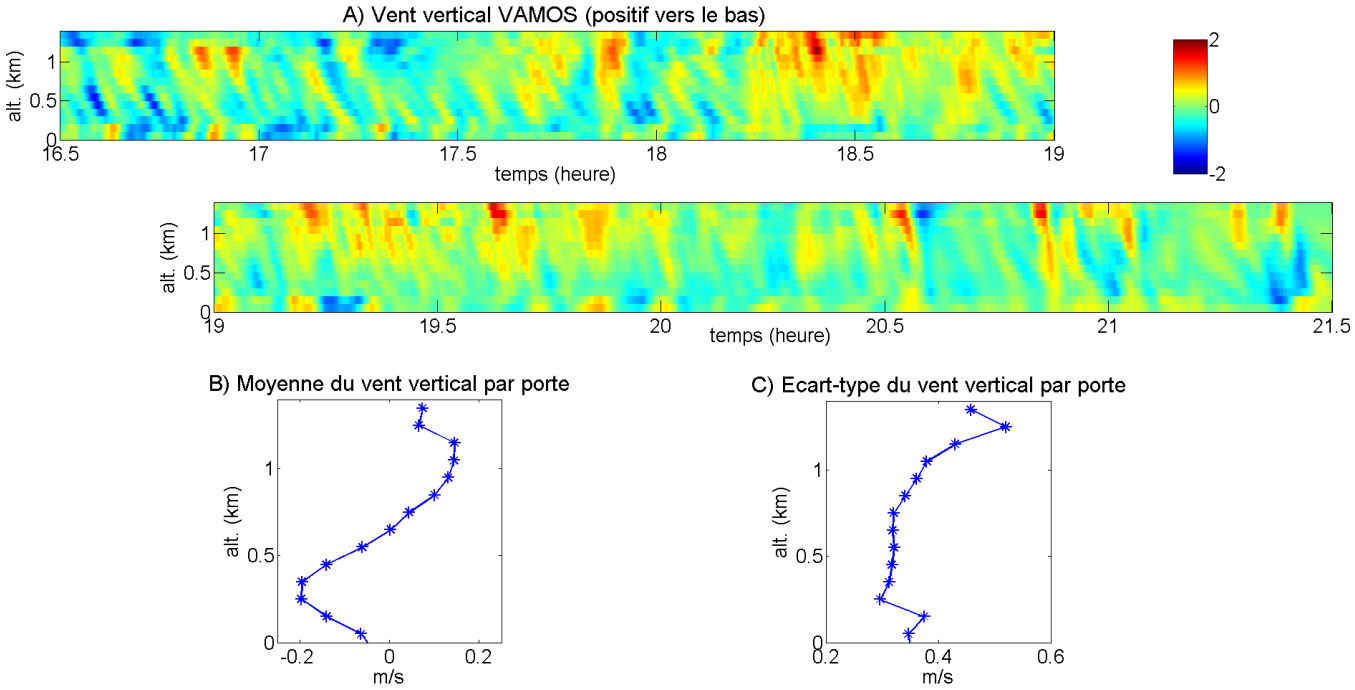

FIGURE V.27 – A) Vent vertical w assimilé (moyenné sur $2min$) sur tout l'évènement. B) Moyenne selon l'altitude. C) Écart-type selon l'altitude. Vent positif vers le bas.

On rappelle que le MRR ne fournit pas de spectres Doppler exploitables pour les 200 mètres inférieurs de la colonne atmosphérique. Les vent verticaux étant essentiellement estimés à partir

des spectres Doppler, les résultats sont à prendre avec prudence pour ces deux portes. Ils sont d'ailleurs suspects, avec par exemple une structure qui se crée uniquement sur les 200 premiers mètres vers 19:15, ou une inversion subite des vents un peu avant 21:30.

Pour le reste, la structure de ce champ de vent est cohérente avec ce que l'on attend pour un évènement de ce type. Pour ce vent à $2min$, on retrouve des structures (temps et longueurs de corrélation) de quelques minutes et environ $500m$, comparables avec les résultats de Kollias et al. (2002) ou Giangrande et al. (2010), par exemple. Par ailleurs, bien que nous ne puissions pas fournir de validation avec des données indépendantes, nous pouvons noter la robustesse de notre algorithme d'assimilation qui retrouve les mêmes structures quelque soit son initialisation (first guess, sur le vent comme sur la DSD).

Sur la colonne, on trouve un vent moyen de $+0.004m/s$, non significatif. On obtient cependant des vents plutôt orientés vers le bas (en rouge sur la Figure V.27A) sur les couches les plus hautes et vers le haut (en bleu) sur la moitié inférieure de la colonne. Les vents moyens (Figure V.27B) varient en effet d'environ $+0.15m/s$ en haut ($1.4km$) à $-0.2m/s$ en bas ($0.3km$). Les écarts-types (Figure V.27C) sont plus difficiles à analyser, notamment car leur valeur absolue dépend bien sûr de la résolution à laquelle on travaille. Les champs présentés ici sont à $2min$. À cette résolution, on n'explique pas la turbulence atmosphérique (voir section 3.3.6). On note simplement ici que les écart-types sont assez stables en dessous de $1km$, et plus forts tout en haut de la colonne, là où les valeurs absolues du vent semblent également légèrement plus fortes et là où l'on s'approche de la bande brillante.

5.5 Résultats : DSD / comparaison aux produits MRR

Comparaison aux produits MRR - Forme des DSD Nous allons dans un premier temps comparer nos restitutions de la DSD avec celles fournies par le logiciel accolé au MRR. (Produit MRR qui inverse chaque spectre indépendamment et sans tenir compte du vent vertical et de la turbulence.) La Figure V.28 présente des moments de la DSD issus du produit MRR (gauche) et VAMOS (droite). On peut d'ors et déjà noter que ces restitutions sont en accord sur plusieurs points qui confirment ce que l'on avait lu de l'évènement plus haut juste à partir des mesures radar : sur la première partie de l'évènement ($<$ 18:15), taux de pluie faibles avec majorité de petites gouttes (faibles D_m), et ensuite taux de pluie et facteurs de réflectivité plus élevés, avec des plus grosses gouttes (les D_m montent jusqu'à presque $2mm$).

Par ailleurs, on peut noter plusieurs points sur les produits MRR (gauche de la Figure V.28) :

(1) - la bande brillante, qui n'est pas traitée comme telle dans l'inversion des spectres Doppler, est très nettement visible (taux de pluie aberrants vers $1.5km$ par exemple).

(2) - les taux de pluie (RR, Figure V.28C) et les facteurs de réflectivité radar (Z, moment d'ordre 6, Figure V.28E) augmentent très sensiblement lorsqu'on s'approche du sol. Cela est dû au fait que, comme on l'a vu en section 5.3, l'intensité des spectres Doppler mesurés augmente lorsqu'on s'approche du sol plus intensément que ce que l'atténuation peut expliquer. Cette

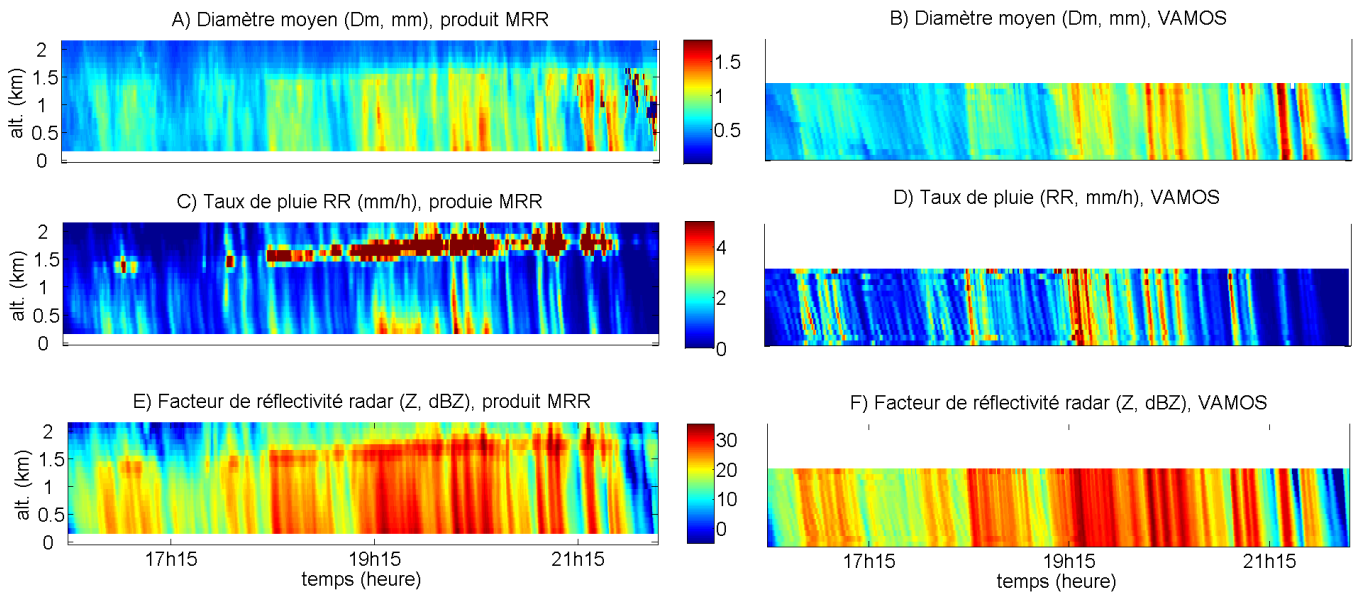

FIGURE V.28 – Comparaison entre produits MRR (gauche) et états assimilés VAMOS (droite) pour différents moments de la DSD. Haut : diamètre volumique moyen (D_m, moment d'ordre 4 sur moment d'ordre 3), mm. Milieu : Taux de pluie RR, mm/h. Bas : Facteur de réflectivité radar Z (moment d'ordre 6), dBZ.

augmentation de réflectivité se traduit logiquement sur les DSD par une augmentation des taux de pluie et facteurs de réflectivité.

(3) - les diamètres volumiques moyens sont moins sensibles à ce phénomène. Ils semblent très perturbés (valeurs aberrantes) en altitude vers la fin de l'évènement (21:15 - 21:30, $0.5 - 1.5km$).

On peut reprendre ces trois points à la lueur de nos résultats (droite de la Figure V.28) :

(1) - la bande brillante n'est pas traitée par notre algorithme, qui se limite à une altitude maximale fixée à $1400m$.

(2) - les DSD sont globalement conservées sur la colonne (modulo le vent vertical qui altère la vitesse de chute des gouttes). Par conséquent, de la même manière que notre algorithme ne produisait pas de variabilité verticale notable (hors atténuation) sur les réflectivités spectrales (Figure V.25D), il restitue des taux de pluie, D_m et facteurs de réflectivité radar assez cohérent sur la colonne. Là encore, ne pouvant déterminer si la variabilité vue par le MRR est réelle (coalescence / break-up par exemple) ou un artefact de mesure, nous ne pouvons conclure sur la pertinence de notre restitution comparée aux produits MRR, mais simplement noter que notre modèle de propagation assure une cohérence verticale stricte à tous les champs restitués.

(3) - les D_m sont assez cohérents verticalement également, et ne sont pas perturbés à la fin de l'évènement.

La Figure V.29 confirme ces résultats. Elle présente les DSD DBS, produits MRR et assimilées (VAMOS) en haut et en bas de la colonne (voir légende). On remarque qu'au sol

(traits continus), VAMOS sous-estime légèrement les DSD sur les diamètres intermédiaires ($1 - 2.5mm$) par rapport aux mesures du DBS, qui sont elles-même plus basses que les estimations du produit MRR. En haut de la colonne au contraire (pointillés), VAMOS restitue des DSD supérieures à celles du produit MRR, confirmant encore une fois que l'on effectue une moyenne sur la colonne.

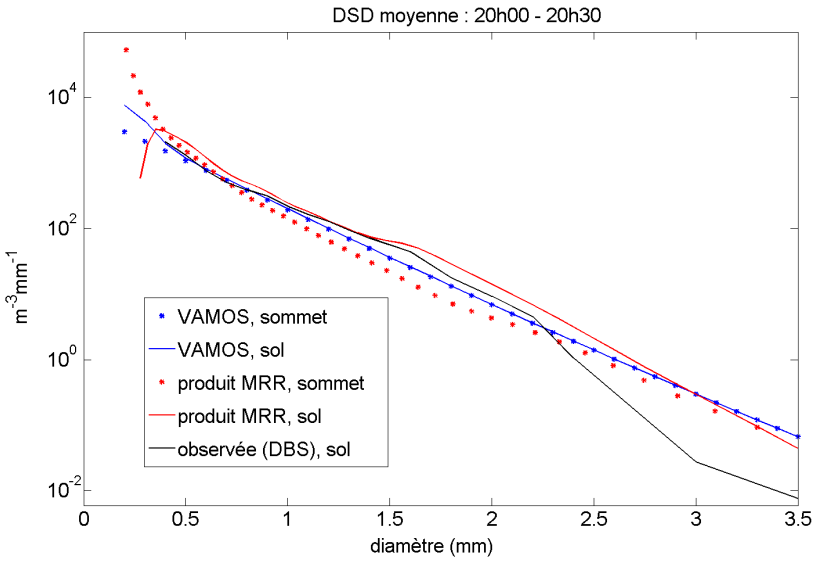

FIGURE V.29 – DSD moyenne entre 20:00 et 20:30. On présente des DSD en bas de la colonne (traits continus) : observées par le DBS (noir), assimilées dans VAMOS (rouge) et fournies par le produit MRR (bleu) et en haut de la colonne (traits pointillés marqués) assimilées (VAMOS, noir) et fournies par le produit MRR (bleu).

On note également que pour les petits diamètres ($< 0.5mm$) les produits MRR semblent très suspects, avec un nombre de petites gouttes particulièrement impressionnant en haut de la colonne, et au contraire très faible en bas. Ceci n'est guère surprenant dans la mesure où ces diamètres correspondent à des zones très peu énergétiques du spectre Doppler (mauvais rapport signal sur bruit). Les résultats de notre algorithme sur ces diamètres sont a priori suspects. En effet, on a vu sur données simulées (section 3.3.5) qu'il était très peu contraint pour les petites gouttes. Ici, cependant, les observations du modèle correspondent bien aux mesures DBS quand elles sont disponibles (jusqu'à $0.4mm$, courbe noire) et ne présentent pas de cassure ensuite (jusqu'à $0.2mm$).

Concernant les gros diamètres ($> 2 - 2.5mm$), les mesures du DBS sont très suspectes et ne correspondent pas aux autres restitutions. Cela s'explique sans doutes par le fait que pour un tel évènement pluvieux (faible), les grosses gouttes sont rares. Ainsi la forme d'une DSD peut être altérée par la présence ou l'absence d'une grosse goutte particulière dans les mesures DBS. De plus, les vitesses terminales de chute de ces gouttes sont $> 7.3m/s$, si bien que le temps d'intégration du disdromètre ($2min$ ici) correspond à des volumes sondés très importants (en 2 minutes, une goutte tombant à $7.3m/s$ parcourt $850m$), volume sondés ne correspondant

absolument pas au volume sondé par le MRR ($100m$). Cependant, ces gouttes étant rares ici et peu déterminantes dans l'estimation des moments de la DSD (taux de pluie en particulier), nous choisissons de conserver un temps d'intégration de $2min$. Pour des évènements convectifs, ce choix devrait bien sur être discuté. Pour gérer la turbulence, nous avons vu qu'il importait de conserver une résolution des spectres Doppler assez supérieure à la résolution du modèle ($5s$). Pour les mesures disdromètre, on pourrait cependant utiliser un temps d'intégration plus petit, ou dépendant de la classe de diamètre (et donc de la vitesse de chute), afin d'avoir pour chaque classe de diamètre des volumes sondés comparables.

Structure des champs de DSD Intéressons nous à présent à la structure des champs de DSD restitués (ou du moins de ses moments), indépendamment des produits MRR.

La Figure V.30 présente des champs temps/altitude de diamètres volumiques moyens D_m (A), de nombre de gouttes par unité de volume N_{tot} (B) et d'inverses de pentes des DSD (C). Ces dernières sont calculées en effectuant, pour chaque DSD, une régression linéaire sur son logarithme. Cela suppose naturellement qu'elles soient "suffisamment" proches d'une forme exponentielle, ce "suffisamment" pouvant être quantifié par le coefficient de corrélation de la régression. (On peut voir sur la Figure V.29 que les DSD restituées sont en moyenne assez proches d'une exponentielle, à l'exception des très petites gouttes). La Figure V.30D représente le taux de pluie RR et la pente de la DSD au sol.

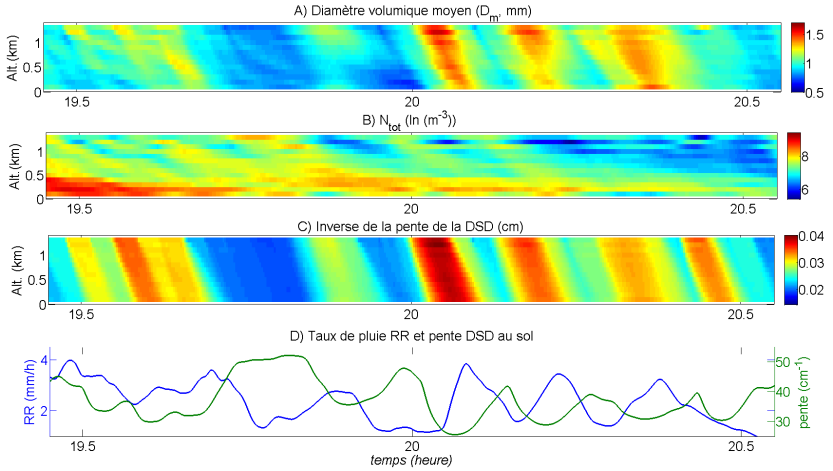

FIGURE V.30 – Champs issus de la DSD assimilée (VAMOS), entre 19:30 et 20:30. A) Diamètre volumique moyen D_m, mm. B) Nombre volumique de gouttes N_{tot}, $ln(m^{-3})$. C) Inverse de la pente de la DSD, cm. La pente est calculée en faisant une régression linéaire sur le logarithme de la DSD (en la supposant suffisamment exponentielle). D) Au sol, taux de pluie assimilé VAMOS (RR, mm/h) et pente de la DSD (cm^{-1}).

On constate tout d'abord que l'on obtient pour D_m et pour la pente de la DSD des valeurs cohérentes pour un évènement de pluie faible (voir par exemple pour la pente l'étude de Giangrande et al. (2010), Fig. 8 et pour D_m celle de Peters et al. (2005), Fig. A4).

On note par ailleurs que le champ des pentes des DSD présente des profils diagonaux très marqués. Cela reste à peu près vrai avec D_m alors que cela ne l'est plus sur le nombre de gouttes N_{tot}. Le fait qu'il existe des structures diagonales est bien sûr un marqueur de la chute verticale des gouttes de pluie dans notre modèle (voir section 4.3, Figure V.19 par exemple). On rappelle que la vitesse de chute des gouttes n'est pas constante mais modulée par le vent vertical. Celui-ci agit peu (relativement à leur vitesse terminale) sur les grosses gouttes. La pente de la DSD est principalement déterminée par la queue des distributions et donc par les grosses gouttes. Cela explique les profils diagonaux très nets. On peut noter qu'il en est de même pour le moment d'ordre 6 de la DSD (Figure V.28F). Le D_m est le rapport des moments d'ordres 4 et 3 de la DSD. Pour un évènement faible comme celui-ci, présentant très peu de grosses gouttes (on a un D_m entre 0.5 et $1.5mm$), sa valeur est principalement déterminée par des petites gouttes. La structure du champ de D_m est donc affectée dans une certaine mesure par le vent vertical.

Le champ du nombre de gouttes N_{tot} est lui principalement déterminé par le vent vertical. Les DSD étant largement exponentielles ici (Figure V.29, bleu), N_{tot} est principalement déterminé par les très petites gouttes, très sensibles au vent vertical. Celui-ci va former des paquets de gouttes. Par exemple, en se reportant à la Figure V.27A, on voit que l'on a une petite zone de vents verticaux négatifs, donc vers le haut (bleu), vers 19:55 entre 100 et $400m$. Cela limite la chute des petites gouttes (les garde en l'air) et fait augmenter N_{tot} entre 300 et $500m$. On note que si les valeurs absolues de N_{tot} sont à considérer prudemment (algorithme peu contraint sur les petits diamètres), ces paquets de gouttes sont probablement réels puisqu'ils sont induits par le vent.

L'effet du vent vertical est également ce qui explique que les champs de taux de pluie restitués par notre algorithme (Figure V.28D) sont relativement plus bruités que ceux de D_m par exemple, le taux de pluie n'étant pas à proprement parler un moment de la DSD. C'est un flux qui dépend directement de la vitesse de chute des gouttes, et donc du vent vertical. Calculer le taux de pluie comme le moment d'ordre 3.67 de la DSD (Atlas and Ulbrich (1977)) ne permettrait pas de restituer cet effet du vent vertical.

On voit sur la Figure V.30D que les pentes de la DSD semblent anticorrélées aux taux de pluie. Cela est caractéristique de notre évènement stratiforme. Une DSD exponentielle à forte pente signifie très peu de grosses gouttes, et donc un taux de pluie faible. Au contraire, une pente faible implique plus de grosses gouttes. Là encore, ces résultats sont très proches de ceux de Giangrande et al. (2010) (Fig. 9 par exemple).

5.6 Conclusions et ouverture sur le deuxième cas d'étude

Dans cette section, nous avons présenté une première application, consacrée à un évènement de pluie faible, long et régulier, au sein duquel les effets de l'évaporation et du vent horizontal apparaissaient assez limités. L'objectif était de valider la faisabilité de notre approche sur un cas simple et d'étudier quelques aspects clé de son comportement (capacité à

reproduire les observations et à s'adapter au mieux lorsqu'elles étaient incompatibles entre elles ou avec le modèle de propagation utilisé, restitution du vent vertical et interaction avec la DSD).

Nous avons montré que notre algorithme parvenait à restituer de manière satisfaisante la variabilité temporelle et les structures de réflectivités radar sur la colonne entière (périodes de pluie forte ou moins forte, signatures diagonales). Nous avons aussi montré que lorsque le radar enregistrait une tendance verticale non reproductible par notre modèle de propagation (forte augmentation des réflectivités en s'approchant du sol), l'algorithme s'adaptait au mieux en restituant un comportement moyen sur la colonne verticale. Son forçage conjoint par le modèle de propagation et par les observations le force à produire une solution cohérente physiquement (cohérente avec la physique introduite dans le modèle, qui ne suffit pas forcément à expliquer tous les comportements observés). Les champs de DSD (étudiés ici principalement via les diamètres volumiques moyens, taux de pluie ou facteurs de réflectivité) et de vents verticaux retrouvés sont cohérents avec ce que l'on attend pour un évènement de ce type (au vu d'autres études réalisées sur des évènements similaires).

Il est toutefois bien clair que nous ne disposions ici d'aucun moyen de valider ces résultats par des mesures indépendantes. Nous allons donc dans la section 6 traiter un autre cas ou l'on dispose d'un jeu de données plus complet. Les sections 6 et 7 nous permettront également d'éprouver et de valider notre approche VAMOS dans différentes conditions météorologiques (pluie froide en Finlande et pluie chaude tropicale en Australie). Nous ajouterons ainsi l'effet de l'évaporation et étudierons l'impact d'un vent horizontal plus conséquent sur la qualité de nos restitutions. Nous nous attarderons également sur la reproduction des effets de la turbulence atmosphérique. Ces études nous permettront aussi de travailler avec un autre radar Doppler, à une fréquence légèrement plus élevée ($35Ghz$ contre $24GHz$ ici), et possédant de meilleures résolutions verticale, temporelle, et Doppler.

6 Deuxième cas d'étude : Finlande

6.1 Présentation et enjeux

Nous avons précédemment appliqué notre algorithme de restitution VAMOS à un évènement de pluie faible et régulière en Ardèche. Nous avons ainsi pu étudier la capacité de notre algorithme à expliquer les différentes observations disponibles tout en restituant la structure et la cohérence des vents verticaux et des DSD. Toutes les données disponibles ayant été utilisées pour produire ces restitutions, nous n'avons cependant pas pu valider nos résultats indépendamment.

Ici, nous allons appliquer notre algorithme à un évènement de pluie aux caractéristiques assez comparables (pluie longue, froide, faible et régulière) mais couvert par davantage d'instruments. En particulier, nous disposerons ici de deux radars enregistrant les spectres Doppler, à deux fréquences différentes. Le premier, à $35GHz$, sera utilisé comme source d'observations (toujours en combinaison avec un disdromètre au sol) tandis que le second, à $95GHz$, sera utilisé en validation. Nous disposons également pour cet évènement des restitutions de vents verticaux et de DSD (notamment) fournis par Tridon and Battaglia (2015). Nous pourrons donc, toujours dans l'idée de valider nos restitutions, comparer nos résultats à ces restitutions.

L'objectif principal de cette section est donc de valider nos résultats, afin de pouvoir ensuite appliquer notre algorithme à un évènement un peu plus complexe. Bien entendu, nous étudierons toujours également, mais plus rapidement, les capacités de notre méthode à expliquer les observations.

Dans un premier temps (section 6.2), nous présentons le site d'étude, les instruments et données disponibles, l'évènement de pluie et la configuration qu'il implique pour notre algorithme. Puis nous présentons les résultats dans la section 6.3. La section 6.4 résume ces résultats, propose des pistes d'amélioration ou d'autres études possibles, et ouvre vers le $3^{ème}$ cas d'étude.

6.2 Instruments disponibles / Présentation de l'évènement

Dans cette section, nous traitons un cas de pluie ayant été enregistré dans le sud-ouest de la Finlande au printemps 2014, plus précisément à Hyytiälä ($179m$ d'altitude, coordonnées $61.8430N$, $24.2880E$). La Figure V.31 présente l'emplacement de ce site (marqueur rouge). Les instruments déployés ici l'ont été dans le cadre du programme ARM (Atmospheric Radiation Measurement) Climate Research Facility, un projet financé par le département américain de l'énergie et déployant en divers lieux dans le monde des instruments de mesures de l'atmosphère. On trouvera plus de détails sur ce programme sur son site internet : https://www.arm.gov. Concernant la campagne de mesures d'Hyytiälä et ses enjeux, on se reportera à sa page internet : https://www.arm.gov/sites/amf/tmp/.

Comme dans la section 5 précédente, nous présenterons d'abord les instruments et produits

FIGURE V.31 – Localisation (marqueur rouge) du site ou sont disposés les instruments ARM utilisés dans la section 6. (source Google Maps)

disponibles pour ce site de mesures (section 6.2.1), puis l'évènement spécifique étudié (section 6.2.2) et la configuration qu'il implique pour notre algorithme d'assimilation.

6.2.1 Instruments et produits disponibles

Données d'entrée de VAMOS Nous utilisons toujours comme données d'entrée pour notre algorithme de restitution des spectres Doppler d'un radar à visée verticale et des mesures disdromètre au sol. Dans cette étude, le radar utilisé est le KAZR (Ka-band Zenith Radar) et le disdromètre un 2D-vidéo. Ces deux instruments ont été présentés dans le chapitre I (voir Tableau I.2 en section 2.3.3 pour le radar et section 2.2.2 pour le disdromètre).

Rappelons en quelques mots leurs caractéristiques principales. Le KAZR est un radar Doppler pulsé à visée verticale à $35GHz$. Ses résolutions, pour le site Finlandais, sont de $30m$ verticalement, environ $2s$ en temps et $0.024m/s$ en vitesse Doppler (résolutions sensiblement plus fines sur tous les paramètres que le MRR utilisé en section 5). Pour l'évènement traité ici, il fournit des mesures utilisables à partir de sa $5^{ème}$ porte (correspondant à la gamme $[205-235]m$). Dans cette étude, les spectres Doppler seront intégrés sur une durée de $30s$. La fine résolution du radar nous permet en effet de décroître sensiblement le temps d'intégration par rapport au MRR (où l'on intégrait à $2min$) mais on a vu que l'on devait conserver un temps d'intégration plus grand que celui du modèle ($5s$) pour pouvoir représenter la turbulence atmosphérique éventuelle. Le 2DVD (2D-Video Disdrometer) est un disdromètre ayant une surface de collecte de $0.01m^2$ (comme le DBS utilisé dans la section 5). On utilisera ses mesures pour les gouttes $> 0.4mm$. On travaillera sur des données intégrées à $1min$.

Dans la suite, les résultats présentés seront par défaut intégrés à $30s$.

Instruments et produits de validation La différence principale par rapport à l'étude précédente (section 5) vient de la présence à Hyytiälä d'un second radar à visée verticale enregistrant les spectres Doppler. Il s'agit d'un instrument en bande W ($95GHz$), le $WACR$, déjà présenté également dans le chapitre I (Tableau I.2 en section 2.3.3). Ses caractéristiques et performances (résolution, ouverture de faisceau) sont comparables à celles du KAZR. La présence de cet instrument va nous permettre d'évaluer les performances de notre algorithme de deux manières :

- d'abord en simulant les spectres Doppler à $95GHz$ que produiraient nos DSD et vents verticaux restitués et en les comparant aux mesures effectives du WACR. On évalue ainsi la capacité de nos restitutions à représenter les observations faites à une autre fréquence, non utilisées dans la phase d'apprentissage. Le passage de 35 à $95GHz$ se traduisant (notamment) par le fait d'accorder plus d'importance aux vitesses Doppler (vitesses de chute) plus faibles, correspondant à des diamètres plus petits, on peut évaluer la capacité de notre algorithme à produire des restitutions cohérentes sur une large gamme de diamètres.

- ensuite en comparant nos restitutions (de DSD et de vents verticaux) à celles obtenues pour le même cas par l'algorithme présenté dans Tridon and Battaglia (2015). Cette méthode restitue DSD, vents verticaux et d'autres paramètres influant sur la mesure radar (intensité de la turbulence, atténuation due à la pluie, au radôme, etc.) à partir des spectres Doppler à 35 et $95GHz$. Frédéric Tridon et Alessandro Battaglia nous ayant fourni les restitutions de leur algorithme pour une partie du cas étudié ici, nous pourrons comparer les restitutions de VAMOS aux leurs. Tout ce qui se rapporte à cette méthode sera par la suite désigné par "TB2015". Les restitutions de notre algorithme seront elles généralement désignées par "VAMOS".

Autres données Toujours dans le cadre de ARM ont été réalisées à Hyytiälä des mesures continues de divers paramètres (température, pression, vent horizontal, humidité, etc.) au sol et toutes les $6h$ en altitude via des radiosondages. Comme précédemment, ces mesures nous permettront de régler les paramètres de notre algorithme d'assimilation (hauteur de chute sous la bande brillante, appréciation a priori de l'importance du vent horizontal, de l'évaporation, etc.).

6.2.2 Présentation de l'évènement / Configuration de l'assimilation

Evènement L'évènement étudié dans ce travail a eu lieu le 12 juin 2014. Il s'agit d'un évènement particulièrement long et régulier, avec un peu plus de $30mm$ d'eau entre 2:30 et 16:00 et des taux de pluie à $1min$ constamment entre 0.5 et $6.5mm/h$. La Figure V.32 présente la réflectivité et la vitesse moyenne obtenues lors d'une partie de cet évènement entre 9:00 et 14:00, par le KAZR (A et B), selon l'altitude, ainsi que des mesures au sol : température et humidité (C) et taux de pluie et cumul d'eau (D).

Sur ces figures, on voit bien la régularité de la pluie durant l'évènement (la courbe du

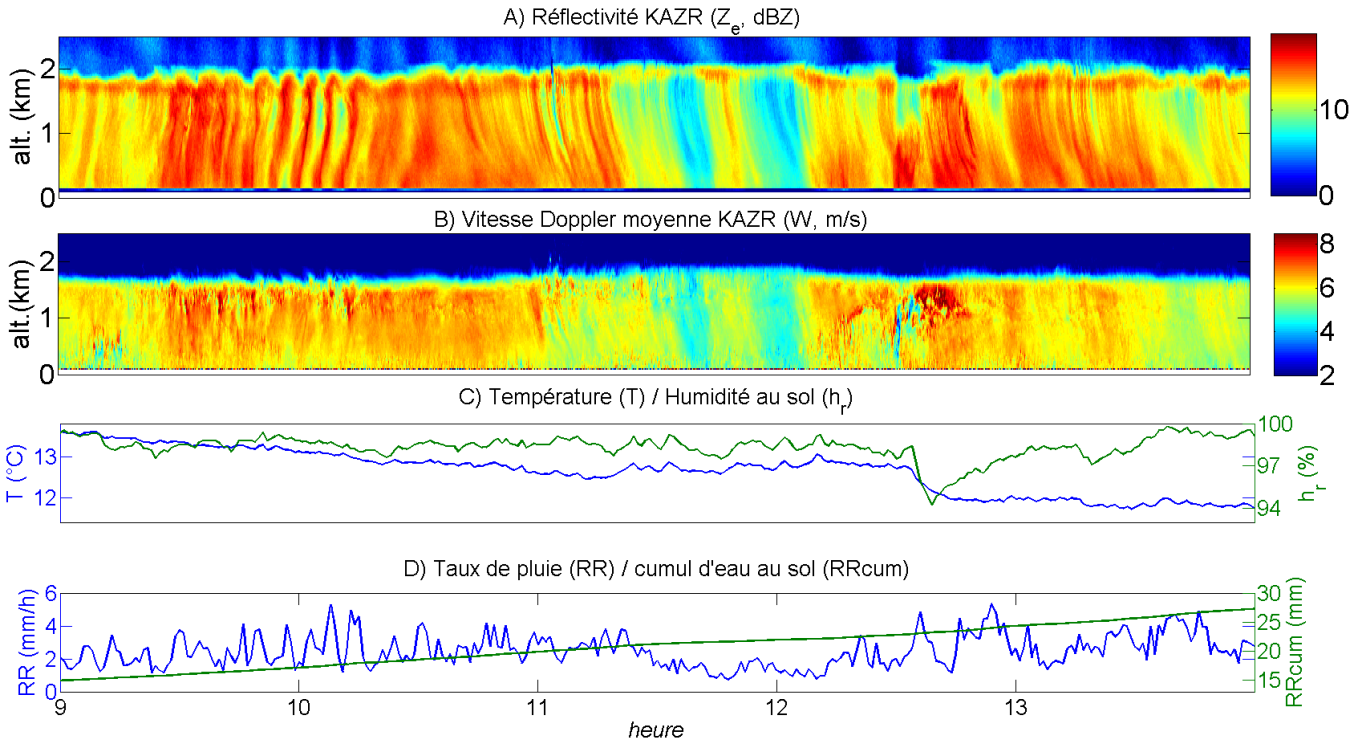

FIGURE V.32 – A-B) Facteur de réflectivité radar équivalent (Z_e, proportionnel à la réflectivité η) et vitesse Doppler moyenne (W) mesurées par le KAZR à Hyytiälä le 12/06/2014. C) Température et humidité relative au sol. D) Taux de pluie et cumul d'eau au sol (pluviomètre).

cumul d'eau au sol semble presque linéaire). On remarque cependant diverses phases et signatures sur la carte de réflectivités. Entre 9:00 et 10:30, on reconnaît les profils inversés de réflectivités mentionnées et représentés en section 4.2.5 (échos radars provenant d'abord des portes près du sol puis des portes supérieures, signe a priori de vent horizontaux plus forts près du sol qu'en altitude). Vers 10:30, les signaux s'inversent et on retrouve des signatures plus habituelles. Celles-ci ne sont cependant pas diagonales mais impliqueraient plutôt un vent horizontal plus fort en altitude que près du sol. On se reportera pour plus de détails à la section 4.2.5, et en particulier à la Figure V.12, où l'on se basait déjà sur ces observations. La carte des vitesses Doppler moyennes semble correspondre globalement à celle de réflectivités (vitesses augmentant et diminuant comme la réflectivité), avec tout de même des différences sensibles, notamment vers 12:30 (en particulier au dessus de $0.8km$ d'altitude), où l'on peut supposer une influence importante du vent vertical et de la turbulence perturbant la vitesse moyenne indépendamment de la réflectivité.

L'altitude de la bande brillante, sur les deux cartes radar, semble relativement constante et située vers $1.6 - 1.8km$ d'altitude, bien que la température baisse légèrement au niveau du sol durant l'épisode (d'environ $14\,^\circ C$ à un peu moins de $12\,^\circ C$).

Concernant l'humidité et le vent horizontal, on va aussi se reporter à la Figure V.33, montrant des résultats du radiosondage effectué à 11:36. Durant son élévation du sol jusqu'à

1800m d'altitude (qui a pris 5min30), la sonde a parcouru 1.1km vers le nord-ouest. Comme on pouvait s'y attendre (vu son déplacement), elle enregistre un vent de sud-sud-est de direction très homogène (à 20 ° près) jusqu'à l'approche de la bande brillante (non représenté ici). La direction du vent varie ensuite davantage au-delà de 1500m d'altitude. Ceci nous rapproche donc du cas étudié en section 4 (advection dans une seule direction horizontale). Au niveau des vitesses du vent (Figure V.33B), on observe une augmentation du sol jusqu'à environ $600 - 800m$, puis une diminution jusqu'à 1400m. Un tel profil expliquerait assez bien les signatures de réflectivités vues sur la Figure V.32A jusqu'à 10:45, mais moins bien ceux observés à l'heure du radiosondage (11:36). Peut-être le déplacement horizontal de la sonde dans la direction de propagation de l'évènement durant son élévation explique-t-il en partie le "retard" que l'on semble ainsi constater de ses mesures sur les observations radar.

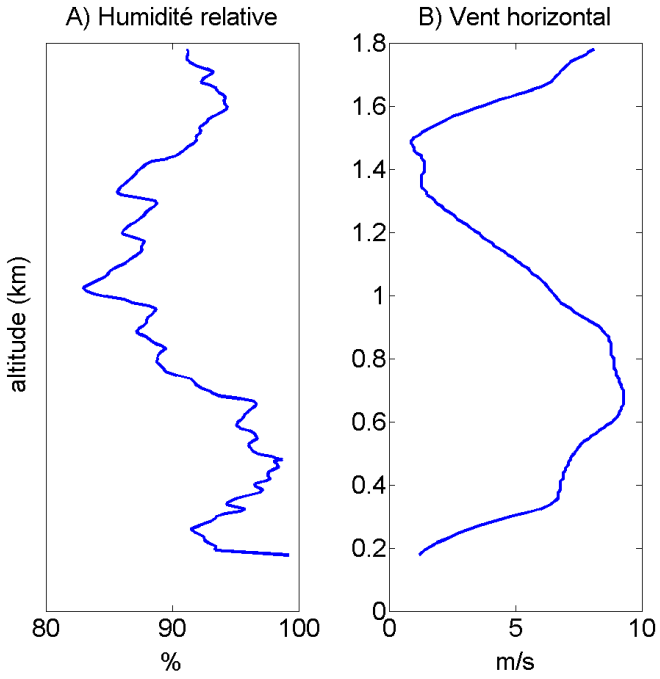

FIGURE V.33 – Humidité relative (A) et vent horizontal (B) mesurés par le radiosondage effectué à Hyytiälä à 11:36.

Au niveau de l'humidité relative (Figure V.33), on remarque que si l'atmosphère est quasiment saturée au sol ($> 98\%$), ce n'est pas le cas sur toute la colonne, avec des humidités relatives descendant jusqu'à environ 82% vers 1km d'altitude.

Voyons à présent comment l'on transcrit toutes ces constatations dans le réglage des paramètres de notre algorithme VAMOS.

Assimilation Nous avons appliqué notre algorithme d'assimilation à (quasiment) l'intégralité de ce long évènement, en le subdivisant en 4 sous-périodes (pour diminuer le temps de calcul en lançant plusieurs opérations simultanées). Néanmoins, afin de pouvoir

analyser en détails les résultats et parce que nous ne disposions des résultats de TB2015 que pour cette période, nous nous focaliserons principalement dans le suite sur la phase 10:30-12:30. On note néanmoins qu'il pourrait être intéressant d'évaluer la qualité de nos restitutions de vents verticaux sur la période 9:00-10:30 (avec les profils inversés de réflectivités mentionnés ci-dessus qui eux ne pourront être expliqués par notre modèle en l'état).

L'atmosphère ne semblant pas saturée sur toute la colonne (Figure V.33A), on choisit contrairement au cas précédent (section 5) de ne pas négliger l'évaporation. Et ne disposant d'autres informations sur les profils d'humidité que celles apportées par le radiosondage de 11:36, nous paramétrons l'humidité avec ce profil pour toute la période 10:30-16:00. Pour la température, nous supposons l'isotherme $0\,^{\circ}C$ à $1800m$ (haut de la bande brillante, estimé à partir des mesures radars et du profil de température du radiosondage, non présenté ici). Nous appliquons ensuite un gradient linéaire entre les deux. Vues les températures au sol (Figure V.32C), ce gradient est d'environ $-1\,^{\circ}/150m$.

Pour éviter les problèmes de bande brillante, notre algorithme couvrira la gamme $[0-1400]m$. Notre modèle aura donc $NZ = 14$ couches verticales avec un pas de $\Delta z = 100m$.

Un dernier problème non encore mentionné subsiste : l'atténuation du signal radar dû à son radôme (abri protégeant l'appareil, voir la photographie du KAZR Figure I.13). Lorsqu'il pleut, de l'eau est présente sur ce radôme, et atténue très significativement le signal émis. Ce phénomène n'étant pas pris en compte dans le modèle, nous le corrigeons sur les observations fournies à l'algorithme, grâce à une méthode simple basée sur les mesures disdromètre. Cette méthode est décrite dans l'Annexe K. Pour l'évènement considéré, on obtient des atténuations radôme fortes, fluctuant autour de $16dB$. On note que l'imprécision de cette méthode suggérerait de corriger cette atténuation radôme directement dans l'algorithme d'assimilation. Le travail requis pour cela fait qu'on le laisse cependant pour plus tard.

6.3 Résultats

L'objectif de cette section est principalement de comparer nos restitutions (DSD et vents verticaux) avec celles de l'algorithme TB2015 et avec les spectres du WACR. Nous allons néanmoins revenir auparavant (section 6.3.1) sur la capacité de notre algorithme à reproduire les observations, afin d'évaluer ses performances sur cet évènement par rapport à l'évènement en Ardèche où nous ne disposions pas d'instruments de validation. Nous ne nous attarderons cependant pas trop sur cet aspect. Ensuite en section 6.3.2, nous comparerons nos restitutions à celles de TB2015. Enfin, nous verrons les implications du changement de fréquence en section 6.3.3.

6.3.1 Résultats généraux

Afin d'estimer visuellement la capacité de VAMOS à restituer les observations qu'on lui a fournies, nous considérons d'abord la Figure V.34. Elle présente les champs temps/altitude de réflectivités et vitesses Doppler assimilées (VAMOS) qui doivent être comparées à la Figure V.32 (A et B) présentant les mêmes paramètres observés par le KAZR.

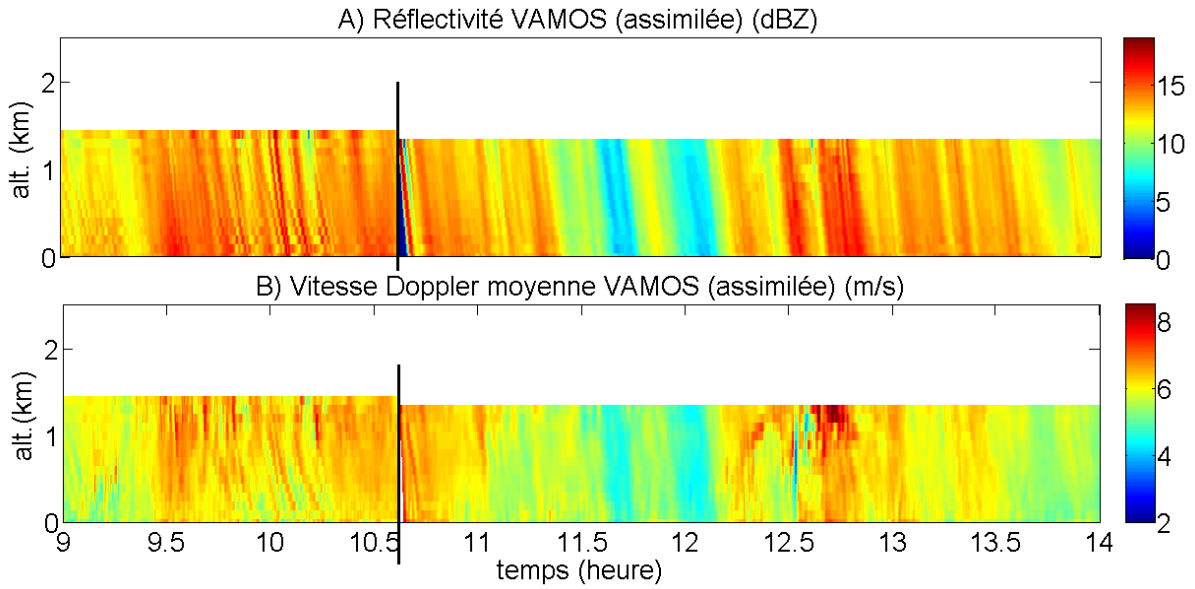

FIGURE V.34 – A-B) Facteur de réflectivité radar équivalente et vitesses Doppler moyennes assimilées pour l'évènement du 12/06/2014. Le trait noir vertical matérialise la date où l'évènement est découpé lors de l'assimilation.

Notons tout d'abord que la rupture (trait vertical noir) visible vers 10:35 sur les deux champs de la Figure V.34 s'explique par le fait que l'évènement a été découpé en sous-parties pour l'assimilation, et que l'une des transitions a lieu à ce moment-là. Au début d'une période couverte par une expérience d'assimilation, il faut que toutes les gouttes aient le temps de tomber (i.e. de parcourir les $1.4km$ verticaux) avant d'espérer obtenir des résultats cohérents.

De manière générale, on peut remarquer :

- que les profils inversés de réflectivité observés par le KAZR (Figure V.32A, vers 10:00) ne sont bien sûr pas reproduits par notre algorithme, forcé de faire tomber les gouttes.

- que pour le reste, les champs de réflectivité et, surtout, de vitesses Doppler moyennes, sont bien restitués. Notamment la structure très perturbée du champ de vitesses Doppler moyennes (vers 12:30, $1km$) est très bien restituée. Nous y reviendrons plus loin (forte influence du vent vertical ici).

Les performances statistiques de l'algorithme sont fournies dans le Tableau V.4. Ce Tableau présente les mêmes statistiques que le Tableau V.3 pour l'expérience d'assimilation en Ardèche de la section 5, à savoir des indicateurs statistiques d'erreurs sur les trois premiers moments centrés des spectres Doppler et le nombre de gouttes disdromètre (voir légende).

TABLE V.4 – Comparaison entre états observés et assimilés pour les trois premiers moments centrés des spectres Doppler et le nombre de gouttes disdromètre pour le cas en Finlande du 12/06/2014. On distingue le cas où l'on considère toutes les altitudes, et où l'on sépare les données en trois gammes d'altitudes. Les indicateurs présentés sont la MAE (Eq.(V.20)), la MAPE (Eq.(V.18)) et le biais relatif (Eq.(V.19)). Biais positif = surestimation par l'assimilation. Statistiques calculées sur la période 11:00-12:00.

		toutes altitudes	sommet $(1100-1400m)$	milieu $(700-1000m)$	bas $(300-500m)$
η	MAE (dB)	0.65	0.52	0.60	0.87
	bias (dB)	-0.45	-0.12	-0.51	-0.78
W	MAPE (%)	2.6	2.5	2.9	2.5
	rbias (%)	-2.2	-2.1	-2.5	-2.1
	moyenne(obs) (m/s)	5.53	5.56	5.54	5.50
LS	MAPE (%)	16	11	17	22
	rbias (%)	$+13$	$+5.2$	$+14$	$+21$
	moyenne(obs) (m/s)	1.15	1.31	1.13	1.02
disdro.	MAE (nbre de gouttes : mm^{-1})	48			
	moyenne(obs) (nbre de gouttes : mm^{-1})	319			

On constate que les résultats agrégés pour toutes les altitudes (colonne 3) sont assez comparables à ceux de la section 5 (Tableau V.3). La MAE pour les réflectivités est plus faible ($0.65dB$ contre $1.2dB$) tout comme la $MAPE$ sur les vitesses moyennes qui est très faible (2.6% contre 4.4%). La MAPE sur les largeurs spectrales est comparable (16% contre 14%). Les erreurs au sol sont aussi similaires (erreur de 48 gouttes/mm contre 102 gouttes/mm pour des valeurs moyennes observées respectives de 319 gouttes/mm et 1080 gouttes/mm).

Les différences principales proviennent des biais. Alors que dans le cas précédent, on n'observait pas de biais significatifs sur la réflectivité et la vitesse Doppler moyenne, on constate ici un biais moyen négatif (sous-estimation par VAMOS) de $0.45dB$ sur la réflectivité et 2.2% sur W (ce qui correspond à $0.12m/s$). Les largeurs spectrales sont davantage surestimées que précédemment par notre algorithme (de 13% contre 5.9%).

Il est très possible que ces biais s'expliquent par plusieurs phénomènes distincts. L'erreur sur la correction du radôme peut être suspectée de créer un biais systématique sur les réflectivités restituées, au moins dans une certaine mesure. En effet, on voit sur le Tableau V.4 (colonnes 4 à 6) que si l'on observe également un accroissement (en valeur absolue) du biais à mesure que l'on se rapproche du sol, les réflectivités sont néanmoins sous-estimées quelque soit l'altitude.

L'origine de cet accroissement (négatif) du biais au cours de la chute n'est pas claire, et n'est pas forcément la même que pour le cas précédent en Ardèche. En effet, on verra dans la suite que ni notre algorithme ni celui de TB2015 ne produit d'augmentation des valeurs des

paramètres liés à la DSD (taux de pluie par exemple) lors de la chute des gouttes (au contraire, voir Figures V.37). Or le phénomène principal affectant la réflectivité sur la colonne (et uniquement elle) est l'atténuation due à la pluie (et aux gaz). Une sous-estimation de cette atténuation dans notre algorithme produirait bien un biais positif de plus en plus fort à mesure que l'on s'élève dans la colonne (à la limite, sans aucune atténuation, la réflectivité reste constante sur la colonne alors qu'elle devrait diminuer quand on monte en présence d'atténuation). C'est bien ce que l'on constate sur la troisième ligne du Tableau V.4. Le fait que l'on semble sous-estimer légèrement la réflectivité de manière générale et que l'on néglige l'atténuation due aux gaz va dans le sens d'une erreur sur les atténuations.

Le léger biais sur la vitesse Doppler moyenne pourrait lui venir de l'impact du vent horizontal. On a en effet montré sur données simulées (voir section 4.3 et en particulier section 4.3.5) qu'en présence de vent horizontal, notre algorithme aurait tendance à surestimer le nombre de grosses gouttes et, pour compenser, à décaler légèrement les spectres Doppler vers les petites vitesses. Ainsi, il diminue le vent vertical et donc la vitesse Doppler moyenne. (On évoquait alors la valeur de $0.2m/s$.) On constate bien ici une sous-estimation de la vitesse Doppler moyenne de cet ordre de grandeur. On constatera effectivement une sous-estimation du vent vertical du même ordre de grandeur par rapport aux résultats de TB2015, et une (légère) tendance à la surestimation du nombre de grosses gouttes.

Concernant la largeur spectrale (Tableau V.4) et sa tendance verticale (surestimation d'autant plus forte que l'on s'approche du sol), une partie du problème vient probablement de la diffusion numérique. Plus la hauteur de chute est grande (donc plus on s'approche du sol) et plus les DSD, donc les spectres, sont lissés par ce phénomène. Si l'évènement est calme, et présente peu de turbulence atmosphérique, comme il semble que ce soit globalement le cas ici, notre algorithme ne peut que surestimer les largeurs spectrales dans les couches inférieures.

6.3.2 Comparaison avec la méthode de Tridon and Battaglia (2015)

Nous allons dans cette section comparer nos restitutions avec celles de TB2015. Nous étudierons d'abord les restitutions de vent vertical puis celles concernant la DSD et ses moments (contenu en eau liquide, taux de pluie, facteur de réflectivité radar). On note tout d'abord que l'algorithme TB2015 fournit des restitutions à la résolution du radar, soit $30m$ verticalement et environ $2s$. Nos restitutions sont moins fines puisque les observations sont moyennées à $100m$ et $30s$. Il est donc normal que les structures fines - sur le vent en particulier, qui varie plus rapidement - soient moins bien restituées.

Restitution du vent vertical Concernant le vent vertical, on présente la Figure V.35, qui représente les champs de vent restitués par VAMOS (A) et TB2015 (B) entre 10:45 et 12:30 et entre le sol et $1.4km$ d'altitude. La Figure V.36 présente, pour trois altitudes, les séries temporelles de vent vertical obtenues par les deux méthodes.

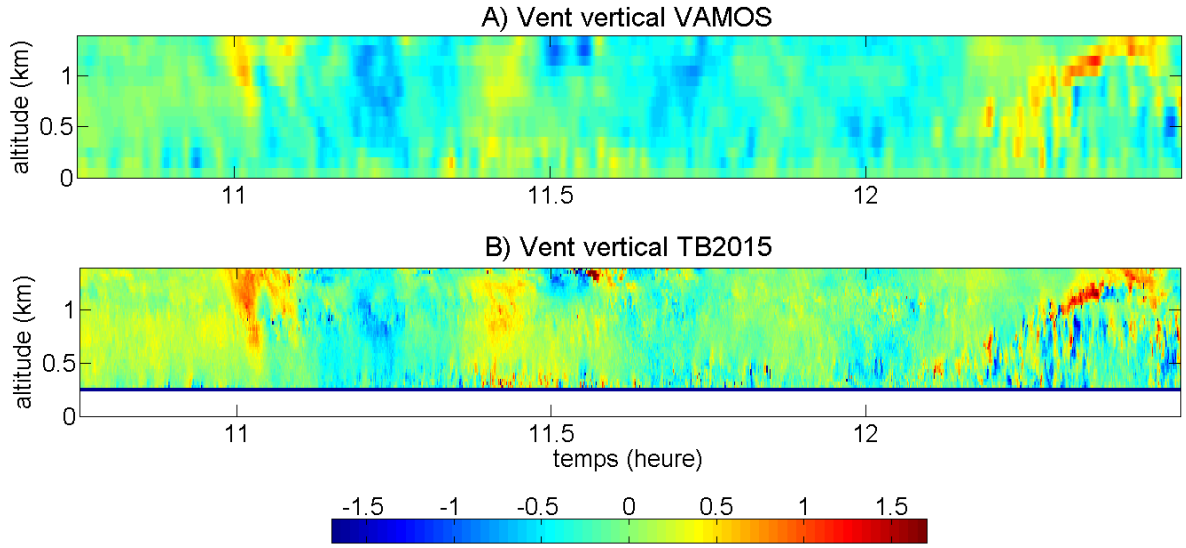

FIGURE V.35 – Vents verticaux VAMOS (sorties d'assimilation, A) et TB2015 (B) entre 10:45 et 12:30. Positif vers le bas. Unité : m/s.

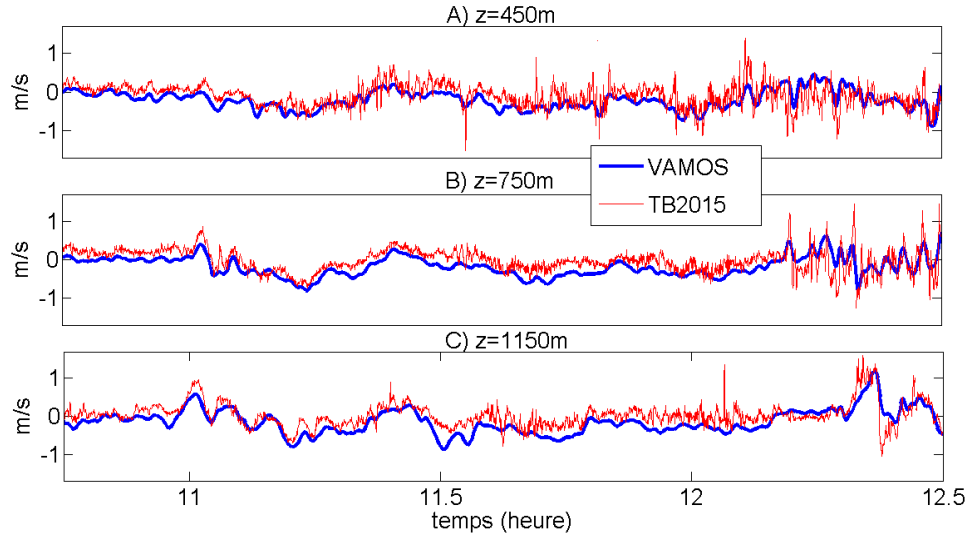

FIGURE V.36 – Séries temporelles de vents verticaux VAMOS (sorties d'assimilation) (bleu) et TB2015 (rouge) pour trois altitudes. A) 450m. B) 750m. C) 1150m. Positif vers le bas.

On constate que :

- les structures principales sont les mêmes. Notamment les vents descendant (vent positif, orange) sur le haut de la colonne vers 11:00 ou vers 12:15-12:30 ou les vents ascendant vers 11:15 ou vers 12:00 (bas de la colonne).

- notre algorithme sous-estime légèrement les vents verticaux par rapport à TB2015 (c'est-à-dire produit des vents verticaux davantage orientés vers le bas). Ce résultat ne semble pas dépendre de l'altitude (Figure V.36). Le fait que notre algorithme sous-estimait aussi les vitesses Doppler moyennes à toutes altitudes (Tableau V.4) nous conduit à penser que nous sous-

estimons bien le vent vertical par rapport à la réalité. Nous avons déjà fourni une explication possible à ce phénomène (influence du vent horizontal, voir section 6.3.1)

- il ne semble pas y avoir de biais particulier à une période donnée (par exemple, il ne semble pas y avoir de période où l'on surestime significativement le vent). Ce résultat doit être mis en parallèle avec les résultats du paragraphe suivant sur la DSD (on verra que nos résultats sur la DSD semblent meilleurs vers 11:30-12:15 que vers 11:00-11:30 par exemple). Cela montre que nos résultats sur le vent ne dépendent pas directement de ceux sur la DSD, ce qui est un bon résultat (notre algorithme n'utilise pas le vent pour expliquer une variabilité des spectres Doppler due à la DSD).

Restitution des DSD Nous allons à présent examiner les restitutions de DSD et de ses moments associés. Intéressons nous d'abord à 3 des moments de la DSD : le taux de pluie (RR, pas formellement un moment mais proche du moment d'ordre 3.67 modulé par le vent vertical), le facteur de réflectivité radar (Z, moment d'ordre 6) et le contenu en eau liquide (LWC, moment d'ordre 3). La Figure V.37 présente les champs temps/altitude de ces 3 moments assimilés (gauche) et TB2015 (droite). La Figure V.38 présente les profils verticaux moyens de ces 3 mêmes moments, moyennes calculées entre 10:45 et 12:30 (période correspondant à la Figure V.37). La Figure V.39 présente les séries temporelles de ces trois moments pour deux altitudes, en bas de la colonne ($450m$, A, C, E) et en haut ($1050m$, B, D, F).

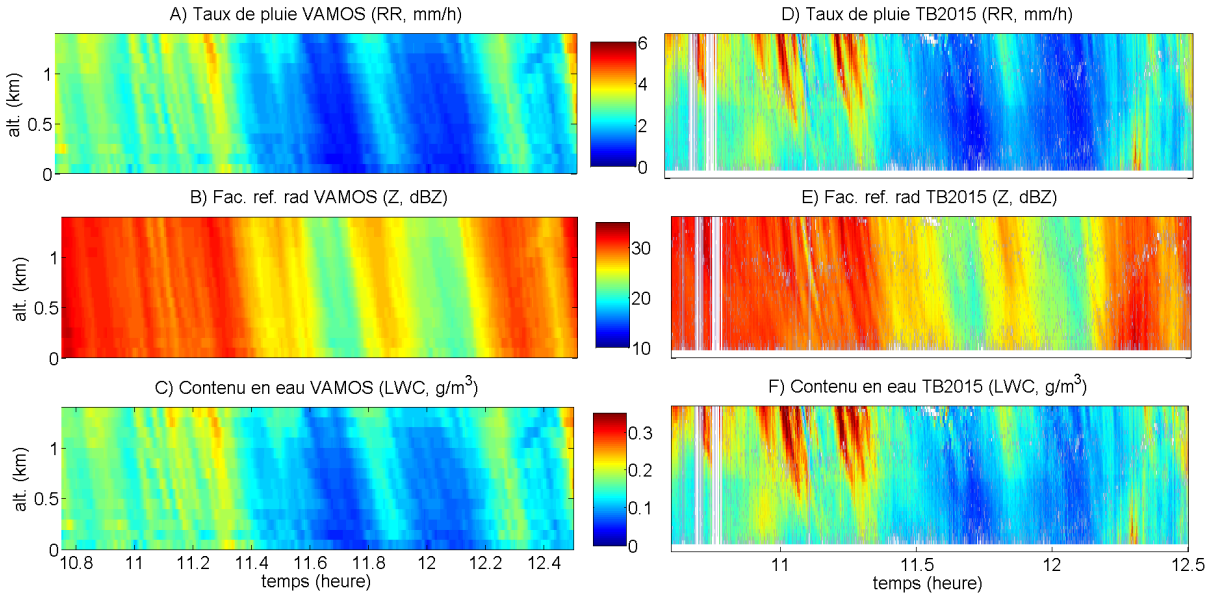

FIGURE V.37 – Moments de la DSD restitués par notre algorithme d'assimilation (VAMOS, gauche) ou par TB2015 (droite). On présente : Taux de pluie RR (A et D) ; Facteur de réflectivité radar Z (B et E) ; Contenu en eau liquide LWC (C et F).

Vue la Figure V.37, on peut subdiviser l'évènement en deux parties : une première entre 10:45 et 11:30, où l'on trouve les valeurs les plus fortes pour les moments, et une période plus

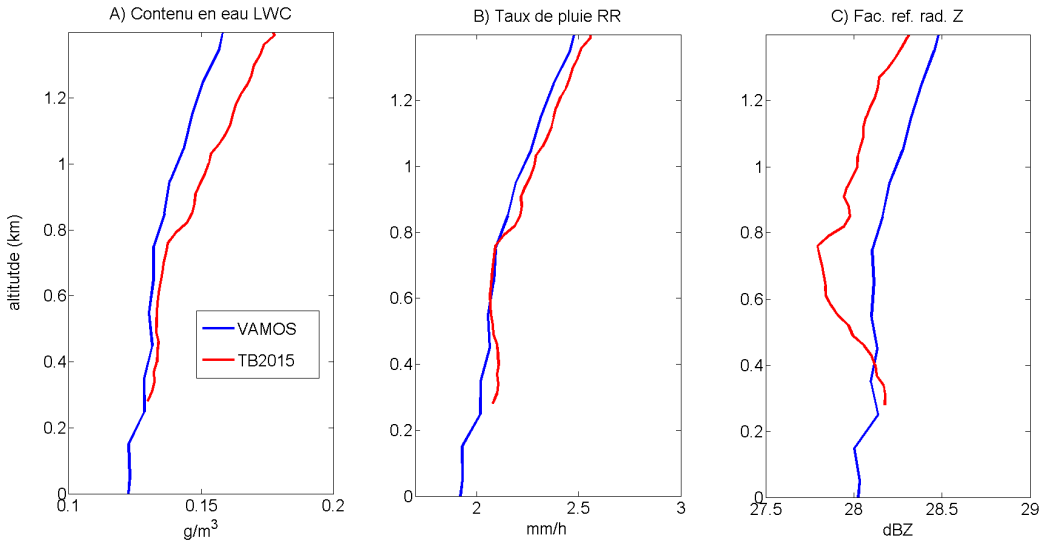

FIGURE V.38 – Profils verticaux moyens (sur la période 10:45-12:30) des contenus en eau liquide (A), taux de pluie (B) et facteurs de réflectivité radar (C) VAMOS (bleu) et TB2015 (rouge).

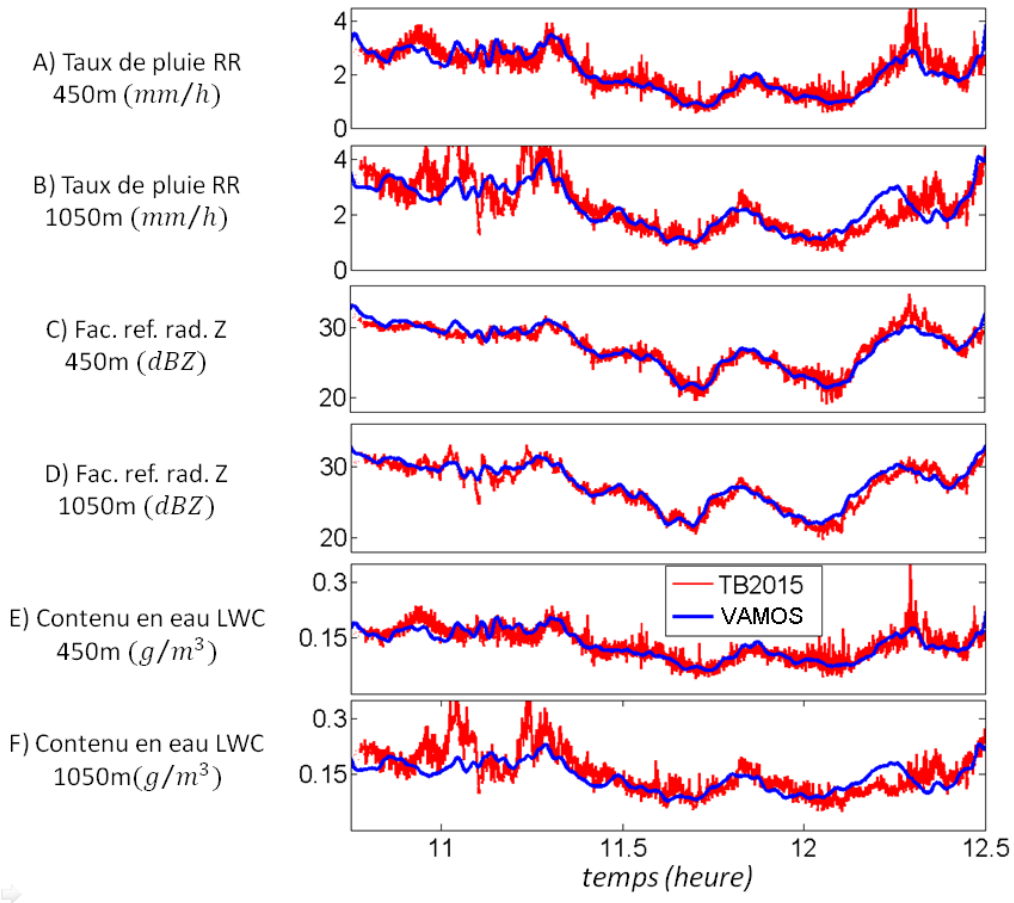

FIGURE V.39 – Séries temporelles des taux de pluie (A, B), des facteurs de réflectivités radar (C, D) et des contenus en eau liquide (E,F) pour deux altitudes (450 et 1050m), tels qu'assimilés (VAMOS, bleu) et restitués par TB2015 (rouge).

calme entre 11:30 et 12:15. Durant la première période, on observe sur les restitutions TB2015

une diminution sensible des taux de pluie et contenu en eau liquide au cours de la chute (les taux de pluie lors des plus fortes intensités passent de $6mm/h$ en haut de la colonne à environ $2.5mm/h$ en bas de la colonne). Notre algorithme restitue également une diminution du taux de pluie à mesure que l'on s'approche du sol (due dans le modèle à l'évaporation), mais moins forte. Cela se voit bien aussi sur la Figure V.39 (par exemple A et B pour RR). Si les taux de pluie sont assez conformes à $450m$ d'altitude, notre algorithme sous-estime assez fortement les pics à $1150m$ (cela est vrai aussi pour LWC, moins pour Z). L'explication la plus naturelle à ce comportement est une sous-estimation de l'évaporation dans notre modèle de propagation. On peut noter que le profil d'humidité inclus dans ce modèle a été obtenu par un radiosondage effectué à 11:36. Peut-être l'humidité était-elle un peu plus basse avant 11:30. On note également que bien qu'il semble que l'on sous-estime légèrement les taux de pluie en haut de la colonne vers ce moment et que notre algorithme présente des restitutions plus lisses (du fait également de sa résolution), les résultats, tant en ordre de grandeur qu'en variabilité temporelle et en tendance de variabilité verticale sont tout à fait compatibles.

Durant la seconde période de l'évènement (11:30-12:15), les restitutions sont encore plus proches, les courbes bleues de la Figure V.39 semblant quasiment des moyennes mobiles des courbes rouges. On remarque que cette période correspond à celle à laquelle a été effectué le radiosondage, ce qui expliquerait que la paramétrisation de l'évaporation corresponde mieux.

On note enfin une erreur un peu plus significative vers 12:15 en haut de la colonne (Figure V.39, B et F). Si l'on se réfère à la Figure V.37 (D et F), il s'agit du seul moment de l'évènement où les signatures des champs de RR ou LWC ne seraient pas diagonales ou exponentielles mais quasiment verticales, avec de plus une intensification de la pluie près du sol. Ce phénomène, peut-être dû à un mouvement horizontal local, n'est pas expliqué par le modèle de propagation inclus dans notre algorithme d'assimilation VAMOS.

Revenons à présent sur les profils verticaux moyens de la Figure V.38. On remarque qu'en moyenne, la variabilité verticale du taux de pluie est bien restituée (B), et que l'on perd dans les deux cas environ 20% du taux de pluie sur la colonne. On vient néanmoins de voir que ce comportement moyen cachait des différences selon la date. Concernant LWC (A), la sous-estimation supposée de l'importance de l'évaporation dans notre algorithme est plus visible sur la moyenne. Les contenus en eau sont conformes près du sol (où l'on dispose aussi des mesures disdromètre), puis divergent légèrement ensuite (écart de $10-15\%$ en haut de la colonne). On note tout de même que dans les deux cas, l'impact de l'évaporation semble tout à fait significatif malgré la faible hauteur de chute et le froid. Le contenu en eau sur la colonne baisse en effet d'environ 30% avec TB2015 et 23% avec notre algorithme.

Les résultats sont un peu différents en ce qui concerne le facteur de réflectivité radar Z. On note tout d'abord que celui-ci, a priori moins influencé par l'évaporation car dicté par les plus grosses gouttes, varie peu sur la colonne (entre 28 et $28.5dBZ$ pour VAMOS, soit moins de 10% de perte sur la colonne). Cependant, on note qu'alors que VAMOS avait plutôt tendance à légèrement sous-estimer les deux autres paramètres par rapport à TB2015, il

semble que l'on surestime plutôt Z. On peut logiquement penser que cette surestimation est liée à une surestimation du nombre de grosses gouttes. Nous allons donc à présent regarder directement les DSD notamment pour évaluer cette hypothèse.

La Figure V.40 présente une représentation temps/diamètres des DSD (logarithmiques) assimilées (A) et TB2015 (B), toujours entre 10:45 et 12:30, pour l'altitude de $750m$ (c'est-à-dire au milieu de la colonne atmosphérique). La Figure V.41 présente pour la même altitude les DSD moyennes sur deux courtes périodes de $5min$: entre 11:10 et 11:15 (A) et entre 12:15 et 12:20 (B). La première période se trouve au milieu de la partie de l'évènement à intensité assez forte mentionnée plus haut. La deuxième période correspond au second pic visible notamment sur la Figure V.40.

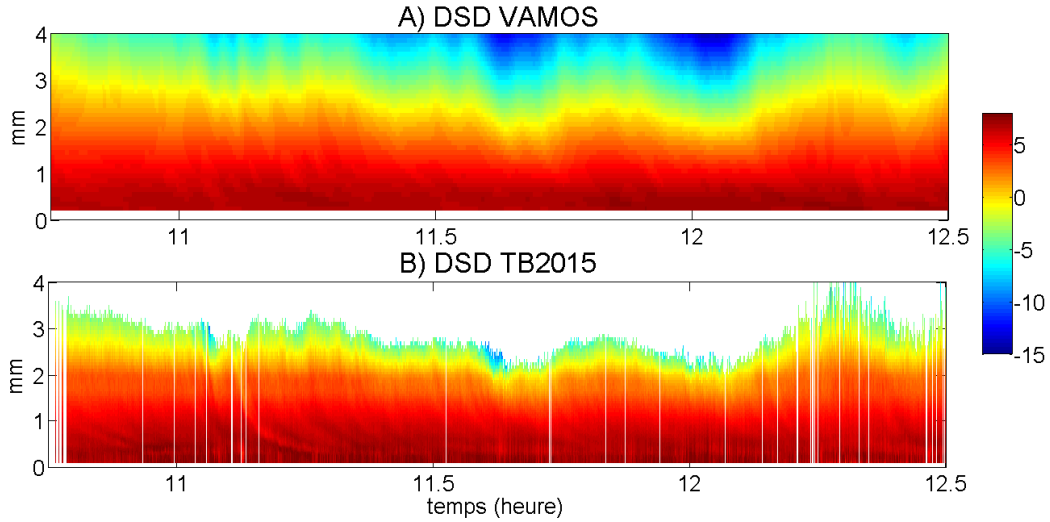

FIGURE V.40 – DSD à $750m$ d'altitude assimilées (A) et TB2015 (B). Unité : $ln(m^{-3}mm^{-1})$.

On note tout d'abord les résultats généraux suivants :

- Comme on pouvait s'y attendre vus les résultats précédents sur les moments, la variabilité temporelle et l'intensité des DSD sont conformes entre les deux méthodes de restitution.

- Notre algorithme restituant des DSD paramétrées en haut de la colonne, on obtient ensuite des valeurs non nulles quelque soit le diamètre. TB2015 restituant au contraire des histogrammes (bin DSD), on ne voit aucune goutte pour des diamètres au-delà de $4mm$, et même parfois au-delà de $2 - 2.5mm$. Cette différence (et la surestimation du nombre de grosses gouttes par notre restitution qui en découle mécaniquement) peut expliquer en partie la surestimation du facteur de réflectivité radar par notre algorithme constaté à partir de la Figure V.38C.

Concernant la forme des DSD (Figure V.41), on peut également noter plusieurs points :

- la forme générale est la même, à savoir queues des distributions exponentielles à pentes relativement proches et distribution tendant à s'aplatir sur les petits diamètres.

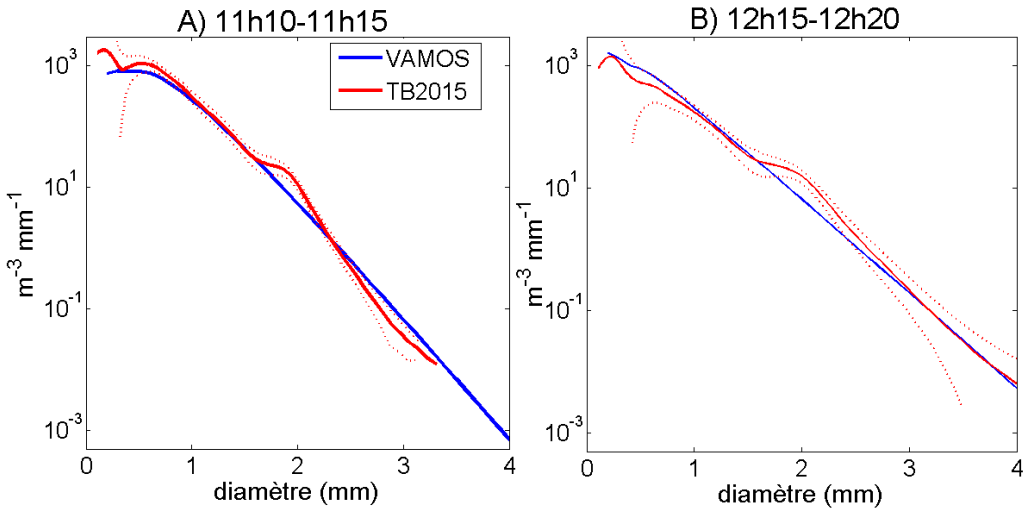

FIGURE V.41 – DSD moyennes à $750m$ d'altitude entre 11:10 et 11:15 (A) et entre 12:15 et 12:20 (B). On présente les DSD assimilées (bleu) et TB2015 (rouge continus) ainsi que les barres d'erreurs associées aux DSD TB2015 (pointillés rouges).

- les distributions TB2015 sont moins régulières et en particulier présentent des inflexions (quasiment des pics) vers $1.8 - 2.1mm$, qui ne sont absolument pas restitués par notre algorithme. Cette inflexion peut sans doutes en partie expliquer la légère sous-estimation de notre algorithme constatée sur les moments d'ordre pas trop élevé (LWC) fortement affectés par les DSD à ces diamètres. L'origine de cette différence reste à éclaircir mais pourrait être liée à des phénomènes de coalescence / break-up (bien qu'on ait ici moins de $2km$ de chute et des taux de pluie faibles).

- la queue des distributions, concernant les grosses gouttes, très peu nombreuses, est plus sujette à caution (on voit bien la forte extension de l'erreur pour TB2015 sur la Figure V.41B au-delà de $2.5mm$). Dans le premier cas (A), TB2015 sous-estime légèrement le nombre de grosses gouttes par rapport à VAMOS (cohérent avec les résultats déjà mentionnés sur les grosses gouttes). Ce n'est pas vrai pour le second cas (B) avec des queues très similaires.

- sur les petites gouttes ($< 0.7mm$), les barres d'erreur sont plus larges également (notamment pour les restitutions basées seulement sur les radars, qui reçoivent peu d'énergie de ces diamètres). Là encore cependant, les restitutions sont assez conformes entre elles, avec comme on l'a vu une nette inflexion de la DSD.

6.3.3 Changement de fréquence : étude des spectres Doppler à $95GHz$

Nous allons à présent étudier la conformité de nos restitutions de DSD et vents verticaux avec les spectres Doppler mesurés par le WACR à $95GHz$ (contrairement à TB2015 nous ne l'avons pas utilisé pour la restitution). Pour cela, nous calculons les spectres à $95GHz$ induits par les sorties de notre algorithme VAMOS. Disposant des DSD en haut de la colonne et des vents verticaux en sortie d'assimilation, il suffit d'exécuter le modèle direct et les opérateurs d'observations des spectres Doppler (voir section 2.3) avec les sections efficaces de

rétrodiffusion et d'extinction de Mie calculées non plus à $35GHz$ mais à $95GHz$.

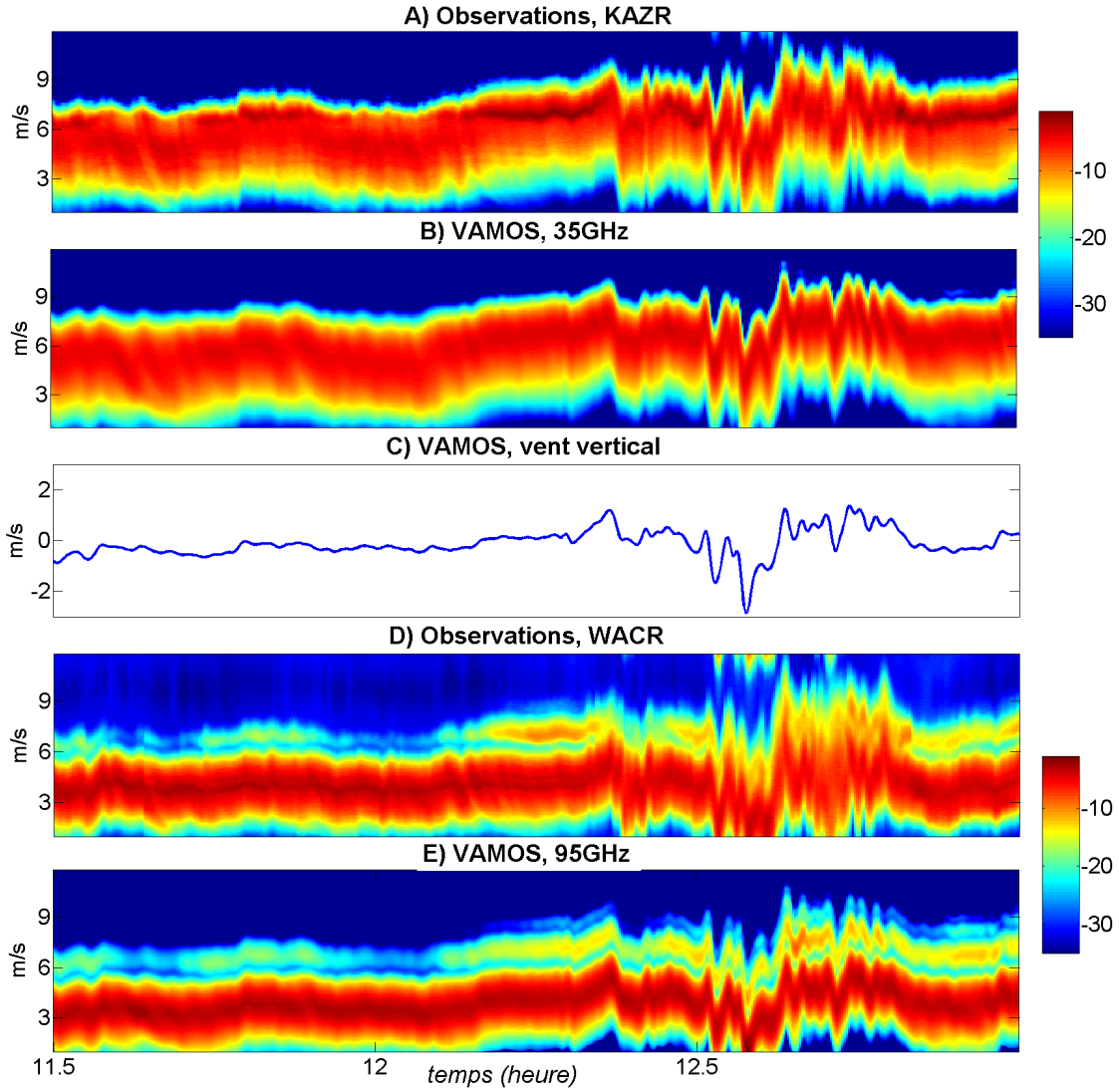

FIGURE V.42 – Spectres normalisés observés (A, D) et assimilés (VAMOS) (B, E) à 35 et $95GHz$. Vent vertical assimilé (VAMOS, C). Résultats pour l'altitude $750m$. Voir texte pour commentaires détaillés.

Sur la Figure V.42, on a représenté, pour une altitude de $750m$ (milieu de la colonne) :

- à $35GHz$, les spectres normalisés (plans temps/vitesses Doppler) observés par le KAZR (A) et restitués par notre algorithme VAMOS (B).

- le vent vertical restitué par notre VAMOS (C).

- à $95GHz$, les spectres normalisés (plans temps/vitesses Doppler) observés par le WACR (D) et restitués par VAMOS (E). Les spectres ici présentés sont normalisés, c'est-à-dire divisés chacun par leur intégrale. La réflectivité (η) de chaque spectre vaut donc 1. Cela permet de s'affranchir des biais déjà mentionnés (en particulier dûs à l'atténuation radôme, biais donnés dans les Tableaux V.4 et V.5) pour se concentrer sur leur forme.

On note bien sûr que les données présentées sur la Figure V.42A (KAZR) ont été utilisées pour produire nos restitutions. La comparaison A-B évalue donc la capacité de notre

algorithme à reproduire les observations qu'on lui a fournies (et pas la qualité intrinsèque de ces restitutions). Les données du E (WACR) n'ont en revanche pas été utilisées pour produire nos restitutions.

Il est tout d'abord intéressant de constater que travailler à une fréquence ou à une autre permet de faire ressortir ("d'allumer") une gamme ou une autre du spectre Doppler. Cela est dû à la variation de la section efficace de rétrodiffusion de Mie (σ) avec la fréquence et le diamètre. Le maximum d'énergie est ainsi vers $5-8m/s$ à $35GHz$ et vers $2.5-5m/s$ à $95GHz$ (avec un second maximum moins prononcé correspondant au maximum à $35GHz$). Les deux fréquences sont cependant sensibles au vent vertical de la même manière.

On note ainsi que notre algorithme est capable de restituer de manière très satisfaisante la forme des spectres observés à $94GHz$, notamment sur les petites vitesses Doppler, peu couvertes par le KAZR. Cela signifie que nous obtenons des DSD cohérentes sur une large gamme de diamètres.

Au niveau des différences, on peut noter que :

- on obtient (Figure V.42E) des spectres sensiblement moins lissés durant la période la plus agitée de l'évènement (après 12:30). En particulier, le minimum d'énergie vers $6m/s$ dû à la résonance de Mie est toujours visible sur nos spectres restitués alors qu'il disparaît fréquemment dans le lissage sur les observations. Cela pourrait laisser penser que l'on sous-estime la turbulence atmosphérique à cette période (ce qui n'est pas aussi clairement visible au même moment sur les spectres à $35GHz$, voir Figure V.42A et B). Nous reviendrons plus en détails sur cette question de la turbulence et de sa gestion dans notre algorithme sur le cas plus perturbé traité en section 7.

- on note sur les observations WACR (D) un peu de repliement spectral dans les périodes où l'on observe les vents verticaux les plus forts (juste après 12:30). En l'occurrence, le vent vertical étant orienté vers le haut d'environ $2.5m/s$, une partie des gouttes (les plus petites) possèdent une vitesse négative (c'est-à-dire remontent). Elles sont enregistrées par le radar modulo la vitesse de Nyquist, donc ici avec une vitesse positive très forte (autour de $11m/s$). Le KAZR étant moins sensible aux petites énergies, ce phénomène est presque transparent pour lui (voir Figure V.42A). Pour des évènements plus perturbés, la question pourrait cependant se poser quelques soit la fréquence. La présence d'un tel repliement - non géré par le modèle - dans des observations utilisées pour l'apprentissage causerait au mieux un coût résiduel fort, au pire des résultats erronés. Il pourrait éventuellement être intéressant d'ajouter un module simulant ce repliement spectral dans le modèle (dans l'opérateur d'observation des spectres Doppler). Ajouter un tel module dans notre algorithme est assez simple, et dans tous les cas beaucoup plus facile que de déplier les spectres a priori.

Le Tableau V.5 présente les mêmes statistiques que le Tableau V.4 (écart entre les trois premiers moments centrés des spectres Doppler) mais évaluées entre les observations WACR et

TABLE V.5 – Mêmes statistiques que Tableau V.4 calculés entre les moments centrés des spectres à $95GHz$ déduits de l'assimilation et mesurés par le WACR.

		toutes altitudes	sommet $(1100-1400m)$	milieu $(700-1000m)$	bas $(300-500m)$
η	MAE (dB)	0.63	0.77	0.60	0.39
	bias (dB)	+0.09	+0.50	−0.17	−0.09
W	MAPE (%)	8.6	8.7	8.8	8.9
	rbias (%)	−8.4	−8.3	−8.7	−8.9
	moyenne(obs) (m/s)	3.99	4.00	3.99	3.96
LS	MAPE (%)	17	19	15	16
	rbias (%)	−17	−19	−15	−16
	moyenne(obs) (m/s)	0.93	0.98	0.91	0.89

celles simulées à $95GHz$ à partir de nos restitutions VAMOS.

La première constatation est que bien que l'on estime ici des écarts de nos restitutions avec des données non utilisées pour l'assimilation, ces écarts sont très comparables à ceux du Tableau V.4 entre données d'apprentissage et sorties d'assimilation. L'erreur sur la réflectivité (MAE) est quasiment la même ($0.63dB$ contre 0.65), celle sur les largeurs spectrales également. Seule l'erreur sur la vitesse moyenne W (MAPE) est plus forte (8.6% contre 2.6%) mais reste raisonnable.

Sur les réflectivités (intégrale des spectres Doppler), on note que l'on observe la même tendance verticale que pour le KAZR (plus on monte dans la colonne, plus on observe un biais positif), dû peut-être au fait que l'on néglige l'atténuation due aux gaz, assez forte à cette fréquence. On n'obtient cependant plus ici de biais systématique. La correction de l'atténuation radôme que l'on a estimée est plus modeste, de l'ordre de $5-6dB$ à $95GHz$ (contre $15-17dB$ à $35GHz$ comme on l'a vu). Elle est sûrement moins impactante à cette fréquence.

Concernant les largeurs spectrales, on note que contrairement au cas précédent (Tableau V.4) où notre algorithme surestimait les largeurs observées, celles du WACR sont ici assez sensiblement sous-estimées. Nos spectres restitués sont moins lisses comme mentionné plus haut. Il semblerait que la turbulence - à certaines périodes de l'évènement - soit plus directement visible sur les spectres à $95GHz$. Elle est donc assez mal restituée par VAMOS qui ne se base que sur les spectres à $35GHz$.

On peut enfin noter que notre algorithme, comme précédemment, sous-estime systématiquement les vitesses moyennes. Le fait que cette sous-estimation soit un peu plus forte en pourcentage que pour le KAZR s'explique en partie naturellement par le fait que la vitesse moyenne mesurée est plus basse ($3.99m/s$ contre $5.53m/s$). Mais cette explication ne suffit pas à expliquer la différence observée (écart moyen de l'ordre de $0.3m/s$ contre $0.12m/s$ précédemment). Le repliement spectral déjà mentionné explique sans doute une partie de l'écart supplémentaire. Le fait d'avoir sur les observations WACR des échos de très petites gouttes enregistrées à $11m/s$ biaise fortement (bien que ponctuellement) les vitesses

moyennes.

6.4 Conclusions, perspectives et ouverture sur le troisième cas d'étude

Dans cette section, nous avons appliqué VAMOS, notre algorithme de restitution des DSD et vents verticaux à un second évènement de pluie calme et stratiforme en Finlande, évènement pour lequel nous disposions de davantage d'instruments de mesures que pour le premier cas en France, et notamment d'un second radar Doppler à une fréquence différente ($95GHz$). Nous avons ainsi pu évaluer nos restitutions par rapport à celles obtenues par l'algorithme présenté dans Tridon and Battaglia (2015).

Nous avons montré que nous obtenions des résultats très cohérents avec ceux de Tridon and Battaglia (2015). La structure du champ des vents verticaux est ainsi très bien restituée, notamment dans les périodes assez perturbées (valeurs absolues au-dessus de $0.5m/s$). Bien que pour l'instant à moins bonne résolution spatiale et temporelle, notre algorithme semble assez robuste et présente l'avantage de n'utiliser les données que d'un radar Doppler à $35GHz$ pouvant fournir des mesures pour des évènements de pluie plus forts qu'un radar à $95GHz$. Sur la restitution des DSD et de leurs moments, on a également montré la cohérence de nos restitutions.

On a néanmoins remarqué un faible biais systématique sur nos restitutions du vent vertical (de l'ordre de $-0.1/-0.2m/s$) associé à une légère surestimation du nombre de grosses gouttes, phénomène déjà constaté sur données simulées et associé alors à l'erreur modèle due au fait de négliger le vent horizontal. On note également que la qualité de nos restitutions est fortement conditionnée à la pertinence de la paramétrisation du modèle de propagation, par exemple du module d'évaporation, et ici à la bonne prise en compte de l'atténuation due au radôme du radar. Le modèle de propagation présente l'avantage d'apporter une cohérence physique permettant d'obtenir de bons résultats avec un seul radar mais doit être correctement paramétré pour ne pas induire de biais dans les résultats.

En comparant les spectres observés à $95GHz$ (et non utilisés dans la production de nos résultats) avec les spectres à la même fréquence qu'induiraient nos résultats, nous avons montré que l'on obtenait des résultats cohérents sur une large gamme de vitesses Doppler (et donc de diamètres), et que l'on obtenait des écarts comparables à ceux calculés par rapport aux observations d'apprentissage (spectres à $35GHz$). Cela est un bon résultat, montrant également une aptitude du modèle à extrapoler les observations grâce à la prise en compte de la physique et aux expression paramétriques des DSD.

À partir de ces résultats, on peut proposer les études et améliorations suivantes :

- contrôle dans l'algorithme d'assimilation de l'atténuation du signal due au radôme du radar.

- contrôle dans l'algorithme des profils verticaux d'humidité relative (en supposant une variabilité temporelle et verticale suffisamment lente pour ne pas trop augmenter le nombre de

degrés de liberté).

- diminution du temps d'intégration des spectres Doppler pour raffiner les restitutions. Cela devrait être possible - dans une certaine mesure - pour un évènement comme celui-ci pas trop perturbé.

- Inclusion de l'atténuation due aux gaz (en particulier si l'on assimile des mesures à $95GHz$).

- Gestion du repliement spectral dans l'algorithme d'assimilation.

- assimilation à la fois des données KAZR et WACR, pour couvrir directement toute la gamme de vitesses Doppler, ou assimilation des données WACR seules puis comparaison avec les mesures KAZR pour évaluer les performances de notre algorithme aux deux fréquences.

Nos résultats ayant été validés, nous allons à présent - et pour terminer - étudier un évènement un peu plus complexe, c'est-à-dire avec une plus grande hauteur de chute, des taux de pluie légèrement plus forts (inférieurs à $10 - 12mm/h$ tout de même) et des vents verticaux un peu plus forts également. Nous nous attarderons en particulier sur la gestion de la turbulence par notre algorithme sur un cas réel.

7 Troisième cas d'étude : Darwin

7.1 Présentation et enjeux

Dans cette troisième et dernière étude de cas, nous allons appliquer notre algorithme VAMOS à un contexte météorologique légèrement différent. Nous utiliserons en effet des données provenant de Darwin, ville du nord de l'Australie soumise à un climat tropical.

Contrairement aux deux cas précédents (Ardèche et Finlande), caractérisés par des pluies froides (faible hauteur de chute sous la bande brillante, environ $2km$) très régulières, nous travaillerons ici sur une journée de pluie chaude (bande brillante vers $4.5km$) marquée par des évènements successifs présentant toujours des pluies faibles ($<10mm/h$) mais relativement complexes, au sein desquels l'évaporation, la turbulence atmosphérique et le vent horizontal sont susceptibles de jouer des rôles importants.

Nous analyserons d'abord les évènements considérés à la lumière des différentes observations disponibles. Les enjeux du travail présenté ici sont d'analyser les évènements à partir des résultats fournis par notre algorithme afin de mettre en lumière les phénomènes physiques prévalant (évaporation, mouvements horizontaux, vent vertical, turbulence, coalescence/break-up). À partie des observations et des sorties de notre algorithme, nous ferons également une analyse extensive, au sol comme en altitude, de l'un des évènements enregistré, afin de montrer sur un exemple détaillé tous les phénomènes que notre méthode permet de mettre à jour. Le dernier objectif de cette étude est d'expliciter la façon dont notre algorithme représente les effets de la turbulence sur les spectres Doppler dans un cas réel et les perspectives que cela ouvre.

Dans la section 7.2, nous présentons le site d'étude, les instruments et les données disponibles ainsi que les évènements de pluie considérés. Nous reviendrons plus en détails sur les données permettant d'évaluer l'humidité relative. En section 7.3, nous présentons les résultats. Nous étudions plus spécifiquement l'impact de l'évaporation et la représentation des effets de la turbulence par notre algorithme. La section 7.4 conclut cette étude.

7.2 Instruments disponibles / Présentation de l'évènement

Dans cette section, nous traitons un évènement pluvieux ayant été enregistré dans le nord de l'Australie, près de la ville de Darwin, localisée sur la Figure V.43 (croix rouge). Il s'agit d'une région tropicale, avec une saison humide durant l'été austral, de novembre à mars. D'après les données climatiques du Bureau of Meteorology (`http://www.bom.gov.au/climate/averages/tables/cw_014015.shtml`), la pluviométrie est maximale en janvier, avec en moyenne 18.9 jours avec plus de $1mm$ d'eau et $424mm$ de cumul mensuel. Les températures moyennes durant ces 5 mois sont peu variables et comprises entre $24.9\,^{\circ}C$ (minimales) et $32.2\,^{\circ}C$ (maximales). On peut trouver des données climatiques complètes sur le site du Bureau.

Nous nous intéresserons ici à un évènement ayant eu lieu au début de l'été (novembre)

2014 et enregistré sur un site situé aux coordonnées (12.425S, 130.892E) et à 29.9m d'altitude (voir Figure V.43). Il s'agit d'un site où sont installés des instruments à la fois dans le cadre du programme ARM déjà présenté (https://www.arm.gov/sites/twp) et par le Bureau of Meteorology (http://www.bom.gov.au/). On le désigne par "site ARM" dans la suite.

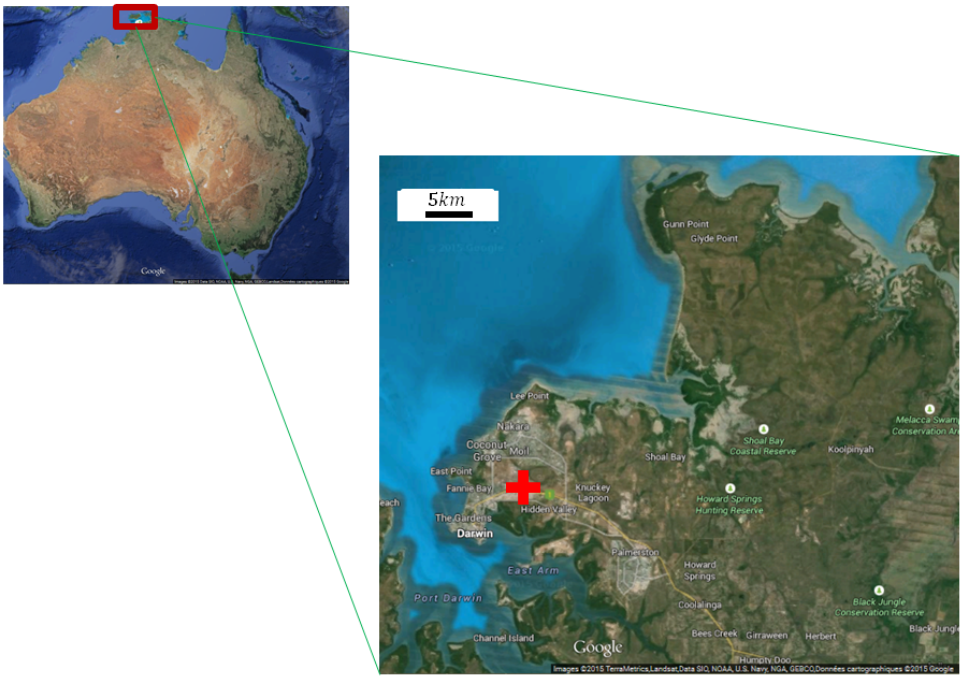

FIGURE V.43 – Localisation (croix rouge) du site ou sont disposés les instruments utilisés dans cette section (image Google).

Nous allons d'abord présenter les instruments et produits disponibles en section 7.2.1 puis nous détaillerons les caractéristiques de l'évènement étudié en section 7.2.2.

7.2.1 Instruments et produits disponibles

Données d'entrée de VAMOS Les données utilisées en entrée de l'algorithme d'assimilation proviennent des mêmes instruments que pour l'évènement en Finlande traité précédemment en section 6. Ces instruments sont le KAZR, radar Doppler pulsé à visée verticale à 35GHz et un disdromètre 2D-vidéo (2DVD dans la suite). Comme précédemment, leurs données seront intégrées spatialement et temporellement avant utilisation dans notre algorithme à 100m et 30s pour le premier, à 1min pour le second.

Par défaut (sans mention contraire), tous les champs présentés seront intégrés à la résolution des observations radar (30s).

Instruments et produits de validation Contrairement au cas précédent, on ne dispose pas ici de spectres Doppler à une autre fréquence. (On ne pourrait donc pas utiliser l'algorithme TB2015 mentionné dans la section précédente.) Nous disposons par contre des mesures du radar BASTA, introduit dans le chapitre I (voir Tableau I.2 en section 2.3.3). Il

s'agit d'un radar Doppler FMCW à $94GHz$ enregistrant la réflectivité (Z_e) et la vitesse Doppler moyenne (W). Nous l'utiliserons ici en particulier pour contrôler la restitution des vitesses Doppler moyennes de notre algorithme, et pour évaluer si celles-ci sont majoritairement déterminées par la DSD ou par le vent vertical.

On signale que l'on dispose également d'un second disdromètre, de type Joss-Waldvogel (JW dans la suite), enregistrant uniquement les DSD (et non les gouttes particulières) à $1min$. Il sera utilisé pour évaluer la variabilité des mesures de taux de pluie au sol (voir Figure V.45E par exemple). Cet instrument a été présenté au chapitre I, section 2.2.1. Nous utiliserons également (Figure V.44) les images du radar opérationnel en bande C du Bureau of Meteorology situé à Berrimah, dans la banlieue est de Darwin à quelques kilomètres du site ARM. On peut trouver les images de réflectivités de ce radar sur le site : http://www.bom.gov.au/products/IDR632.loop.shtml.

On dispose également de séries temporelles de données météorologiques au sol (température, pression, humidité relative).

On dispose enfin d'un produit, nommé mergesonde, calculé par l'ARM. Il fournit des champs temps/altitude de divers paramètres atmosphériques (vent, température, humidité, etc.), résultant de la fusion de radiosondages et de sorties du modèle numérique global ECMWF (European Centre for Medium-Range Weather Forecasts : http://www.ecmwf.int/). Ce produit peut éventuellement être corrigé par d'autres mesures (radiomètre pour la quantité d'eau dans la colonne). La fusion est opérée en privilégiant le radiosondage par rapport au modèle dès lors que l'on est près de sa date d'occurrence. Une description plus précise de ce produit peut être trouvée en ligne : https://www.arm.gov/data/vaps/mergesonde. Dans un premier temps, lors de l'étude générale de l'évènement (section 7.2.2), nous étudions les champs vent horizontal et température. Notre algorithme nécessitant des champs d'humidité relative (h_r) en entrée, nous reviendrons sur le champ humidité fourni par ce produit en section 7.2.3.

7.2.2 Présentation de l'évènement / Configuration de l'assimilation

Évènement L'évènement étudié dans cette section a eu lieu à Darwin le 04 novembre 2014, au début de l'été austral, et donc de la saison des pluies. Cet évènement, assez complexe, couvre une période assez longue - $3h02$ entre 17:58 et 21:00. Il est subdivisable en plusieurs parties. Dans la suite, nous ne le considérerons donc pas comme un tout mais comme 5 évènements successifs, que nous décrirons. La Figure V.44 présente une succession (toutes les 30 minutes) d'imagettes des précipitations sur la région comme vues par le radar opérationnel de Berrimah. Le site ARM est représenté par le triangle rouge et noir. On voit d'ors et déjà qu'un premier front pluvieux, circulant vers le nord-nord-ouest, atteint le site entre 17:30 et 18:30 environ. Il s'en suit une période plus calme, quasiment sans échos radar (19:00 - 20:00).

Finalement, entre 20:00 et 21:00, le site est survolé par un deuxième front circulant, cette fois, plutôt vers le nord-est.

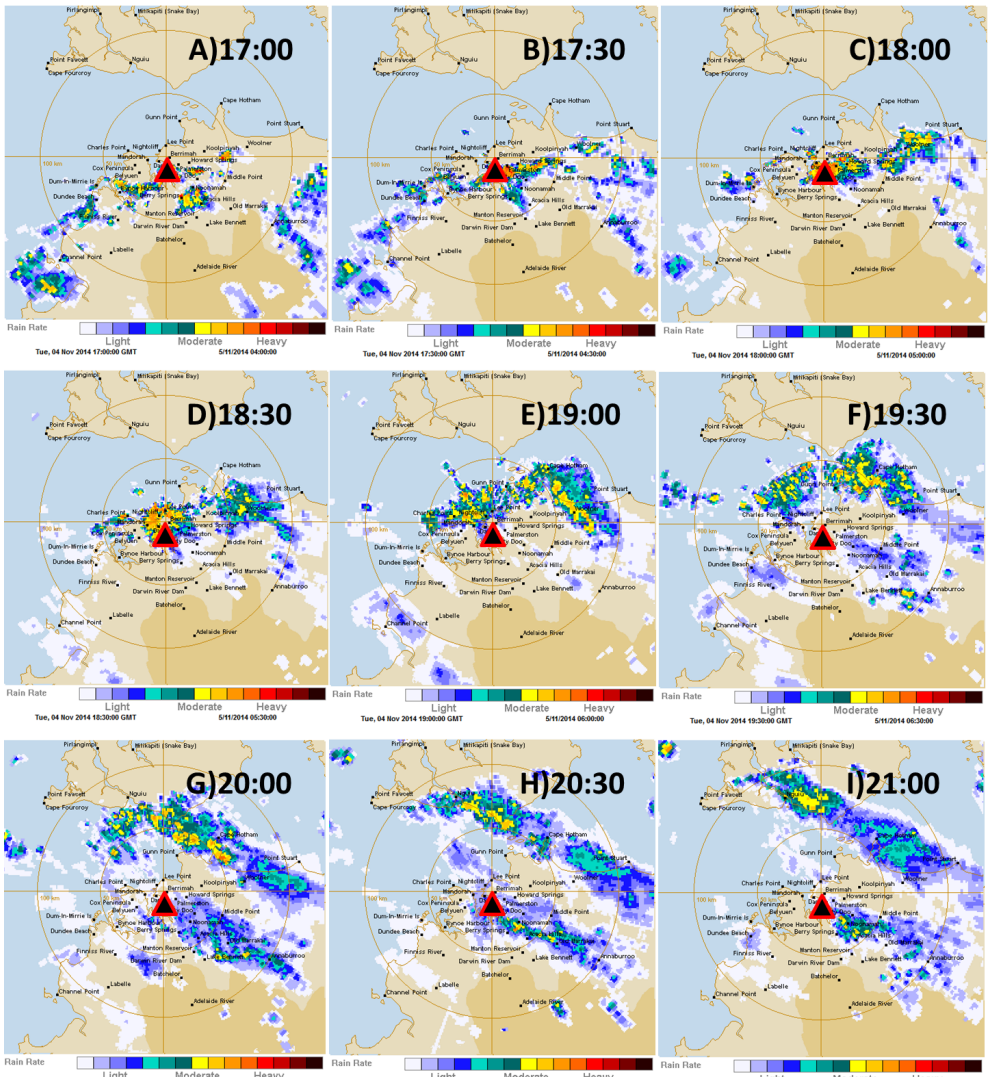

FIGURE V.44 – Imagettes des évènements pluvieux du 04 novembre 2014 observés par le radar opérationnel de Berrimah du Bureau of Meteorology. Le site ARM est indiqué par le triangle rouge.

Nous allons, à présent, analyser plus en détails ces évènements tels qu'ils sont vus depuis le site ARM.

La Figure V.45 présente des mesures du radar KAZR, à savoir la réflectivité et la vitesse Doppler moyenne (A et C respectivement), ainsi que les taux de pluie au sol, mesurés par les disdromètres 2DVD (rouge) et JW (noir). La Figure V.45 (B et D) montre aussi les restitutions de ces champs par VAMOS : nous y reviendrons plus loin. La Figure V.46 montre des profils verticaux de vent horizontal (direction en A et module en B) et de température (C) fournis par le produit mergesonde. On rappelle que ce produit effectue une fusion entre des sorties de modèles et des radiosondages. Sur la Figure V.46, on présente les profils verticaux moyens sur le début de la période (18:00-19:00, bleu), sur la fin (20:00-21:00, rouge) et tels

que mesurés par le radiosondage de 23:15 (noir). Ce radiosondage a eu lieu assez longtemps après la fin de l'évènement (21:00), ce qui fait que le produit mergesonde est largement basé sur les sorties de modèle, ce qui complique singulièrement l'analyse.

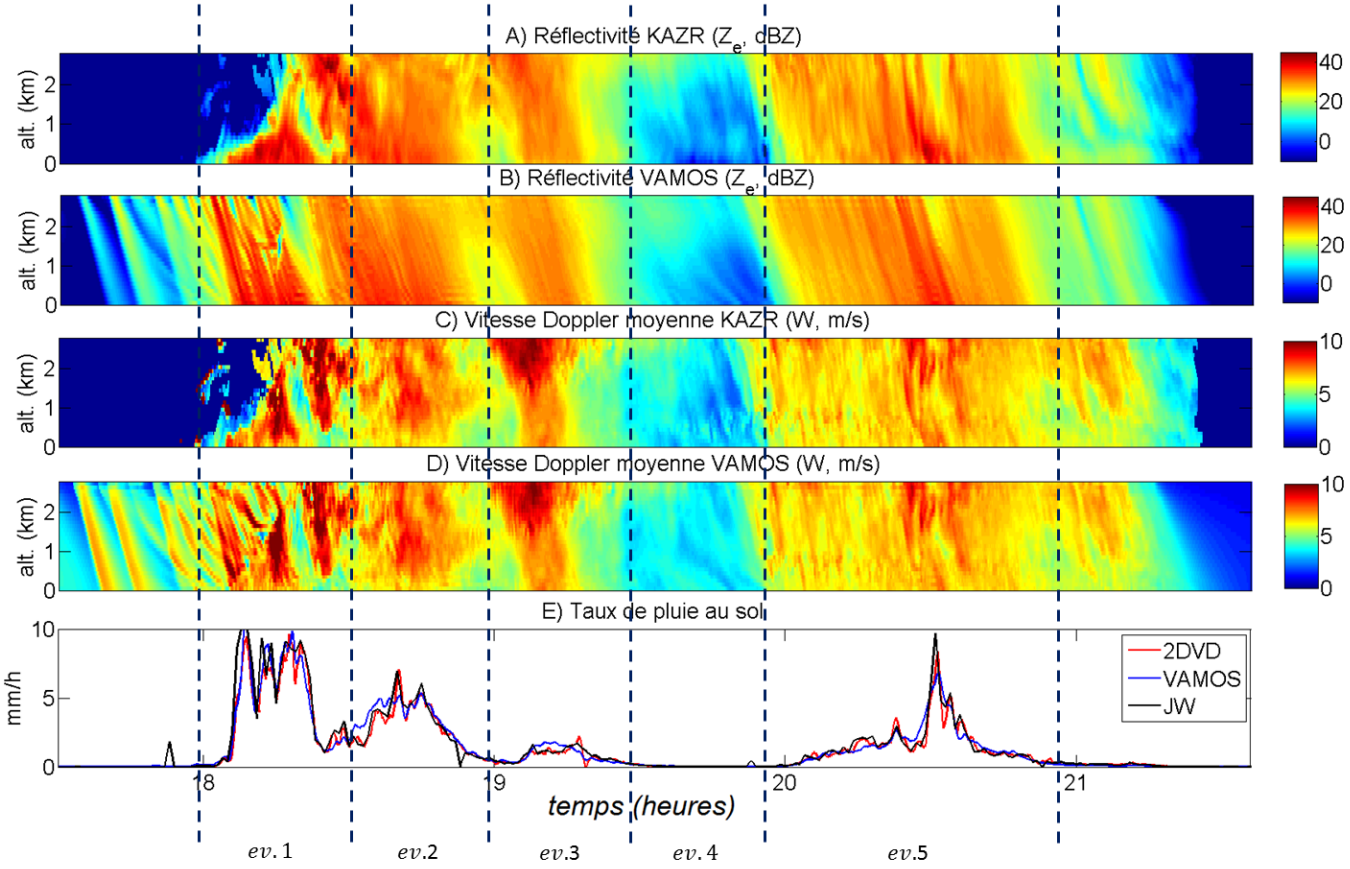

FIGURE V.45 – A-B) Facteurs de réflectivité radar équivalents Z_e observés par le KAZR (A) et assimilés dans VAMOS (B). C-D) Idem pour la vitesse Doppler moyenne W. E) Taux de pluie au sol tels que : mesurés par le 2DVD (rouge) et par le JW (noir) ainsi qu'assimilés dans VAMOS (bleu). Les pointillés délimitent de manière indicative les différents évènements au sens du Tableau V.6.

On constate bien sur la Figure V.45 l'hétérogénéité de l'évènement étudié. Nous allons donc le découper en cinq évènements successifs. On note que ce découpage, résumé dans le Tableau V.6, n'est pas basé sur des considérations physiques précises mais principalement sur l'analyse visuelle de la Figure V.45. Il a pour principale vocation de structurer les analyses qui suivront.

Avant d'étudier chacun des évènements, nous allons nous intéresser au produit mergesonde.

Revenons rapidement d'abord sur les vents horizontaux. Ceux-ci ne semblent pas trop forts (cf. Figure V.46B, moyenne sur la colonne autour de $6m/s$, assez proche des valeurs testées sur données simulées en section 4.3). Mais la variabilité sur la colonne de la direction des vents (Figure V.46A) nous donne moins de garanties que précédemment sur nos restitutions. Nous étudierons donc attentivement celles-ci pour évaluer leur capacité à

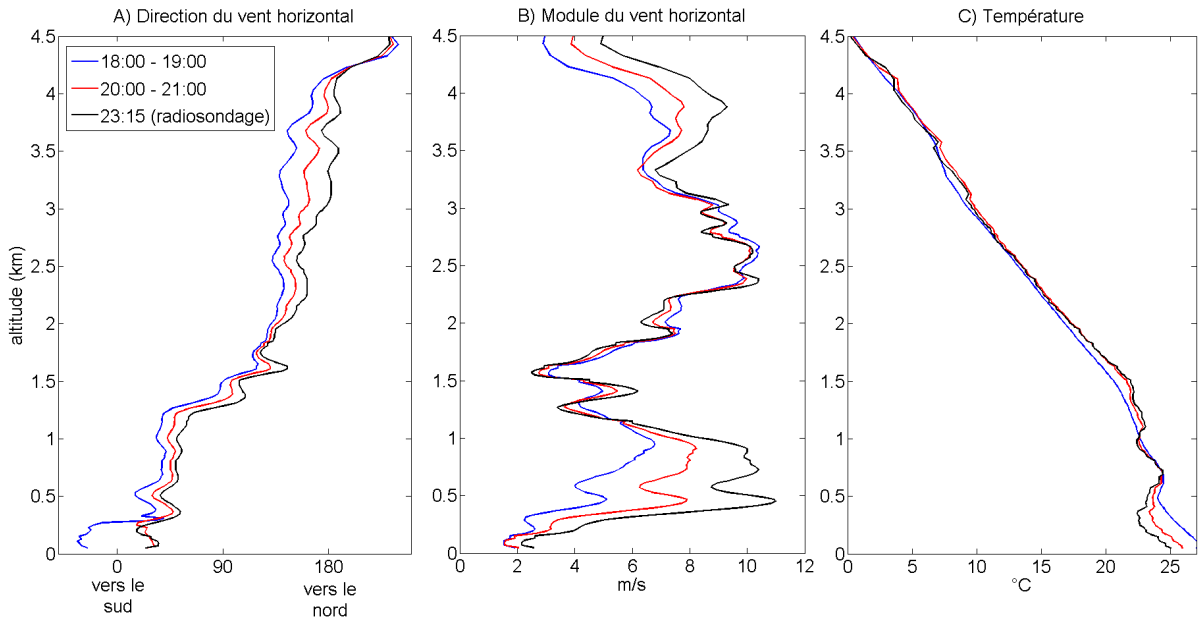

FIGURE V.46 – Profils verticaux de vents horizontaux et de températures fournis par le produit mergesonde. A) Direction du vent (0=vent vers le sud, 90=vent vers l'ouest, etc.) B) Module du vent. C) Température. Profils moyens pour la période 18:00-19:00 (bleu), 20:00-21:00 (rouge) ainsi que pour le radiosondage de 23:15 (noir).

expliquer les observations. On peut encore noter que d'après le produit mergesonde, la direction du vent (pour une altitude donnée) ne semble pas varier durant l'évènement (Figure V.46A, les 3 courbes sont presque superposées). Cela peut sembler assez contradictoire avec les observations du radar de Berrimah (Figure V.44, différence sensible de direction de déplacement des cellules selon la période), et nous incite encore une fois à considérer avec prudence les sorties de modèle qui ont fourni l'essentiel des données d'entrée du produit mergesonde sur la période couverte par les 5 évènements.

Au niveau des températures (Figure V.46C), on note que :

- l'iso-0 est très haut (autour de $4.5km$). Cette valeur est cohérente avec les mesures du KAZR (non représentées ici) qui voient bien la bande brillante vers cette altitude.

- il semble y avoir deux parties dans la colonne atmosphérique, avec en-dessous de $1500m$ un gradient de température plus faible qu'au-dessus.

Examinons à présent les évènements successifs que l'on peut voir sur la Figure V.45.

Le premier de ces évènements commence assez brutalement vers 18:00 avec les taux de pluie les plus forts (environ $10mm/h$, Figure V.45E). Il dure jusqu'à 18:30 environ et est caractérisé par des champs de réflectivités et de vitesses Doppler moyennes très perturbés. Les échos radar observés semblent a priori peu explicables par notre modèle de propagation, dans la mesure où ils proviennent d'abord (vers 18:00) des portes les plus proches du sol pour "remonter" ensuite vers le haut de la colonne (vers 18:15, Figure V.45A). Ce phénomène

pourrait être dû à de la condensation ayant lieu près du sol, ce qui est peu probable vues les cartes $2D$ de la Figure V.44 qui montrent bien le passage d'un front à ce moment. Il s'agit donc plus vraisemblablement de mouvements horizontaux intenses (vent plus fort près du sol qu'en altitude, cisaillements importants dans l'intensité et/ou la direction du vent). On note d'après le produit mergesonde, sur le radiosondage comme sur les sorties de modèle (Figure V.46A, toutes courbes) une inversion de la direction des vents sur la colonne durant toute la période couverte et jusqu'à la date du radiosondage (vent vers le sud près du sol et vers le nord vers $4km$). Ce phénomène, qui n'a jamais été modélisé dans ce travail (dans l'étude sur le vent horizontal nous avons supposé une direction unique de déplacement quelque soit la date et l'altitude) complique également l'étude, dans la mesure où il est susceptible de déplacer les gouttes dans un plan $2D$ (le nuage peut ne jamais passer à la verticale du radar tandis que les gouttes qui en sont issues peuvent y être ramenées par le vent horizontal).

On distingue ensuite une seconde phase entre 18:30 et 19:00 (Figure V.45), avec des taux de pluie plus faibles ($2-7mm/h$) et un champ KAZR de réflectivités plus régulier dans le temps et sur la colonne. Les vitesses Doppler (Figure V.45C) restent néanmoins assez perturbées et peu corrélées aux réflectivités, signe a priori de mouvements verticaux significatifs.

La troisième phase a lieu entre 19:00 et 19:20. On relève des taux de pluie au sol faibles ($< 2mm/h$) mais des réflectivités encore relativement fortes, notamment sur le haut de la colonne (autour de $30dBZ$). Cette phase correspond au début de la période où l'on ne relevait pas clairement d'échos sur le radar à visée horizontale (Figure V.44E, F).

Le quatrième évènement couvre la période 19:25-20:00. On ne relève quasiment pas de précipitations au sol (cohérent avec la Figure V.44F, G). Les échos radars provenant des plus hautes porte (Figure V.45A) sont cependant un peu plus forts ($20dBZ$ environ) et diminuent ensuite sensiblement sur la colonne (autour de $0dBZ$ sur le bas de la colonne). On suspecte ici un effet important de l'évaporation, dont nous présenterons la paramétrisation (section 7.2.3) et les effets (section 7.3).

La cinquième et dernière phase couvre la période 20:00-21:00. Il s'agit de la deuxième période de pluie vue (Figure V.44, G, H et I), avec déplacement de la cellule vers le nord-est. Au sol, elle est caractérisée par des taux de pluie variables, entre 1 et $10mm/h$, avec un pic principal vers 20:30. Le champ de réflectivité est relativement régulier verticalement, malgré un vent horizontal semble-t-il assez fort (Figure V.46B). (Cette remarque doit être considérée avec précaution vu l'éloignement temporel du radiosondage.)

Le Tableau V.6 synthétise ce découpage de l'évènement complet en 5 évènements successifs. Il donne pour chacun sa durée et les taux de pluie et cumuls d'eau enregistrés au

sol par le 2DVD.

TABLE V.6 – Dates et caractéristiques (cumul d'eau $RRcum$ et taux de pluie moyen au sol $RRmean$) des 5 évènements successifs du 04 novembre 2014 à Darwin.

	total	ev 1	ev 2	ev 3	ev 4	ev 5
dates	17:58-21:00	17:58-18:31	18:31-18:56	18:58-19:21	19:25-19:51	19:55-20:53
durée	$3h02$	$33min$	$25min$	$23min$	$26min$	$58min$
$RRcum$ (mm)	5.43	2.26	1.33	0.32	0.037	1.48
$RRmean$ (mm/h)	1.81	4.05	3.19	0.82	0.085	1.53

Assimilation Nous avons appliqué notre algorithme d'assimilation aux 5 évènements du Tableau V.6 (en une fois, le découpage est fait a posteriori lors de l'interprétation). La bande brillante étant très haute ($4.5km$), nous décidons de fixer la limite haute de notre modèle plus bas, à $3km$ d'altitude. Cela permet de limiter les problèmes de diffusion numérique (lissage des DSD modèle durant la chute), qui croît avec la hauteur de chute, et de limiter les problèmes, notamment dûs à la non prise en compte du vent horizontal. Les inconnues de l'algorithme seront donc les DSD à $3km$ et les vents verticaux jusqu'à cette altitude. Nous utilisons toujours un pas vertical de $\Delta z = 100m$ et temporel de $\Delta t = 5s$.

Nous détaillerons dans la section suivante la paramétrisation de l'évaporation dans notre algorithme.

7.2.3 Réglage de l'évaporation

Durant la période considérée (Figure V.46C), on a une température élevée ($26\,^\circ C$ au sol) et une grande hauteur de chute des gouttes. Ce sont deux facteurs augmentant l'impact de l'évaporation. Comme on l'a vu (Figure V.45A), durant l'évènement 4, celle-ci est suspectée de jouer un rôle prépondérant sur les profils verticaux de réflectivités. Nous allons donc devoir paramétrer cette évaporation. Il nous faut pour cela fixer d'abord un profil de températures. Vue la Figure V.46C, nous utilisons une température au sol de $27\,^\circ C$, et un gradient de $-1\,^\circ/360m$ jusqu'à $1.5km$ et de $-1\,^\circ/130m$ au-dessus, ce qui nous donne une température de $11\,^\circ C$ en haut de la colonne sur laquelle on applique notre algorithme (à $3km$), et un isotherme $0\,^\circ$ à $4.45km$.

Intéressons nous à présent aux profils d'humidité relative.

La Figure V.47 présente le produit mergesonde d'humidité relative (h_r) en fonction du temps et de l'altitude (A), le champ que nous utiliserons finalement dans notre algorithme

d'assimilation, VAMOS (B), et les mesures au sol de taux de pluie et d'humidité (C).

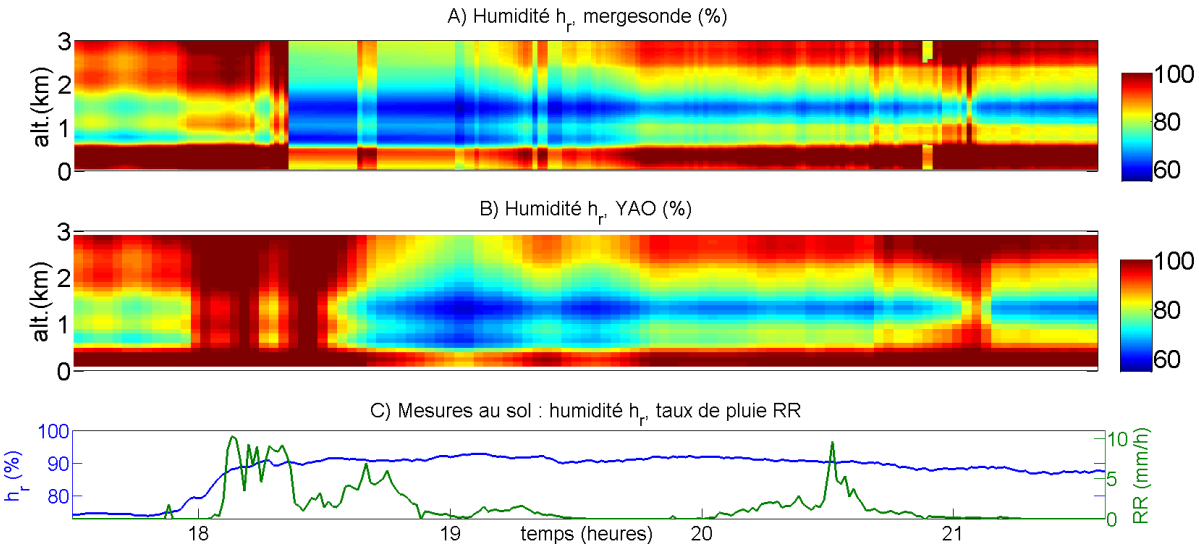

FIGURE V.47 – A) Champ d'humidité relative (h_r) du produit mergesonde. B) Idem corrigé pour utilisation dans notre algorithme d'assimilation (voir texte). C) Taux de pluie et humidités relatives mesurées au sol. Les pointillés délimitent de manière indicative les différents évènements au sens du Tableau V.6.

Le champ d'humidités relatives du produit mergesonde est calculé en deux phases. Dans un premier temps sont fusionnées les sorties du modèle et les radiosondages (comme pour les autres champs du produit). Ensuite, pour chaque profil vertical, on réajuste les humidités relatives (en appliquant un coefficient multiplicateur) pour que la quantité d'eau sur la colonne soit la même que celle mesurée par un radiomètre, lorsque les données de celui-ci sont disponibles. C'est ce champ qui est présenté Figure V.47A. On comprend que cette méthode pose deux soucis :

(1) - lorsqu'une partie de la colonne était déjà quasiment saturée avant correction par les données radiométriques, ou que la correction est très fortement positive (ajout de beaucoup de vapeur d'eau), le fait de multiplier tout le profil correspondant entraîne localement des humidités relatives très supérieures à 100%.

(2) - lorsqu'il n'y a pas de données radiomètre disponibles, les champs ne sont pas corrigés. Il s'en suit des transitions brutales et injustifiées, par exemple vers 18:20.

Pour corriger (rapidement) ces deux problèmes, on a dans ce travail :

(1) - Repris la correction par les données radiométriques, en distribuant l'eau ajoutée de manière homogène sur la colonne, mais uniquement dans les zones non saturées.

(2) - Lorsqu'une période n'est pas couverte par le radiomètre, supprimé totalement les données correspondantes, remplacées par une interpolation basée sur les valeurs corrigées les plus proches.

On obtient ainsi le champ de la Figure V.47B.

On remarque que ce champ reste néanmoins suspect. Au niveau de la variabilité temporelle, le premier évènement (au sens du Tableau V.6, vers 18:00) correspond à une atmosphère presque totalement saturée. Ensuite (à partir de 18:30), on a des taux d'humidité sensiblement plus faibles (jusqu'à 60%) malgré la pluie persistante. Ce phénomène d'assèchement est possible mais serait ici d'une ampleur surprenante. Au niveau de la variabilité verticale, le fait d'avoir des humidités plus fortes au-dessus de $2km$ peut s'expliquer par la présence du nuage. Le fait que l'atmosphère soit quasiment saturée sur les 500 premiers mètres et bascule subitement ensuite vers des taux d'humidité beaucoup plus faibles est davantage suspect. Cela l'est d'autant plus que les mesures directes d'humidité au sol (Figure V.47C), si elles montrent bien une augmentation forte en début d'évènement (vers 18:00) donnent ensuite des humidités relativement constantes (autour de 90%) et souvent plus basses que sur le champ de la Figure V.47B. Le fait que ces champs (en particulier leur variabilité verticale) soient principalement déterminés par des sorties de modèle diminue la confiance que l'on peut leur accorder (résolutions différentes).

Par conséquent, dans la suite, nous étudierons les résultats d'expériences d'assimilation réalisées d'une part avec le champ d'humidité de la Figure V.47B et d'autre part sans aucune évaporation. On note que dans tous les cas, il pourrait être intéressant soit de faire d'autres expériences d'assimilation avec d'autres profils d'humidité (et de tester la capacité de chacun à expliquer les observations), soit de contrôler directement ces profils d'humidité dans l'algorithme d'assimilation, comme on l'avait suggéré dans la section 6. (Mais cela représente un travail conséquent qui ne pouvait être fait dans les délais de cette thèse. Il s'agit néanmoins d'une perspective intéressante de ce travail).

L'évaporation étant néanmoins supposée jouer un rôle important ici, par défaut (sans indication contraire), les résultats présentés dans la suite seront ceux obtenus avec le champ d'humidité relative de la Figure V.47B.

7.3 Résultats de l'assimilation (VAMOS)

Nous avons vu que nous traitions ici une période de pluie longue et complexe, divisée en 5 évènements successifs. Afin de rester raisonnablement concis, nous ne fournirons pas l'analyse détaillée de tous les phénomènes régissant ces 5 évènements. Nous procéderons comme suit. Dans la section 7.3.1, nous étudierons de manière assez générale les résultats de notre algorithme sur ces évènements (capacité à reproduire les observations et les structures restituées pour les champs de vents verticaux). Nous verrons aussi l'impact de l'évaporation sur les résultats. On rappelle que par défaut (c'est-à-dire partout en dehors du paragraphe dédié à ce sujet), les résultats présentés sont produits avec un module d'évaporation paramétré par le champ d'humidité de la Figure V.47B. Ensuite, dans la section 7.3.2, nous nous focaliserons uniquement sur l'évènement 3 (18:58 - 19:21, cf. Tableau V.6) que nous analyserons en détails. Enfin, dans la section 7.3.3 enfin, nous reviendrons sur la gestion de la turbulence par notre algorithme. (Ce problème a déjà été traité sur données simulées en section 3.3.6.)

7.3.1 Résultats généraux / Impact de l'évaporation

Analyse générale Ici, nous allons donc analyser les performances générales de notre algorithme VAMOS pour les évènements considérés. Pour cela, nous nous reporterons d'abord à la Figure V.45, déjà mentionnée, qui présente les deux premiers moments centrés des spectres Doppler (réflectivité Z_e et vitesse Doppler moyenne W) observés (par le KAZR) et assimilés (VAMOS), et les taux de pluie au sol, observés (par les deux disdromètres) d'une part et assimilés d'autre part. Le Tableau V.7 fournit les performances statistiques de notre algorithme pour toute la période (colonne 3), toute la période sauf l'évènement 1 (colonne 4), ou évènement par évènement (colonnes 5 à 8). Comme le Tableau V.4 sur l'évènement finlandais, il donne des statistiques d'erreurs sur les trois premiers moments centrés des spectres Doppler et sur les mesures disdromètres (mais ici sans dépendance verticale, voir légende).

TABLE V.7 – Comparaison entre états observés et assimilés pour les trois premiers moments centrés des spectres Doppler et le nombre de gouttes disdromètre (flux de gouttes) pour le cas à Darwin du 05/11/2014. On présente les résultats pour les 5 évènements de la journée (au sens du Tableau V.6), colonne 3, puis seulement pour les 4 derniers (colonne 4) et finalement pour les évènements 2, 3, 4 et 5 individuellement (colonnes 5 à 8). Les indicateurs présentés sont la MAE, la MAPE et le biais relatif (cf. section 3 Eq.(V.20), Eq.(V.18) et Eq.(V.19) respectivement). On rappelle qu'un biais positif correspond à une surestimation par VAMOS (par l'assimilation).

		evs 1 à 5	evs 2 à 5	ev 2	ev 3	ev 4	ev 5
η	MAE (dB)	2.7	1.6	1.4	1.4	2.1	1.5
	bias (dB)	+0.71	+0.0	−0.14	−0.51	+0.38	+0.24
W	MAPE (%)	7.0	3.7	2.4	2.3	9.6	2.1
	rbias (%)	+1.1	+0.29	+0.53	−0.084	+2.3	−0.30
	moyenne(obs) (m/s)	6.2	6.1	6.8	6.7	3.8	6.6
LS	MAPE (%)	-	36	23	24	89	21
	rbias (%)	−0.73	+20	+10	+14	+67	+11
	moyenne(obs) (m/s)	1.4	1.1	1.2	1.2	0.75	1.1
disdro	MAE (flux de gouttes : mm^{-1})	34	29	50	28	12	28
	mean(obs) (flux de gouttes : mm^{-1})	163	124	303	85	32	113

Tout d'abord, comme on l'avait anticipé, l'évènement 1 (avant 18:30) est particulièrement mal restitué par l'algorithme d'assimilation (Figure V.45). En effet, on avait vu que les échos radars de cet évènement provenaient d'abord des portes près du sol, puis remontaient ensuite. Ces échos sont impossibles à restituer par notre modèle, contraint de faire tomber les gouttes du haut vers le bas de la colonne. Le fait que le radar, durant $15 - 20min$ (avant 18:15), voit des gouttes en bas de la colonne et pas en haut, rend toutes les restitutions absurdes sur cette période (y compris les restitutions de vitesses Doppler moyennes, et donc de vent vertical).

C'est ce qui explique que les statistiques globales présentées dans le Tableau V.7 sont évaluées avec et sans ce premier évènement (colonnes 3 et 4). On voit d'ailleurs que l'on obtient des erreurs largement plus élevées sur les réflectivités (η) et les vitesses Doppler moyennes (W) si l'on prend en compte l'évènement 1. Sans lui (colonne 4), on obtient pour le reste de la période pluvieuse des résultats un peu moins bons que pour les évènements HyMeX et Finlande (typiquement, MAE sur la réflectivité de $1.6dB$ contre respectivement 0.65 et $1.2dB$ dans les cas précédents et largeurs spectrales largement surestimées, ici de 20% en moyenne). Néanmoins, comme en témoigne par ailleurs la Figure V.45, ces résultats restent cohérents (les variabilités temporelle et verticale des réflectivités et vitesse Doppler et les taux de pluie au sol sont bien reproduits pour la plupart des périodes). La complexité de ces évènements explique ces résultats un peu moins bons. La hauteur de chute plus importante induit une augmentation de la diffusion numérique. Cela implique donc une surestimation plus forte qu'habituellement de la largeur spectrale. Nous verrons cependant que cette moyenne cache des variations locales selon la turbulence (voir section 7.3.3).

Afin de pousser un peu plus loin l'analyse, nous présentons la Figure V.48, qui représente des champs temps/altitude de restitutions de notre algorithme : le vent vertical (A), le diamètre volumique moyen D_m (B), le facteur de réflectivité radar Z (moment d'ordre 6 de la DSD), le contenu en eau liquide LWC et le nombre total de gouttes N_{tot}. La Figure V.49 présente elle les vitesses Doppler moyennes (W) enregistrées par le radar BASTA à $94GHz$ (A), les vitesses Doppler moyennes correspondantes simulées à partir de nos restitutions VAMOS de DSD et vents verticaux (B) et le vent vertical VAMOS présenté cette fois-ci à la résolution de notre modèle, $5s$ (alors qu'il est par défaut présenté à la résolution des observations : $30s$).

Sur le vent vertical à $30s$ d'abord (Figure V.48A), on constate à nouveau que les restitutions de VAMOS sur l'évènement 1 (avant 18:30) sont très suspectes. Elles correspondent à une zone où l'algorithme est inapte à expliquer la structure des champs de réflectivité. Les vents qu'il produit sont fortement corrélés aux structures observées de réflectivités, ce qui signifie que notre algorithme d'assimilation utilise le vent vertical pour expliquer ce que le modèle l'empêche d'expliquer avec les DSD. Dans la suite, on se concentrera uniquement sur les évènements 2 à 5 (après 18:30).

Pour tenter d'évaluer la pertinence de nos restitutions de vents verticaux, nous utilisons le radar BASTA. Indépendamment des différences de fréquences des mesures, on constate que la variabilité des vitesses Doppler BASTA (Figure V.49A) est très proche de celle des vitesses Doppler KAZR (Figure V.45C, attention, la gamme de temps, en abscisse, est légèrement différente). Cela nous laisse penser que cette variabilité est principalement due à la variabilité du vent vertical, puisque, contrairement aux DSD, il affecte de la même manière les W aux deux fréquences. Le fait que cette variabilité soit particulièrement bien reproduite par notre

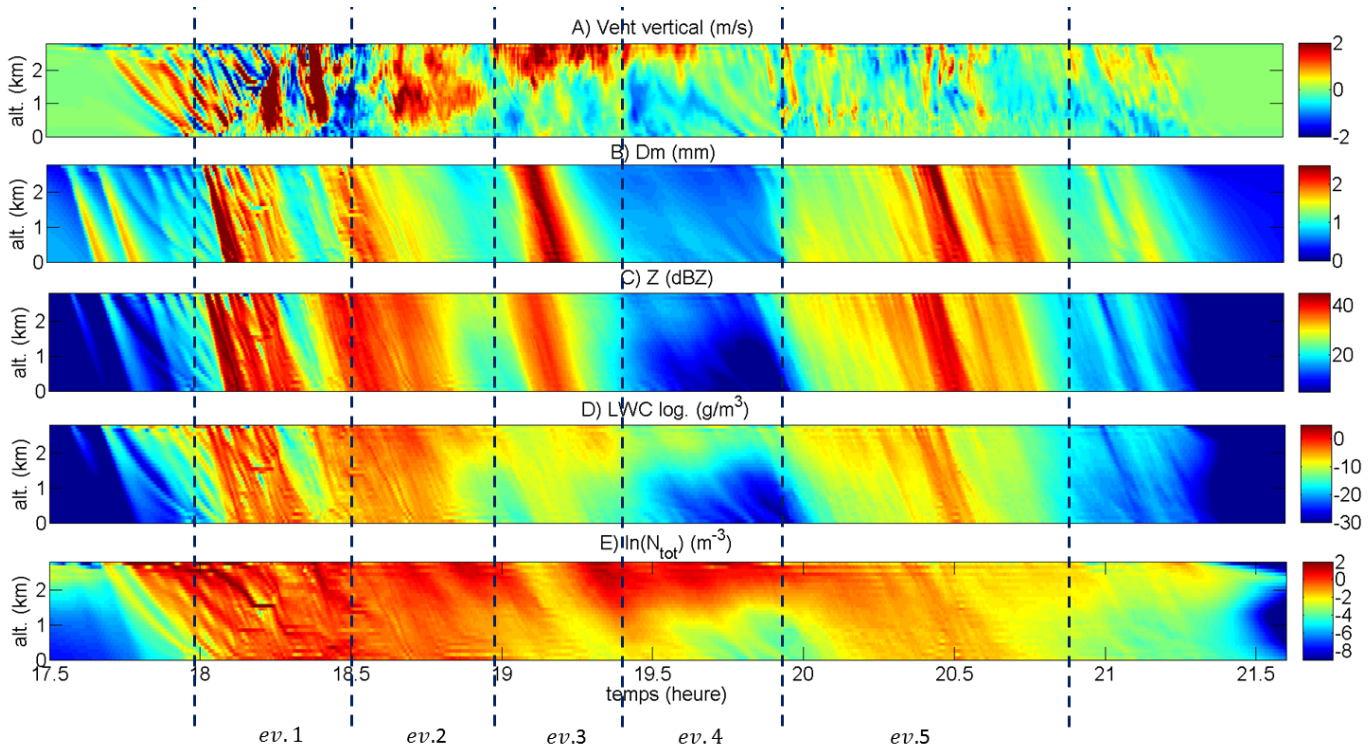

FIGURE V.48 – Moments de la DSD et vents verticaux restitués par l'algorithme d'assimilation. A) Vent vertical, positif vers le bas. B) Diamètre volumique moyen D_m. C) Facteur de réflectivité radar Z. D) Contenu en eau liquide LWC. E) Nombre de gouttes N_{tot}. Les pointillés délimitent de manière indicative les différents évènements au sens du Tableau V.6.

algorithme (Figure V.49B et V.45D) nous laisse penser que nous restituons des vents verticaux valides. Le fait que ces vents verticaux (Figure V.48A) soient également généralement plus forts (en valeur absolue) que dans les deux cas précédents (HyMeX et Finlande) rend également moins prégnants (relativement) les biais constatés sur l'évènement en Finlande (de l'ordre de $-0.2m/s$).

La structure du champ des vents verticaux assimilés (Figure V.48A) est assez variable dans le temps. Sur l'évènement 2 (18:30-19:00) on obtient des vents dirigés vers le bas (en rouge) assez significatifs ($1 - 2m/s$) sur une bonne partie de la colonne. Dans la période suivante (19:00-20:00 , évènements 3 et 4), cette zone de vents descendants se retrouve uniquement sur le haut de la colonne. Là où on a plutôt des vents montants (négatifs) en bas de la colonne. Durant l'évènement 5 (20:00-21:00), le champ est plus perturbé avec des "cellules" de vent plus localisées en temps/altitude. Pour les vitesses Doppler moyennes KAZR (W, Figure V.45C) et surtout BASTA (Figure V.49A), on constate vers $0.5 - 1km$ d'altitude une zone très perturbée entre 19:30 et 20:30 (cadre blanc de la Figure V.49A). Cela est reproduit de manière atténuée sur les sorties correspondant de notre algorithme (Figure V.45D et V.49B) mais est plus visible entre 20:00 et 20:30 sur le champ de vent assimilé à $5s$ (Figure V.49C). Signe de la turbulence atmosphérique, nous reviendrons sur ce phénomène en conclusion de la section 7.3.3.

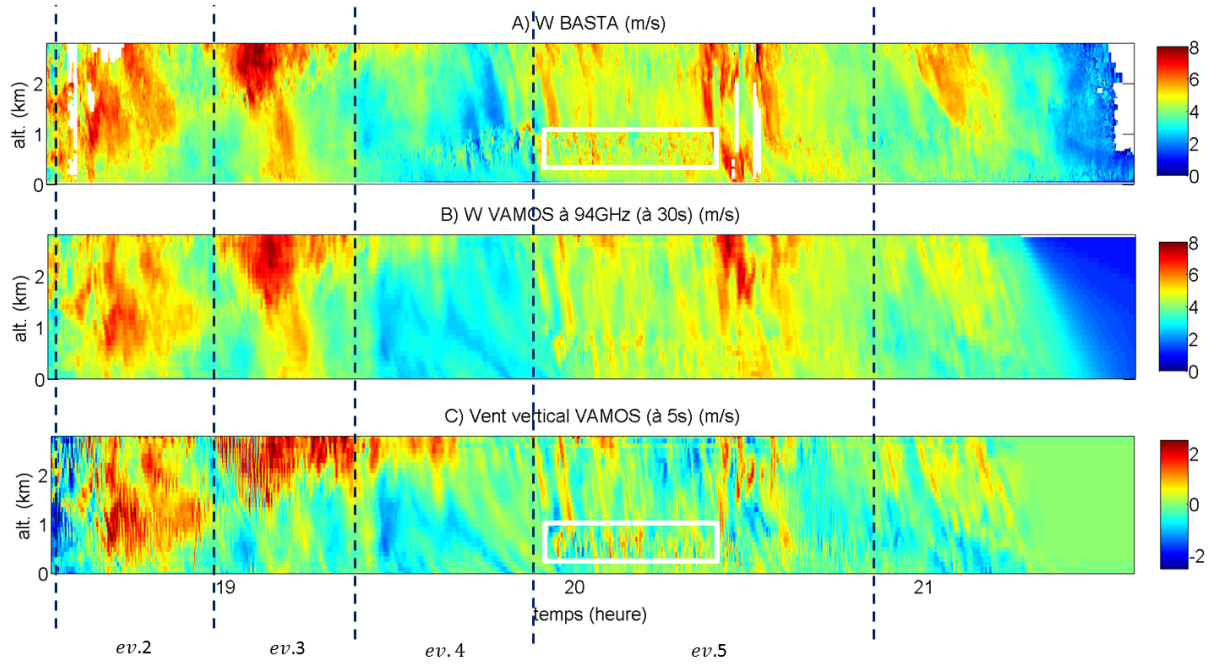

FIGURE V.49 – A) Vitesse Doppler moyenne W enregistrée par le radar BASTA. B) Vitesse Doppler moyenne W à $94GHz$ simulée à partir de nos restitutions de DSD et vents verticaux. C) Vent vertical restitué à résolution $5s$, positif vers le bas. Les pointillés délimitent de manière indicative les différents évènements au sens du Tableau V.6. Les cadres blancs délimitent la zone de turbulence évoquée plus particulièrement en section 7.3.3.

Donnons pour terminer quelques considérations évènement par évènement.

L'évènement 2 (18:30-19:00) présente des facteurs de réflectivités radar Z et contenus en eau LWC supérieurs à ceux de l'évènement 3 (19:00-19:30). Les D_m très supérieurs (Figure V.48) sur ce dernier laissent penser que les petites gouttes sont moins nombreuses durant le cœur de l'évènement 3, que nous étudions plus en détails en section 7.3.2 (faibles réflectivité mais forts D_m suggèrent un évènement composé de grosses gouttes).

On peut aussi noter l'interdépendance entre vent vertical et nombre de gouttes (et dans une moindre mesure contenu en eau liquide). Par exemple en haut de la colonne (au-dessus de $2km$, vers 19:20), on voit un patch important de gouttes (augmentation de N_{tot}, Figure V.48E). Ce patch se traduit par une augmentation du contenu en eau liquide (Figure V.48D) mais n'engendre pas vraiment de variations des diamètres moyens et facteurs de réflectivité radar. Il est vraisemblablement provoqué par les vents verticaux. Ceux-ci vers cette date sont fortement positifs au-dessus de $2km$ (vents descendants de l'ordre de $2m/s$) et négatifs (ascendants) en-dessous ($-0.5/-1m/s$). Les petites gouttes, chassées des plus hautes altitudes par le vent descendant, sont ensuite "bloquées" vers $2-2.2km$ par les vents ascendant. On note cependant que le nombre de gouttes (N_{tot}) et dans une moindre mesure le contenu en eau (LWC) sont aussi fortement impactés par l'évaporation sur la colonne, compliquant ces analyses.

On constate enfin sur le Tableau V.7 que l'évènement 4 (19:30-20:00) présente des erreurs plus fortes en terme de vitesse Doppler moyenne W (légèrement surestimée, de 2.3%) et surtout

de largeur spectrale LS (surestimée de 67%). L'erreur au sol est aussi assez significative vu le peu de gouttes présentes (erreur de 12 gouttes/m^3 pour une moyenne de 32 gouttes/m^3). On note cependant que cet évènement est celui où les taux de pluie au sol sont quasiment nuls et où l'on suspectait un fort impact de l'évaporation. Vue la Figure V.45 (A et B), on constate qu'il semble que l'on a bien reproduit la tendance générale à la perte de réflectivité durant la chute. La légère surestimation de la vitesse Doppler moyenne s'explique peut-être par une évaporation encore insuffisante (dont l'impact est discuté dans le paragraphe suivant). On note (Figure V.48D) que pour cet évènement, quasiment toute l'eau est perdue sur la colonne, ce qui rend les résultats plus anecdotiques.

Impact de l'évaporation Nous allons à présent commenter l'influence de l'évaporation sur nos restitutions. On a vu que celle-ci dépendait du champ d'humidités relatives et que les informations dont nous disposions à ce sujet (Figure V.47) étaient suspectes. Nous avons donc effectué deux expériences d'assimilation. Une avec évaporation, utilisant pour l'humidité le champ de la Figure V.47B ; et une sans évaporation (humidité à 100%).

La Figure V.50 montre les profils verticaux moyens de trois paramètres du problème (facteur de réflectivité radar équivalent Z_e, diamètre volumique moyen D_m et contenu en eau liquide LWC) pour trois évènements (2, 4, 5, voir Tableau V.6), avec (bleu) ou sans (vert) évaporation, ainsi que les profils d'humidité pour le cas avec évaporation.

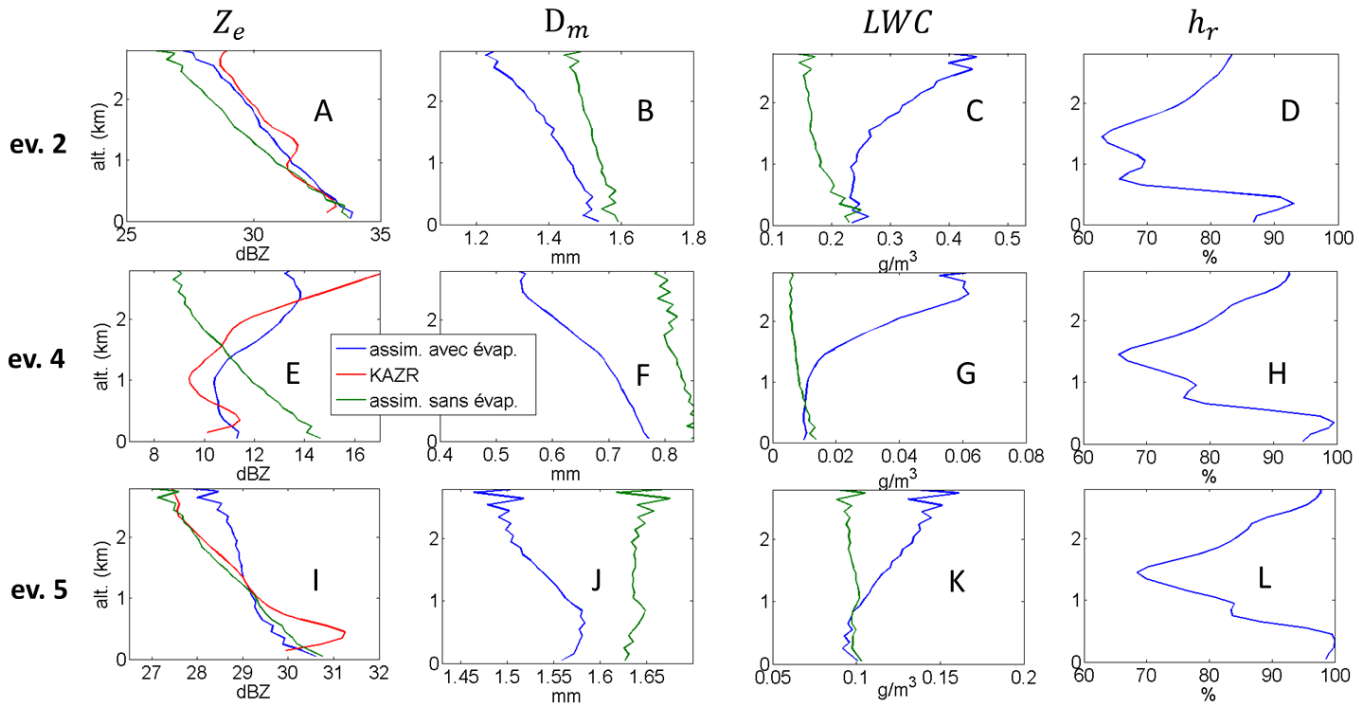

FIGURE V.50 – Pour trois évènements, profils verticaux moyens de 3 paramètres du problème (facteur de réflectivité radar équivalent Z_e, diamètre volumique moyen D_m et contenu en eau liquide LWC), avec (bleu) ou sans (vert) évaporation. Pour Z_e, on présente aussi les observations KAZR (rouge). Profils d'humidités relatives moyennes pour le cas avec évaporation.

Intéressons nous tout d'abord au facteur de réflectivité radar (A, E, I). Deux phénomènes modélisés dans notre algorithme sont susceptibles de le faire varier sur la colonne. Le premier est l'atténuation du signal radar, qui fait décroître la réflectivité à mesure que l'on s'élève dans la colonne, le second est l'évaporation, qui fait décroître la réflectivité au contraire lorsqu'on s'approche du sol. Il s'agit donc d'une "compétition" entre les deux phénomènes. Pour les évènements 2 et 5, on voit sur les mesures du KAZR que la réflectivité diminue lorsqu'on monte en altitude (Figure V.50A et I, rouge). L'effet de l'atténuation prévaut donc sur celui de l'évaporation. On obtient l'effet inverse pour l'évènement 4, en particulier sur le haut de la colonne atmosphérique (E). Mais le fait que l'atténuation prévale ne signifie pas qu'il n'y a pas d'évaporation. Ainsi, on remarque que pour l'évènement 2, notre algorithme avec évaporation semble mieux reproduire les observations que celui sans évaporation (Figure V.50A, bleu et vert). Sans évaporation (avec donc juste l'atténuation), la baisse de réflectivité en montant dans la colonne est trop forte.

Sur l'évènement 4 où l'évaporation prédomine sur l'atténuation (Figure V.50E), les résultats sans évaporation présentent naturellement un profil inversé par rapport aux mesures. Avec évaporation, la variation se fait dans la bonne direction, mais la structure verticale ne correspond en revanche que moyennement aux mesures. On voit que sur notre profil d'humidité (Figure V.50H), l'atmosphère est presque saturée sur le bas et le haut de la colonne et que l'évaporation ne peut qu'avoir lieu sur la partie médiane ($0.5-2km$). C'est bien ce que l'on voit sur le profil de réflectivité (Figure V.50E, bleu) : lorsqu'on descend la colonne, la réflectivité croît légèrement de 2.5 à $2km$, décroît sur la partie médiane (2 à $1km$) et croît à nouveau près du sol. Le profil enregistré est différent.

Sur l'évènement 5 (Figure V.50I), on explique mieux le profil observé sans évaporation qu'avec. VAMOS n'explique cependant pas le pic de réflectivité (à $31dBZ$) observé par le KAZR en bas de la colonne (vers $0.4km$). Celui-ci est probablement lié au maximum de réflectivité observé vers 20:30 à cette altitude (voir Figure V.45A), et qui correspond peut-être à un effet du vent horizontal, dont les effets ne sont pas pris en compte dans notre algorithme.

On a donc trois situations différentes sur ces trois évènements.

Voyons à présent comment ces résultats sur les spectres se transcrivent sur la DSD.

On note tout d'abord que lorsqu'on n'inclut pas d'évaporation dans notre algorithme, les quantités d'eau sont conservées sur la colonne en moyenne (des vents verticaux négatifs, par exemple, peuvent retenir un temps en l'air des petites gouttes). Sur les restitutions de la Figure V.50 (vert), les diamètres moyens D_m (B, F, J) et contenu en eau LWC (C, G, K) varient donc peu en moyenne sur les différents évènements. Les légères variations néanmoins constatées viennent probablement des débuts et fin d'évènements (les limites d'évènements sont définies à date fixe et ne suivent pas la chute des gouttes : par exemple, la fin de l'évènement 1 - plus fort que l'évènement 2 - peut affecter le bas de la colonne au début de l'évènement 2 et augmenter légèrement les contenus en eau à ce niveau).

En comparant les profils de contenu en eau avec et sans évaporation (Figure V.50C, G, K), on note :

- qu'ils se correspondent au niveau du sol, grâce au forçage par les données disdromètres.

- que les différences en haut de la colonne sont très significatives. Sur l'évènement 4 par exemple (évènement très faible au sol et où l'on suspectait dès le départ une forte influence de l'évaporation), la quantité d'eau est divisée par 6 sur la colonne lorsqu'on prend en compte l'évaporation, passant de $0.06g/m^3$ à $3km$ à $0.01g/m^3$ au sol. Cet évènement, faible, est le plus affecté relativement par l'évaporation, mais on perd encore la moitié de l'eau pour l'évènement 2 (Figure V.50C), où les contenus en eau sont pourtant 10 fois plus forts.

Concernant D_m (Figure V.50B, F, J), on remarque :

- que l'évaporation augmente le diamètre moyen D_m lors de la chute (on détruit davantage de petites gouttes). En haut de la colonne, les D_m calculés avec évaporation sont donc largement plus petits que sans.

- Lorsqu'on prend en compte l'évaporation, les D_m sont également un peu plus élevés en bas de la colonne, mais la différence n'est pas très significative (pas plus de $0.1mm$).

On remarque enfin que dans tous les cas, la structure verticale du champ d'humidité impacte directement D_m et LWC, avec des variations de ces paramètres d'autant plus faibles que l'humidité est forte.

7.3.2 Analyse détaillée de l'évènement 3

Nous allons à présent analyser en détail l'un des évènements du 04 novembre 2014, l'évènement 3 (cf. Tableau V.6), qui couvre la courte période 18:58-19:21. On note qu'il s'agit d'un évènement pour lequel VAMOS fournit des résultats mitigés en terme de nombre de gouttes (flux) au sol (Tableau V.7, colonne 6) et où les profils verticaux de réflectivité observés ne semblent pas parfaitement bien reproduits (Figure V.45, A et B). Cette étude permettra d'expliquer ces résultats. Nous analyserons dans un premier temps les résultats au sol, où nous comparerons les mesures disponibles (disdromètres, spectres KAZR sur la première porte disponible) entre elles et par rapport aux sorties de notre algorithme. Puis nous étendrons l'étude à toute la colonne pour évaluer l'impact de la dynamique de la pluie durant la chute sur ces données sol.

Etude au sol Intéressons nous tout d'abord aux paramètres intégrés au sol ou près du sol, le taux de pluie RR et la réflectivité Z_e. La Figure V.51 présente des séries temporelles de ces paramètres (voir légendes).

Sur les taux de pluie (A), on remarque que :

- il semble y avoir deux phases dans cet évènement : une couvrant les dix premières minutes (18:58-19:07) avec des taux de pluie faibles et réguliers (environ $0.5mm/h$) et une entre 19:07 et 19:18 avec des taux de pluie un peu plus forts ($1 - 2mm/h$).

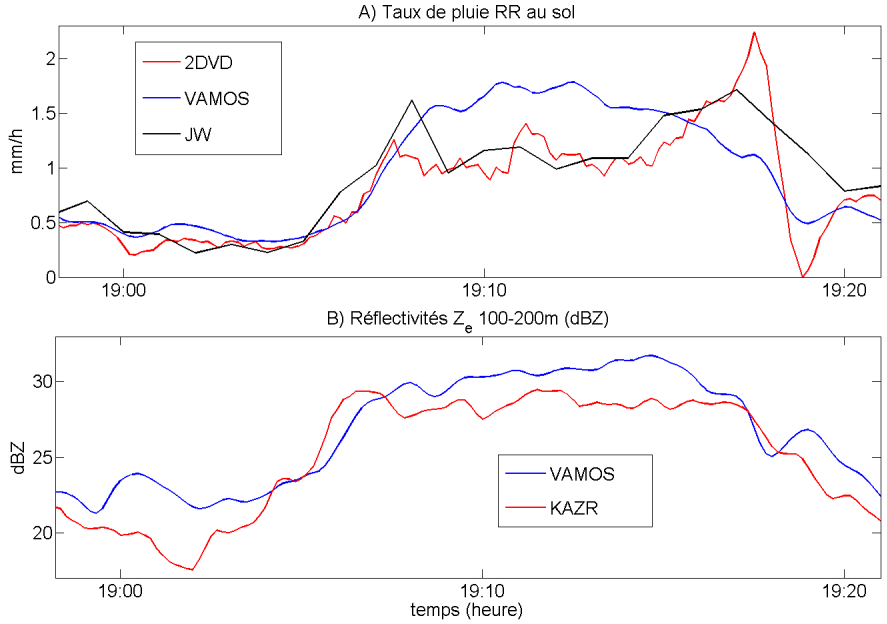

FIGURE V.51 – A) Taux de pluie au sol (mm/h) durant l'évènement 3 enregistrés par les disdromètres 2DVD (rouge) et JW (noir) et restitués par l'algorithme d'assimilation (bleu). B) Réflectivités (dBZ) enregistrées par le KAZR (rouge) et restituées par l'algorithme d'assimilation (bleu) pour la porte $100 - 200m$.

- les mesures des taux de pluie par les deux instruments disponibles (rouge : 2DVD, noir : JW) présentent des différences sensibles (premier pic vers 19:08 plus important par JW et second pic vers 19:17 moins important), là où VAMOS (bleu) a plus de mal à restituer la dynamique.

- durant la deuxième phase de l'évènement (entre 19:08 et 19:16), l'assimilation surestime la pluviométrie.

- le second pic de pluviométrie de 19:17 (du 2DVD, fourni à l'algorithme), n'est absolument pas reproduit par l'assimilation.

On regarde à présent les mesures de réflectivités du KAZR (rouge) sur sa première porte ($100 - 200m$) comparées aux sorties d'assimilation (bleu) correspondantes (Figure V.51B). On note que :

- les deux phases de l'évènement sont également bien visibles par le radar à cette altitude.

- l'assimilation surestime également la réflectivité sur la deuxième phase, comme il surestimait la pluviométrie au sol. Ceci nous laisse penser que cette erreur (comme dans le cadre d'HyMeX dans la section 5) n'est pas due à une incohérence entre les mesures des deux instruments (on peut alors suspecter une incapacité du modèle de propagation à reproduire la variabilité de la réflectivité sur la colonne atmosphérique, nous y reviendrons).

- le pic important de pluviométrie de 19:17 n'apparaît pas sur la mesure de réflectivité du KAZR. Il semble donc que sur ce point, l'algorithme d'assimilation ait choisi de "suivre" le radar plutôt que le disdromètre, probablement car sur un évènement aussi faible, le nombre

de gouttes disdromètre reste malgré tout limité (Tableau V.7, colonne 6) : le disdromètre a dans la fonction de coût un poids assez faible.

Afin d'approfondir cette étude, nous allons à présent regarder non plus des paramètres intégrés (taux de pluie, réflectivité) mais les mesures du nombre de gouttes disdromètre et les spectres Doppler. La Figure V.52 présente des comparaisons de ces paramètres entre observations (gauche) et sorties d'assimilation (droite). En A, on présente un champ temps/diamètre des mesures brutes du disdromètre (flux de gouttes, tel qu'inséré dans la fonction de coût du 4D-VAR : il s'agit du nombre de gouttes enregistrées par l'instrument). En C, on présente les DSD (volumiques) correspondantes (logarithmiques). En E, on présente les spectres hydrologiques déduits de ces DSD. En G, on présente les spectres Doppler mesurés par le KAZR sur la première porte disponible ($100 - 200m$). Là encore, il s'agit de données d'entrée de l'algorithme d'assimilation.

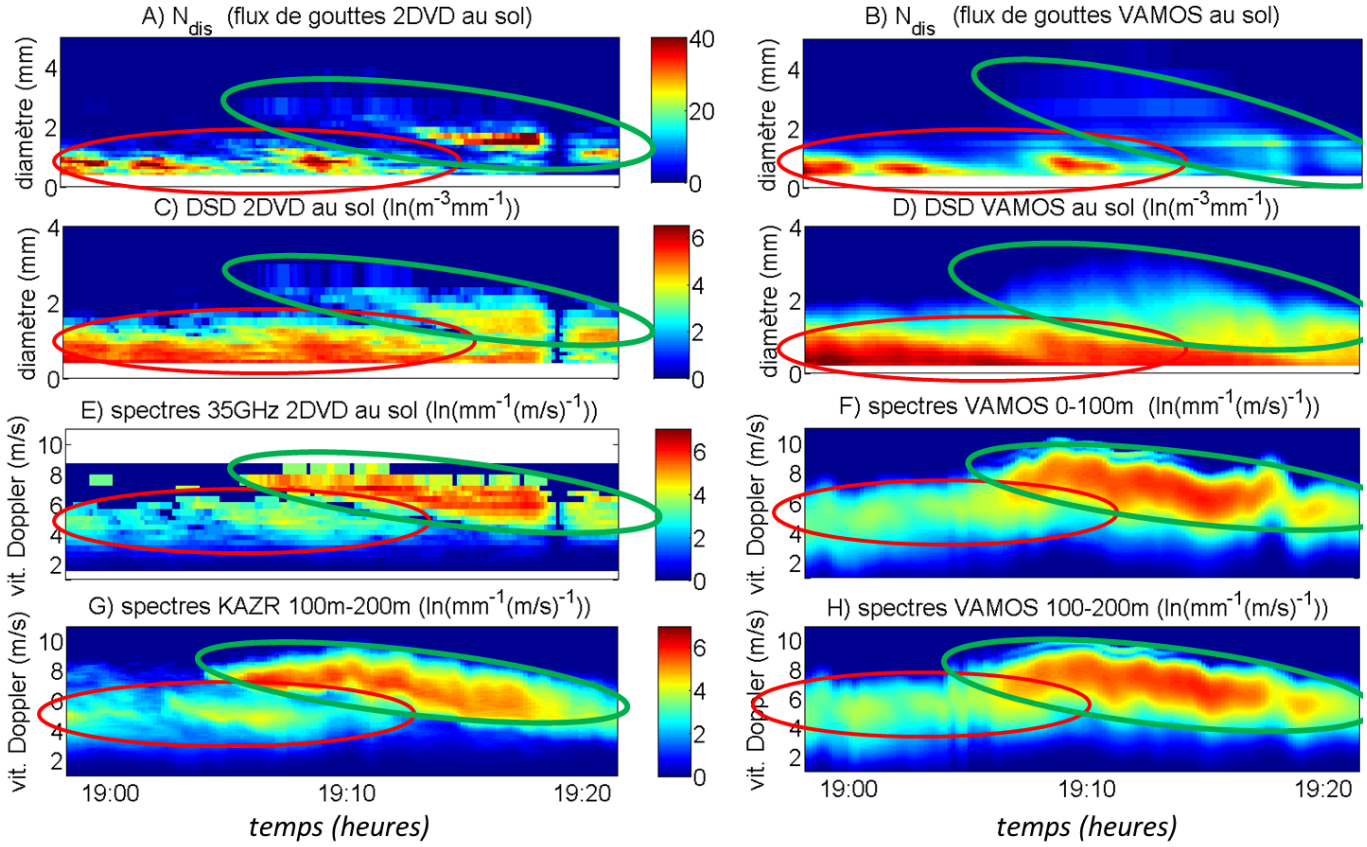

FIGURE V.52 – A) Nombre de gouttes (flux) enregistrées par le disdromètre 2DVD, selon la date et le diamètre. B) Idem restitution algorithme d'assimilation. C) DSD (volumique) correspondant à A. D) Idem restitution algorithme d'assimilation (VAMOS). E) Spectres Doppler hydrologiques à $35GHz$ correspondant à C. F) Idem restitution algorithme d'assimilation (VAMOS). G) Spectres Doppler enregistrés par le KAZR pour la porte $100 - 200m$. H) Idem restitution algorithme d'assimilation (VAMOS). Un spectre Doppler ou une DSD particulière correspond sur chaque image à une "ligne" verticale (date fixée). Unités : pour les DSD : $ln\left[m^{-3}mm^{-1}\right]$. Pour les spectres : $ln\left[mm^{-1}\left(m/s\right)^{-1}\right]$.

On peut noter plusieurs points à partir de cette Figure V.52 :

- Intéressons nous d'abord aux mesures disdromètres brutes (A). On note qu'il semble toujours y avoir deux phases dans l'évènement. La première (ovale rouge, nommée ensuite phase "rouge") est composée presque exclusivement de petites gouttes ($< 1mm$). Ces gouttes sont nombreuses mais correspondent à un faible volume d'eau (LWC) et donc à peu de pluviométrie (d'où les faibles taux précipitant de la Figure V.51). Durant cette phase, on voit par ailleurs clairement trois paquets avec beaucoup de gouttes (qui, on le répète, n'engendrent pas de forts taux de pluie). Cette phase est bien reproduite par notre algorithme d'assimilation (B), qui montre clairement les trois patchs. La deuxième phase (ovale vert, phase "verte") commence avant même la fin de la première avec d'abord des grosses gouttes ($2 - 3mm$) dont la taille décroît petit à petit (jusqu'à $1mm$ à la fin de l'évènement). On peut penser qu'il s'agit en quelque sorte de deux évènements distincts : les petites gouttes de la phase rouge tombant a priori plus lentement que les grosses du début de la phase verte, elles viennent d'endroits d'autant plus distants (elles ont été lâchées bien avant même si elles arrivent presque en même temps au sol). (Nous y reviendrons dans la suite, lors de l'examen du problème sur la colonne entière.) L'assimilation (B) reproduit également cette deuxième phase, mais de manière moins précise : la différence entre les deux phases est moins nette et surtout on constate bien que le pic vu sur les observations vers 19:17 (gouttes de $2mm$ environ) est mal restitué (ce qui fait que le pic de pluviométrie du 2DVD, cf. Figure V.51, est également mal restitué).

- Lorsqu'on passe en DSD (Figure V.52C), on note que le logarithme bien sûr lisse le champ. Le pic de 19:17 est nettement moins visible et l'erreur de l'assimilation sur ce point (D) paraît moins critique.

- Lorsqu'on transcrit ces DSD en spectres Doppler hydrologiques (E), on renverse les poids relatifs accordés aux phases rouges et vertes. La section efficace de rétrodiffusion étant très largement plus grande pour les grosses gouttes, on trouve les réflectivités spectrales les plus fortes pour la phase verte, quand bien même elle est composés de moins de gouttes. Les trois paquets de gouttes de la phase rouge ne sont plus aussi visibles, et le pic de 19:17, qui était composé de gouttes plus petites que celles le précédent dans la phase verte, non plus. L'assimilation reproduit assez bien ces observations, même si les deux phases de l'évènement semblent moins clairement distinctes que sur les observations (à cause probablement de la diffusion numérique).

- Si l'on regarde à présent les mesures KAZR près du sol (Figure V.52G), on constate qu'elles sont globalement en très bonne adéquation avec les spectres issus du disdromètre 2DVD. La différence principale se trouve sur la fin de la phase verte où le radar sous-estime un peu la réflectivité par rapport au disdromètre. (Il faut garder à l'esprit le fait que les deux capteurs ne sondent pas exactement les mêmes volumes).

Un des buts de l'analyse de cette Figure était de montrer la complémentarité entre les mesures radar et disdromètre. Les deux instruments ne se focalisent pas sur la même partie des DSD : le disdromètre va accorder plus d'importance à la gamme de diamètres où l'on

trouve le plus de gouttes (i.e. pour les plus petits diamètres, typiquement en-dessous de $1mm$) tandis que le radar va se focaliser sur les zones à énergies de Mie les plus fortes, et donc à cette fréquence et pour un tel évènement calme plutôt sur des gouttes vers $1-3mm$. Les mesures des deux instruments, apparemment contradictoires, peuvent donc être expliquées par une même distribution. Dans notre cas, cependant, on constate bien une différence à la fin de l'évènement vers le pic de 19:17. Cette différence peut être due à un phénomène local mesuré par le disdromètre (localisé soit au niveau du sol, le radar ne mesurant qu'au-dessus de $100m$, ou localisé dans le plan, le disdromètre couvrant une surface très petite), ou à une hétérogénéité spatiale de la pluie.

Sur cet évènement, les principaux enseignements de cette étude au sol sont :

- le passage successif de deux phases de pluie, l'une (rouge) composée de très petites gouttes exclusivement, l'autre (verte) de plus grosses gouttes dans un premier temps et de plus petites ensuite. Ces deux phases se mélangent en bas de la colonne et font apparaitre des spectres et des DSD bimodales, comme l'avait suggéré par exemple Sauvageot and Koffi (2000).
- la capacité de notre algorithme à restituer ces 2 phases et leurs caractéristiques principales.
- la légère surestimation des observations par notre algorithme durant le début de la deuxième phase (verte).

Nous allons à présent examiner les profils verticaux de précipitations pour étudier la dynamique de ces phénomènes.

Et en altitude ? Intéressons nous d'abord aux deux premiers moments centrés des spectres Doppler et à la capacité de notre algorithme à les restituer. La Figure V.53 présente ces deux moments centrés observés par le KAZR (A : Z_e, C : W) et restitués par notre algorithme : (B : Z_e, D : W). La Figure V.53E présente le vent vertical restitué par VAMOS.

On note que la deuxième phase (verte) définie à partir des données sol (19:10-19:20) correspond bien, que ce soit sur les observations ou l'assimilation, à une zone de forte réflectivité (autour de $30dBZ$) qui a descendu la colonne atmosphérique : on retrouve cette zone de forte réflectivité à $3km$ d'altitude vers le début de notre évènement (19:00-19:10).

Néanmoins, on constate clairement que la réflectivité en haut de la colonne est sous-estimée par VAMOS. Alors que la réflectivité observée décroît assez sensiblement durant la chute des gouttes sur les données KAZR (Figure V.53A : maxima vers $35dBZ$ en haut et $30dBZ$ en bas), elle a plutôt tendance à croître sur nos restitutions. Comme on l'a vu plus haut sur d'autres évènements (section 7.3.1), cela est du au fait que pour notre algorithme, l'atténuation prédomine sur l'évaporation. Expliquer les mesures du KAZR par la seule évaporation nécessiterait cependant des humidités très basses. Cela n'est pas impossible mais peut-être observe-t-on ici un effet combiné de l'évaporation et de mouvements horizontaux particuliers. En tout état de cause, cette variabilité verticale de la réflectivité observée explique la sous-estimation par notre algorithme constatée au sol sur les réflectivités et les taux de pluie pour la phase verte de l'évènement. VAMOS effectue ici une moyenne sur la

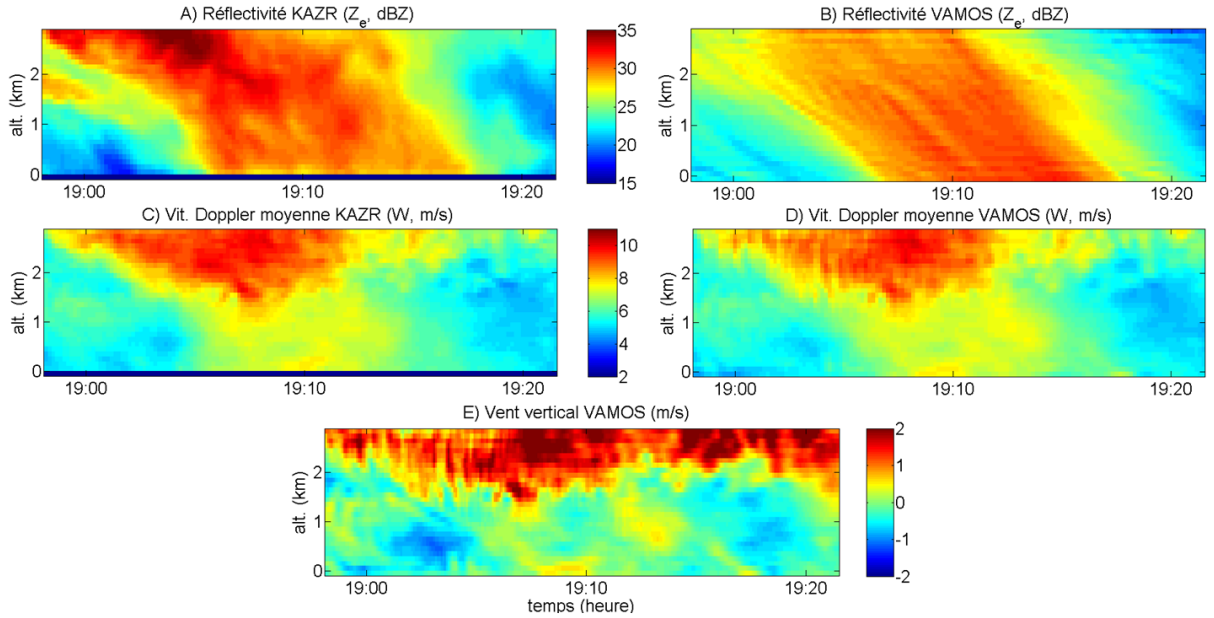

FIGURE V.53 – Pour l'évènement 3 du 04 novembre 2014 : A, B) Facteurs de réflectivité radar équivalents (Z_e) mesurés par le KAZR (A) et assimilés (VAMOS, B). C, D) Idem vitesses Doppler moyennes (W). E) Vent vertical assimilé (VAMOS).

colonne (sur la période de chute des gouttes, correspondant à des diagonales sur les Figures : V.53B par exemple), et par conséquent surestime les pluviométries au sol et les sous-estime en haut de la colonne.

Sur les vitesses Doppler moyennes (Figure V.53C, D), on constate une baisse très sensible lorsqu'on s'approche du sol (d'environ 10 à $7m/s$). Cette baisse ne peut être due seulement à l'effet de la dépendance à l'altitude de la vitesse de chute ou à l'évaporation et est (très bien) expliquée par notre algorithme par le champ de vents verticaux, avec des mouvements descendants assez forts en haut de la colonne. On note que ces mouvements descendants persistent même lorsque le gros de l'évènement est passé (vers 19:20).

La Figure V.54 présente les profils (temps/altitude) de différents moments de la DSD, en dB, du moment d'ordre 0 (A), affecté principalement par les très petites gouttes, au moment d'ordre 6 (F), affecté principalement par les grosses gouttes.

On distingue bien sur cette Figure V.54 les caractéristiques des deux phases de notre évènement (en pointillés, on entoure une partie de ces deux phases). La phase rouge est composée de petites gouttes. Ces petites gouttes tombent lentement (voir la pente de la structure entourée en rouge par exemple sur la Figure V.54A) et sont très affectées par l'évaporation, ce qui explique la diminution très forte des valeurs des moments d'ordre 0, 1 et 2 durant la chute. On note bien que la structure entourée en pointillés rouges ici ne correspond pas exactement aux ovales rouges de la Figure V.52 (qui couvrent au sol la période 19:00-19:10) mais tous deux correspondent à des périodes avec essentiellement des petites gouttes.

La deuxième phase (verte) est composée majoritairement de grosses gouttes (moment d'ordre

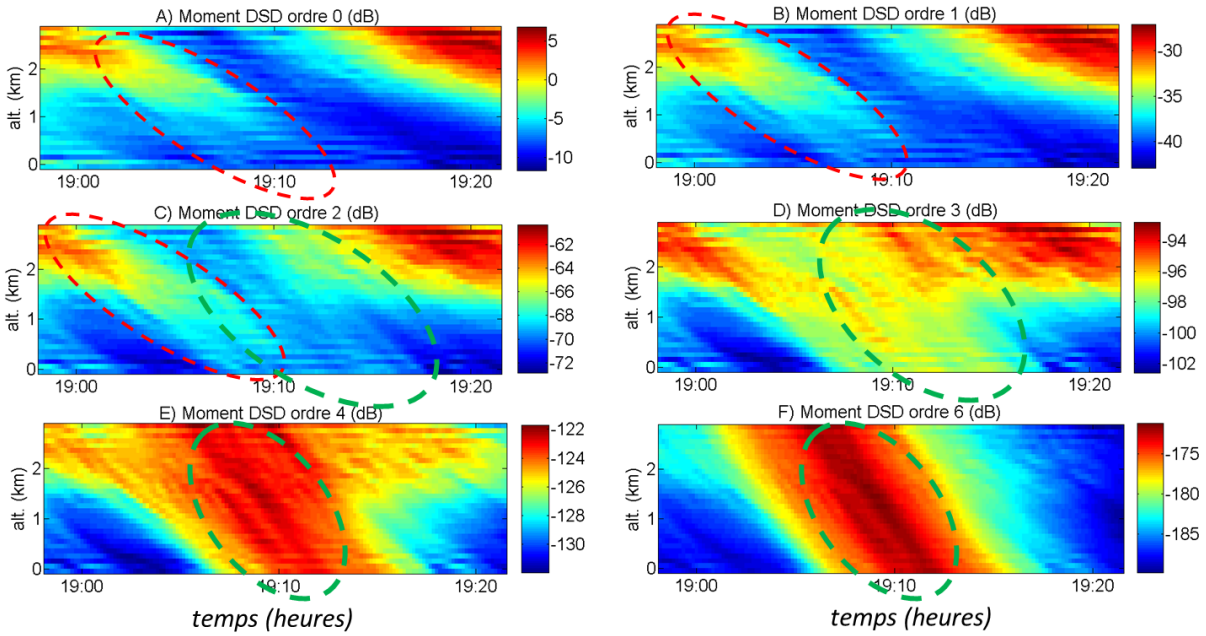

FIGURE V.54 – Pour l'évènement 3 du 04 novembre 2014 : moments d'ordre 0, 1, 2, 3, 4 et 6 des DSD assimilées (VAMOS). Formellement, pour le moment d'ordre n, l'unité est : $10log_{10}\left(m^{n-3}mm^{-1}\right)$. Les pointillés rouges et verts représentent une partie des phases rouges et vertes définies par exemple sur la Figure V.52.

6 fort et d'ordre 0 faible) qui tombent vite (pente de la structure verte de la Figure V.54F forte) et de gouttes de tailles intermédiaires qui tombent moins vite (moment d'ordre 3, Figure V.54D) et sont largement évaporées.

On note enfin que la différentiation entre les deux phases sur cet évènement est particulièrement claire. La phase rouge ne comprend presque que des petites gouttes (elle correspond à un minimum du moment d'ordre 6), et la phase verte presque que des grosses (minimum du moment d'ordre 0), alors que rien n'empêche a priori que les moments évoluent en phase (ou en léger décalage).

Pour terminer cette analyse fine, nous allons présenter la Figure V.55, qui montre les spectres Doppler (observés, KAZR et assimilés, VAMOS), ainsi que les vents verticaux et les DSD VAMOS pour différentes altitudes. Le haut de la Figure représente le haut de la colonne atmosphérique, et le bas de la Figure la première porte radar disponible ($100 - 200m$). On montre pour chaque altitude (chaque ligne) les spectres observés par le KAZR (colonne A, images temps/vitesse Doppler), VAMOS (colonne B, idem), ainsi que les séries temporelles de vents verticaux VAMOS (colonne C) et les DSD VAMOS (colonne D, images temps/diamètre).

Regardons d'abord les spectres KAZR (colonne A). On a entouré les deux phases (verte et rouge) de notre évènement. On constate qu'en haut de la colonne (deux premières lignes), la distinction n'est pas très claire, avec une évolution temporelle des spectres relativement

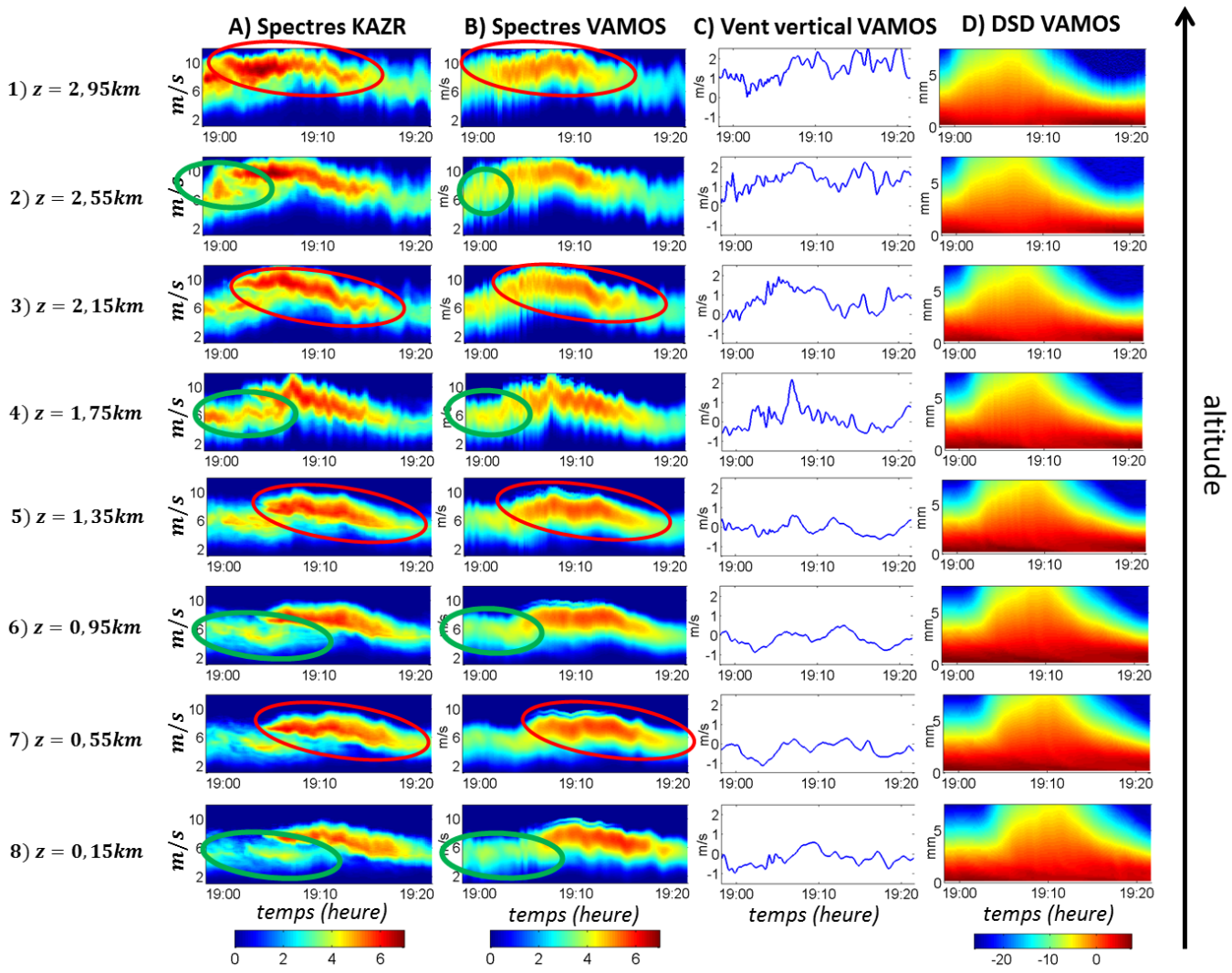

FIGURE V.55 – Pour l'évènement 3 du 04 novembre 2014 : pour différentes altitudes, spectres Doppler mesurés par le KAZR (colonne A) et assimilés (VAMOS, colonne B), vents verticaux VAMOS (colonne C) et DSD VAMOS (colonne D). Vents verticaux en m/s, positifs vers les bas. Spectres Doppler en $ln\,(10^{10}m^{-1}(m/s)^{-1}+1)$, DSD en $log_{10}\,(m^{-3}mm^{-1})$. Les phases rouge (resp. vertes) sont représentées sur les lignes impaires (resp. paires).

régulière. Ce n'est qu'à partir des troisième/quatrième portes représentées (ligne 3 et 4, 2.15km et 1.75km d'altitude) qu'il semble y avoir une différentiation plus nette des deux phases (vers 19:05), se traduisant en particulier par des spectres bimodaux (avec deux maximums vers 6 et 9.5m/s). Il est possible que cette différentiation soit due au vent vertical, puisqu'on est à ce moment là vers l'altitude où le fort mouvement descendant (vent positif) disparaît (voir colonne C de la Figure V.55 ou Figure V.53E). Sur le reste de la colonne (Figure V.55 colonne A, lignes 5 à 8), les deux phases de l'évènement continuent à évoluer de manière relativement indépendante. La phase verte perd beaucoup d'énergie, perte sans doutes due à l'évaporation.

Concernant les spectres assimilés (colonne B), on retrouve globalement les mêmes phénomènes, même si les cisaillements sont moins nets (diffusion numérique et régularité de la solution). Les DSD restituées (colonne D) traduisent bien ces deux phases (beaucoup de petites gouttes en début et fin d'évènement, beaucoup de grosses gouttes en milieu

d'évènement) même si là encore on restitue une solution en apparence assez lisse.

7.3.3 Turbulence / Explication des moments centrés des spectres Doppler

Nous nous intéressons dans cette dernière section à la modélisation dans notre algorithme des effets de la turbulence atmosphérique sur les spectres Doppler. Rappelons d'abord quelques points vus précédemment là-dessus :

- la turbulence lisse les spectres Doppler. Mathématiquement, elle est représentée pour un spectre donné comme une convolution de ce spectre avec une gaussienne centrée et possédant un écart-type (en m/s) caractérisant son intensité (voir chapitre I, section 2.3.2).

- cette convolution se traduit par une augmentation de la largeur spectrale (moment centré d'ordre deux du spectre Doppler), sans modification de la réflectivité (moment d'ordre 0) et de la vitesse Doppler moyenne (moment centré d'ordre 1).

- il n'existe pas dans notre algorithme un tel opérateur de convolution. La turbulence n'est donc pas représentée explicitement. Il n'existe pas d'inconnue de notre algorithme correspondant directement à l'écart-type de la gaussienne mentionné plus haut.

- cependant, VAMOS parvient à restituer les effets de la turbulence sur les spectres Doppler en utilisant les fluctuations du vent vertical à la résolution du modèle ($5s$) sur des fenêtres temporelles correspondant à la résolution des observations ($30s$ ici). Ce phénomène a été détaillé et justifié en section 3.3.6.

Nous allons illustrer ce phénomène sur l'évènement 3 déjà analysé dans la section précédente. Nous avons vu (Tableau V.7) que notre algorithme d'assimilation VAMOS surestimait assez sensiblement la largeur spectrale LS (de 14% pour cet évènement). Étudions à présent plus en détails la structure de LS. La Figure V.56 présente les champs temps/altitude de largeurs spectrales observées par le KAZR (A) et restituées par VAMOS (B).

Considérons d'abord le champ observé (Figure V.56A). On constate plusieurs phases assez différentes. Le début d'évènement (19:00-19:05) est caractérisé par une phase (appelée "zone turbulente" dans la suite) présentant une largeur spectrale importante ($2-3m/s$) entre 3 et $2km$ d'altitude. On a ensuite une phase ("zone calme") avec des largeurs spectrales faibles ($< 1m/s$), en particulier sur le bas de la colonne, puis à nouveau en fin d'évènement des largeurs un peu plus importantes en haut de la colonne. Ce champ ne semble à première vue pas particulièrement corrélé avec ceux représentant les deux premiers moments du spectre Doppler (voir Figure V.53). On peut donc suspecter ici une zone de turbulence plus forte vers 19:00-19:05 en haut de la colonne. En effet, la largeur spectrale n'est pas uniquement déterminée par la turbulence. Elle l'est d'abord par la DSD et sa forme. Mais on peut penser que la variabiité des champs de largeurs spectrales non visible aussi sur les champs de réflectivité n'est pas liée à la DSD et l'est donc à la turbulence. L'effet de cette turbulence sur les largeurs spectrales est bien reproduit par notre algorithme d'assimilation (Figure V.56B).

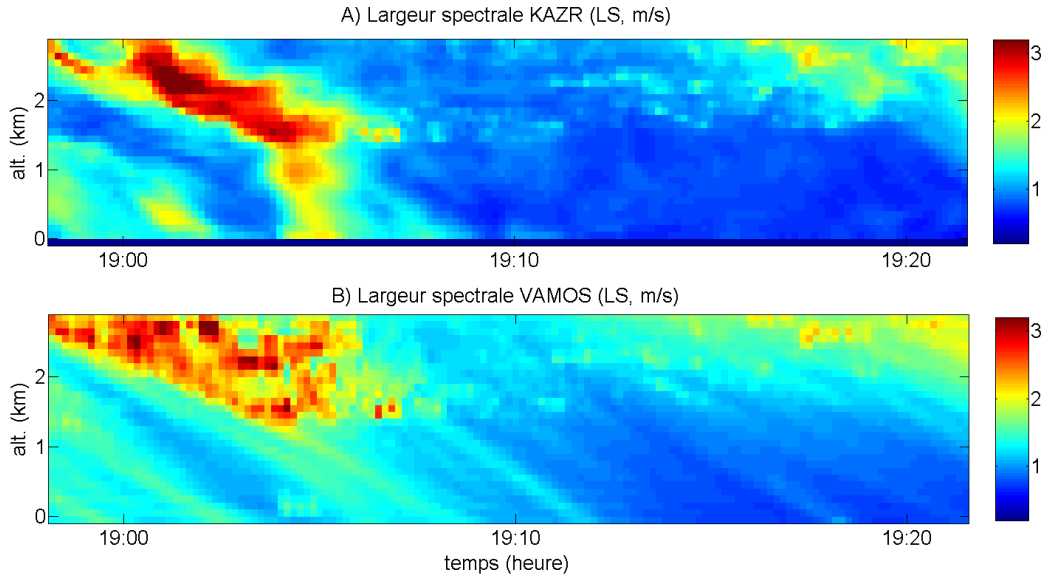

FIGURE V.56 – Pour l'évènement 3 du 04 novembre 2014 : largeurs spectrales mesurées par le KAZR (A) et assimilées (VAMOS, B).

Celui-ci tend effectivement à surestimer LS dans la zone calme (19:05-19:20 en bas de la colonne) - et de manière assez forte en pourcentage vues les petites valeurs mesurées - mais pas sur la zone turbulente. Seule la variabilité de la largeur spectrale observée tout en haut de la colonne (au-dessus de $2.5km$, maxima et minima successifs entre 18:55 et 19:05) n'est pas restituée par notre algorithme (qui restituait également mal la variabilité de la réflectivité à ce niveau, voir Figure V.53A et B), signe peut-être d'une influence du vent horizontal à ce niveau.

On peut cependant noter que cette zone turbulente de plus fortes largeurs spectrales correspond à une zone avec un plus grand nombre de très petites gouttes (moment d'ordre 0 des DSD, Figure V.54A). Peut-être qu'au delà d'une turbulence atmosphérique plus forte, le fait qu'elle s'applique à des plus petites gouttes, que l'on peut penser plus réactives aux variations rapides de vents verticaux, fait qu'elle se traduit plus visiblement sur les spectres Doppler. Néanmoins, ce lien forte largeur spectrale / présence de beaucoup de petites gouttes n'est pas systématique (voir plus loin Figure V.59 et son analyse). Et notre algorithme (comme d'ailleurs les papiers considérant un opérateur de convolution pour modéliser la turbulence) ne prenant pas en compte le temps de réaction des gouttes aux variations du vent vertical, nous continuons à considérer cette zone comme une zone "turbulente", quitte à classer comme "calmes" des zones où la turbulence existe mais où les gouttes présentes sont trop grosses (tombent trop vite) pour qu'elle impacte le champ des largeurs spectrales.

Nous allons maintenant montrer que l'augmentation de la largeur spectrale dans la zone "turbulente" de la Figure V.56B (assimilation) est produite par des fluctuations haute fréquence du vent vertical à $5s$. Pour cela, reportons nous à la Figure V.57. En A, on

représente le champ temps/altitude des vents VAMOS à $5s$ (résolution minimale, celle du modèle). En B, on représente le vent VAMOS dégradé à la résolution des observations ($30s$). On constate bien dans notre zone turbulente (haut de la colonne, 19:00-19:05), des fluctuations fortes du vent à $5s$ (variations dans -2 / $+3m/s$). Ces fluctuations disparaissent si l'on dégrade le champ à $30s$ (B). Sur la zone calme, on ne constate pas de telles fluctuations (les cartes à 5 et $30s$ sont quasiment les mêmes). L'algorithme a donc bien utilisé le vent à $5s$ pour représenter la turbulence.

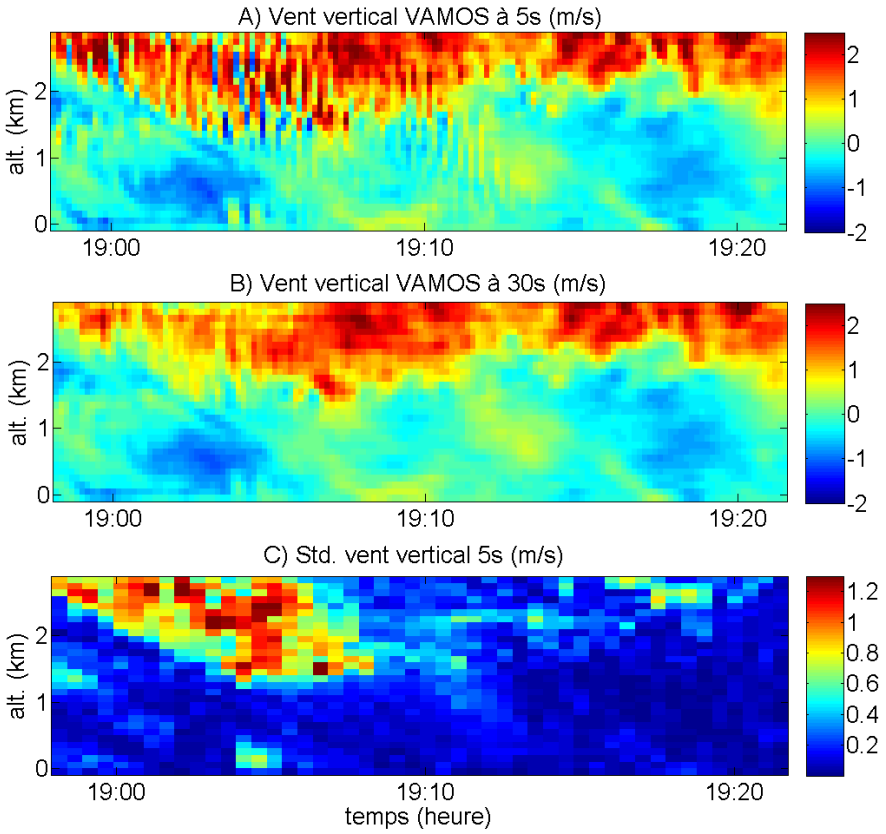

FIGURE V.57 – Pour l'évènement 3 du 04 novembre 2014 : A) Vent assimilé, résolution $5s$; B) Vent assimilé, résolution $30s$; C) Ecart-type des vents à $5s$ sur $30s$ (voir texte). Unité : m/s, positif vers le bas.

Fixons à présent une altitude et une date. On peut calculer l'écart-type de la série temporelle composée des 6 valeurs des vents verticaux (à $5s$) couvrant la période de $30s$ autour de la date considérée. En répétant cette opération pour chaque date et chaque altitude, on peut calculer le champ des écart-types des vents à $5s$ sur des périodes de $30s$. C'est ce champ qui est représenté sur la Figure V.57C. On constate très clairement que cet écart-type est fort dans la zone turbulente et faible (quasiment nul) dans la zone calme. Les valeurs obtenues dans la zone turbulente (de l'ordre de $1m/s$) correspondent par ailleurs assez bien aux valeurs obtenues pour l'écart-type de l'opérateur de convolution modélisant la turbulence dans d'autres études (voir Tridon and Battaglia (2015) par exemple).

L'idée est bien sûr ici de tenter de déterminer une valeur d'intensité de la turbulence à

partir de la variabilité des vents verticaux restitués par notre algorithme. Ne disposant pas de données de validation pour des évènements assez turbulents (l'évènement en Finlande de la section 6 présentait des turbulences très faibles), nous n'irons cependant pas plus loin sur ce sujet pour l'instant.

Nous allons à présent généraliser et conclure quelque peu ce travail sur la turbulence et les moments des spectres Doppler. Tout d'abord, intéressons nous à la Figure V.58. Elle présente, toujours pour l'évènement 3, les trois premiers moments centrés des spectres Doppler hydrologiques (ligne 1), atténués (ligne 2) et "obs" (ligne 3). Les spectres hydrologiques sont ceux définis par l'Eq.(V.7). Ils ne dépendent que de la DSD et de la longueur d'onde (et ne prennent donc en particulier pas en compte l'atténuation et l'effet du vent vertical). Les spectres atténués ajoutent aux précédents l'effet de l'atténuation due à la pluie présente dans les couches inférieures. Les spectres "obs" (au sens de l'Eq. (V.11)) sont les spectres atténués auxquels on a ajouté l'effet du vent vertical (à $5s$).

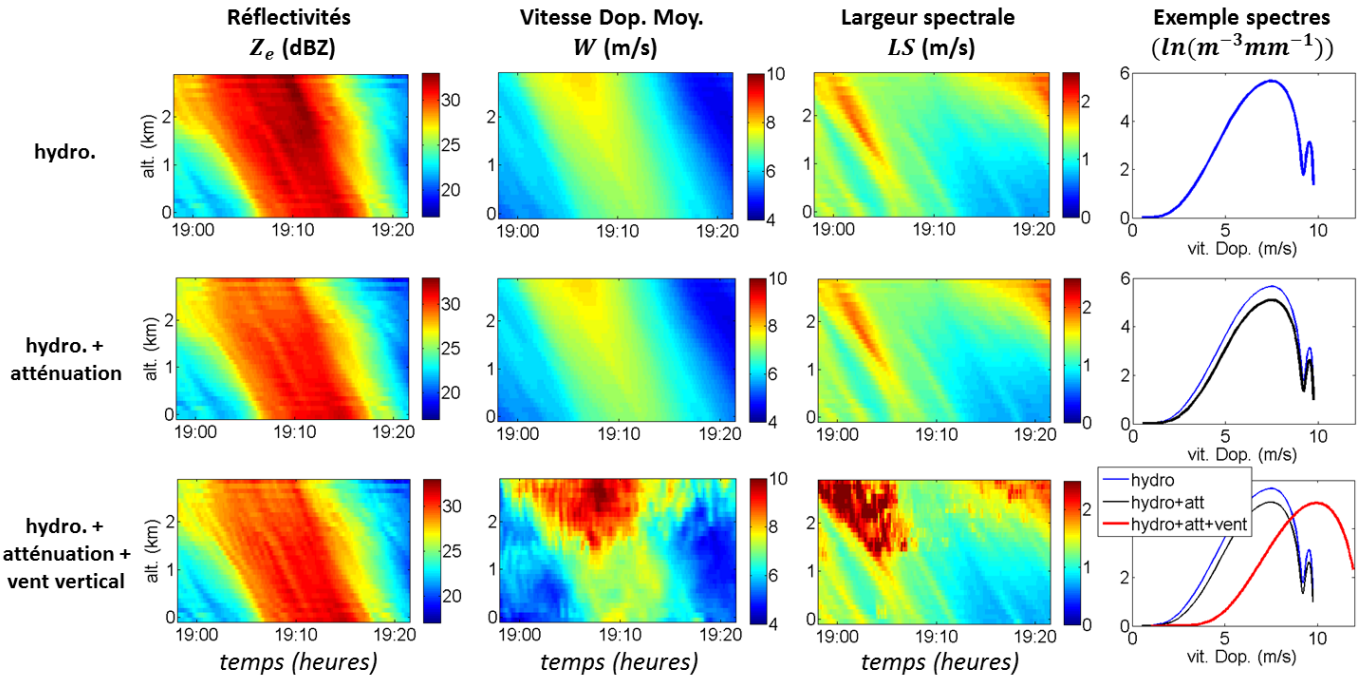

FIGURE V.58 – Pour l'évènement 3 du 04 novembre 2014 : Trois premiers moments centrés des spectres Doppler VAMOS (réflectivité Z_e, gauche ; vitesse Doppler moyenne W, milieu gauche, largeurs spectrales LS, milieu droite) et exemple des spectres correspondants (moyennés sur la période 19:01-19:06). Haut : spectres hydrologiques ; Milieu : ajout de l'atténuation ; Bas : Ajout du vent vertical (voir texte pour les détails et commentaires).

On constate que la DSD (ligne 1) détermine largement la forme générale des champs, en particulier celui de réflectivités.

Ensuite, l'atténuation impacte uniquement la réflectivité (intégrale du spectre), et pas la vitesse Doppler moyenne et la largeur spectrale.

Le vent vertical impacte lui à la fois la vitesse Doppler moyenne et la largeur spectrale, puisqu'on a vu qu'il servait aussi (à $5s$) dans notre algorithme à reproduire les effets de la turbulence.

Notre algorithme dispose donc de moyens d'expliquer les trois premiers moments du spectre Doppler relativement indépendants : la DSD explique la réflectivité, le vent à $30s$ explique la variabilité du champ de vitesses Doppler moyennes non expliquée par la DSD, et le vent à $5s$ explique la variabilité de largeur spectrale non expliquée par la DSD.

Pour terminer, nous revenons sur la zone de turbulence suspectée sur le bas de la colonne $(0.3 - 1km)$ vers 20:00-20:30 (durant l'évènement 5) à partir des images de la Figure V.49 (A et C, cadres blancs). On remarquait ici une forte variabilité de la vitesse Doppler moyenne enregistrée par le radar BASTA et une forte variabilité également de notre vent vertical VAMOS à $5s$ (signe de turbulence comme on vient de le voir).

La Figure V.59 présente les largeurs spectrales observées par le KAZR et restituées par VAMOS (B) pour cette période, avec toujours un cadre blanc repérant la zone de turbulence, ainsi que la moyenne de ces largeurs spectrales selon l'altitude (C, calculées sur la période 20:00-20:30).

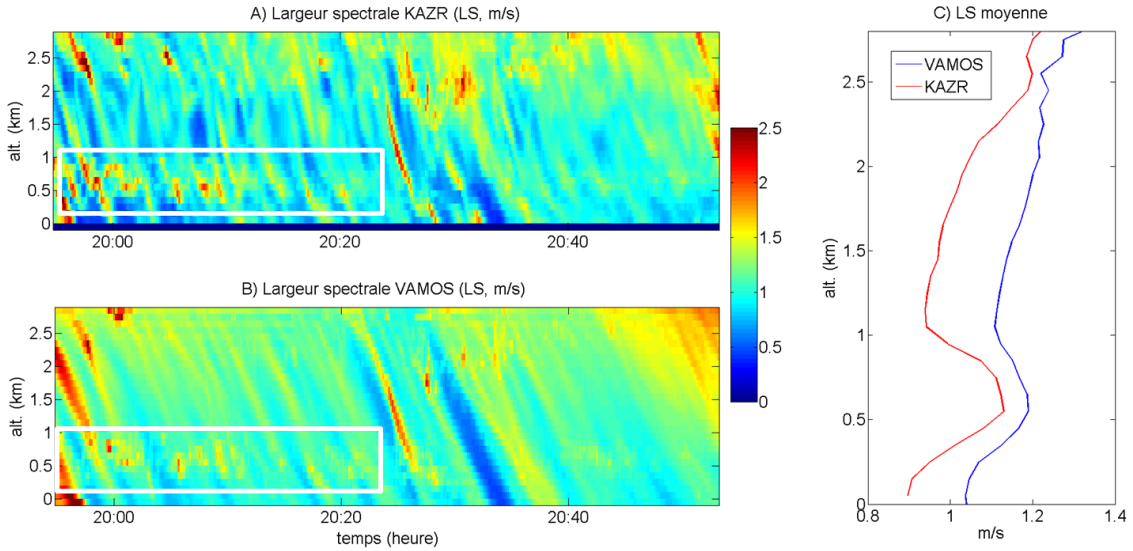

FIGURE V.59 – Pour l'évènement 5 du 04 novembre 2014 : Largeurs spectrales observées par le KAZR (A) et assimilés (B). Profils verticaux moyens sur la période 20:00-20:30. Les cadres blancs, comme sur la Figure V.49, délimitent la zone de turbulence évoquée dans cette section.

L'augmentation de la largeur spectrale autour de $800m$ d'altitude est bien visible sur les observations du KAZR (Figure V.59A). Cela est moins visible sur le champ restitué par notre algorithme (Figure V.59B). Dans la mesure où celui-ci tend à surestimer globalement la largeur spectrale, l'augmentation apportée par cette zone de turbulence se détache moins sur nos restitutions. En regardant les profils moyens de largeurs spectrales (Figure V.59C), on remarque néanmoins bien un "pic" de largeurs spectrales vers $700m$ d'altitude, que ce soit sur

les données KAZR ou VAMOS. Cette couche de turbulence est particulièrement bien visible sur cet évènement et n'est pas liée à la présence d'un très grand nombre de petites gouttes comme dans l'évènement 3, comme on peut le voir sur la Figure V.48B et E.

7.4 Conclusions

Nous avons dans cette étude appliqué notre algorithme d'assimilation VAMOS à un cas plus complexe que les précédents, ayant eu lieu à Darwin, dans le nord de l'Australie, sous un climat tropical. Le cas de pluie considéré, s'il était toujours constitué de taux de pluie assez faibles ($< 10mm/h$) était composé de plusieurs évènements présentant des caractéristiques assez différentes. Ces évènements variaient par les taux de pluie au sol qu'ils impliquaient, mais aussi par la structure des champs de réflectivité radar associés, laissant supposer selon l'évènement un impact plus ou moins prépondérant des vents horizontaux, de l'évaporation, des vents verticaux, de la turbulence ou des collisions entre gouttes.

Nous avons donc analysé ces différents évènements et la capacité de notre algorithme à les expliquer, à partir de données parfois suspectes (humidités relatives en particulier). Nous avons ainsi analysé l'importance de l'évaporation sur les profils verticaux de réflectivités, de diamètres volumiques moyens ou de contenus en eau liquide. En comparant nos restitutions aux mesures d'un autre radar Doppler en bande W, nous avons montré que la variabilité du champ de vitesses Doppler moyennes était principalement déterminée par le vent vertical (pour ces évènements). Nous avons ensuite analysé en détails l'un des évènements considéré ici, puis nous nous sommes attardés sur la turbulence atmosphérique, son impact sur les mesures radar et la restitution de cet impact par VAMOS via l'utilisation des fluctuations rapides du vent vertical.

Une fois de plus, nous avons mis en évidence l'intérêt qu'il y aurait à contrôler les profils d'humidité relative (toujours sous réserve que celle-ci ne soit pas trop variable dans le temps pour ne pas ajouter trop de degrés de liberté à notre problème). Nous avons aussi montré qu'il pourrait être possible de déterminer des intensités de turbulence atmosphérique à partir de nos restitutions (ce que nous ne faisons pas pour l'instant par manque de possibilités de validation).

8 Conclusions et perspectives

8.1 Conclusions

Nous avons développé VAMOS, un algorithme d'assimilation de données 4D-VAR restituant les profils verticaux de DSD et les vents verticaux dans la pluie à partir d'observations radar et disdromètre contraintes par un modèle de propagation verticale des gouttes de pluie. Cet algorithme et ses restitutions ont été utilisées sur trois cas d'étude, couvrant des conditions météorologiques différentes, afin d'étudier la variabilité spatiotemporelle des DSD et des vents verticaux et les phénomènes physiques la conditionnant (évaporation, vents verticaux et horizontaux, etc.).

Nous nous sommes limités dans ce travail à l'utilisation d'un modèle de propagation assez simple, prenant en compte les effets du vent vertical et de l'évaporation, mais pas des mouvements horizontaux, ni des collisions entre gouttes.

En ce qui concerne le vent horizontal, nous avons réalisé une étude (expérience jumelle d'assimilation de données) consistant à simuler des observations radar et disdromètres en utilisant un modèle de propagation avec vent horizontal, puis d'assimiler ces observations simulées dans un code d'assimilation sans vent horizontal. Cette méthode a permis de montrer que l'on parvenait à extraire des informations cohérentes, y compris sur le vent vertical, malgré l'erreur commise sur le vent horizontal. Nous nous sommes néanmoins limités à un vent pas trop fort ($< 6m/s$).

Afin de limiter les erreurs commises par le fait de négliger le vent horizontal et les collisions entre gouttes, nous avons restreint notre étude à des évènements de pluie modérés (typiquement avec des taux de pluie en-dessous de $10mm/h$).

Sur le premier cas d'étude (section 5), concernant des données provenant du sud de la France dans le cadre de la campagne HyMeX, nous avons montré la capacité de notre algorithme VAMOS à fusionner les mesures radar et disdromètre, a priori très différentes et difficiles à comparer (de par leur localisation comme de par leur nature). Nous avons aussi noté que cette capacité à restituer les observations était bien sûr limitée par l'aptitude du modèle de propagation (et de la physique qu'il contient) à représenter la variabilité verticale et temporelle des observations. Par exemple sur ce premier cas d'étude, nous avons constaté qu'une tendance visible sur les observations radar (croissance de la réflectivité en s'approchant du sol) ne peut être restituée par notre modèle. En l'absence de données indépendantes, nous n'avons cependant pas pu conclure sur l'origine de cette tendance (problème de calibration du radar, microphysique non représentée dans le modèle, mouvements horizontaux ?). Nous nous sommes bornés à constater que l'inclusion d'un modèle de propagation apportait une cohérence physique stricte à nos restitutions (contrainte éventuellement trop forte) et que ses restitutions (DSD, vents verticaux) semblaient cohérentes avec ce que l'on attendait pour un tel évènement de pluie stratiforme.

Le deuxième cas d'étude traitait un évènement enregistré en Finlande dans le cadre de la campagne ARM. Il s'agissait d'un évènement de pluie assez semblable au précédent (pluie faible et régulière), mais présentant tout de même a priori une évaporation plus importante. Nous avons pu comparer et valider nos restitutions (vents verticaux et DSD) avec celles de Tridon and Battaglia (2015). Nous avons ainsi montré la cohérence de nos restitutions malgré l'utilisation d'un seul radar Doppler à $35GHz$ (et d'un disdromètre). Nous avons néanmoins montré que VAMOS produisait un léger biais sur les restitutions de vents verticaux (de l'ordre de $0.2m/s$), biais déjà constaté sur données simulées et attribué au fait que l'on néglige le vent horizontal. On note également qu'en l'état (restitution uniquement du vent vertical et des DSD avec humidité et atténuation du signal radar par le radôme estimés a priori), la qualité de nos restitutions dépendait fortement de la paramétrisation du modèle de propagation (une erreur sur l'atténuation due au radôme induisant par exemple un biais sur les restitutions de taux de pluie).

Le troisième cas d'étude traitait des évènements plus complexes, enregistrés en climat tropical à Darwin, en Australie. Sur ces évènements, nous nous sommes davantage concentrés sur les phénomènes physiques influençant la DSD sur la colonne atmosphérique, et avons montré que les évènements semblaient marqué par un impact plus fort du vent horizontal, de l'évaporation, ou de l'atténuation, selon le cas. Nous avons également accordé une importance particulière à la gestion de la turbulence atmosphérique par notre algorithme. À ce sujet, nous avons montré que bien qu'elle ne soit pas paramétrée explicitement, celui-ci pouvait restituer ses effets sur les spectres Doppler (élargissement spectral) en utilisant les fluctuations du vent vertical sur de courtes périodes temporelles.

8.2 Perspectives

Les perspectives d'amélioration et d'utilisation de cette méthode sont nombreuses, tant d'un point du vue algorithmique que scientifique (les deux étant liés).

D'abord, nous avons plusieurs fois noté que la qualité des résultats était fortement liée pour l'instant à la paramétrisation dans notre modèle de propagation de l'évaporation (et en particulier des profils d'humidité relative). Ces profils proviennent généralement soit de radiosondages, qui n'ont pas forcément lieu au moment idéal, soit de sorties de modèles, qui ne prévoient pas forcément exactement la situation effective. Il pourrait donc être intéressant de contrôler ces profils d'humidité, i.e. de les considérer comme des inconnues de notre algorithme d'assimilation. Le même raisonnement est applicable à l'atténuation du signal radar dû à son radôme. Nous avons pour l'instant estimé et corrigé cette dernière a priori à partir des données disdromètres. Elle pourrait également être considérée comme une inconnue de l'algorithme. Ces améliorations permettraient de se débarrasser de certains biais et de se focaliser davantage sur les autres phénomènes physiques impliqués.

Il pourrait également être intéressant de pousser plus loin le traitement de la turbulence commencé ici. Nous avons vu que VAMOS restituait grâce aux fluctuations haute fréquence du vent vertical l'effet produit par la turbulence atmosphérique sur les spectres Doppler, et qu'il semblait que nous puissions caractériser l'intensité de cet effet par le calcul des écart-type locaux des séries temporelles de vents verticaux restitués. Ce travail reste néanmoins à perfectionner. D'abord car les écarts-types ainsi calculés semblent a priori dépendants de la résolution à laquelle sont les observations fournies en entrée de l'algorithme. Ensuite car un élargissement spectral peut aussi être produit dans notre modèle par d'autres phénomènes (diffusion numérique en particulier). Il serait donc intéressant de comparer nos restitutions à ce niveau avec celles d'autres algorithmes. La validation de notre travail ayant été effectuée ici pour un évènement particulièrement calme et peu turbulent, ce travail reste à faire. On note enfin à ce sujet qu'il pourrait être intéressant de modéliser le fait que la turbulence atmosphérique affecte davantage les petites gouttes que les grosses (qui réagissent a priori moins vite à des fluctuations du vent vertical).

Il paraîtrait également naturel d'ajouter au modèle de propagation les effets des collisions entre gouttes (coalescence / break-up), et aussi éventuellement une voire deux dimensions horizontales. Ce perfectionnement permettrait de traiter des évènements de pluie plus intenses, convectifs. Cela pourrait être fait sous réserve de disposer de suffisamment d'observations pour que le problème ne soit pas drastiquement sous-déterminé. On remarque cependant que le formalisme utilisé (assimilation de données et formalisme de graphe modulaire du logiciel YAO) rend ces perfectionnements assez simples algorithmiquement (au-delà des problèmes de coûts de calcul et de sous-détermination qu'ils pourraient engendrer).

L'intérêt scientifique d'inclure ces différents phénomènes dans le modèle de propagation serait de contrôler (de régler) leur paramétrisation directement grâce aux observations dans l'algorithme d'assimilation (c'est-à-dire de passer un certain nombre de leurs paramètres en inconnues de l'algorithme), quitte à les contraindre fortement ou à ne les "activer" qu'un par un selon l'évènement considéré. Pour un tel travail d'étude physique et de paramétrisation, l'utilisation de spectres Doppler à deux fréquences différentes comme observations (données d'entrée) dans l'algorithme d'assimilation pourrait être d'une grande aide afin de couvrir toute la gamme de vitesses de chute. L'utilisation combinée ou voisine de données provenant d'un radar en bande Ka et de profileurs de vent UHF et/ou VHF (instruments disponibles sur des sites voisins à Darwin dans le cadre de la campagne TWP/ICE de 2006 - http://acrf-campaign.arm.gov/twpice/) paraît également prometteuse.

Conclusion générale

Dans cette thèse, nous avons travaillé sur des algorithmes d'assimilation de données d'observation des précipitations afin d'étudier leur physique et leur variabilité spatiotemporelle à différentes échelles.

Les précipitations peuvent en effet être envisagées à la fois à l'échelle de la cellule de pluie (échelle macrophysique) ou à celle de la goutte (échelle microphysique). Différents instruments existent pour évaluer les précipitations à chaque échelle via des mesures directes ou indirectes. Mais ces instruments, de par leurs natures différentes, sont souvent difficilement intercomparables. Des modèles existent également pour prédire l'évolution de la pluie. Mais ces modèles sont souvent complexes et difficiles à paramétrer et à valider. L'assimilation de données, dans la mesure où elle permet de lier modèles et observations, apparaissait donc comme une théorie intéressante pour développer les connaissances sur les précipitations. L'idée dans cette thèse était de fusionner des observations hétérogènes en utilisant l'assimilation de données qui assure, via un modèle, une cohérence physique permanente. Nous avons dans un premier temps travaillé à l'échelle macrophysique, en développant un algorithme restituant des cartes de précipitations à partir de mesures d'atténuations d'ondes TV-SAT. Ensuite, nous avons étudié la pluie à échelle microphysique. Pour cela, nous avons développé une méthode permettant de restituer les profils verticaux de DSD (distribution en taille des gouttes de pluie - Drop Size Distribution) et de vents verticaux. Nous avons appliqué cette méthode à trois cas d'étude afin de valider notre approche et d'étudier la fusion des observations et l'analyse physique des évènements qu'elle permet.

Les deux premiers chapitres de ce manuscrit ont introduit les notions associées à ce travail (définition, observation et modélisation de la pluie). Le troisième chapitre a présenté l'assimilation de données et son application à ce travail. Les chapitres 4 et 5 ont présenté les deux applications. Résumons ici leurs principaux enseignements.

Dans le chapitre I, nous avons tout d'abord défini les précipitations et les principales grandeurs les caractérisant. Partant de la goutte d'eau individuelle, nous avons défini la distribution en taille des gouttes de pluie (DSD), et les paramètres intégrés associés (taux de pluie, contenu en eau liquide, facteur de réflectivité radar). Ensuite, nous sommes revenus sur les instruments permettant de mesurer la pluie utilisés dans ce travail et sur la nature de leur mesure. Nous avons ainsi évoqué les instruments au sol mesurant directement des flux d'eau :

pluviomètres et disdromètres, et les instruments de télédétection mesurant indirectement la pluie : radars météorologiques et capteurs d'atténuation. Nous avons évoqué, en particulier pour ces deux derniers instruments, les difficultés d'inversion qui se posent (difficulté à évaluer un taux de pluie ou une DSD à partir de leur mesure brute).

Dans le chapitre II, nous sommes revenus sur la dynamique de la pluie et les modèles développés pour la simuler. Nous nous sommes d'abord concentrés sur l'échelle macrophysique, et avons présenté les modèles d'advection-diffusion représentant l'évolution macroscopique des cellules de pluie, indépendamment des phénomènes physiques affectant les gouttes individuelles. Ensuite, nous nous sommes attardés sur l'équation aux dérivées partielle modélisant l'évolution des DSD, qui simule les différents phénomènes physiques agissant sur les gouttes durant leur chute (effet du vent, évaporation, collisions entre gouttes). Nous avons montré sa complexité et la difficulté qu'il y a à la paramétrer et à la valider. Nous avons donné les relations permettant de lier diamètre des gouttes et vitesse terminale de chute, ainsi que quelques schémas permettant de résoudre numériquement les modèles continus.

Le chapitre III concernait l'assimilation de données, théorie mathématique permettant de coupler des observations dans des modèles numériques de propagation. Après avoir rapidement passé en revue les différents algorithmes d'assimilation existant et leurs principales caractéristiques, nous sommes revenus plus en détails sur le 4D-VAR, algorithme de minimisation permettant de déterminer un état initial (ou les paramètres) d'un modèle, qui une fois propagé sera le mieux à-même d'expliquer les observations disponibles. Nous présentons également YAO, un logiciel permettant de faciliter la programmation et l'entretien d'algorithmes d'assimilation 4D-VAR.

Le chapitre IV a présenté la première application de ce travail. Il s'agissait de restituer des cartes de précipitations à partir de mesures d'atténuations d'ondes électromagnétiques provenant de satellites de télévision en bande Ku. Pour cela, nous utilisions l'assimilation 4D-VAR avec un modèle supposant les cellules de pluie déplacées sans déformation par le vent (advection pure). Dans ce travail, nous ne disposions que d'un récepteur permettant de mesurer les atténuations le long de 4 liaisons Terre-satellite. Ce dispositif expérimental ne nous permettant pas de calculer les vitesses de déplacement des cellules, nous les avons déterminées à partir des mesures radar. Mais nous montrons (sur données simulées) que l'utilisation d'au moins 2 capteurs permettrait de retrouver les conditions d'advection directement à partir des séries temporelles d'atténuations.

Cet algorithme a été d'abord testé sur données simulées. On a ainsi montré sa capacité à restituer des champs de pluie (à résolution $10s$ / $500 * 500m^2$) reproduisant très précisément les quantités d'eau tombées sur la zone couverte, et à spatialiser correctement les zones de précipitations intenses. Il semble néanmoins avoir tendance à lisser les champs de pluie (à la fois à cause du caractère intégré des mesures et de la diffusion numérique inhérente au modèle de

propagation), et à être assez sensible à la direction d'advection (la direction du vent). Cela est dû à la géométrie particulière des liaisons Terre-satellite, toutes orientées vers le sud (satellites géostationnaires au-dessus de l'équateur). L'algorithme a ensuite été appliqué à des cas réels, enregistrés en Ardèche à l'été et à l'automne 2013. Il s'agit d'évènements de pluie intenses. On a montré que notre méthode était applicable sous réserve que les évènements présentent un caractère suffisamment advectif, c'est-à-dire que le modèle de transport sans déformation les représente assez bien, selon un critère statistique précis évaluable a priori. Ce critère était vérifié pour 4 évènements étudiés sur 8. On a alors montré que pour ces évènements, on obtenait des cartes de pluie restituées dont les caractéristiques étaient comparables à celles obtenues par mesure radar (à résolution 500m / 5 min), notamment lorsqu'on les comparait à des mesures ponctuelles de pluviomètres.

Nous avons donc utilisé un dispositif peu coûteux (comparé aux radars ou à l'entretien d'un réseau dense de pluviomètres) pour produire des cartes de pluie à fine échelle. Des bases de données plus conséquentes pourraient être obtenues en installant des récepteurs directement sur les récepteurs personnels existant. Cette méthode donne ainsi des résultats très prometteurs pour la gestion des ressources en eau dans des zones non couvertes par les radars météorologiques (zones montagneuses, pays en développement), ou pour l'hydrologie urbaine (où un grand nombre de récepteurs est potentiellement disponible, permettant une résolution très fine).

Le chapitre V a présenté la seconde application. Il s'agissait de fusionner des mesures directes et indirectes de la DSD au sol (flux de gouttes de disdromètres) et en altitude (spectres Doppler de radars à visée verticale) pour étudier sa variabilité et les phénomènes physiques l'affectant. Pour cela, on a utilisé un algorithme 4D-VAR baptisé VAMOS (Variational Assimilation of Multi-scale Observations for rain Studies) couplant ces observations dans un modèle de propagation des DSD. Ce modèle simule la chute des gouttes dans la colonne atmosphérique sous l'isotherme $0\,^{\circ}C$ (on ne travaille que dans la pluie) sous l'influence de la gravité (chute des gouttes), du vent vertical et de l'évaporation. En particulier, on néglige le vent horizontal et les collisions entre gouttes. L'algorithme 4D-VAR restitue des profils de DSD et de vents verticaux compatibles avec les observations, et dont la cohérence physique est assurée par le modèle.

On néglige donc les mouvements horizontaux des gouttes, en réduisant le problème à une colonne atmosphérique. Les implications de cette simplification ont été étudiées séparément. On a ainsi d'abord travaillé sur un modèle théorique simple prédisant la signature du vent horizontal (selon son profil vertical) sur les mesures radar. On a ensuite utilisé un modèle numérique de propagation des DSD avec vent horizontal, d'une part pour vérifier sur données simulées les sorties de ce modèle théorique simple, et d'autre part pour simuler des observations radar et disdromètre dans des conditions plus réalistes. En appliquant ensuite notre algorithme d'assimilation VAMOS sans vent horizontal à ces observations simulées (expérience jumelle d'assimilation), on montre que l'on restitue malgré cette erreur modèle des profils de DSD et

de vents verticaux raisonnables (sous réserve que l'évènement soit suffisamment régulier et le vent pas trop fort, typiquement en-dessous de $6m/s$).

Nous avons ensuite appliqué VAMOS à 3 cas d'étude. Ces trois cas d'étude avaient pour point commun de concerner des évènements pas trop intenses (taux de pluie sous les $10mm/h$) afin de limiter les erreurs dues au fait de négliger le vent horizontal et les collisions entre gouttes. Ils couvraient cependant des régions différentes (Ardèche, Finlande, nord de l'Australie) et des conditions météorologiques (température, vent, humidité relative) variées. Les données d'entrée (utilisées pour restituer les profils de DSD) provenaient de deux types de disdromètres et de radars Doppler à visée verticale à 24 ou $35GHz$.

On a d'abord montré l'aptitude générale de l'algorithme à expliquer les observations qu'on lui fournit. On montre aussi que lorsque les observations sont incompatibles entre elles (du fait généralement qu'elles ne couvrent pas les mêmes échelles), ou lorsque le modèle inclus dans l'algorithme de restitution est incapable de représenter correctement les profils observés, on restitue des champs représentant un comportement moyen sur la colonne atmosphérique. La qualité de nos restitutions est également sensible à la paramétrisation des phénomènes inclus directement ou indirectement dans le modèle (humidité relative, pertes d'énergie due à l'eau présente sur les radômes des radars), et des erreurs sur ces paramètres peuvent induire des biais. On note que ces paramétrisations pourraient être contrôlées directement par l'algorithme.

Les restitutions de VAMOS en elles-mêmes ont été validées sur le cas Finlandais par rapport à une autre méthode utilisant les spectres de deux radars Doppler à deux fréquences différentes. On montre ainsi notamment que l'on est capables de restituer les mouvements verticaux de l'air à partir d'un seul radar à $35GHz$ et d'un disdromètre, en utilisant la cohérence spatiotemporelle des DSD fournie par le modèle.

Nous montrons également que bien que non paramétrée explicitement dans l'algorithme, celui-ci restitue également l'effet de la turbulence atmosphérique sur les spectres Doppler (élargissement spectral) en utilisant les fluctuations du vent vertical à haute fréquence. Cela n'est cependant possible que si les observations sont intégrées à une résolution plus large que le modèle (typiquement $30s$ vs. $5s$).

Enfin, nous avons montré qu'à la lumière de notre méthode nous pouvions, pour un évènement de pluie donné, déterminer l'importance relative des phénomènes physiques représentés dans l'algorithme (évaporation, atténuation, vent vertical, turbulence) et l'influence potentielle de ceux ne l'étant pas (vent horizontal, collisions entre gouttes).

Ce travail pourrait être étendu à des évènements convectifs (sujets a priori à une influence plus forte du vent horizontal et des collisions) en ajoutant la paramétrisation de ces phénomènes dans le modèle de propagation. Ce travail, grâce au formalisme de YAO, est assez simple à mettre en oeuvre mais doit être effectué prudemment pour éviter un nombre trop important de degrés de liberté (le nombre d'inconnues de l'algorithme devant être en adéquation avec la quantité d'observations disponible).

Bibliographie

(1998). *MMR, Physical Basis*. Available from METEK GmbH, Fritz-Straßmann-Straße 4, D-25337 Elmshorn, Germany.

Akrour, N. (2015). *Simulation stochastique des précipitations à fine échelle : Application à l'observation en milieu urbain*. PhD thesis, Université Versailles-St-Quentin-en-Yvelines / Université Paris-Saclay.

Akrour, N., Mercier, F., Mallet, C., and Barthès, L. (2015). Rainfall at urban hydrological scales: Opportunist use of tv-sat receivers network. In *Raingain Conference" Researchers & Water Managers Preparating Cities for a Changing Climate" in the Framework of the COP21 Preparation*.

Allen, D. J., Douglass, A. R., Rood, R. B., and Guthrie, P. D. (1991). Application of a monotonic upstream-biased transport scheme to three-dimensional constituent transport calculations. *Monthly Weather Review;(United States)*, 119.

Arakawa, A. and Lamb, V. R. (1977). Computational design of the basic dynamical processes of the ucla general circulation model. *Methods in computational physics*, 17:173–265.

Atlas, D., Srivastava, R. C., and Sekhon, R. S. (1973). Doppler radar characteristics of precipitation at vertical incidence. *Reviews of Geophysics*, 11(1):1–35.

Atlas, D. and Ulbrich, C. W. (1977). Path-and area-integrated rainfall measurement by microwave attenuation in the 1-3 cm band. *Journal of Applied Meteorology*, 16(12):1322–1331.

Barros, A. P., Prat, O. P., Shrestha, P., Testik, F. Y., and Bliven, L. F. (2008). Revisiting low and list (1982): Evaluation of raindrop collision parameterizations using laboratory observations and modeling. *Journal of the Atmospheric Sciences*, 65(9).

Barros, A. P., Prat, O. P., and Testik, F. Y. (2010). Size distribution of raindrops. *Nature Physics*, 6(4):232–232.

Barthès, L. and Mallet, C. (2013). Rainfall measurement from the opportunistic use of an earth–space link in the ku band. *Atmospheric Measurement Techniques*, 6(8):2181–2193.

Barthes, L. and Mallet, C. (2013). Vertical evolution of raindrop size distribution: Impact on the shape of the DSD. *Atmospheric Research*, 119:13–22.

Beard, K. V. (1976). Terminal velocity and shape of cloud and precipitation drops aloft. *Journal of the atmospheric sciences*, 33(5):851–864.

Beard, K. V. (1985). Simple altitude adjustments to raindrop velocities for doppler radar analysis. *Journal of Atmospheric and Oceanic Technology*, 2(4):468–471.

Beard, K. V. and Ochs III, H. T. (1995). Collisions between small precipitation drops. part II: Formulas for coalescence, temporary coalescence, and satellites. *Journal of the atmospheric sciences*, 52(22):3977–3996.

Beheng, K. D. (2010). The evolution of raindrop spectra: A review of microphysical essentials. *Rainfall: State of the Science*, page 29–48.

Beheng, K. D., Jellinghaus, K., Sander, W., Roth, N., and Weigand, B. (2006). Investigation of collision-induced breakup of raindrops by numerical simulations: First results. *Geophysical research letters*, 33(10).

Berne, A. and Krajewski, W. F. (2013). Radar for hydrology: Unfulfilled promise or unrecognized potential? *Advances in Water Resources*, 51:357–366.

Best, A. C. (1950). Empirical formulae for the terminal velocity of water drops falling through the atmosphere. *Q.J.R. Meteorol. Soc.*, 76(329):302–311.

Bianchi, B., Jan van Leeuwen, P., Hogan, R. J., and Berne, A. (2013). A variational approach to retrieve rain rate by combining information from rain gauges, radars, and microwave links. *J. Hydrometeor*, 14(6):1897–1909.

Bocquet, M. (2005). Introduction aux principes et méthodes de l'assimilation de données en géophysique. *Notes de cours de l'ENSTA*.

Bousquet, O., Berne, A., Delanoë, J., Dufournet, Y., Gourley, J., Van-Baelen, J., Augros, C., Besson, L., Boudevillain, B., Caumont, O., et al. (2015). Multifrequency radar observations collected in southern france during hymex-sop1. *Bulletin of the American Meteorological Society*, 96(2):267–282.

Brazier-Smith, P. R., Jennings, S. G., and Latham, J. (1972). The interaction of falling water drops: coalescence. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 326(1566):393–408.

Brenguier, J. (1991). Parameterization of the condensation process: A theoretical approach. *Journal of the atmospheric sciences*, 48(2):264–282.

Brisseau, O. (2005). *Développement de modèles de similitude en fréquence pour les liaisons Terre-Espace en EHF*. PhD thesis, Rennes 1.

Brutsaert, W. (2013). *Evaporation into the atmosphere: theory, history and applications*, volume 1. Springer Science & Business Media.

Chwala, C., Gmeiner, A., Qiu, W., Hipp, S., Nienaber, D., Siart, U., Eibert, T., Pohl, M., Seltmann, J., Fritz, J., et al. (2012). Precipitation observation using microwave backhaul links in the alpine and pre-alpine region of southern germany. *Hydrology & Earth System Sciences*, 16(8).

Chwala, C., Keis, F., and Kunstmann, H. (2015). Real time data acquisition of commercial microwave link networks for hydrometeorological applications. *Atmospheric Measurement Techniques Discussions*, 8(11).

Chèze, J.-L., Moisselin, J.-M., Tzanos, R., and Brovelli, P. (2014). Radar-based products for nowacasting in france. Presented at the Rain-Gain international workshop, Antwerp, Belgium.

Cuccoli, F., Facheris, L., and Gori, S. (2009). Radio base network and tomographic processing for real time estimation of the rainfall rate fields. In *Geoscience and Remote Sensing Symposium, 2009 IEEE International, IGARSS 2009*, volume 3, pages III–121. IEEE.

Daget, N. (2007). Revue des méthodes d'assimilation.

Delahaye, J.-Y., Barthès, L., Golé, P., Lavergnat, J., and Vinson, J.-P. (2006). A dual-beam spectropluviometer concept. *Journal of Hydrology*, 328(1):110–120.

Delanoë, J., Vinson, J.-P., Brett, W., Bertrand, F., Caudoux, C., Protat, A., Haeffelin, M., Du Chatelet, J. P., et al. (2014). Basta: a 95ghz fm–cw cloud radar. In *Journée Scientifique SIRTA 2014*.

Desa, M. and Niemczynowicz, J. (1997). Dynamics of short rainfall storms in a small scale urban area in coly limper, malaysia. *Atmospheric research*, 44(3):293–315.

Doumounia, A., Gosset, M., Cazenave, F., Kacou, M., and Zougmore, F. (2014). Rainfall monitoring based on microwave links from cellular telecommunication networks: First results from a west african test bed. *Geophysical Research Letters*, 41(16):6016–6022.

Doviak, R. J. and Zrnic, D. S. (2014). *Doppler Radar & Weather Observations*. Academic press.

Feingold, G., Tzivion, S., and Leviv, Z. (1988). Evolution of raindrop spectra. part i: Solution to the stochastic collection/breakup equation using the method of moments. *Journal of the atmospheric sciences*, 45(22):3387–3399.

Flossmann, A. I. and Wobrock, W. (2010). A review of our understanding of the aerosol–cloud interaction from the perspective of a bin resolved cloud scale modelling. *Atmospheric Research*, 97(4):478–497.

Foote, G. B. and Du Toit, P. S. (1969). Terminal velocity of raindrops aloft. *Journal of Applied Meteorology*, 8(2):249–253.

Gates, D. M. (1980). Biophysical ecology, 611 pp.

Giangrande, S. E., Luke, E. P., and Kollias, P. (2010). Automated retrievals of precipitation parameters using non-rayleigh scattering at 95 ghz. *Journal of Atmospheric and Oceanic Technology*, 27(9):1490–1503.

Giangrande, S. E., Luke, E. P., and Kollias, P. (2012). Characterization of vertical velocity and drop size distribution parameters in widespread precipitation at arm facilities. *Journal of Applied Meteorology and Climatology*, 51(2):380–391.

Gilbert, J. C. and Lemaréchal, C. (2006). The module m1qn3. *INRIA Rocquencourt & Rhone-Alpes*.

Giuli, D., Facheris, L., and Tanelli, S. (1999). Microwave tomographic inversion technique based on stochastic approach for rainfall fields monitoring. *Geoscience and Remote Sensing, IEEE Transactions on*, 37(5):2536–2555.

Giuli, D., Toccafondi, A., Biffi Gentili, G., and Freni, A. (1991). Tomographic reconstruction of rainfall fields through microwave attenuation measurements. *Journal of applied meteorology*, 30(9):1323–1340.

Gossard, E. E. (1988). Measuring drop-size distributions in clouds with a clear-air-sensing doppler radar. *Journal of Atmospheric and Oceanic Technology*, 5(5):640–649.

Gossard, E. E. (1994). Measurement of cloud droplet size spectra by doppler radar. *Journal of Atmospheric and Oceanic Technology*, 11(3):712–726.

Gunn, R. and Kinzer, G. D. (1949). THE TERMINAL VELOCITY OF FALL FOR WATER DROPLETS IN STAGNANT AIR. *J. Meteor.*, 6(4):243–248.

Gupta, V. K. and Waymire, E. (1987). On taylor's hypothesis and dissipation in rainfall. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 92(D8):9657–9660.

Hall, W. D. (1980). A detailed microphysical model within a two-dimensional dynamic framework: Model description and preliminary results. *Journal of the Atmospheric Sciences*, 37(11):2486–2507.

Hauser, D. and Amayenc, P. (1981). A new method for deducing hydrometeor-size distributions and vertical air motions from doppler radar measurements at vertical incidence. *Journal of Applied Meteorology*, 20(5):547–555.

Hu, Z. and Srivastava, R. C. (1995). Evolution of raindrop size distribution by coalescence, breakup, and evaporation: Theory and observations. *Journal of the atmospheric sciences*, 52(10):1761–1783.

Ide, K., Courtier, P., Ghil, M., and Lorenc, A. (1997). Uni
ed notation for data assimilation: Operational, sequential and variational.

Jalaal, M. and Mehravaran, K. (2012). Fragmentation of falling liquid droplets in bag breakup mode. *International Journal of Multiphase Flow*, 47:115–132.

Jarraud, M. (2008). Guide to meteorological instruments and methods of observation (wmo-no. 8). *World Meteorological Organisation: Geneva, Switzerland*.

Jinno, K., Kawamura, A., Berndtsson, R., Larson, M., and Niemczynowicz, J. (1993). Real-time rainfall prediction at small space-time scales using a two-dimensional stochastic advection-diffusion model. *Water resources research*, 29(5):1489–1504.

Joss, J. and Waldvogel, A. (1967). Ein spektrograph für niederschlagstropfen mit automatischer auswertung. *Pure and Applied Geophysics*, 68(1):240–246.

Kalnay, E. (2003). *Atmospheric modeling, data assimilation, and predictability*. Cambridge university press.

Kawamura, A., Jinno, K., Berndtsson, R., and Furukawa, T. (1997). Real-time tracking of convective rainfall properties using a two-dimensional advection-diffusion model. *Journal of Hydrology*, 203(1):109–118.

Ko, G. H. and Ryou, H. S. (2005). Modeling of droplet collision-induced breakup process. *International journal of multiphase flow*, 31(6):723–738.

Kollias, P., Albrecht, B., and Marks, F. (2001). Raindrop sorting induced by vertical drafts in convective clouds. *Geophysical research letters*, 28(14):2787–2790.

Kollias, P., Albrecht, B., and Marks Jr, F. (2002). Why mie? accurate observations of vertical air velocities and raindrops using a cloud radar. *Bulletin of the American Meteorological Society*, 83(10):1471–1483.

Kollias, P., Bharadwaj, N., Widener, K., Jo, I., and Johnson, K. (2014). Scanning arm cloud radars. part i: Operational sampling strategies. *Journal of Atmospheric and Oceanic Technology*, 31(3):569–582.

Kollias, P., Clothiaux, E., Miller, M., Albrecht, B., Stephens, G., and Ackerman, T. (2007). Millimeter-wavelength radars: new frontier in atmospheric cloud and precipitation research, b. am. meteorol. soc., 88, 1608–1624, doi: 10.1175. Technical report, BAMS-88-10-1608.

Kostinski, A. B. and Shaw, R. A. (2009). Droplet dynamics: Raindrops large and small. *Nature Physics*, 5(9):624–625.

Krajewski, W. F., Kruger, A., Caracciolo, C., Golé, P., Barthes, L., Creutin, J.-D., Delahaye, J.-Y., Nikolopoulos, E. I., Ogden, F., and Vinson, J.-P. (2006). Devex-disdrometer evaluation experiment: Basic results and implications for hydrologic studies. *Advances in water resources*, 29(2):311–325.

Kruger, A. and Krajewski, W. F. (2002). Two-dimensional video disdrometer: A description. *Journal of Atmospheric and Oceanic Technology*, 19(5):602–617.

Larsen, M., Kostinski, A., and Jameson, A. (2014). Further evidence for superterminal raindrops. *Geophysical Research Letters*, 41(19):6914–6918.

Lavergnat, J. and Golé, P. (1998). A stochastic raindrop time distribution model. *Journal of Applied Meteorology*, 37(8):805–818.

Laws, J. O. and Parsons, D. A. (1943). The relation of raindrop-size to intensity. *Eos, Transactions American Geophysical Union*, 24(2):452–460.

Leijnse, H., Uijlenhoet, R., and Stricker, J. (2007). Rainfall measurement using radio links from cellular communication networks. *Water Resources Research*, 43(3).

Leijnse, H., Uijlenhoet, R., and Stricker, J. (2008). Microwave link rainfall estimation: Effects of link length and frequency, temporal sampling, power resolution, and wet antenna attenuation. *Advances in Water Resources*, 31(11):1481–1493.

Lemaréchal, C. and Panier, E. (2000). Les modules m2qn1 et mqhess.

Leroy, D. (2007). *Développement d'un modèle de nuage tridimensionnel à microphysique détaillée-Application à la simulation de cas de convection profonde*. PhD thesis, Université Blaise Pascal-Clermont-Ferrand II.

Leroy, M. (2000). Estimation de l'incertitude de mesure des précipitations. *Note interne Meteo-France/DSO*, 42.

Levin, Z. and Machnes, B. (1977). Experimental evaluation of the coalescence efficiencies of colliding water drops. *pure and applied geophysics*, 115(4):845–867.

Lhermitte, R. M. (1988). Observation of rain at vertical incidence with a 94 ghz doppler radar: An insight on mie scattering. *Geophysical research letters*, 15(10):1125–1128.

Liebe, H., Hufford, G., and Cotton, M. (1993). Propagation modeling of moist air and suspended water/ice particles at frequencies below 1000 ghz. In *In AGARD, Atmospheric Propagation Effects Through Natural and Man-Made Obscurants for Visible to MM-Wave Radiation 11 p (SEE N94-30495 08-32)*, volume 1.

Lin, C. L. and Lee, S. C. (1975). Collision efficiency of water drops in the atmosphere. *Journal of the Atmospheric Sciences*, 32(7):1412–1418.

List, R. (1988). A linear radar reflectivity-rainrate relationship for steady tropical rain. *Journal of the atmospheric sciences*, 45(23):3564–3572.

List, R. and McFarquhar, G. M. (1990). The role of breakup and coalescence in the three-peak equilibrium distribution of raindrops. *Journal of the atmospheric sciences*, 47(19):2274–2292.

Lotka, A. J. et al. (1925). Elements of physical biology.

Lovejoy, S. and Schertzer, D. (1990). Multifractals, universality classes and satellite and radar measurements of cloud and rain fields. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 95(D3):2021–2034.

Low, T. B. and List, R. (1982a). Collision, coalescence and breakup of raindrops. part i: Experimentally established coalescence efficiencies and fragment size distributions in breakup. *Journal of the Atmospheric Sciences*, 39(7):1591–1606.

Low, T. B. and List, R. (1982b). Collision, coalescence and breakup of raindrops. part II: Parameterization of fragment size distributions. *Journal of the Atmospheric Sciences*, 39(7):1607–1619.

Maitra, A. and Chakravarty, K. (2005). Ku-band rain attenuation observations on an earth–space path in the indian region. *Proc. URSI Gen. Assembly F*, 1(3).

Marshall, J., Langille, R., and Palmer, W. M. K. (1947). Measurement of rainfall by radar. *Journal of Meteorology*, 4(6):186–192.

Marshall, J. S. and Palmer, W. M. K. (1948). The distribution of raindrops with size. *Journal of meteorology*, 5(4):165–166.

McFarquhar, G. M. (2004). A new representation of collision-induced breakup of raindrops and its implications for the shapes of raindrop size distributions. *Journal of the atmospheric sciences*, 61(7):777–794.

Mcfarquhar, G. M. (2010). Raindrop size distribution and evolution. In Testik, F. Y. and Gebremichael, M., editors, *Rainfall: State of the Science*, pages 49–60. American Geophysical Union.

McTaggart-Cowan, J. D. and List, R. (1975). Collision and breakup of water drops at terminal velocity. *J. Atmos. Sci.*, 32(7):1401–1411.

Mercier, F., Barthès, L., and Mallet, C. (2015a). Estimation of finescale rainfall fields using broadcast tv satellite links and a 4dvar assimilation method. *Journal of Atmospheric and Oceanic Technology*, 32(10):1709–1728.

Mercier, F., Chazottes, A., Barthès, L., and Mallet, C. (2015b). 4-d-var assimilation of disdrometer data and radar spectral reflectivities for rain drop size distribution and vertical wind retrievals. *Atmospheric Measurement Techniques Discussions*, 8(11):12383–12431.

METEK (1998). *MMR, Physical Basis*. Available from METEK GmbH, Fritz-Straßmann-Straße 4, D-25337 Elmshorn, Germany.

Molinié, G., Ceresetti, D., Anquetin, S., Creutin, J. D., and Boudevillain, B. (2012). Rainfall regime of a mountainous mediterranean region: Statistical analysis at short time steps. *Journal of Applied Meteorology and Climatology*, 51(3):429–448.

Montero-Martínez, G., Kostinski, A. B., Shaw, R. A., and García-García, F. (2009). Do all raindrops fall at terminal speed? *Geophysical Research Letters*, 36(11).

Munnannur, A. and Reitz, R. D. (2007). A new predictive model for fragmenting and non-fragmenting binary droplet collisions. *International journal of multiphase flow*, 33(8):873–896.

Nardi, L., Sorror, C., Badran, F., and Thiria, S. (2009). Yao: a software for variational data assimilation using numerical models. In *Computational Science and Its Applications–ICCSA 2009*, pages 621–636. Springer.

Orme, M. (1997). Experiments on droplet collisions, bounce, coalescence and disruption. *Progress in Energy and Combustion Science*, 23(1):65–79.

Overeem, A., Leijnse, H., and Uijlenhoet, R. (2013). Country-wide rainfall maps from cellular communication networks. *Proceedings of the National Academy of Sciences*, 110(8):2741–2745.

Overeem, A., Leijnse, H., and Uijlenhoet, R. (2015). Retrieval algorithm for rainfall mapping from microwave links in a cellular communication network. *Atmospheric Measurement Techniques Discussions*, 8(8).

Parent-du Châtelet, J. (2003). Aramis, le réseau français de radars pour la surveillance des précipitations.

Peters, G., Fischer, B., Münster, H., Clemens, M., and Wagner, A. (2005). Profiles of raindrop size distributions as retrieved by microrain radars. *Journal of Applied Meteorology*, 44(12):1930–1949.

Pinsky, M., Khain, A., and Shapiro, M. (2001). Collision efficiency of drops in a wide range of reynolds numbers: Effects of pressure on spectrum evolution. *Journal of the atmospheric sciences*, 58(7):742–764.

Prat, O. P. and Barros, A. P. (2007). A robust numerical solution of the stochastic collection-breakup equation for warm rain. *Journal of Applied Meteorology and Climatology*, 46(9):1480–1497.

Prat, O. P., Barros, A. P., and others (2007). Exploring the use of a column model for the characterization of microphysical processes in warm rain: Results from a homogeneous rainshaft model. *Advances in Geosciences*, 10:145–152.

Prat, O. P., Barros, A. P., and Testik, F. Y. (2012). On the influence of raindrop collision outcomes on equilibrium drop size distributions. *J. Atmos. Sci.*, 69(5):1534–1546.

Reich, S. and Cotter, C. (2015). *Probabilistic Forecasting and Bayesian Data Assimilation*. Cambridge University Press.

Rogers, R. (1964). An extension of the zr relation for doppler radar. In *11th Weather Radar Conference*.

Rysman, J.-F. (2013). *Caractérisations multi-échelles de la pluie et des processus associés dans l'Eurorégion Alpes-Méditerranée: de l'observation radar à la prévision*. PhD thesis, Paris 6.

Sauvageot, H. (1982). *Radarmétéorologie*. Eyrolles.

Sauvageot, H. and Koffi, M. (2000). Multimodal raindrop size distributions. *Journal of the atmospheric sciences*, 57(15):2480–2492.

Schleiss, M. and Berne, A. (2010). Identification of dry and rainy periods using telecommunication microwave links. *Geoscience and Remote Sensing Letters, IEEE*, 7(3):611–615.

Schlottke, J., Straub, W., Beheng, K. D., Gomaa, H., and Weigand, B. (2010). Numerical investigation of collision-induced breakup of raindrops. part i: Methodology and dependencies on collision energy and eccentricity. *J. Atmos. Sci.*, 67(3):557–575.

Seifert, A. (2008). On the parameterization of evaporation of raindrops as simulated by a one-dimensional rainshaft model. *Journal of the Atmospheric Sciences*, 65(11).

Seifert, A., Khain, A., Blahak, U., and Beheng, K. D. (2005). Possible effects of collisional breakup on mixed-phase deep convection simulated by a spectral (bin) cloud model. *Journal of the atmospheric sciences*, 62(6):1917–1931.

Seinfeld, J. H. and Pandis, S. N. (2012). *Atmospheric chemistry and physics: from air pollution to climate change*. John Wiley & Sons.

Smolarkiewicz, P. K. (1983). A simple positive definite advection scheme with small implicit diffusion. *Monthly Weather Review*, 111(3):479–486.

Srivastava, R. (1971). Size distribution of raindrops generated by their breakup and coalescence. *Journal of the Atmospheric Sciences*, 28(3):410–415.

Straub, W., Beheng, K. D., Seifert, A., Schlottke, J., and Weigand, B. (2010). Numerical investigation of collision-induced breakup of raindrops. part II: Parameterizations of coalescence efficiencies and fragment size distributions. *Journal of the Atmospheric Sciences*, 67(3):576–588.

Sutherland, W. (1893). Lii. the viscosity of gases and molecular force. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 36(223):507–531.

Tabary, P. (2007). The new french operational radar rainfall product. part i: Methodology. *Weather and forecasting*, 22(3):393–408.

Testik, F. Y. (2009). Outcome regimes of binary raindrop collisions. *Atmospheric Research*, 94(3):389–399.

Testik, F. Y., Barros, A. P., and Bliven, L. F. (2011). Toward a physical characterization of raindrop collision outcome regimes. *Journal of the Atmospheric Sciences*, 68(5):1097–1113.

Testud, J., Oury, S., Black, R. A., Amayenc, P., and Dou, X. (2001). The concept of "normalized" distribution to describe raindrop spectra: A tool for cloud physics and cloud remote sensing. *Journal of Applied Meteorology*, 40(6):1118–1140.

Thiria, S., Badran, F., and Sorror, C. (2006). Yao: Un logiciel pour les modèles numériques et l'assimilation de données. *Rapport technique, LOCEAN, France*.

Tokay, A., Bashor, P. G., and Wolff, K. R. (2005). Error characteristics of rainfall measurements by collocated joss-waldvogel disdrometers. *Journal of Atmospheric and Oceanic Technology*, 22(5):513–527.

Tokay, A., Kruger, A., and Krajewski, W. F. (2001). Comparison of drop size distribution measurements by impact and optical disdrometers. *Journal of Applied Meteorology*, 40(11):2083–2097.

Tokay, A., Wolff, D. B., Wolff, K. R., and Bashor, P. (2003). Rain gauge and disdrometer measurements during the keys area microphysics project (kamp). *Journal of Atmospheric and Oceanic Technology*, 20(11):1460–1477.

Tran, V. D. (2009). *Assimilation de données: les propriétés asymptotiques du filtre de Kalman d'ensemble*. PhD thesis, Université de Bretagne Sud.

Tridon, F. (2011). *Mesure des précipitations à l'aide d'un radar en bande X non-cohérent à haute résolution et d'un radar en bande K à visée verticale. Application à l'étude de la variabilité des précipitations lors de la campagne COPS*. PhD thesis, Université Blaise Pascal-Clermont-Ferrand II.

Tridon, F. and Battaglia, A. (2015). Dual-frequency radar doppler spectral retrieval of rain drop size distributions and entangled dynamics variables. *Journal of Geophysical Research: Atmospheres*.

Tridon, F., Battaglia, A., and Kollias, P. (2013). Disentangling mie and attenuation effects in rain using a ka-w dual-wavelength doppler spectral ratio technique. *Geophysical Research Letters*, 40(20):5548–5552.

Ulaby, F. T., Moore, R. K., and Fung, A. K. (1981). Microwave remote sensing: Active and passive. volume 1-microwave remote sensing fundamentals and radiometry.

Ulbrich, C. W. (1983). Natural variations in the analytical form of the raindrop size distribution. *Journal of climate and applied meteorology*, 22(10):1764–1775.

Valdez, M. P. and Young, K. C. (1985). Number fluxes in equilibrium raindrop populations: A markov chain analysis. *Journal of the atmospheric sciences*, 42(10):1024–1036.

VanZanten, M., Stevens, B., Vali, G., and Lenschow, D. (2005). Observations of drizzle in nocturnal marine stratocumulus. *Journal of the atmospheric sciences*, 62(1):88–106.

Verrier, S., De Montera, L., Barthès, L., and Mallet, C. (2010). Multifractal analysis of african monsoon rain fields, taking into account the zero rain-rate problem. *Journal of hydrology*, 389(1):111–120.

Villermaux, E. and Bossa, B. (2009). Single-drop fragmentation determines size distribution of raindrops. *Nature Physics*, 5(9):697–702.

Villermaux, E. and Eloi, F. (2011). The distribution of raindrops speeds. *Geophysical Research Letters*, 38(19).

Volterra, V. (1926). Fluctuations in the abundance of a species considered mathematically. *Nature*, 118:558–560.

Whelpdale, D. and List, R. (1971). The coalescence process in raindrop growth. *Journal of Geophysical Research*, 76(12):2836–2856.

Widener, K., Bharadwaj, N., and Johnson, K. (2012). Ka-band arm zenith radar (kazr) instrument handbook. Technical report, DOE Office of Science Atmospheric Radiation Measurement (ARM) Program (United States).

Widener, K. and Mead, J. (2004). W-band arm cloud radar–specifications and design. In *Proc. 14th ARM Science Team Meeting*.

Williams, C. R. (2002). Simultaneous ambient air motion and raindrop size distributions retrieved from uhf vertical incident profiler observations. *Radio Science*, 37(2).

Williams, C. R. (2012). Vertical air motion retrieved from dual-frequency profiler observations. *Journal of Atmospheric and Oceanic Technology*, 29(10):1471–1480.

Williams, C. R., Kruger, A., Gage, K. S., Tokay, A., Cifelli, R., Krajewski, W. F., and Kummerow, C. (2000). Comparison of simultaneous rain drop size distributions estimated from two surface disdrometers and a uhf profiler. *Geophysical Research Letters*, 27(12):1763–1766.

Wilson, J. W., Crook, N. A., Mueller, C. K., Sun, J., and Dixon, M. (1998). Nowcasting thunderstorms: A status report. *Bulletin of the American Meteorological Society*, 79(10):2079–2099.

Zawadzki, I. (1973). Statistical properties of precipitation patterns. *Journal of Applied Meteorology*, 12(3):459–472.

Zinevich, A., Alpert, P., and Messer, H. (2008). Estimation of rainfall fields using commercial microwave communication networks of variable density. *Advances in Water Resources*, 31(11):1470–1480.

Zinevich, A., Messer, H., and Alpert, P. (2009). Frontal rainfall observation by a commercial microwave communication network. *Journal of Applied Meteorology and Climatology*, 48(7):1317–1334.

Zohidov, B., Andrieu, H., Servières, M., and Normand, N. (2014). Rainfall retrieval in urban areas using commercial microwave links from mobile networks: A modelling feasibility study. In *EGU General Assembly Conference Abstracts*, volume 16, page 11911.

Zwiebel, J. (2015). *Etude de l'impact orographique sur la structure microphysique horizontale et verticale des précipitations*. PhD thesis, Université Blaise-Pascal.

Table des figures

I.1 Photographies des trois disdromètres utilisés dans ce travail. a) disdromètre Joss-Walvogel, de la société Distromet Ltd. (http://distromet.com/), image provenant du site de l'Atmospheric Measurement Research (https://www.arm.gov/instruments/disdrometer) ; b) disdromètre 2D vidéo, même source (http://www.arm.gov/instruments/vdis) ; c) spectropluviomètre du LATMOS. 11

I.2 Figure extraite de Kruger and Krajewski (2002) illustrant le principe de fonctionnement du disdromètre 2D-vidéo. 12

I.3 Figure extraite de Delahaye et al. (2006) illustrant le principe de fonctionnement du spectropluviomètre. Voir le texte pour les commentaires. . 13

I.4 Volume sondé par un disdromètre. Schéma représentant un plan altitude/dimension horizontale. Le disdromètre est symbolisé par le trait épais noir au sol, sa surface de collecte est notée S. On considère deux classes de gouttes, des petites tombant à une vitesse constante v_1 et des grosses tombant à une vitesse constante $v_2 > v_1$. Toutes les gouttes ont aussi une même vitesse horizontale V_h parallèle à l'axe x (correspondant à un vent horizontal). Pendant la durée Δt suivant cette image, le disdromètre enregistre les gouttes provenant du cadre vert (gouttes lentes) ou rouge (gouttes rapides). On remarque qu'en augmentant le temps d'intégration pour les petites gouttes (cadre vert pointillé), on peut obtenir la même hauteur sondée ($v_2 \Delta t$), mais pas les volumes (à moins d'avoir $V_h = 0$). 15

I.5 *Haut* : signal émis par un radar pulsé. On représente la puissance émise P_e en fonction du temps. Il s'agit d'une fonction créneau. Le temps d'impulsion est τ_{on}, le temps entre 2 impulsions τ_{off}. *Milieu* : Représentation schématique du signal renvoyé au radar par des cibles réparties dans la zone visée par le radar (typiquement des gouttes d'eau). Il s'agit de la puissance reçue P. On discrétise ensuite l'axe temporel (dates successives t_i entre 2 impulsions), et on estime la puissance moyenne reçue par le radar sur chaque période $[t_k : t_{k+1}]$. Chaque fenêtre temporelle correspond à une porte radar (zone de l'espace dans laquelle sont comprises les cibles provoquant l'écho enregistré). On définit ici la distance du milieu de chaque porte par $r_k = c\,(t_{k+1} + t_k)\,/2 - ct_0$ (c la vitesse de la lumière). *Bas* : Visualisation de la puissance reçue en fonction de la porte (distance au radar). 16

I.6 Définition de l'ouverture de faisceau radar θ_0. On considère un radar qui émet une onde. On représente ici la densité de puissance émise en fonction de la direction. La densité de puissance émise est maximale ($p_{e,max}$) dans une direction donnée. On définit ensuite l'angle θ_0 comme celui englobant toutes les directions dans lesquelles le radar émet une densité de puissance supérieure à la moitié de sa valeur maximale (d'où l'expression "ouverture à $-3dB$" pour désigner θ_0). On note que l'on suppose ici que l'émission est la même dans tous les plans contenant la direction d'émission maximale (on devrait formellement, en coordonnées sphériques, définir deux angles θ_0 et ϕ_0). 17

I.7 Visualisation du volume (noté V) sondé par un radar, pour une porte et une ouverture de faisceau données. On représente ici un radar à visée verticale (maximum de densité de puissance émis à la verticale). La distance r caractérisant la porte du radar considérée est définie sur la Figure I.5. L'ouverture du faisceau θ_0 est définie sur la Figure I.6. 18

I.8 Pour différentes longueurs d'onde, représentation des surfaces efficaces de rétrodiffusion de particules sphériques en fonction de leur diamètre, calculées dans le cadre de la théorie de Mie. Les fréquences utilisées sont : $6GHz$ (bande C), $10GHz$ (bande X), $25GHz$ (bande K), $35GHz$ (bande Ka), $90GHz$ (bande W). 20

I.9 Selon la fréquence d'éclairage, atténuation spécifique des ondes due à l'oxygène atmosphérique (bleu) et à la vapeur d'eau pour différentes humidités relatives (rouge), à $1013HPa$ et $20\,^{\circ}C$. Figure adaptée de Liebe et al. (1993) et Brisseau (2005). 22

I.10 a) Influence du diamètre et de la fréquence sur les surfaces efficaces d'extinction. b) Atténuation spécifique produite par une DSD de Marshall-Palmer (avec un taux de pluie de $20mm/h$) selon la fréquence. c) Atténuation totale produite par cette même DSD selon l'altitude et la fréquence. 23

I.11 DSD simulée (haut) et réflectivités spectrales prenant en compte divers phénomènes. Voir texte pour la description détaillée. 26

I.12 Le radar opérationnel en bande S de Météo France de Bollène. Photographie issue du site de l'institut (http://education.meteofrance.fr/la-mesure-des-precipitations-par-radar 28

I.13 Photographies des 4 radars à visée verticale utilisés dans ce travail. a) le Micro-Rain Radar (extrait de METEK (1998)). b) le KAZR (extrait de Widener et al. (2012)). c) le radar BASTA (extrait du site du constructeur : http://www.meteomodem.com/basta.php). d) le WACR. 29

I.14 Représentation des différents satellites de télévision en orbite géostationnaire, colorés sur leur bande de fréquence d'émission. Figure extraite de http://fr.questmachine.org/Applications_des_satellites. 32

I.15 Schéma présentant le fonctionnement du capteur Ku. Une onde est émise en bande Ku par un satellite géostationnaire de télévision ("Satellite TV"). A cette fréquence, elle est atténuée essentiellement par les gouttes de pluie, donc en dessous de la bande brillante et du nuage. Un récepteur situé au sol mesure la puissance reçue, ce qui, connaissant la puissance émise, permet d'estimer l'atténuation. Projeté au sol et pour des conditions typiques en France en été, on obtient ainsi des mesures d'atténuations linéiques sur des segments de quelques kilomètres de longueur. β désigne l'élévation, angle selon lequel le satellite est vu depuis le sol. 33

II.1 Produit lame d'eau (champ de taux de pluie à 5 minutes, en mm/h) issu du radar Météo France de Bollène, pour le 20 octobre 2013 entre 6:35 et 7:00 autour de Mirabel, Ardèche. 37

II.2 a) Taux d'évaporation massique selon l'altitude et le diamètre des gouttes sphériques. Calculé selon la paramétrisation proposée par Seifert (2008). Le taux d'humidité est homogène à 80%, la température est de $27\,^{\circ}C$ au sol, avec un gradient de $-1\,^{\circ}/150m$. L'unité est ici (pour des questions de lisibilité) le milligramme par heure. b) Impact selon l'altitude sur une DSD d'un tel taux d'évaporation. On propage verticalement, sans vent ni collisions, une DSD de Marshall and Palmer (1948) (pour un taux de pluie de $6mm/h$). La hauteur de chute est de $3000m$. 44

II.3 Impact de l'évaporation sur 3 paramètres issus de la DSD, selon l'altitude. On fait ici tourner un modèle de dynamique des DSD dans les mêmes conditions que sur la Figure II.2 (ni vent ni collisions, évaporation de Seifert (2008), DSD au sommet de la colonne de Marshall and Palmer (1948) pour un taux de pluie de $6mm/h$). Expérience réalisée pour 3 niveaux d'humidité : 80% (bleu), 90% (rouge) et 95% (noir). On représente : a) le diamètre volumique moyen D_m ($\mathcal{M}_4(n)/\mathcal{M}_3(n)$, en mm) ; b) le nombre de gouttes N_{tot} ($\mathcal{M}_0(n)$, en m^{-3}) ; c) le taux de pluie RR (proportionnel à $\int n(D)v_{atlas}(D)D^3dD$, en mm/h). 45

II.4 Évolutions possibles pour deux gouttes de masses m_1 et m_2 susceptibles de se percuter. Voir texte pour les définitions des noyaux de collision Φ_{coll}, des efficacités de coalescence et de break-up E_{coal} et E_{bu}, et de la fonction de break-up χ_{bu}. 46

II.5 Croquis représentant les résultats possibles d'une collision entre deux gouttes, à savoir : a) coalescence ; b) rebond ; c) "disruption" (séparation à nouveau en deux gouttes) ; d) fragmentation. Figure adaptée de Orme (1997). Définition du break-up à partir de ces types d'impacts. 47

II.6 Figure extraite de Mcfarquhar (2010) présentant différentes distributions stationnaires. Distributions stationnaires obtenues pour un noyau de coalescence de Low and List (1982a) (traits continues et points) ou de Straub et al. (2010) (tirets) et avec une fonction de break-up de Low and List (1982b) (trait continu), de McFarquhar (2004) (points), de Straub et al. (2010) (tirets). Dans tous les cas, la condition aux limites en haut est une DSD de Marshall and Palmer (1948) pour un taux de pluie de $54mm/h$. . . . 50

II.7 Figure extraite de Prat and Barros (2007). Distributions aux limites (marques) et stationnaires (pointillés) obtenues dans le cadre de la paramétrisation de Low and List (1982a) / Low and List (1982b) (efficacité de coalescence et fonction de break-up) pour différentes distributions aux limites. Celles-ci sont des distributions de Marshall and Palmer (1948) pour des taux de pluie de $50mm/h$ ("MP50"), $15mm/h$ ("MP15") et $5mm/h$ ("MP5"). 51

II.8 Figure extraite de Prat and Barros (2007). Evolution temporelle des moments d'ordre 0 et 1 de la DSD lors de la chute des gouttes, dans le cadre d'une efficacité de coalescence et d'une fonction de break-up de Low and List (1982a) et Low and List (1982b). Les conditions aux limites sont des distributions de Marshall and Palmer (1948) pour des taux de pluie de $50mm/h$ ("MP50"), $15mm/h$ ("MP15") et $5mm/h$ ("MP5"). 51

II.9 Représentation de la grille C d'Arakawa en une dimension. Le champ scalaire advecté ($\tilde{\phi}$) est évalué au centre des mailles (c'est-à-dire que l'on considère sa valeur homogène dans tout le point de maille) ; les vitesses d'advection ($\tilde{v}$) sont évaluées aux bords des mailles (permettant de définir naturellement les flux d'une maille à l'autre). 55

II.10 Expériences d'advection numérique avec les schémas upstream et de Smolarkiewicz. La résolution spatiale est de $0.4 * 0.4km^2$, la résolution temporelle de $1s$. La vitesse d'advection est de $13.6m/s$ vers le sud et $10.4m/s$ vers l'est. a) champ initial. b) champ initial advecté sur $5min$ par le schéma upstream. c) champ initial advecté sur $10min$ par le schéma upstream. d) champ initial. e) champ initial advecté sur $5min$ par le schéma de Smolarkiewicz. f) champ initial advecté sur $10min$ par le schéma de Smolarkiewicz. 58

III.1 Schéma du résultat d'une expérience d'assimilation 4D-VAR. 69

III.2 Schéma représentant le fonctionnement de l'algorithme d'assimilation 4D-VAR. 70

III.3 Détail des étapes d'une itération de l'algorithme 4D-VAR appliqué au problème proies/prédateurs. 75

III.4 Illustration rapide du concept de graphe modulaire. Figure adaptée de Nardi et al. (2009), fig.3. 77

III.5 Schéma représentant la discrétisation du plan temps/altitude sur lequel sont définis des états du système dynamique. La direction de propagation de ce système est indiquée (propagation temporelle et du haut vers le bas), ainsi que les paramètres de contrôle de l'algorithme 4D-VAR associé, à savoir les conditions initiales (en jaune) et les conditions limites (en rouge). 81

III.6 Exemple arbitraire d'une fonction de pénalisation d'un paramètre x_0. Ici, on veut contraindre x_0 dans l'intervalle $[\alpha : \beta]$, et le coût associé est $\mathcal{C}^1$. 83

III.7 Exemple de problème justifiant l'utilisation d'un filtre. Voir texte pour les commentaires. 84

IV.1 Dispositif expérimental autour de Mirabel, Ardèche. Le capteur Ku est positionné dans le village de Mirabel, et capte 4 liaisons satellites (caractéristiques dans le Tableau IV.1). On représente ici ces liaisons en-dessous de l'isotherm 0 ° (là où la pluie peut atténuer le signal) avec un iso-0 vers $2800m$, et projetées au sol (traits bleus). On représente aussi la position des 10 pluviomètres du réseau HPiconet (triangles rouges). 90

IV.2 Pour la n-ième liaison Ku projetée au sol (définie arbitrairement pour cet exemple), définition de la longueur traversée par la liaison sur chaque point de maille d'une grille de discrétisation. ($L^n_{i,j}$ est la longueur traversée dans le pixel (i,j).) $L^n = \sum_{i,j} L^n_{i,j}$ est la longueur totale de la liaison (km). Appliquée, à la date t_k, au champ d'atténuations linéiques A_k (dB/km), la n-ième composante de l'opérateur d'observation fournira ici la quantité $L^n_{i,j}A_{k,i,j} + L^n_{i+1,j}A_{k,i+1,j} + L^n_{i+1,j+1}A_{k,i+1,j+1} + L^n_{i+1,j+2}A_{k,i+1,j+2}$. 94

IV.3 Description du processus de test de l'algorithme de restitution des champs de pluie sur données simulées (expérience jumelle). Voir texte pour la description. 98

IV.4 Définition des zones d'assimilation pour l'évaluation des performances de l'algorithme de restitution des cartes de pluie. Voir texte pour les commentaires. 100

IV.5 Pour chacune des 15 simulations de la section 5.2. A) Champs initiaux vrais (A_0^{vrai}) restreints à la grande zone d'assimilation. B) Champs initiaux assimilés (A_0^{assim}) dans le cadre d'une expérience d'assimilation avec 1 capteur Ku. C) Idem avec 2 capteurs. Dans tous les cas, le(s) capteur(s) Ku est (sont) situé(s) au nord-est de la grande zone d'assimilation. Les vitesses d'advection correspondant à chaque simulation sont données dans le Tableau IV.4. Simulations 1 à 5 sur cette Figure. 102

IV.6 Idem Figure IV.5 pour les simulations 6 à 10. 103

IV.7 Idem Figure IV.5 pour les simulations 11 à 15. 104

IV.8 Pour chaque évènement (1 à 8, lignes) et chaque liaison Ku (1 à 4, colonne) : observations du capteur Ku ($\mathcal{A}^{real}$, en dB, traits continus) et reproduction des observations par l'algorithme d'assimilation ($\mathcal{A}^{assim}$, traits pointillés, en dB). Évènements 1 à 4 sur cette Figure. 109

IV.9 Idem Figure IV.8, évènements 5 à 8. Unité : dB. 109

IV.10 Cartes radar successives pour 2 évènements de l'année 2013. A) Évènement peu advectif : 28/09, 14:40 à 15:20. A) Évènement davantage advectif : 20/10, 6:35 à 7:05. Les cartes présentées sont centrées sur la position du capteur Ku. 110

IV.11 Evènement 4 (28/09/2013). Champs de pluie à résolution $1km^2$ tel que restitués : A) par le radar Météo France de Bollène. B) par notre algorithme d'assimilation des mesures d'atténuation du capteur Ku. On restreint les champs à la grande zone d'assimilation définie par la Figure IV.4. Cartes toutes les 5 minutes sur tout l'évènement. 115

IV.12 Idem Figure IV.11 pour l'évènement 5 (29/09/2013). Cartes toutes les 10 minutes sur tout l'évènement. 115
IV.13 Idem Figure IV.11 pour l'évènement 7 (20/10/2013). Cartes toutes les 5 minutes sur tout l'évènement. 116
IV.14 Idem Figure IV.11 pour l'évènement 8 (20/10/2013). Cartes toutes les 5 minutes sur tout l'évènement. 116
IV.15 Boîtes à moustache des différences entre cartes radar et assimilées pour les 4 évènements advectifs à 1 ou $4km^2$ de résolution. On représente dans chaque cas la médiane (traits rouges), les quartiles (boîtes bleues) et les déciles (traits noirs). 117
IV.16 Cumuls d'eau pour les 4 évènements advectifs considérés pour le pluviomètre de Mirabel Pradel fermes (pluviomètre 5 du Tableau IV.2). Pour chaque évènement, on présente les mesures effectuées par le pluviomètre (bleu), et les données restituées sur le pixel lui correspondant par le radar de Bollène (vert) et par l'algorithme d'assimilation (rouge). 119
IV.17 Idem Figure IV.16 pour les pluviomètres de St Germain école (traits pleins) et de Villeneuve de Berg (pointillés) (respectivement pluviomètres 8 et 10 du Tableau IV.2). 120

V.1 Schéma représentant l'opérateur d'observation pour un disdromètre, décomposé en trois étapes. L'entrée de l'opérateur d'observation est la DSD en sortie du modèle de propagation (n). On transpose ensuite la DSD au niveau du sol ($i = NZ - 1$) en flux de gouttes (noté ici $\tilde{N}^1$), via la surface de collecte du disdromètre S_{dis}, le pas de temps du modèle Δt, le pas de diamètre ΔD et la vitesse terminale de chute des gouttes de diamètre D_p. On obtient ensuite $\tilde{N}^2$ (deuxième intermédiaire de calcul) en intégrant $\tilde{N}^1$ sur la durée T_{dis}, temps d'intégration du disdromètre. Pour obtenir la sortie N_{dis} de l'opérateur d'observation, on intègre $\tilde{N}^2$ sur les classes de diamètre du disdromètre. 136
V.2 Schéma résumant les opérations successives effectuées par l'opérateur d'observation des spectres Doppler. Ses entrées sont les DSD en sortie de modèle de propagation ($n_{k,i,p}$) et le vent vertical ($w_{k,i}$). Ses sorties sont les spectres Doppler simulés (logarithmiques) $\eta_{obs,k,i,pdop}$. Voir description détaillée dans le texte. 137

V.3 Schéma résumant l'algorithme d'assimilation 4D-VAR utilisé dans ce travail. Les inconnues de l'algorithme (ce que l'on veut restituer) sont encadrées en rouge. Il s'agit du vent vertical (w) à toutes dates et toutes altitudes et des trois coefficients d'une DSD gamma (α, k, θ) caractérisant les DSD en haut de la colonne atmosphérique à toutes dates. Ces DSD sont ensuite propagées à toutes dates et altitudes par le modèle numérique de propagation. Elles sont transformées en grandeurs comparables aux observations (cadres violets) par les opérateurs d'observation, que ce soit pour les mesures disdromètres (en bas) ou radars (à droite). Le coût est calculé en évaluant l'écart entre ces grandeurs et les observations effectivement disponibles (cadres verts). On ajoute à ce coût des termes de pénalisation et de régularisation des inconnues (haut de la figure). 142

V.4 Description du processus d'expériences jumelles pour la section 3. Voir texte pour la description. 144

V.5 Exemple de profil vertical de vent vertical (gauche) et de spectres Doppler (logarithmiques, droite). Pour le vent, on présente les données vraies (pointillés) et assimilées (continu). Pour les spectres, on présente les données vraies (étoiles), réalistes (incluant la turbulence, tirets) et assimilées (continus). Le vent est positif vers le bas. 152

V.6 DSD et taux de pluie vrais et assimilés. A) C) DSD vraies en fonction du temps en haut et en bas de la colonne atmosphérique. B) D) Idem assimilées. Unité : $ln(m^{-3}mm^{-1})$. E) Taux de pluie au sol vrais (bleu) et assimilés (rouge) selon la date. 153

V.7 Nombre de gouttes dans la gamme $0.2-0.4mm$ (haut) et $3-4mm$ (bas) selon la date dans la couche $900-1000m$. Données vraies (bleu) et assimilées (rouge). 155

V.8 Restitution du vent vertical (résolution : $2min$). A) Vent vertical vrai. B) Assimilé. C) Selon l'altitude, moyenne des vents verticaux vrais (bleu) et assimilés (rouge). D) Idem écart-type. E) Fonction temporelle d'autocorrélation pour les vents verticaux vrais (bleu) et assimilés (rouge). Vents positifs vers le bas. 156

V.9 Profils verticaux de vent horizontaux utilisés dans la section 4.2. Vent nul au sol, de $3m/s$ à l'altitude maximale ($z_0 = 2km$). Rouge : vent horizontal linéaire : $u(z) = 1.5z$. Noir : vent horizontal logarithmique : $u(z) = -4.03 + 0.92ln(80 + 1000z)$. ($z$ en km, u en m/s). 161

V.10 Pour deux diamètres de gouttes et trois profils de vent horizontal, position dans un plan (x, z) à une date donnée (T) de gouttes lâchées en continu par une cellule de pluie ponctuelle située à l'altitude $z_0 = 2km$. Traits continus : gouttes de diamètre $1mm$; pointillés : $6mm$. Bleu : vent horizontal nul (mais la cellule se déplace tout de même à la vitesse horizontale $u_{cell} = 3m/s$). Rouge : Vent horizontal linéaire (cf Fig. V.9). Noir : Vent horizontal logarithmique (cf Fig. V.9). On place le 0 de l'axe horizontal à la position de la cellule à la date $t = T$. 163

V.11 Simulation théorique simple de la vision qu'un radar en visée verticale aurait de l'arrivée de fronts de pluie, selon la taille des gouttes et le profil de vent horizontal. On représente, pour des dates successives, la position verticale des gouttes situées au-dessus d'une abscisse donnée, pour une cellule de pluie ponctuelle située à l'altitude $z_0 = 2km$ et lâchant continument des gouttes de diamètre fixe. Traits continus : gouttes de diamètre $1mm$; pointillés : $6mm$. Bleu : vent horizontal nul (mais la cellule se déplace tout de même à la vitesse horizontale $u_{cell} = 3m/s$). Rouge : Vent horizontal linéaire (cf Fig. V.9). Noir : Vent horizontal logarithmique (cf Fig. V.9). 164

V.12 Réflectivités (dBZ) sur des périodes d'environ $20min$ mesurées par le radar Doppler à visée verticale KAZR (voir chapitre I, section 2.3.3, Tableau I.2) en Finlande le 12/06/2014 (voir section 6 la description de l'évènement de pluie complet correspondant). 165

V.13 Description du processus d'expériences jumelles pour la section 4.3. Voir texte pour la description. 168

V.14 États vrais simulés nécessaires à l'exécution du modèle 4D de propagation des DSD. A) Taux de pluie en haut de la colonne atmosphérique à l'instant initial selon la position horizontale. B) Taux de pluie selon la date à la verticale des instruments de mesure, correspondant à la cellule du A propagée par le profil 1 de vent horizontal (G, rouge). C) Idem, profil 2 de vent horizontal (H, rouge). D) Cellules de vent vertical à l'instant initial selon l'altitude et la position horizontale. E) Vent vertical selon la date et l'altitude à la verticale des instruments de mesure, correspondant à la cellule du D propagée par le profil 1 de vent horizontal (G, rouge). Positif vers le bas. F) Idem, propagation par le profil 2 de vent horizontal (H, rouge). G) Vent horizontal selon l'altitude tel que : pointillés bleus : mesuré par radiosondage en Finlande le 12/06/2014, 11:36 ; pointillés noirs : estimé par fusion de radiosondages et de sorties de modèle à Darwin, Australie, le 04/11/2014, moyenné sur $17h$-$21h$; continu rouge : simulé dans le cadre de ce travail ("profil 1"). H) Idem G tel que : pointillés bleus : mesuré par radiosondage en Finlande le 12/06/2014, 17:23 ; continu rouge : simulé dans le cadre de ce travail ("profil 2"). Voir aussi le texte pour la description de ces champs. 170

V.15 Pour le profil 1 de vent horizontal : Nombre (logarithmique) de gouttes au-dessus du radar simulé, pour l'état vrai (simulé) et pour trois classes de diamètres : A) $0.9 - 1.0mm$; B) $1.8 - 1.9mm$; C) $4.1 - 4.2mm$. Formellement, on représente ici $n^{vrai}_{k,i,mr,p}$, $\forall k$ (temps), $\forall i$ (altitude), pour un mr arbitrairement fixé (position du radar) et pour trois p (classes de diamètres) différents. Unité : $10log_{10}\,(m^{-1}mm^{-1})$. 172

V.16 Pour le profil 1 de vent horizontal : facteur de réflectivité radar équivalent (proportionnel à la réflectivité η, moment d'ordre 0 de la réflectivité spectrale η_v, voir Eq.(I.18)) au-dessus du radar simulé. A) État vrai. B) État assimilé sans vent horizontal. Unité : dBZ. 173

V.17 Pour le profil 1 de vent horizontal : Vitesse Doppler moyenne (moment centré d'ordre 1 de la réflectivité spectrale η_v, voir Eq.(I.31)) au-dessus du radar simulé. A) État vrai. B) État assimilé sans vent horizontal. Unité : m/s, positif vers le bas. 173

V.18 Pour le profil 1 de vent horizontal : A-B) nombre de gouttes en fonction du temps et du diamètre vues par le disdromètre simulé : A) État vrai. B) État assimilé. C) Série temporelle des taux de pluie correspondant : bleu : état vrai ; rouge : état assimilé. 174

V.19 Pour le profil 2 de vent horizontal : idem Figure V.16. 174

V.20 Pour le profil 2 de vent horizontal : idem Figure V.17. 175

V.21 Comportement moyen durant l'évènement de plusieurs grandeurs vraies (bleu) et assimilées (rouge), pour les simulations avec profil de vent horizontal 1 (traits pointillés) ou 2 (traits pleins). On représente : A) le facteur de réflectivité moyen selon l'altitude. B) Le Doppler moyen selon l'altitude. C) Le vent vertical moyen selon l'altitude. D-E) La DSD moyenne respectivement en haut ($z = 2km$) et en bas de la colonne ($z = 0$). Les moyennes temporelles sont calculées sur le coeur des évènements, soit entre les heures 0.83 et 1.28 pour le profil de vent horizontal 1, et entre les heures 2.08 et 3.19 pour le 2. 176

V.22 Vent vertical au-dessus du radar simulé, états vrais et assimilés, pour les deux profils de vent. Profil de vent horizontal 1, vent vertical vrai (A) et assimilé (B). Profil de vent horizontal 2, vent vertical vrai (C) et assimilé (D). Sur les représentation des vents verticaux assimilés, les deux traits noirs délimitent la zone où la structure des champs vrais est retrouvée. Le vent est positif vers le bas. 177

V.23 Localisation de la zone d'étude dans le sud-est de la France sur une carte topographique (source IGN). Les deux instruments utilisés dans l'algorithme d'assimilation sont au Pradel. Les données de stations Météo France au sol sont disponibles au Pradel, à Aubenas et à Berzème. 180

V.24 Taux de pluie au sol tels qu'observés par le DBS (rouge) et restitué par l'algorithme d'assimilation VAMOS (bleu). 182
V.25 3 premiers moments observés et assimilés des spectres Doppler. Gauche : Observations MRR ; Droite : Assimilation (VAMOS). Haut : Facteur de réflectivité radar équivalent Z_e en dBZ (proportionnel à la réflectivité η, moment d'ordre 0 du spectre Doppler). Milieu : Vitesse Doppler moyenne W (espérance). Bas : Largeur spectrale LS (écart-type). 182
V.26 A) Pour toutes les altitudes, spectres Doppler observés (MRR, traits continus) et assimilés (VAMOS, traits pointillés) moyennés sur la période 20:00 - 20:23. On remarque que pour les deux portes les plus basses on ne dispose pas d'observations MRR. B) Part de la fonction de coût J après minimisation due à l'erreur commise sur les spectres Doppler pour chaque altitude (sur l'évènement entier). On rappelle que la fonction de coût comprend aussi un terme d'écart aux mesures disdromètres et des termes de filtrage / pénalisation. La somme des parts présentées ici ne fait donc pas 100% (mais 51%, voir texte). 186
V.27 A) Vent vertical w assimilé (moyenné sur $2min$) sur tout l'évènement. B) Moyenne selon l'altitude. C) Écart-type selon l'altitude. Vent positif vers le bas. 187
V.28 Comparaison entre produits MRR (gauche) et états assimilés VAMOS (droite) pour différents moments de la DSD. Haut : diamètre volumique moyen (D_m, moment d'ordre 4 sur moment d'ordre 3), mm. Milieu : Taux de pluie RR, mm/h. Bas : Facteur de réflectivité radar Z (moment d'ordre 6), dBZ. 189
V.29 DSD moyenne entre 20:00 et 20:30. On présente des DSD en bas de la colonne (traits continus) : observées par le DBS (noir), assimilées dans VAMOS (rouge) et fournies par le produit MRR (bleu) et en haut de la colonne (traits pointillés marqués) assimilées (VAMOS, noir) et fournies par le produit MRR (bleu). . . 190
V.30 Champs issus de la DSD assimilée (VAMOS), entre 19:30 et 20:30. A) Diamètre volumique moyen D_m, mm. B) Nombre volumique de gouttes N_{tot}, $ln(m^{-3})$. C) Inverse de la pente de la DSD, cm. La pente est calculée en faisant une régression linéaire sur le logarithme de la DSD (en la supposant suffisamment exponentielle). D) Au sol, taux de pluie assimilé VAMOS (RR, mm/h) et pente de la DSD (cm^{-1}). 191
V.31 Localisation (marqueur rouge) du site ou sont disposés les instruments ARM utilisés dans la section 6. (source Google Maps) 195
V.32 A-B) Facteur de réflectivité radar équivalent (Z_e, proportionnel à la réflectivité η) et vitesse Doppler moyenne (W) mesurées par le KAZR à Hyytiälä le 12/06/2014. C) Température et humidité relative au sol. D) Taux de pluie et cumul d'eau au sol (pluviomètre). 197
V.33 Humidité relative (A) et vent horizontal (B) mesurés par le radiosondage effectué à Hyytiälä à 11:36. 198

V.34 A-B) Facteur de réflectivité radar équivalente et vitesses Doppler moyennes assimilées pour l'évènement du 12/06/2014. Le trait noir vertical matérialise la date où l'évènement est découpé lors de l'assimilation. 200
V.35 Vents verticaux VAMOS (sorties d'assimilation, A) et TB2015 (B) entre 10:45 et 12:30. Positif vers le bas. Unité : m/s. 203
V.36 Séries temporelles de vents verticaux VAMOS (sorties d'assimilation) (bleu) et TB2015 (rouge) pour trois altitudes. A) $450m$. B) $750m$. C) $1150m$. Positif vers le bas. 203
V.37 Moments de la DSD restitués par notre algorithme d'assimilation (VAMOS, gauche) ou par TB2015 (droite). On présente : Taux de pluie RR (A et D) ; Facteur de réflectivité radar Z (B et E) ; Contenu en eau liquide LWC (C et F). 204
V.38 Profils verticaux moyens (sur la période 10:45-12:30) des contenus en eau liquide (A), taux de pluie (B) et facteurs de réflectivité radar (C) VAMOS (bleu) et TB2015 (rouge). 205
V.39 Séries temporelles des taux de pluie (A, B), des facteurs de réflectivités radar (C, D) et des contenus en eau liquide (E,F) pour deux altitudes (450 et $1050m$), tels qu'assimilés (VAMOS, bleu) et restitués par TB2015 (rouge). 205
V.40 DSD à $750m$ d'altitude assimilées (A) et TB2015 (B). Unité : $ln(m^{-3}mm^{-1})$. 207
V.41 DSD moyennes à $750m$ d'altitude entre 11:10 et 11:15 (A) et entre 12:15 et 12:20 (B). On présente les DSD assimilées (bleu) et TB2015 (rouge continus) ainsi que les barres d'erreurs associées aux DSD TB2015 (pointillés rouges). . 208
V.42 Spectres normalisés observés (A, D) et assimilés (VAMOS) (B, E) à 35 et $95GHz$. Vent vertical assimilé (VAMOS, C). Résultats pour l'altitude $750m$. Voir texte pour commentaires détaillés. 209
V.43 Localisation (croix rouge) du site ou sont disposés les instruments utilisés dans cette section (image Google). 215
V.44 Imagettes des évènements pluvieux du 04 novembre 2014 observés par le radar opérationnel de Berrimah du Bureau of Meteorology. Le site ARM est indiqué par le triangle rouge. 217
V.45 A-B) Facteurs de réflectivité radar équivalents Z_e observés par le KAZR (A) et assimilés dans VAMOS (B). C-D) Idem pour la vitesse Doppler moyenne W. E) Taux de pluie au sol tels que : mesurés par le 2DVD (rouge) et par le JW (noir) ainsi qu'assimilés dans VAMOS (bleu). Les pointillés délimitent de manière indicative les différents évènements au sens du Tableau V.6. 218
V.46 Profils verticaux de vents horizontaux et de températures fournis par le produit mergesonde. A) Direction du vent (0=vent vers le sud, 90=vent vers l'ouest, etc.) B) Module du vent. C) Température. Profils moyens pour la période 18:00-19:00 (bleu), 20:00-21:00 (rouge) ainsi que pour le radiosondage de 23:15 (noir). 219

V.47 A) Champ d'humidité relative (h_r) du produit mergesonde. B) Idem corrigé pour utilisation dans notre algorithme d'assimilation (voir texte). C) Taux de pluie et humidités relatives mesurées au sol. Les pointillés délimitent de manière indicative les différents évènements au sens du Tableau V.6. 222

V.48 Moments de la DSD et vents verticaux restitués par l'algorithme d'assimilation. A) Vent vertical, positif vers le bas. B) Diamètre volumique moyen D_m. C) Facteur de réflectivité radar Z. D) Contenu en eau liquide LWC. E) Nombre de gouttes N_{tot}. Les pointillés délimitent de manière indicative les différents évènements au sens du Tableau V.6. 226

V.49 A) Vitesse Doppler moyenne W enregistrée par le radar BASTA. B) Vitesse Doppler moyenne W à $94GHz$ simulée à partir de nos restitutions de DSD et vents verticaux. C) Vent vertical restitué à résolution $5s$, positif vers le bas. Les pointillés délimitent de manière indicative les différents évènements au sens du Tableau V.6. Les cadres blancs délimitent la zone de turbulence évoquée plus particulièrement en section 7.3.3. 227

V.50 Pour trois évènements, profils verticaux moyens de 3 paramètres du problème (facteur de réflectivité radar équivalent Z_e, diamètre volumique moyen D_m et contenu en eau liquide LWC), avec (bleu) ou sans (vert) évaporation. Pour Z_e, on présente aussi les observations KAZR (rouge). Profils d'humidités relatives moyennes pour le cas avec évaporation. 228

V.51 A) Taux de pluie au sol (mm/h) durant l'évènement 3 enregistrés par les disdromètres 2DVD (rouge) et JW (noir) et restitués par l'algorithme d'assimilation (bleu). B) Réflectivités (dBZ) enregistrées par le KAZR (rouge) et restituées par l'algorithme d'assimilation (bleu) pour la porte $100 - 200m$. 231

V.52 A) Nombre de gouttes (flux) enregistrées par le disdromètre 2DVD, selon la date et le diamètre. B) Idem restitution algorithme d'assimilation. C) DSD (volumique) correspondant à A. D) Idem restitution algorithme d'assimilation (VAMOS). E) Spectres Doppler hydrologiques à $35GHz$ correspondant à C. F) Idem restitution algorithme d'assimilation (VAMOS). G) Spectres Doppler enregistrés par le KAZR pour la porte $100 - 200m$. H) Idem restitution algorithme d'assimilation (VAMOS). Un spectre Doppler ou une DSD particulière correspond sur chaque image à une "ligne" verticale (date fixée). Unités : pour les DSD : $ln\left[m^{-3}mm^{-1}\right]$. Pour les spectres : $ln\left[mm^{-1}\left(m/s\right)^{-1}\right]$. 232

V.53 Pour l'évènement 3 du 04 novembre 2014 : A, B) Facteurs de réflectivité radar équivalents (Z_e) mesurés par le KAZR (A) et assimilés (VAMOS, B). C, D) Idem vitesses Doppler moyennes (W). E) Vent vertical assimilé (VAMOS). . . 235

V.54 Pour l'évènement 3 du 04 novembre 2014 : moments d'ordre 0, 1, 2, 3, 4 et 6 des DSD assimilées (VAMOS). Formellement, pour le moment d'ordre n, l'unité est : $10log_{10}\left(m^{n-3}mm^{-1}\right)$. Les pointillés rouges et verts représentent une partie des phases rouges et vertes définies par exemple sur la Figure V.52. 236

V.55 Pour l'évènement 3 du 04 novembre 2014 : pour différentes altitudes, spectres Doppler mesurés par le KAZR (colonne A) et assimilés (VAMOS, colonne B), vents verticaux VAMOS (colonne C) et DSD VAMOS (colonne D). Vents verticaux en m/s, positifs vers les bas. Spectres Doppler en $ln\left(10^{10}m^{-1}(m/s)^{-1}+1\right)$, DSD en $log_{10}\left(m^{-3}mm^{-1}\right)$. Les phases rouge (resp. vertes) sont représentées sur les lignes impaires (resp. paires). 237

V.56 Pour l'évènement 3 du 04 novembre 2014 : largeurs spectrales mesurées par le KAZR (A) et assimilées (VAMOS, B). 239

V.57 Pour l'évènement 3 du 04 novembre 2014 : A) Vent assimilé, résolution $5s$; B) Vent assimilé, résolution $30s$; C) Ecart-type des vents à $5s$ sur $30s$ (voir texte). Unité : m/s, positif vers le bas. 240

V.58 Pour l'évènement 3 du 04 novembre 2014 : Trois premiers moments centrés des spectres Doppler VAMOS (réflectivité Z_e, gauche ; vitesse Doppler moyenne W, milieu gauche, largeurs spectrales LS, milieu droite) et exemple des spectres correspondants (moyennés sur la période 19:01-19:06). Haut : spectres hydrologiques ; Milieu : ajout de l'atténuation ; Bas : Ajout du vent vertical (voir texte pour les détails et commentaires). 241

V.59 Pour l'évènement 5 du 04 novembre 2014 : Largeurs spectrales observées par le KAZR (A) et assimilés (B). Profils verticaux moyens sur la période 20:00-20:30. Les cadres blancs, comme sur la Figure V.49, délimitent la zone de turbulence évoquée dans cette section. 242

A.1 Justification de l'équation (A.2) pour le noyau de collision. On a représenté une grosse goutte (diamètre D_1, vitesse v_1) et deux petites (diamètres D_2, vitesses v_2) positionnées en deux points différents dans un volume V. Pendant la durée Δt, la grosse goutte parcourt la distance $v_1\Delta t$, les petites la distance $v_2\Delta t$. Indépendamment des effets dûs à l'air ambiant, la grosse goutte va alors percuter la petite si celle-ci se trouve à l'instant initial dans le cylindre rouge, de hauteur $(v_1-v_2)\,\Delta t$ et de diamètre D_1+D_2. La petite goutte située en bas à gauche du cylindre sera percutée à la fin de la période Δt, celle en haut à droite du volume V ne le sera pas. Ces considérations sont valables si le cylindre défini ici a un volume petit devant V. xxiii

A.2 Figure extraite de Seifert et al. (2005). Valeur de E_{coal}, efficacité de coalescence, en fonction du diamètre de deux gouttes lors d'une collision, et pour différentes paramétrisations. a) Paramétrisation proposée par Beard and Ochs III (1995) pour des petites gouttes ; b) Paramétrisation proposée par Low and List (1982a) ; c) Tentative d'unification des deux théories précédentes proposée par Seifert et al. (2005). xxvii

A.3 Figure extraite de Straub et al. (2010). Valeur de E_{coal}, efficacité de coalescence, en fonction du diamètre de deux gouttes lors d'une collision, dans le cadre de la paramétrisation proposée par Straub et al. (2010) (voir Eq.(A.15)). xxviii

A.4 Figure et légende extraite de Barros et al. (2008). Séquence d'une caméra à grande vitesse montrant différents types de break-up : (a) - filament ; (b) - sheet ; (c) - disc. Le temps entre les images est de $2ms$. xxx

A.5 Figure extraite de Schlottke et al. (2010). Nombre de satellites produits par une collision en fonction de l'énergie en jeu (traits), de l'excentricité du choc (abscisses) et du type de rupture (ellipses) d'après les simulations numériques de Schlottke et al. (2010) xxxii

A.6 Figure extraite de Testik et al. (2011). Graphe des devenir des gouttes après collisions en fonction du nombre de Weber We et du ratio de taille des gouttes $p = D_2/D_1$ à partir d'expérimentations présentées dans Testik et al. (2011). "Neck" break-up est synonyme de rupture en filament. "Crown" break-up est un type de rupture en disque dans le cas d'une plus faible énergie en jeu (Testik et al. (2011)) xxxv

C.1 Vitesses terminales de chute de gouttes de pluie. rouge) relation (C.4), $C_D = 0.5$; bleu) relation (C.5), $C_D = 1$ (le diamètre du pancake est défini comme celui de la sphère de même volume) ; noir) relation de Foote and Du Toit (1969) approchant les mesures de Gunn and Kinzer (1949) avec une expression polynomiale (degré 5 ici) xliii

C.2 Diverses relations vitesse-diamètre appliquées à des gouttes sphériques sur la gamme de diamètres $[0 : 7.5mm]$. Les relations utilisées sont décrites dans le texte par : bleu) Eq.(C.10) ; noir) Eq.(C.11) ; vert) Eq.(C.9) ; magenta) Eq.(C.13) ; rouge) Eq.(C.12). xlvi

C.3 Figure extraite de Foote and Du Toit (1969) : vitesse terminale de chute des gouttes telle que mesurée par Gunn and Kinzer (1949) (courbe du bas) et modélisée selon les conditions de pression et de température par Foote and Du Toit (1969). xlvii

D.1 Représentation du schéma upstream sous forme de flux entrant et sortant des mailles. À la date $k+1$, la quantité OUT va être advectée hors de la maille i et la quantité IN va y entrer. Voir le texte pour les équations complètes. . . . l

G.1 Définition du $\tilde{d}_{pq}$ de l'Eq.(G.4). Une cellule pluvieuse, présentant un front rectiligne (pointillés bleus), progresse dans la direction $\boldsymbol{v}$. Elle touche deux liaisons p et q, de même longueur L, partant du même capteur Ku positionné en O, et faisant un angle β_{pq} entre elles. La liaison pointant vers l'est fait un angle γ par rapport au sud. L'angle δ caractérise la direction de propagation. lxiv

G.2 Idem Figure G.1 avec l'introduction de notations complémentaires pour la preuve de l'Eq.(G.5). lxv

G.3 Idem Figure G.1 mais avec des liaisons p et q provenant de deux capteurs Ku différents, positionnés en O_1 et O_2. Pour exprimer $\tilde{d}_{pq}$, on introduit $\tilde{d}_{O_1O_2}$, distance entre les deux capteurs, et une liaison fictive $q^{'}$, parallèle à q d'origine O_1. lxvii

G.4 Idem Figure G.3 avec l'introduction de notations complémentaires pour la preuve de l'Eq.(G.16). lxviii

G.5 Représentation d'une situation où les points milieux A, B et C de trois liaisons Ku quelconques sont alignés selon une direction caractérisée par un angle v. . lxx

G.6 Visualisation des 6 simulations réalisées pour tester la méthode de triangulation de restitution du vent avec 1 capteur Ku et 4 liaisons. $A1$ à $A6$: les liaisons (simulées, voir Tableau G.1) sont représentées en blanc au nord-est de chaque carte. La cellule de pluie à l'instant initiale est également représentée, tout comme la vitesse d'advection (flèches blanches). $A1$ à $A6$: séries temporelles d'atténuations enregistrées par chaque liaison Ku lors de la l'advection de la cellule pluvieuse. On a sur chaque graphe : rouge : liaison pointant le plus à l'ouest ; bleu : liaison pointant sud-sud-ouest ; magenta : liaison pointant sud-sud-est ; noir : liaison pointant le plus à l'est. lxxiii

G.7 Evaluation de la fonction de coût permettant l'évaluation des conditions d'advection (Eq.(G.1)) avec 1 (a) ou 2 (b) capteurs Ku, dans le cadre de la simulation 2 (voir Tableau G.2), et de la configuration A pour le cas à deux capteurs (voir Tableau G.3). lxxiv

G.8 Idem Figure G.6 mais avec deux capteurs Ku, dans le cadre de la simulation 2 (Tableau G.2), et de la configuration A (Tableau G.3). Les liaisons 1 à 4 proviennent du 1^{er} capteur et sont dirigées d'ouest en est. Les liaisons 5 à 8 sont celles issues du 2^{nd} capteur. lxxv

H.1 Expérience jumelle parfaite. A) Champ initial vrai (A_0^{vrai}, cellule de pluie simulée), géométrie des liaisons Ku projetées au sol (traits blanc au nord-est), conditions d'advection (flèche et vitesse) et grande zone d'assimilation (cadre blanc). B) Champ initial assimilé (A_0^{assi}), géométrie des liaisons et grande zone d'assimilation. Les liaisons Ku sont dessinées à titre indicatif (angles et longueurs ne sont pas très précisément respectés). lxxix

H.2 Expérience jumelle parfaite. Projection (i.e. pluviométrie moyenne) des champs vrais et assimilés (A_0^{vrai} et A_0^{assi}) restreints à la grande zone d'assimilation selon les axes est-ouest (A) et nord-sud (B). lxxx

H.3 Expérience jumelle parfaite. A) Pour chaque liaison, séries temporelles d'observations vraies ($\mathcal{A}^{vrai}$, trait continu) et assimilées ($\mathcal{A}^{assi}$, trait pointillé, superposé). B) Séries temporelles des différences entre observations vraies et assimilées ($\mathcal{A}^{vrai} - \mathcal{A}^{assi}$). Les liaisons sont numérotées d'ouest en est. lxxx

H.4 Expérience jumelle avec erreur sur la longueur des liaisons Ku. A) Champ initial vrai (A_0^{vrai}), géométrie des liaisons Ku (traits blanc à l'est) et grande zone d'assimilation (cadre blanc). B) Champ initial assimilé (A_0^{assi}), géométrie des liaisons et grande zone d'assimilation. Les liaisons Ku sont dessinées à titre indicatif. lxxxi

H.5 Expérience jumelle avec erreur sur la vitesse d'advection v. A) Champ initial vrai (A_0^{vrai}), géométrie des liaisons Ku (traits blanc à l'est), conditions d'advection vraies (flèche et vitesse) et grande zone d'assimilation (cadre blanc). B) Champ initial assimilé (A_0^{assi}) avec sous-estimation de v, géométrie des liaisons et grande zone d'assimilation. C) Idem avec surestimation de v. Les liaisons Ku sont dessinées à titre indicatif. lxxxiii

H.6 Expérience jumelle avec erreur sur la vitesse d'advection v. Pour chaque liaison, séries temporelles des différences entre observations vraies et assimilées dans le cas d'une sous-estimation de la vitesse d'advection (traits continus) ou d'une surestimation (traits pointillés). Les liaisons sont numérotées d'ouest en est. lxxxiv

H.7 Expérience jumelle avec advection vers le nord. A) Champ initial vrai (A_0^{vrai}), géométrie des liaisons Ku (traits blanc au nord), conditions d'advection vraies (flèche et vitesse) et grande zone d'assimilation (cadre blanc). B) Champ initial assimilé (A_0^{assi}) sans terme de filtrage. C) Idem avec terme de filtrage. Les liaisons Ku sont dessinées à titre indicatif. lxxxiv

J.1 Illustration du processus de minimisation numérique de la fonction de coût du 4D-VAR pour le vent vertical. Expérience jumelle décrite au chapitre V, section 4.3, pour le profil 2 de vent horizontal. A) État vrai (w^{vrai}). B à D) Vent vertical pour les itérations 0, 12 et 74 de la minimisation. E) Etat assimilé (w^{assim}). Vents à la résolution des observations ($30s$) et positifs vers le bas. . . xci

J.2 Selon l'itération du minimiseur numérique du 4D-VAR, erreur sur la DSD en haut de la colonne atmosphérique (A), sur le vent vertical (B) et partie observations de la fonction de coût (C). Même expérience jumelle que Figure J.1. xcii

K.1 Estimation de l'atténuation due au radôme du KAZR à partir des mesures radar et disdromètre. Voir dans le texte pour les commentaires détaillés. . . . xcv

Liste des tableaux

I.1 Bandes de fréquence micro-ondes classiquement utilisées pour la télédétection de la pluie ou des nuages. On indique, pour chaque bande, la lettre la désignant, les gammes de fréquence et de longueur d'onde λ correspondant, la limite haute de l'approximation Rayleigh pour la diffusion par des particules sphériques (voir texte) et les instruments utilisés dans ce travail émettant ou recevant dans cette bande. 19

I.2 Caractéristiques principales des 4 radars à visée verticale utilisés dans ce travail. Voir commentaires dans le texte. 29

IV.1 Caractéristiques des 4 liaisons microondes Terre-satellites géostationnaires enregistrées par le capteur Ku installé à Mirabel en 2013. Les indices les référençant sont ceux mentionnés sur la Figure IV.1. Pour les azimuth, 180° = vers le sud ; 270° = vers l'est. 89

IV.2 Position des pluviomètres du réseau HPiconet. Les indices sont ceux définis sur la Figure IV.1. 91

IV.3 Description des 4 liaisons Terre-satellite simulées et projetées au sol utilisées dans cette section. 99

IV.4 Conditions d'advections pour les 15 simulations présentées dans la section 5.2. On rappelle $\delta = 180^\circ$ = déplacement vers le nord ; $\delta = 270^\circ$ = déplacement vers l'est. 101

IV.5 Statistiques d'erreurs entre états initiaux vrais (A_0^{vrai}) et assimilés (A_0^{assim}) pour chacune des 15 simulations réalisées dans le section 5.2, pour des expériences d'assimilation avec 1 ou 2 capteurs Ku. Les indicateurs utilisés sont définis dans la section 5.1 105

IV.6 Statistiques d'erreurs entre états initiaux vrais (A_0^{vrai}) et assimilés (A_0^{assim}) sur les 15 simulations réalisées dans le section 5.2, pour des expériences d'assimilation avec 1 ou 2 capteurs Ku. Les indicateurs utilisés sont définis dans la section 5.1. On distingue également les évènements avec une advection majoritairement orientée vers l'est ($\delta > 225^\circ$) et vers le nord ($\delta < 225^\circ$). 106

IV.7 Caractéristiques des évènements étudiés de l'année 2013. Dates (colonnes 2, 3, 4), atténuations (dB) maximales et moyennes relevées le long des 4 liaisons Ku (colonnes 5 et 6), estimation de l'iso-0 (colonne 7), et cumuls d'eau (mm) enregistrés par les 10 pluviomètres du réseau HPiconet : pluviomètre ayant vu le moins d'eau (colonne 8), le plus d'eau (colonne 9) et moyenne sur les 10 (colonne 10). 108

IV.8 Restitution des paramètres d'advection par analyse de la corrélation entre cartes radar successives pour les évènements de l'année 2013. On calcule entre chaque couple de cartes successives une valeur maximale de corrélation croisée, puis la moyenne (colonne 3) et l'écart-type (colonne 4) de ces valeurs sur tous les couples. On définit la vitesse d'advection (colonnes 5 et 6) comme la moyenne sur tous les couples des décalages permettant de maximiser les corrélations croisées. Les évènements présentant les caractères advectifs les plus marqués sont en gras. 111

IV.9 Comparaison entre cartes radar et restituées par notre algorithme d'assimilation pour chaque évènement de l'année 2013, pour 2 résolutions différentes ($1 * 1$ et $2 * 2km^2$), et pour les deux zones d'assimilation (voir Figure IV.4). Coefficients de corrélation, droites de régression, biais (biais positif = surestimation de la pluie par notre algorithme par rapport au radar), taux de pluie radar moyen. 113

IV.10 Estimation ponctuelle de la pluie. Coefficients de corrélation, équation des droites de régression à différentes résolutions temporelles (5 à $25min$), biais moyen entre pluviomètres et sorties d'assimilation (colonne 3) et entre pluviomètres et cartes radar (colonne 4). Taux de pluie moyen enregistré par les pluviomètres. On mélange ici les données des 4 évènements advectifs (évènements 4, 5, 7, 8) et des 7 pluviomètres situés dans la grande zone d'assimilation. 121

V.1 Comparaisons entre champs vrais, réalistes et assimilés, pour les deux premiers moments des spectres Doppler (vitesse Doppler moyenne W et largeur spectrale LS) et pour le nombre de gouttes enregistrées par le disdromètre. On présente pour les deux premiers cas la $MAPE$ (Eq.(V.18)) et le biais relatif (Eq.(V.19)) et pour le troisième cas la MAE ((V.20)). La dernière colonne donne la moyenne des champs vrais. 151

V.2 Comparaison (MAE, voir Eq.(V.20)) entre champs vrais et assimilés pour le vent vertical (colonne 3) et différents paramètres tirés de la DSD : taux de pluie (RR), facteur de réflectivité radar (Z), diamètre volumique moyen (D_m), contenu en eau liquide (LWC) et nombre de gouttes (N_{tot}). On présente les résultats pour des observations intégrées à $10s$ ou $2min$ (colonne 2) et pour des champs dégradés à différentes résolution ($10s$ à $8min$). On présente aussi le biais absolu et la valeur absolue moyenne. Pour chaque situation, la meilleure performance selon la résolution des observations est en gras. 154

V.3 Comparaison entre états observés et assimilés pour les trois premiers moments des spectres Doppler (réflectivité η, vitesse moyenne W, largeur spectrale LS) et le nombre de gouttes disdromètre. On distingue le cas où l'on considère toutes les altitudes, et où l'on sépare les données en trois gammes d'altitudes. Les indicateurs présentés (et introduits en section 3.3.4) sont la MAE (moyenne des erreurs absolues, Eq.(V.20)), la MAPE (moyenne - en % - des valeurs absolues des erreurs relatives, Eq.(V.18)) et le biais relatif $rbias$ (Eq.(V.19)). Biais positif = surestimation par l'assimilation. 184

V.4 Comparaison entre états observés et assimilés pour les trois premiers moments centrés des spectres Doppler et le nombre de gouttes disdromètre pour le cas en Finlande du 12/06/2014. On distingue le cas où l'on considère toutes les altitudes, et où l'on sépare les données en trois gammes d'altitudes. Les indicateurs présentés sont la MAE (Eq.(V.20)), la MAPE (Eq.(V.18)) et le biais relatif (Eq.(V.19)). Biais positif = surestimation par l'assimilation. Statistiques calculées sur la période 11:00-12:00. 201

V.5 Mêmes statistiques que Tableau V.4 calculés entre les moments centrés des spectres à $95GHz$ déduits de l'assimilation et mesurés par le WACR. 211

V.6 Dates et caractéristiques (cumul d'eau $RRcum$ et taux de pluie moyen au sol $RRmean$) des 5 évènements successifs du 04 novembre 2014 à Darwin. 221

V.7 Comparaison entre états observés et assimilés pour les trois premiers moments centrés des spectres Doppler et le nombre de gouttes disdromètre (flux de gouttes) pour le cas à Darwin du 05/11/2014. On présente les résultats pour les 5 évènements de la journée (au sens du Tableau V.6), colonne 3, puis seulement pour les 4 derniers (colonne 4) et finalement pour les évènements 2, 3, 4 et 5 individuellement (colonnes 5 à 8). Les indicateurs présentés sont la MAE, la MAPE et le biais relatif (cf. section 3 Eq.(V.20), Eq.(V.18) et Eq.(V.19) respectivement). On rappelle qu'un biais positif correspond à une surestimation par VAMOS (par l'assimilation). 224

G.1 Description des 4 liaisons Terre-satellite simulées et projetées au sol pour évaluer la pertinence de la méthode de triangulation pour restituer les conditions d'advection avec un seul capteur Ku. lxxi

G.2 Conditions d'advection pour les 6 simulations réalisées ici : colonne 2 : direction d'advection (180 ° correspond à une propagation sud-nord, 270 ° à une propagation ouest-est, etc.) ; colonne 3 : vitesse d'advection. Colonne 4 : valeurs de vitesse restituées par la méthode de triangulation avec un seul capteur Ku et 4 liaisons en supposant connue la direction d'advection. lxxii

G.3 Positionnement du second capteur Ku par rapport au premier dans le cadre des 4 configurations définies ici. lxxv

G.4 Valeurs des directions et des vitesses de vent restituées par la méthode de triangulation avec deux capteurs Ku, pour les simulations décrites dans le Tableau G.2 et les configurations décrites dans le Tableau G.3. On calcule aussi, pour chaque configuration, la moyenne des écarts absolus entre valeurs restituées et réalité simulée. lxxvi

Annexe A: Coalescence / break-up : paramétrisation

Dans le chapitre II, nous revenons sur le modèle d'équations aux dérivées partielles (EDP) développé pour représenter l'évolution spatiotemporelle de la DSD lors de la chute des gouttes. Ce modèle permet de représenter l'effet des collisions entre gouttes, avec un formalisme introduit dans la section 2.5. Dans cette annexe, nous référençons un certain nombre de paramétrisations proposées dans la littérature pour représenter explicitement ces phénomènes. On note bien que dans le travail présenté dans cette thèse l'effet des collisions entre gouttes n'a pas été pris en compte. Les paramétrisations présentées ici le sont donc pour comprendre la difficulté du problème et pour d'éventuelles évolutions futures de ce travail.

Les grandeurs utilisées ici (noyau de collision, efficacité de coalescence, etc.) ont été introduites dans le chapitre II, section 2.5 (voir sinon dans la nomenclature au début du manuscrit).

Noyau de collision Considérons deux gouttes de masses, diamètres et vitesses de chute $m_1 > m_2$, $D_1 > D_2$, $v_1 > v_2$, réparties dans un volume V. Voyons sous quelles conditions ces deux gouttes vont se percuter durant la durée Δt. La Figure A.1 illustre cette situation. Pendant la durée Δt, la grosse goutte va parcourir verticalement une distance $v_1 \Delta t$ et la petite $v_2 \Delta t$. La grosse goutte, si elle part de plus haut que la petite, la "rattrapera" (c'est-à-dire finira au moins aussi bas qu'elle) si celle-ci est à moins de la distance $(v_1 - v_2)\,\Delta t$ d'elle. Par ailleurs, projetées sur le plan horizontal, les deux gouttes se touchent si le centre de la petite goutte se situe dans un disque de diamètre $D_1 + D_2$ autour du centre de la grosse goutte. La grosse goutte "couvre" ainsi un volume $\pi\,(D_1 + D_2)^2\,(v_1 - v_2)\,\Delta t/4$, dans le sens où, en supposant aucun effet d'aspiration ou d'éjection dû à l'air ambiant (voir plus bas), elle va percuter la petite goutte sous réserve qu'elle se trouve à l'instant initial dans ce volume immédiatement en-dessous d'elle.

En supposant les deux gouttes réparties uniformément dans le volume V, la probabilité de collision pendant Δt est alors :

$$\frac{1}{V}\frac{\pi}{4}\,(D_1 + D_2)^2\,(v_2 - v_1)\,\Delta t \tag{A.1}$$

Dans la mesure où l'on travaille ici avec des grandeurs normalisées en volume (DSD n_m) et

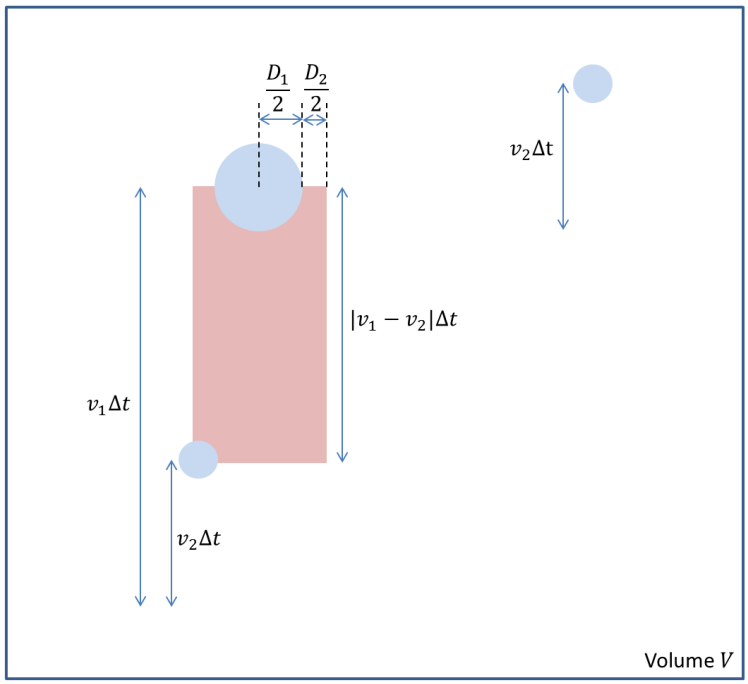

Figure A.1 – Justification de l'équation (A.2) pour le noyau de collision. On a représenté une grosse goutte (diamètre D_1, vitesse v_1) et deux petites (diamètres D_2, vitesses v_2) positionnées en deux points différents dans un volume V. Pendant la durée Δt, la grosse goutte parcourt la distance $v_1\Delta t$, les petites la distance $v_2\Delta t$. Indépendamment des effets dûs à l'air ambiant, la grosse goutte va alors percuter la petite si celle-ci se trouve à l'instant initial dans le cylindre rouge, de hauteur $(v_1 - v_2)\,\Delta t$ et de diamètre $D_1 + D_2$. La petite goutte située en bas à gauche du cylindre sera percutée à la fin de la période Δt, celle en haut à droite du volume V ne le sera pas. Ces considérations sont valables si le cylindre défini ici a un volume petit devant V.

que l'on estime des dérivées, le noyau de collision Φ_{coll} s'écrit :

$$\Phi_{coll}(m_1, m_2) = \frac{\pi}{4}\left(D_1 + D_2\right)^2 (v_2 - v_1) \tag{A.2}$$

En supposant les gouttes sphériques et tombant à leur vitesse terminale de chute, il ne dépend que de la masse des gouttes (ou de leur diamètre) et éventuellement de l'altitude (pour la relation vitesse-diamètre). Il s'exprime bien en m^3/s.

Cette expression suppose comme on l'a vu que les gouttes qui se percutent sont composées des paires percuteur / percuté dont le percuté est dans le "chemin" du percuteur. On trouve ainsi dans la littérature les noms de noyau de collision géométrique (Beheng (2010), Lin and Lee (1975)) ou gravitationnel (Feingold et al. (1988), Prat and Barros (2007)) pour cette expression. Mais il existe deux objections à cette affirmation. Tout d'abord on peut imaginer qu'une petite goutte se trouvant dans le chemin du percuteur soit chassée de ce chemin par la couche d'air en mouvement entourant le percuteur. Au contraire, on pourrait imaginer des cas où une goutte hors du chemin du percuteur pourrait être aspirée dans ce chemin. On peut se référer à Beheng (2010) pour une explication plus détaillée de ces phénomènes. Pour les prendre en compte, on

ajoute à l'expression (A.2) un facteur sans unité E_{coll}, appelé efficacité de collision. On a alors :

$$\Phi_{coll}(m_1, m_2) = \frac{\pi}{4} (D_1 + D_2)^2 (v_2 - v_1) E_{coll} \tag{A.3}$$

On retrouve l'expression (A.3) pour le noyau de collisions dans plusieurs études, notamment Hu and Srivastava (1995) ou Prat and Barros (2007) sur la pluie, Leroy (2007), Flossmann and Wobrock (2010) ou Pinsky et al. (2001) sur les nuages ou des publications plus générales (Hall (1980), Beheng (2010)).

Sur le domaine des nuages (taille de gouttes de l'ordre de quelques micromètres), de nombreuses études ont été menées pour estimer le noyau de collision. Beheng (2010) les référence. Néanmoins, il ressort de ces études (voir les données de Hall (1980) ou de Pinsky et al. (2001)), même si elles ne traitent généralement pas de paires de gouttes de la taille de gouttes de pluie, que le noyau de collision est très proche de 1 dans la pluie. Par exemple, Pinsky et al. (2001) donne un $E_{coll} = 1.0024$ pour un couple $(250\mu m - 39\mu m)$, et tendant vers 1 lorsque les diamètres augmentent encore.

Les études travaillant du côté pluie (Hu and Srivastava (1995), Prat and Barros (2007), Flossmann and Wobrock (2010)) prennent ainsi $E_{coll} = 1$, se limitant au noyau de collision géométrique (A.2).

Efficacité de coalescence L'objet de ce paragraphe est d'étudier la paramétrisation de l'efficacité de coalescence $E_{coal}(m_1, m_2)$, probabilité qu'une collision entre deux gouttes de masses m_1 et m_2 résulte en coalescence (et non en break-up). On rappelle que le noyau de break-up E_{bu} sera ensuite simplement défini par $E_{bu} = 1 - E_{coal}$.

Un certain nombre de paramétrisations existent dans la littérature pour cette efficacité de coalescence. Celles proposées par Low and List (1982a), Brazier-Smith et al. (1972) et Straub et al. (2010) seront détaillées ici, celles proposées par Beard and Ochs III (1995) et Seifert et al. (2005), et proposant des synthèses et généralisations d'autres paramétrisations, seront vues plus rapidement.

Toutes ces paramétrisations se basent sur des considérations énergétiques. Lors d'un choc entre deux gouttes, une certaine énergie cinétique, susceptible de faire exploser les gouttes, est mise en jeu, tandis que la tension de surface des gouttes tend au contraire à les garder entières. Plus l'énergie cinétique sera grande par rapport à la tension de surface, plus les gouttes auront tendance à exploser, et plus E_{coal} sera petit.

Définissons donc tout d'abord quelques grandeurs qui seront utilisées par les différentes paramétrisations. Supposons comme précédemment deux gouttes, de masses, diamètres et vitesses de chute $m_1 > m_2$, $D_1 > D_2$, $v_1 > v_2$. L'énergie cinétique CKE (en J) mise en jeu dans la collision s'écrit :

$$CKE = \frac{\rho_e \pi}{12} \frac{D_1^3 D_2^3}{D_1^3 + D_2^3} (v_1 - v_2)^2 \tag{A.4}$$

Avec ρ_e la masse volumique de l'eau.

Pour l'une des deux gouttes, par exemple la première, la tension de surface S_1 (en J) s'écrit :

$$S_1 = \pi \gamma_e D_1^2 \tag{A.5}$$

Avec γ_e la tension superficielle de l'eau (typiquement $\gamma_e \approx 73.10^{-3} N/m$ pour de l'eau à $20\,^\circ C$ en contact avec l'air).

Avant le choc, la tension de surface est la somme de celle de chaque goutte :

$$S_{1+2} = \pi \gamma_e (D_1^2 + D_2^2) \tag{A.6}$$

Après le choc, la tension de surface est celle de l'équivalent sphérique de la somme des deux gouttes. Le volume de cette goutte "coalescée" est $2\pi(D_1^3 + D_2^3)/3$ et le diamètre équivalent $(D_1^3 + D_2^3)^{1/3}$. La tension de surface après le choc est donc :

$$S_c = \pi \gamma_e (D_1^3 + D_2^3)^{2/3} \tag{A.7}$$

On définit alors un nombre de Weber, We^*, par :

$$We^* = \frac{CKE}{S_c} \tag{A.8}$$

Plus We^* est grand, plus l'énergie cinétique est prépondérante et plus on s'attend à avoir un E_{coal} petit.

On note par ailleurs qu'on trouve souvent défini un autre nombre de Weber, We, par :

$$We = \frac{(v_1 - v_2)^2 D_2 \rho_e}{\gamma_e} \tag{A.9}$$

Il traduit les mêmes phénomènes, en croissant avec la vitesse relative des gouttes, donc leur énergie cinétique. Schlottke et al. (2010) donne la relation entre les deux nombres de Weber ainsi définis.

Nous allons à présent voir l'utilisation de ces grandeurs pour paramétrer effectivement E_{coal}.

Paramétrisation Low and List (1982a) Le modèle de Low and List (1982a) part de l'idée que le E_{coal} correspond à la capacité des gouttes après leur choc à dissiper l'énergie "en trop" sans exploser, c'est-à-dire en oscillant et en se déformant (voir détails dans Low and List (1982a) ou Orme (1997)). L'énergie en trop est composée de l'énergie cinétique de collision et de la différence entre la tension de surface avant et après le choc (la surface étant plus grande avant avec deux gouttes qu'après avec une seule). Ils bâtissent ensuite un modèle empirique à partir d'expérimentations faites sur des couples de gouttes en chute libre en laboratoire de tailles $(1.8, 0.395)mm$ à $(4.6, 1.8)mm$.

La tension de surface rendue disponible par la collision est :

$$\Delta S_{\gamma_e} = S_{1+2} - S_c \tag{A.10}$$

Avec S_{1+2} la tension de surface avant le choc (Eq.(A.6)) et S_c celle après le choc (Eq.(A.7)).

L'énergie totale libérée par la collision est ensuite :

$$En_{tot} = CKE + \Delta S_{\gamma_e} \tag{A.11}$$

Le modèle empirique paramétrant E_{coal} proposé par Low and List (1982a) s'écrit alors :

$$E_{coal} = 0.778(1 + \frac{D_2}{D_1})^{-2} e^{-\frac{2.6.10^{-6}\sigma En_{tot}^2}{S_c}} \tag{A.12}$$

Sa pertinence est évaluée positivement par exemple dans Beheng et al. (2006). Il est très utilisé dans les études cherchant à estimer des distributions d'équilibres pour l'EDP (voir Eq. (II.5) du chapitre II), par exemple dans List and McFarquhar (1990), Hu and Srivastava (1995), Prat and Barros (2007) ou Mcfarquhar (2010). Il est confronté à d'autres paramétrisations dans Valdez and Young (1985), Prat et al. (2012) ou Barthes and Mallet (2013).

Paramétrisation Brazier-Smith et al. (1972) Le modèle de Brazier-Smith et al. (1972) est également basé sur des considérations sur les énergies mises en jeu (inertie vs. tension de surface) et utilise directement le nombre de Weber We (Eq.(A.9)). Il utilise ensuite une fonction empirique basée sur des expérimentations sur des gouttes de diamètre 0.15 à $0.75mm$ projetées l'une contre l'autre (on est donc ici à la limite inférieure de la gamme de diamètres de la pluie) avec un angle donné par un courant (pas en chute libre donc). Orme (1997) détaille aussi cette paramétrisation.

Ses équations sont :

$$E_{coal} = 2.4\frac{g(D_1/D_2)}{We} \tag{A.13}$$

$g(D_1/D_2)$ est une fonction empirique s'écrivant :

$$g\left(\frac{D_1}{D_2}\right) = \frac{4.8\left(1 + {\frac{D_1}{D_2}}^2 - \left[1 + {\frac{D_1}{D_2}}^3\right]^{2/3}\right)\left(1 + {\frac{D_1}{D_2}}^3\right)^{11/3}}{{\frac{D_1}{D_2}}^6\left(1 + \frac{D_1}{D_2}\right)^2} \tag{A.14}$$

Plusieurs études (Munnannur and Reitz (2007), Ko and Ryou (2005)) détaillent la paramétrisation de Brazier-Smith et al. (1972) avant de remarquer qu'elle ne prévoit qu'un type de break-up (rebond) et pas la fragmentation (voir paragraphe sur le break-up). Ils développent alors des modèles pour intégrer la fragmentation à partir de ces théories. Valdez and Young (1985) ou Barthes and Mallet (2013) confrontent ce modèle à celui de Low and List (1982a).

Synthèses Beard and Ochs III (1995) et Seifert et al. (2005) Plus globalement Beard and Ochs III (1995) compare plusieurs paramétrisations pour E_{coal}, en distinguant les études et expérimentations réalisées sur des gouttes posées ou dans des flux (comme Brazier-Smith et al. (1972)) d'avec celles réalisées avec des gouttes en chute libre (comme Low and List (1982a)).

On note que Beard and Ochs III (1995) prend en compte d'autres paramétrisations que celles de Low and List (1982a) ou Brazier-Smith et al. (1972), voir notamment Whelpdale and List (1971), pour des mesures avec des couples de gouttes de tailles très différentes (typiquement 0.7vs.0.04mm) ou Levin and Machnes (1977) pour des gouttes de tailles proches et entre 0.1 et 0.6mm. Ils notent des incompatibilités entre ces différentes paramétrisations malgré des tendances similaires. Beard and Ochs III (1995) remarque également que le fait que les expérimentations à l'origine des paramétrisations soient réalisées dans des conditions différentes peut expliquer ces incompatibilités (par exemple entre Brazier-Smith et al. (1972) et de Low and List (1982a)).

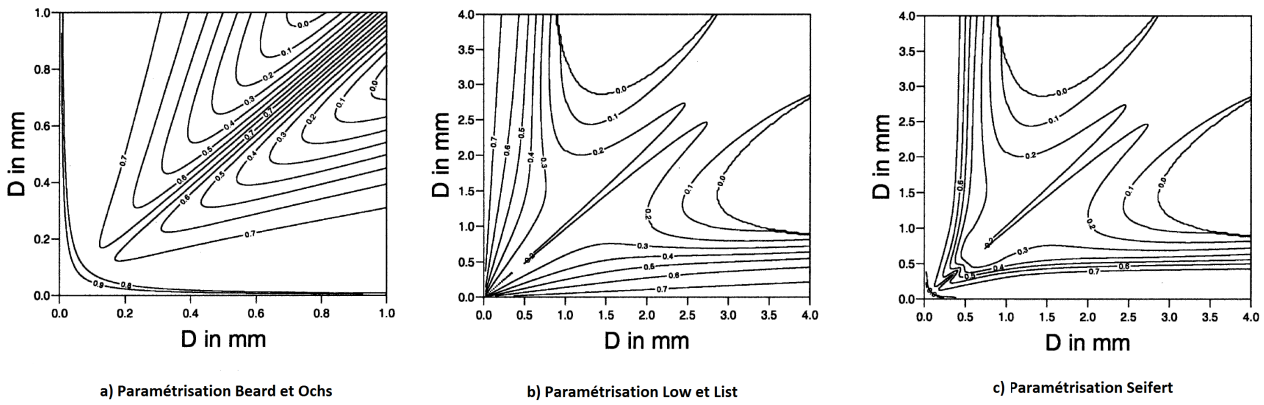

Figure A.2 – Figure extraite de Seifert et al. (2005). Valeur de E_{coal}, efficacité de coalescence, en fonction du diamètre de deux gouttes lors d'une collision, et pour différentes paramétrisations. a) Paramétrisation proposée par Beard and Ochs III (1995) pour des petites gouttes ; b) Paramétrisation proposée par Low and List (1982a) ; c) Tentative d'unification des deux théories précédentes proposée par Seifert et al. (2005).

Beard and Ochs III (1995) propose ensuite une paramétrisation en fonction de la taille des gouttes considérées, paramétrisation résumée et unifiée par exemple par Seifert et al. (2005). Ce dernier note cependant encore une incompatibilité entre les paramétrisations de Beard and Ochs III (1995) et de Low and List (1982a) (voir Figure A.2 a et b). Il note que la paramétrisation de Low and List (1982a) semble incohérente pour les petits diamètres (surestimation du break-up par rapport à la coalescence). Ils proposent finalement d'utiliser la paramétrisation de Beard and Ochs III (1995) pour les gouttes avec $D_2 < 0.3mm$ (D_2 diamètre de la petite goutte impliquée dans la collision), celle de Low and List (1982a) pour celles avec $D_2 > 0.6mm$ et propose une formule assurant la cohérence entre les deux pour

$0.3 < D_2 < 0.6mm$ (voir leur Eq.5). La Figure A.2c montre la sortie de cette formule.

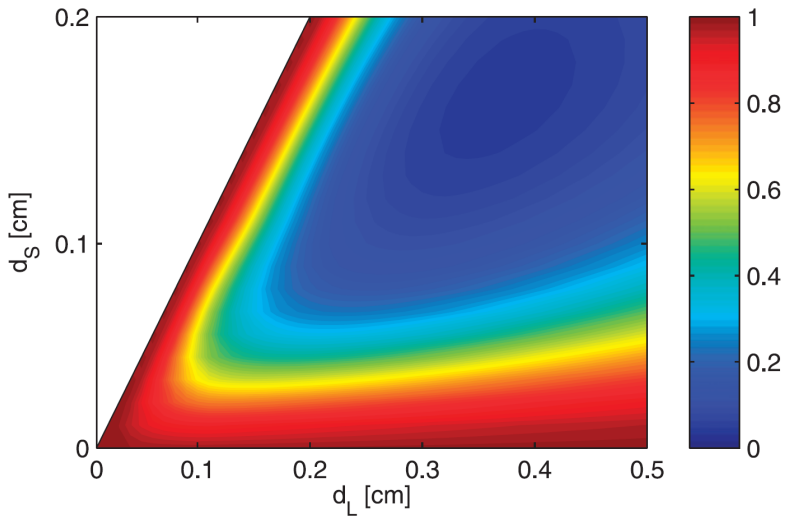

Figure A.3 – Figure extraite de Straub et al. (2010). Valeur de E_{coal}, efficacité de coalescence, en fonction du diamètre de deux gouttes lors d'une collision, dans le cadre de la paramétrisation proposée par Straub et al. (2010) (voir Eq.(A.15)).

Paramétrisation Straub et al. (2010) Straub et al. (2010) propose un nouveau modèle pour l'expression de E_{coal}, basé sur la simulation numérique de la collision de 32 paires de gouttes (simulation qui sera aussi utilisée pour paramétrer la fonction de break-up, voir paragraphe suivant). Leur modèle est une expression en exponentielle décroissante du nombre de Weber We^*, à savoir :

$$E_{coal} = e^{-1.15We^*} \tag{A.15}$$

La Figure A.3 présente la sortie de la paramétrisation de Straub et al. (2010) pour des collisions entre gouttes de 0 à $5mm$.

Comportement général Dans tous les cas (Figures A.2 et A.3), l'efficacité de coalescence est proche de 1 (coalescence systématique) lorsque l'on a affaire à des très petites gouttes, typiquement moins de $0.3mm$, ou lorsque la petite goutte est suffisamment petite et beaucoup plus petite que la grosse (par exemple un couple $0.5mm$; $4mm$). Elle est au contraire proche de 0 (explosion systématique) pour des gouttes de tailles voisines et suffisamment proches ($2mm$; $2mm$), ou lorsque la grosse goutte est très grosse et la petite suffisamment grosse également ($1.5mm$; $3mm$).

Fonction de break-up L'objet de ce paragraphe est d'étudier la fonction de break-up, $\chi_{bu}(m; m_1, m_2)$, distribution des masses des gouttes résultant d'une collision de gouttes de masses m_1 et m_2 débouchant sur un break-up (gouttes satellites). Pour alléger les notations, χ_{bu} sera simplement notée χ dans cette section.

Généralités / Mesures Il existe là encore dans la littérature un certain nombre d'études proposant des paramétrisations pour χ. Beaucoup sont basées sur les mesures en laboratoire de Low and List (1982a). Celui-ci se base sur les travaux de McTaggart-Cowan and List (1975), qui avait étudié le devenir de deux gouttes lâchées l'une au dessus de l'autre de manière à ce qu'elles se percutent, et ce pour 5 couples de gouttes de diamètres différents. Low and List (1982a) étend cette base de données de manière à avoir 10 couples de gouttes. Pour chaque couple, pour avoir des résultats représentatifs, Low and List (1982a) étudie entre 109 et 173 collisions (en incluant les collisions aboutissant à une coalescence). Les gouttes étudiées ont des diamètres entre 0.395 et 1.8mm pour la plus petite, entre 1.8 et 4.6mm pour la plus grosse. Dans la foulée de ces travaux, Low and List (1982b) propose une paramétrisation pour χ, détaillée ci-après. Constatant certaines faiblesses dans cette paramétrisation, plusieurs études ont travaillé à la perfectionner (McFarquhar (2004), Testik et al. (2011)). Nous évoquerons plus rapidement ces travaux ici.

Plus récemment, une équipe allemande a réalisé une étude basée non plus sur des mesures en laboratoire mais sur des modélisations numériques utilisant un programme de mécanique des fluides modélisant la collision (voir Beheng et al. (2006) et Schlottke et al. (2010)). Ces simulations ont été effectuées sur 32 paires de gouttes (Schlottke et al. (2010) et Straub et al. (2010)). Ces paires de gouttes reprennent les 10 étudiées en laboratoire par Low and List (1982a) et couvrent tout le spectre en diamètre (0.35 à 1.8mm pour les petites ; 0.6 à 4.6mm pour les grosses). Par ailleurs, pour chaque paire de gouttes, les simulations ont été réalisées avec différentes valeurs d'excentricités (excentricité = 0 si les centres des deux gouttes sont alignés sur la verticale ; = 1 si les gouttes ne se touchent que par le bord). Ces excentricités sont ensuite prises en compte pour la paramétrisation de la fonction de break-up. Nous détaillerons également cette paramétrisation ici.

Notations Comme précédemment, nous considérerons deux gouttes de masses, diamètres et vitesses de chute $m_1 > m_2$, $D_1 > D_2$, $v_1 > v_2$ et réutiliserons les définitions données pour l'énergie cinétique de collision CKE (Eq.(A.4)) et pour le nombre de Weber We^* (Eq.(A.8)). Par ailleurs, comme Straub et al. (2010), nous introduisons l'énergie CW :

$$CW = CKE.We^* \tag{A.16}$$

Et d'après Low and List (1982b), nous définissons un troisième nombre de Weber, Web, par :

$$Web = \frac{CKE}{S_{1+2}} \tag{A.17}$$

Avec S_{1+2} la tension de surface avant contact (Eq.(A.6)).

Les distributions en diamètre des gouttes de pluie satellites seront notées ici $\chi^i(D)$ avec D

le diamètre (ou $\chi^i(m)$, les gouttes étant sphériques, on a une équivalence masse-diamètre). La dépendance à la masse des deux gouttes d'origine (m_1 et m_2) est ainsi implicite. i est un indice (lettre) indiquant la paramétrisation considérée (par exemple l = Low and List (1982b)...). $\chi^i(D)$ est en m^{-1} et $\chi^i(D)dD$ représente le nombre de gouttes satellites dans l'intervalle $[D, D + dD]$.

Types de break-up De nombreuses études (McTaggart-Cowan and List (1975), Low and List (1982a), McFarquhar (2004)) séparent le break-up en trois catégories : feuille (sheet), filament et disque (disc).

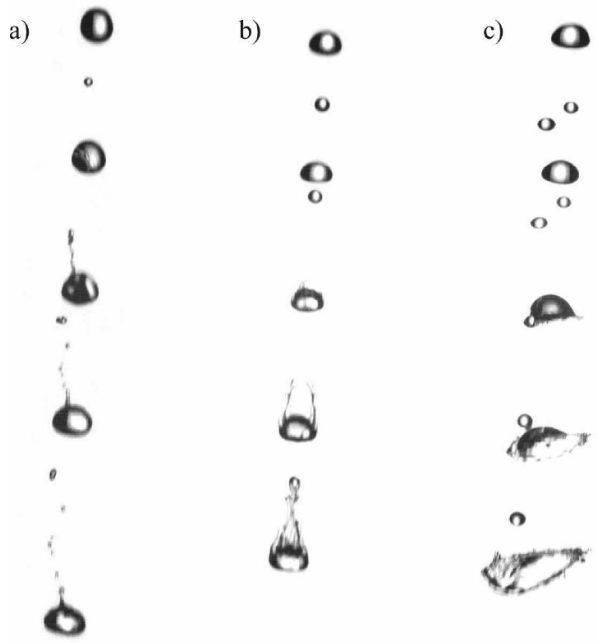

Figure A.4 – Figure et légende extraite de Barros et al. (2008). Séquence d'une caméra à grande vitesse montrant différents types de break-up : (a) - filament ; (b) - sheet ; (c) - disc. Le temps entre les images est de $2ms$.

La Figure A.4 définit visuellement ces trois types de break-up en montrant des photographies de collisions entre gouttes. La paramétrisation de la fonction de break-up peut ensuite se faire en utilisant cette distinction entre les différents types de break-up (Low and List (1982b), McFarquhar (2004), Testik et al. (2011)) ou sans les utiliser (Straub et al. (2010)).

Paramétrisation de Low and List (1982b) La distribution des gouttes de pluie satellites proposée par Low and List (1982b) sera dans la suite nommée $\chi^l(D)$.

Lorsqu'on développe une paramétrisation dépendant du type de rupture (filament / feuille / disque), il convient de définir une première relation donnant, en fonction de la taille des deux

gouttes impliquées dans la collision, la proportion de collisions débouchant sur chacun des types de rupture. On note R_f^l, R_d^l et R_s^l les proportions de collisions débouchant respectivement sur des ruptures en filament, en disque et en feuille. On a bien sûr : $R_f^l + R_d^l + R_s^l = 1$ (on ne traite pas la coalescence ici). On note ensuite χ_f^l, χ_d^l et χ_s^l les distributions des satellites pour chaque type de rupture. On a alors :

$$\chi^l = R_f^l \chi_f^l + R_d^l \chi_d^l + R_s^l \chi_s^l \tag{A.18}$$

On note ensuite N_f^l, N_d^l et N_s^l les nombres de gouttes pour chacune des trois distributions définies plus haut et f_f^l, f_d^l et f_s^l les distributions de probabilités correspondant (ainsi, on aura par exemple $\int \chi_f^l(D)dD = N_f^l$ et $\int f_f^l(D)dD = 1$. On a donc finalement :

$$\chi^l = R_f^l N_f^l f_f^l + R_d^l N_d^l f_d^l + R_s^l N_s^l f_s^l \tag{A.19}$$

Sous ces conditions, donner une paramétrisation de χ revient à donner une expression pour chacun des 9 termes de l'équation (A.19) ci-dessus.

On définit d'abord les trois paramètres R_f^l, R_d^l et R_s^l. Low and List (1982b) propose :

$$R_f^l = \begin{cases} 1 & \text{si } CKE < CKE_0 \\ 1.11.10^{-4} CKE^{-0.654} & \text{si } CKE > CKE_0 \end{cases} \tag{A.20}$$

Avec l'énergie cinétique critique $CKE_0 = 89.3\mu J$.

$$R_s^l = \begin{cases} 0.685(1 - e^{-1.63(Web - W_0)}) & \text{si } Web > W_0 \\ 0 & \text{si } Web < W_0 \end{cases} \tag{A.21}$$

Avec $W_0 = 0.86$.

R_d^l est enfin définie par rapport aux deux autres paramètres :

$$R_d^l = \begin{cases} 1 - (R_f + R_s) & \text{si } (R_f + R_s) < 1 \\ 0 & \text{si } (R_f + R_s) > 1 \end{cases} \tag{A.22}$$

On note que ces relations supposent que les énergies faibles produisent exclusivement des ruptures en filaments et que les ruptures en disques ne se produisent que pour les énergies les plus fortes.

La Figure A.5, tirée des expérimentations numériques de Schlottke et al. (2010), confirme la logique de cette paramétrisation. On remarque que pour les énergies les plus faibles (carrés et "losanges"), les collisions se traduisent soit en coalescence (lorsque les excentricités sont faibles), soit en ruptures en filaments, tandis qu'avec les énergies plus fortes (ronds), on va de ruptures en filament (lorsque l'excentricité est la plus forte, donc le choc d'autant moins violent a priori) vers des ruptures en disques (excentricité faible = choc frontal). On note

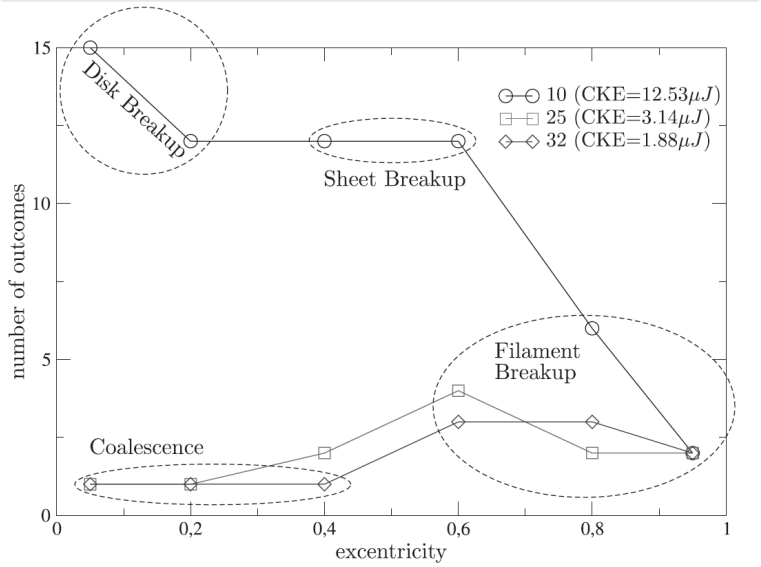

Figure A.5 – Figure extraite de Schlottke et al. (2010). Nombre de satellites produits par une collision en fonction de l'énergie en jeu (traits), de l'excentricité du choc (abscisses) et du type de rupture (ellipses) d'après les simulations numériques de Schlottke et al. (2010)

aussi sur cette figure une grande dépendance du nombre de satellites à l'excentricité, prise en compte dans les études de Schlottke et al. (2010) et Straub et al. (2010) (voir plus bas) mais pas dans Low and List (1982b).

On définit ensuite les trois paramètres N_f^l, N_d^l et N_s^l donnant le nombre de satellites produits selon le type de rupture. Low and List (1982b) propose :

$$N_f^l = \begin{cases} [-2.25.10^4(D_1 - 0.403)^2 - 37.9]D_2^{2.5} + 9.67(D_1 - 0.17)^2 + 4.95 & \text{si } D_2 > D_{S_0} \\ 1.02.10^4 D_2^{2.83} + 2 & \text{si } D_2 < D_{S_0} \end{cases} \quad \text{(A.23)}$$

Avec D_{S_0} un diamètre critique défini de manière à ce que l'expression ci-dessus soit continue. Low and List (1982b) note que le nombre de satellites (dans le cadre de la rupture en filaments) tend vers 2 (c'est-à-dire que 2 gouttes forment 2 gouttes) lorsque l'énergie diminue, c'est-à-dire lorsque $D_2 \longrightarrow 0$ ou lorsque $D_2 \longrightarrow D_1$. (On rappelle que D_1 et D_2 sont les diamètres de la grosse et de la petite goutte)

$$N_s^l = 5.erf\left(\frac{S_{1+2} - 2.53.10^{-6}}{1.85.10^{-6}}\right) + 6 \quad \text{(A.24)}$$

Avec erf la fonction d'erreur. Le nombre de satellites augmente ici avec le diamètre de chacune des gouttes originelles.

$$N_d^l = max\left(297.5 + 23.7ln(CKE), 2\right) \quad \text{(A.25)}$$

Encore une fois, le nombre de satellites augmente avec l'énergie mise en jeu. À faible énergie, 2 gouttes avant collision forment 2 gouttes après collision.

On définit enfin les trois distributions de probabilités pour les satellites : f_f^l, f_d^l et f_s^l.

Pour la rupture en filament, Low and List (1982b) propose une mixture de trois distributions.

La première est une gaussienne tronquée. Elle représente le devenir de la plus grosse des deux gouttes originelles, qui reste plus ou moins intacte après la collision. Cette gaussienne est centrée sur le diamètre D_1 de cette grosse goutte et tronquée à droite pour les diamètres supérieurs à ceux des deux gouttes originelles coalescées (i.e. pour $D > (D_1^3 + D_2^3)^{1/3}$). Le coefficient devant l'exponentielle dans l'expression de la gaussienne est déterminé empiriquement, voir expressions dans Low and List (1982b), équations (4.1) à (4.3).

La deuxième partie de la mixture est une gaussienne. Elle représente le devenir de la plus petite des deux gouttes originelles, elle aussi plus ou moins intacte après la collision. Elle est centrée sur le diamètre D_2 de cette petite goutte, et, de même, le coefficient devant l'exponentielle dans l'expression mathématique de la distribution est défini empiriquement. Voir Low and List (1982b) équations (4.11) à (4.13).

La troisième partie de la mixture est une loi log-normale. Elle représente la distribution des petits satellites créés par la queue du filament (voir Figure A.4a). L'emplacement du pic de cette loi est situé à $0.241D_2 + 0.0129$. Voir Low and List (1982b) équations (4.14) à (4.23).

Pour la rupture en feuille (Figure A.4b), Low and List (1982b) propose une mixture de deux distributions.

La première est définie de la même manière que dans le cas de la rupture en filament, c'est-à-dire par une gaussienne centrée sur D_1 et tronquée pour $D > (D_1^3 + D_2^3)^{1/3}$ (Low and List (1982b) équations (4.25) et (4.26)).

La deuxième goutte n'est plus reconnaissable après le choc et donc la deuxième distribution est, comme la troisième distribution dans le cas filament, une loi log-normale dont l'emplacement du pic est cette fois-ci fonction des deux diamètres originels (Low and List (1982b) équations (4.27) à (4.32)).

Pour la rupture en disque (Figure A.4c), Low and List (1982b) propose une mixture de deux distributions.

La première est à nouveau une gaussienne, mais elle n'est plus centrée sur D_1. Le choc étant plus violent, même le plus gros fragment résultant n'est pas le même que la grosse goutte initiale. Le pic de la gaussienne est donc décalé légèrement à gauche de D_1. Mathématiquement, il est placé à une certaine fraction de D_1, fraction fonction du nombre de Weber W^*. Voir Low and List (1982b) équations (4.33) à (4.36).

Les petits fragments sont ensuite distribués comme dans le cas d'une rupture en feuille selon une loi log-normale (avec des paramètres différents mais du même type). Voir Low and List

(1982b) équations (4.37) à (4.39).

Adaptations de Low and List (1982b) Il a rapidement été remarqué que la paramétrisation de Low and List ne conservait généralement pas la masse d'eau entre avant et après la collision (Valdez and Young (1985)). Le même article remarque d'autres erreurs (proportion de fragments filaments et feuilles $R_s^l + R_f^l$ possiblement supérieure à 1 ; pic de la distribution log-normale pour le break-up en filaments possiblement négatif...) et y remédie par des ajustements locaux (c'est-à-dire sans remettre en cause le schéma global de Low and List (1982b) et les types de distribution utilisées dans les mixtures). Ce type de corrections est ensuite utilisé dans d'autres études (Hu and Srivastava (1995) par exemple).

Tout en restant dans un cadre proche de celui de Low and List (1982b) et en se basant toujours sur les mesures en laboratoire de Low and List (1982a), McFarquhar (2004) a ensuite proposé une nouvelle paramétrisation pour les distributions des fragments pour chaque type de break-up. A savoir : pour les ruptures en feuilles et en disques, mixture d'une loi log-normale (pour les plus petites gouttes) et d'une loi normale (pour ce qui reste de la plus grosse goutte). Pour les ruptures en filaments, mixture d'une loi log-normale (pour les plus petites gouttes), d'une gaussienne pour ce qui reste de la petite goutte originelle et d'un Dirac pour la grosse goutte originelle. Il ajoute à cela des équations de fermeture diminuant le nombre de paramètres à ajuster aux données expérimentales. Ces équations de fermeture incluent la conservation de la masse et du nombre de gouttes (par exemple, les distributions gaussiennes modélisant le devenir de la grosse goutte après collision doivent donner 1 goutte). Voir McFarquhar (2004) pour l'expression analytique de la paramétrisation proposée. Nous reviendrons rapidement dans le paragraphe suivant sur l'effet de ces modifications sur les DSD stationnaires obtenues via l'EDP (II.5) (chapitre II).

Par ailleurs, plusieurs papiers (Testik (2009), Testik et al. (2011), Prat et al. (2012)) développent une classification des résultats des collisions en fonction du nombre de Weber et du ratio des diamètres des deux gouttes originelles, à partir de simulations numériques (Prat et al. (2007), Prat and Barros (2007)) et de mesures en laboratoire (décrites dans Barros et al. (2008) et Testik (2009)). Elles incluent également le break-up en "rebond", correspondant à une collision entre deux gouttes n'affectant finalement aucune des deux gouttes (obtenue en cas de faible nombre de Weber et non prise en compte par Low and List (1982b)).

La figure A.6 présente, en fonction du nombre de Weber et du ratio de la taille des gouttes, des résultats expérimentaux sur les différents types de collisions et de break-up consécutifs, et inclut des frontières entre les différents types de collisions basées sur ces résultats expérimentaux. Notons que l'on trouve le même type d'étude (en fonction du nombre de Weber) dans des travaux sur les nuages (par exemple Orme (1997), figure 7) avec naturellement des gouttes plus petites (donc des We plus petits) impliquant principalement rebond et coalescence.

Prat et al. (2012) paramétrise ensuite à partir de ces études les différents paramètres de

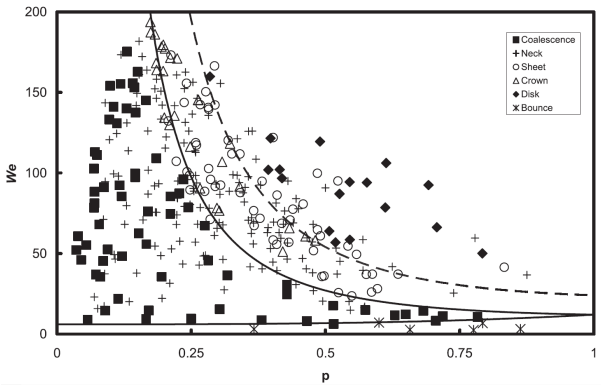

Figure A.6 – Figure extraite de Testik et al. (2011). Graphe des devenir des gouttes après collisions en fonction du nombre de Weber We et du ratio de taille des gouttes $p = D_2/D_1$ à partir d'expérimentations présentées dans Testik et al. (2011). "Neck" break-up est synonyme de rupture en filament. "Crown" break-up est un type de rupture en disque dans le cas d'une plus faible énergie en jeu (Testik et al. (2011))

break-up évoqués plus haut, et compare sa paramétrisation (avec ou sans prise en compte du rebond) à celle de Low and List (1982b).

Paramétrisation de Straub et al. (2010) Nous allons à présent présenter la paramétrisation proposée par Straub et al. (2010) à partir des expérimentations et analyses proposées dans les articles précédents de la même équipe (Beheng et al. (2006) et Schlottke et al. (2010)). On rappelle que ces travaux ne sont plus basés sur des mesures expérimentales en laboratoire mais sur des simulations numériques de collisions entre gouttes.

Contrairement à Low and List (1982a), Straub et al. (2010) ne distingue pas les différents types de break-up (filaments, feuilles, disques) et propose une distribution des gouttes satellites plus générale.

La distribution des gouttes de pluie satellites proposée par Straub et al. (2010) sera dans la suite notée $\chi^b(D)$. Dans ce cadre, D sera en cm et donc χ^b en cm^{-1}. La distribution proposée est une mixture de 4 distributions, chacune représentant la distribution des gouttes satellites sur une classe de diamètres. On a donc pour la distribution sur l'ensemble du spectre :

$$\chi^b(D) = \sum_{i=1}^{4} \chi_i^b(D) \tag{A.26}$$

Par ailleurs, pour chacune des 4 distributions, on peut écrire comme précédemment :

$$\chi_i^b(D) = N_i^b f_i^b(D) \tag{A.27}$$

Avec N_i^b le nombre de gouttes satellites impliquées par la i-ème distribution (qui correspond, comme on va le voir, au nombre de fragments sur un certain intervalle de diamètres) et $f_i^b(D)$ la distribution de probabilités correspondant (avec donc $\int_0^\infty f_i^b(D)dD = 1$).

Les trois premières composantes de la mixture ($i = 1$ à 3) sont des lois normales ou log-normales (voir plus bas les détails), dépendant donc de deux paramètres, par exemple leur moyenne et leur variance. Pour ces trois composantes, la distribution χ_i^b est donc définie par trois paramètres : le nombre de gouttes, l'espérance et la variance. Straub et al. (2010) propose des espérances fixes ou fonctions de D_2 (diamètre de la petite goutte), à savoir $0.4mm$ pour la première composante, $0.95mm$ pour la deuxième et $0.9D_2$ pour la troisième. Les deux premières composantes de la mixture représentent donc les petits satellites créés par la collision tandis que la troisième représente les satellites de l'ordre de grandeur de la plus petite des deux gouttes impliquées. La variance des distributions et le nombre de gouttes sont quant à eux exprimés en fonction de l'énergie mise en jeu lors de la collision (voir les détails plus bas), énergie, comme on l'a vu, fonction elle-même de la taille et de la vitesse des deux gouttes impliquées.

La dernière composante de la mixture ($i = 4$) est un Dirac (création d'une goutte à diamètre fixe) placé là où il faut pour qu'il y ait conservation de la masse d'eau (avant et après break-up) en fonction du diamètre des deux gouttes originelles et des trois autres composantes de la mixture de distributions. On verra que ce Dirac se situe généralement légèrement à gauche sur le spectre des diamètres par rapport à la grosse goutte initiale, qui perd donc une petite partie de sa masse durant le choc.

Nous allons à présent présenter les équations correspondant à la paramétrisation de Straub et al. (2010) du break-up. Nous verrons successivement la forme de chacune des quatre composantes de la mixture de distribution proposée.

Première composante

La première distribution de la mixture s'écrit donc $\chi_1^b(D) = N_1^b f_1^b(D)$ et la distribution de probabilités f_1^b suit une loi log-normale qui s'écrit :

$$f_1^b(D) = \frac{1}{D\sigma_1^b\sqrt{2\pi}} e^{-\frac{(ln(D)-\mu_1^b)^2}{2\left(\sigma_1^b\right)^2}} \tag{A.28}$$

Avec $\mathbb{E}$ l'espérance et Var la variance, on a :

$$\left(\sigma_1^b\right)^2 = ln\left(\frac{Var}{\mathbb{E}^2} + 1\right) \tag{A.29}$$

$$\mu_1 = ln(\mathbb{E}) - \frac{\left(\sigma_1^b\right)^2}{2} \tag{A.30}$$

La paramétrisation proposée par Straub et al. (2010) est ensuite :

$$\mathbb{E} = 0.4mm \tag{A.31}$$

$$Var = \frac{0.0125^2}{12} CW \tag{A.32}$$

$$N_1^b = 0.088(\frac{D_1}{D_2}CW - 7) \text{ si } \frac{D_1}{D_2}CW > 7\mu J \text{ (0 sinon)} \tag{A.33}$$

Avec CW l'énergie définie plus haut (Eq.(A.16)).

Deuxième composante

La deuxième distribution de la mixture s'écrit également $\chi_2^b(D) = N_2^b f_2^b(D)$ et la distribution de probabilités f_2^b suit une loi normale qui s'écrit :

$$f_2^b(D) = \frac{1}{\sigma_2^b\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{D-\mu_2^b}{\sigma_2^b}\right)^2} \tag{A.34}$$

Avec μ_2^b l'espérance et $\left(\sigma_2^b\right)^2$ la variance. La paramétrisation est ensuite :

$$\mu_2^b = 0.4mm \tag{A.35}$$

$$\left(\sigma_2^b\right)^2 = \frac{0.007^2}{12}(CW - 21)^2 \text{ si } CW > 21\mu J \text{ (0 sinon)} \tag{A.36}$$

$$N_2^b = 0.22(\frac{D_1}{D_2}CW - 21) \text{ si } CW > 21\mu J \text{ (0 sinon)} \tag{A.37}$$

Troisième composante

La troisième distribution de la mixture s'écrit également $\chi_3^b(D) = N_3^b f_3^b(D)$ et la distribution f_3^b suit une loi normale qui s'écrit :

$$f_3^b(D) = \frac{1}{\sigma_3^b\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{D-\mu_3^b}{\sigma_3^b}\right)^2} \tag{A.38}$$

Avec μ_3^b l'espérance et $\left(\sigma_3^b\right)^2$ la variance. La paramétrisation est ensuite :

$$\mu_3^b = 0.9D_1 \tag{A.39}$$

$$\left(\sigma_3^b\right)^2 = \frac{0.01^2}{12}(1 + 0.76\sqrt{CW})^2 \tag{A.40}$$

$$N_3^b = \begin{cases} 1 \text{ si} & CW < 21\mu J \\ 0.04(46 - CW) & \text{si } 21 < CW < 46\mu J \\ 0 \text{ si} & CW > 46\mu J \end{cases} \tag{A.41}$$

Quatrième composante

La quatrième distribution, comme on l'a dit, consiste en un Dirac placé sur un certain diamètre critique D_4. Elle s'écrit donc :

$$\chi_4^b(D) = N_4^b \delta(D - D_4) \tag{A.42}$$

En suivant leurs expérimentations, Straub et al. (2010) fixe $N_4^b = 1$. Il reste donc à déterminer D_4. Or, le système total doit conserver la masse d'eau, ce qui fournit une relation définissant notre dernier paramètre. La masse de la i^{eme} composante de la mixture de DSD s'écrit, via un coefficient de proportionnalité, comme le moment d'ordre 3 de la distribution des gouttes, à savoir :

$$\mathcal{M}_i^3 = \frac{\pi}{6} \int_0^\infty D^3 \chi_i^b(D) dD \tag{A.43}$$

Connaissant les trois premières distribution, on connaît $\mathcal{M}_1^3$, $\mathcal{M}_2^3$ et $\mathcal{M}_3^3$ (Straub et al. (2010) donne leurs expressions analytiques). Pour la quatrième distribution, on a :

$$\mathcal{M}_4^3 = \frac{\pi}{6} D_4^3 \tag{A.44}$$

La masse initiale $m_1 + m_2$ étant la somme de la masse des deux gouttes :

$$m_1 + m_2 = \frac{\pi}{6}(D_1^3 + D_2^3) \tag{A.45}$$

Oon a finalement D_4 via l'équation :

$$D_4^3 = \frac{\pi}{6}(D_1^3 + D_2^3) - \mathcal{M}_1^3 - \mathcal{M}_2^3 - \mathcal{M}_3^3 \tag{A.46}$$

Conclusion

Straub et al. (2010) présente, pour chacune des 32 paires de gouttes dont ils ont simulé la collision, des graphes des DSD résultantes mesurées et données par la paramétrisation ci-dessus, et compare cela avec la paramétrisation de Low and List (1982a).

Pour Straub et al. (2010), cette paramétrisation présente de nombreuses similarités avec celle de Low and List (1982a) mais l'améliore sur trois points notamment :

1. Conservation de la masse.
2. Développement sur un plus grand nombre de paires de diamètres (32 vs. 10) couvrant une plus large part du spectre total.
3. Développement d'un modèle de DSD ne dépendant pas du type de break-up (sheet, filament, disk).

Annexe B: Auto break-up

Dans le chapitre II, nous revenons sur le modèle d'équations aux dérivées partielles (EDP) développé pour représenter l'évolution spatiotemporelle de la DSD lors de la chute des gouttes. Ce modèle permet de représenter l'effet des collisions entre gouttes mais aussi celui de leur désintégration sous l'effet de leur propre vitesse de chute. Dans cette Annexe, nous revenons sur les études réalisées à ce sujet. On note que ce phénomène (comme d'ailleurs l'effet des collisions entre gouttes) n'est pas pris en compte dans le travail présenté dans cette thèse. Il permet néanmoins d'illustrer la complexité du phénomène et d'anticiper d'éventuelles évolutions futures de ce travail.

Une goutte seule peut donc également exploser du fait d'une trop grande vitesse par rapport à sa tension superficielle. On peut alors parler d'auto break-up. Un tel phénomène a déjà été étudié et paramétrisé empiriquement, notamment dans Srivastava (1971). En travaillant empiriquement, il s'agit de donner une probabilité de désintégration (dans un temps donné et selon le diamètre ou la masse de la goutte considérée) et, en cas de désintégration, une distribution pour les satellites. Srivastava (1971) propose ainsi une probabilité de désintégration augmentant exponentiellement avec le diamètre de la goutte, et une distribution exponentielle pour les satellites, et intègre cette expression dans l'EDP (II.5) (voir chapitre II) en prenant seulement en compte la chute des gouttes et la coalescence.

Plus récemment, Villermaux and Bossa (2009) a proposé une étude donnant d'une part une base plus physique au problème de l'auto break-up, et affirmant d'autre part que ce phénomène était prépondérant (par rapport à la coalescence et au break-up provoqués par des collisions entre gouttes) et déterminait la forme des distributions. Il affirme enfin que ce phénomène a lieu à la base des nuages et que les distributions restent ensuite essentiellement stationnaires durant toute la chute des gouttes.

Villermaux and Bossa (2009) donne ainsi l'équation différentielle (non justifiée ici mais basée sur l'étude des différences de pression exercées par l'air sur la goutte) exprimant l'évolution du diamètre D_g d'une goutte en chute libre dans l'air. Cette goutte n'est pas supposée sphérique (puisque l'on modélise justement ses déformations) mais en forme de "pancake" (cylindre défini par un diamètre et une hauteur). Cette équation différentielle est fonction : 1) d'une constante de temps τ_g, fonction elle-même du diamètre équivalent sphérique de la goutte (caractérisant sa masse) et de sa vitesse terminale de chute ; 2) du nombre de Weber We (voir Eq.(A.9) dans

l'Annexe A). Cette équation différentielle s'écrit :

$$D_g{}'' = \frac{1}{\tau_g^2}\left(1 - \frac{6}{We}\right) D_g \tag{B.1}$$

On reconnaît une équation différentielle du second ordre à coefficients constants. Deux cas de figure peuvent se présenter : 1) $(1 - 6/We) < 1$, c'est-à-dire $We < 6$, alors la solution consiste en des oscillations autour d'un rayon moyen et la goutte est stable, elle n'explose pas (ce qui est cohérent physiquement : un nombre de Weber faible implique une tension de surface suffisamment forte par rapport à la gravité pour empêcher l'explosion). 2) $We > 6$: la solution est exponentielle : le rayon augmente exponentiellement et la goutte explose dans un temps caractéristique τ_{burst} donné analytiquement dans Villermaux and Bossa (2009). Constatant τ_{burst} très petit devant le temps caractéristique de collisions entre gouttes, Villermaux and Bossa (2009) conclut à une importance prépondérante de l'auto break-up face aux collisions.

Le travail de Villermaux and Bossa (2009) a été défendu (Kostinski and Shaw (2009), Villermaux and Eloi (2011), Larsen et al. (2014)) ou critiqué (Barros et al. (2010), Mcfarquhar (2010)), notamment sur le fait qu'il travaille sur un modèle théorique dans des conditions idéales ne tenant pas compte de phénomènes naturels potentiellement importants (turbulence qui peut augmenter significativement le nombre volumique de gouttes, augmentant ainsi les probabilités de collision).

Sur la distribution des tailles des gouttes satellites crées en cas d'auto break-up, le problème est différent du cas d'une collision (pas de "restes" des deux gouttes initiales) et les distributions résultantes aussi. Il existe une littérature assez conséquente sur le sujet. On cite simplement ici Villermaux and Eloi (2011), qui propose une distribution gamma pour les fragments, paramétrée en fonction du diamètre de la goutte initiale. Jalaal and Mehravaran (2012) réalise lui une étude de l'auto break-up d'une goutte via un modèle numérique et compare ses résultats à une loi log-normale (sans proposer de paramétrisation formelle).

Annexe C: Discussion sur la paramétrisation de la vitesse terminale de chute des gouttes.

Le travail présenté au chapitre V utilise des relations liant la vitesse terminale de chute des gouttes de pluie à leur diamètre (et à l'altitude). Dans le chapitre II, section 3, nous avons présenté et rapidement justifié la relation vitesse/diamètre utilisée dans ce travail. Dans cette annexe, nous présentons une étude bibliographique des nombreuses relations développées dans la littérature.

Nous cherchons donc ici une expression pour la vitesse terminale de chute v_{th}. Toutes les notations utilisées ici sont celles du chapitre II.

La section 1 tentera d'expliciter la fonction v_{th} à partir de considérations physiques. La section 2 explorera les différentes relations empiriques proposées dans la littérature pour v_{th}. La section 3, enfin, étudiera la dépendance à l'altitude de v_{th}.

1 Un peu de physique

Dans cette section, on suppose qu'il n'y a ni vent, ni turbulence, que les gouttes tombent dans l'air et que celui-ci a une masse volumique constante (les forces exercées sur la goutte qui chute ne changent pas).

Une goutte lâchée dans l'air atteint sa vitesse terminale de chute lorsqu'est atteint l'équilibre entre les forces de gravité F_{grav} et de traînée F_x (résistance de l'air). La première s'écrit :

$$F_{grav} = m_g g = \rho_e V_g g \tag{C.1}$$

Avec m_g la masse de la goutte (kg), V_g son volume (m^3), ρ_e la masse volumique de l'eau (kg/m^3) et $g \approx 9.81 m/s^2$ l'accélération de la pesanteur.

La deuxième s'écrit :

$$F_x = \frac{1}{2}\rho_a v_g^2 S_g C_D \tag{C.2}$$

Avec ρ_a la masse volumique de l'air (kg/m^3), v_g la vitesse de la goutte (m/s), S_g sa section horizontale (m^2) (pour une sphère $S_g = \frac{\pi}{4} D_g^2$) et C_D un coefficient de frottement (sans unité)

fonction de sa géométrie et éventuellement de sa vitesse (voir plus bas). On peut déjà noter que plus C_D est grand, plus la traînée, la résistance de l'air, est forte (et donc plus la goutte est freinée).

A l'équilibre, on a $F_{grav} = F_x$. Dans le cas le plus général, cela conduit à :

$$v_g^2 = \frac{2\rho_e}{C_D \rho_a} \frac{V_g}{S_g} g \tag{C.3}$$

Si l'on suppose que la goutte est une sphère, on obtient à partir de la formule précédente, la relation vitesse/diamètre :

$$v_{g,sphère}^2 = \frac{4}{3C_D} \frac{\rho_e}{\rho_a} g D_g \tag{C.4}$$

Avec toujours D_g le diamètre de la goutte. On retrouve notamment cette expression dans Villermaux and Eloi (2011).

Pour une goutte en forme de "pancake" (forme générée par un cercle intégré sur une hauteur h_g), on trouve, avec h_g cette hauteur :

$$v_{g,pancake}^2 = \frac{2}{C_D} \frac{\rho_e}{\rho_a} g h_g \tag{C.5}$$

Villermaux and Eloi (2011) donne pour C_D des valeurs d'environ 0.5 dans le cas d'une sphère et 1 dans le cas d'un "pancake". Il précise également que ces valeurs sont valables à suffisamment grand nombre de Reynolds Re (i.e. $Re > 1000$), sans quoi il convient de les modifier pour prendre en compte les effets de la viscosité de l'air (qui tend à ralentir les gouttes). Le nombre de Reynolds évalue le rapport des forces de gravité sur celles de viscosité et peut s'écrire :

$$Re = \frac{v_g D_g}{\nu_a} \tag{C.6}$$

Avec ν_a la viscosité cinématique de l'air ($15.6.10^{-6} m^2/s$ à $25\,^\circ C$) (Villermaux and Eloi (2011)).

On note que pour des diamètres ($1mm$) et des vitesses ($4m/s$) intermédiaires, on trouve avec l'expression précédente $Re = 256$, une valeur assez sensiblement inférieure à 1000. Cette hypothèse ($Re > 1000$) ne serait donc valable que pour d'assez grosses gouttes (typiquement $> 2mm$). Pour une sphère et pour des Reynolds inférieurs à 1000, on trouve d'autres relations pour exprimer le C_D. Par exemple pour la gamme $1 < Re < 1000$ (Seinfeld and Pandis (2012)) :

$$C_D = \frac{18.5}{Re^{0.6}} \tag{C.7}$$

Ou encore l'expression pour $Re < 1000$ (http://nte.mines-albi.fr/CristalGemme/co/uc_Trainee.html) :

$$C_D = \frac{24}{Re}(1 + 0.15 Re^{0.687}) \tag{C.8}$$

Ces deux relations donnent respectivement $C_D = 1.52$ et 1.35 pour des couples vitesse/diamètres de $(0.5mm - 2m/s)$; 0.66 et 0.73 pour $(1mm - 4m/s)$; 0.33 et 0.47 pour $(2mm - 6.5m/s)$. Logiquement, plus le nombre de Reynolds diminue, plus les forces de viscosité deviennent importantes par rapport à celles de gravité, plus les frottements sont importants (c'est-à-dire plus le C_D augmente) et plus la goutte est freinée.

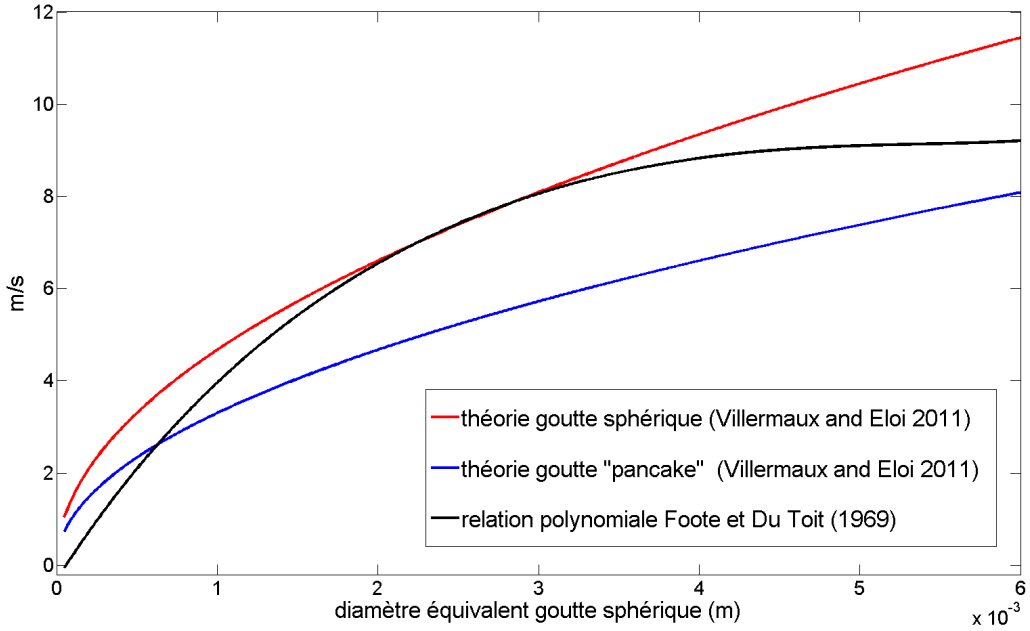

Figure C.1 – Vitesses terminales de chute de gouttes de pluie. rouge) relation (C.4), $C_D = 0.5$; bleu) relation (C.5), $C_D = 1$ (le diamètre du pancake est défini comme celui de la sphère de même volume) ; noir) relation de Foote and Du Toit (1969) approchant les mesures de Gunn and Kinzer (1949) avec une expression polynomiale (degré 5 ici)

La Figure C.1 présente un graphe avec, pour des diamètres entre 0 et $6mm$, les vitesses terminales de chute estimées, pour des gouttes sphériques, par la relation (C.4), avec un coefficient de traînée de 0.5 (rouge) ; et pour des gouttes en forme de pancake par la relation (C.5), avec un coefficient de traînée de 1 (bleu). On calcule la hauteur h_g des pancakes par la relation $h_g = 4V_g/\pi/D_g^2$, avec V_g le volume de la sphère de diamètre D_g. La Figure C.1 présente aussi (courbe noire) les vitesses de chute estimées par une formule analytique proposée par Foote and Du Toit (1969) pour approcher les mesures réalisées par Gunn and Kinzer (1949) (polynôme à l'ordre 5 en D, voir plus bas).

On remarque sur cette figure que la formule pour une sphère (rouge) semble plus appropriée, par rapport aux mesures de Gunn and Kinzer (1949), pour les diamètres intermédiaires $(2 - 4mm)$ et que la formule des pancakes (bleu) semble meilleure ensuite, ce qui semble cohérent avec le fait que les gouttes se déforment et s'éloignent d'une sphère d'autant plus que leur taille, et donc leur vitesse, augmente. A petit diamètre, les vitesses mesurées par Gunn and Kinzer (1949) sont assez sensiblement plus faibles que celles pour une goutte sphérique sans viscosité (courbe rouge), ce qui semble là encore cohérent avec ce qu'on

a vu plus haut : pour des petits diamètres, on doit prendre en compte l'effet de la viscosité, ce qui revient à augmenter le coefficient de traînée C_D de l'équation (C.3) et donc à diminuer la vitesse de chute. On peut donc en quelques sorte diviser la gamme des diamètres en 3 parties : sphère avec viscosité, sphères sans viscosité et pancakes.

La difficulté à modéliser la forme de la goutte et à exprimer l'influence de la viscosité ont cependant conduit à chercher des relations vitesses/diamètres empiriques. Une large littérature existe sur le sujet, explorée dans la section qui suit.

2 Relations empiriques

Dans cette section, nous allons donner quelques unes des paramétrisations empiriques développées pour exprimer la vitesse terminale de chute des gouttes au niveau du sol. On peut tout d'abord noter qu'un grand nombre de ces relations prennent pour référence les mesures réalisées par Gunn and Kinzer (1949).

On peut citer tout d'abord la relation proposée par Best (1950) (citée par exemple dans Foote and Du Toit (1969)) :

$$v_{th,best}(D_g) = 9.43\left(1 - e^{-\left(\frac{D}{1.77}\right)^{1.147}}\right) \tag{C.9}$$

Avec D_g en mm. Cette relation est notamment utilisée dans les expériences déjà mentionnées sur le break-up de Low and List (1982a) et Low and List (1982b).

Plusieurs études utilisent également les relations proposées par Beard (1976) et décomposant le spectre de diamètre en 3 classes (avec une relation analytique par classe), classes basées sur le nombre de Reynolds et les phénomènes prépondérants qu'il implique, et couvrant la gamme $(0.5\mu m - 7mm)$ (voir les expressions dans Beard (1976)). L'étude de Valdez and Young (1985) ou le travail plus récent sur la paramétrisation numérique du break-up (Beheng et al. (2006), Schlottke et al. (2010), Straub et al. (2010)) utilisent ces relations.

On peut aussi citer l'expression largement utilisée proposée par Atlas et al. (1973). Cette relation est celle utilisée dans cette thèse. Comme on l'a vu au chapitre II, section 3, elle conduit à des écarts inférieurs à $\pm 2\%$ par rapport aux mesures de Gunn and Kinzer (1949) sur la gamme de diamètre $(0.6 - 5.8mm)$:

$$v_{atlas}(D_g) = 9.65 - 10.3e^{-6D_g} \tag{C.10}$$

Avec D_g en cm. Les restitutions de DSD à partir de mesures radars en UHF de Williams et al. (2000) utilisent cette relation, ainsi que l'étude de Tridon and Battaglia (2015) sur les

restitutions de vent et de DSD à partir de spectres Dopper en bande Ka et W, avec une correction pour l'altitude, et pour les gouttes de diamètre au-dessus de $0.86mm$.

Atlas and Ulbrich (1977) a simplifié cette relation, pour la gamme de diamètres $[0.5 : 5]mm$, sous la forme :

$$v_{atlas2}(D_g) = 17.67 D_g^{0.67} \tag{C.11}$$

Avec D_g en cm. On note qu'en écrivant la vitesse de chute des gouttes comme proportionnelle à une puissance de leur diamètre, on obtient le taux de pluie comme un moment de la DSD, en l'occurrence le moment d'ordre 3.67. L'étude réalisée par Lavergnat and Golé (1998) utilise cette expression.

Avant cela, Foote and Du Toit (1969) avait proposé une relation polynomiale de la forme :

$$v_{foote}(D_g) = \sum_{j}^{j_{max}} (a_j D_g^j) \tag{C.12}$$

Foote and Du Toit (1969) fournit les coefficients a_j et l'erreur maximale par rapport aux données de Gunn and Kinzer (1949) sur la gamme de diamètre $[0.1 - 5.8mm]$ en fonction de j_{max}, le degré du polynôme (par exemple $0.07m/s$ pour $j_{max} = 5$).

Citons enfin l'expression utilisée par Lhermitte (1988), qui s'écrit :

$$v_{lhermitte}(D_g) = 9.25 \left(1 - e^{-6.8 D_g^2 - 4.88 D_g}\right) \tag{C.13}$$

Avec D_g en cm. Cette relation a été utilisée dans plusieurs études restituant vent vertical et DSD à partir de spectres Doppler en bande W, notamment Lhermitte (1988) ou Giangrande et al. (2010).

La Figure C.2 représente ces différentes relations pour les diamètres couverts par la pluie. On voit que dans leurs domaines de validité (voir texte plus haut), les relations donnent des résultats très similaires. On veillera cependant à ne pas utiliser l'expression d'Atlas (Eq.(C.10)) pour de trop petites gouttes (elle donne des vitesses négatives).

3 Vitesse terminale de chute : dépendance à l'altitude

On a vu (section 1 de cette Annexe) que la force de traînée freinait les gouttes d'eau, et qu'elle dépendait de la masse volumique de l'air (Eq.(C.2)). Or celle-ci dépend des conditions de pression et de température, qui elles-même varient fortement avec l'altitude. On comprend donc que le freinage subi par les gouttes dépendra de l'altitude. La pression atmosphérique baissant lorsqu'on s'élève dans l'atmosphère, on peut supposer que la traînée baissera également, et que donc la vitesse terminale de chute des gouttes augmentera.

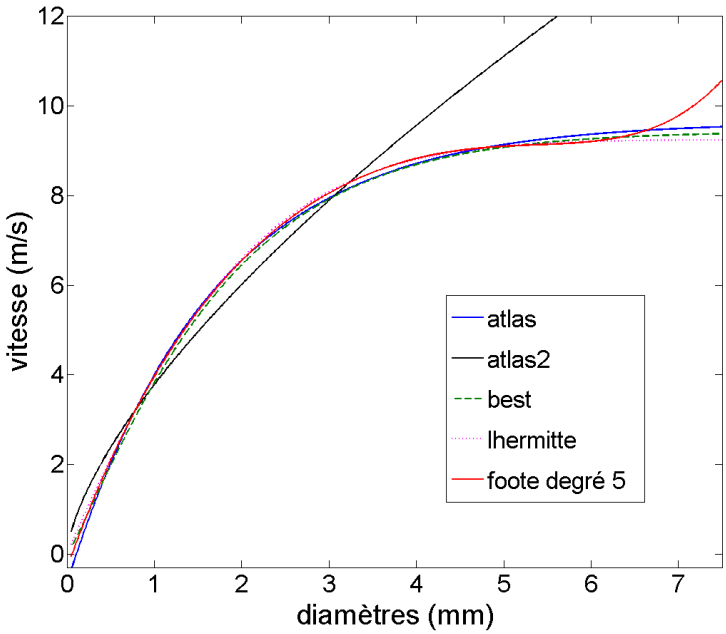

Figure C.2 – Diverses relations vitesse-diamètre appliquées à des gouttes sphériques sur la gamme de diamètres $[0 : 7.5mm]$. Les relations utilisées sont décrites dans le texte par : bleu) Eq.(C.10) ; noir) Eq.(C.11) ; vert) Eq.(C.9) ; magenta) Eq.(C.13) ; rouge) Eq.(C.12).

Best (1950), le premier, fournit une relation empirique estimant la vitesse terminale corrigée par l'altitude (notée avec une étoile, ici v^*_{best}) :

$$v^*_{best}(D_g, z) = v_{best}(D_g)e^{bz} \tag{C.14}$$

Avec z l'altitude (m) et b une constante dépendant du type d'atmosphère choisi (par exemple 0.0405 pour un atmosphère américaine standard, ou 0.0354 pour une atmosphère tropicale, voir Foote and Du Toit (1969)).

En utilisant l'Eq.(C.4), on obtient cependant que la vitesse terminale de chute d'une goutte sphérique est proportionnelle à $1/\sqrt{\rho_a}$, avec ρ_a la masse volumique de l'air. On obtient alors que la vitesse à l'altitude z s'écrit :

$$v^*_{g,sphere}(D_g, z) = v_{g,sphere}(D_g)\left(\frac{\rho_a(z=0)}{\rho_a(z)}\right)^{0.5} \tag{C.15}$$

Plusieurs études proposent des relations empiriques mais basées sur ces considérations. Foote and Du Toit (1969) propose ainsi une relation plus complexe (voir le papier), dépendant directement de la masse volumique de l'air et de la température, et qui peut éventuellement

être simplifiée par :

$$v^*_{foote,simple}(D_g, z) = v_{foote}(D_g) \left(\frac{\rho_a(0)}{\rho_a(z)} \right)^{0.4} \tag{C.16}$$

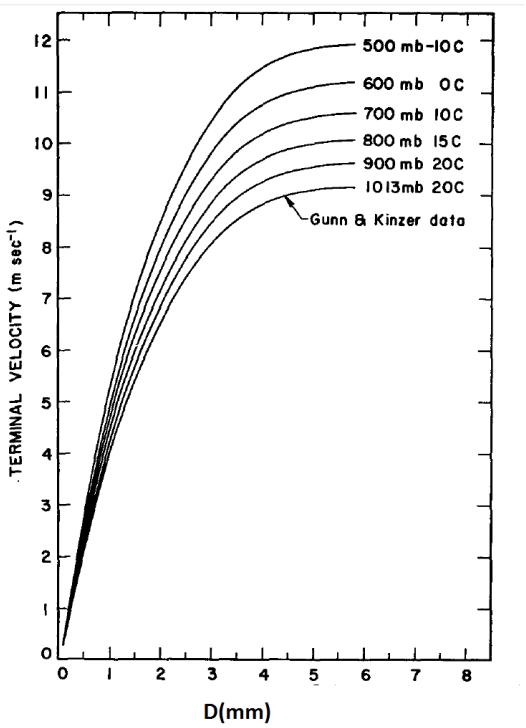

Figure C.3 – Figure extraite de Foote and Du Toit (1969) : vitesse terminale de chute des gouttes telle que mesurée par Gunn and Kinzer (1949) (courbe du bas) et modélisée selon les conditions de pression et de température par Foote and Du Toit (1969).

Cette relation, appliquée à différentes conditions de pression et de température, fournit les courbes vitesse/diamètre représentées sur la Figure C.3. On voit bien que la vitesse de chute augmente à mesure que la pression de l'air diminue. A $800mb$, correspondant à une altitude d'environ $2000m$, les gouttes les plus grosses tombent autour de $10m/s$ (contre $9m/s$ au sol). A $600mb$ (environ $4000m$ d'altitude), elles tombent à presque à $12m/s$.

Beard (1985) enfin préconise une expression de la forme :

$$v^*_{beard}(D_g, z) = v_{beard}(D_g) \left(\frac{\rho_a(0)}{\rho_a(z)} \right)^{m_b} \tag{C.17}$$

Avec m_b un coefficient dépendant du diamètre des gouttes, et variant de 0.4 pour les petites gouttes ($1mm$) à 0.55 pour les plus grosses ($7mm$). Dans ce modèle, les gouttes sont donc, proportionnellement, d'autant plus accélérées avec l'altitude qu'elles sont grosses.

Au niveau des applications, Tridon and Battaglia (2015) utilise la correction "physique" (Eq.(C.15)). Giangrande et al. (2010) se base sur la correction de Beard (1985) (Eq.(C.17)). Hu and Srivastava (1995) utilise lui l'expression de Foote and Du Toit (1969). MRR (1998) adapte

sous forme polynomiale cette même expression pour une atmosphère standard, en la combinant avec la vitesse terminale au sol de Atlas et al. (1973) :

$$v^*_{MRR}(D_g, z) = \left[1 + 3.68.10^{-5} z + 1.71.10^{-9} z^2\right] v_{atlas}(D_g) \tag{C.18}$$

Comme on l'a vu au chapitre II, section 3, dans ce travail, nous utiliserons également cette dernière expression.

Annexe D: Discrétisation des EDP

Le chapitre II, section 4, revient sur les schémas numériques permettant de discrétiser les équations aux dérivées partielles (EDP) d'advection. Il introduit deux de ces schémas, le schéma upstream et le schéma de Smolarkiewicz (Smolarkiewicz (1983)), et compare rapidement leurs performances sur une application simple. Nous avons montré en particulier la forte diffusivité du schéma upstream, diffusivité plus limitée dans le schéma de Smolarkiewicz. Nous avons également donné les conditions de stabilité des deux schémas. Ces schémas ont été présentés pour une dimension d'espace.

Dans cette Annexe, nous justifions la condition de stabilité sur le schéma d'Euler. Ensuite, nous étudions la diffusion numérique qu'il induit. Cette étude nous permet d'introduire plus naturellement le schéma de Smolarkiewicz et de justifier ses équations. Enfin, nous donnons les équations de ces deux schémas pour des problèmes à deux dimensions d'espace. Les notations utilisées dans cette Annexe sont celles du chapitre II, section 4.

1 Stabilité du schéma upstream

Nous rappelons que le schéma upstream permet de propager un champ discrétisé initial $(\tilde{\phi}_i^0)$ à toutes les dates, d'itération temporelle en itération temporelle. Supposons connu le champ à la k-ième itération. L'équation permettant d'évaluant le champ à la $(k+1)$-ième itération s'écrit :

$$\tilde{\phi}_i^{k+1} = \tilde{\phi}_i^k - F\left(\tilde{\phi}_i^k, \tilde{\phi}_{i+1}^k, \tilde{v}_i^k, \Delta x\right) + F\left(\tilde{\phi}_{i-1}^k, \tilde{\phi}_i^k, \tilde{v}_{i-1}^k, \Delta x\right) \tag{D.1}$$

Avec la fonction F définie dans ce cadre par :

$$F(a,b,v,\Delta x) = \begin{cases} \frac{v\Delta t}{\Delta x} a & \text{si } v > 0 \\ \frac{v\Delta t}{\Delta x} b & \text{si } v < 0 \end{cases} \tag{D.2}$$

On a vu (chapitre II, section 4) qu'en se plaçant dans le cas où les vitesses sont positives aux deux bords de la maille considérée ($\tilde{v}_i^k > 0$ et $v_{i-1}^k > 0$), l'Eq.(D.1) se réécrit sous la forme plus simple :

$$\tilde{\phi}_i^{k+1} = \tilde{\phi}_i^k + \tilde{v}_{i-1}^k \frac{\Delta t}{\Delta x} \tilde{\phi}_{i-1}^k - \tilde{v}_i^k \frac{\Delta t}{\Delta x} \tilde{\phi}_i^k \tag{D.3}$$

Cette équation peut encore s'écrire :

$$\frac{\tilde{\phi}_i^{k+1} - \tilde{\phi}_i^k}{\Delta t} + \frac{\tilde{v}_i^k \tilde{\phi}_i^k - \tilde{v}_{i-1}^k \tilde{\phi}_{i-1}^k}{\Delta x} = 0 \tag{D.4}$$

On comprend alors que le schéma upstream est une résolution approchée de l'équation d'advection (Eq.(II.37) du chapitre II) selon la méthode d'Euler en temps et en espace, c'est-à-dire en écrivant, pour une date et un point de l'espace quelconques :

$$\begin{cases} \frac{\partial \phi}{\partial t}(t,x) & \approx \frac{\phi(t+\Delta t)-\phi(t)}{\Delta t} \\ \frac{\partial (v\phi)}{\partial t}(t,x) & \approx \frac{v(t,x+\Delta x)\phi(t,x+\Delta x)-v(t,x)\phi(t,x)}{\Delta x} \end{cases} \tag{D.5}$$

On a vu (chapitre II, section 4) qu'une condition nécessaire et suffisante pour que le schéma upstream soit stable est que la discrétisation vérifie la condition CFL (Courant-Friedrichs-Lewy), qui s'écrit dans ce cadre :

$$\max_{k,i}(\tilde{v}_i^k)\frac{\Delta t}{\Delta x} < 1 \tag{D.6}$$

Nous allons à présent rapidement justifier cette condition pour ce schéma numérique.

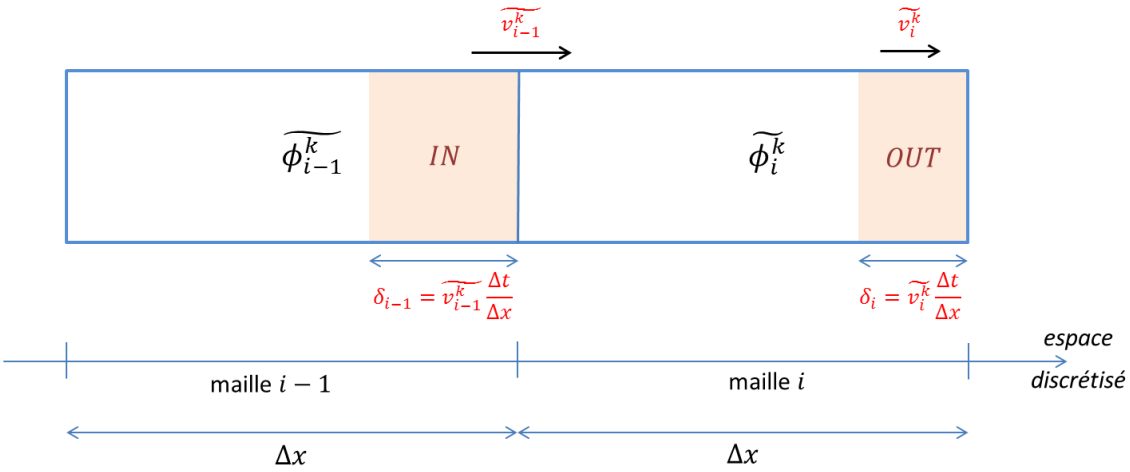

FIGURE D.1 – Représentation du schéma upstream sous forme de flux entrant et sortant des mailles. À la date $k+1$, la quantité OUT va être advectée hors de la maille i et la quantité IN va y entrer. Voir le texte pour les équations complètes.

Le schéma upstream peut facilement être vu, indépendamment de l'équation continue d'advection, comme un problème de flux. Reprenons l'équation (D.3). Elle peut s'écrire :

$$\tilde{\phi}_i^{k+1} = \tilde{\phi}_i^k + IN - OUT \tag{D.7}$$

Avec IN la quantité entrant dans la maille i à la date $k+1$ et OUT la quantité en sortant. IN est alors une fraction δ_{i-1} de la quantité présente dans la maille $i-1$ à la date k, et OUT est

une fraction δ_i de la quantité présente dans la maille i à la date k. On a donc (puisque l'on a supposé les concentrations homogènes dans chaque maille) :

$$\begin{cases} IN & = \delta_{i-1}\tilde{\phi}_{i-1}^k \\ OUT & = \delta_i \tilde{\phi}_i^k \end{cases} \tag{D.8}$$

Et d'après l'Eq.(D.3), les fractions entrantes et sortantes s'écrivent :

$$\delta_i = \tilde{v}_i^k \frac{\Delta t}{\Delta x} \tag{D.9}$$

Toutes ces notions sont résumées de manière plus visuelle sur la Figure D.1.

On comprend alors que si, par exemple, $\delta_{i-1} > 1$, on va faire entrer dans la maille i (et donc faire sortir de la maille $i-1$) plus de matière qu'il n'y en avait de disponible. D'une part cela n'a pas de sens physiquement. D'autre part, si la condition initiale $(\tilde{\phi}_i^1)$ n'est pas monotone, cela crée des oscillations amplifiées. Par exemple, prenons une condition initiale $\left(\tilde{\phi}_{2k}^0 = 0; \tilde{\phi}_{2k+1}^0 = 1\right)$ (alternance de 1 et de 0). Utilisons $\Delta t = 1$, $\Delta x = 1$, $\tilde{v}_i^k = 2$. On a alors : $\left(\tilde{\phi}_{2k}^1 = 2; \tilde{\phi}_{2k+1}^1 = -1\right)$, $\left(\tilde{\phi}_{2k}^2 = -4; \tilde{\phi}_{2k+1}^2 = 5\right)$, $\left(\tilde{\phi}_{2k}^3 = 14; \tilde{\phi}_{2k+1}^3 = -13\right)$, etc. Il conviendra donc dans la suite de bien veiller à respecter l'Eq.(D.6). On admettra par ailleurs les formes que prend la condition CFL pour les autres schémas, et pour d'autres dimensions.

2 Diffusion numérique

Nous allons dans cette partie évoquer le principal désavantage du schéma upstream, à savoir la forte diffusion numérique qu'il présente. Nous allons introduire cette notion sur l'exemple simplifié de l'équation d'advection à vitesse constante et homogène v. Elle s'écrit :

$$\frac{\partial \phi}{\partial t} + v\frac{\partial \phi}{\partial x} = 0 \tag{D.10}$$

En ajoutant la condition initiale suivante (avec une fonction ϕ_0 dérivable quelconque) :

$$\phi(0, x) = \phi_0(x) \tag{D.11}$$

L'équation D.10 possède l'unique solution :

$$\phi(t, x) = \phi_0(x - vt) \tag{D.12}$$

On remarque au passage que l'Eq.(D.12) traduit bien le fait que la condition initiale est alors transportée sans déformation. On peut alors montrer, en dérivant simplement ϕ_0, que ϕ est aussi solution de :

$$\frac{\partial^2 \phi}{\partial t^2} + v^2\frac{\partial^2 \phi}{\partial x^2} = 0 \tag{D.13}$$

Écrivons par ailleurs la formule de Taylor à l'ordre 2 en temps et en espace pour ϕ :

$$\begin{cases} \phi(t+\Delta t,x) = \phi(t,x) + \Delta t \frac{\partial \phi}{\partial t}(t,x) + \frac{\Delta t^2}{2}\frac{\partial^2 \phi}{\partial t^2}(t,x) + o(\Delta t^2) \\ \phi(t,x-\Delta x) = \phi(t,x) - \Delta x \frac{\partial \phi}{\partial x}(t,x) + \frac{\Delta x^2}{2}\frac{\partial^2 \phi}{\partial x^2}(t,x) + o(\Delta x^2) \end{cases} \tag{D.14}$$

De l'équation (D.14), on tire :

$$\begin{aligned} & \frac{\phi(t+\Delta t,x)-\phi(t,x)}{\Delta t} + v\frac{\phi(t,x)-\phi(t,x-\Delta x)}{\Delta x} \\ = \; & \frac{\partial \phi}{\partial t}(t,x) + \frac{\Delta t}{2}\frac{\partial^2 \phi}{\partial t^2}(t,x) + v\frac{\partial \phi}{\partial x}(t,x) - v\frac{\Delta x}{2}\frac{\partial^2 \phi}{\partial x^2}(t,x) + o(\Delta t) + o(\Delta x) \end{aligned} \tag{D.15}$$

Le schéma upstream approxime à 0 le premier membre de l'équation (D.15). Il résout donc bien l'équation d'advection à l'ordre 0 (en négligeant les termes en $o(1)$ dans (D.15), le membre de droite se ramène bien à l'équation d'advection). A l'ordre 1 en revanche, c'est-à-dire en négligeant seulement les termes en $o(\Delta x)$ et $o(\Delta t)$, le schéma upstream résout :

$$\frac{\partial \phi}{\partial t}(t,x) + \frac{\Delta t}{2}\frac{\partial^2 \phi}{\partial t^2}(t,x) + v\frac{\partial \phi}{\partial x}(t,x) - v\frac{\Delta x}{2}\frac{\partial^2 \phi}{\partial x^2}(t,x) = 0 \tag{D.16}$$

En reportant l'Eq.(D.13) dans l'Eq.(D.16), on obtient que le schéma upstream résout :

$$\frac{\partial \phi}{\partial t}(t,x) + v\frac{\partial \phi}{\partial x}(t,x) = \frac{\partial}{\partial x}\left[\frac{v\Delta x - v^2\Delta t}{2}\frac{\partial \phi}{\partial x}\right](t,x) \tag{D.17}$$

En notant $K_{up} = (v\Delta x - v^2\Delta t)/2$ (en m^2/s), on résout finalement :

$$\frac{\partial \phi}{\partial t}(t,x) + v\frac{\partial \phi}{\partial x}(t,x) = \frac{\partial}{\partial x}\left[K_{up}\frac{\partial \phi}{\partial x}\right](t,x) \tag{D.18}$$

Il s'agit d'une équation d'advection-diffusion, avec v la vitesse d'advection et K_{up} le taux de diffusion. On en conclut tout d'abord que le schéma upstream ajoute une diffusion (numérique), généralement non désirée lorsque l'on modélise un phénomène uniquement avec un terme d'advection. On peut quantifier cette diffusion via l'expression obtenue pour K_{up}. Le respect de la condition CFL (Eq.(D.6)) impose $K_{up} > 0$. Par ailleurs, on peut limiter K_{up} en améliorant la résolution spatiale (en diminuant Δx), mais le respect de la condition CFL impose, en une dimension, de diminuer Δt d'autant. Chaque doublement de résolution spatiale, par exemple, multipliera donc par 4 le nombre de points de grille. On peut enfin noter que lorsque l'on veut modéliser un phénomène par une équation d'advection-diffusion, si l'on utilise le schéma upstream, il n'y aura pas d'intérêt à ajouter effectivement un terme de diffusion si le taux "physique" de diffusion est inférieur à K_{up}.

Ce terme de diffusion numérique étant souvent non désiré, on peut essayer de le corriger. C'est ce que se propose de faire Smolarkiewicz (1983).

3 Justification du schéma de Smolarkiewicz

L'idée générale du schéma de Smolarkiewicz est d'utiliser le schéma upstream dans un premier temps, puis de corriger la diffusion numérique en appliquant un second schéma upstream avec une certaine vitesse d'"antidiffusion". Pour cela, réécrivons l'Eq.(D.18) sous la forme :

$$\frac{\partial \phi}{\partial t}(t,x) + v\frac{\partial \phi}{\partial x}(t,x) = -\frac{\partial}{\partial x}\left[v_d \phi\right](t,x) \tag{D.19}$$

Avec :

$$v_d(t,x) = -\frac{K_{up}}{\phi(t,x)}\frac{\partial \phi}{\partial x}(t,x) \tag{D.20}$$

v_d est appelée vitesse de diffusion. L'Eq.(D.19) est à présent une équation d'advection (à vitesse non constante). Dans le schéma de Smolarkiewicz, on appliquera d'abord le schéma upstream "classique". On construira ainsi, pour une itération temporelle donnée, un état intermédiaire qui résoudra à l'ordre 1 l'Eq.(D.19). Puis on appliquera une deuxième fois le schéma upstream à l'équation d'advection avec la vitesse $-v_d$, dans le but de corriger la diffusion numérique ajoutée précédemment. On note qu'un tel schéma n'est pas totalement exempte de diffusion numérique pour autant, ne serait-ce que parce que la deuxième étape ajoute elle-même de la diffusion numérique qui n'est pas corrigée par une troisième étape. Enfin, on note que si l'on a ici justifié empiriquement la construction du schéma avec une vitesse d'advection constante (v), on l'utilisera dans la suite également avec des vitesses d'advection non constante. On se reportera à Smolarkiewicz (1983) pour des évaluations des performances de ce schéma.

Si l'on se reporte à présent aux équations du schéma de Smolarkiewicz (voir Eq.II.42 dans le II, section 4), on comprend que le troisième point (qui définit $\tilde{v}^*$), est une résolution numérique simple de l'Eq.(D.20), à laquelle on a ajouté un coefficient Sc. Ce coefficient a été ajouté afin d'augmenter légèrement la vitesse d'antidiffusion pour (partiellement) corriger également la diffusion apportée par l'étape de correction.

4 Schémas pour l'équation d'advection 2D

Pour terminer cette Annexe, nous allons donner les équations des 2 schémas (upstream et Smolarkiewicz) pour un problème à 2 dimensions spatiales. On note j la coordonnée de la deuxième dimension d'espace, et Δy le pas correspondant. $\tilde{\phi}$ est toujours le champ scalaire discrétisé (matrice $3D$ désormais). Le champ de vitesse comprend désormais deux composantes, notées $\tilde{v}_x$ et $\tilde{v}_y$. Sous ces conditions, et toujours avec la définition de la fonction F de l'Eq.(D.2), le schéma upstream s'écrit, pour l'évaluation au $k+1$-ième pas de temps :

$$\begin{aligned}\tilde{\phi}_{i,j}^{k+1} = \tilde{\phi}_{i,j}^{k} - F\left(\tilde{\phi}_{i,j}^{k}, \tilde{\phi}_{i+1,j}^{k}, \tilde{v}_{x_{i,j}}^{k}, \Delta x\right) + F\left(\tilde{\phi}_{i-1,j}^{k}, \tilde{\phi}_{i,j}^{k}, \tilde{v}_{x_{i-1,j}}^{k}, \Delta x\right) \\ -F\left(\tilde{\phi}_{i,j}^{k}, \tilde{\phi}_{i,j+1}^{k}, \tilde{v}_{y_{i,j}}^{k}, \Delta y\right) + F\left(\tilde{\phi}_{i,j-1}^{k}, \tilde{\phi}_{i,j}^{k}, \tilde{v}_{y_{i-1,j}}^{k}, \Delta y\right)\end{aligned} \tag{D.21}$$

La condition CFL pour ce schéma $2D$ s'écrit :

$$\max_{i,j}\left(\Delta t\sqrt{\frac{v_{x_{i,j}}^2}{\Delta x^2} + \frac{v_{y_{i,j}}^2}{\Delta y^2}}\right) \leq \frac{1}{\sqrt{2}} \tag{D.22}$$

Le schéma de Smolarkiewicz s'écrit quant à lui formellement :

$$\begin{cases} \tilde{\phi}_{i,j}^{*} & = \tilde{\phi}_{i,j}^{k} - F\left(\tilde{\phi}_{i,j}^{k}, \tilde{\phi}_{i+1,j}^{k}, \tilde{v}_{x_{i,j}}^{k}, \Delta x\right) + F\left(\tilde{\phi}_{i-1,j}^{k}, \tilde{\phi}_{i,j}^{k}, \tilde{v}_{x_{i-1,j}}^{k}, \Delta x\right) \\ & -F\left(\tilde{\phi}_{i,j}^{k}, \tilde{\phi}_{i,j+1}^{k}, \tilde{v}_{y_{i,j}}^{k}, \Delta y\right) + F\left(\tilde{\phi}_{i,j-1}^{k}, \tilde{\phi}_{i,j}^{k}, \tilde{v}_{y_{i-1,j}}^{k}, \Delta y\right) \\ \tilde{\phi}_{i,j}^{k+1} & = \tilde{\phi}_{i,j}^{*} - F\left(\tilde{\phi}_{i,j}^{*}, \tilde{\phi}_{i+1,j}^{*}, \tilde{v}_{x_{i,j}}^{*}, \Delta x\right) + F\left(\tilde{\phi}_{i-1,j}^{*}, \tilde{\phi}_{i,j}^{*}, \tilde{v}_{x_{i-1,j}}^{*}, \Delta x\right) \\ & -F\left(\tilde{\phi}_{i,j}^{*}, \tilde{\phi}_{i,j+1}^{*}, \tilde{v}_{y_{i,j}}^{*}, \Delta y\right) + F\left(\tilde{\phi}_{i,j-1}^{*}, \tilde{\phi}_{i,j}^{*}, \tilde{v}_{y_{i-1,j}}^{*}, \Delta y\right) \\ \tilde{v}_{x_{i,j}}^{*} & = S_c \frac{\left(|\tilde{v}_{x_{i,j}}^{k}|\Delta x - \Delta t\left(\tilde{v}_{x_{i,j}}^{k}\right)^2\right)\left(\tilde{\phi}_{i+1,j}^{*} - \tilde{\phi}_{i,j}^{*}\right)}{\left(\tilde{\phi}_{i,j}^{*} + \tilde{\phi}_{i+1,j}^{*} + \epsilon\right)\Delta x} \\ \tilde{v}_{y_{i,j}}^{*} & = S_c \frac{\left(|\tilde{v}_{y_{i,j}}^{k}|\Delta y - \Delta t\left(\tilde{v}_{y_{i,j}}^{k}\right)^2\right)\left(\tilde{\phi}_{i,j+1}^{*} - \tilde{\phi}_{i,j}^{*}\right)}{\left(\tilde{\phi}_{i,j}^{*} + \tilde{\phi}_{i,j+1}^{*} + \epsilon\right)\Delta y} \end{cases} \tag{D.23}$$

La condition CFL est la même que pour le schéma upstream (Eq.(D.22)) sous réserve d'avoir $Sc < 2$.

Annexe E: Contrôle des paramètres d'un modèle dans un algorithme 4D-VAR.

Le chapitre III décrit le principe de l'algorithme d'assimilation de données 4D-VAR utilisé dans cette thèse. Cette algorithme vise généralement à restituer les conditions initiales d'un modèle numérique de propagation les mieux à-même d'exliquer des observations réparties dans le temps. On a néanmoins vu au chapitre III section 4 que l'on pouvait aisément étendre cet algorithme à la restitution de conditions limites. On a également mentionné le fait que l'on pouvait restituer les paramètres du modèle de propagation eux-mêmes. Dans cette Annexe, nous montrons sur un exemple simple que cette dernière modification ne pose pas de difficulté algorithmique majeure, dès lors que les paramètres du modèle que l'on cherche à contrôler apparaissent explicitement dans les équations reliant les sorties du modèle (états à toutes dates / lieux) à ses entrées (conditions initiales et/ou aux limites).

Nous reprenons ici l'exemple simple du système proies/prédateurs présenté en détails au chapitre III section 3. Les notations utilisées sont celles du chapitre III.

On suppose donc ici x_0 et x_L (conditions initiales et aux limites) connus, et on cherche à retrouver (à ajuster) un paramètre du modèle, noté C (sans a priori sur la nature de C, qui peut être un scalaire, un vecteur, une matrice, avec ou sans dimension spatiale et temporelle). Le terme de rappel à l'ébauche sur x_0 et x_L (voir Eq. (III.23) du chapitre III) n'a plus lieu d'être. On peut cependant en ajouter un sur C. En notant C^b l'ébauche sur le paramètre et $\boldsymbol{B_c}$ sa matrice de covariance d'erreur, il s'écrit :

$$J^b(C) = \left(C - C^b\right) \boldsymbol{B_c}^{-1} \left(C - C^b\right)^T \tag{E.1}$$

De même que précédemment, le rappel aux observations s'écrit :

$$J^{obs}(x) = \sum_{k=0}^{NT} \left(y_k - H_k(x_k)\right) \boldsymbol{R_k}^{-1} \left(y_k - H_k(x_k)\right)^T \tag{E.2}$$

On peut écrire explicitement la dépendance de x_k à C :

$$x_k = M_k(x_0, x_L, C) \tag{E.3}$$

Cette notation est abusive, le C étant inclus dans le modèle M_k lui-même. Néanmoins, ce que l'on a besoin d'évaluer pour minimiser J par rapport à C est la grandeur $\nabla_C J(x)$. Si C apparaît dans l'expression analytique des sorties du modèle par rapport aux entrées dans le module général défini pour tous points de maille et toutes dates, on va pouvoir calculer sa jacobienne et rétropropager le gradient comme on le faisait pour les états du modèle.

Nous allons détailler ce processus sur l'exemple simple déjà évoqué du système proies/prédateurs. Modifions légèrement notre modèle, défini par l'Eq. (III.12) du chapitre III, en ajoutant (totalement arbitrairement) un coefficient C, scalaire ne dépendant pas du temps :

$$x_{k+1} = \begin{pmatrix} x^L_{k+1} \\ x^R_{k+1} \end{pmatrix} = M_{k,k+1}(x_k) = \begin{pmatrix} Cx^L_k(1 + dt + dtx^R_k) \\ x^R_k(1 + dt + dtx^L_k) \end{pmatrix} \tag{E.4}$$

On note que cette modification n'a pas de sens mathématiquement. On remarque néanmoins qu'ici, on a supposé x_0 connu, et on a toujours des observations à t_2 portant sur x^L et x^R. On peut donc remarquer que l'on va chercher la valeur de C (c'est-à-dire que l'on va ajuster le modèle) pour que la propagation de x_0 soit la plus proche possible de y_2, les observations.

En pratique, le coût s'exprime de la même façon, selon l'Eq.(III.13) du chapitre III :

$$J(x_0) = \left(y^L_2 - x^L_2\right)^2 + \left(y^R_2 - x^R_2\right)^2 \tag{E.5}$$

Son gradient par rapport à x_2 a également la même expression que dans le chapitre III, soit :

$$\nabla_{x_2} J(x) = \begin{pmatrix} \frac{\partial J}{\partial x^L_2} \\ \frac{\partial J}{\partial x^R_2} \end{pmatrix} = 2 \begin{pmatrix} y^L_2 - x^L_2 \\ y^R_2 - x^R_2 \end{pmatrix} \tag{E.6}$$

À chaque itération de la minimisation, on dispose d'une valeur pour C. En effet, à la première itération, on part du first guess. Par la suite, le résultat d'une itération servira de départ à la suivante. On peut ainsi évaluer numériquement l'expression précédente. On cherche à présent à évaluer le gradient du coût par rapport au paramètre : $\partial J/\partial C$. On va pour cela utiliser le théorème des fonctions composées :

$$\frac{\partial J}{\partial C} = \frac{\partial J}{\partial x^L_2}\frac{\partial x^L_2}{\partial C} + \frac{\partial J}{\partial x^R_2}\frac{\partial x^R_2}{\partial C} \tag{E.7}$$

D'après l'Eq.(E.4) :

$$\frac{\partial x^R_2}{\partial C} = 0 \tag{E.8}$$

Et :

$$\frac{\partial x_2^L}{\partial C} = x_1^L \left(1 + dt + dt x_1^R\right) + C \left(1 + dt + dt x_1^R\right) \frac{\partial x_1^L}{\partial C} \tag{E.9}$$

(on a utilisé $\partial x_1^R / \partial C = 0$ pour écrire (E.9)).

En procédant de même pour le pas de temps précédent, on obtient :

$$\frac{\partial x_1^L}{\partial C} = x_0^L \left(1 + dt + dt x_0^R\right) \tag{E.10}$$

x_0 étant connu, il n'est pas défini par rapport à un état précédent. Il ne dépend donc pas de C. On a ainsi tout ce qu'il faut pour évaluer numériquement $\partial J / \partial C$. On peut ensuite utiliser le minimiseur numérique pour actualiser C et passer à l'itération suivante.

Comme précédemment, on procède donc à rebours (backward) pour calculer le gradient de la fonction de coût relativement à C. Cela est possible car C apparaît explicitement dans l'expression du modèle (Eq.(E.4)). Le problème est aisément généralisable à une ou plusieurs dimensions spatiales.

Annexe F: Ecriture du terme de filtrage du chapitre IV sous forme de terme d'ébauche

Dans le chapitre IV, section 3, sur la description de l'algorithme 4D-VAR utilisé pour restituer des cartes de précipitations, nous avons introduit un terme de filtrage, afin d'ajouter de la corrélation entre les points de maille voisins du champ d'atténuations. Dans cette Annexe, nous allons montrer que ce terme de filtrage peut s'écrire sous la forme d'un terme d'ébauche, au sens de l'assimilation 4D-VAR. On note que les calculs qui suivent s'appliquent également au filtre à une dimension d'espace présenté dans le chapitre Assimilation de Données (chapitre III), utilisé ensuite dans le chapitre V.

Ici, nous reprenons les notations introduites au chapitre IV, section 3. Pour un champ d'atténuation A, le terme de filtrage introduit J^f s'écrit :

$$J^f(A) = c^f \sum_{k=0}^{NT} J_k^f(A_k) \tag{F.1}$$

Avec c^f un coefficient permettant de pondérer les deux termes de la fonction de coût, et pour la date t_k :

$$J_k^f(A_k) = ||A_k - m_k||^2 \tag{F.2}$$

Formellement, A_k désigne toujours ici le champ d'atténuations à la date t_k réécrit sous forme vectorielle. m_k désigne la matrice (elle aussi réécrite sous la même forme vectorielle) contenant les atténuations moyennées localement. Son terme général, $(m_k)_{i,j}$ s'écrit :

$$(m_k)_{i,j} = \frac{1}{9} \sum_{\substack{p=-1,0,1 \\ q=-1,0,1}} A_{k,i+p,j+q} \tag{F.3}$$

Formellement, J_k^f peut être réécrit comme un terme d'ébauche (à toute date et pas seulement à t_0) avec $A^b = 0$. C'est-à-dire que l'on peut définir une matrice $\boldsymbol{B}$ telle que :

$$J_k^f(A_k) = A_k \boldsymbol{B}^{-1} A_k^T \tag{F.4}$$

La matrice $\boldsymbol{B}$ ne dépend pas de la date. On peut l'écrire explicitement comme la somme de

matrices définies pour chaque point de la maille de discrétisation. Soit formellement :

$$\boldsymbol{B}^{-1} = \sum_{\substack{i\in[|1;NX|]\\ j\in[|1;NY|]}} \boldsymbol{B}^{ij} \tag{F.5}$$

Avec $\boldsymbol{B}^{ij}$ une matrice dont le terme général, $\left(\boldsymbol{B}^{ij}\right)_{p,q}$, est défini par :

$$\left(\boldsymbol{B}^{ij}\right)_{p,q} = c^f \begin{cases} \frac{64}{81} & \text{si } (p,q) = (i,j) \\ -\frac{8}{81} & \text{si } (p,q) = (i,j\pm 1) \text{ ou } (p,q) = (i\pm 1,j) \\ \frac{1}{81} & \text{si } (p,q) = (i\pm 1, j\pm 1) \\ 0 & \text{sinon} \end{cases} \tag{F.6}$$

Preuve

On va démontrer la possibilité d'écrire le terme de filtrage sous la forme d'un terme de covariance d'erreur d'ébauche, c'est-à-dire démontrer l'Eq.(F.4), avec la matrice $\boldsymbol{B}$ définie par l'Eq.(F.6). Nous allons d'abord simplifier les notations. Oublions l'indice de temps k, qui devient implicite. Écrivons le champ A (à une date t_k quelconque donc), sous une forme pleinement vectorielle, soit $A = (A_i)_{i\in[|1:NX*NY|]}$. Pour un point d'espace i quelconque, notons $\left(a_p^i\right)_{p\in[|1:8|]}$ les indices des 8 points de maille voisins du point i (ils sont voisins dans l'écriture matricielle mais pas dans l'écriture vectorielle). La moyenne locale m voit ainsi sont terme général s'écrire :

$$m_i = \frac{1}{9}\left[A_i + \sum_{p=1}^{8} A_{a_p^i}\right] \tag{F.7}$$

On a alors :

$$J_k^f(A) = \sum_i J_i^f(A_i) \tag{F.8}$$

Avec (la dépendance à k est à nouveau pleinement implicite et A_i et m_i sont des scalaires) :

$$J_i^f(A_i) = (A_i - m_i)^2 \tag{F.9}$$

Soit :

$$J_i^f(A_i) = \left(\frac{8}{9}A_i - \frac{1}{9}\sum_{p=1}^{8} A_{a_p^i}\right)^2 \tag{F.10}$$

Et en développant :

$$J_i^f(A_i) = \frac{64}{81}A_i^2 - 2\frac{8}{81}A_i\sum_{p=1}^{8} A_{a_p^i} + \frac{1}{81}\left(\sum_{p=1}^{8} A_{a_p^i}\right)^2 \tag{F.11}$$

Par ailleurs, avec la matrice $\boldsymbol{B^i}$ définie par son terme général :

$$\left(\boldsymbol{B^i}\right)_{p,q} = c_f \begin{cases} \frac{64}{81} & \text{si } p = q = i \\ -\frac{8}{81} & \text{si } \begin{cases} p = i \text{ et } q \in \{a_1^i, ..., a_8^i\} \\ \text{ou } q = i \text{ et } p \in \{a_1^i, ..., a_8^i\} \end{cases} \\ \frac{1}{81} & \text{si } p, q \in \{a_1^i, ..., a_8^i\} \\ 0 & \text{sinon} \end{cases} \tag{F.12}$$

On définit :

$$\tilde{J}_k^f(A) = \sum_i \tilde{J}_i^f(A) \tag{F.13}$$

Avec :

$$\tilde{J}_i^f(A) = A\boldsymbol{B^i}A^T \tag{F.14}$$

On doit donc maintenant montrer :

$$\tilde{J}_i^f(A) = J_i^f(A_i) \tag{F.15}$$

Développons pour cela l'Eq.(F.14) avec pour $\boldsymbol{B^i}$ les coefficients de l'Eq.(F.12). On a ainsi de manière générale :

$$\left(\boldsymbol{B^i}A^T\right)_p = \sum_q \left(\boldsymbol{B^i}\right)_{p,q} A_q \tag{F.16}$$

Soit :

$$\boldsymbol{B^i}A^T = \begin{pmatrix} ... \\ -\frac{8}{81}A_i + \frac{1}{81}\sum_p A_{a_p^i} \\ ... \\ \frac{64}{81}A_i - \frac{8}{81}\sum_p A_{a_p^i} \\ ... \\ -\frac{8}{81}A_i + \frac{1}{81}\sum_p A_{a_p^i} \\ ... \end{pmatrix} \begin{matrix} \\ \leftarrow \text{ ligne } a_1^i \\ \\ \leftarrow \text{ ligne } i \\ \\ \leftarrow \text{ ligne } a_8^i \\ \end{matrix} . \tag{F.17}$$

(sans préjuger des valeurs de a_1^i, ..., a_8^i : on peut avoir $a_8^i < a_1^i$ par exemple).

On a ensuite en multipliant A par le vecteur colonne précédent :

$$A\boldsymbol{B^i}A^T = A_{a_1^i}\left[-\frac{8}{81}A_i + \frac{1}{81}\sum_p A_{a_p^i}\right] + ... + A_i\left[\frac{64}{81}A_i - \frac{8}{81}\sum_p A_{a_p^i}\right] + ... + A_{a_8}\left[-\frac{8}{81}A_i + \frac{1}{81}\sum_p A_{a_p^i}\right] \tag{F.18}$$

Et en réorganisant cette expression :

$$A\boldsymbol{B^i}A^T = \frac{64}{81}A_i^2 - 2\frac{8}{81}A_i\sum_{p=1}^{8}A_{a_p^i} + \frac{1}{81}\left(\sum_{p=1}^{8}A_{a_p^i}\right)^2 \tag{F.19}$$

Ce qui clôt le preuve.

Annexe G: Détermination par triangulation de la vitesse d'advection des champs de pluie

Dans le travail sur la reconstitution de cartes de précipitations à partir de mesures d'atténuations présenté au chapitre IV, nous utilisons un modèle d'advection des champs de pluie. Un tel modèle nécessite en entrée une vitesse d'advection (vitesse de propagation des champs). Nous avons vu au chapitre IV section 4 la méthode utilisée dans cette thèse pour déterminer cette vitesse d'advection à partir de mesures radar. Il serait néanmoins souhaitable de pouvoir appliquer notre algorithme sans utiliser de données radar. Durant cette thèse, nous avons développé une méthode originale de triangulation pour déterminer la vitesse de déplacement de champs de pluie directement à partir de mesures d'atténuations par capteur Ku. Cette méthode ne fonctionnant qu'avec au minimum les données de deux capteurs Ku voisins (tandis que nous ne disposions que d'un pour ce travail), elle n'a pas été présentée dans le corps du manuscrit. Nous la présentons et la discutons en détails dans cette Annexe.

L'idée générale de la méthode de triangulation développée ici pour estimer la vitesse d'advection directement à partir des relevés d'atténuation le long de liaisons Ku est de trouver les deux paramètres du vent (vitesse et direction ou deux composantes) qui minimiseront l'écart entre la mesure du temps mis par la cellule pour toucher successivement les différentes liaisons et un temps théorique à définir. De telles méthodes ont déjà été développées et utilisées dans la même optique. Zinevich et al. (2009) par exemple utilise une méthode de triangulation développée par Desa and Niemczynowicz (1997). Le fait que l'on travaille ici avec des liaisons fortement non isotropes rend néanmoins difficilement applicable cette méthode. Nous allons présenter une autre méthode développée pour notre problème.

Les notations utilisées dans cette Annexe sont celles du chapitre IV.

Notations / Définitions Dans cette optique, notons $\boldsymbol{v} = (v_x; v_y)$ le vecteur vitesse du vent. C'est l'inconnue du problème. On note $v = \sqrt{\left(v_x^2 + v_y^2\right)}$ sa vitesse et δ sa direction. Dans toutes le chapitre IV, on a utilisé systématiquement pour δ la même convention que pour les azimuts, à savoir 180° = vent de sud ; 270° = vent d'ouest, etc. Cependant dans cette Annexe, pour

simplifier les expressions, nous utilisons la convention définie sur la Figure G.1. Pour deux liaisons Ku d'indices p et q, on note τ_{pq} le temps effectivement mis par la cellule pour passer de l'une à l'autre (il faut bien sûr définir une méthode pour évaluer τ_{pq} à partir des mesures). On note d_{pq} une distance entre les deux liaisons pour la cellule considérée, caractérisée par $\boldsymbol{v}$ et la géométrie des liaisons. On note $\boldsymbol{d}$ le vecteur contenant tous les d_{pq}, et $\boldsymbol{\tau}$ celui contenant tous les τ_{pq}. On va alors chercher le $\boldsymbol{v}$ minimisant l'écart entre $\boldsymbol{d}$ et $v\boldsymbol{\tau}$, normalisé par $\boldsymbol{d}$. Notre problème consiste donc à minimiser une fonction de coût J^{vent} qui s'écrit :

$$J^{vent}(\boldsymbol{v}) = \frac{\|\boldsymbol{d} - v\boldsymbol{\tau}\|^2}{\|\boldsymbol{d}\|^2} \tag{G.1}$$

Reste bien sûr à définir une distance $\boldsymbol{d}$ et une méthode pour estimer $\boldsymbol{\tau}$.

La section 1 de cette Annexe revient sur la définition de ces grandeurs et donne quelques résultats et considérations théoriques sur ces questions. La section 2 donne des résultats d'applications de cette méthode sur des données simulées.

1 Théorie

Estimation de $\boldsymbol{\tau}$

Notons $A_1, A_2, ..., A_{NL}$ les séries temporelles d'atténuations mesurées le long des NL liaisons Ku, et NT le nombre de mesures pour chaque liaison (le nombre de dates). La fonction de corrélation croisée entre deux liaisons d'indices p et q, notée Co_{pq}, s'écrit, $\forall m \in [|-NT; NT|]$:

$$Co_{pq}(m) = \begin{cases} \sum_{k=1}^{NT-m} A_p(k+m)A_q(k) & \text{if } m > 0 \\ Co_{pq}(-m) & \text{if } m < 0. \end{cases} \tag{G.2}$$

On définit alors τ_{pq} comme le décalage temporel (le m) qui maximise cette corrélation croisée, soit formellement :

$$\tau_{pq} \ / \ \max_m [Co_{pq}(m)] = Co_{pq}(\tau_{pq}) \tag{G.3}$$

Notons que le vecteur contenant les valeurs $Co_{pq}(\tau_{pq})$ pour chaque couple (p, q) indique à quel point la corrélation est forte entre les enregistrements de chaque couples de liaisons, et pourrait être utilisé pour pondérer l'importance à donner à chaque composante des vecteurs $\boldsymbol{d} - v\boldsymbol{\tau}$ dans le coût défini par l'Eq.(G.1).

On s'en tiendra néanmoins ici à un cas simple sans pondération.

Estimation de $\boldsymbol{d}$

Dans ce paragraphe, nous allons définir la distance d_{pq} de l'Eq.(G.1) entre deux liaisons d'indices p et q.

Cas simple : 1 capteur, liaisons de même longueur Jusqu'à nouvel ordre, on suppose que les deux liaisons ont la même longueur, L, et partent du même capteur Ku, placé en un point O du plan. On définit alors d_{pq} par :

$$d_{pq}(\beta_{pq}, \gamma, \delta) = \frac{1}{L}\int_0^L \tilde{d}_{pq}(r, \beta_{pq}, \gamma, \delta)dr \tag{G.4}$$

Avec β_{pq} l'angle entre les deux liaisons, γ l'angle par rapport au sud de la liaison pointant le plus vers l'est (q sur la Figure G.1), δ la direction d'advection et $\tilde{d}_{pq}$ défini par la Figure G.1. Sur cette figure, on voit que l'on suppose que la cellule pluvieuse présente un front rectiligne orthogonal à sa direction d'advection (vecteur $\boldsymbol{v}$). Les distances définies ici correspondront également d'autant plus aux décalages temporels mesurés précédemment ($\boldsymbol{\tau}$) que la cellule de pluie est homogène (on suppose que les liaisons voient la même chose à différents moments). On définit alors $\tilde{d}_{pq}$ comme la distance que devra parcourir une telle cellule pour passer d'un point sur la liaison p à la distance r du capteur Ku à un point sur la liaison q à la même distance r du capteur Ku. La distance d_{pq} entre les 2 liaisons est ensuite la moyenne sur toute la longueur L de cette distance $\tilde{d}_{pq}$.

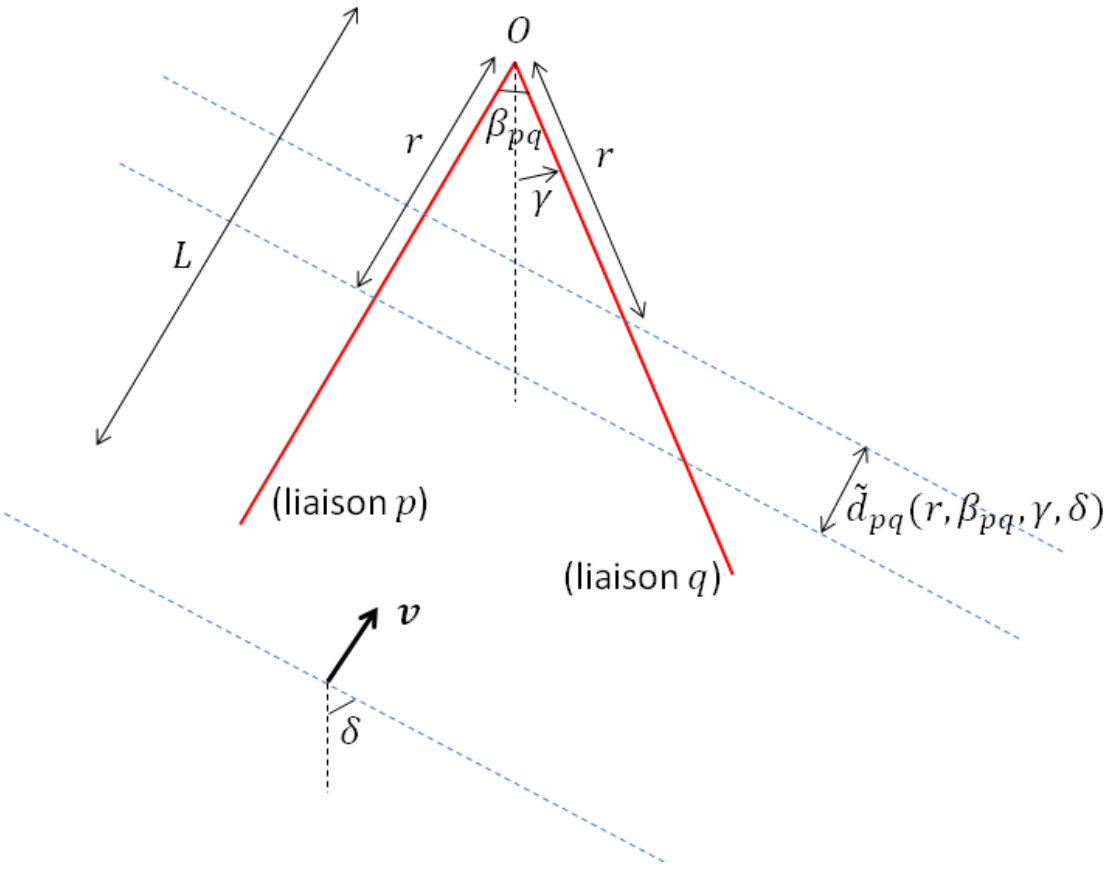

Figure G.1 – Définition du $\tilde{d}_{pq}$ de l'Eq.(G.4). Une cellule pluvieuse, présentant un front rectiligne (pointillés bleus), progresse dans la direction $\boldsymbol{v}$. Elle touche deux liaisons p et q, de même longueur L, partant du même capteur Ku positionné en O, et faisant un angle β_{pq} entre elles. La liaison pointant vers l'est fait un angle γ par rapport au sud. L'angle δ caractérise la direction de propagation.

On va montrer que la distance $\tilde{d}_{pq}$ s'écrit :

$$\tilde{d}_{pq}(r, \beta_{pq}, \gamma, \delta) = 2rsin\left(\frac{\beta_{pq}}{2}\right) cos\left(\delta + \gamma + \frac{\beta_{pq}}{2}\right) \tag{G.5}$$

En intégrant ensuite cette expression selon l'Eq.(G.4), on obtient que la distance globale d_{pq} entre les deux liaisons s'écrit :

$$d_{pq}(\beta_{pq}, \gamma, \delta) = \tilde{d}_{pq}\left(\frac{L}{2}, \beta_{pq}, \gamma, \delta\right) \tag{G.6}$$

Soit finalement :

$$d_{pq}(\beta_{pq}, \gamma, \delta) = Lsin\left(\frac{\beta_{pq}}{2}\right) cos\left(\delta + \gamma + \frac{\beta_{pq}}{2}\right) \tag{G.7}$$

Au-delà de l'expression analytique, le point important est que l'on montre que l'on peut réduire les liaisons à leur point milieu, ce qui va simplifier très considérablement la mise en pratique de cette méthode.

Preuve de l'Eq.(G.5)

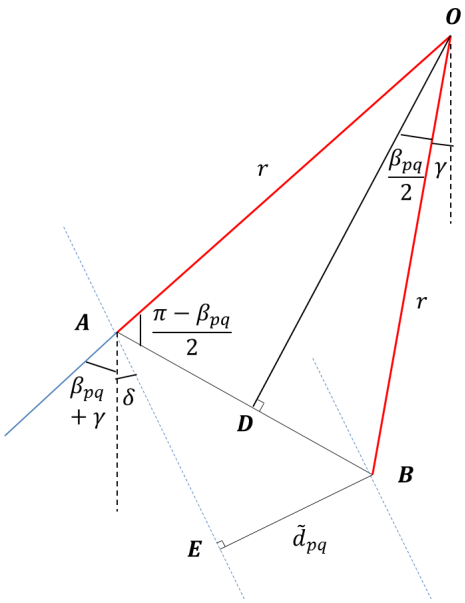

FIGURE G.2 – Idem Figure G.1 avec l'introduction de notations complémentaires pour la preuve de l'Eq.(G.5).

Il s'agit de montrer :

$$\tilde{d}_{pq}(r, \beta_{pq}, \gamma, \delta) = 2rsin\left(\frac{\beta_{pq}}{2}\right) cos\left(\delta + \gamma + \frac{\beta_{pq}}{2}\right) \tag{G.8}$$

La Figure G.2 va nous aider. Elle représente la même situation que la Figure G.1 mais

introduit en plus un certain nombre d'angles et de notations. Il s'agit d'estimer $\tilde{d}_{pq}$, c'est à dire la longueur $\overline{EB}$ de la Figure G.2. (OAB) étant isocèle, on a, dans (ODB) :

$$\overline{DB} = rsin\left(\frac{\beta_{pq}}{2}\right) \tag{G.9}$$

Et donc :

$$\overline{AB} = \overline{DB} = 2rsin\left(\frac{\beta_{pq}}{2}\right) \tag{G.10}$$

Dans (ABE) :

$$\overline{EB} = \overline{AB}cos\left(\widehat{ABE}\right) \tag{G.11}$$

Avec :

$$\widehat{ABE} = \frac{\pi}{2} - \widehat{EAB} \tag{G.12}$$

Or :

$$\widehat{EAB} = \pi - \left(\frac{\pi - \beta_{pq}}{2} + \delta + \gamma + \beta\right). \tag{G.13}$$

D'où :

$$\widehat{ABE} = \delta + \gamma + \frac{\beta_{pq}}{2} \tag{G.14}$$

Et donc :

$$\overline{EB} = 2rsin\left(\frac{\beta_{pq}}{2}\right)cos\left(\delta + \gamma + \frac{\beta_{pq}}{2}\right) \tag{G.15}$$

Ce qui clôt la preuve.

Généralisation à deux capteurs Ku Dans le cas où l'on a deux capteurs Ku, et toujours des liaisons de même longueur L, on peut toujours définir la distance d_{pq} entre deux liaisons p et q par l'Eq.(G.4). $\tilde{d}_{pq}(r, \beta, \delta)$ est défini de la même manière, voir Figure G.3.

Exprimer analytiquement $\tilde{d}_{pq}$ requiert cependant désormais d'introduire une nouvelle distance, $\tilde{d}_{O_1O_2}$, entre les positions des deux capteurs Ku. Cette distance est définie de la même manière que $\tilde{d}_{pq}$ (c'est-à-dire comme la distance que doit parcourir la cellule de pluie pour toucher successivement O_1 et O_2) et est représentée sur la Figure G.3. Elle ne dépend que de δ, la direction du vent, et peut être exprimée à partir des coordonnées de O_1 et O_2. Nous allons montrer :

$$\tilde{d}_{pq}(r, \beta_{pq}, \gamma, \delta) = \tilde{d}_{pq'}(r, \beta_{pq}, \gamma, \delta) + \tilde{d}_{O_1O_2} \tag{G.16}$$

Avec q' la liaison parallèle à la liaison q et d'origine O_1 (voir Figure G.3). On a donc :

$$d_{pq}(\beta_{pq}, \gamma, \delta) = \frac{1}{L}\int_0^L \left[\tilde{d}_{pq'}(r, \beta_{pq}, \gamma, \delta)dr + \tilde{d}_{O_1O_2}dr\right] \tag{G.17}$$

Soit encore :

$$d_{pq}(\beta_{pq}, \gamma, \delta) = d_{pq'}(\beta_{pq}, \gamma, \delta) + \tilde{d}_{O_1O_2}(\delta) \tag{G.18}$$

Les liaisons p et q' provenant ensuite du même capteur Ku, on peut appliquer l'Eq.(G.7)

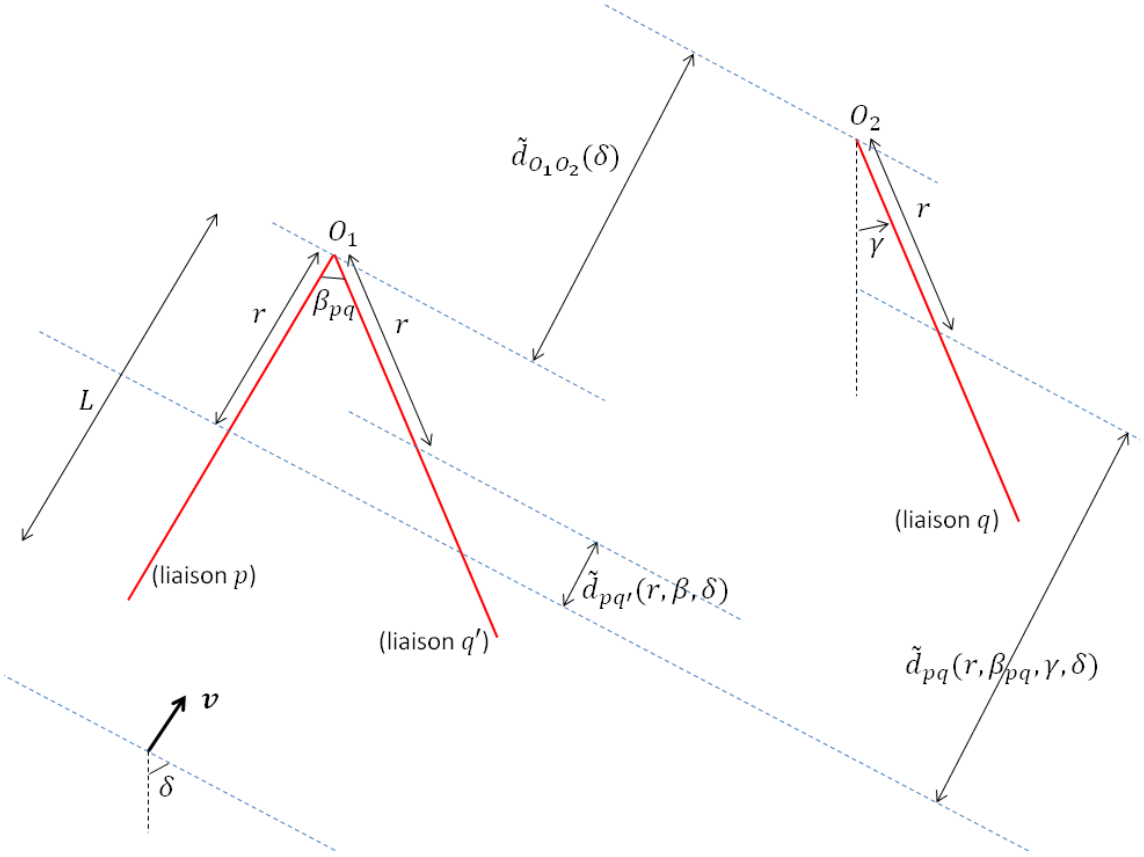

FIGURE G.3 – Idem Figure G.1 mais avec des liaisons p et q provenant de deux capteurs Ku différents, positionnés en O_1 et O_2. Pour exprimer $\tilde{d}_{pq}$, on introduit $\tilde{d}_{O_1O_2}$, distance entre les deux capteurs, et une liaison fictive $q^{'}$, parallèle à q d'origine O_1.

pour évaluer $d_{pq'}$ et on a finalement :

$$d_{pq}(\beta_{pq},\gamma,\delta) = Lsin\left(\frac{\beta_{pq}}{2}\right)cos\left(\delta+\gamma+\frac{\beta_{pq}}{2}\right)+\tilde{d}_{O_1O_2}(\delta) \tag{G.19}$$

On peut donc là encore assimiler les liaisons à leurs points milieux dans le calcul de la distance d_{pq}.

Preuve de l'Eq.(G.16)

Il s'agit de montrer (sans les dépendances) :

$$\tilde{d}_{pq} = \tilde{d}_{pq'} + \tilde{d}_{O_1O_2} \tag{G.20}$$

La Figure G.4 va nous aider. Elle représente la même situation que la Figure G.3 mais introduit en plus un certain nombre de distances. Tous les fronts de pluie (pointillés bleus) étant parallèles, toutes les perpendiculaires représentées ici (lignes noires) le sont aussi. Et

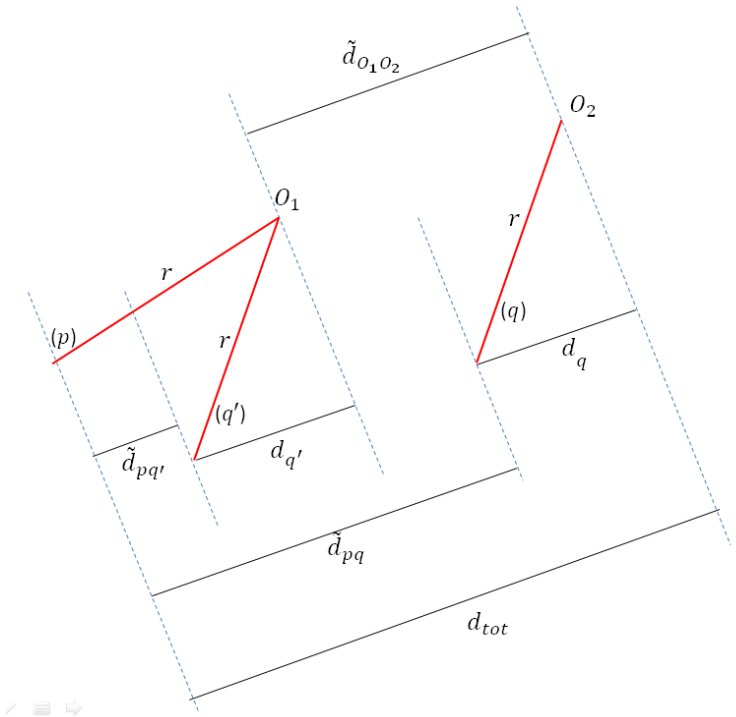

Figure G.4 – Idem Figure G.3 avec l'introduction de notations complémentaires pour la preuve de l'Eq.(G.16).

donc :

$$\begin{aligned} d_{tot} &= \tilde{d}_{pq} + d_q \\ &= \tilde{d}_{pq'} + d_{q'} + \tilde{d}_{O_1O_2} \end{aligned} \tag{G.21}$$

Les liaisons q et $q^{'}$ étant par ailleurs parallèles (par définition de $q^{'}$), on a :

$$d_{q'} = d_q \tag{G.22}$$

Et donc d'après l'Eq.(G.21) :

$$\tilde{d}_{pq} = \tilde{d}_{pq'} + \tilde{d}_{O_1O_2} \tag{G.23}$$

Ce qui clôt la preuve.

Quelques conclusions intermédiaires

- Après avoir défini, via l'Eq.(G.1), un cadre général pour déterminer la vitesse d'advection à partir des mesures d'atténuations, on a donné des formules analytiques permettant d'évaluer la fonctionnelle de l'Eq.(G.1).

- Ces formules sont basées sur des suppositions fortes sur la forme de la cellule de pluie (front rectiligne et perpendiculaire à son déplacement, cellule suffisamment homogène).

- Sous l'hypothèse que les liaisons ont les mêmes longueurs, et en utilisant la distance définie par l'Eq.(G.4), la pertinence de l'assimilation des liaisons à leur points milieux a été démontrée.

- C'est cette dernière propriété qui est importante. Connaissant les positions des capteurs et des points milieu des liaisons, il sera ensuite très facile en pratique d'évaluer numériquement $\boldsymbol{d}(\delta)$ pour un δ donné.

À partir de là, on peut aussi noter :

- que rien ne nous empêche d'utiliser ce travail (c'est-à-dire d'assimiler les liaisons à leurs points milieu) dans un cadre plus général (cellules et liaisons quelconque). On n'a dans ce cas plus de base théorique mais on peut tester sur données réelles ou simulées la pertinence de cette méthode.

- que dans la réalité, si les capteurs sont plus ou moins à la même altitude, et pour des latitudes comparables à celles de la France, les liaisons projetées au sol peuvent avoir des longueurs relativement proches. Par exemple, pour les liaisons décrites dans le Tableau IV.1 du chapitre IV et avec un iso-0 à $3000m$, on obtient des longueurs de 5.5, 4.1, 3.9 et $4.5km$.

- que pour des cellules de pluie de taille conséquente ($>>5km$ en France), l'hypothèse d'une ligne de front rectiligne sur la zone autour des liaisons et perpendiculaire au déplacement paraît raisonnable.

- que rien ne prouve que la minimisation de la fonctionnelle définie par l'Eq.(G.1) fournira un minimum unique. Il peut y avoir ambiguïté entre vitesse et direction du vent. Le paragraphe suivant montre théoriquement (avec les mêmes hypothèses que précédemment sur la géométrie de la cellule de pluie) qu'avec un seul capteur Ku, il y a effectivement ambiguïté.

Limites avec un seul capteur

Nous allons montrer ici que la méthode décrite ne permet pas de retrouver conjointement la vitesse et la direction du vent dans le cas où les milieux des liaisons Ku sont alignés.

Pour cela, on se place toujours dans le cas précédent, où les liaisons ont la même longueur (ce qui permet de les réduire à leurs points milieu via l'Eq.(G.7)) et où les cellules de pluie présentent un front rectiligne perpendiculaire à leur vecteur de déplacement.

La Figure G.5 présente une telle situation avec trois liaisons, dont seuls les centres (points A, B et C) sont représentés. Ils sont alignés selon une direction quelconque caractérisée ici par l'angle υ. La cellule (traits pointillés bleus) se déplace toujours selon une direction δ. Disposant de 3 liaisons (notées par les indices A, B et C correspondant à leurs milieux), le vecteur $\boldsymbol{d}$ aura trois composantes : d_{AB}, d_{AC} et d_{BC}. En plus de la géométrie des liaisons, il est donc fonction de δ et s'écrit, avec les notations de la Figure G.5 :

$$\boldsymbol{d}(\delta) = \begin{pmatrix} \overline{AB} sin(\upsilon - \delta) \\ \overline{AC} sin(\upsilon - \delta) \\ \overline{BC} sin(\upsilon - \delta) \end{pmatrix}. \tag{G.24}$$

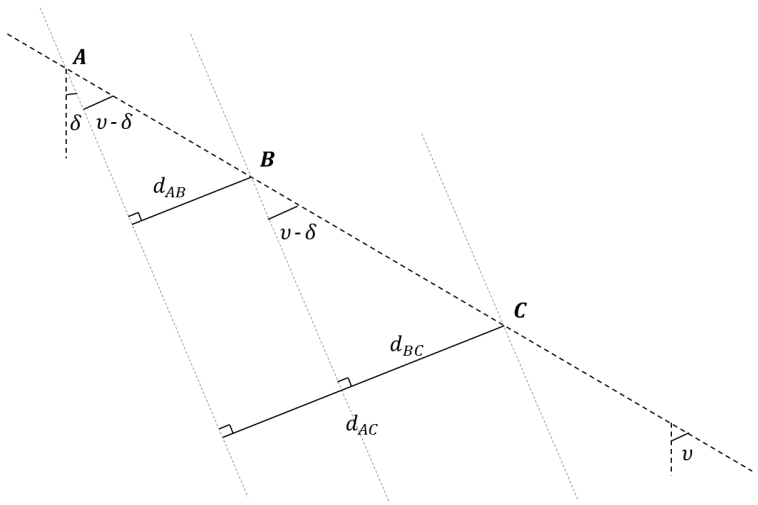

FIGURE G.5 – Représentation d'une situation où les points milieux A, B et C de trois liaisons Ku quelconques sont alignés selon une direction caractérisée par un angle υ.

Pour deux directions d'advection quelconques δ_1 et δ_2, on a donc :

$$\boldsymbol{d}(\delta_2) = \frac{sin(\upsilon - \delta_2)}{sin(\upsilon - \delta_1)}\boldsymbol{d}(\delta_1) \tag{G.25}$$

La fonction de coût (Eq.(G.1)) permettant de déterminer le vent possède un minimum, noté J^{vent*}, atteint (au moins) pour un certain couple (δ_{min}, v_{min}). Soit formellement :

$$J^{vent}(\delta_{min}, v_{min}) = \min_{\delta,v} J^{vent}(\delta, v) \overset{def}{=} J^{vent*} \tag{G.26}$$

Considérons à présent un couple (δ, v) quelconque. On a :

$$J^{vent}(\delta, v) = \frac{\|\boldsymbol{d}(\delta) - v\boldsymbol{\tau}\|^2}{\|\boldsymbol{d}(\delta)\|^2} \tag{G.27}$$

Et avec l'Eq.(G.25), on peut écrire :

$$J^{vent}(\delta, v) = \frac{\left\|\frac{sin(\alpha-\delta)}{sin(\alpha-\delta_{min})}\boldsymbol{d}(\delta_{min}) - v\boldsymbol{\tau}\right\|^2}{\left\|\frac{sin(\alpha-\delta)}{sin(\alpha-\delta_{min})}\boldsymbol{d}(\delta_{min})\right\|^2} \tag{G.28}$$

Et avec :

$$v_\delta = \frac{sin(\upsilon - \delta)}{sin(\upsilon - \delta_{min})}v_{min} \tag{G.29}$$

On a :

$$J^{vent}(\delta, v_\delta) = J^{vent*} \tag{G.30}$$

On montre ainsi que quelque soit la valeur de δ, on peut trouver une valeur pour v permettant à J^{vent} d'atteindre son minimum global. Il n'y a donc pas unicité de la solution du problème mais ambiguïté entre vitesse et direction du vent. Si les milieux des liaisons sont alignés, on ne

peut déterminer la vitesse d'advection que si l'on connaît la direction, et vice-versa.

Dans le travail présenté ici, on dispose de 4 liaisons provenant toutes du même capteur Ku. Les liaisons ayant des longueurs (une fois projetées au sol) légèrement différentes, leurs milieux ne sont pas rigoureusement alignés mais le sont quasiment, sur un axe ouest-est. On pressent donc que cette méthode ne sera pas très performante dans ce cas. La section 2 suivante va confirmer sur données simulées qu'elle ne pourra être utilisée dans ce cas-là. On confirmera en revanche sa pertinence dans le cas où l'on dispose de 2 capteurs Ku.

2 Résultats

Dans cette partie, nous allons évaluer, avec des données simulées, la pertinence de la méthode décrite ci-dessus pour des cellules et des liaisons Ku réalistes. Ainsi, les cellules seront bien advectées dans l'espace à vitesse constante et homogène, mais ne présenteront pas un front strictement rectiligne et perpendiculaire à leur déplacement. Travaillant sur données simulées, les conditions d'advection (vecteur vitesse $\boldsymbol{v}$ caractérisé par son module v et sa direction δ) seront parfaitement connues. L'objectif sera de voir quelles valeurs notre algorithme restitue, selon le nombre de capteurs Ku disponibles (1 ou 2), le positionnement relatif des capteurs Ku (s'il y en a deux) et la vitesse, la direction et l'emplacement initial de la cellule de pluie.

Dans tous les cas, nous utilisons la même cellule de pluie, générée avec le simulateur évoqué dans la section 5 du chapitre IV. On peut voir cette cellule notamment sur la Figure G.6. Elle présente une pluviométrie maximale d'environ $85mm/h$ et couvre une zone d'environ $7 * 7km^2$. Dans le premier paragraphe suivant, nous étudierons notre méthode dans le cas où un seul capteur Ku est disponible, enregistrant les atténuations le long de 4 liaisons (comme dans le dispositif expérimental du chapitre IV). Dans le second paragraphe, nous travaillerons avec deux capteurs, chacun captant les émissions des 4 mêmes satellites.

Table G.1 – Description des 4 liaisons Terre-satellite simulées et projetées au sol pour évaluer la pertinence de la méthode de triangulation pour restituer les conditions d'advection avec un seul capteur Ku.

	liaison 1	liaison 2	liaison 3	liaison 4
azimut ($^\circ$)	210	190	170	150
longueur (km)	6.5	6.1	6.1	6.0

Dans tous les cas, les caractéristiques des 4 liaisons captées par le capteur Ku sont données dans le Tableau G.1. Les azimuts sont tous séparés de 20 $^\circ$, et les longueurs des liaisons projetées au sol sont d'environ $6km$. Ces caractéristiques correspondent assez bien à celles obtenues avec notre dispositif expérimental.

1 capteur Ku : ambiguïté !

Nous allons donc dans un premier temps évaluer notre méthode dans le cas où un seul capteur Ku est disponible. 6 simulations sont réalisées dans ce cas-là. Entre elles, on modifie la position initiale de la cellule de pluie, la vitesse, et la direction d'advection. La Figure G.6 présente les caractéristiques de ces 6 simulations (A), ainsi que les enregistrements simulés du capteur sous ces conditions (B). Les deux premières simulations ($A1$ et $A2$) voient la cellule à l'instant initial placée immédiatement à l'ouest-sud-ouest des liaisons Ku, et advectée vers l'est-nord-est (flèche blanche), de manière à ce qu'elle survole ensuite bien les 4 liaisons. La vitesse d'advection est de $10m/s$ (1) ou de $20m/s$ (2). Logiquement, les enregistrements réalisés par le capteur Ku présentent les mêmes variations ($B1$ et $B2$) mais sont contractés lorsque la cellule passe plus vite ($B2$). Dans les simulations 3 et 4, on advecte la cellule vers le nord-est à $14m/s$. Elle est néanmoins placée à l'instant initial de manière à survoler de manière plus homogène les 4 liaisons dans la simulation 4. Dans les simulations 5 et 6 enfin, on advecte la cellule vers le nord-nord-est à $13m/s$. Dans la simulation 5, la cellule est placée de manière à ne toucher que très marginalement la liaison 1. Le Tableau G.2 (colonnes 1 à 3) présente également les caractéristiques d'advection de ces réalités simulées.

Table G.2 – Conditions d'advection pour les 6 simulations réalisées ici : colonne 2 : direction d'advection (180 ° correspond à une propagation sud-nord, 270 ° à une propagation ouest-est, etc.) ; colonne 3 : vitesse d'advection. Colonne 4 : valeurs de vitesse restituées par la méthode de triangulation avec un seul capteur Ku et 4 liaisons en supposant connue la direction d'advection.

	réalité simulée		restitution
simulation	δ (°)	$v(m/s)$	$v(m/s)$
1	250	10	8.8
2	250	20	16.6
3	225	14	9.4
4	225	14	13.2
5	200	13	2.3
6	200	13	8.2

Nous avons vu que dans le cas où les milieux des capteurs d'atténuation étaient alignés (et où les cellules présentaient un front rectiligne), il était théoriquement impossible de restituer à la fois la vitesse et la direction d'advection. Nous allons constater ici que pour un cas plus réaliste, on aboutit à la même conclusion. Ici, les 4 liaisons provenant du même capteur Ku, leurs milieux sont en effet plus ou moins alignés.

La Figure G.7, partie gauche, présente le coût obtenu (i.e. la valeur de la fonction J^{vent} de l'Eq.(G.1)), en fonction de v et δ, pour la simulation 2 présentée ci-dessus (advection vers l'est-nord-est (250 °) à $20m/s$). On remarque que l'on n'obtient pas un minimum unique mais une courbe minimisante. C'est-à-dire que pour chaque valeur de direction (i.e. chaque ligne horizontale sur la figure), on peut trouver une vitesse fournissant le minimum global. Ceci confirme les résultats théoriques vus plus haut.

On constate par ailleurs que la méthode induit une légère erreur. On sait en effet que l'on a

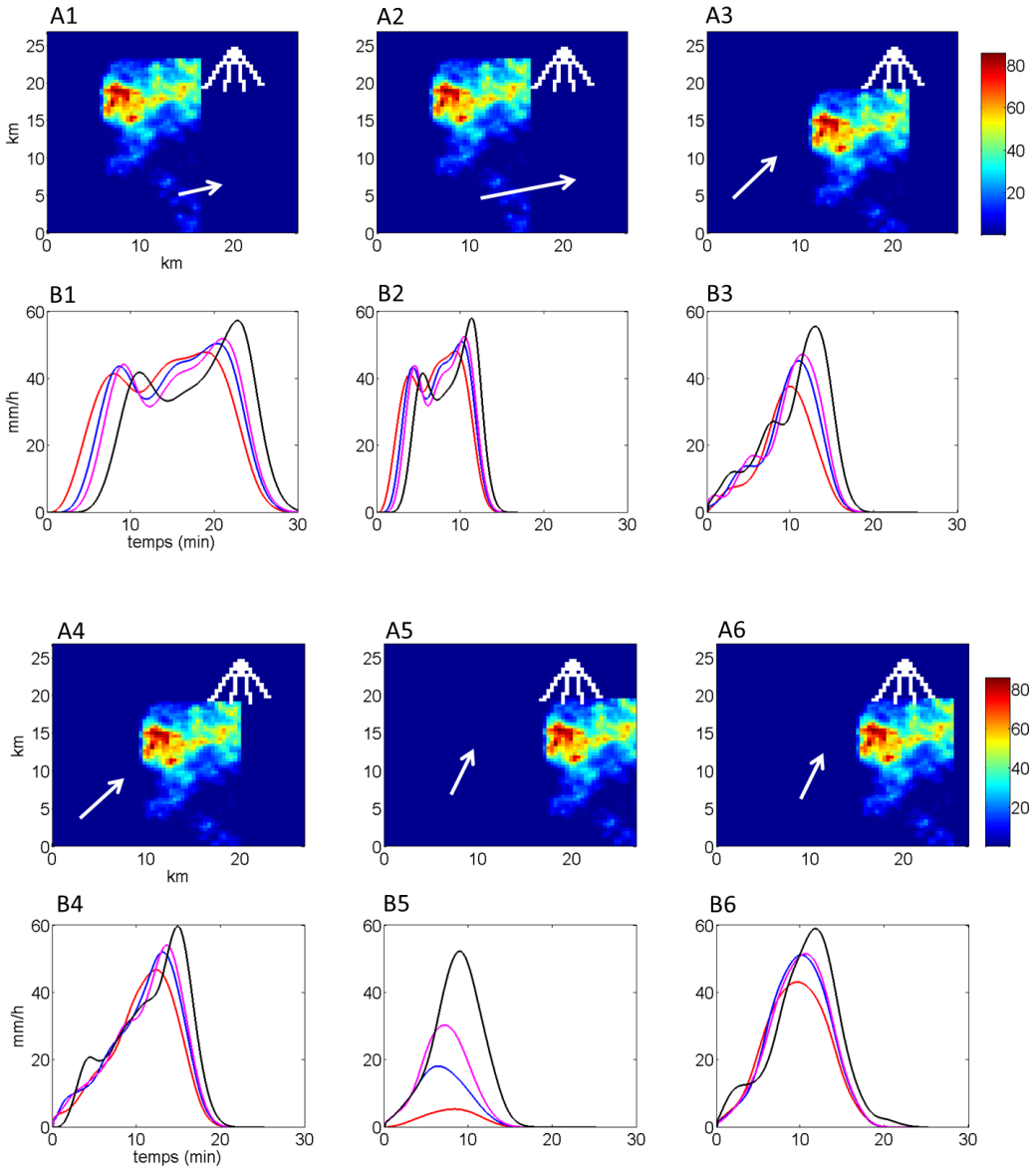

FIGURE G.6 – Visualisation des 6 simulations réalisées pour tester la méthode de triangulation de restitution du vent avec 1 capteur Ku et 4 liaisons. $A1$ à $A6$: les liaisons (simulées, voir Tableau G.1) sont représentées en blanc au nord-est de chaque carte. La cellule de pluie à l'instant initiale est également représentée, tout comme la vitesse d'advection (flèches blanches). $A1$ à $A6$: séries temporelles d'atténuations enregistrées par chaque liaison Ku lors de la l'advection de la cellule pluvieuse. On a sur chaque graphe : rouge : liaison pointant le plus à l'ouest ; bleu : liaison pointant sud-sud-ouest ; magenta : liaison pointant sud-sud-est ; noir : liaison pointant le plus à l'est.

simulé ces données avec $\delta = 250^\circ$ et $v = 20m/s$. Or si on regarde la ligne correspondant à 250°, on constate que le minimum pour J^{vent} se situe autour de $16.5m/s$. Un résultat raisonnable mais pas rigoureusement exact.

On a par ailleurs tracé sur cette figure la courbe définie par l'Eq.(G.29). Cette équation fournissait (toujours dans un cadre théorique idéal bien sûr), pour toute direction (tout δ), et connaissant au préalable un couple (δ_{min}, v_{min}) minimisant J^{vent}, la valeur de v fournissant le mimimum global de $J^{vent}(\delta, v)$. Cette courbe est tracée en blanc sur la figure, soit en prenant

pour (δ_{min}, v_{min}) les valeurs réelles simulées $(250^\circ, 20m/s)$ (courbe blanche de droite), soit en prenant pour (δ_{min}, v_{min}) une valeur lue sur la figure elle-même $(\delta_{min} = 250^\circ ; v_{min} = 16.6m/s)$. On remarque aussi que l'Eq.(G.29) fournit aussi une bonne estimation de la courbe minimisante dans un cas non idéal (courbe de gauche) ; et que la méthode donne des résultats légèrement décalés par rapport à la réalité simulée (courbe de droite).

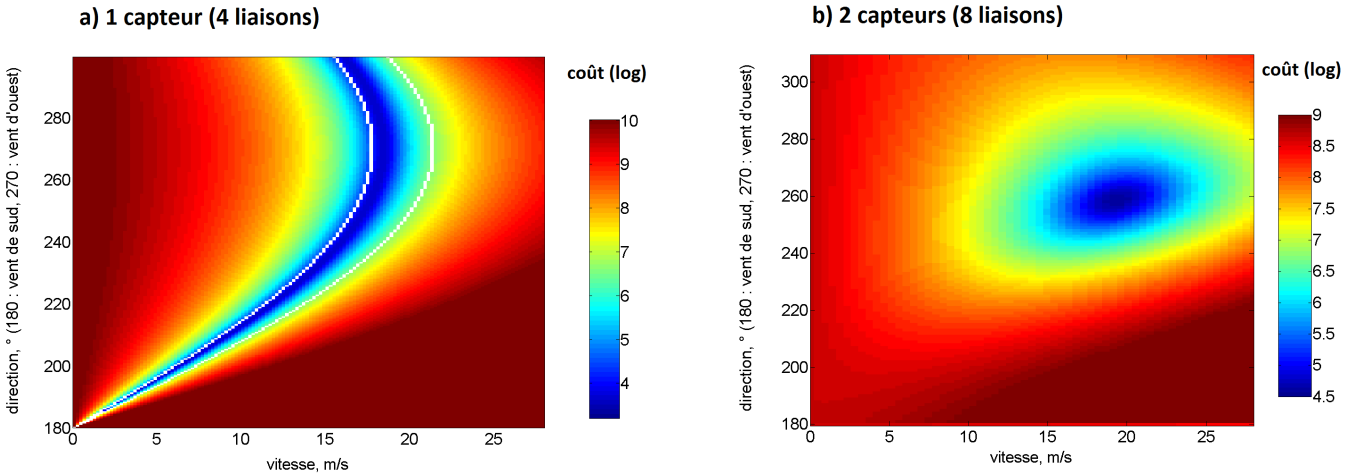

FIGURE G.7 – Evaluation de la fonction de coût permettant l'évaluation des conditions d'advection (Eq.(G.1)) avec 1 (a) ou 2 (b) capteurs Ku, dans le cadre de la simulation 2 (voir Tableau G.2), et de la configuration A pour le cas à deux capteurs (voir Tableau G.3).

On constate donc bien qu'il est vain d'espérer retrouver vitesse et direction du vent avec un seul capteur. Nous pouvons néanmoins tenter de déterminer l'un en supposant l'autre connu. Ces résultats sont présentés dans le Tableau G.2 (colonne 4). On suppose connues les directions, et on tente de restituer les vitesses en minimisant la fonction $J^{vent}(v)$. On retrouve ainsi bien pour la simulation 2 la valeur de $16.6m/s$ évoquée plus haut. On remarque à partir de ces simulations (bien que leur faible nombre limite les généralisations possibles) que la restitution de la vitesse est d'autant plus difficile que les cellules ne survolent qu'imparfaitement les liaisons (simulations 3 et 5) et que l'advection se fait selon l'axe sud-nord (simulations 5 et 6). Ces résultats sont peu étonnants. Les corrélations entre les séries temporelles (vecteur $\boldsymbol{\tau}$) sont moins bonnes dans ces cas-là. Ces résultats seront confirmés (quoique moins marqués) avec l'emploi de 2 capteurs.

2 capteurs Ku : succès !

Nous allons maintenant effectuer le même travail en supposant que l'on dispose de deux capteurs Ku identiques, enregistrant chacun les émissions des 4 mêmes satellites. On dispose donc de $2*4$ liaisons, dont les caractéristiques sont toujours celles données dans le Tableau G.1.

La première question à se poser est naturellement de savoir où placer le second capteur Ku par rapport au premier. Nous allons tester ici 4 configurations (A, B, C et D), dont les caractéristiques sont fournies dans le Tableau G.3.

TABLE G.3 – Positionnement du second capteur Ku par rapport au premier dans le cadre des 4 configurations définies ici.

	config A	config B	config C	config D
décalage sud	$2km$	$2.4km$	$1.2km$	$1.6km$
décalage est	$0.8km$	$1.2km$	$0.4km$	$1.6km$

Le second capteur est ainsi toujours placé au sud-est du premier, pas trop loin pour qu'il voit les mêmes cellules de pluie (sans quoi la corrélation évaluée par le vecteur $\boldsymbol{\tau}$ serait nulle entre les liaisons des deux capteurs et l'on ne gagnerait rien par rapport à l'emploi d'un seul capteur). On réalise ensuite, pour chacune de ces 4 configurations, les 6 mêmes simulations que précédemment (décrites dans le Tableau G.2). La Figure G.8 présente, comme la Figure G.6, la position des liaisons, direction d'advection et cellule à l'instant initial (a) et les enregistrements simulés le long des 8 liaisons sous ces conditions (b). Ceci n'est présenté sur cette Figure G.8 que pour la simulation 2 et la configuration A.

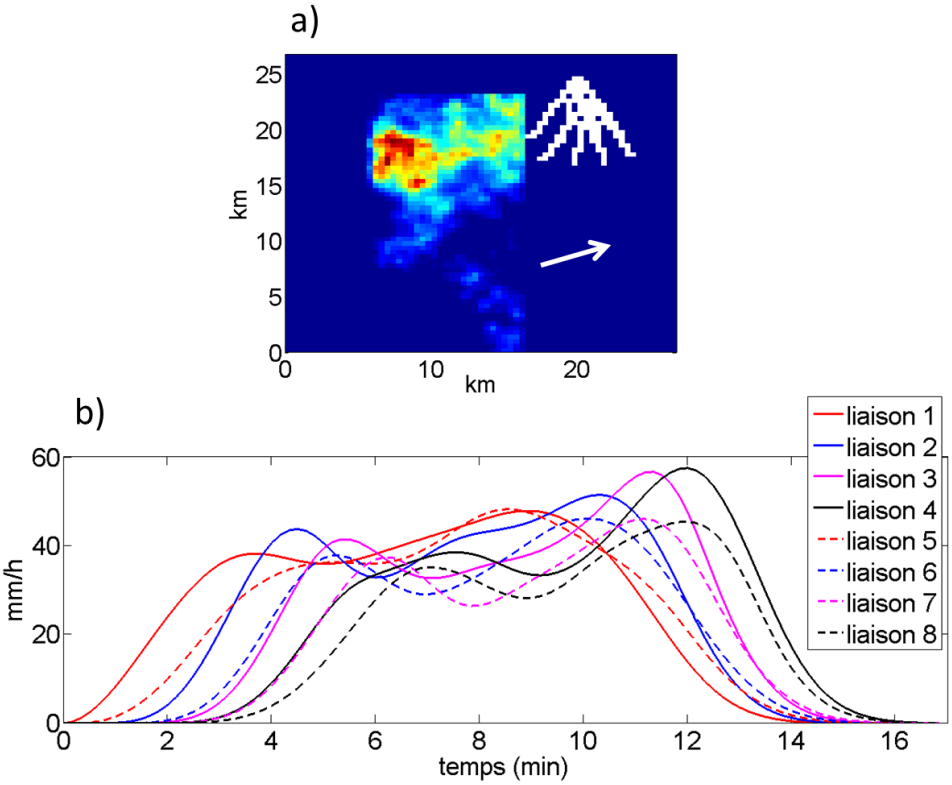

FIGURE G.8 – Idem Figure G.6 mais avec deux capteurs Ku, dans le cadre de la simulation 2 (Tableau G.2), et de la configuration A (Tableau G.3). Les liaisons 1 à 4 proviennent du 1^{er} capteur et sont dirigées d'ouest en est. Les liaisons 5 à 8 sont celles issues du 2^{nd} capteur.

Le Tableau G.4 présente les restitutions de vent pour chacune des 6 simulations et chacune des 4 configurations. De manière générale, on arrive ici à restituer à la fois la vitesse et la direction du vent de manière très satisfaisante.

Au niveau des différentes simulations, on constate, comme précédemment, que l'on obtient

les meilleurs résultats avec un vent d'ouest (simulations 1 et 2), et que l'augmentation de la vitesse d'advection ne perturbe pas la restitution (simulation 2 par rapport à 1). On constate ensuite que dans le cadre des simulations 3 et 5, où les cellules ne survolent qu'imparfaitement les liaisons, les restitutions de direction sont plus difficiles (par rapport aux simulations 4 et 6) mais que les restitutions de vitesse semblent moins perturbées (simulation 4 surtout), sûrement car même si toutes les liaisons ne sont pas touchées par les cellules, un nombre suffisamment grand l'est. Les restitutions sont également toujours un peu moins bonnes pour une advection majoritairement sud-nord (simulations 5 et 6), même si là aussi l'ajout d'un second capteur Ku limite les dégâts.

Au niveau des différentes configurations, on a ajouté dans le Tableau G.4, dernière ligne, l'écart moyen (en valeur absolue pour chaque configuration) entre les valeurs restituées et les valeurs réelles simulées. On obtient des résultats légèrement meilleurs (pour la direction comme pour la vitesse) avec la configuration B, correspondant au cas où les deux capteurs sont les plus éloignés. Les configurations avec des capteurs très rapprochés (configuration C) ou où le second est très décalé vers l'est (configuration D) donnent de légèrement moins bons résultats. On note par ailleurs que pour la restitution des cartes de pluie par assimilation de données, il est également avantageux d'avoir des capteurs suffisamment proches pour que les liaisons se croisent (ce qui est limité dans la configuration B). Nous concluons donc ici à l'utilisation de la configuration A, avec un second capteur décalé de $2km$ vers le sud et $0.8km$ vers l'est par rapport au premier.

Table G.4 – Valeurs des directions et des vitesses de vent restituées par la méthode de triangulation avec deux capteurs Ku, pour les simulations décrites dans le Tableau G.2 et les configurations décrites dans le Tableau G.3. On calcule aussi, pour chaque configuration, la moyenne des écarts absolus entre valeurs restituées et réalité simulée.

	réalité simulée		config A		config B		config C		config D	
simulation	δ ($^\circ$)	$v(m/s)$	δ ($^\circ$)	$v(m/s)$	δ ($^\circ$)	$v(m/s)$	δ ($^\circ$)	$v(m/s)$	δ ($^\circ$)	$v(m/s)$
1	250	10	258	9.3	258	9.8	259	9.3	259	9.8
2	250	20	258	19.1	258	19.6	261	19.6	258	19.1
3	225	14	236	14.5	232	14.0	239	14.0	235	14.5
4	225	14	225	14.0	226	13.1	226	14.9	225	13.1
5	200	13	214	10.3	213	11.7	209	8.4	220	9.3
6	200	13	208	13.1	200	13.5	210	12.6	200	12.1
erreur absolue moyenne			5.6	0.74	4.7	0.43	6.0	1.2	6.4	1.0

On note cependant pour terminer que dans toutes ces simulations numériques, l'hypothèse fondamentale de cette méthode de restitution du vent, à savoir que les cellules se déplacent effectivement sans déformation au gré du vent, était "presque" vérifiée, "presque" à la diffusion numérique induite par le modèle d'advection près (voir chapitre II). Dans les cas réels, notre objectif est de restituer une vitesse d'advection constante et homogène alors même que les

cellules ne restent pas strictement intactes, et ne se déplacent pas strictement à vitesse constante et homogène. L'erreur induite par ces phénomènes n'est pas évaluée ici et demande à être testée sur des cas réels où les données de deux capteurs Ku seraient disponibles. On note à nouveau que l'on aurait pu tenter d'améliorer cette méthode par exemple en pondérant le poids donné à chaque couple de liaison dans l'Eq.(G.1) par la valeur du maximum de corrélation entre les séries temporelles mesurées.

Annexe H: Évaluation rapide de l'impact de quelques erreurs modèle sur la restitution de cartes de précipitations

La section 5 du chapitre IV évalue l'algorithme de restitution des cartes de précipitations à partir de mesures d'atténuations d'ondes TV-SAT sur données simulées. En particulier, elle propose une étude sur 15 cas simulés, différenciant les cas où l'on dispose d'1 ou de 2 capteurs Ku. Dans cette section, nous n'évaluions cependant pas l'impact potentiel d'erreurs modèles. Celles-ci peuvent par exemple porter sur la longueur des liaisons Ku projetées au sol, mal connue (il existe une erreur sur l'altitude de l'isotherme 0 $^\circ$), ou encore sur la vitesse d'advection, déterminée à partir de la corrélation entre cartes radar successives. Dans cette Annexe, nous testons la pertinence de notre algorithme dans diverses situations, notamment en cas d'erreurs sur ces paramètres.

Le fait d'effectuer une expérience d'assimilation sur données simulées est appelé expérience jumelle. Son principe est décrit au chapitre IV, section 5.1 (voir notamment la Figure IV.3).

Dans cette annexe, nous effectuerons d'abord une expérience jumelle dans le cas où il n'y a aucune erreur. Dans ce cas-là, avec les notations du chapitre IV, section 5, on a : $\mathcal{A}^{vrai} = \mathcal{A}^{real}$. Ensuite, nous introduirons une erreur sur la longueur des liaisons Ku et la vitesse d'advection. Nous verrons enfin l'intérêt d'ajouter un terme de filtrage (terme décrit au chapitre IV, section 3.3).

Dans toute cette Annexe, nous utilisons la même carte de pluie initiale (A_0^{vrai}). Celle-ci est présentée par exemple sur la Figure H.1A. Par ailleurs, dans les sections 1, 2 et 3 de cette Annexe, cette carte sera toujours advectée vers l'est (soit formellement $\delta = 180^\circ$). Ceci correspond à la situation présentée sur la Figure H.1A. Dans la section 4, elle est advectée vers le nord ($\delta = 270^\circ$).

1 Expérience jumelle parfaite

Dans un premier temps donc, nous supposons parfaitement connues la géométrie des liaisons (décrites dans le Tableau IV.3 du chapitre IV et représentées sur la Figure H.1) et la vitesse d'advection (fixée à $16.5m/s$ vers l'est). Les observations ne sont pas bruitées ($\mathcal{A}^{real} = \mathcal{A}^{vrai}$). Cette expérience a deux objectifs. Tout d'abord, il s'agit de vérifier la fonctionnalité de l'algorithme (erreurs de code). Ensuite, elle permet d'étudier si le problème est sous-déterminé, c'est-à-dire si plusieurs solutions (plusieurs A_0^{assim}) sensiblement différentes peuvent être de "bonnes" solutions (c'est-à-dire produire un coût très faible après assimilation).

La Figure H.1 présente l'état initial vrai (A_0^{vrai}, A) et assimilé (A_0^{assim}, B), ainsi que les paramètres du problème (géométrie des liaisons et vecteur d'advection) et le tracé de la grande zone d'assimilation (cadres blancs). On remarque que dans cette zone, on reproduit bien les caractéristiques générales du champ vrai, même s'il semble exister des différences résiduelles.

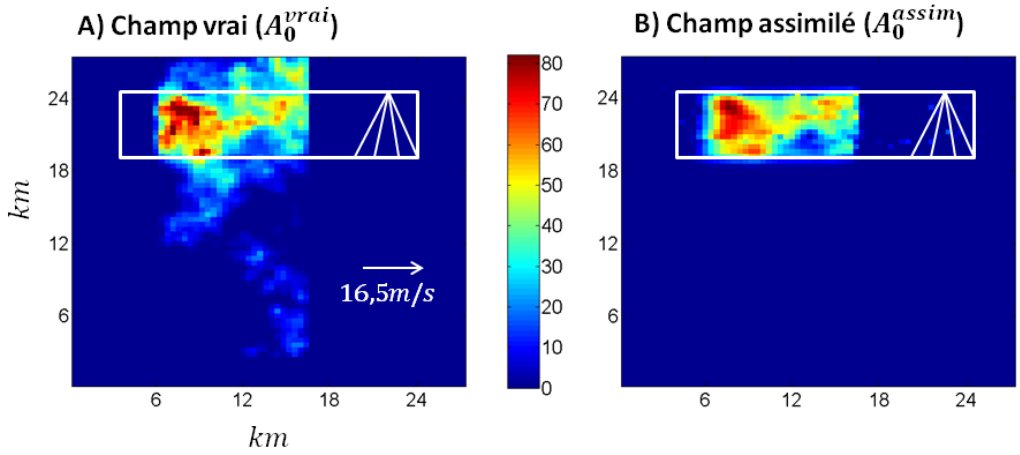

Figure H.1 – Expérience jumelle parfaite. A) Champ initial vrai (A_0^{vrai}, cellule de pluie simulée), géométrie des liaisons Ku projetées au sol (traits blanc au nord-est), conditions d'advection (flèche et vitesse) et grande zone d'assimilation (cadre blanc). B) Champ initial assimilé (A_0^{assi}), géométrie des liaisons et grande zone d'assimilation. Les liaisons Ku sont dessinées à titre indicatif (angles et longueurs ne sont pas très précisément respectés).

La Figure H.2 projette les champs vrais et assimilés (restreints à la grande zone d'assimilation) sur les axes est-ouest (A) et nord-sud (B). On remarque que la variabilité du champ vrai selon l'axe est-ouest est très bien restituée, mais que les résultats sont moins concluants selon l'axe nord-sud. Cela s'explique par la géométrie des liaisons. Les observations sont en effet intégrées le long des liaisons, qui pointent toutes globalement vers le sud. L'algorithme a donc beaucoup plus de mal à discriminer la pluie dans cette direction que selon l'axe est-ouest, où les liaisons sont touchées les unes après les autres.

La Figure H.3 enfin présente, pour chacune des 4 liaisons Ku, les observations ($\mathcal{A}^{vrai}$ et $\mathcal{A}^{assim}$, A) et l'écart entre les observations vraies et assimilées (B). On constate que les observations vraies sont presque parfaitement restituées par l'algorithme (le coût final est très

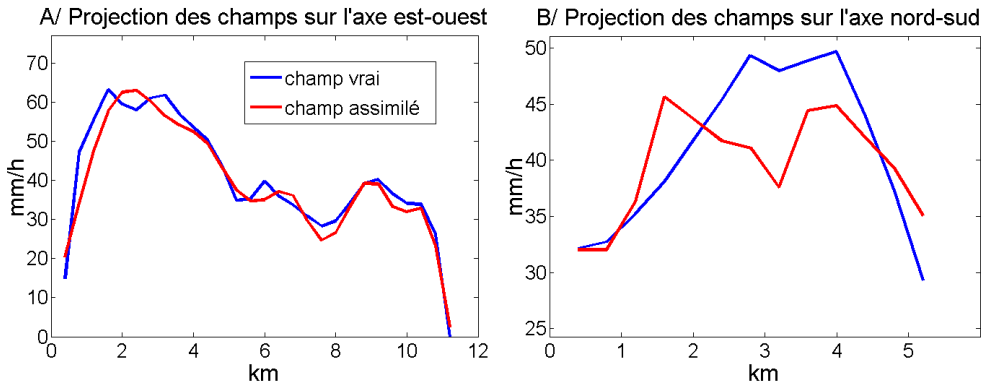

FIGURE H.2 – Expérience jumelle parfaite. Projection (i.e. pluviométrie moyenne) des champs vrais et assimilés (A_0^{vrai} et A_0^{assi}) restreints à la grande zone d'assimilation selon les axes est-ouest (A) et nord-sud (B).

faible). Le coût résiduel ne montre également pas de structure particulière non expliquée, mais des oscillations résiduelles. On en conclut que l'algorithme a été performant (dans ce cas idéal) mais que plusieurs solutions (plusieurs champs A_0^{assim}) étaient susceptibles d'expliquer les observations. Tous ces résultats (sous-détermination du problème par ambiguïté sur l'axe nord-sud) nous incitent à suggérer l'utilisation d'un second capteur Ku.

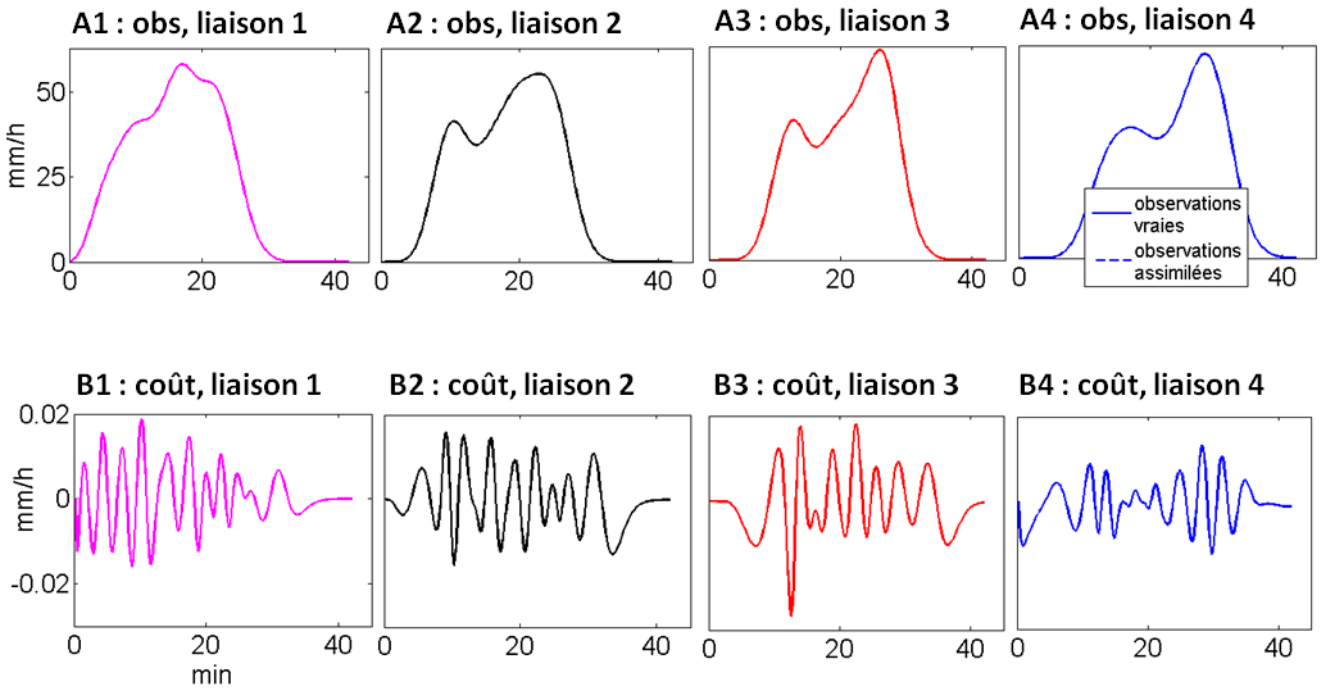

FIGURE H.3 – Expérience jumelle parfaite. A) Pour chaque liaison, séries temporelles d'observations vraies ($\mathcal{A}^{vrai}$, trait continu) et assimilées ($\mathcal{A}^{assi}$, trait pointillé, superposé). B) Séries temporelles des différences entre observations vraies et assimilées ($\mathcal{A}^{vrai} - \mathcal{A}^{assi}$). Les liaisons sont numérotées d'ouest en est.

2 Erreur sur la longueur des liaisons

On a vu que la longueur des liaisons Ku dépendait de l'iso-0. Or celui-ci n'est généralement connu qu'avec une certaine erreur, soit via une mesure au sol et un gradient type de température, soit via des images radar. Il paraît donc intéressant de voir la réaction de l'algorithme d'assimilation à une erreur sur la longueur des liaisons projetées au sol. On suppose toujours ici que les observations sont parfaites ($\mathcal{A}^{vrai} = \mathcal{A}^{real}$). On suppose également toujours parfaitement connue la vitesse d'advection (en l'occurrence $16.5m/s$ vers l'est). En revanche, les observations vraies seront produites avec les liaisons décrites plus haut (Tableau IV.3) et l'assimilation sera réalisée avec des liaisons plus courtes de $0.5km$. On introduit donc une erreur d'un peu moins de 10% sur la longueur des liaisons. Cela correspond à une situation où l'on estime l'iso-0 à environ $3300m$, alors qu'il est en fait à environ $3600m$.

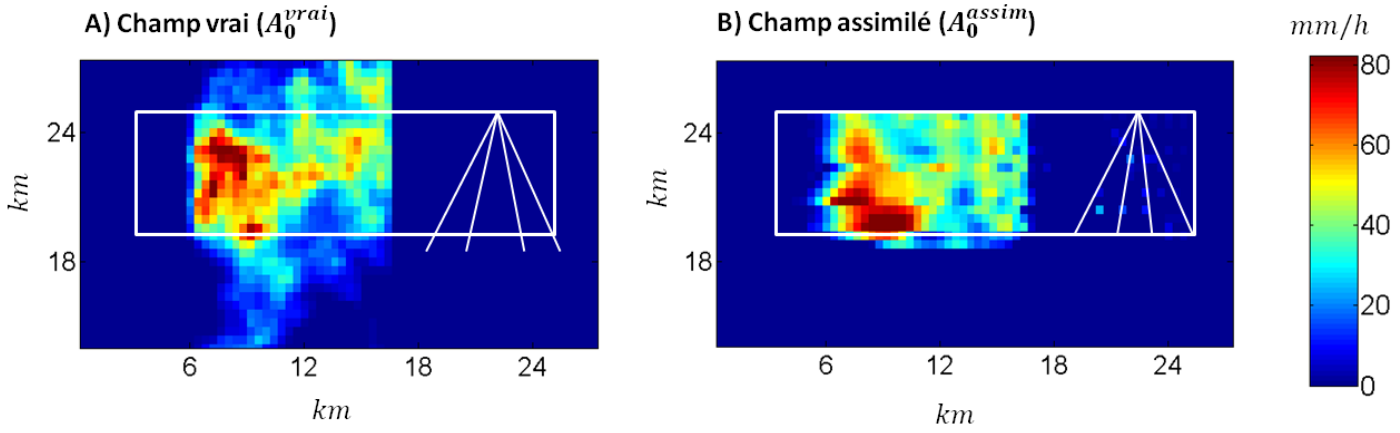

FIGURE H.4 – Expérience jumelle avec erreur sur la longueur des liaisons Ku. A) Champ initial vrai (A_0^{vrai}), géométrie des liaisons Ku (traits blanc à l'est) et grande zone d'assimilation (cadre blanc). B) Champ initial assimilé (A_0^{assi}), géométrie des liaisons et grande zone d'assimilation. Les liaisons Ku sont dessinées à titre indicatif.

La Figure H.4 présente les champs initiaux vrai et assimilé sous ces conditions. On remarque :

- que la pluviométrie moyenne sur la grande zone d'assimilation (cadre blanc, défini par rapport aux liaisons estimées et non vraies) est bien reproduite par l'algorithme ($41.7mm/h$ assimilé contre $40.7mm/h$ vrai sur la zone couverte par la cellule).

- que, comme précédemment, la variabilité ouest-est est également bien reproduite.

- que la restitution de la variabilité nord-sud est rendue encore plus difficile par l'erreur sur la longueur des liaisons. En particulier, l'algorithme surestime largement la pluviométrie au sud de la cellule (notamment vers la coordonnée $9km$ est / $20km$ nord). Ceci est aisément explicable. On a fourni à l'algorithme d'assimilation (via $\mathcal{A}^{vrai}$) des mesures de pluie venant d'une zone plus étendue vers le sud que celle effectivement couverte par les liaisons dont il dispose. En conséquence, il "place" cette eau en plus le plus au sud possible (puisque sur le champ vrai elle vient bien du sud de la zone) à l'intérieur de la zone d'assimilation, surestimant la pluviométrie à ce niveau-là.

- qu'en conséquence, la pluviométrie maximale est surestimée ($105.9mm/h$ au lieu de $85.6mm/h$).

Avoir une bonne estimation de l'altitude de l'iso-0 est donc un point important du travail préparatoire à réaliser. Dans le travail sur données réelles réalisé ici (chapitre IV, section 6), cependant, ce point ne devrait pas être critique puisque l'on dispose à la fois de mesures de température au sol et de relevés d'un radar à visée verticale.

3 Erreur sur la vitesse de déplacement

Même en supposant le modèle parfait, il n'est pas évident de restituer la vitesse d'advection (voir par exemple en Annexe G). Il est donc intéressant d'estimer la performance de l'algorithme en cas d'erreur sur ce paramètre. Comme précédemment, nous supposerons toujours les observations non bruitées ($\mathcal{A}^{vrai} = \mathcal{A}^{real}$). Nous supposerons connue la géométrie des liaisons. Les observations vraies seront produites en advectant le champ initial (A_0^{vrai}) à $16.5m/s$ vers l'est. Nous assimilerons ensuite les observations vraies en utilisant une vitesse d'advection toujours vers l'est mais à $14m/s$ (sous-estimation de v) ou à $18m/s$ (surestimation de v) .

La Figure H.5 présente les champs initiaux vrais et assimilés sous ces conditions, à savoir : sous-estimation de v (B) et surestimation de v (C). On remarque :

- que la pluviométrie moyenne est bien reproduite dans tous les cas.

- pour une sous-estimation de v, que la cellule de pluie assimilée est décalée vers l'est, contractée, et que les zones à forte pluviométrie sont (légèrement) décalées vers le nord.

- pour une surestimation de v, que la cellule de pluie assimilée est décalée vers l'ouest, étirée, et que les zones à forte pluviométrie sont décalées vers le sud.

Ces phénomènes sont là encore aisément explicables. Lorsque la vitesse d'advection fournie à l'algorithme est plus faible que la vitesse réelle (B), la cellule va passer sur les liaisons Ku plus tard (d'où un décalage vers l'est) et plus lentement (d'où la contraction). Lorsque la vitesse est surestimée, on a l'effet inverse.

Concernant le décalage nord / sud, il s'explique par les mêmes phénomènes. Les portions des liaisons Ku au sud de la zone d'assimilation sont plus éloignées les unes des autres que les portions au nord. Le temps mis par la cellule pour passer de l'une à l'autre sera donc le même au nord avec une vitesse de déplacement faible qu'au sud avec une vitesse de déplacement forte. On note que cela implique à nouveau que l'algorithme a du mal à discriminer la variabilité nord/sud de la cellule.

La Figure H.6 montre les écarts entre les observations vraies et assimilées (la somme du carré de ces séries temporelles donne la valeur finale de J^{obs}). Elle est donc comparable à la Figure H.3B. On note cette fois-ci, surtout dans le cas d'une surestimation de la vitesse

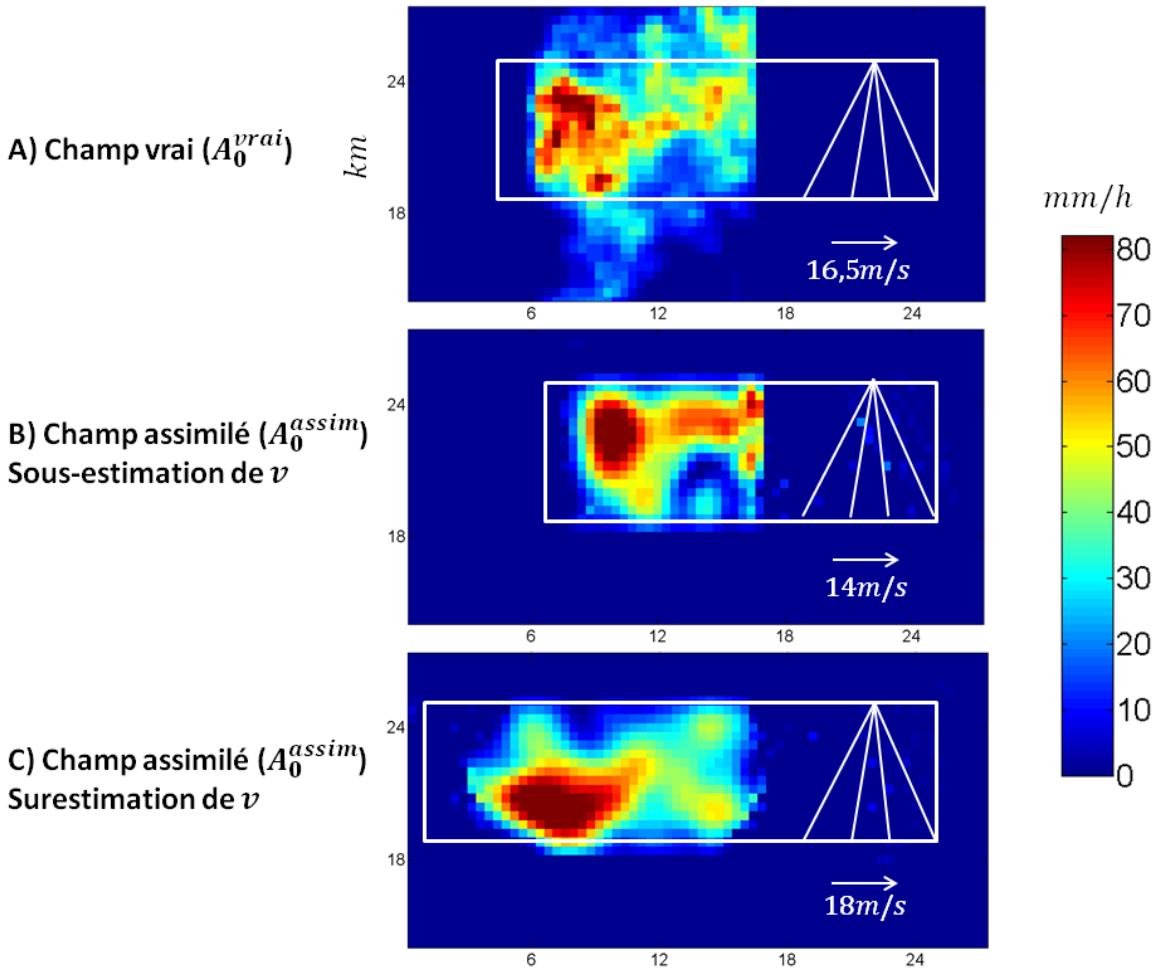

Figure H.5 – Expérience jumelle avec erreur sur la vitesse d'advection v. A) Champ initial vrai (A_0^{vrai}), géométrie des liaisons Ku (traits blanc à l'est), conditions d'advection vraies (flèche et vitesse) et grande zone d'assimilation (cadre blanc). B) Champ initial assimilé (A_0^{assi}) avec sous-estimation de v, géométrie des liaisons et grande zone d'assimilation. C) Idem avec surestimation de v. Les liaisons Ku sont dessinées à titre indicatif.

d'advection (liaison 3 notamment), qu'il existe un peu de signal (variabilité non expliquée) dans ces écarts et plus seulement de très faibles oscillations. L'écart reste cependant faible comparé aux valeurs de pluviométrie. Cela implique qu'une erreur significative sur la vitesse d'advection peut conduire à une erreur forte sur le champ initial restitué (Figure H.5) mais une erreur très faible sur l'écart aux observations. Le problème est donc bien sous-contraint, et la qualité des restitutions dépendra fortement de la qualité de l'estimation des paramètres du problème.

Ces remarques montrent qu'avec un seul capteur Ku disponible, la qualité des restitutions sur une zone étendue sera fortement conditionnée à la qualité de la vitesse d'advection estimée, et donc, on peut l'imaginer, au caractère advectif ou non de l'évènement.

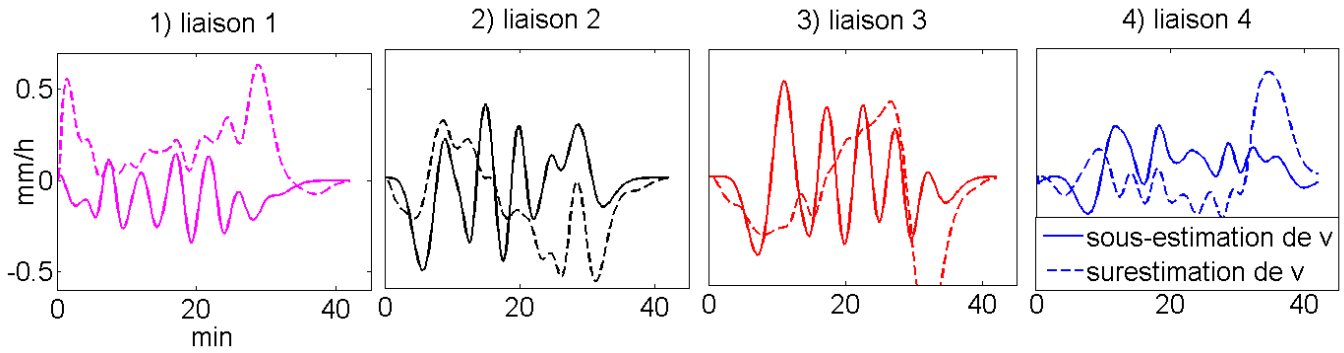

Figure H.6 – Expérience jumelle avec erreur sur la vitesse d'advection v. Pour chaque liaison, séries temporelles des différences entre observations vraies et assimilées dans le cas d'une sous-estimation de la vitesse d'advection (traits continus) ou d'une surestimation (traits pointillés). Les liaisons sont numérotées d'ouest en est.

4 Intérêt du filtrage

Nous avons jusqu'à présent étudié des cas où l'advection se faisait d'ouest en est. Nous allons maintenant travailler avec une advection du sud vers le nord ($\delta = 180^\circ$). Nous utilisons les mêmes liaisons Ku, et déplaçons donc la cellule de pluie vers le nord, avec $v = 13m/s$. Le champ initial vrai (A_0^{vrai}) est présenté sur la Figure H.7A, avec les liaisons et les conditions d'advection.

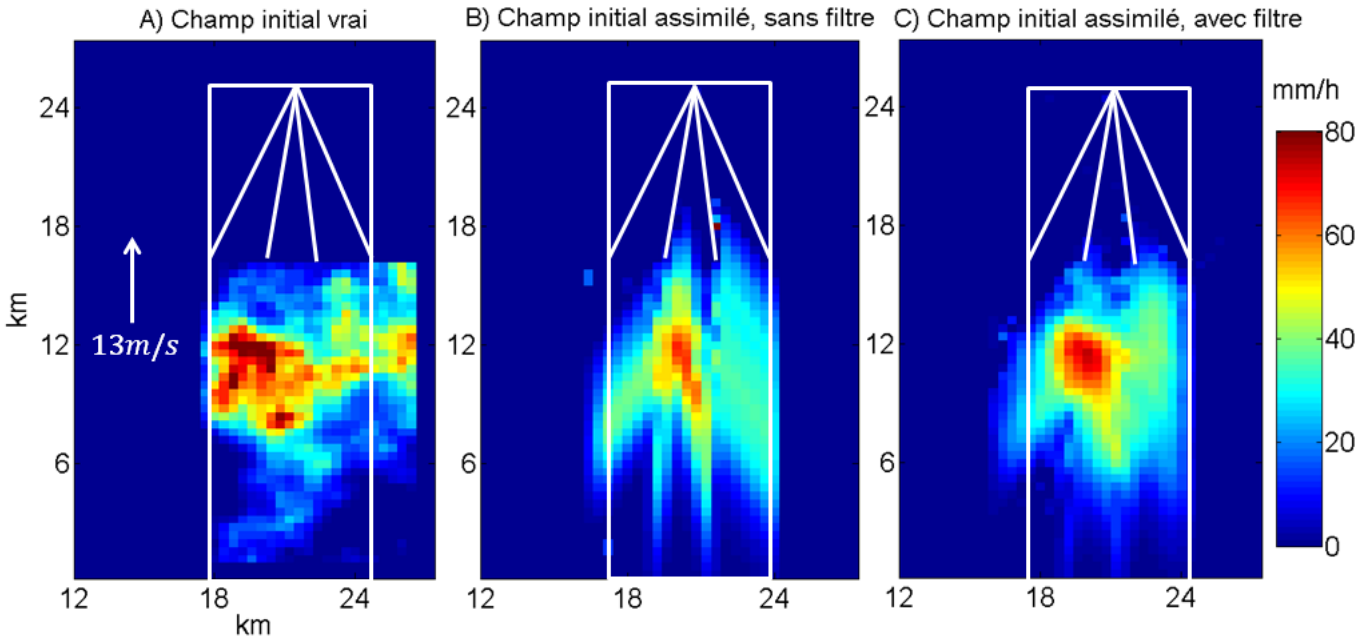

Figure H.7 – Expérience jumelle avec advection vers le nord. A) Champ initial vrai (A_0^{vrai}), géométrie des liaisons Ku (traits blanc au nord), conditions d'advection vraies (flèche et vitesse) et grande zone d'assimilation (cadre blanc). B) Champ initial assimilé (A_0^{assi}) sans terme de filtrage. C) Idem avec terme de filtrage. Les liaisons Ku sont dessinées à titre indicatif.

Nous assimilons ensuite les observations vraies toujours sans introduire de bruit ($\mathcal{A}^{vrai} = \mathcal{A}^{real}$) et sans erreur sur la longueur des liaisons ou la vitesse d'advection. Le résultat (A_0^{assi}) est présenté sur la Figure H.7B. L'assimilation est ici rendue plus difficile par la géométrie du problème. En effet, certaines portions du champ de pluie ne traversent une liaison que tardivement et durant un temps court (les portions du champ qui durant l'advection passent entre deux liaisons quasiment jusqu'au moment où elles touchent le capteur Ku lui-même). En conséquence, la forme du champ restitué est directement influencée par la géométrie des liaisons. (On note que cela n'empêche pas la bonne restitution de la quantité d'eau moyenne sur la zone couverte.)

Afin de limiter l'influence de la géométrie des liaisons, nous introduisons dans la fonction de coût le terme de filtrage (J^f) décrit dans le chapitre IV, section 3.3 (voir les Eqs. (IV.6) et (IV.7)). Celui-ci permet d'introduire de la corrélation directement dans la fonction de coût. On rappelle que le coefficient c^f (Eq. (IV.6)) permet de régler les poids relatifs du filtre et de l'écart aux observations ($J = J^{obs} + c^f J^f$). On règle ici empiriquement c^f de manière à ce que, appliqué au champ assimilé sans filtre (champ de la Figure H.7B), on ait environ : $c^f J^f \approx 1/5 J^{obs}$. Le résultat de l'expérience d'assimilation qui en découle est présenté sur la Figure H.7C. On remarque que le filtre a agi comme on s'y attendait, c'est-à-dire en propageant l'information apportée par les zones bien couvertes par les liaisons dans les zones voisines (en quelque sorte en "bouchant les trous" du champ assimilé sans filtre de la Figure H.7B). Le champ résultant est plus lisse que le champ vrai (par exemple, pluviométrie maximale à $74mm/h$ contre $86mm/h$, quantile 75% à $36mm/h$ contre $41mm/h$), mais la moyenne n'est pas affectée et la projection du champ sur les axes nord/sud comme ouest/est est assez satisfaisante.

Dans les expériences présentées dans le corps du manuscrit (chapitre IV, section 5.2 et section 6), nous ajoutons systématiquement un tel terme de filtrage. La valeur c^f_{cas} de c^f pour un cas donné sera réglée de la manière suivante. En notant c^f_{simu} le c^f utilisé ici et $J^{obs}_{0,simu}$ l'écart aux observations obtenu pour un champ initial nul ($A_0 = 0$), on prendra :

$$c^f_{cas} = \frac{J^{obs}_{0,cas}}{J^{obs}_{0,simu}} c^f_{simu} \tag{H.1}$$

Avec $J^{obs}_{0,cas}$ l'écart aux observations obtenu pour un champ initial nul pour le cas (réel ou simulé) considéré.

On définit donc le poids du filtre selon une valeur réglée empiriquement sur données simulées puis pondérée ensuite par la "masse" d'observations disponibles.

5 Conclusions

Nous avons illustré ici le comportement de l'algorithme d'assimilation dans quelques situations typiques. Nous avons ainsi montré que la géométrie des liaisons (orientation vers le

sud) rendait plus difficile la restitution de la variabilité des champs de pluie selon l'axe nord/sud pour des évènements se déplaçant d'ouest en est. Nous avons aussi montré qu'une erreur sur la longueur des liaisons Ku ou la vitesse d'advection n'empêchait pas la reproduction des observations mais entrainaît des erreurs assez significatives sur la répartition de la pluie dans la zone d'assimilation. Nous avons enfin illustré la nécessité qu'il peut y avoir d'ajouter un terme de filtrage.

Annexe I: Compléments sur la paramétrisation de la DSD en haut de la colonne dans VAMOS

Dans le chapitre V, nous cherchons à restituer (notamment) des DSD en haut de la colonne atmosphérique. Nous avons vu (section 2.2 dans ce chapitre) que ces DSD, afin de réduire le nombre de degrés de liberté du problème, étaient paramétrées sous forme de DSD gamma.

En se plaçant à une date t_k quelconque, la DSD correspondant, notée $(n_{k,0,p})_{p\in[|0:ND-1|]}$, s'écrit alors, pour la *p-ième* classe de diamètres :

$$n_{k,0,p} = \frac{\alpha}{\Delta D}\int_{D_p}^{D_p+\Delta D} pdf_\gamma(x;k,\theta)dx \tag{I.1}$$

Avec α un coefficient traduisant le nombre total de gouttes par unité de volume (unité : m^{-3}) et pdf_γ la densité de probabilités de la loi gamma, qui dépend de deux coefficients k et θ (voir chapitre V, section 2.2).

Dans cette annexe, nous donnerons d'abord la méthode utilisée pour implémenter cette paramétrisation dans le logiciel YAO. Ensuite, nous discuterons rapidement du lien entre les coefficients introduits ici et ceux utilisés par Ulbrich (1983) ou ceux utilisés par Testud et al. (2001).

1 Implémentation dans YAO

En reformulant l'Eq. (I.1), on peut également écrire :

$$n_{k,0,p} = \frac{\alpha}{\Delta D}\left[cdf_\gamma\left(D_p+\Delta D,k,\theta\right) - cdf_\gamma\left(D_p,k,\theta\right)\right] \tag{I.2}$$

Avec cdf_γ la fonction de répartition de la loi gamma, qui s'écrit :

$$cdf_\gamma(x;k,\theta) = \frac{1}{\Gamma(k)}\gamma\left(\frac{x}{\theta},k\right) \tag{I.3}$$

Avec γ la fonction gamma incomplète :

$$\gamma(x,k) = \int_0^x e^{-t} t^{k-1} dt \tag{I.4}$$

Forward Pour une date et une classe de diamètre données, $n_{k,0,p}$ sera donc évaluée à partir de trois entrées, les trois coefficients α_k, k_k et θ_k à la date t_k, via l'expression :

$$n_{k,0,p} = \frac{\alpha_k}{\Delta D} \frac{1}{\Gamma(k_k)} \left[\gamma\left(\frac{D_p + \Delta D}{\theta_k}, k_k \right) - \gamma\left(\frac{D_p}{\theta_k}, k_k \right) \right] \tag{I.5}$$

Dans YAO, la fonction gamma incomplète est implémentée via son expression sous forme de série :

$$\gamma(x,k) = x^k e^{-x} \sum_{m=0}^{m_{max}} \frac{x^m}{k(k+1)...(k+m)} \tag{I.6}$$

Avec m_{max} un nombre de termes assez grand pour assurer une convergence à 10^{-7} près.

Backward Afin de rétropropager le gradient de la fonction de coût jusqu'aux entrées (les trois coefficients à déterminer α_k, k_k et θ_k), il s'agit d'évaluer les dérivées partielles de $n_{k,0,p}$ par rapport à chacune d'entre elle, soit $\partial n_{k,0,p}/\partial \alpha_k$, $\partial n_{k,0,p}/\partial k_k$ et $\partial n_{k,0,p}/\partial \theta_k$.

La première s'écrit simplement sous la forme :

$$\frac{\partial n_{k,0,p}}{\partial \alpha_k} = \frac{1}{\Delta D \Gamma(k_k)} \left[\gamma\left(\frac{D_p + \Delta D}{\theta_k}, k_k \right) - \gamma\left(\frac{D_p}{\theta_k}, k_k \right) \right] \tag{I.7}$$

La dérivée par rapport à k_k s'écrit :

$$\frac{\partial n_{k,0,p}}{\partial k_k} = \frac{\alpha_k}{\Delta D} \frac{1}{\Gamma(k_k)} \left[\frac{\partial \gamma}{\partial k_k}\left(\frac{D_p + \Delta D}{\theta_k}, k_k \right) - \frac{\partial \gamma}{\partial k_k}\left(\frac{D_p}{\theta_k}, k_k \right) - \frac{\Gamma'(k)}{\Gamma(k)} \left(\gamma\left(\frac{D_p + \Delta D}{\theta_k}, k_k \right) - \gamma\left(\frac{D_p}{\theta_k}, k_k \right) \right) \right] \tag{I.8}$$

Dans YAO, la dérivée de la fonction gamma et la dérivée partielle de la fonction gamma incomplète par rapport à k_k seront évaluées numériquement au voisinage de k_k (par exemple par $\Gamma'(k_k) \approx \left(\Gamma(k_k + dk_k) - \Gamma(k_k)\right)/dk_k$ pour $dk_k = k_k/1000$).

La dérivée par rapport à θ_k s'écrit :

$$\frac{\partial n_{k,0,p}}{\partial \theta_k} = \frac{\alpha_k}{\Delta D} \frac{1}{\Gamma(k_k)} \left[f_1'(\theta_k) \frac{\partial \gamma}{\partial \theta_k}\left(f_1(\theta)\right) - f_2'(\theta_k) \frac{\partial \gamma}{\partial \theta_k}\left(f_2(\theta)\right) \right] \tag{I.9}$$

Avec :

$$\begin{cases} f_1(\theta) = \frac{D + \Delta D}{\theta} \\ f_2(\theta) = \frac{D}{\theta} \end{cases} \tag{I.10}$$

Comme précédemment, dans YAO, nous évaluerons numériquement la dérivée partielle de la fonction gamma incomplète par rapport à θ_k.

2 Lien expressions de Ulbrich (1983) et Testud et al. (2001)

Ulbrich (1983) exprime également la DSD (notée ici $n_{ulb}(D)$) sous forme de DSD gamma dépendant de 3 coefficients μ_{ulb}, Λ_{ulb} et No_{ulb} :

$$n_{ulb}(D) = No_{ulb} D_{ulb}^{\mu} e^{-\Lambda_{ulb} D} \tag{I.11}$$

Ces trois coefficients sont alors liés aux trois coefficients α, k et θ introduits précédemment :

$$\begin{cases} \mu_{ulb} = k - 1 \\ \Lambda_{ulb} = \frac{1}{\theta} \\ No_{ulb} = \frac{\alpha \Lambda_{ulb}^{k}}{\Gamma(k)} \end{cases} \tag{I.12}$$

Le diamètre volumique moyen D_m et le nombre normalisé de gouttes N_0^* introduit par Testud et al. (2001) s'écrivent eux :

$$\begin{cases} D_m = (k+3)\theta \\ N_0* = \frac{\alpha 4^4}{\Gamma(4)(k+3)} \frac{k(k+1)(k+2)}{\theta} \end{cases} \tag{I.13}$$

Annexe J: Surapprentissage dans VAMOS

Nous allons dans cette annexe rapidement revenir sur la minimisation numérique effectuée dans l'algorithme d'assimilation 4D-VAR VAMOS (restitution des DSD et vents verticaux) et les problèmes de surapprentissage qu'elle peut provoquer. Cette annexe se rapporte donc au chapitre V en général et aux sections 3 et 4 en particulier (sections où l'on présente, comme dans cette Annexe, des résultats sur données simulées).

Nous reprenons ici l'expérience jumelle effectuée dans la section 4.3.3 du chapitre V. On rappelle que l'on avait alors simulé nos états vrais avec un modèle de propagation $4D$ prenant en compte un vent horizontal, puis que l'on avait assimilé les observations ainsi simulées dans notre algorithme d'assimilation VAMOS sans vent horizontal. On présente ici des résultats obtenus pour le profil 2 de vent horizontal (croissance logarithmique sur la colonne atmosphérique, voir par exemple Figure V.14H). Les comportements généraux de l'algorithme présentés ici ne sont cependant pas spécifiques à ce problème et se retrouvent sur diverses expériences jumelles.

L'algorithme 4D-VAR fournit un champ de vent vertical et des DSD en haut de la colonne atmosphérique. Pour cela, il minimise la fonction de coût J grâce à un minimiseur numérique (voir chapitre III section 3.2.3). Celui-ci est itératif. L'algorithme part donc du first guess (vent vertical nul partout et DSD constante homogène très faible) et converge progressivement vers un minimum local de la fonction de coût. Nous pouvons donc, pour chaque itération de ce minimiseur, afficher les vents verticaux et les DSD correspondant.

La Figure J.1 présente cela pour le champ temps/altitude de vents verticaux. La première image (A) présente les vents verticaux vrais (w^{vrai}), ce que l'on cherche à restituer. Les images B à D présentent le même vent vertical pour les itérations 0 (first guess), 12 et 74 de la minimisation. L'image E présente le vent vertical assimilé (w^{assim}), correspondant à la $240^{ème}$ et dernière itération du minimiseur. La Figure J.2A présente elle, en fonction de l'itération du minimiseur, l'erreur (normalisée à 1 pour le first guess) commise sur la restitution de la DSD en haut de la colonne (formellement $\sum_k \sum_p \left(n^{vrai}_{k,0,p} - n^{mod}_{k,0,p}\right)^2$), avec n^{mod} la DSD de l'algorithme pour un itération quelconque du minimiseur. La Figure J.2B présente

le même résultat pour le vent vertical. La Figure J.2C présente l'évolution de la partie J^{obs} de la fonction de coût au cours de la minimisation.

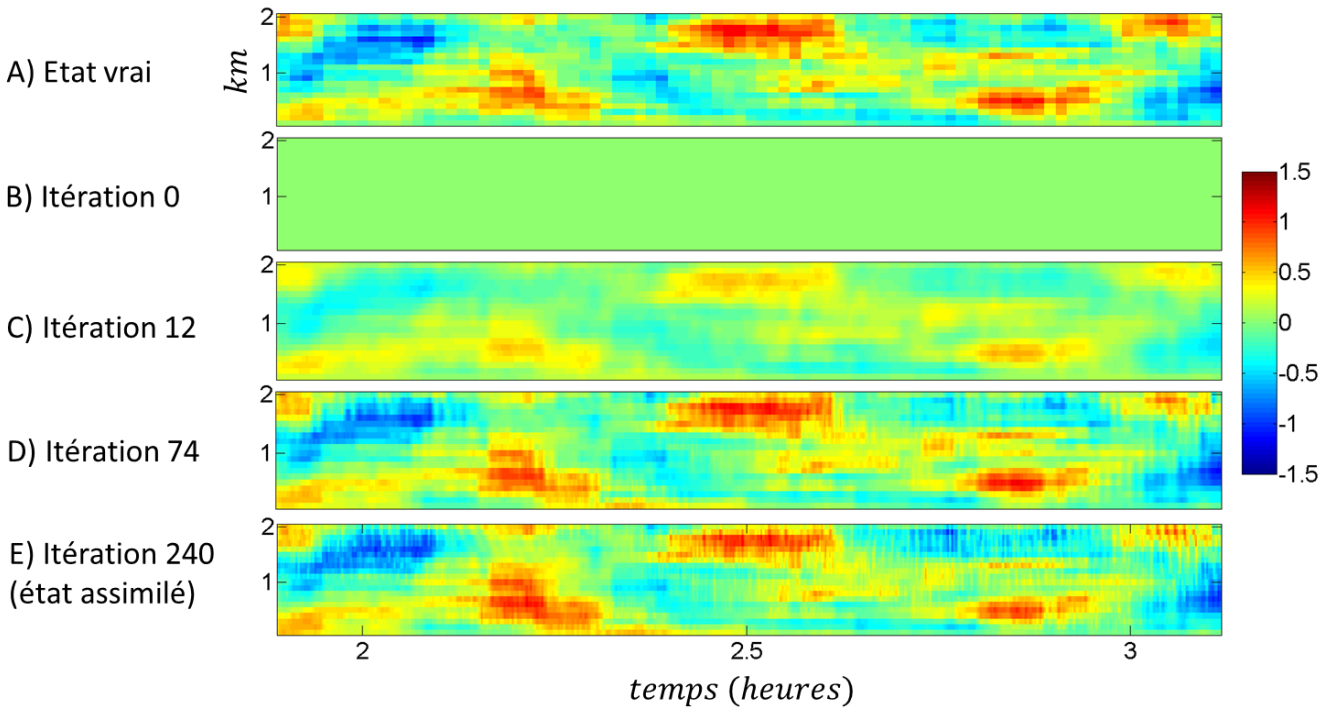

Figure J.1 – Illustration du processus de minimisation numérique de la fonction de coût du 4D-VAR pour le vent vertical. Expérience jumelle décrite au chapitre V, section 4.3, pour le profil 2 de vent horizontal. A) État vrai (w^{vrai}). B à D) Vent vertical pour les itérations 0, 12 et 74 de la minimisation. E) Etat assimilé (w^{assim}). Vents à la résolution des observations ($30s$) et positifs vers le bas.

On remarque (Figure J.2) que si le coût décroît bien sûr jusqu'à la fin de la minimisation, les erreurs commises sur le vent vertical et sur la DSD atteignent leur minimum aux alentours de l'itération 75 du minimiseur. Ce comportement est caractéristique d'un problème de surapprentissage. Pour mieux comprendre ce qu'il se passe, on va s'intéresser à la Figure J.1 qui fait ressortir les différentes phases de la minimisation :

- on part du first guess avec un vent vertical nul partout (Figure J.1B).

- on apprend ensuite très rapidement les structures du champ de vent vertical, sans restituer correctement encore leurs intensités (Figure J.1C).

- on ajuste ensuite relativement bien cette intensité (Figure J.1D), qui correspond à l'erreur minimale du vent vertical (voir Figure J.2C).

- on distord ensuite ce champ, faisant ainsi augmenter l'erreur du vent vertical (Figure J.1E) pour continuer à minimiser la fonction de coût.

La perturbation du champ de vent (et des DSD, même si ce n'est pas représenté ici explicitement) durant la deuxième phase de l'assimilation (après l'itération 75), diminue le coût (marginalement, voir Figure J.2C) en exploitant le grand nombre de degrés de liberté du modèle. Le modèle n'ayant pas la possibilité d'expliquer le vent horizontal, il dénature les

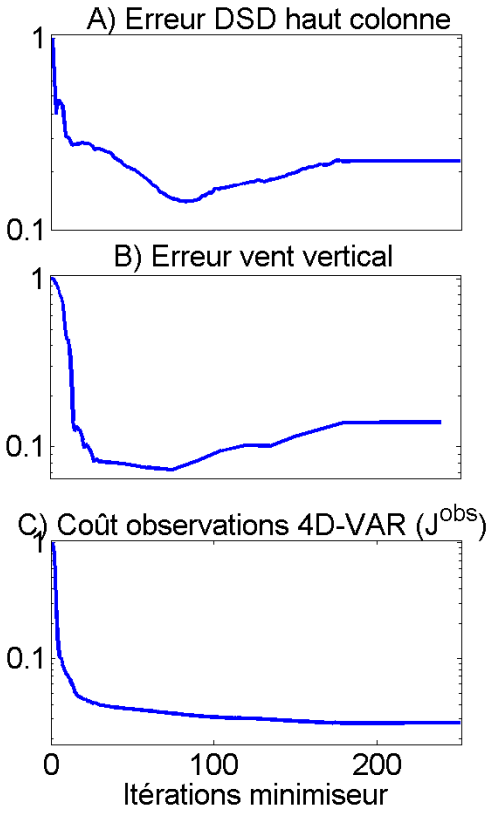

Figure J.2 – Selon l'itération du minimiseur numérique du 4D-VAR, erreur sur la DSD en haut de la colonne atmosphérique (A), sur le vent vertical (B) et partie observations de la fonction de coût (C). Même expérience jumelle que Figure J.1.

champs de vent et de DSD pour expliquer (très partiellement) le bruit modèle et le bruit ajouté aux observations.

On note :

- que la question de la perturbation des champs de vents restitués se pose également dans la section 3.3.6 du chapitre V sur la turbulence (voir cette section). Il ne s'agit cependant pas du même problème, la turbulence étant reproduite comme on le verra via des oscillations du vent vertical à la résolution du modèle ($5s$) tandis que l'on regarde bien ici des perturbations à la résolution des observations (typiquement entre $30s$ et $2min$).

- que l'on peut tenter d'évaluer le moment dans la minimisation (i.e. l'itération) à partir duquel on ne fait plus que perturber les restitutions en regardant la variabilité (l'écart-type) des champs normalisés. Par exemple sur la Figure J.1, si l'on normalise les champs représentés, leur écart-type au fil de la minimisation augmentera naturellement au début (passage de B à C), puis restera plus ou moins stable lors de la phase d'ajustement des intensités (passage de C à D), puis augmentera à nouveau dans la phase de surapprentissage (passage de D à E).

- que les filtres introduits (voir chapitre V, section 2.4) limitent ce phénomène. Le terme de filtrage va en effet pénaliser par exemple le champ de la Figure J.1E par rapport à celui de la Figure J.1D. En conséquence, la minimisation s'arrêtera lorsque le fait pour l'algorithme de

mieux expliquer le bruit en perturbant les champs sera compensé par le coût supplémentaire provoqué par le filtre. La section 3.2 du chapitre V revient sur le réglage des poids relatifs des termes de filtrage et de pénalisation dans la fonction de coût. Ceci dit, il est assez difficile de régler précisément le poids de ce terme de filtrage a priori. L'expérience jumelle présentée a par exemple été réalisée avec la procédure de la section 3.2, sans que le problème ne soit réglé comme on le voit.

Lorsque l'assimilation produit, que ce soit sur des expériences jumelles ou sur des données réelles, des sorties très perturbées, il faut garder à l'esprit le surapprentissage avant de chercher une explication physique aux résultats obtenus.

Annexe K: Correction de l'atténuation due au radôme des radars à visée verticale

Certains des radars à visée verticale utilisés dans le chapitre V sont couverts par un radôme (abri protégeant l'appareil). C'est notamment le cas du KAZR (utilisé en Finlande dans la section 6 et à Darwin dans la section 7).

Lorsqu'il pleut, de l'eau est présente sur ce radôme, et atténue très significativement le signal émis. Ce phénomène n'étant pas pris en compte dans le modèle, il convient de le corriger sur les observations fournies à l'algorithme.

Pour cela, nous utilisons les données du disdromètre. À partir de ses mesures, nous calculons (en appliquant l'Eq. (I.26), chapitre I, section 2.3.2), pour chaque date (chaque minute), la réflectivité à $35GHz$ correspondant (en supposant aucun vent vertical). Cette réflectivité est réprésentée sur la Figure K.1A, en bleu. Nous représentons également la réflectivité effectivement mesurée par le radar à sa plus basse porte disponible (rouge). En supposant que les DSD sont les mêmes entre les 2 altitudes et pour les 2 volumes sondés (et le vent et la turbulence nuls), les deux courbes devraient se correspondre, modulo l'atténuation subie par le signal durant son trajet $0 - 220m$ aller-retour. Cette atténuation est composée de trois parts : une due aux gaz, une due à la pluie, et une due au radôme. On néglige la première ($< 0.15dB$, cf. Tridon and Battaglia (2015)) et on estime la deuxième à partir des mesures disdromètres (via l'Eq.(I.22), chapitre I, section 2.3.1). Celle-ci (atténuation due à la pluie sur la couche $0 - 220m$) est représentée sur la Figure K.1B. On estime ensuite l'atténuation due au radôme, qui est représentée sur la Figure K.1C. Les résultats à $1min$ sont en bleu, la courbe noir correspondant à une moyenne mobile sur $7min$. On constate que l'atténuation est très forte (autour de $16dB$) et relativement variable dans le temps ($\pm 1dB$).

On note encore que cette façon de procéder possède deux désavantages majeurs :

- elle corrèle les mesures radar et disdromètre (données d'entrée de l'algorithme de restitution) avant même la phase d'assimilation. Cette corrélation ne concerne cependant que les intensités des spectres (et non leur forme) et leur variabilité temporelle (proche du sol) et non verticale (la correction est la même pour toutes les portes).

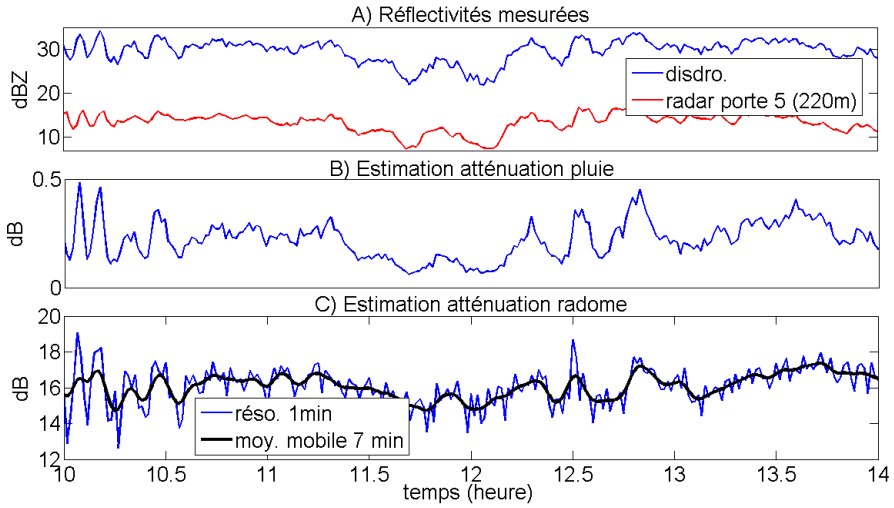

Figure K.1 – Estimation de l'atténuation due au radôme du KAZR à partir des mesures radar et disdromètre. Voir dans le texte pour les commentaires détaillés.

- par ses imprécisions (on suppose l'atmosphère figé sur les 200 premiers mètres, on néglige l'atténuation dûe aux gaz, etc.), elle est susceptible de créer un biais systématique sur chaque profil vertical restitué. Nous reviendrons plus loin sur cette question.

Il paraitrait donc très intéressant de ne plus corriger les atténuations radômes a priori mais de les contrôler dans l'algorithme d'assimilation. Dans la mesure où cette correction est la même pour toutes les portes et semble tout de même assez fortement corrélée dans le temps (malgré un bruit significatif, voir Figure K.1C), cette proposition ne parait pas déraisonnable (elle n'ajouterait a priori pas trop de degrés de liberté au système). Le travail requis pour cela (construction des opérateurs de corrélation en particulier) fait qu'on laisse cependant ce travail pour plus tard.

Annexe L: Article "Estimation of Finescale Rainfall Fields Using Broadcast TV Satellite Links and a 4DVAR Assimilation Method", paru Journal of Atmospheric and Oceanic Technology.

Estimation of Finescale Rainfall Fields Using Broadcast TV Satellite Links and a 4DVAR Assimilation Method

FRANÇOIS MERCIER, LAURENT BARTHÈS, AND CÉCILE MALLET

Université de Versailles Saint-Quentin-en-Yveline,Versailles, and Sorbonne Universites, UPMC University Paris 06, Paris, and CNRS/INSU, LATMOS-IPSL, Guyancourt, France

(Manuscript received 22 May 2014, in final form 23 April 2015)

ABSTRACT

This study proposes a method based on the use of a set of commercial satellite-to-Earth microwave links to rebuild finescale rainfall fields. Such microwave links exist all over the world and can be used to estimate the integrated rain attenuation over the links' first 5–7 km with a very high temporal resolution (10 s in the present case). The retrieval algorithm makes use of a four-dimensional variational data assimilation (4DVAR) method involving a numerical advection scheme. The advection velocity is recovered from the observations or from radar rainfall fields at successive time steps.

This technique has been successively applied to simulated 2D rain maps and to real data recorded in the autumn of 2013 during the Hydrological Cycle in the Mediterranean Experiment (HyMeX), with one sensor receiving microwave signals from four different satellites. The performance of this system is assessed and is compared to an operational Météo-France radar and a network of 10 rain gauges. Because of the limitations of the propagation model, this study is limited to the events with strong advective characteristics (four out of eight recorded events). For these events (only), the method produces rainfall fields that are highly correlated with the radar maps at spatial resolutions greater than $2 \times 2\,\mathrm{km}^2$. The point-scale results are also satisfactory for temporal resolutions greater than 10 min (mean correlation with rain gauge data equal to approximately 0.8, similar to the correlation between radar and rain gauge data).

This method can also be adapted to the fusion of a rain gauge with microwave link measurements and, through the use of several sensors, it has the potential of being applied to larger areas.

1. Introduction

The estimation of small-scale rainfall is an important issue in several domains, such as urban hydrology, flash-flood forecasting, water cycle studies, or climate change. Rain gauge networks or weather radars are commonly used for the observation of rain cells. The latter system allows large areas to be observed from one site with a typical temporal resolution of 5–15 min and a spatial resolution of $1\,\mathrm{km}^2$. Rain gauge systems have a temporal resolution in the range between 5 min and 1 day and require a large number of rain gauges in order to achieve good spatial sampling. Both techniques have various disadvantages, such as the cost of weather radars and the cost of maintaining a large number of rain gauges.

Microwaves are affected by rain, especially at frequencies above 10 GHz. Their specific attenuation K ($\mathrm{dB\,km^{-1}}$) can be related to the rainfall rate R ($\mathrm{mm\,h^{-1}}$) by the well-known power-law equation

$$K = aR^b, \tag{1}$$

where a and b are two parameters depending on the frequency, the polarization, and the drop size distribution (see, e.g., Leijnse et al. 2008).

Several studies have evaluated the accuracy with which localized or integrated rainfall rates can be derived from satellite-to-Earth attenuation data (see, for instance, Maitra and Chakravarty 2005) and have explored various applications of this technique, often using cellular communication networks (Schleiss and Berne 2010; Chwala et al. 2012; Overeem et al. 2013). In addition, a considerable number of studies have investigated the reconstruction of rainfall fields from microwave attenuation measurements. These studies make use of both simulated data (Giuli et al. 1991, 1999), with a predefined chosen geometry, and real data (Overeem et al. 2013; Zinevich et al. 2009) provided by

Corresponding author address: François Mercier, LATMOS, 11 Boulevard d'Alembert, 78280 Guyancourt, France.
E-mail: francois.mercier@latmos.ipsl.fr

DOI: 10.1175/JTECH-D-14-00125.1

© 2015 American Meteorological Society

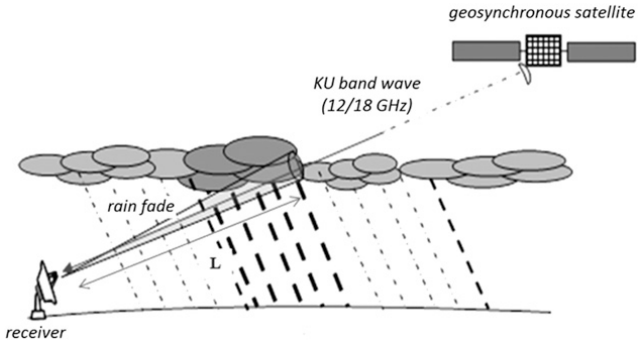

FIG. 1. Measurement setup used for the experiments described in this paper. A receiver on the ground measures Ku-band microwave signals transmitted by a geosynchronous satellite, and a spectrum analyzer determines the rain-induced attenuation of the recorded signals. See Barthès and Mallet (2013) for further details.

cellular communication networks. The reconstruction of rainfall fields from microwave attenuation data employs various techniques, such as tomography (Giuli et al. 1991; Zinevich et al. 2008) and ordinary kriging (Overeem et al. 2013). Zinevich et al. (2009) first uses a data assimilation technique (extended Kalman filter) based on a rainfall advection model.

The present study differs from previous research in this field with respect to two major aspects of the data retrieval and processing. First, we use attenuation measurements performed on a set of geostationary broadcast television (TV) satellites. These allow the integrated rain attenuation to be estimate along the link's first 5–7 km [depending on the freezing level and the direction (azimuth and elevation) in which the satellite is viewed from the receiver] with a high temporal resolution. This measurement system is presented in detail in Barthès and Mallet (2013) and schematically in Fig. 1. Second, the retrieval algorithm is based on the four-dimensional variational data assimilation (4DVAR) of observations in an advection model (see section 3).

In this paper, we use data recorded in the southeastern area of France, in the Cévennes–Vivarais region, during the extreme meteorological events that occurred in 2013 in the context of the Hydrological Cycle in the Mediterranean Experiment (HyMeX; see, for instance, Drobinski et al. 2013). The measuring device comprises a single receiver and four low-noise block downconverters mounted on a multifocus antenna, allowing quasi-instantaneous attenuation measurements along four azimuths corresponding to the positions of four TV satellites operating in the Ku band (12 GHz). By using just one receiver to measure the signals from four satellite links (satlinks), the spatial sampling of the rainfall is very sparse. Conversely, this device has a very good temporal resolution (10 s), and it is thus possible to rebuild rainfall fields from these measurements and then to connect the measurements recorded at successive discrete time steps through the use of an advection model to propagate the rainfall fields. A simple triangulation algorithm was developed to deduce the direction and speed of advection from the measurements. The rainfall field is then determined with a spatiotemporal resolution of $0.5 \times 0.5\,\text{km}^2$ and 10 s, based on a combination of measurements and values retrieved from the advection model with the 4DVAR algorithm. The performance of the measurement system and its associated retrieval algorithm is assessed and compared with radar and rain gauge observations.

The datasets are described in section 2. We then present the retrieval algorithm in section 3. Section 4 describes the algorithm developed to compute the advection velocity and the simulated data used to evaluate its performance. In section 5 we provide initial results and estimate our model's performance through the use of simulated data. Finally, section 6 presents the results obtained with several case studies, using real measurements recorded during rainfall events. Section 7 provides our conclusions.

2. Datasets

In this section, we first present the studied area, its location, and principal characteristics, as well as the size and features of the attenuation measuring device. We then present the radar and rain gauge data used to validate the method.

The Ku receiver is located at Mirabel, in the Ardèche department of the southwestern area of France. Mirabel is situated in a mountainous area called the Cévennes–Vivarais (see Fig. 2). This region experiences intense Mediterranean rain events with strong spatial and temporal variabilities, which are amplified by the mountainous topography. A description of the region and its characteristics is provided in Delrieu et al. (2009) and Molinié et al. (2012). The Ku sensor can record four different satellite-to-Earth microwave links simultaneously. The data made available for our study span several different time periods in 2013. The characteristics (azimuth and elevation) of the microwave links are provided in Table 1. As described above (see section 1; Fig. 1), the effective pathlengths depend on the altitude of the freezing level, which is calculated using the temperature recorded at Météo-France's Aubenas weather station, located 9 km to the west of the sensor, and assuming a temperature decrease of 1°C every 150 m. As none of the studied events was observed in winter, the freezing level could be considered to be consistently higher than 2.5 km. The length of the link is thus given by $H_0/\tan(\alpha)$, where H_0 is the freezing level and α is the elevation angle. The link's pathlength thus has a minimum value of 4 km and can exceed 7 km during warm-rain events.

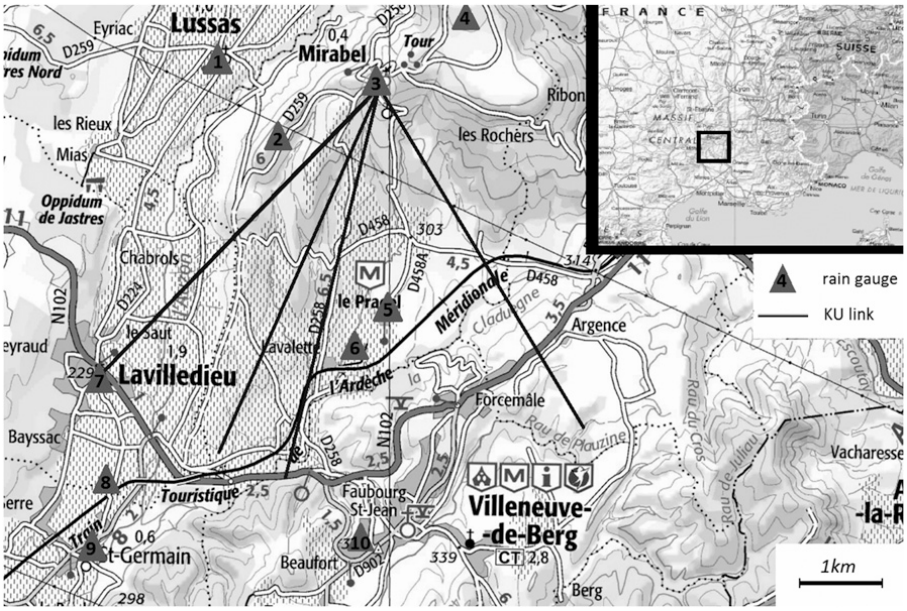

FIG. 2. Location of the measuring devices used during our experiments in 2013. The black frame indicates the position of the working area (Cévennes–Vivarais region) in the south of France. The enlarged section of this map (approximately 10 km in the east–west direction) indicates the location and geometry of the monitored Ku-band links (black lines) and of the HPiconet network rain gauges (gray triangles) used to validate the model. The pathlength of the microwave links is approximately 5 km.

The retrieval method described here was validated using two different rain measurement systems: weather radar and rain gauges. The Météo-France radar situated at Bollène, 40 km from the sensor, provided us with rainfall fields with a spatial resolution of $1 \times 1\ \text{km}^2$ and a temporal resolution of 5 min. The radar data were processed as described by Tabary (2007) and were then used to assess the performance of our technique, especially for space- and time-integrated quantities.

The rain gauge data are provided by the HPiconet network (http://www.ohmcv.fr/hpiconet/index.html), developed by the Observatoire Hydro-Météorologique Méditerranéen Cévennes–Vivarais (OHMCV) and HyMeX. This network comprises 10 rain gauges covering the same surface area as that analyzed by the Ku links, which was equal to approximately $60\ \text{km}^2$. The locations of these rain gauges are shown in Fig. 2. Their temporal resolution is 5 min. The data from these gauges were used to validate our field rebuilding method at certain points and to determine the Ku data integration times required to obtain sufficiently accurate rainfall values for point-scale locations.

3. Retrieval algorithm

In this section we describe the algorithm used to rebuild the rainfall fields. First, we present the numerical model used to propagate the rainfall fields. Then, we describe the 4DVAR data assimilation algorithm developed for this application.

a. *Space–time direct model*

To simplify the description of the spatial dynamics of rainfall, an advection model is used to propagate the rainfall fields. The aim is not to physically describe the atmospheric phenomena, since the conserved properties that are generally advected are heat or humidity (Allen et al. (1991)), but to describe the evolution of the rainfall cells during short periods of time (less than 45 min). This approach to the modeling of the transport of rainfall fields has already been applied and discussed in Zinevich et al. (2009). Diffusion is not taken into account for two main reasons: first, the network of satellite links is not

TABLE 1. Azimuth and elevation angles of the Ku-band microwave links observed during rain field experiments in the southwestern area of France in 2013. The link numbers correspond to those indicated in Fig. 2.

2013 events		
Link number (from west to east)	Azimuth	Elev angle
1	223.8°	28.5°
2	203.5°	35.9°
3	193.4°	37.7°
4	148.2°	33.5°

sufficiently dense to provide a satisfactory estimation of the diffusion parameters; second, we made our analyses over short time periods, during which the rainfall cells were assumed to remain relatively underformed (the consequences of this hypothesis are discussed in section 6).

Three important hypotheses are made.

(i) Contrary to Zinevich et al. (2009), we work directly with specific attenuations, rather than rainfall rates ($mm\,h^{-1}$) in order to simplify the observation operator. The specific attenuations are then converted into rainfall rates using Eq. (1) at the very end of the process.

(ii) The analysis is performed in two dimensions, meaning that measurements made at ground level (horizontal projection of the links path) are considered. This implies the assumption of vertical homogeneity of the rain field.

(iii) In all of our experiments, the wind vector field over the working area (29 km × 29 km) is assumed to be constant over time and homogeneous in space during the displacement of the rain cell across the observation area. Although the theoretical assimilation algorithm remains operational with a variable wind field, our measurement network is too sparse to evaluate such a field. For an area with a steep topography, as in the Cévennes, this hypothesis is reasonable for brief events only. For this reason, our analyses of rainfall events were restricted to short periods only (less than 45 min). For longer durations the events were split into several distinct parts (each part being treated as an independent event with, for instance, its own advection velocity).

The advection equation for a specific attenuation field $K(t, x, y)$ is given by

$$\frac{\partial K}{\partial t} = u\frac{\partial K}{\partial x} + v\frac{\partial K}{\partial y}, \tag{2}$$

where u and v denote the two horizontal components of the wind.

To discretize Eq. (2), we use the finite difference scheme developed by Smolarkiewicz (1983). The rain fields generally have a strong spatial variability, especially in mountainous areas such as the Cévennes–Vivarais region (see, e.g., Delrieu et al. 2009). To conserve the strong rainfall gradients when the fields are propagated through time by the numerical advection scheme, a scheme without strong numerical diffusion is needed. The Smolarkiewicz scheme, developed for the modeling of atmospheric phenomena, has this characteristic.

In the present study, although their values are not critical, the parameters of the Smolarkiewicz (1983) scheme Sc and ϵ [see Eqs. (15) and (23) in Smolarkiewicz 1983] are set to 1.04 and 10^{-15}, respectively.

The stability condition for this scheme (Courant–Friedrichs–Lewy condition) takes the form

$$\Delta t\sqrt{\frac{u^2}{\Delta x^2} + \frac{v^2}{\Delta y^2}} < \frac{1}{\sqrt{2}}, \tag{3}$$

where $(\Delta x, \Delta y)$ is the spatial resolution and Δt is the temporal resolution.

In our experiments on real data, we work with a 1-s time step and with a spatial resolution of. 0.5 km × 0.5 km. These parameters respect Eq. (3) for wind velocities up to $350\,m\,s^{-1}$.

b. 4DVAR data assimilation

In this section, we briefly recall the principles of the 4DVAR. We then develop the cost function used in this paper and include the filter term to ensure that the rebuilt rainfall fields are realistic. Finally, we present the YAO software used to implement the assimilation algorithm.

1) 4DVAR THEORY

Variational data assimilation consists of minimizing a cost function depending on an unknown initial field (and eventually on some model parameters). This function generally has two parts. The first part of these evaluates the gap between the available observations and the unknown initial field, propagated through time by a numerical model. The second part of the function evaluates the gap between the unknown initial field and a background field (first guess). Kalnay (2003) presents the principles of variational data assimilation. Navon (2009) presents a history of data assimilation and a review of the variational assimilation methods, including 4DVAR. Figure 3 presents the different steps of the assimilation algorithm developed in this study.

The specific attenuation field at time $t_0 = 0$ is noted $\mathbf{x}^0$. The unknown field at time $t_n = n\Delta t$ is $\mathbf{x}^n = \mathbf{x}(t_n)$, and the variable of the cost function (control parameter), which is the initial field in this study, is $\mathbf{x} = \mathbf{x}^0$. The cost function thus has the form

$$J(\mathbf{x}) = J_o(\mathbf{x}) + J_b(\mathbf{x}) \tag{4}$$

with

$$J_b(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}_b)^{\mathrm{T}}\mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) \tag{5}$$

and

$$J_o(\mathbf{x}) = \frac{1}{2}\sum_{n=0}^{N}(\mathbf{y}^n - \mathbf{H}_n\mathbf{x}^n)^{\mathrm{T}}\mathbf{R}_n^{-1}(\mathbf{y}^n - \mathbf{H}_n\mathbf{x}^n), \tag{6}$$

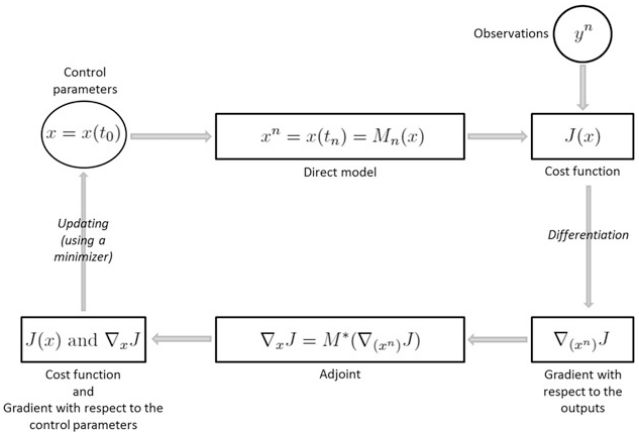

FIG. 3. Block diagram presenting the different steps of the 4DVAR algorithm described in section 3. The notations are the same as those used in this section.

where $N\Delta t$ is the duration of the simulation, x_b is the background, $\mathbf{B}$ is the background error covariance matrix, $\mathbf{y}^n$ represents the observations available at time t_n, $\mathbf{R}_n$ is the observations' error covariance matrix, and $\mathbf{H}_n$ is an operator used to project the field $\mathbf{x}^n$ on the observation space at time t_n.

If the forecast model is assumed to be deterministic (which means that an initial field $\mathbf{x}_0$ propagated by the model gives a unique field at time t_n), with $\mathbf{M}_{n-1,n}$ representing the forecast step propagating a field from time t_{n-1} to time t_n (the model being nonlinear), then we have

$$\mathbf{x}^n = \mathbf{M}_{n-1,n}(\mathbf{x}^{n-1}). \tag{7}$$

Then, noting $\mathbf{M}_n = \mathbf{M}_{n-1,n} \circ \mathbf{M}_{n-2,n-1} \circ \ldots \circ \mathbf{M}_{0,1}$, we have

$$\mathbf{x}^n = \mathbf{M}_n(\mathbf{x}). \tag{8}$$

The gradient of the cost function (4) is then given by

$$\nabla J(x) = \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) - \sum_{n=0}^{N} \mathbf{M}_n^{\mathrm{T}} \mathbf{H}_n^{\mathrm{T}} \mathbf{R}_n^{-1} [\mathbf{y}^n - \mathbf{H}_n \mathbf{M}_n(x)], \tag{9}$$

where $\mathbf{M}_n^{\mathrm{T}}$ is the adjoint of the linearized operator $\mathbf{M}_n$.

The difficulty in using the 4DVAR data assimilation method lies in the implementation of the adjoint model. In this study we use the YAO software developed by Laboratoire d'Océanographie et du Climat: Expérimentations et Approches Numériques (LOCEAN) and described in Nardi et al. (2009). This provides a simple method for deriving the adjoint. A brief description of this software is provided in section 3b(3).

2) COST FUNCTION

The background field $\mathbf{x}_b$ is usually the analyzed state of a previous assimilation cycle. However, in this study, as we work with very brief, highly variable events, the assumption is made that no background field is available. Moreover, a filter term J_f is added to the cost function in order to provide a certain degree of correlation between neighboring pixels, similarly to the approach of Giuli et al. (1991), that thus smooths the rainfall fields. This filter term is assumed to have the value 0 for a constant field. The cost function of the assimilation is then

$$J(\mathbf{x}) = J_o(\mathbf{x}) + J_f(\mathbf{x}). \tag{10}$$

The filter term J_f takes the form

$$J_f(\mathbf{x}) = c_f \|\mathbf{x} - \mathbf{m}\|^2, \tag{11}$$

where c_f is a scalar weighting coefficient and the vector $\mathbf{m}$ at grid point (i, j) is the average value of the field around point (i, j), namely,

$$m_{ij} = \frac{1}{9} \sum_{\substack{a=-1,0,1 \\ b=-1,0,1}} x_{i+a,j+b}, \tag{12}$$

where $x_{i+a,j+b}$ is the component of the field x at point $[(i + a)\Delta x, (j + b)\Delta y]$.

It should be noted that J_f is a filter term, in the sense that its value increases when the gradients of the attenuation field increase. Its aim is to prevent the assimilation algorithm from producing physically incoherent attenuation fields (meaning, with excessively large values of attenuation or gradients). Also that J_f can be written in the assimilation formalism as $J_f(\mathbf{x}) = \|\mathbf{x}\|_{\mathbf{B}^{-1}} = J_b(\mathbf{x})$, where $\|\mathbf{u}\|_{\mathbf{A}} = \mathbf{u}^{\mathrm{T}}\mathbf{A}\mathbf{u}$. This implies a background field $\mathbf{x}_b$ of 0 and that a background error covariance matrix ensuring $J_f(\mathbf{x})$ of 0 is obtained whenever $\mathbf{x}$ is constant. The background error covariance matrix $\mathbf{B}$, such as $c_f\|\mathbf{x} - \mathbf{m}\|^2 = \|\mathbf{x}\|_{\mathbf{B}^{-1}}$, is given by $\mathbf{B}^{-1} = \sum \mathbf{B}_k$, where k represents one of the points of the discretized domain [one of the (i, j) points]. Then, as described in Giuli et al. (1991), the value of the matrix $\mathbf{B}_k$ at grid point (i, j) (written b_{ij}^k) is given by

$$b_{ij}^k = c_f \begin{cases} \dfrac{64}{81} & \text{if} \quad k = (i,j) \\ -\dfrac{8}{81} & \text{if} \quad k = (i, j \pm 1) \quad \text{or} \quad k = (i \pm 1, j). \\ \dfrac{1}{81} & \text{if} \quad k = (i \pm 1, j \pm 1) \end{cases} \tag{13}$$

3) THE YAO SOFTWARE

The YAO software was developed by LOCEAN to assist with the implementation of the adjoint model for

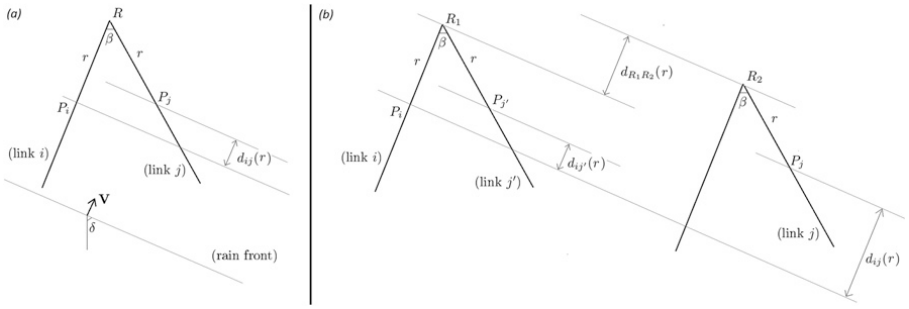

FIG. 4. (a) Notations used in section 4b, defining the distance $d_{ij}(r)$ between two points P_i and P_j on satellite links i and j, at a distance r from receiver R. The vector $\mathbf{V}$ defines the advection velocity (velocity of the rain cell), defined by its norm V and argument δ. The angle β is between links i and j. (b) For two links i and j incident on two different receivers R_1 and R_2: $\overline{d_{ij}(r)} = \overline{d_{ij'}(r)} + \overline{d_{R_1R_2}}$, where $\overline{d_{ab}}$ is the algebraic distance from link a to link b, and j' is parallel to link j, incident on receiver R_1. The distance D_{ij} [see Eq. (18)] is thus given by $D_{ij}(\delta) = L\sin(\delta/2)\cos(\beta) + \overline{d_{R_1R_2}}$, such that the links can also be assimilated to their center points.

4DVAR algorithms. An accurate description of YAO can be found in Thiria et al. (2006) and Nardi et al. (2009), and on the LOCEAN website (http://www.locean-ipsl.upmc.fr/~yao/). The aforementioned papers provide a complete description of this software and the implicit mathematical principles.

The cost function is minimized with M1QN3, an algorithm developed by the Institut National de Recherche en Informatique et en Automatique (INRIA), which uses a quasi-Newton technique to solve unconstrained optimization problems. This algorithm is presented in Gilbert and Lemaréchal (2006) and is interfaced with the YAO software.

4. Determining the advection velocity

In this section, we present the technique used to produce simulated data. We then describe the triangulation algorithm used to determine the advection velocity. Finally, we evaluate the performance of this algorithm using the simulated data.

a. Generation of simulated data

The aim of this method is to use 2D rainfall maps to produce Ku attenuation time series along the links. Simulated rainfall fields are generated by a 2D multifractal model [developed by Lovejoy and Schertzer (1990), using the rain parameters determined by Verrier et al. (2010)] that is able to reproduce the natural spatial variability of rainfall. In each simulation, the maximum rainfall rate is set to 100 mm h^{-1} and the rain percentage is set to 60%. Then, for a given (homogeneous and constant, as described in section 3a) advection velocity, the rainfall field is propagated through time at this velocity using the numerical method described in section 3a. During this advection phase, the attenuations produced by the field along the paths of the microwave links are determined using Eq. (1).

b. Advection velocity estimation

The aim is to estimate the advection velocity of the attenuation fields, using the recorded satellite link measurements. Zinevich et al. (2009) estimated such a velocity. But because their measuring device consists of a large number of microwave links, they were able to use a method developed in Desa and Niemczynowicz (1997). In the present study we present a new triangulation method, adapted to the case of a network of sparse microwave links with a specific anisotropic geometry. This approach involves the numerical minimization of a cost function J_1:

$$J_1(\delta, \mathbf{V}) = \frac{\|\mathbf{D}(\delta) - \mathbf{V}\boldsymbol{\tau}\|^2}{\|\mathbf{D}(\delta)\|^2}, \tag{14}$$

where δ is the direction (azimuth) and $\mathbf{V}$ is the velocity of the advection (see Fig. 4). Vector $\mathbf{D}(\delta)$ comprises the distances between each pair of satellite links (see below) and $\boldsymbol{\tau}$ is the vector comprising the experimental time lags between the arrival of rainfall events at each pair of links. The latter terms are calculated using the method presented in the next section.

1) TIME LAGS

Let $\mathbf{A}_1, \ldots, \mathbf{A}_p$ be the integrated rain attenuation vectors measured along p satellite-to-Earth microwave links during a rain event and let N be the number of measurements available for this event (and thus the size of vectors $\mathbf{A}_i$). The cross-correlation function between links i and j is then defined as

$$\forall m \in [|-N; N|]:$$

$$C_{ij}(m) = \begin{cases} \sum_{k=1}^{N-m} \mathbf{A}_i(k+m)\mathbf{A}_j(k) & \text{if} \quad m > 0 \\ C_{ji}(-m) & \text{if} \quad m < 0. \end{cases} \tag{15}$$

It is assumed that for each value of the pair (i, j), the value τ_{ij} that maximizes the cross-correlation function corresponds to the time lag between time series i and j, whereas the corresponding value of the cross-correlation function $\widetilde{C_{ij}} = C_{ij}(\tau_{ij})$ indicates the strength of the correlation between the two series.

2) DISTANCES

The distances $\mathbf{D}(\delta)$ between the satellite links defined by Eq. (14) depend on the advection direction. In this section, we define $\mathbf{D}(\delta)$ and show that the distance between the Ku links is the distance between their centers.

It is assumed that the rain front is perpendicular to the wind direction (characterized by the angle δ; see Fig. 4) and is moving in this direction. If $d_{ij}(r)$ is the distance between the two parallels to the rain front passing through the points P_i and P_j, located on the links i and j at a distance r from the receiver, as shown in Fig. 4a, then the mean distance between links i and j is given by

$$D_{ij} = \frac{1}{L} \int_0^L d_{ij}(r)\, dr, \tag{16}$$

where L is the pathlength. Note that here, the lengths of the links are assumed to be equal. But we could show that the results of this section are still verified in the case of links of unequal length.

The following expression is derived for the function d_{ij}:

$$d_{ij}(r) = 2r \sin\left(\frac{\beta}{2}\right) \cos(\delta), \tag{17}$$

where β is the angle between links i and j.

This leads to the following expression for the distance D_{ij}:

$$D_{ij}(\beta) = L \sin\left(\frac{\beta}{2}\right) \cos(\delta). \tag{18}$$

Finally, from Eq. (18) it follows that the links can be assimilated to their center points [because $D_{ij}(\beta) = d_{ij}(L/2)$].

In the above-mentioned calculations, it is assumed that just one receiver monitors all of the p satellite links. However, the final result of this section (that the links can be assimilated to their center points) remains valid when determining the distance between two links monitored by two different receivers. Further details of this result are provided in Fig. 4b.

c. *Results*

When only one receiver is used, the preceding result leads to nonunicity of the solution, since the center points of the links are almost aligned. In such a case, the pair (δ_m, V_m) minimizing the cost function J_1 defined by Eq. (14) is not unique. For each value $\tilde{\delta}$ of wind direction, there is a corresponding wind speed $V(\tilde{\delta})$, which satisfies $J_1[\tilde{\delta}, V(\tilde{\delta})] = \min_{\delta,V}[J_1(\delta, V)]$. Under these conditions, only the wind speed can be deduced from the wind direction, or vice versa. Figure 5a illustrates this case with a simulated example (further details concerning the link's characteristics are provided in the caption). This difficulty can be avoided by adding a second receiver: under these conditions, there is only one pair of wind parameters (δ_m, V_m) that can minimize the cost function of Eq. (14). This result is illustrated in Fig. 5b.

Note that when using simulated data, the results can be improved by positioning the second receiver as far as possible from the first receiver (provided both receivers are simultaneously affected by the same rain cells). However, when running the assimilation algorithm, it is preferable for the second receiver to be positioned relatively close to the first. A compromise separation of approximately 2.5 km was found to provide satisfactory results.

d. *Method used for observations with a single sensor*

The experiments on real data presented in section 6 of this paper were recorded with just one receiver that was able to simultaneously monitor the signals emitted by four satellites. This configuration led to the measurement of four distinct Earth–satellite links. Under these conditions, as described above, the advection parameters cannot be directly estimated from the attenuation measurements. A different approach was thus used in order to determine these parameters from the radar data (see section 2): this involved calculating the distance and direction leading to the strongest correlation between successive radar maps. This method allows the distance between the two maps to be retrieved, from which the associated velocity vector can be determined.

5. Results based on simulated data

In this section, we evaluate our reconstruction method with simulated data. Section 5a describes the features of the simulated Ku links. In section 5b, we define the validation areas. Section 5c provides a statistical evaluation of the model's performance in the absence of

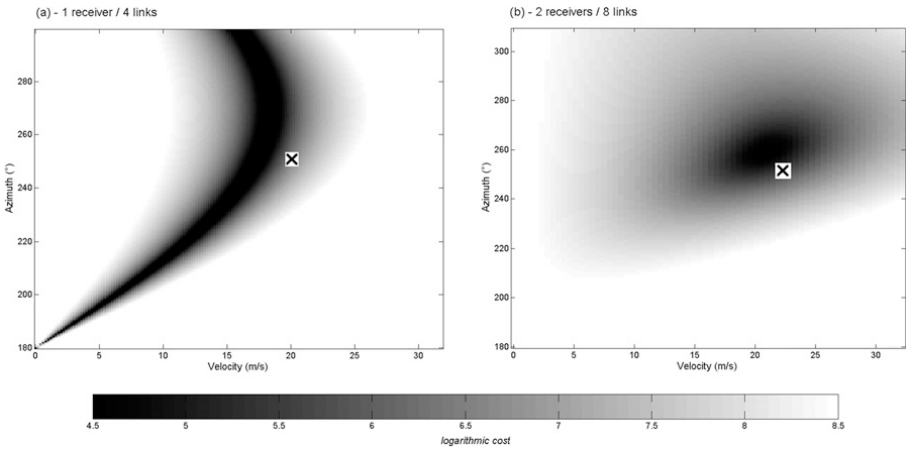

FIG. 5. Values of the cost function defined by Eq. (14) for simulated attenuation data with a wind velocity $V = 20\,\mathrm{m\,s^{-1}}$ and azimuth $\delta = 250°$ (real values indicated by a black cross). The simulated measuring device comprises (a) one receiver receiving four satellite-to-Earth microwave links and angularly separated from one another by 20° and (b) two receivers of the same type, separated by 2.15 km. Whereas the first system does not allow both advection speed and azimuth to be rebuilt (1D cost function minimum), the second system allows a relatively accurate solution to be found: the rebuilt advection is $21\,\mathrm{m\,s^{-1}}$ and 259°. Section 4a provides a description of the method used to generate simulated data.

perturbations, together with some results. We comment on the influence of the geometry of the simulated context on these results. Finally, in section 5d, we provide a qualitative evaluation of our algorithm's response when errors are added to the link pathlengths and advection velocity estimations.

a. Link characteristics

The method used to produce simulated attenuation time series is described in section 4a. Here, we simply add white noise to these attenuation time series. The noise has a uniform distribution centered around 0 and a width of 0.5 dB, corresponding to the uncertainty typically associated with satellite microwave attenuation measurements (Barthès and Mallet 2013). The resulting attenuation time series were used as input data for the assimilation algorithm.

We simulate the use of either one or two Ku sensors, with each of these receiving microwave signals from four different satellites. For each sensor, the azimuths of the four satellite links are 150°, 170°, 190°, and 210° (with 180° corresponding to the north/south axis). All eight links have the same length, equal to 7 km. The second sensor is located 2 km east and 0.8 km south of the first sensor, thus at a distance of 2.2 km compatible with the estimated optimal distance (2.5 km; see section 4c).

b. Evaluation area (assimilation area)

As described above, the algorithm produces rain fields over a rectangular $N \times N$ grid [covering an area of $(N\Delta x)^2\,\mathrm{km}^2$]. However, the algorithm is not able to rebuild realistic fields over the full grid. Although some portions of the rain fields are positioned on the grid, they never pass above the Ku links and are thus never recorded. For the purposes of evaluating the algorithm's performance, we thus reduce the grid to a "large assimilation area," defined as the area "seen" by the Ku links, as shown in Fig. 6. For a total simulation time of T seconds and a rainfall map considered t seconds after the beginning of the simulation, our large assimilation area is defined as that portion of the rain field that will pass through the links before the end of the simulation [rectangle of length $\mathbf{V}(T - t)$, where $\mathbf{V}$ is the advection velocity; see Fig. 6].

We also introduce a "small assimilation area," defined as the area directly below the Ku links (see Fig. 6). This second area is introduced for two main reasons. First, the large assimilation area is not fixed, since it depends on the advection speed and direction, and on the simulation time. The small assimilation area has a fixed position. Second, in the present section (simulated data), the model is assumed to be perfect. However, in real situations the model can be affected by large errors (for instance, numerical diffusion, or nonadvective events; see section 6). The small assimilation area thus allows the algorithm to be evaluated over a more constrained area, situated just below the links.

In the present section, all of the figures and statistics are restricted to the large assimilation area.

c. Statistical results

Fifteen initial rainfall fields were produced by the multifractal model, with a spatial resolution of 0.4 km × 0.4 km

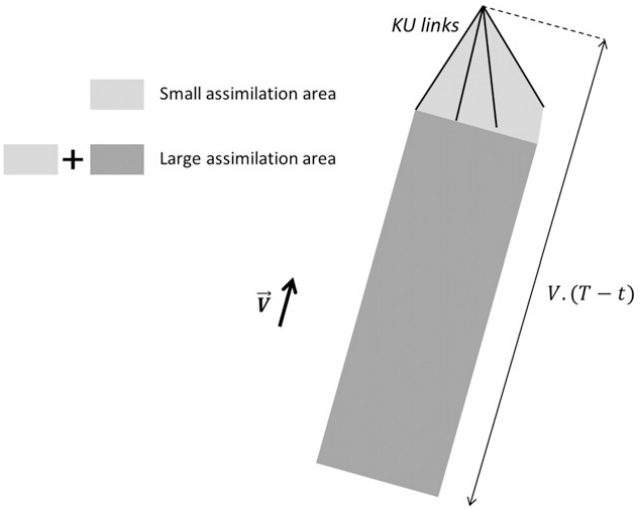

FIG. 6. Small and large assimilation areas as defined in section 5b. The advection velocity is represented by **V**, T is the total simulation time, and t is the time of the rain field under consideration.

on a 68×68 grid, so that the total area covered was close to 27 km $\times$ 27 km. For each of these 15 fields, we then created the corresponding attenuation time series using the method described in section 4a, with an advection velocity chosen to ensure that the rainiest portion of the field passes above the microwave links during the advection phase (phase during which the attenuation time series are created). The velocities, which were chosen in order to limit the total duration of the simulations to approximately 30 min, thus ranged from 9 to 21 m s^{-1}. The advection azimuths ranged from 180° (south wind) to 270° (west wind). These characteristics are coherent with those encountered during all of the strong rain events monitored during our case studies.

Finally, the attenuation series are assimilated. The value of the cost function weighting coefficient c_f [Eq. (11)] was set to 7×10^{-3}. This value was chosen in order to ensure that, after a few iterations of the minimizer, the observation and filter components of the cost function would be nearly balanced. Otherwise, the minimizer would focus on minimizing the most strongly weighted component of the cost function, and the other component of this function would contribute nothing to the results. Note that the value assigned to this parameter depends on the weighting of the observation component of the cost function, and thus on the total quantity of rain falling above the microwave links during the experiment (in the case of a greater quantity of rain, the attenuation is stronger, a higher number of observations is available, and the observation component of the cost function has a higher weight). In these simulated experiments, the total rainfall remains almost constant (all of the initial fields are generated with a multifractal model using the same parameters), such that the value of c_f does not change. With real data, the total rainfall above the links can vary strongly from one event to another. In such a case, an initial weighting coefficient is selected and the cost function is evaluated after a small number of iterations of the minimizer. The weighting coefficient is then adjusted so that the two components of the cost function are balanced. This method could probably be improved, for example, by automatically adjusting c_f during the minimization process.

We define the following indicators to quantify the method's performance. The first of these is the absolute bias b_a^n defined by

$$b_a^n = |m_n^{\text{as}} - m_n^0|, \tag{19}$$

where m_n^{as} and m_n^0 indicate the mean rainfall rates of the assimilated and original initial fields, respectively, for the nth experiment (with n ranging from 1 to 15).

Using the same notations, the relative bias b_r^n can be defined as a percentage by

$$b_r^n = 100\left(\frac{m_n^{\text{as}} - m_n^0}{m_n^0}\right). \tag{20}$$

To evaluate the model's ability to reproduce the strongest parts of the rain cells, we also calculate the mean bias at the 95% percentile, defined as

$$b_{95\%} = \frac{1}{15}\sum_{n=1}^{15}(m_{95,n}^{\text{as}} - m_{95,n}^0), \tag{21}$$

where $m_{95,n}^{\text{as}}$ and $m_{95,n}^0$ indicate the value of the 95% percentile of the field for the nth experiment for the assimilated and original fields, respectively.

Finally, the root-mean-square error rmse^n is defined as

$$\text{rmse}^n = \sqrt{\frac{1}{N}\sum_{(S)}[(R_n^{\text{as}})_{ij} - (R_n^0)_{ij}]^2}, \tag{22}$$

where $(R_n^{\text{as}})_{ij}$ and $(R_n^0)_{ij}$ are the assimilated and original initial rainfall fields of the nth experiment at grid point (i, j), respectively, where (S) is the large assimilation area and N is the number of grid points in the (S) area.

Table 2 provides a summary of all these indicators, in terms of bias and standard deviation over the 15 simulations, when one or two Ku sensors are available.

Table 2 shows that the assimilation algorithm accurately reproduces the total rainfall rate over the area, with a mean absolute bias of only 2.0% associated with the use of eight microwave links. Moreover, the algorithm does not produce any significant systematic

TABLE 2. Statistical results on 15 simulated assimilation experiments: mean and standard deviation of the absolute bias b_a^n [Eq. (19)], mean of the absolute relative bias $|b_r^n|$ [see Eq. (20)], mean relative bias b_r^n [Eq. (20)], mean bias to the 95% quantile $b_{95\%}$ [Eq. (21)], and mean and standard deviation of the root-mean-square error rmse^n [Eq. (22)].

		2 Ku sensors (2 × 4 microwave links)	1 Ku sensor (4 microwave links)
Absolute bias (mm h^{-1})	Mean	0.8	2.1
	Std dev	0.5	1.8
Absolute relative bias	Mean	2.0%	5.2%
Relative bias	Mean	+0.3%	+3.6%
$b_{95\%}$ (mm h^{-1})		+4.0	+11.1
RMSE (mm h^{-1})	Mean	12.4	21.6
	Std dev	6.3	6.6
Mean rainfall rate (mm h^{-1})		40	

deviation since the relative bias over the 15 simulations is only +0.3%. The absolute bias is very small (0.8 mm h^{-1}) compared to the mean rainfall rate in the assimilation area, which is approximately 40 mm h^{-1} (ranging from 27.6 to 48.4 mm h^{-1}). The standard deviation for the combined 15 simulations is equal to 0.5 mm h^{-1}, which thus indicates that the algorithm accurately reproduces the total rain. When only one receiver is used (corresponding to just four satellite-to-Earth microwave links), there is an associated decrease in the quality of the results (mean absolute error = 2.1 mm h^{-1} with a standard deviation of 1.8 mm h^{-1} and a mean overestimation of the rainfall of 3.6%). These results remain satisfactory and do not show any reason to not support the use of this algorithm in the current experiment with only one receiver (see section 6).

Finally, a pixel-to-pixel comparison reveals a mean RMSE of 12.4 mm h^{-1} with a standard deviation of 6.3 mm h^{-1}. This value is quite satisfactory (the average rain rate is close to 40 mm h^{-1}), but it must be used with care: when working with real data, many sources of error are introduced such that these point-to-point comparisons will no longer be realistic. Nevertheless, these results show that the use of only one receiver does not strongly disturb the model, since it finds a mean RMSE of 21.6 mm h^{-1} and a similar standard deviation.

One notable geometric characteristic of the microwave links used in this study is their anisotropy. As these links are directed mainly southward from the sensor (since they correspond to geostationary satellites observed from the Northern Hemisphere), it is reasonable to expect the results to be better with westerly winds than with southerly winds. In the first case, the links will be successively traversed by the rain front, whereas in the second case they will be traversed simultaneously. Moreover, in the first case, each portion of the rain cell will necessarily advance successively through each of the links. On the other hand, if the rain field is advected northward, then some portions of the rain cells may advance between two links without touching either of these for a significant length of time, thus making the minimization and determination of field features more difficult.

Our simulations confirm this assumption. We recalculate the statistics of Table 2, while distinguishing the events arriving mainly from south (meaning, with an advection azimuth between 180° and 225°) from the events arriving mainly from west (advection azimuth between 225° and 270°). We note that from a total of 15 simulations, 8 had a northward velocity and 7 had an eastward velocity. In the case of the northward events, the mean absolute bias is 1.0 mm h^{-1} (2.5%) and the root-mean-square error is 15.7 mm h^{-1}, whereas these values are 0.6 mm h^{-1} (1.4%) and 8.7 mm h^{-1}, respectively, for the eastward events. This confirms the expected outcome of mainly west/east advected events being easier to rebuild with our algorithm than the north/south advected events.

We also note (Table 2) that the model overestimates the 95% quantile by 4 mm h^{-1} (note that the mean value of the 95% quantile over the 15 simulations is approximately 73 mm h^{-1}). Although this overestimation is quite small, it is significant, especially as all of the simulations overestimate the value of the 95% quantile. This outcome is probably due to the propagation model used and its associated numerical diffusion. Very strong, localized gradients on the initial field are very quickly smoothed by the numerical model during advection, meaning that they do not significantly increase the cost function, even with a strong filter term [see section 3b(2)]. However, by placing very high rainfall rates on the borders of the assimilation area, the algorithm gains new degrees of freedom. These high rainfall rates, which occur during the advection phase, are also numerically smoothed, thereby helping the algorithm to adjust the cost function, even at the center of the assimilation area. This phenomenon occurs mainly with rainfall events accompanied by a southerly wind, during which it is

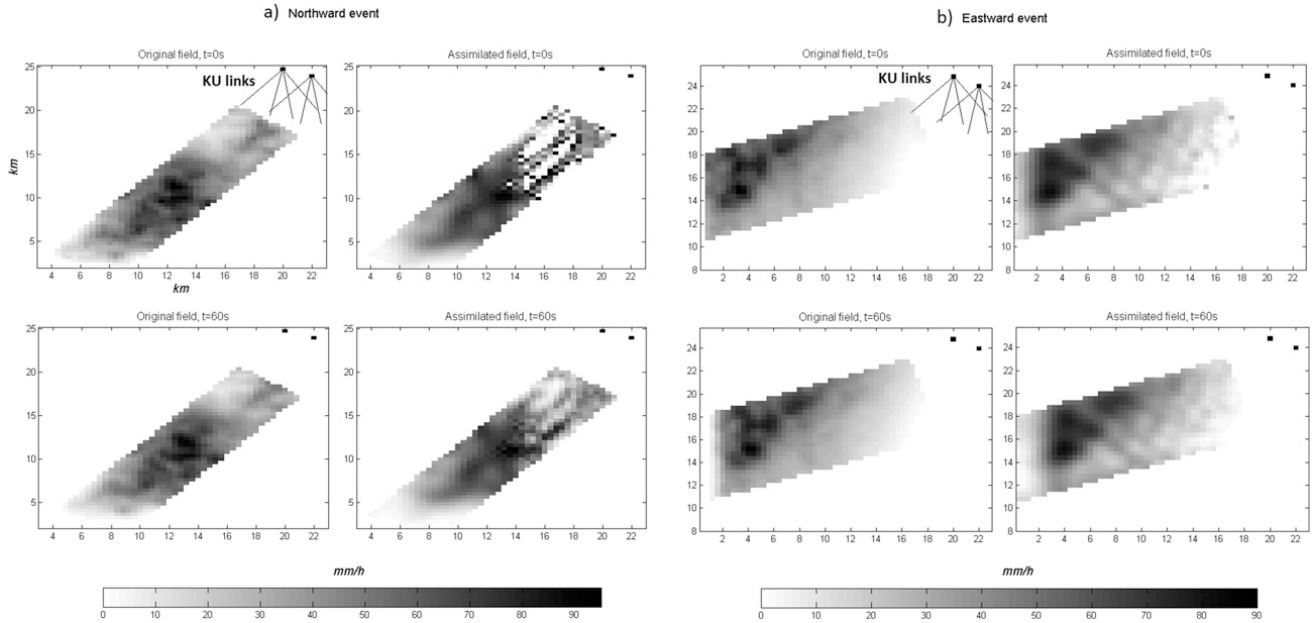

FIG. 7. Results of two assimilation experiments run with simulated data (recorded by two Ku receivers), using different original rain fields and advection parameters. (a) Results for a mainly northward event. The advection azimuth is 217° and the advection speed is 17.1 m s^{-1}. (left column) The original field (to be rebuilt) at $t = 0$ (initial field, top row) and after 60 s (advected by the numerical scheme, bottom row). (right) The rain field rebuilt by the assimilation algorithm at $t = 0$ and after 60 s. (b) As in (a), but for a mainly eastward event. The advection azimuth is 253° and the advection speed is 17.9 m s^{-1}.

more difficult for the microwave link geometry to accurately locate the rain cells. Figure 7 provides a good illustration of this phenomenon. Figure 7a shows the results of an assimilation experiment for a mainly southerly event. The upper part of the figure compares the original field with the rebuilt field at $t = 0$. Although the total rainfall over the area is reproduced with reasonable accuracy, the assimilated field has a significant number of very steep gradients, with some pixels having rainfall rates more than 90 mm h^{-1} in excess of that of dry neighboring pixels. The lower part of the figure represents the same fields, 1 min later. The steep gradients of the assimilated field have been strongly smoothed, so that the error produced by these gradients does not have a strong influence on the cost function. This phenomenon is not observed in the case of events accompanied by a westerly (Fig. 7b), because each rain cell passes successively through the path of each microwave link. This phenomenon could be reduced by carefully adjusting the weighting of the filter term for each simulation, or by adding various constraints at the level of the cost function to the rain field's characteristics (and replacing the filter by these constraints). We also note that decreasing the spatial resolution of the rebuilt fields (e.g., by averaging these fields on a 1 km × 1 km grid) smooths the fields and reduces the extent of this problem. As already described, this phenomenon is amplified by the use of a single receiver, which leads to a small overestimation of the 95% quantile (mean excess of 11.1 mm h^{-1}), which nevertheless remains reasonable.

These statistical results show that the algorithm is efficient for the processing of simulated data with no error model. Overall, the rain field features are satisfactorily reproduced (no systematic under- or overestimation of the total rainfall, a small mean absolute bias, and point-to-point errors that are small in comparison with the rain rates). It is also shown that, although the results are deteriorated by the use of only one receiver, they remain acceptable. Nevertheless, these results show that our model could be applied to real data. But under these conditions, various sources of error can be expected to affect the estimations.

d. *Impact of uncertainties of link length and advection parameters*

In this section, the algorithmic parameters, such as the advection velocity vector or the length of the Ku links, are called input parameters. In a case study with real data, these parameters are evaluated with a certain degree of uncertainty. The aim of this section is to qualitatively assess the algorithm's response when these parameters are not accurately known. Only qualitative results are given here: during real measurements, larger errors can be expected as a consequence of inaccuracies in the model itself (real events are not purely advective;

TABLE 3. Mean (ρ) and standard deviation (σ) of the normalized maximum value of cross correlations between successive radar maps, for eight rain events observed in 2013. Advection parameters determined from the analysis of radar maps. Mean and maximum attenuations recorded by the Ku sensor during the rain events. Estimated freezing level, derived from the temperature at the Aubenas Météo-France weather station.

No.	Date	Beginning time (UTC)	End time (UTC)	ρ	σ	Advection		Attenuation (dB)		Freezing level
						Velocity ($m\,s^{-1}$)	Direction (°)	Mean	Max	
1	7 Sep	1400	1435	0.90	0.025	11.5	210	1.64	5.29	3.5 km
2	7 Sep	1735	1810	0.81	0.026	12.6	199	1.19	5.00	3.0 km
3	8 Sep	1440	1525	0.81	0.082	6.2	238	2.51	12.78	3.9 km
4	28 Sep	1805	1840	0.90	0.015	13.6	191	1.20	9.89	3.5 km
5	29 Sep	0205	0250	0.89	0.019	7.9	208	4.37	13.97	2.8 km
6	4 Oct	1745	1825	0.81	0.030	12.1	194	1.95	7.49	3.0 km
7	20 Oct	0635	0705	0.92	0.018	16.2	220	2.56	11.28	2.5 km
8	20 Oct	0820	0850	0.89	0.015	16.2	220	5.32	12.48	2.5 km

see below). Furthermore, the statistical results corresponding to simulated idealized cases would not be representative of the results obtained during a real rainfall scenario.

We first studied the influence of an error on the length of the links. For this, the observations (i.e., attenuations) were initially simulated with the lengths of the Ku links set to $L = 7$ km. During the assimilation process, the parameter L was then decreased to 6.5 km, leading to an overestimation of the rainfall in the southern portions of the links: the algorithm assigns the same volume of rain to a smaller area (since the assimilation area is reduced when the length of its associated link is decreased). The algorithm allocates the rain "excess," recorded by those portions of the links that have been removed, to the southern portion of the new shortened links. However, the total rain mass is correctly reproduced. We also studied the influence of an error in advection speed. Simulations show that a decrease in advection speed acts like a contraction operator on the resulting field, in addition to shifting it toward the location of the microwave links, whereas an increase in advection speed produces the opposite effects. This outcome can be explained by the fact that when the advection speed is decreased, the resulting rainfall field has to be shifted closer to the microwave links (in order to transit above them at the correct time) and must be compressed in order to take the same time to pass over the links. Once again, the total rain mass is well reproduced.

6. Results with experimental data

In this section, we present the rainfall events monitored in 2013 with the measuring device described in section 2. This is followed by a discussion of the results.

a. Description of the events

Eight rain events were observed with the Ku links during the autumn of 2013. Table 3 presents the main features of these events. They were generally quite intense (the maximum attenuation is always greater than 5 dB) and moved mainly northward (mean advection direction of 210°). The mean attenuation rates along the microwave links ranged between 1.20 and 5.3 dB. Events longer than 45 min were split into separate independent components.

As described in section 3b, our 4DVAR algorithm is designed to rebuild the initial (t_0) attenuation field, so that the differences between the latter field (when it is propagated through time by the numerical model) and the observations are minimized. In section 3a the numerical model used in this study is shown to be purely advective. This assumption is clearly a broad simplification that is not always verified, especially in a mountainous area such as the Cévennes–Vivarais region (Molinié et al. 2012). The upper part of Fig. 8 shows radar maps corresponding to an event that occurred on 8 September 2013 in the vicinity of the Ku links. Propagation of the rainfall cells during this event is clearly not advective, to the extent that the total rain over the area increases significantly during the event, due to orographic precipitation caused by vertical ascent of the air mass. Conversely, the lower part of Fig. 8 shows an event that occurred on 20 October 2013, during which the total rain rate appears to be quite well conserved and the propagation is relatively advective.

To evaluate the advective characteristic of the events, we introduce a statistical criterion that is the value of the maximum cross correlation between successive weather radar recordings. Table 3 presents, for our eight events, the mean and standard deviation of the normalized maximum cross correlations between all of the

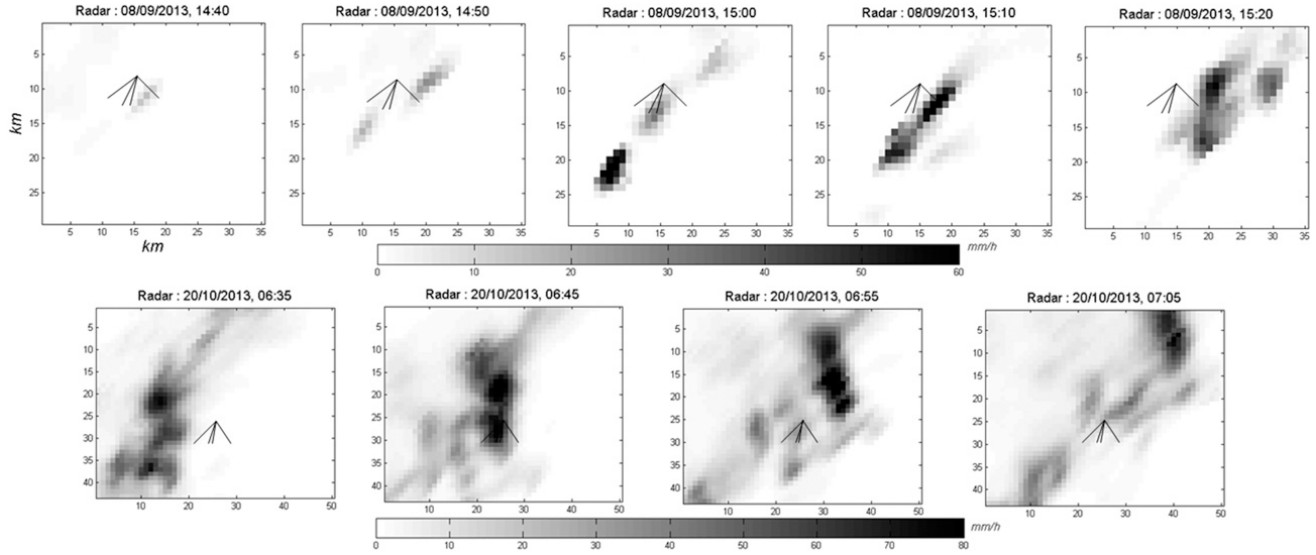

FIG. 8. Weather radar images around the Ku sensor, for two events recorded in 2013: (top) on 8 Sep and (bottom) on 20 Oct. The time lag between successive snapshots is 10 min. The four Ku links are represented by the continuous black lines.

successive 5-min radar maps. As it is reasonable to assume that the events with a large mean value and a small standard deviation are more "advective," better results can be expected from our rebuilding algorithm when these events are analyzed. We thus selected four advective events (events 4, 5, 7, and 8).

b. *Results*

In this section, Table 4 presents the results of numerical comparisons between Ku and radar rainfall fields for the eight aforementioned events and for the small and large assimilation areas (see section 5b) at

TABLE 4. Correlation coefficients r between radar and Ku-band (microwave) rain fields; regression line equations for two different spatial resolutions; bias (corresponding to the difference between the mean values of microwave and radar rainfall measurements); average radar rain rates; and relative bias (sixth column divided by seventh column) for the eight events studied in 2013. These statistics have been computed for both the small and large assimilation areas defined in section 5b.

Small assimilation area							
	r		Regression line		Bias	Avg rain rate	Relative bias
Event (see Table 3)	1 km × 1 km	2 km × 2 km	1 km × 1 km	2 km × 2 km	(mm h^{-1})	(mm h^{-1})	(%)
4 (28 Sep)	0.85	0.92	$0.98x + 1.3$	$1.08x + 0.5$	+1.1	5.2	+21
5 (2 Sep)	0.77	0.84	$0.95x + 0.5$	$0.94x + 1.1$	−0.7	24.9	−2.8
7 (20 Oct)	0.81	0.88	$0.44x + 6.1$	$0.55x + 2.5$	−7.9	25.1	−31
8 (20 Oct)	0.39	0.82	$0.40x + 26$	$0.83x + 3.5$	−2.7	47.1	−5.7
1 (7 Sep)	0.69	0.67	$0.19x + 4.4$	$0.24x + 2.6$	−12.8	21.3	−60
2 (7 Sep)	0.17	0.32	$0.13x + 0.3$	$0.19x + 1.5$	−5.5	9.0	−61
3 (8 Sep)	0.62	0.66	$0.47x + 9.9$	$0.54x + 6.9$	+2.7	13.8	+20
6 (4 Oct)	0.42	0.63	$0.36x + 8.9$	$0.54x + 4.8$	−0.3	14.4	−2.1
Large assimilation area							
	r		Regression line		Bias	Avg rain rate	Relative bias
Event (see Table 3)	1 km × 1 km	2 km × 2 km	1 km × 1 km	2 km × 2 km	(mm h^{-1})	(mm h^{-1})	(%)
4 (28 Sep)	0.48	0.52	$0.68x + 1.3$	$0.74x + 0.97$	−0.1	3.6	−2.7
5 (29 Sep)	0.51	0.53	$0.53x + 3.8$	$0.51x + 4.7$	−6.9	34.7	−20
7 (20 Oct)	0.81	0.87	$0.58x + 3.5$	$0.68x + 2.0$	−2.1	12.2	−17
8 (20 Oct)	0.64	0.78	$0.72x + 4.7$	$0.85x + 1.3$	−6.5	35.4	−18
1 (7 Sep)	0.72	0.70	$0.21x + 2.0$	$0.23x + 1.3$	−10.8	16.0	−68
2 (7 Sep)	0.41	0.49	$0.26x + 0.34$	$0.30x + 0.14$	−6.6	9.0	−73
3 (8 Sep)	0.52	0.67	$0.45x + 14$	$0.61x + 8.1$	+3.5	19.5	+18
6 (4 Oct)	0.43	0.61	$0.33x + 5.6$	$0.45x + 3.5$	−7.9	19.9	−40

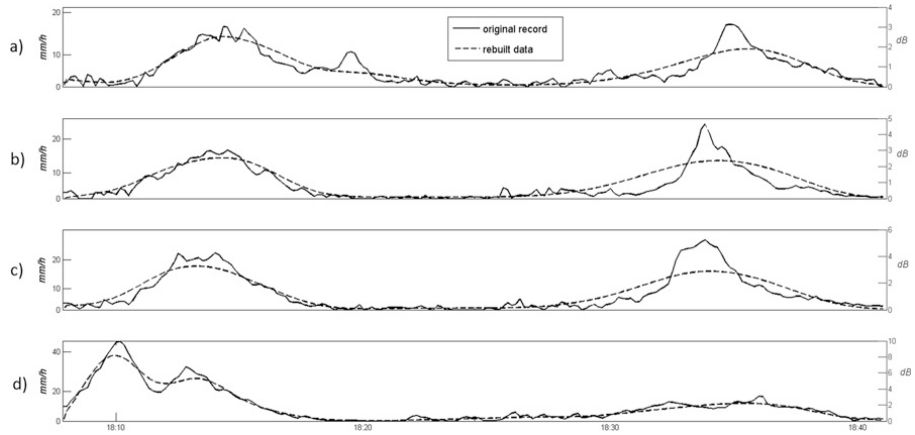

FIG. 9. Integrated rainfall along the four microwave links as recorded by the Ku sensor (solid lines) and reproduced by the assimilation algorithm (dashed lines), during the event that occurred from 1810 to 1840 UTC on 28 Sep 2013. Plots (a)–(d) represent data from the most westward link to data from the most eastward link (see Fig. 2).

different spatial resolutions. Figure 9 reveals the algorithm's ability to reproduce the observations above the Ku links. For the case of the rain event that occurred on 28 September, Fig. 10 provides a comparison between the Ku and radar fields. Figure 11 presents the boxplots of the "error" (i.e., the difference between the radar and Ku rainfall fields) for 1 and 4 km^2 rainfall fields. Figure 12 presents the cumulative rainfall recorded by

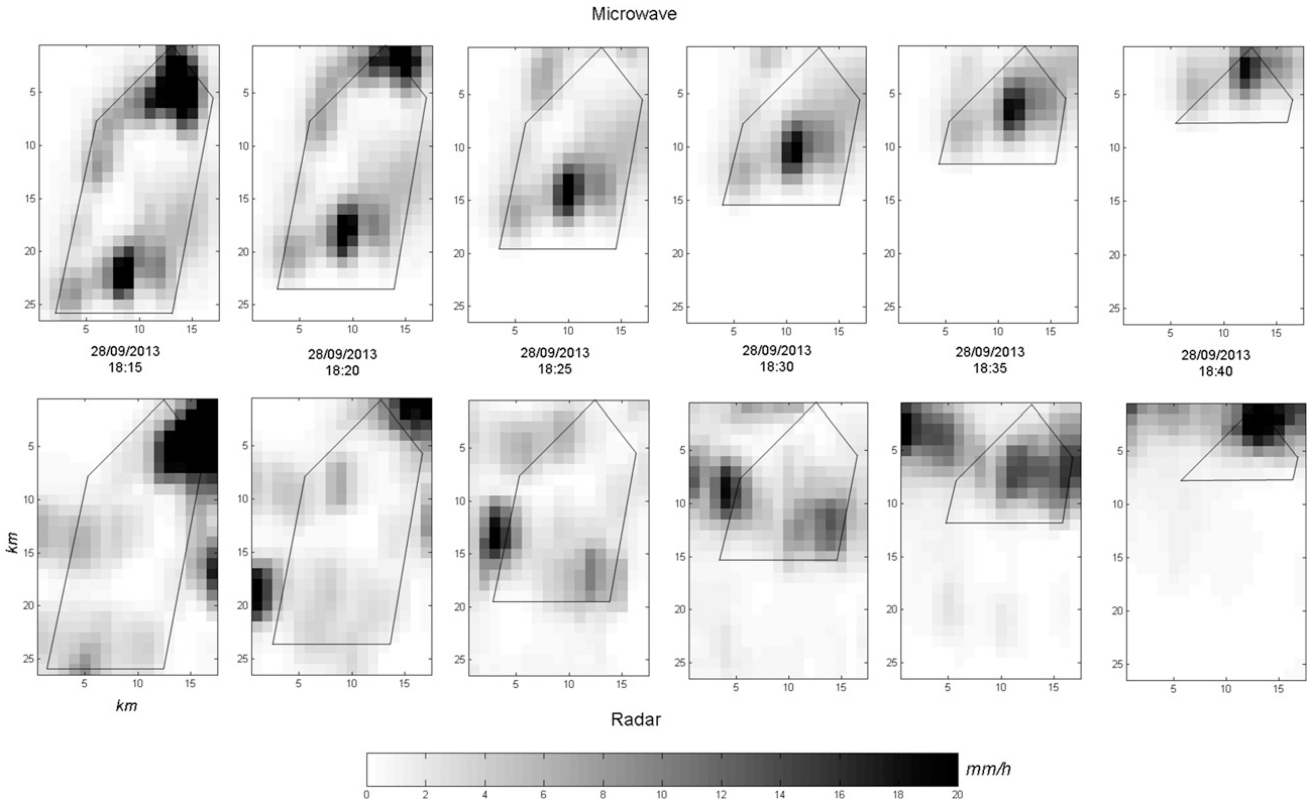

FIG. 10. (top) The 5-min integrated rain fields, rebuilt every 5 min by the assimilation algorithm, using Ku data recorded during the rainfall event of 28 Sep 2013. Following the hypothesis made in this study, the initially rebuilt field (at 1815 UTC, in the top left-hand corner) is advected at the next time steps only, which means that the rain cell passing over the Ku links at 1835 UTC was positioned to the southwest of the links at 1815 UTC. (bottom) Images of the same rainfall fields, when observed by the Bollène weather radar. The area delineated in red is the large assimilation area defined in section 5b. The small assimilation area is indicated by a triangle situated above the large area (below the Ku links).

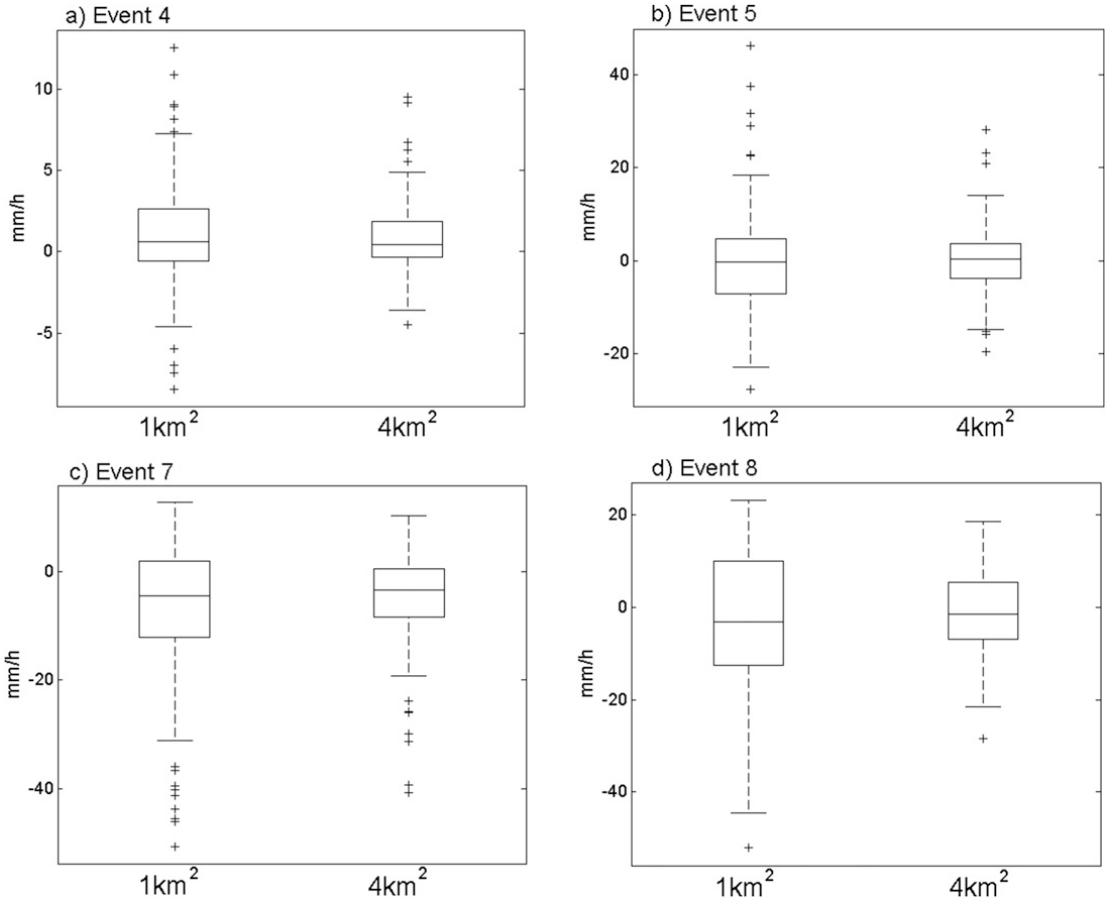

FIG. 11. Boxplots of the differences between radar and Ku rainfall rates in the assimilation area for the four studied advective events: (a) event 4, 28 Sep; (b) event 5, 29 Sep; (c) event 7, 20 Oct; (d) event 8, 20 Oct. For all of these events, we present the quantiles of the pixel-to-pixel differences on (left) a 1 km × 1 km grid and (right) a 2 km × 2 km grid.

the rain gauge at Mirabel, and the values of rainfall recorded by the radar and reproduced by the assimilation algorithm at the same location. The last two figures are restricted to the four advective events. Finally, Table 5 presents the correlation coefficients and regression line equations of Ku observations versus rain gauge data and radar measurements versus rain gauge data for various integration times ranging between 5 and 25 min.

1) RADAR COMPARISON

The Ku rainfall fields are first integrated in space and time, so that they have the same resolution as the radar data (i.e., 5 min in time and 1 km^2 in space). In this section, we first perform numerical comparisons between radar and Ku fields and confirm the assumption that four of the eight events could not be correctly represented by our advection model. In the following, only the four advective events will be described. We then focus on the case study of the event that occurred on 28 September 2013 to demonstrate the performance of the algorithm. Finally, more general conclusions are presented.

The main parameter (among those presented in Table 4) involved in the evaluation of the ability of our advection model to correctly represent the propagation of rain fields is the mean bias found in the large assimilation area. If the actual propagation is not advective, then there must be some ''creation'' or ''destruction'' of rainfall during the simulation. Thus, since the total rain is conserved by our model, large biases can be expected, especially in the areas not directly covered by the Ku links. In three events the biases are large: events 1 (−68%), 2 (−73%), and 6 (−40%). Also note that these three events show very flat regression lines (with slopes generally below 0.5). This means that the Ku fields are strongly smoothed out in comparison with the radar fields. Their associated regression coefficients are also rather small. The two latter remarks also apply to event 3. For instance, in the small assimilation area and over a 2 km × 2 km grid, only these four events (1, 2, 3, and 6) have regression coefficients of less than 0.7. On the contrary, the regression coefficients of the other events are above 0.8 (note that event 8 has a small regression

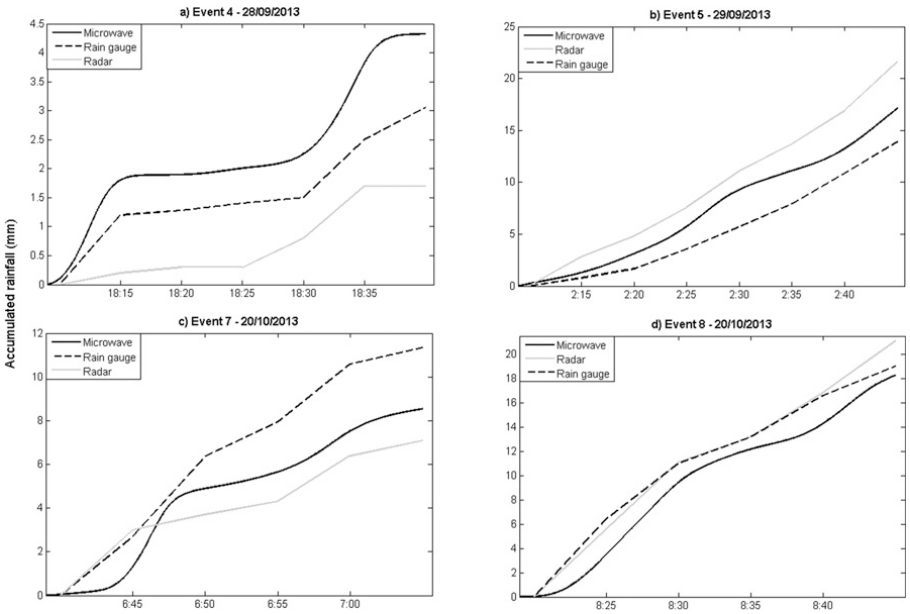

FIG. 12. Cumulative rainfall for the four rain events studied at Mirabel, Le Pradel (44°58′N, 4°50′E; see Fig. 2). The data extracted from the 500 m × 500 m instantaneous fields, rebuilt by the assimilation algorithm (continuous black lines), are plotted against rain gauge data from Le Pradel (continuous gray lines) and data extracted from the 1 km × 1 km radar fields.

coefficient of 0.39 at 1 km^2, but that this is no longer the case at 4 km^2, suggesting a scaling problem, as explained below).

The statistical criteria defined for radar maps in the previous section will now be analyzed in more detail in order to detect the advective properties of events (see Table 3, columns 5 and 6). The four events for which poor results (1, 2, 3, and 6) were obtained are the only ones whose cross-correlation mean and standard deviation are less than 0.85 and greater than 0.025, respectively. These events will be considered as nonadvective and will not be used again in the rest of this paper.

We will now focus on the 28 September event. As seen from the attenuation time series measured by the Ku sensor (see Fig. 9, solid lines) there are two distinct peaks along all of the four Ku links, with the first one occurring at approximately 1810 UTC and the second one at approximately 1835 UTC. The radar maps for this event (Fig. 10, bottom) show that the second peak corresponds to a rain cell that gradually intensified from 1825 to 1840 UTC as it moved northward while approaching the Ku sensor.

When looking at the results of the assimilation algorithm for this event, the following four main points can be noted:

TABLE 5. Point-scale rainfall estimations. Correlation coefficients, regression line equations at different time scales, and average biases between rain gauge and Ku data (third column) and between rain gauge and radar data (fourth column). Average rain rate recorded by the rain gauges. Four rain events and seven rain gauges are aggregated for these estimations (see section 6b).

	Time resolution (min)	Ku–rain gauges	Radar–rain gauges
r	5	0.63	0.69
	10	0.76	0.79
	15	0.77	0.83
	20	0.78	0.83
	25	0.82	0.84
Regression line equation	5	$0.53x + 0.64$	$0.61x + 0.64$
	10	$0.63x + 0.79$	$0.69x + 0.85$
	15	$0.62x + 1.08$	$0.69x + 1.16$
	20	$0.61x + 1.52$	$0.68x + 1.58$
	25	$0.57x + 1.58$	$0.64x + 1.56$
Bias (mm h^{-1})		−3.8	−1.9
Rain gauges avg rain rate (mm h^{-1})		24.5	

(i) The mean rainfall rate near to the microwave links (the small assimilation area) is suitably reproduced (with a mean bias of 1.1 mm h^{-1}), as is the mean rainfall rate in the large assimilation area (with a bias of −0.1 mm h^{-1}). This shows once again that one advantage of this measuring device and reconstruction method is that all of the rain falling above the microwave links is seen (as a result of the very good temporal resolution of the device: 10 s). Thus, because of the space-integrated feature of the data and the assimilation algorithm, whose effect is to connect the measurements recorded at successive time steps, the total rain falling in the area of the satlinks during an event is well reproduced by the algorithm.

(ii) The northern rain cell (corresponding to the first rainfall peak in Fig. 9) is correctly reproduced (see Fig. 10), while the southern cell is highly overestimated in the initial field (Fig. 10, left). This is a consequence of the model error. Pure advection is assumed by the propagation model, so that rain is perfectly conserved, whereas it has been seen above that the strength of the second rain cell increased while advancing. Consequently, the algorithm overestimates the initial rainfall so that, once advected above the microwave links, it provides realistic attenuation rates, albeit different from the actual ones. This model error can also be seen when comparing the radar/Ku correlation in the small assimilation area (0.85) with that obtained in the large assimilation area (only 0.48).

(iii) The projection of the second rainfall peak on the Ku links is smoothed out by the assimilation algorithm (Fig. 9). This is another consequence of the model error. As seen before (section 5c), the model's numerical diffusion is significant, and therefore the strong rain gradients cannot be preserved through advection: the rebuilt time series are then smoothed in comparison with the original ones. However, it may be noted that this smoothing is rather small for this event, to the extent that it does not appear when performing a regression between rebuilt and radar rainfall fields (Table 4: the slope of the regression line is 0.98, which is very close to 1).

(iv) The north/south variations of the rainfall fields seem to be better located and reproduced than the west/east variations (see Fig. 10). The event is mainly advected northward (azimuth: 191°), so that the Ku links are almost perpendicular to the rain front direction. The rain front then gradually moves northward along the links, helping the algorithm to distinguish between the north/south rain variations. (For northward advection, the northern part of the rain cell passes over the links before the southern part. For westward advection, both the northern and the southern parts of the rain cell pass simultaneously above the links.) However, it has already been pointed out above (section 5c) that this case (northward advection) is more difficult to solve numerically for the algorithm than is the other case (eastward advection). Additionally, all of the cases selected among the events shown in Table 3 are mainly advected northward.

The other events studied in this paper confirm the first point. The total rainfall is satisfactorily reproduced (Table 4; Fig. 11) both in the small and large assimilation areas.

The tendency of the algorithm to smooth the rain fields (point iii) is very clear for two events, namely, the fifth (29 September 2013) and the seventh (20 October 2013) events. This tendency appears both on the time series of rainfall above the microwave links and on the regression slopes between the radar and microwave fields (Table 4), with regression slopes of 0.44 and 0.40 at 1 km^2 (small assimilation area), respectively.

The radar/microwave pixel-to-pixel (1 km^2) comparisons performed in the small area show a good agreement for three events (4, 5, 7, with correlation coefficients of 0.85, 0.77, and 0.81, respectively), but this is not the case for event 8, with $r = 0.39$, and a y-intercept value of 26 mm h^{-1}, which is excessively large. The latter failure also appears clearly from Fig. 11, which shows large values of the quartiles of the radar/microwave difference (around ±10 mm h^{-1}). However, when the resolution of the radar and microwave fields are reduced to a 4 km^2 grid (Table 4; Fig. 11), the correlation coefficients significantly increase for all the events—to 0.92, 0.84, and 0.88 for events 4, 5, and 7, respectively, and from 0.39 to 0.82 for event 8—and the widths of the boxplots systematically decrease. This means that, even if the microwave field can be quite different from the radar field at the minimum resolution (1 km^2), the rebuilding algorithm seems to provide very accurate rain fields as soon as the resolution is set to less than 4 km^2. However, both the small size of the studied areas and the small number of rain events presented here require that more experiments be carried out to statistically confirm these conclusions.

Moreover, point-to-point comparisons are naturally less satisfactory in the large assimilation area (due to the model errors) but note the following characteristics:

(i) The covered area can be much larger, provided the actual rainfall propagation can be described by an advection model. For instance, results are very good in both assimilation areas for the 20 October events.

(ii) The model performance and the surface area covered can be increased by using more than one receiver.

2) RAIN GAUGE COMPARISON

The rain gauge data are point scale in space and time integrated (5-min resolution). To compare these data with the rainfall fields rebuilt from microwave measurements, the Ku rainfalls must be integrated over 5-min time periods and rain gauges must be assumed to provide good estimates of the rain fallen within the 0.5 km × 0.5 km cells of the assimilation algorithm, where the rain gauges are located. Figure 12 shows, for the four events studied, the cumulative rainfall measured by one of these rain gauges, located at Le Pradel (rain gauge 5, black dashed lines). The cumulative rainfalls measured by the Météo-France radar of Bollène on the 1 km^2 pixel, where the rain gauge is located, are also plotted (gray solid lines). There are significant differences between the total radar and rain gauge rainfall values—for instance, for the 20 October event, approximately 11.5 mm is recorded by the rain gauge and only 7 mm by the radar. However, note that the correlation coefficient between the data of the two rain gauges located near Le Pradel (rain gauges 5 and 6; see Fig. 2) is 0.97. The distance between these rain gauges is 415 m, which is close to the size of the grid boxes of our rebuilding algorithm. Figure 12 also shows the cumulative rainfalls at Le Pradel as reproduced by the algorithm (black solid lines). It should be noted that the microwave rainfall seems to be satisfactory: for all four events, the microwave rainfall values are closer to the rain gauge values than those provided by the radar. Therefore, the algorithm appears to provide a good assessment of the total rain fallen at particular ground points.

Another way to assess the performance of the algorithm on point-scale locations is to calculate the correlation coefficient between Ku and rain gauges data. Table 5 (third column) shows the correlation coefficients and the regression line equations between rain gauges and Ku rainfalls for different integration times, from 5 to 25 min. The data used are the values obtained for the four advective events and at the locations of the seven rain gauges within the large assimilation area (our algorithm cannot be used to estimate rainfall over rain gauges 1, 2, and 4, as seen in Fig. 2).

The Ku-link algorithm slightly underestimates the mean rain rate (the bias is −3.8 mm h^{-1} compared with the average rain gauge rain rate of 24.5 mm h^{-1}). Moreover, the radar also underestimates the point-scale mean rain rate but with a smaller difference (−1.9 mm h^{-1}). The reasons for this underestimation is difficult to elucidate (with only four events), although these results are consistent with Table 4, which shows that the Ku links slightly underestimate the mean rain rates in comparison with those derived from radar data for three out of four events.

For an integration time of 5 min, the correlation coefficient is rather small (0.63) and the slope of the regression line is also small (0.50). This confirms the tendency of our algorithm to smooth the rain fields, which is more pronounced because we compare space-integrated (microwave) and point-scale (rain gauges) data. A time scale is then sought at which a better fit between the two kinds of data would be obtained (by integrating the data over longer time periods). The correlation coefficients between Ku and rain gauge data are 0.76 and 0.77 at 10 and 15 min, respectively, while the slope of the regression lines are 0.63 and 0.62, respectively. Thus, the algorithm seems to lead to good results on point-scale locations for integration times larger than 10 min. For longer times, the correlation coefficient does not significantly increase (0.82 for 25 min). It should be noted that the correlation coefficients between radar and rain gauge data (Table 5, right column) are in good agreement, ranging from 0.69 for a 5-min integration time to 0.79 for 10 min and 0.84 for 25 min. This similarity between radar gauges and satlink gauges is likely because the measurements compared are of a different nature. Radar and Ku data are spatially integrated. When they are compared with point-scale data (rain gauges), the obtained regression slopes are substantially smaller than 1.

7. Conclusions

The use of a 4DVAR data assimilation algorithm allows certain local rain fields to be rebuilt from space-integrated rainfall data recorded by a very sparse network of Ku-band satellite-to-Earth microwave links. We produce local rainfall fields with a 10-s temporal resolution and a 0.5 km × 0.5 km spatial resolution, with features similar to those detected on radar maps having a temporal resolution of 5 min and a spatial resolution of 1 km^2.

By working first on simulated data, we evaluate some of the consequences of the recording device's geometry on the results, including the algorithm's difficulty in reproducing northward advected events. We also note the difficulties resulting from numerical diffusion of the advection model used to propagate the rain fields through time. This numerical diffusion tends to smooth the rain fields after a small number of iterations over time and, on the contrary, to allow the algorithm to produce excessively strong gradients on the initial rain

fields, despite the addition of a filter term to the assimilation algorithm's cost function. These effects were also observed on real data recorded in the southeastern area of France in 2013. By comparing rebuilt fields with radar maps and point-scale rain gauge data, the algorithm generally obtains good results in the area directly below the Ku links, with a very satisfactory reproduction of the average rain rate in particular. However, we note that the performance of the model depends strongly on the realism of the advection scheme used to model rain field propagation in time and space. In a mountainous area such as the Cévennes–Vivarais region studied in this paper, the propagation of rain events cannot be reduced to a pure translation for more than a short period of time. We therefore worked with short events only (less than 45 min), selected according to their advective characteristics. The rebuilt areas (the small and large assimilation areas mentioned above) are thus quite small. However, a higher proportion of observed events (and the analysis of larger areas) could be correctly processed through the use of several Ku receivers. Moreover, it is shown that the use of two receivers allows the advection velocity to be determined using attenuation time series only.

Our results have been validated by comparing them with radar maps and rain gauge data. In cases where rain gauge data are available, it would be straightforward to adapt this algorithm to the assimilation of both microwave and rain gauge data (data fusion). A data fusion experiment (involving radar, rain gauge, and microwave link data) has been proposed by Bianchi et al. (2013). The addition of point-scale data (such as that provided by rain gauges) to spatially integrated data could enable the extrema, shapes, and characteristics of the rain fields to be more accurately determined.

In the present study, rain is assumed to fall vertically, and the attenuation determined with the Ku sensor is projected orthogonally onto the ground. A further improvement in the rebuilding algorithm could be achieved by taking into account the drops' horizontal displacements, produced by the wind field during their fall.

A further improvement in the algorithm's numerical ability to minimize the cost function, and to physically constrain the resulting rain fields, could be achieved by replacing the filter term of the cost function—for example, by including additional constraints on the rain field features, derived from their known meteorological characteristics.

Finally, it is interesting to note that the method described in this study could be a credible alternative to the use of radars for the observation of unmonitored zones (e.g., mountainous catchment areas, or the countryside of developing countries). In the long term, rain monitoring networks could be set up in urban areas, through the addition of a small dedicated device to existing home TV satellite antennas.

Acknowledgments. This work was supported by the French Programme National de Télédétection Spatiale (PNTS, http://www.insu.cnrs.fr/pnts) Grant PNTS-2013-01 and by the Institut Pierre-Simon Laplace (IPSL). The authors wish to thank Drs. G. Molinié and B. Boudevillain from the Laboratoire d'étude des Transferts en Hydrologie et Environnement for supplying us with the HPiconet data.

REFERENCES

Allen, D. J., A. R. Douglass, R. B. Rood, and P. D. Guthrie, 1991: Application of a monotonic upstream-biased transport scheme to three-dimensional constituent transport calculations. *Mon. Wea. Rev.*, **119**, 2456–2464, doi:10.1175/1520-0493(1991)119<2456:AOAMUB>2.0.CO;2.

Barthès, L., and C. Mallet, 2013: Rainfall measurement from opportunistic use of earth-space link in Ku Band. *Atmos. Measur. Tech. Discuss.*, **6**, 2113–2150, doi:10.5194/amtd-6-2113-2013.

Bianchi, B., P. Jan van Leeuwen, R. J. Hogan, and A. Berne, 2013: A variational approach to retrieve rain rate by combining information from rain gauges, radars, and microwave links. *J. Hydrometeor.*, **14**, 1897–1909, doi:10.1175/JHM-D-12-094.1.

Chwala, C., and Coauthors, 2012: Precipitation observation using microwave backhaul links in the alpine and pre-alpine region of southern Germany. *Hydrol. Earth Syst. Sci.*, **16**, 2647–2661, doi:10.5194/hess-16-2647-2012.

Delrieu, G., B. Boudevillain, J. Nicol, B. Chapon, P.-E. Kirstetter, H. Andrieu, and D. Faure, 2009: Bollène-2002 Experiment: Radar quantitative precipitation estimation in the Cévennes–Vivarais region, France. *J. Appl. Meteor. Climatol.*, **48**, 1422–1447, doi:10.1175/2008JAMC1987.1.

Desa, M., and J. Niemczynowicz, 1997: Dynamics of short rainfall storms in a small scale urban area in Coly Limper, Malaysia. *Atmos. Res.*, **44**, 293–315, doi:10.1016/S0169-8095(97)00017-3.

Drobinski, P., and Coauthors, 2013: HyMeX, le cycle de l'eau méditerranéen à la loupe. *Meteorologie*, **80**, 23–36, doi:10.4267/2042/48792.

Gilbert, J. C., and C. Lemaréchal, 2006: The module M1QN3. INRIA Rep., version 3.3, 21 pp.

Giuli, D., A. Toccafondi, G. Biffi Gentili, and A. Freni, 1991: Tomographic reconstruction of rainfall fields through microwave attenuation measurements. *J. Appl. Meteor.*, **30**, 1323–1340, doi:10.1175/1520-0450(1991)030<1323:TRORFT>2.0.CO;2.

——, L. Facheris, and S. Tanelli, 1999: Microwave tomographic inversion technique based on stochastic approach for rainfall fields monitoring. *IEEE Trans.* Geosci. Remote Sens., **37**, 2536–2555, doi:10.1109/36.789649.

Kalnay, E., 2003: *Atmospheric Modeling, Data Assimilation, and Predictability*. Cambridge University Press, 364 pp.

Leijnse, H., R. Uijlenhoet, and J. Stricker, 2008: Microwave link rainfall estimation: Effects of link length and frequency, temporal sampling, power resolution, and wet antenna attenuation. *Adv. Water Resour.*, **31**, 1481–1493, doi:10.1016/j.advwatres.2008.03.004.

Lovejoy, S., and D. Schertzer, 1990: Multifractals, universality classes and satellite and radar measurements of cloud and rain fields. *J. Geophys. Res.*, **95**, 2021–2034, doi:10.1029/JD095iD03p02021.

Maitra, A., and K. Chakravarty, 2005: Ku-band rain attenuation observations on an earth–space path in the Indian region. *Proc. XXVIIIth General Assembly*, URSI, New Delhi, India, F01.4(0650)]. [Available online at http://www.ursi.org/Proceedings/ProcGA05/pdf/F01.4%280650%29.pdf.]

Molinié, G., D. Ceresetti, S. Anquetin, J. D. Creutin, and B. Boudevillain, 2012: Rainfall regime of a mountainous Mediterranean region: Statistical analysis at short time steps. *J. Appl. Meteor. Climatol.*, **51**, 429–448, doi:10.1175/2011JAMC2691.1.

Nardi, L., C. Sorror, F. Badran, and S. Thiria, 2009: YAO: A software for variational data assimilation using numerical models. *Computational Science and Its Applications—ICCSA 2009*, O. Gervasi et al. Eds., Lecture Notes in Computer Science, Vol. 5593, 621–636.

Navon, I. M., 2009: Data assimilation for numerical weather prediction: A review. *Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications*, S. K. Park and L. Xu, Eds., Vol. 1, Springer, 21–65.

Overeem, A., H. Leijnse, and R. Uijlenhoet, 2013: Country-wide rainfall maps from cellular communication networks. *Proc. Natl. Acad. Sci. USA*, **110**, 2741–2745, doi:10.1073/pnas.1217961110.

Schleiss, M., and A. Berne, 2010: Identification of dry and rainy periods using telecommunication microwave links. *IEEE Geosci. Remote Sens. Lett.*, **7**, 611–615, doi:10.1109/LGRS.2010.2043052.

Smolarkiewicz, P. K., 1983: A simple positive definite advection scheme with small implicit diffusion. *Mon. Wea. Rev.*, **111**, 479–486, doi:10.1175/1520-0493(1983)111<0479:ASPDAS>2.0.CO;2.

Tabary, P., 2007: The new French operational radar rainfall product. Part I: Methodology. *Wea. Forecasting*, **22**, 393–408, doi:10.1175/WAF1004.1.

Thiria, S., F. Badran, and C. Sorror, 2006: Yao: Un logiciel pour les modèles numériques et lassimilation de données. LOCEAN Research Rep., 60 pp.

Verrier, S., L. De Montera, L. Barthès, and C. Mallet, 2010: Multifractal analysis of African monsoon rain fields, taking into account the zero rain-rate problem. *J. Hydrol.*, **389**, 111–120, doi:10.1016/j.jhydrol.2010.05.035.

Zinevich, A., P. Alpert, and H. Messer, 2008: Estimation of rainfall fields using commercial microwave communication networks of variable density. *Adv. Water Resour.*, **31**, 1470–1480, doi:10.1016/j.advwatres.2008.03.003.

——, H. Messer, and P. Alpert, 2009: Frontal rainfall observation by a commercial microwave communication network. *J. Appl. Meteor. Climatol.*, **48**, 1317–1334, doi:10.1175/2008JAMC2014.1.

Copyright of Journal of Atmospheric & Oceanic Technology is the property of American Meteorological Society and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.

Annexe M: Article "4-D-VAR assimilation of disdrometer data and radar spectral reflectivities for rain drop size distribution and vertical wind retrievals", révision Atmospheric Measurement Technique.

Manuscript prepared for Atmos. Meas. Tech. Discuss.
with version 2015/04/24 7.83 Copernicus papers of the LaTeX class copernicus.cls.
Date: 11 April 2016

4-D-VAR assimilation of disdrometer data and radar spectral reflectivities for rain drop size distribution and vertical wind retrievals

F. Mercier, A. Chazottes, L. Barthès, and C. Mallet

Université de Versailles-Saint-Quentin, LATMOS-IPSL, CNRS-INSU,
Laboratoire Atmosphères, Milieux, Observations Spatiales, Université Paris-Saclay, France

Correspondence to: F. Mercier (francois.mercier@latmos.ipsl.fr)

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

Abstract

This paper presents a novel framework for retrieving the vertical raindrop size distribution (DSD) and vertical wind profiles during light rain events. This is also intended as a tool to better characterize rainfall microphysical processes. It consists in coupling K band Doppler spectra and ground disdrometer measurements (raindrop fluxes) in a 2-D numerical model propagating the DSD from the clouds to the ground level. The coupling is made via a 4-D-VAR data assimilation algorithm. As a first step, in this paper, the dynamical model and the geometry of the problem are quite simple. They do not allow to encompass the complexity implied by all rain microphysical processes (evaporation, coalescence break-up and horizontal air motion are not taken into account). In the end, the model is limited to the fall of droplets under gravity, modulated by the effects of vertical winds. The framework is thus illustrated with light, stratiform rain events.

We firstly use simulated data sets (data assimilation twin experiment) to show that the algorithm is able to retrieve the DSD profiles and vertical winds. It also demonstrates the ability of the algorithm to deal with the atmospheric turbulence (broadening of the Doppler spectra) and the instrumental noise. The method is then applied to a real case study which happened in the south-west of France during the autumn 2013. The data set collected during a long, quiet event (6 h duration, rain rate between 2 and 7 $\mathrm{mm\,h^{-1}}$) comes from an optical disdrometer and a 24 GHz vertically pointing Doppler radar. We show that the algorithm is able to reproduce the observations and retrieve realistic DSD and vertical wind profiles, when compared to what could be expected for such a rain event.

A perspective for this study is to apply it to extended data sets for a validation with independent data, which cannot have been conducted with our available 2013 data. Other data sets would also help to parameterize more phenomena needed in the model (evaporation, coalescence/break-up) to apply the algorithm to convective rain and to evaluate the adequacy of the model's parameterization.

1 Introduction

The knowledge of the raindrop size distribution (DSD) both at ground and throughout the vertical atmospheric profile is an important topic. In remote sensing, parameters of the Z–R relationship, allowing to convert a radar reflectivity into a rain rate, depend on rain microphysics. These parameters are generally based on a particular DSD distribution and are supposed constant through the entire atmospheric column. In practice, rain rates estimated from reflectivity through the Z–R relationship can vary at least by a factor of two depending on the DSD (List, 1988). In telecommunications, and especially for a ground to space telecommunication, rain induced attenuation also depends on rain microphysics, especially in the Ku band and beyond. More generally the observation of vertical DSD profiles is fundamental to address many questions regarding the numerous processes involved during a rain event. A large number of investigations have been conducted in order to model the vertical evolution of the DSD during rain for different meteorological situations (List and McFarquhar, 1990; Brown Jr, 1993; Hu, 1995; Prat and Barros, 2007a, b; Barros et al., 2008; Mcfarquhar, 2010). The parameterization of the different phenomena occurring during a droplet fall is a complex task due to the great number of processes involved. Among them, some are considered as predominant, notably a raindrop can breakup, coalesce or evaporate. It is also subject to sorting by gravity along with updrafts, downdrafts, and horizontal wind. All these processes, playing an important role in the nature of the DSD, are the subject of numerous studies. More specifically, raindrop sorting induced by vertical drafts can play an important role in the vertical DSD variability. For instance, Kollias et al. (2001) shows that updraft structures can cause horizontal and vertical sorting of raindrops with a lack of large raindrops ($> 3\,\mathrm{mm}$) in the updraft core and an increase at the periphery. Moreover, concerning evaporation, drizzle can never reach the ground in specific meteorological situations (VanZanten et al., 2005). On the other hand, Mcfarquhar (2010) summarizes the main results on models concerning the coalescence and break-up processes. However, he explains that these models, built on laboratory or simulation experiments generally suffer from a lack of validation under real conditions. Nevertheless, a very large number of di-

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

rect or indirect observations of DSD are available either on the ground, from disdrometers (among others, Joss and Waldvogel, 1967; Kruger and Krajewski, 2002; Delahaye et al., 2006), or at various heights, from the use of Doppler radars (Kollias et al., 2002; Peters et al., 2005). Several studies have been carried out to compare these different observations (Bringi et al., 2003; Tokay et al., 2009; Schleiss and Smith, 2015). However, the combination of these sources of information about the DSD is a difficult task given their very different natures. It is even more demanding when you add the different spatio-temporal scales and the various locations of the measurements.

In order to improve vertical DSD profile retrievals, this study establishes a framework allowing the assimilation at fine scale ($< 5\,\mathrm{min}$, $100\,\mathrm{m}$) of heterogeneous observations in a dynamic model. The developed method allows the use of observations from different instruments, each with its own spatio-temporal resolution. Compared to direct observation retrievals, merging data in a dynamic model includes spatio-temporal consistency to the DSD retrievals. Moreover, it could help to assess and improve the model parameterization.

In this paper, we will focus on disdrometer and vertical-radar-reflectivity collocated observations. The data are merged through the use of a 4-D-VAR assimilation algorithm. The evolution in time and space of the DSD is based on a very simplified version of the Hu and Srivastava (1995) propagation model (see Eq. 3). It is actually limited to the fall of droplets under gravity, solely modulated by the effect of vertical wind. The goal of this study is to develop a new mathematical framework in a simplified context. Once introduced, it will be tested by evaluating its performances, sensitivity and limits on both simulated data and real data (a calm stratiform real event of light rain). Notice that, even if simple, this model insures a resulting DSD in adequacy (through the model) with all the instruments (and their different scales). This aspect is not negligible when, as mentioned previously, you consider all the work done to establish consistent $Z{-}R$ relationships.

Improvements of this framework (encompassing evaporation, horizontal air motion, coalescence / break-up) as well as validation, both requiring more measurements, are deferred to future work allowed by extended data sets.

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

The paper is organized as follows. In Sect. 2, we present the simple model used and discuss the terms of the complete Hu and Srivastava (1995) propagation model, namely: wind, collision and evaporation terms. Then model simplifications are retained considering different hypotheses. General principles behind 4-D-VAR data assimilation are presented and a focus is made on the cost function and regularization terms. Disdrometer and radar observation operators are then given. Finally we describe the algorithm used to retrieve the DSD and vertical wind fields from radar and disdrometer data. Then, in Sect. 3, the instruments available and the case study are described. In Sect. 4, through a twin experiment (i.e. on simulated data), we explore the impact of observation errors in addition to non-modeled phenomena on the algorithm performances as well as the quality of the retrievals. Section 5 gives some results obtained in a stratiform real case study. Finally, Sect. 6 concludes the study. Perspectives are also drawn concerning the improvements needed to extend the application of the method to a more general context.

2 Assimilation method

The aim of this work consists in retrieving information about the drop size distribution (DSD) vertical profiles by linking the measurements made by different instruments at different scales and heights. To carry out this relationship, we will use a dynamic model which propagates the DSD through space and time (mentioned below as “propagation model”). In this section, firstly, we describe the corresponding partial differential equation (PDE) model used. We as well introduce the associated discretization scheme, and the unknowns of our model. Then, we explain how the 4-D-VAR data assimilation algorithm combines the data available through the model to retrieve its unknowns. The second part of this section details this algorithm.

2.1 Propagation model

As explained before, this study develops a simple framework and so uses a very simple model presented in Sect. 2.1.1. This simple model has the merit to add a spatio-temporal

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

coherence to the DSD field. Since it only incorporates the effects of gravity and vertical wind it lacks the complexity to fully model rain.

In Sect. 2.1.2, we introduce a complete PDE governing the evolution of the DSD (Hu and Srivastava (1995)). This propagation model is not in the purview of this article. It is solely used to discuss the different terms of the complete PDE (Eq. 3) and see how their inclusion or exclusion could affect the results. Finally, in Sect. 2.1.3, we present the effective "simplistic" numerical model and the discretization underlying hypotheses.

2.1.1 Simple propagation model

The DSD, noted N, is the number of raindrops by unit of volume and diameter (unit: m^{-4}). It is a function of time, position and diameter. In the framework presented here, we will work in an air column, and so limit the study to a time-height-diameter space, so that $N = N(t, z, D)$. Then, the PDE modelling the vertical evolution of DSD has the form :

$$\frac{\partial N}{\partial t}(t,z,D) + \frac{\partial}{\partial z}\left(\left[v(D) + w(t,z) \right] N(t,z,D) \right) = 0 \tag{1}$$

The first term of Eq. (1) represents the instantaneous variation of N. The second term is the vertical advection of droplets : w is the vertical wind (component of the 3D-wind **W** along the vertical axis $\boldsymbol{u}_z$), v is the terminal velocity of droplets under gravity. In this study, the parameter v is supposed to depend only on the diameter D, according to the Atlas et al. (1973) relation:

$$v(D) = 9.65 - 10.3e^{-600D}, \tag{2}$$

with D in m and v in $m\,s^{-1}$. This relation was fitted to be in good agreement with Gunn and Kinzer (1949) measurements (which are often used as a reference) in the diameter range [0.6–5.8 mm] (Atlas et al., 1973).

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

Eq. (1) can be considered as a simplified version of the model proposed for instance in Hu and Srivastava (1995). We will now present this complete model and discuss its different terms.

2.1.2 Complete propagation model (Hu and Srivastava (1995)) and discussion of its different terms

In the Hu and Srivastava (1995) study, N depends on time, 3 coordinates of space, and diameter ($N = N(t, x, y, z, D)$) and obey the PDE :

$$\frac{\partial N}{\partial t} + \mathrm{div}\left[\left(v\boldsymbol{u}_z + \mathbf{W}\right)N\right] + \frac{\partial}{\partial D}\left(\tau \frac{\mathrm{d}D}{\mathrm{d}m} N\right) = I. \tag{3}$$

We recognize the two first terms of Eq. (1) with the difference that we have a 3D-wind **W** advection here. The third term of Eq. (3) represents the evaporation. τ is the net rate of change of droplets masses (unit $\mathrm{kg\,s^{-1}}$) and $\mathrm{d}D/\mathrm{d}m$ is the derivative of the diameter-mass relation for spheric drops (unit $\mathrm{m\,kg^{-1}}$). The last term I represents the collision effects (drop mass changes not due to evaporation).

3-D wind W

The vertical wind is considered to add an offset to the terminal velocity of drops. Besides, the knowledge of the real droplet vertical speed is critical to have information about the DSD from the return power spectra of a vertically pointing Doppler radar. This question has been widely treated in the literature (Lhermitte, 1988; Giangrande et al., 2012; Tridon et al., 2013) and it shows that we have to take the vertical wind into account to properly deal with Doppler radar data.

The horizontal wind is the main force which makes the droplets move in the horizontal plane. However, with only one vertically-pointing radar and colocalized ground point measurements of the DSD, it is not possible to record the horizontal variability of the DSD. Consequently, as mentioned above, we limit the study to a (z, t) plane. This simplification

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

is consistent if the horizontal air motions are weak or if the DSD is homogeneous in the horizontal plane. (Of course the DSD can be inhomogeneous with height).

To limit the errors caused by this simplification, we will only treat quiet, stratiform light rain events (see the case study in Sect. 3.3) and we will study the ability of the algorithm to reproduce the observations.

Evaporation (3rd term of Eq. 3)

The evaporation is mathematically modeled as a term of advection along the mass of droplets coordinate. The advection speed (parameter τ of Eq. 3) depends on the drop diameter, and on several meteorological variables, including pressure, temperature and mainly humidity. Seifert (2008) proposes a parameterization of this term.

According to this parameterization and for long cold stratiform rain events similar to the one treated in this study (see Sect. 3.3), we verified (not shown) that evaporation can be neglected.

Collisions (4th term of Eq. 3)

Collisions between drops can lead to coalescence (two drops producing one drop) or break-up (two drops producing at least two drops). Coalescence/break-up phenomena are generally assumed to be an important factor governing the DSD temporal evolution (Mcfarquhar, 2010; Prat et al., 2012), even if some studies reconsider this assumption (Villermaux and Bossa, 2009; Villermaux and Eloi, 2011). The phenomena have been widely investigated. Numerous parameterizations have been proposed. Among them, some characterize the ability of 2 drops to coalesce, depending on the energy involved in the collision (Low and List, 1982b; Brazier-Smith et al., 1972 or Straub et al., 2010). Others are more interested in the distribution of resulting drops in case of break-up. (The resulting distributions can be based on laboratory experiments: Low and List, 1982a, b, or on numerical fluid dynamics models: Schlottke et al., 2010; Beheng, 2010.)

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

Since our first objective is to retrieve DSD profiles using data assimilation, we restrict ourselves to a simple framework, and we do not take into account the collisions between drops. Moreover, among the wide literature mentioned above, many authors show that the coalescence/break-up processes are less critical for low rain rates (List et al., 1987; Prat and Barros, 2007a; Barthes and Mallet, 2013). Thus, this paper will focus only on a light rain event in order to limit the errors caused by this simplification. And as for the horizontal wind exclusion, we will study the ability of the algorithm to reproduce the observations despite this simplification.

2.1.3 Numerical model (discretization)

This simplified PDE used in this study (Eq. (1)) is discretized on a (time/height/diameter) Arakawa C-grid (Arakawa and Lamb, 1977), meaning that N is evaluated at the grid-box centers while w is evaluated at the grid-box faces. For this discretization of the PDE, we use the finite difference scheme of Smolarkiewicz (1983), developed for advective problems in atmospheric modeling requiring to deal with a large number of variables. For this scheme, the Courant–Friedrichs–Lewy stability condition is:

$$\Delta t \sqrt{\frac{(v+w)^2}{\Delta z^2}} < \frac{1}{\sqrt{2}} \tag{4}$$

with Δt and Δz the time and height steps, respectively. In order to be consistent with radar observations (see Sect. 3.1), we choose $\Delta z = 100\,\mathrm{m}$. Then, according to the usual range of drop terminal velocities ($v \in [0;\ 10]\,\mathrm{m\,s^{-1}}$) and of stratiform vertical wind speeds ($w \in \pm 2\,\mathrm{m\,s^{-1}}$), we choose $\Delta t = 5\,\mathrm{s}$. The number of time and height discretization steps are noted NT and NZ, respectively. In order to have a good coverage of the range of diameters, they are discretized from $D_{\mathrm{min}} = 0.2\,\mathrm{mm}$ to $D_{\mathrm{max}} = 7.5\,\mathrm{mm}$, with a diameter step $\Delta D = 0.1\,\mathrm{mm}$. The resulting number of diameter bins ND is thus equal to 73.

We note $\tilde{\boldsymbol{N}} = (\tilde{\boldsymbol{N}}_i^n)_{n \in [\![1:\mathrm{NT}]\!], i \in [\![1:\mathrm{NZ}]\!]}$ the discretized DSD field. $\tilde{\boldsymbol{N}}_i^n$ indicates a specific discretized DSD at grid point (n, i), namely, at time $t^n = t^0 + n\Delta t$ and height $z_i = z_0 - i\Delta z$.

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

Similarly, the notations $\tilde{\mathbf{w}}$ and $\tilde{v}$ stand respectively for the discretized wind field and terminal velocity vector.

The unknowns of such a model are:

1. The initial (t^0) DSD field (initial condition): $(\tilde{\boldsymbol{N}}_i^0)_{i\in[\![1:\mathrm{NZ}]\!]}$. However, if we suppose that the model starts running before the beginning of the rain event, we can suppose that this initial field is 0 everywhere. This is the assumption used hereafter.
2. The top (z_0) DSD field (boundary condition): $(\tilde{\boldsymbol{N}}_0^n)_{n\in[\![1:\mathrm{NT}]\!]}$. This corresponds to NT DSD bins, leading to NT $\times$ ND unknowns.
3. The vertical wind time-height field $\tilde{\mathbf{w}}$ (which is, in our case, a parameter of the model) represents NT $\times$ NZ unknowns.

In order to reduce the dimension of the problem and thus to substantially reduce the number of degrees of freedom, we have to add some a priori information. Since the density of raindrops in the atmosphere is well approximated by a gamma distribution (Mcfarquhar (2010)), we parameterize the top DSD $(\tilde{\boldsymbol{N}}_0^n)_{n\in[\![1:\mathrm{NT}]\!]}$ under a gamma form:

$$N(D) = \alpha \cdot f(D;\theta,k) \tag{5}$$

with α the total volumetric number of droplets (unit m^{-3}), and f the gamma probability density function (unit m^{-1}):

$$f(D;\theta,k) = \frac{D^{k-1}}{\Gamma(k)\theta^k} e^{-\frac{D}{\theta}} \tag{6}$$

with k the shape coefficient and θ the scale coefficient.

Thanks to this parameterization, for each time step n, we limit the top DSD $(\tilde{\boldsymbol{N}}_0^n)_{n\in[\![1:\mathrm{NT}]\!]}$ determination to the choice of 3 parameters and so divide the resulting number of unknowns by almost 25 (from 73 to 3). We note $\tilde{\mathbf{x}}$ the corresponding 3 parameters field (dimension: $3\times$NT): $\tilde{\mathbf{x}} = (\tilde{\mathbf{x}}^n)_{n\in[\![1:\mathrm{NT}]\!]}$. Concerning the implementation of the algorithm, at each time step

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

n, $\tilde{\mathbf{x}}^n$ is directly converted into a top DSD $\tilde{\boldsymbol{N}}_0^n$ using Eqs. 5 and 6, so that the propagation model can use it as an input data.

Finally, the problem is reduced to the determination of the parameter field $\tilde{\mathbf{x}}$ along with the vertical wind field $\tilde{\mathbf{w}}$, resulting in a total number of unknowns of $(\mathrm{NZ}+3)\times\mathrm{NT}$.

2.2 Data assimilation algorithm

2.2.1 General principles

In the previous section, we have seen that the model propagates the top DSD $(\tilde{\boldsymbol{N}}_0^n)_{n\in[\![1:\mathrm{NT}]\!]}$ through space and time according to the vertical wind $\tilde{\mathbf{w}}$. Besides the numerical model, we have a full set of observations at various heights and times (from radar and disdrometer, see Sects. 3.1 and 3.2).

The retrieval of the top DSD and of the vertical wind is thus carried out through the use of variational data assimilation which basically consists in minimizing the distance between the model and the observations (see Fig. 1 which details our contextualized assimilation process).

In this study, the well-known 4-D-VAR data assimilation algorithm is used to allow the fusion of observations thanks to our propagation model. Detailed descriptions of its theoretical principles can be found in Kalnay (2003); Bocquet (2005), or online on the European Centre for Medium-Range Weather Forecasts (ECMWF) website: http://www.ecmwf.int/sites/default/files/Data%20assimilation%20concepts%20and%20methods.pdf.

To perform a 4-D-VAR assimilation, we still have to build a 4-D-VAR cost function which is crucial since it is used to evaluate the gap between the observations and the top DSD $(\tilde{\boldsymbol{N}}_0^n)_{n\in[\![1:\mathrm{NT}]\!]}$ propagated by the model to different times and heights. The numerical minimization of the cost function allows to retrieve a set of unknowns (namely, $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{w}}$) which minimize the cost function.

In the data assimilation context, the cost function has usually the form:

$$J(\tilde{\boldsymbol{N}}) = J_{\mathrm{b}}(\tilde{\boldsymbol{N}}) + J_{\mathrm{o}}(\tilde{\boldsymbol{N}}), \tag{7}$$

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

with $\tilde{\boldsymbol{N}}$ the discretized DSD as previously defined. J_o is the observation term of the cost function, presented later. J_b is the background term (also called first guess) of the cost function. J_b keeps the solution in the vicinity of a given a priori state. In our case, we suppose that there is no background available, so that J_b will be dropped from the cost function. Moreover, we add regularization and penalization terms (see Sect. 2.2.4).

The cost function finally used has the form:

$$J(\tilde{\mathbf{x}}, \tilde{\mathbf{w}}) = J_\mathsf{o}(\tilde{\mathbf{x}}, \tilde{\mathbf{w}}) + J_\mathsf{r}(\tilde{\mathbf{x}}, \tilde{\mathbf{w}}) + J_x(\tilde{\mathbf{x}}) + J_w(\tilde{\mathbf{w}}), \tag{8}$$

where the cost function is the sum of a regularization term J_r along with two penalization terms J_x, dedicated to $\tilde{\mathbf{x}}$ the DSD parameter field, and J_w, dedicated to $\tilde{\mathbf{w}}$ the wind field. These three terms will be discussed and explained in Sect. 2.2.4.

In the end, the cost function $J(\tilde{\mathbf{x}}, \tilde{\mathbf{w}})$ is a function of both the top DSD parameter field ($\tilde{\mathbf{x}}$) and the wind field ($\tilde{\mathbf{w}}$) that allows to compute the overall distance between a propagated set of parameters ($\tilde{\mathbf{x}}, \tilde{\mathbf{w}}$) and the observations. By minimizing the cost function, the algorithm will estimate an optimal set of ($\tilde{\mathbf{x}}, \tilde{\mathbf{w}}$) minimizing this distance. Once the assimilation process achieved, for each time step, we will have a top DSD parameter and wind values for all heights.

2.2.2 Observation term of the cost function J_o

J_o, the observation term of the cost function, has the form:

$$J_\mathsf{o}(\tilde{N}) = \frac{1}{2} \sum_{n \in \mathbf{[}1\text{:NT}\mathbf{]}} \sum_{i \in \mathbf{[}1\text{:NZ}\mathbf{]}} \left\| \boldsymbol{y}_i^n - \mathbf{H}_{i,n} \tilde{\boldsymbol{N}}_i^n \right\|^2. \tag{9}$$

For a given grid point (n, i), $\boldsymbol{y}_i^n$ is the vector composed of the observations from radar and disdrometer. $\mathbf{H}_{i,n}$ is the operator projecting the DSD on the observation space. Note that, here, the covariance matrix of the observation error is the identity.

At time t^n, using the vertical wind field $\tilde{\mathbf{w}}$, the model has propagated the top DSD $\tilde{\boldsymbol{N}}_0^{n'}$ from times $t^{n'} < t^n$. We note M_i^n the numerical model propagating all the necessary top

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

DSD $\tilde{\boldsymbol{N}}_0^{n'}$ to compute $\tilde{\boldsymbol{N}}_i^n$, the DSD at time t^n and height z_i. Then, we have:

$$\tilde{\boldsymbol{N}}_i^n = M_i^n(\tilde{\mathbf{x}}, \tilde{\mathbf{w}}). \tag{10}$$

With $\tilde{\mathbf{x}}$ the 3 DSD parameters field as defined above (see Sect. 2.1.3). Consequently, the cost function J_o can be expressed as a function of the unknowns $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{w}}$ only:

$$J_\mathrm{o}(\tilde{\mathbf{x}}, \tilde{\mathbf{w}}) = \frac{1}{2} \sum_{n \in [1:\mathrm{NT}]} \sum_{i \in [1:\mathrm{NZ}]} \left\| \boldsymbol{y}_i^n - \mathbf{H}_i^n M_i^n(\tilde{\mathbf{x}}, \tilde{\mathbf{w}}) \right\|^2. \tag{11}$$

2.2.3 Observation operator H

The observation operator, $\mathbf{H}$, can be considered as an aggregation of 2 sub-operators $\mathbf{H}_\mathrm{dis}$ and $\mathbf{H}_\mathrm{dop}$, modeling the disdrometer and radar observations, respectively.

For disdrometer measurements, we build an operator $\mathbf{H}_\mathrm{dis}$ able to convert the propagation model ground ($i = \mathrm{NZ}$) discretized DSD $\left(\tilde{\boldsymbol{N}}_\mathrm{NZ}^n\right)_{n \in [1:\mathrm{NT}]}$ into disdrometer like measurements (i.e. drop histograms in our case). A disdrometer records a flux of droplets crossing a given area (A, in m^2). Then, at time t^n and for drops in the diameter class p, this operator has the form:

$$\mathbf{H}_\mathrm{dis}(\tilde{N}_{\mathrm{NZ},p}^n) = \tilde{N}_{\mathrm{NZ},p}^n \cdot A \cdot (v_p + \tilde{w}_\mathrm{NZ}^n)\, \Delta t \Delta D \tag{12}$$

For radar measurements, we build an operator $\mathbf{H}_\mathrm{dop}$ converting model DSD into Doppler radar like measurements. Radars operating in a vertically pointing mode provide the distribution of the radar return power as a function of the Doppler velocity, the so-called Doppler spectrum (Giangrande et al., 2010). The retrieval of the DSD directly from these Doppler spectra is an inverse problem (Kollias et al., 2002) mixing various phenomena. Here we build a "direct" operator which acts in reverse order (from DSD to Doppler spectra). It takes into account vertical air motion and rain attenuation. At a given time and location, this operator can be computed according to the 3 following steps:

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

1. Calculate the theoretical non attenuated and non noisy spectra $F(v)$ ($\mathrm{s\,m^{-2}}$) from a DSD $N(D)$:

$$F(v) = \sigma(D)N(D)\frac{\mathrm{d}D}{\mathrm{d}v}(v), \tag{13}$$

where, given a drop diameter D, $\sigma(D)$ is the backscattering cross-section (unit $\mathrm{m^2}$), calculated for the radar frequency according to the Mie scattering theory (Ulaby et al., 1982). $\mathrm{d}D/\mathrm{d}v$ is the derivative of the diameter-velocity relation.

2. Then deduce the final attenuated spectra $F_\mathsf{a}(v)$, encompassing both the effects of the attenuation and of the vertical air motion w.

$$F_\mathsf{a}(v) = F(v+w)e^{-2\int_0^z K(r)\mathrm{d}r}. \tag{14}$$

The exponential part of Eq. (14) adds the effect of rain attenuation (see for instance Peters et al., 2005), with $K(z)$ the attenuation coefficient at height z (unit $\mathrm{m^{-1}}$), calculated through: $K(z) = \int_D \sigma_\mathsf{ext}(D)N(D,z)\mathrm{d}D$, with $\sigma_\mathsf{ext}(D)$ the extinction cross-section (Mie theory).

3. Finally, for minimization purposes, F_a is converted in a logarithmic scale, resulting in a logarithmic bounded attenuated spectra $Z(v)$:

$$Z(v) = ln\left(10^{10}F_\mathsf{a}(v) + 1\right). \tag{15}$$

We recall that the operator run on the computer is necessarily a discretized version of what is presented here.

2.2.4 Regularization and penalization terms

In this section, we present a brief description of the penalization term J_r as well as the two regularization terms J_x and J_w of the cost function (see Eq. 8):

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

1. J_r: we have seen in Sect. 2.1.3 that the time step of the model is set to 5 s. It means that $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{w}}$ have to be retrieved at this resolution. However, the minimal resolution of the instruments is generally superior to this value (10 s in our case, see Sects. 3.1 and 3.2). Then, J_r avoids the assimilation algorithm producing high frequency fluctuations on both of the unknowns $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{w}}$ by giving lower cost to smooth fields.

2. J_x allows to proceed to the minimization under constraints for $\tilde{\mathbf{x}}$, the three DSD parameters penalization term. Because the instruments used are not able to record the complete drops diameter range, it is used to avoid the assimilation algorithm producing unphysical DSD. To this end, we add this penalization term which avoids the components of the $\tilde{\mathbf{x}}$ DSD parameters field to go out of predefined ranges: [0; 8000] m^{-3} for α; [0; 3] for k; [0; 10^{-3}] m for θ.

3. J_w: the vertical wind field $\tilde{\mathbf{w}}$ modifies the advection velocities and shifts the Doppler spectra. Then, the vertical wind is mainly controlled (in the assimilation algorithm) by the Doppler spectra. If no spectra are available for a given vertical layer, the assimilation algorithm could produce unphysical temporal fluctuations of the vertical wind which would not imply significant extra cost without the term J_w. Since there may be no Doppler spectra available at given heights (see Sect. 3.1), we use J_w, the wind penalization term, to force the vertical wind to stay close to 0 by adding extra-cost to strong vertical winds.

2.2.5 Minimization of the cost function and parameter estimation

The numerical minimization of the cost function $J(\tilde{\mathbf{x}}, \tilde{\mathbf{w}})$ goes through the calculation of its gradient. Note that all the regularization terms are easily differentiable. The gradient of the observation term (J_o) of the cost function is:

$$\nabla J_\mathsf{o}(\tilde{\mathbf{x}}, \tilde{\mathbf{w}}) = \sum_{n \in [1:\mathsf{NT}]} \sum_{i \in [1:\mathsf{NZ}]} {}^t\mathbf{M}_i^{nt} \mathbf{H}_i^n \left(\boldsymbol{y}_i^n - \mathbf{H}_i^n M_i^n(\tilde{\mathbf{x}}, \tilde{\mathbf{w}}) \right), \tag{16}$$

with ${}^t\mathbf{M}_i^n$ the adjoint of the linearized version of the operator M_i^n.

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

We use the YAO software (Thiria et al., 2006; Nardi et al., 2009), developed by the LOCEAN (Laboratoire d'Océanographie et du Climat). It provides a simple method for deriving the adjoint of the model. For the numerical minimization itself, it is coupled with M1QN3 (Gilbert and Lemaréchal, 2006), a quasi-Newton algorithm to solve unconstrained optimization problems. Once the propagation model implemented, it allows the use of observations to compute realistic estimations for the unknowns.

We finally note that we have to initialize the unknowns of the problem $(\tilde{\mathbf{x}}, \tilde{\mathbf{w}})$ at the very beginning of the assimilation process. We found that this initialization is not critical. Consequently, we will initialize the wind field $(\tilde{\mathbf{w}})$ at 0 and the 3 gamma parameters time series $(\tilde{\mathbf{x}})$ at constant very low values, namely: $\alpha = 1$, $k = 0.8$, $\theta = 2.10^{-4}$.

3 Measuring instruments and case study

In this section, we firstly give the characteristics of the two instruments used in this study, namely a Doppler 24 GHz radar, and a disdrometer. Then, we present the characteristics of the case study treated.

3.1 Micro-Rain Radar (MRR)

The first instrument used in this study is a micro-rain radar (MRR, Peters et al., 2005), developed by METEK GmbH, Germany, and belonging to Météo France. The MRR is a 24 GHz frequency modulated, continuous wave (FM-CW) vertically pointing Doppler radar, with a small transmit power (50 mW) and a beamwidth of 2°. The raw return power spectrum is processed by the MRR, see METEK (1998), which supplies the spectral reflectivities. In our assimilation process, only the spectral reflectivities will be used as observations (noted F_{MRR}) .

The MRR software also supplies DSD estimations (obtained through an inversion algorithm of the Doppler spectra, including attenuation correction) and estimations of different moments of the DSD (liquid water content, rain rate, reflectivity factor). Those quantities,

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

mentioned as “MRR products”, will be used as a comparison framework for the DSD $\tilde{\boldsymbol{N}}$ retrieved by our algorithm.

The radar gates have a $100\,\mathrm{m}$ length width. The MRR provides data up to a $3100\,\mathrm{m}$ height. The minimal temporal resolution is $10\,\mathrm{s}$. This fine temporal resolution will allow us to integrate the observations over various time periods. This integration time will be referred later as “observation window”. The Doppler velocities used range from 0.56 up to $9.54\,\mathrm{m\,s^{-1}}$, with a step of $0.19\,\mathrm{m\,s^{-1}}$. According to the manufacturer recommendations (METEK, 1998), the MRR records for the 2 first radar gates ($< 200\,\mathrm{m}$) will not be used in the assimilation process.

3.2 Dual-beam spectropluviometer (DBS)

The DBS is an optical disdrometer developed by the LATMOS, whose main advantage is its ability to cover a very large range of raindrop diameters (Delahaye et al., 2006). Its collecting area is $0.01\,\mathrm{m}^2$. For each recorded drop, the DBS supplies a triplet (time of arrival/spherical equivalent diameter/fall velocity), allowing the measurement of drop flux time series, noted N_{DBS}, used as observations in the assimilation algorithm. We apply a minimum diameter threshold of $0.4\,\mathrm{mm}$. $0.4\,\mathrm{mm}$ corresponds to a value over which we are sure that the instrument avoids false detections (Delahaye et al. (2006)) and can be compared to other instruments (Krajewski et al. (2006)).We will only conserve the drops above this diameter.

3.3 Case study data set

The two instruments described above have been deployed in Ardèche (south-west of France), during the autumn 2013, in the context of the HyMeX campaign (HYdrological cycle in the Mediterranean EXperiment, see for instance Bousquet et al., 2014). Figure 2 presents a topographic map of the area. The two instruments (DBS and MRR) were colocalized at Le Pradel (44.6° N, 4.5° E), in a mountainous area called the Cévennes-Vivarais, subject to so-called Cévenol rain events (Molinié et al., 2012). However, for the purpose

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

described in Sect. 2.1.2, we will not work on intense rain events, but rather on a stratiform event of light rain which occurred on 12 October 2013.

This event occurred at Le Pradel in the afternoon, from 16:30 to 21:30 UTC. It is a long rain event characterized by low but sustained rain rates at the ground level (the disdrometer was used to record 2 min rain rates always between 1 and 6 mm h^{-1}, see Fig. 3, red line). The event is also quite homogeneous. All the Météo France automatic weather stations located around Le Pradel (see Fig. 2) record consistent cumulative rainfall: 11.5 mm at Le Pradel, 12 mm at Aubenas, and 12.9 mm at Berzème. From the MRR (Fig. 4, left) and DBS (Fig. 3, red) measurements, we can distinguish two distinct phases in the event. Until 18:15, there is very light rain, with only small drops (rain rates around 2 mm h^{-1} and very low reflectivities and characteristic velocities). After 18:15, there are short periods with higher rain rates (4–5 mm h^{-1}) and, mainly, higher reflectivities and falling velocities (up to 35 dBZ and 8 m s^{-1}).

The temperature was quite low: around 8.5 °C at the ground level (Le Pradel station, 278 m height). Assuming a gradient of $-1\,°C/150\,m$, we found a freezing level at 1550 m, which is consistent with the MRR records (on Fig. 4, left, we see the bright band around 1500 m). As a consequence, the top of the retrieval domain in our assimilation algorithm will be set to 1400 m in order to avoid bright band problems (the propagation model is not able to treat the melting layer). Note that the blank strip on the right of Fig. 4 is explained by this choice. The 10 m horizontal southerly wind recorded by Météo France at Berzème station is light during the event, ranging from 1.0 to 2.3 m s^{-1}. The humidity rates are always close to 100 % (97 % at Berzème station, 99 % at Aubenas', see Fig. 2). This rules in favor of the choice to neglect horizontal wind and evaporation effect in the propagation model (Sect. 2.1.2).

For this case study, the number of time and height discretization steps (NT and NZ, see Sect. 2.1.3) is set to 4900 and 15, respectively. We remind that, for numerical purposes, the time and height steps were chosen to be $\Delta t = 10s$ and $\Delta z = 100m$.

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

4 Simulated data (Twin experiment)

We have seen in the last section that the case study chosen to apply our retrieval algorithm is consistent with the propagation model underlying hypothesis. (It assumes DSD driven by gravity and vertical wind alone. Thus the effects of evaporation, collisions and horizontal wind are discarded). However, other phenomena remain unaccounted for in the model. Since this could impair the ability of the model to satisfactorily manage our real case study, we have to assess the impact of two of these phenomena (instrumental noise and turbulence) on the assimilation retrievals. This is done, in this section, by performing a twin experiment, whose principle is described below.

4.1 Description of the twin experiment process

Twin experiments consist in applying the assimilation process to simulated data with properties analogous to real data.

Figure 5 schematically summarizes all the successive steps of our twin experiment (i.e. to simulate realistic observations and to run the assimilation). Firstly, we simulate realistic series of top DSD parameters and vertical winds (step a on Fig. 5). These parameters are what we would like to retrieve with the algorithm. In Sect. 4.2, we will explain how these fields are simulated. Then, these series are propagated through space and time using the propagation model (step b) and "true" (namely, by supposing the model perfect) DBS and MRR data are produced through the use of the observation operators (step c). All the fields obtained through this process (DSD and vertical wind, as well as observations like Doppler spectra) are mentioned below as "true".

Then, from these fields, we simulate realistic MRR and DBS observations by adding model and instrumental errors (instrumental noise and turbulence, see Sect. 4.3), which cannot be directly produced by the propagation model (step d) since they are not present in the model despite their effects on the measurements. These data are mentioned below as "observed". The process used to simulate these observed fields is detailed in Sect. 4.3.

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

Finally, the assimilation algorithm estimates an optimal set of unknowns explaining these observations (step e). Assimilated $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{w}}$ fields are then propagated by the model (step f) and the observation operators (step g). All these retrieved fields are mentioned below as "assimilated". They will be compared to true and observed fields. The results of this data assimilation twin experiment are detailed in Sect. 4.4.

This twin experiment demonstrates the ability of the assimilation algorithm to produce coherent DSD fields. It also shows how it indirectly handles the turbulence, underlining its efficiency despite the presence of non modeled phenomena. Note that in this section, all the observations (MRR spectra F_{MRR} and DBS DSD N_{DBS}) are integrated over a 2 min observation window (if not, it will be explicitly specified).

4.2 Simulation of top DSD parameters ($\tilde{\mathbf{x}}$) and vertical wind field ($\tilde{\mathbf{w}}$)

First, we have to simulate realistic time series of top DSD parameters ($\tilde{\mathbf{x}}$). To do this, we use the rain rates recorded on the ground by the DBS during the case study presented above (Sect. 3.3). From each of the successive disdrometer rain rates, we calculate the 2 parameters of an exponential DSD according to Marshall and Palmer (1948). This time series of Marshall–Palmer DSD (special case of gamma DSD) is our simulated top DSD $(\tilde{\boldsymbol{N}}_0^n)_{n\in[\![1:\mathrm{NT}]\!]}$, which is used as input in the propagation model.

We also need a time-height field of vertical winds $\tilde{\mathbf{w}}$, (resolution 5 s, 100 m) to run the propagation model and simulate observations. We have seen that our algorithm is, up to now, limited to stratiform quiet rain events. For this kind of events, vertical winds range in $\pm 2\,\mathrm{m\,s^{-1}}$, with temporal correlations around 2–5 min and spatial correlations around 500–1000 m (Giangrande et al., 2010). Here is the process used to simulate such a field. (1) We create a $5\,\mathrm{s}/100\,\mathrm{m}$ field of independent, normally distributed, random variables. (2) We average this field over a mobile $(2\,\mathrm{min}/500\,\mathrm{m})$ window. (3) We scale the field to get winds within $\pm 2\,\mathrm{m\,s^{-1}}$. In the end, we obtain a wind field with triangular correlations corresponding to the expected values. Figure 7 (top, on the left) shows a part of the vertical wind field used in the twin experiment (i.e., on simulated data). Note that the winds are positive downward.

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

4.3 Simulation of realistic observations

We detail in this section the process used to simulate realistic observations (addition of turbulence effects and instrumental noise). The purpose of this step is to build observations consistent with what is expected from real measurements.

4.3.1 Radar Doppler spectra

We want to simulate Doppler spectra as they would be recorded by a MRR in realistic conditions. For this purpose, we run the propagation model and save the attenuated reflectivities calculated by the algorithm (observation operator) from the DSD (variable $F_{\mathsf{a}}(v)$, see Eq. 14). Then, for each spectrum, we apply the following process. (1) Addition of the atmospheric turbulence impact. Turbulence can be seen as a fine scale modulation of the vertical wind. For a time and location, the effective vertical air motion is the sum of the mean vertical wind (defined above and taken into account in the propagation model) and of a random variable, modeling the turbulence (not taken into account). Mathematically, the impact of turbulence on the Doppler spectra can be modeled as the convolution of the Doppler spectra with a gaussian function (Gossard, 1994; Williams, 2002; Tridon et al., 2013). So, the turbulence is added to our spectra by applying a discretized version of the convolution operator:

$$F_t(v) = F_{\mathsf{a}}(v) * \frac{1}{\sqrt{2\pi}\sigma_t} e^{-\frac{v^2}{2\sigma_t^2}}, \tag{17}$$

with $*$ the convolution operator and σ_t the width of the turbulence spectrum. We use $\sigma_t = 0.7\,\mathrm{m\,s^{-1}}$, considered a realistic but large value (Peters et al., 2005; Tridon et al., 2013) implying the use of a large spectral broadening (Williams and Gage, 2009). (2) Addition of instrumental gaussian white noise (intensity 1 dB) to F_t. On Fig. 6, on the right, the dotted lines show observed Doppler spectra at different heights.

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

4.3.2 Disdrometer

We simulate the records of a ground disdrometer. For this purpose, we run the propagation model and save the discretized DSD in the grid box just above the ground. For a given time and diameter class p, we note n_p the corresponding number of drops in a given box. Then, for each of these n_p drops, we choose a diameter (following a uniform law in the interval $[p;\ p+1]\Delta D + D_0$) and a velocity. This velocity is the theoretical Atlas velocity for the diameter modulated by a gaussian white noise (as presented previously) representing the atmospheric turbulence. The height of the drop is finally uniformly drawn in the interval [0; 100] m (height range of the grid box).

In the end, we check which drops, among the n_p, will touch the ground during the time interval Δt (and so will be seen by the disdrometer). The drops touching the ground are recorded as observed by the disdrometer. This process is a way to mimic the natural variability of rain and to run in a stochastic way the observation operator defined by Eq. (12).

4.4 Results – robustness of the algorithm

In this section, we firstly assess the ability of the assimilation process to reproduce the observations. Then, we compare the true and assimilated DSD $\tilde{N}$ and wind fields $\tilde{\mathbf{w}}$. Finally, we investigate how the model response is impacted by the size of the observations window. First of all, we have to introduce the numerical indicators needed after to perform this investigation.

4.4.1 Error Indicators

In this section, we define the indicators used to quantitatively evaluate the performance of the algorithm. Choosing error indicators is rather sensitive to the range of values under study. For fields with magnitude far from 0 and ranging in small intervals, we use relative indicators without any risk of giving too much importance to very low (close to 0) data. We define the two following indicators: the Mean Absolute Percentage Error (MAPE), and the

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

relative bias (rbias), defined by:

$$\text{MAPE}(\%) = 100 \cdot \frac{1}{P} \sum_{i=1}^{P} \frac{|A_i - B_i|}{|A_i|}, \tag{18}$$

$$\text{rbias}(\%) = 100 \cdot \frac{\sum_{i=1}^{P} (A_i - B_i)}{\sum_{i=1}^{P} A_i}, \tag{19}$$

where $\boldsymbol{A}$ and $\boldsymbol{B}$ are the respective reference and evaluated fields rearranged in vector forms. P stands for the dimension of the vectors.

Fields having very low or negative values render the 2 previous indicators inappropriate. For these fields, we use the Mean Absolute Error (MAE), defined by:

$$\text{MAE} = \frac{1}{P} \sum_{i=1}^{P} |A_i - B_i|. \tag{20}$$

4.4.2 Observations reproduction

In the following analysis, we use the denominations “true”, “observed” and “assimilated” (see Sect. 4.1 and Fig. 5). Table 1 gives the MAPE and the relative biases between true, observed and assimilated characteristic velocities and spectral widths.

We recall (Sect. 4.3.1) that the atmospheric turbulence, which broadens the Doppler spectra, is included in the observations but not in the true data. Consequently, for the observed spectra, we obtain spectral widths larger ($+12\,\%$) than for the true ones.

The assimilated spectra are also broadened compared to the true ones. But to a lesser extent (spectral width $9.5\,\%$ larger). They are slightly less broadened than the observed spectra ($-2.3\,\%$). We also note that the MAPE and rbias have almost the same absolute values (Table 1, 3rd and 4th rows). This means that the sign of the differences between the 2 compared fields are the same for almost all the grid points. This implies that all spectra are broadened by turbulence independently of time and height.

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

On the contrary, the characteristic velocity is not highly affected by the addition of turbulence impact on observed fields (bias $-0.36\,\%$, MAPE $0.40\,\%$). Consequently, the characteristic velocities are correctly reproduced by the algorithm (biases and MAPE inferior to $2\,\%$). Obtaining the result that turbulent broadening only impacts the spectral width and not the reflectivity-weighted mean downward velocity in the simulations confirms the robustness of the assimilation procedure and are expected results from prior work (Williams, 2002; Williams and Gage, 2009).

To reproduce the turbulence broadening, the algorithm can act on the 2 unknown fields $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{w}}$. Firstly, it can affect the DSD shape, by adding small and large droplets and removing drops of intermediate diameters. But because the DSD is directly controlled on the ground by the disdrometer (observation N_{DBS}), it forces the algorithm to well reproduce the ground DSD $(\tilde{\boldsymbol{N}}^n_{\mathrm{NZ}})_{n\in[\![1:\mathrm{NT}]\!]}$. Between true and assimilated disdrometer observations, we find a MAE (see Table 1, last row) of $27\,\mathrm{mm}^{-1}$, while it is $42\,\mathrm{mm}^{-1}$ between truth and observations (the mean number of drops being $689\,\mathrm{mm}^{-1}$). Secondly, and because this way of acting (modification of DSD shapes) is too costly, the algorithm uses another way, which consists in using the fast modulations of the vertical wind $\tilde{\mathbf{w}}$ at fine resolution. We have seen that all the observations are integrated over a $2\,\mathrm{min}$ observation window (Sect. 3.1), while the wind is retrieved at the $5\,\mathrm{s}$ model resolution. Consequently, if the $5\,\mathrm{s}$ winds oscillate inside the $2\,\mathrm{min}$ observation window, it will result in shifting the $5\,\mathrm{s}$ corresponding spectra alternatively towards small and large Doppler velocities. Once averaged over the $2\,\mathrm{min}$ time period, these shifted spectra will result in a final $2\,\mathrm{min}$ spectrum with a spectral width larger than all the $5\,\mathrm{s}$ spectral widths, whatever the original DSDs. This is the way used by the algorithm to mimic the turbulence. This phenomenon can be seen in the 3rd column of Table 2, which shows the MAE between true and assimilated wind fields for an observation window of $2\,\mathrm{min}$ and for various field resolutions (spreading from $10\,\mathrm{s}$ to $8\,\mathrm{min}$). We can see that the MAE of the $10\,\mathrm{s}$ resolution wind fields is large ($0.46\,\mathrm{m\,s}^{-1}$) compared to the true absolute mean value ($0.34\,\mathrm{m\,s}^{-1}$, see last row of Table 2) but it sharply decreases for resolutions greater than $2\,\mathrm{min}$. This means that there are fluctuations at the $5\,\mathrm{s}$ model resolution which do not correspond to the physics. They are numerical artifacts produced by the model

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

which uses any degree of freedom available to reproduce the broadening observed in the measurements.

An example of Doppler spectra at a $2\,\mathrm{min}$ resolution is given in the right part of Fig. 6. In this plot, examples of vertical profiles of Doppler spectra are shown. On the observed and assimilated spectra, we can clearly see the turbulence broadening, absent in true ones. Since this last result (broadening on assimilated spectra) is due to a turbulence reproduction (using the wind resolution) and not to a modification of the DSD, the good behavior of the assimilation algorithm for realistic stratiform rain is thus assessed.

4.4.3 Wind and DSD reproduction

Even if we confirmed the ability of the assimilation algorithm to handle the turbulence, we still have to assess some other rain characteristic quantities.

As in the previous section, we compare in Table 2 the MAE between true and assimilated fields for an observation window of $2\,\mathrm{min}$ and for various resolutions. The quantities Z_e and D_m (moments of the DSD, see caption of Table 2) are very well reproduced by the assimilation algorithm, with MAE very small compared to the true absolute mean values (for instance, at $2\,\mathrm{min}$ field resolution, MAE of $0.80\,\mathrm{dB}$ and $24.7\,\mu\mathrm{m}$ compared to true absolute mean values of $27.6\,\mathrm{dBZ}$ and $1.12\,\mathrm{mm}$, respectively).

The results are slightly less satisfactory for the LWC (10.0 vs. $142\,\mu\mathrm{g\,m^{-3}}$) and, mostly, for N_0 (241 vs. 1070). This is due to the fact that the number of drops is mainly driven by the number of very small drops ($< 0.4\,\mathrm{mm}$), which are not seen by the disdrometer. Moreover, with their very low backscattering cross-section, they have almost no weight in the cost function. They are consequently not effectively controlled. However, we can note that these drops represent a very low mass of water and that their non control is not critical. (Moreover, since the top DSD is parameterized, we cannot obtain unreasonable values.)

We have seen that the $2\,\mathrm{min}$ MAE for the wind field is quite satisfactory (meaning that the global trends of the wind field have been reproduced by the algorithm). But we would also like to know if the wind patterns are well reproduced. Figure 7 shows extracts of the true and assimilated wind fields, as well as the mean and standard deviations of these

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

fields according to the height (radar gate) and the temporal autocorrelation functions. The top images confirm a satisfactory reproduction of the wind fields. We also see that the correlation time, the absolute mean values and the standard deviations are well reproduced by the assimilation, even if the standard deviation results are more dubious for the two lowest radar gates, where no Doppler spectra are available (as seen in Sect. 3.1).

The rain rates are also satisfactory for resolutions over $2\,\mathrm{min}$ (Table 2, column 4), with a MAE of $0.14\,\mathrm{mm\,h^{-1}}$ with a mean value of $2.18\,\mathrm{mm\,h^{-1}}$.

4.4.4 Impact of the observations window

In this last subsection, we quickly assess the influence of the integration time of the observations (observation window, see Sect. 3.1). Table 2 presents the MAE between different fields depending on the observations window: $10\,\mathrm{s}$ or $2\,\mathrm{min}$. We have seen that a $2\,\mathrm{min}$ observations window allows the algorithm to use the $5\,\mathrm{s}$ model resolution to reproduce the turbulence spectral broadening by adding artificial synthetic fluctuations. If the observation window is reduced to $10\,\mathrm{s}$, no room is left for fluctuation. Consequently, the MAE for the wind are smaller with $10\,\mathrm{s}$ observations window for fine resolutions (for instance, 0.21 vs. $0.46\,\mathrm{m\,s^{-1}}$ at $10\,\mathrm{s}$).

Because the model fits, at best, the $10\,\mathrm{s}$ resolution, it is natural that the algorithm performs less accurately at the $2\,\mathrm{min}$ resolution and hence beyond.

The results are globally the same for the other parameters of Table 2. Because the algorithm cannot use the wind $\tilde{\mathbf{w}}$ to adjust the assimilated spectra to the observed broadened ones, it has no choice but to fit the observations through a change of the top DSD (excess of small and large drops) in order to produce the best compromise at the cost level. This effect is visible on the N_0 column of Table 2: the results are worse for a $10\,\mathrm{s}$ observation window whatever the resolution (excess of small droplets). We can also note that the spectra are much more noisier at $10\,\mathrm{s}$, implying a tougher minimization.

Finally, we conclude that an observation window of $2\,\mathrm{min}$ is a better choice than $10\,\mathrm{s}$. It allows the algorithm to handle the turbulence and limit the instrumental noise influence. Thus the assimilation in the next section will be run with a $2\,\mathrm{min}$ observation window.

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

5 Results for real case study

In this section, we apply the assimilation algorithm to the real data case study described in Sect. 3.3. Again, we show that the assimilation algorithm is efficient (able to fuse different observations) for this case study. We then show that it produces spatio-temporally coherent wind and DSD fields.

5.1 Observation reproduction

We use again the indicators introduced in Sect. 4.4.1 (see Table 3). Firstly, we note that the disdrometer observations N_{DBS} are well reproduced. The MAE between observed and assimilated data is $102\,\text{mm}^{-1}$, while the mean observed number of drops is $1080\,\text{mm}^{-1}$. These results are very similar to the results of the previous simulated study, meaning that the algorithm successfully reproduces the disdrometer observations. Figure 3 shows the ground rain rates as calculated from the disdrometer observations (red) and as reproduced by the assimilation algorithm (blue). It confirms that the algorithm is efficient for ground data. We will later come back to the remaining differences between the 2 rain rate time series.

Table 3 also presents results concerning the MRR observations. Even if the MAPE for the characteristic velocity W_{c} and spectral width SW are larger than the ones on simulated data, they remain satisfactory (mean error of 4.4 and $14\,\%$, respectively).

The 3 columns on the right of Table 3 present these results at different heights and help us to analyze the vertical structure of MRR observations and assimilated data. When considering reflectivity η (0 order moment, second row), we see that there is globally no bias between observed and assimilated reflectivities ($-0.017\,\text{dB}$). However, the assimilation process overestimates the reflectivities at the top ($+0.60\,\text{dB}$) and underestimates the reflectivities close to the ground ($-0.40\,\text{dB}$), while there is no notable bias at middle heights ($-0.12\,\text{dB}$). Since the model propagates the drops from the top to the bottom, without diameter change, besides the attenuation, no phenomenon taken into account can produce a height-decreasing of the mean reflectivity at the event scale. This means that the MRR observations display more reflectivity decreasing with height than what can be explained by the model atten-

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

uation. This phenomenon is confirmed by the left part of Fig. 4. We clearly see that the rain rate (top) and reflectivity factor (middle) display a decrease with height for the MRR products. This phenomenon cannot be reproduced by our algorithm (right part of Fig. 4) which more or less (depending on the wind field, see Sect. 5.3 below) conserves these quantities with height. We also see that the algorithm reaches a compromise between the different heights, so that the assimilated reflectivities (and hence the rain rates and other DSD parameters displayed in Fig. 4) are globally (compared to MRR products) slightly overestimated on the top and underestimated on the ground. This would also explain why, on Fig. 3, assimilated rain rate peaks are generally underestimated compared to disdrometer data. Figure 8 shows the bottom DSD (continuous lines) for the MRR products (200–300 m, red), for the DBS measurements (on the ground, in black) as well as the assimilated ones (on the ground, in blue). For the intermediate diameters ([0.7; 2.5] mm, corresponding to the range which mainly impacts the rain rate and the reflectivity factor at low rain rates), the assimilation underestimates the number of drops. This could be explained by the MRR observed spectra at the top, which induce a lower number of drops (Fig. 8, red dotted marked line). The origin of this variation of reflectivity with height for the MRR observations is not clear. It could be a calibration problem (height dependent transfer function TF, see METEK, 1998), as well as the effects of the horizontal wind or of some microphysical processes (coalescence/break-up) assumed negligible in our propagation model. Anyway, the assimilation algorithm appears able to merge the different data available to produce a solution making a compromise between the observations available (low biases and MAE / MAPE for the three first moments of the spectra and for disdrometer data, see Table 3). It produces results embedding the spatio-temporal coherence brought by the propagation model.

We can also note that the very small drops (< 0.4 mm) are loosely constrained and that the corresponding results are then dubious (Fig. 8). The disdrometer does not cover these diameters and the MRR does not receive a lot of energy for the corresponding Doppler velocities, so that the MRR observed spectra seem less coherent. Consequently, the assimilation algorithm produces a very large number of small drops (more than an exponential law would) without any physical justification associated.

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

For large drops ($> 2.5\,\mathrm{mm}$), we note (Fig. 8) that the disdrometer provides observed ground DSD N_{DBS} which are not reproduced by the assimilation, probably because they are not consistent with the MRR observed spectra. However, large drops are very rare for this light rain event, so that the presence or absence of a particular drop can deeply affect the tail of the disdrometer DSD. Moreover, the terminal velocity of these drops is $> 7.3\,\mathrm{m\,s^{-1}}$, so that disdrometer measurements with $2\,\mathrm{min}$ observation window correspond to very large probed volumes (corresponding to a height of around $850\,\mathrm{m}$, crossed in $2\,\mathrm{min}$ for a $7.3\,\mathrm{m\,s^{-1}}$ drop) which absolutely do not match the MRR probed volumes ($100\,\mathrm{m}$ height). Because large drops are rare for this event, we nevertheless chose to work with a $2\,\mathrm{min}$ observation window. Naturally, for convective events, this choice could be questioned, at least for the disdrometer (we have seen in Sect. 4.4.4 that a decrease of the observation window width for the Doppler spectra decreases the ability of this algorithm to handle the turbulence).

5.2 Wind

Figure 9 (up) presents the $2\,\mathrm{min}$ vertical winds as retrieved by the algorithm from 19:30 to 20:30.

We remind that there are no Doppler spectra available for the two lowest gates (0–200 m), so that the wind retrievals for these 2 heights are loosely controlled by the algorithm, since the wind also acts on the vertical advection of drops and so on the disdrometer records.

The structure of the assimilated wind field is consistent with what we could expect for such an event. The correlation characteristic time and height are around $1\,\mathrm{min}$ and $300\,\mathrm{m}$ for the $10\,\mathrm{s}$ wind field (which can be used by the algorithm to handle the turbulence as explained in Sect. 4.4.4) and around $3\,\mathrm{min}$ and $500\,\mathrm{m}$ for the $2\,\mathrm{min}$ field. These results, as well as the wind range ($\pm 1.5\,\mathrm{m\,s^{-1}}$) are consistent with the results of Kollias et al. (2002) or Giangrande et al. (2010), for instance. Even if we cannot validate the retrieved field using independent measurements in this study, we could nevertheless notice that our method produces the same wind field structures whatever the algorithm initialization.

We also have an overall mean wind of $+0.004\,\mathrm{m\,s^{-1}}$, a slightly not notable downward air motion. We get slight downward air motion in the upper layers of the atmosphere (mean

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

wind around $+0.15\,\mathrm{m\,s^{-1}}$ above 700 m) and slight upward air motion closer to the ground (around $-0.2\,\mathrm{m\,s^{-1}}$ between 200 and 500 m). The 2 min standard deviation is quite constant with height and around $0.25\,\mathrm{m\,s^{-1}}$.

5.3 DSD

We have seen (Sect. 5.1) that, on average, the MRR products (rain rate, reflectivity factor) are consistent with our assimilated data. There is nevertheless a vertical decreasing tendency on MRR products, which cannot be reproduced by the algorithm, given that the propagation model produces vertically coherent fields (Fig. 4). However, in this section, we focus on the structure of the assimilated DSD fields, independently of the MRR products.

We have seen that the algorithm conserves well the reflectivity factor during a drop fall. It results in diagonal structures on the time-height plane (see Fig. 4, middle, right). The slope of a diagonal structure reflects the velocity fall. We can see that the reflectivity field is consistent with the characteristic velocity field (bottom, right): the lower the velocity, the lower the slope of the reflectivity diagonal structure.

Now we will also assess if these structures are retrieved for some important moments of the DSD. Figure 9 shows the mean volume diameter D_m, and the number of drops N_0, resulting from the assimilated DSD $\tilde{\boldsymbol{N}}$. We note that the retrieved values for D_m (0.5–1.5 mm) are reasonable for light stratiform rain (see Fig. A4 in Peters et al., 2005). D_m is also quite well conserved with height, but the patterns seem noisier than for the reflectivity. So does it for the rain rate field (Fig. 4). This is a consequence of the vertical wind, which modifies the vertical advection and so the DSD. The vertical wind has a greater impact, relatively, on small drops than on large ones (given their smaller terminal velocity). So the low order moments, mainly driven by the small drops, are more affected than the high order moments. This is particularly clear for N_0. We have seen that the algorithm, underconstrained for these diameters, produces a very large number of very small drops (< 0.4 mm, i.e. with terminal velocities $< 1.5\,\mathrm{m\,s^{-1}}$) (see Fig. 8). Consequently: (1) the vertical consistency of N_0 is low. (2) The slope of the diagonal patterns of the N_0 time-height field is low (because N_0 is driven by small drops, with low terminal velocities). (3) Patches of raindrops are formed

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

by upward winds (see Fig. 9), and these patches disappear as soon as the wind in the lower layers turns downward. For instance, around 19:55, there is an upward wind between 100 and 400 m, which results in a patch of drops between 300 and 500 m. We can also note that even if the absolute values of N_0 are dubious, these patches are probably real, since they are driven by the wind.

Because the DSD for this event are globally exponential (see Fig. 8), it is reasonable to estimate a DSD slope (Fig. 9, bottom). And because this slope is calculated over the entire range of diameters, the retrieved slope field is vertically consistent (less affected by the wind). We also note that the retrieved slopes are consistent with the results of Giangrande et al. (2010) (Fig. 8 in this paper) on a similar case study.

6 Conclusions

We have built a 4-D-VAR data assimilation algorithm for retrieving vertical DSD profiles and vertical wind under the bright band from observations coming from a micro rain radar (MRR) and a co-located disdrometer, and associated with a vertical propagation model. In this paper, we focused on the data assimilation algorithm (the algorithm finely handles measurements of various natures collected at different resolutions). We consequently chose to use a simple propagation model taking only into account gravity and vertical wind. Because of the limitations of the model, the retrieval algorithm is fit to study stratiform, light rain events.

The algorithm has been firstly applied to simulated data in a twin experiment context. On simulated observations, despite the addition of instrumental noise and turbulence effects, we showed that the proposed technique is able to retrieve DSD fields along with vertical wind patterns and intensities thanks to radar Doppler spectra and disdrometer drop fluxes. In particular, the algorithm appears able to handle the turbulence impact on Doppler spectra even if the turbulence is not explicitly implemented in the propagation model.

The good behavior of the algorithm is assessed on a light rain event during which co-localized MRR and optical disdrometer were used in the south of France during the HyMeX

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

campaign. According to this work, the algorithm appears able to make the best use of the data from the two instruments. In particular, we noticed (suspicious) vertical trends on the MRR products which cannot be reproduced by the propagation model due to its conservative nature. However, despite this trend the assimilation algorithm is able to produce a good compromise between the observations at different heights. The vertical retrieved wind field, although not independently validated, is spatially and temporally coherent with what can be expected for a rain event such as the one studied here. The DSD fields also appear spatially coherent and self-consistent. By combining ground and radar data the presented algorithm is therefore able to retrieve the vertical wind field and to improve the DSD profile retrieval.

Independent information on the vertical wind and/or on the DSD are nevertheless necessary to estimate the performances of the algorithm. For this purpose, we plan to use dual-frequency wind profilers (see for instance Williams, 2012, which retrieves vertical wind in rain from 2 profilers). Another possibility would be to use the methods developed by Giangrande et al. (2010) or Tridon and Battaglia (2015), which retrieve vertical winds and DSD from K band and/or W band Doppler spectra. We could also study the impact on the quality of the retrievals of the addition of other frequencies directly in the assimilation process.

Some improvements are needed to provide an algorithm suitable for various weather situations (tropical rain and convective events, for instance). In our model, we discarded horizontal motions. This hypothesis is valid only if there is no significant horizontal wind or at least if the DSD is sufficiently homogeneous in the horizontal plane. It will be necessary to evaluate more precisely the impact of such an hypothesis. Another possibility would consist in taking into account the horizontal wind by extending the model with an additional spatial dimension. In our case, this can be done with the use of the YAO assimilation tool that provides a simple way to do it.

Other processes such as evaporation, coalescence/break-up have to be implemented in the propagation model. This enhanced model would allow to investigate the relative importance of the different physical phenomena (wind, evaporation, collisions, ...), and therefore to determine which are really essential for the assimilation algorithm. It would also help to better parameterize the algorithm through its ability to explain the observations.

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

Finally, the method is well suited for merging observations of different types each with its own temporal and spatial resolution. Indeed, when merging new observations, only the observation operator has to be adapted. For instance we could use more disdrometers located in the vicinity or the reflectivity at another frequency from another radar. Any kind of sensor which observes geophysical parameters playing a role during the fall of raindrops could be also easily introduced in the assimilation algorithm (ground wind, pressure, humidity).

Acknowledgements. This work was supported by the French Programme National de Télédétection Spatiale 759 (PNTS, http://www.insu.cnrs.fr/pnts), grant no. PNTS-2013-01, and by the IPSL (Institut Pierre-Simon Laplace). The authors wish to thank Jean-Marie Donier from the "Centre National de Recherches Météorologiques – Groupe d'étude de l'Atmosphère Météorologique" (CNRM-GAME) for supplying us with the MRR data. We also thank the LOCEAN team for providing us the YAO software.

References

Arakawa, A. and Lamb, V. R.: Computational design of the basic dynamical processes of the UCLA general circulation model, Methods in Computational Physics, 17, 173–265, 1977.

Atlas, D., Srivastava, R., and Sekhon, R. S.: Doppler radar characteristics of precipitation at vertical incidence, Rev. Geophys., 11, 1–35, 1973.

Barros, A. P., Prat, O. P., Shrestha, P., Testik, F. Y., and Bliven, L. F.: Revisiting low and list (1982): evaluation of raindrop collision parameterizations using laboratory observations and modeling, J. Atmos. Sci., 65, 2983–2993, 2008.

Barthes, L. and Mallet, C.: Vertical evolution of raindrop size distribution: impact on the shape of the DSD, Atmos. Res., 119, 13–22, 2013.

Beheng, K.: The evolution of raindrop spectra: a review of microphysical essentials, in: Rainfall: State of the Science, 29–48, 2010.

Bocquet, M.: Introduction aux principes et méthodes de l'assimilation de données en géophysique, Notes de cours de l'ENSTA, 2005.

Bousquet, O., Berne, A., Delanoë, J., Dufournet, Y., Gourley, J., Van-Baelen, J., Augros, C., Besson, L., Boudevillain, B., Caumont, O., Defer, E., Grazioli, J., Jorgensen, D. J., Kirstetter, P.-E., Ribaud, J.-F., Beck, J., Delrieu, G., Ducrocq, V., Scipion, D., Schwarzenboeck, A., and Zwiebel,

J.: Multiple-frequency radar observations collected in southern france during the field phase of the hydrometeorological cycle in the mediterranean experiment (hymex), B. Am. Meteorol. Soc., 96, 267–282, doi:10.1175/BAMS-D-13-00076.1, 2015.

Brazier-Smith, P., Jennings, S., and Latham, J.: The interaction of falling water drops: coalescence, P. R. Soc. A, 326, 393–408, 1972.

Bringi, V., Chandrasekar, V., Hubbert, J., Gorgucci, E., Randeu, W., and Schoenhuber, M.: Raindrop size distribution in different climatic regimes from disdrometer and dual-polarized radar analysis, J. Atmos. Sci., 60, 354–365, 2003.

Brown Jr, P. S.: Analysis and parameterization of the combined coalescence, breakup, and evaporation processes, J. Atmos. Sci., 50, 2940–2951, 1993.

Delahaye, J.-Y., Barthès, L., Golé, P., Lavergnat, J., and Vinson, J.-P.: A dual-beam spectropluviometer concept, J. Hydrol., 328, 110–120, 2006.

Krajewski, W., Kruger, A., Caracciolo, C., Golé, P., Barthès, L., Creutin, J.-D., Delahaye, J.-Y., Nikolopoulos, E., Ogden, F., Vinson, J.-P. : DEVEX-disdrometer evaluation experiment: basic results and implications for hydrologic studies, Advances in Water Resources, 29, 311–325, 2006.

Giangrande, S. E., Luke, E. P., and Kollias, P.: Automated retrievals of precipitation parameters using non-Rayleigh scattering at 95 GHz, J. Atmos. Ocean. Tech., 27, 1490–1503, 2010.

Giangrande, S. E., Luke, E. P., and Kollias, P.: Characterization of vertical velocity and drop size distribution parameters in widespread precipitation at ARM facilities, J. Appl. Meteorol. Clim., 51, 380–391, 2012.

Gilbert, J. C. and Lemaréchal, C.: The Module M1QN3, INRIA Rocquencourt & Rhone-Alpes, available at: https://who.rocq.inria.fr/Jean-Charles.Gilbert/preprint/u01-m1qn3.pdf, 2006.

Gossard, E. E.: Measurement of cloud droplet size spectra by Doppler radar, J. Atmos. Ocean. Tech., 11, 712–726, 1994.

Gunn, R. and Kinzer, G. D.: The terminal velocity of fall for water droplets in stagnant air, J. Meteorol., 6, 243–248, 1949.

Hu, Z.: The role of raindrop coalescence and breakup in rainfall modeling, Atmos. Res., 37, 343–359, 1995.

Hu, Z. and Srivastava, R.: Evolution of raindrop size distribution by coalescence, breakup, and evaporation: theory and observations, J. Atmos. Sci., 52, 1761–1783, 1995.

Joss, J. and Waldvogel, A.: Ein spektrograph für niederschlagstropfen mit automatischer auswertung, Pure Appl. Geophys., 68, 240–246, 1967.

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

Kalnay, E.: Atmospheric Modeling, Data Assimilation, and Predictability, Cambridge University Press, Cambridge, UK, 341 pp., 2003.

Kollias, P., Albrecht, B., and Marks Jr, F.: Raindrop Sorting Induced by Vertical Drafts, Geophys. Res. Lett., 28, 2787–2790, 2001.

Kollias, P., Albrecht, B., and Marks Jr, F.: Why Mie? Accurate observations of vertical air velocities and raindrops using a cloud radar, B. Am. Meteorol. Soc., 83, 1471–1483, 2002.

Kruger, A. and Krajewski, W. F.: Two-dimensional video disdrometer: a description, J. Atmos. Ocean. Tech., 19, 602–617, 2002.

Lhermitte, R. M.: Observation of rain at vertical incidence with a 94 GHz Doppler radar: an insight on Mie scattering, Geophys. Res. Lett., 15, 1125–1128, 1988.

List, R.: A linear radar reflectivity-rainrate relationship for steady tropical rain, J. Atmos. Sci., 45, 3564–3572, 1988.

List, R. and McFarquhar, G. M.: The role of breakup and coalescence in the three-peak equilibrium distribution of raindrops, J. Atmos. Sci., 47, 2274–2292, 1990.

List, R., Donaldson, N., and Stewart, R.: Temporal evolution of drop spectra to collisional equilibrium in steady and pulsating rain, J. Atmos. Sci., 44, 362–372, 1987.

Low, T. B., and List, R.: Collision, coalescence and breakup of raindrops. Part I: Experimentally established coalescence efficiencies and fragment size distributions in breakup, J. Atmos. Sci., 39, 1591–1606, doi:10.1175/1520-0469(1982)039%3C1591:CCABOR%3E2.0.CO%3B2, 1982a.

Low, T. B., and List, R.: Collision, coalescence and breakup of raindrops. Part II: Parameterization of fragment size distributions, J. Atmos. Sci., 39, 1607–1619, 1982b.

Marshall, J. S. and Palmer, W. M. K.: The distribution of raindrops with size, J. Meteorol., 5, 165–166, 1948.

Mcfarquhar, G. M.: Raindrop size distribution and evolution, in: Rainfall: State of the Science, edited by: Testik, F. Y., and Gebremichael, M., American Geophysical Union, 49–60, available at: http://onlinelibrary.wiley.com/doi/10.1029/2010GM000971/summary (last access: 12 November 2015), 2010.

METEK: MMR, Physical Basis, Available from METEK GmbH, Fritz-Straßmann-Straße 4, D-25337 Elmshorn, Germany, 1998.

Molinié, G., Ceresetti, D., Anquetin, S., Creutin, J. D., and Boudevillain, B.: Rainfall regime of a mountainous Mediterranean Region: Statistical analysis at short time steps, J. Appl. Meteorol. Clim., 51, 429–448, 2012.

Nardi, L., Sorror, C., Badran, F., and Thiria, S.: YAO: a software for variational data assimilation using numerical models, in: Computational Science and Its Applications – ICCSA 2009, Springer, Berlin Heidelberg, Germany, 621–636, doi:10.1007/978-3-642-02457-3_53, 2009.

Peters, G., Fischer, B., Münster, H., Clemens, M., and Wagner, A.: Profiles of raindrop size distributions as retrieved by microrain radars, J. Appl. Meteorol., 44, 1930–1949, 2005.

Prat, O. P. and Barros, A. P.: Exploring the use of a column model for the characterization of microphysical processes in warm rain: results from a homogeneous rainshaft model, Adv. Geosci., 10, 145–152, doi:10.5194/adgeo-10-145-2007, 2007a.

Prat, O. P. and Barros, A. P.: A robust numerical solution of the stochastic collection-breakup equation for warm rain, J. Appl. Meteorol. Clim., 46, 1480–1497, 2007b.

Prat, O. P., Barros, A. P., and Testik, F. Y.: On the influence of raindrop collision outcomes on equilibrium drop size distributions, J. Atmos. Sci., 69, 1534–1546, 2012.

Schleiss, M. and Smith, J.: A method to estimate the 3D+ time structure of the raindrop size distribution using radar and disdrometer data, J. Hydrometeorol., 16, 1222–1242, 2015.

Schlottke, J., Straub, W., Beheng, K. D., Gomaa, H., and Weigand, B.: Numerical investigation of collision-induced breakup of raindrops. Part I: Methodology and dependencies on collision energy and eccentricity, J. Atmos. Sci., 67, 557–575, 2010.

Seifert, A.: On the parameterization of evaporation of raindrops as simulated by a one-dimensional rainshaft model, J. Atmos. Sci., 65, 3608–3619, 2008.

Smolarkiewicz, P. K.: A simple positive definite advection scheme with small implicit diffusion, Mon. Weather Rev., 111, 479–486, 1983.

Straub, W., Beheng, K. D., Seifert, A., Schlottke, J., and Weigand, B.: Numerical investigation of collision-induced breakup of raindrops. Part II: Parameterizations of coalescence efficiencies and fragment size distributions, J. Atmos. Sci., 67, 576–588, 2010.

Thiria, S., Badran, F., and Sorror, C.: YAO: Un logiciel pour les modèles numériques et l'assimilation de données, Rapport technique, LOCEAN, Paris, France, 2006.

Tokay, A., Hartmann, P., Battaglia, A., Gage, K. S., Clark, W. L., and Williams, C. R.: A field study of reflectivity and ZR relations using vertically pointing radars and disdrometers, J. Atmos. Ocean. Tech., 26, 1120–1134, 2009.

Tridon, F. and Battaglia, A.: Dual-frequency radar Doppler spectral retrieval of rain drop size distributions and entangled dynamics variables, J. Geophys. Res.-Atmos., 120, 5585–5601, 2015.

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

Tridon, F., Battaglia, A., and Kollias, P.: Disentangling Mie and attenuation effects in rain using a Ka-W dual-wavelength Doppler spectral ratio technique, Geophys. Res. Lett., 40, 5548–5552, 2013.

Ulaby, F. T., Moore, R. K., and Fung, A. K.: Microwave Remote Sensing Active and Passive – Volume II: Radar Remote Sensing and Surface Scattering and Enission Theory, Addison–Wesley, Reading, UK, 1982.

VanZanten, M., Stevens, B., Vali, G., and Lenschow, D.: Observations of drizzle in nocturnal marine stratocumulus, J. Atmos. Sci., 62, 88–106, 2005.

Villermaux, E. and Bossa, B.: Single-drop fragmentation determines size distribution of raindrops, Nat. Phys., 5, 697–702, 2009.

Villermaux, E. and Eloi, F.: The distribution of raindrops speeds, Geophys. Res. Lett., 38, L19805, doi:10.1029/2011GL048863, 2011.

Williams, C. R.: Simultaneous ambient air motion and raindrop size distributions retrieved from UHF vertical incident profiler observations, Radio Science, 37, 2002.

Williams, C. R. and Gage, K. S.: Raindrop size distribution variability estimated using ensemble statistics, Annales geophysicae, 27, 555–567, 2009.

Williams, C. R.: Vertical air motion retrieved from dual-frequency profiler observations, J. Atmos. Ocean. Tech., 29, 1471–1480, 2012.

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

Table 1. For the twin experiment: comparisons between truth, observations and assimilation for the moments of order 1 and 2 of the Doppler spectra (characteristic velocity W_c and spectral width SW, respectively) of the MRR and of the total number of drops seen at successive 2 min observation windows by the DBS. For the MRR data, the MAPE (Eq. 18) and the relative bias (Eq. 19) are presented, as well as the mean value of the true fields. For DBS data, the MAE (Eq. 20) are presented, as well as the mean value of the true field.

			obs-true	assim-obs	assim-true	mean(true)
MRR	characteristic velocity (W_c)	MAPE (%)	0.40	1.6	1.6	$6.2\,\mathrm{m\,s^{-1}}$
		rbias (%)	-0.36	$+0.48$	$+0.12$	
	spectral width (SW)	MAPE (%)	12	3.3	9.7	$1.3\,\mathrm{m\,s^{-1}}$
		rbias (%)	$+12$	-2.3	$+9.5$	
DBS	number of drops ($\mathrm{mm^{-1}}$)	MAE	42	38	27	689

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

Table 2. For the twin experiment: MAE (Eq. 20) between true and assimilated fields for different parameters: vertical wind; rain rate, RR; reflectivity factor, Z_e; mean volume diameter, D_m; liquid water content, LWC; number of drops, N_0. The indicators are presented for 2 observation windows (10 s, in italics, and 2 min) and for different temporal field resolutions (from 10 s to 8 min). The absolute bias is also presented, as well as the absolute values of the means of the true fields. In each case, the best result (lowest MAE) has been bolded if it happened in the case of 2 min observations window.

Field resolution	obs. window	wind ($\mathrm{m\,s^{-1}}$)	RR ($\mathrm{mm\,h^{-1}}$)	Z_e (dBZ)	D_m (µm)	LWC ($\mathrm{\mu g\,m^{-3}}$)	N_0 ($\mathrm{m^{-3}}$)
10 s	10 s	*0.21*	*0.27*	*0.96*	*28.5*	*14.5*	*337*
	2 min	0.46	0.51	1.05	43.0	19.2	**257**
30 s	10 s	*0.17*	*0.24*	*0.94*	*27.4*	*14.2*	*334*
	2 min	0.40	0.34	**0.92**	33.4	14.8	**251**
1 min	10 s	*0.17*	*0.24*	*0.91*	*26.4*	*13.8*	*330*
	2 min	0.29	**0.17**	**0.83**	**26.2**	**10.7**	**246**
2 min	10 s	*0.16*	*0.23*	*0.87*	*25.0*	*13.5*	*323*
	2 min	**0.14**	**0.14**	**0.80**	**24.7**	**10.0**	**241**
4 min	10 s	*0.15*	*0.22*	*0.81*	*23.3*	*13.1*	*311*
	2 min	**0.12**	**0.13**	**0.77**	23.5	**9.54**	**234**
8 min	10 s	*0.14*	*0.22*	*0.74*	*21.8*	*12.8*	*296*
	2 min	**0.098**	**0.12**	**0.72**	22.4	**9.13**	**225**
abs. bias	10 s	*0.10*	*0.20*	*0.69*	*22.6*	*13.0*	*305*
	2 min	**0.054**	**0.087**	0.71	22.7	**8.67**	**212**
true absolute mean		0.34	2.18	27.6	1120	142	1070

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

Table 3. For the real case study: comparison between observations and assimilation for 3 moments of the MRR Doppler spectra: order 0 (reflectivity η); order 1 (characteristic velocity W_c); order 2 (spectral width SW). For η, MAE (Eq. 20) and bias are presented. For W_c and SW, MAPE (Eq. 18), rbias (Eq. 19) and mean value of the observed field are presented. The results are shown for every heights, as well as for 3 different height ranges. Positive bias mean overestimation by the assimilation compared to the observations. Comparisons between observed and assimilated DBS measurements are also presented.

		every gates	top (1100–1400 m)	middle (700–1000 m)	bottom (200–500 m)
η	MAE (dB)	1.2	1.8	0.86	1.1
	bias (dB)	-0.017	$+0.60$	-0.12	-0.40
W	MAPE (%)	4.4	5.7	3.5	4.4
	rbias (%)	-0.39	-1.5	$+0.16$	-0.15
	mean(obs) ($m\,s^{-1}$)	5.11	5.28	5.18	4.87
SW	MAPE (%)	14	14	12	17
	rbias (%)	$+5.9$	$+3.1$	$+3.5$	$+12$
	mean(obs) ($m\,s^{-1}$)	1.20	1.25	1.22	1.11
DBS	MAE (number of drops: mm^{-1})	102			
	mean(obs) (number of drops: mm^{-1})	1080			

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

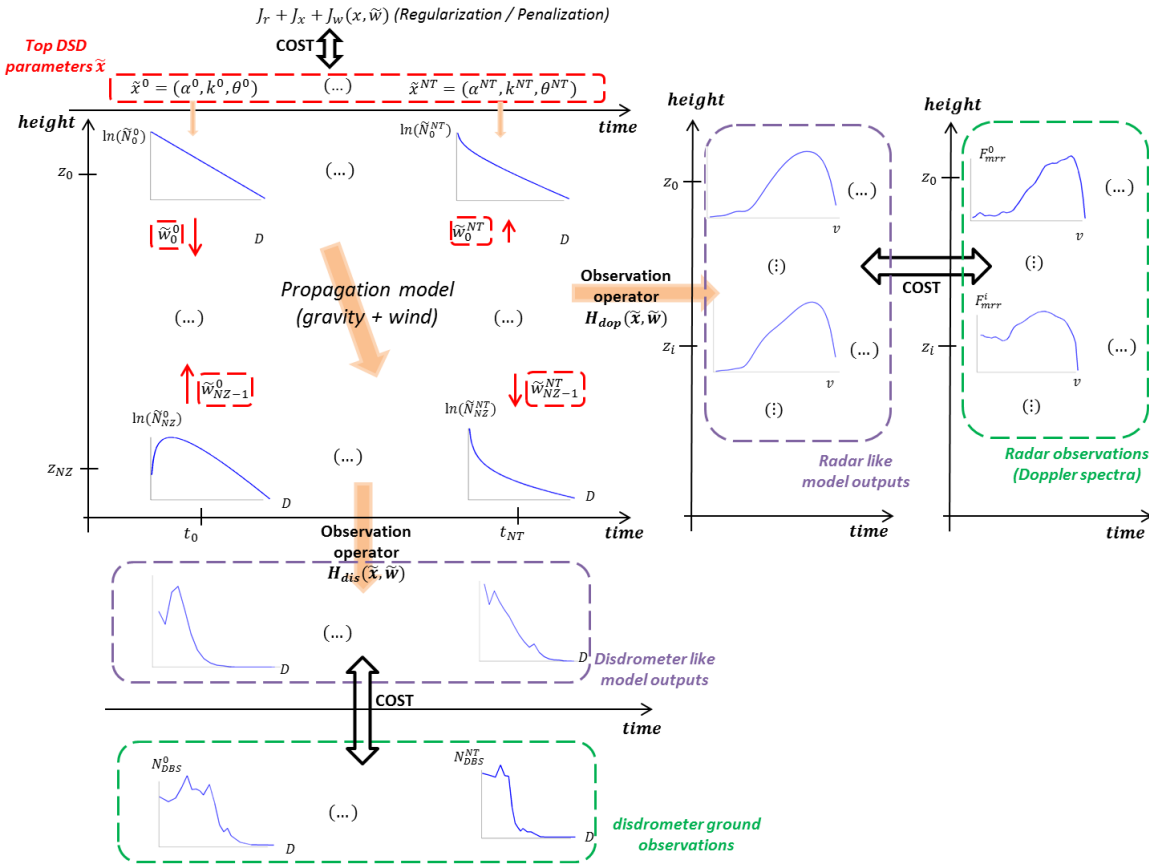

Figure 1. Diagram of the assimilation algorithm. The unknowns of the problem (top DSD parameters and vertical winds) are circled in red. The top DSD parameters are converted into bin DSD and propagated through space and time, under gravity and vertical wind, by the propagation model. Radar and disdrometer observations (circled in green) are available at different places and times. Observation operators are used to convert the model DSD into observations like data (circled in purple). Then, the gap between these data and the observations is the value of the observation term of the cost function. Regularization and penalization terms are added in the cost function. Then, we use a numerical minimizer to estimate a set of unknowns which minimize the cost function. The notations are the ones defined in Sect. 2.

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

Figure 2. Location of the studied area in the south-east of France, on a topographic map. The two instruments used in the assimilation algorithm (MRR and DBS, see Sect. 3.1 and 3.2) are located at Le Pradel, 278 m height. The 3 automatic Météo France weather stations of Le Pradel (same place), Aubenas (7.5 km westward, 180 m height) and Berzème (7.5 km on the north-east, 650 m height) are also shown.

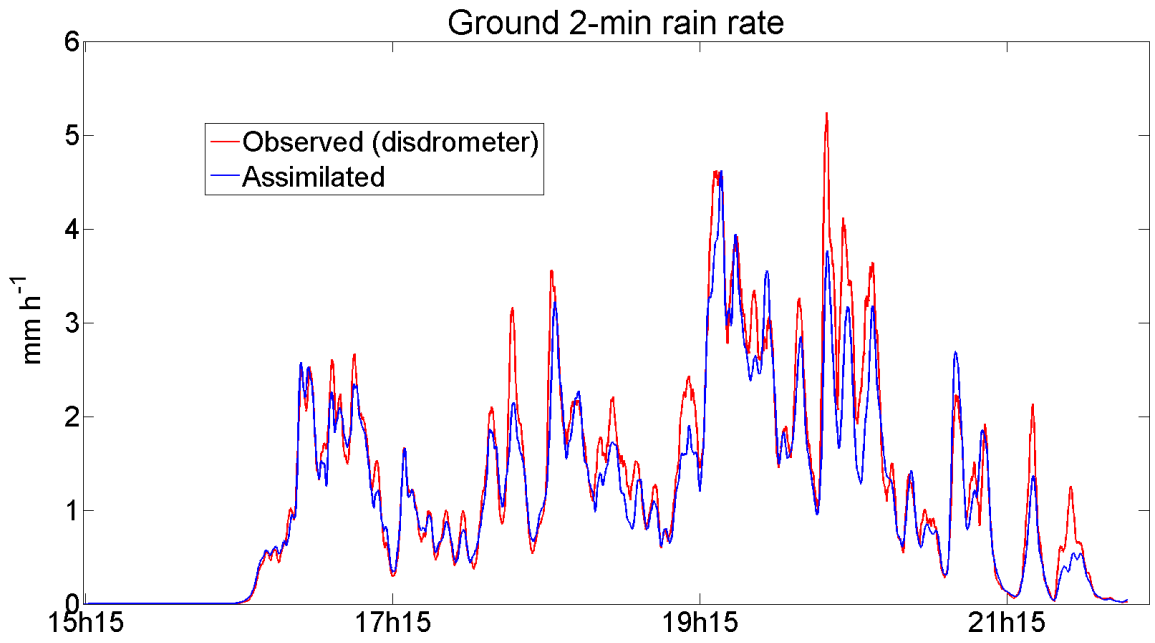

Figure 3. For the real case study of the 12 October 2013, 2 min rain rates as measured by the DBS (red) and as reproduced by the 4-D-VAR assimilation algorithm (blue).

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

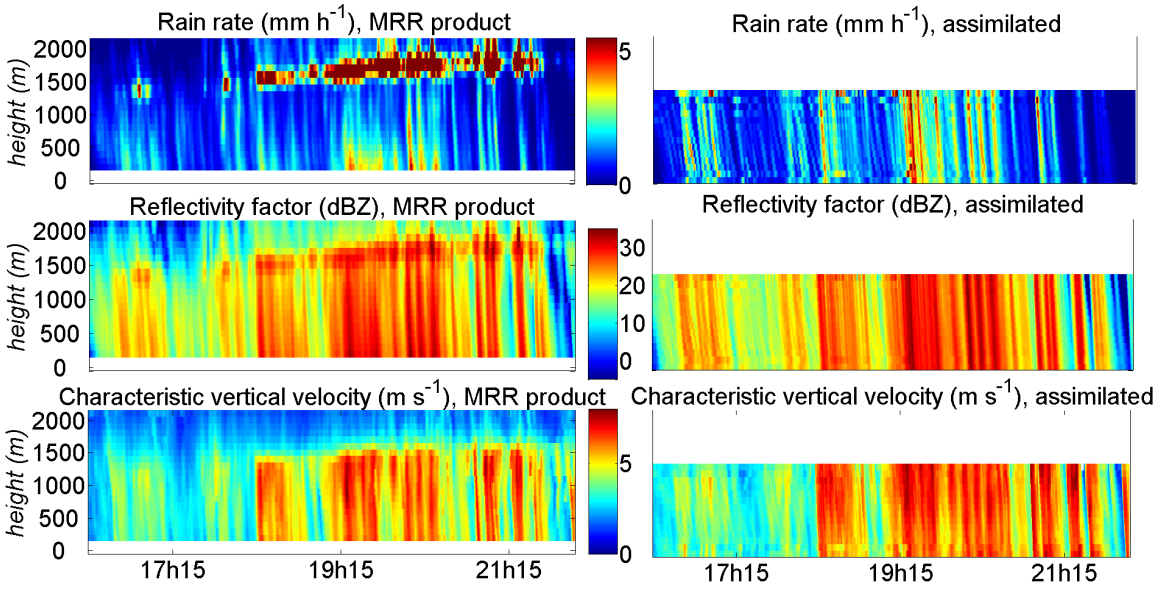

Figure 4. For the real case study: rain rate (up), reflectivity factor (middle) and characteristic velocity (bottom) for MRR products (left) and assimilated fields (right).

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

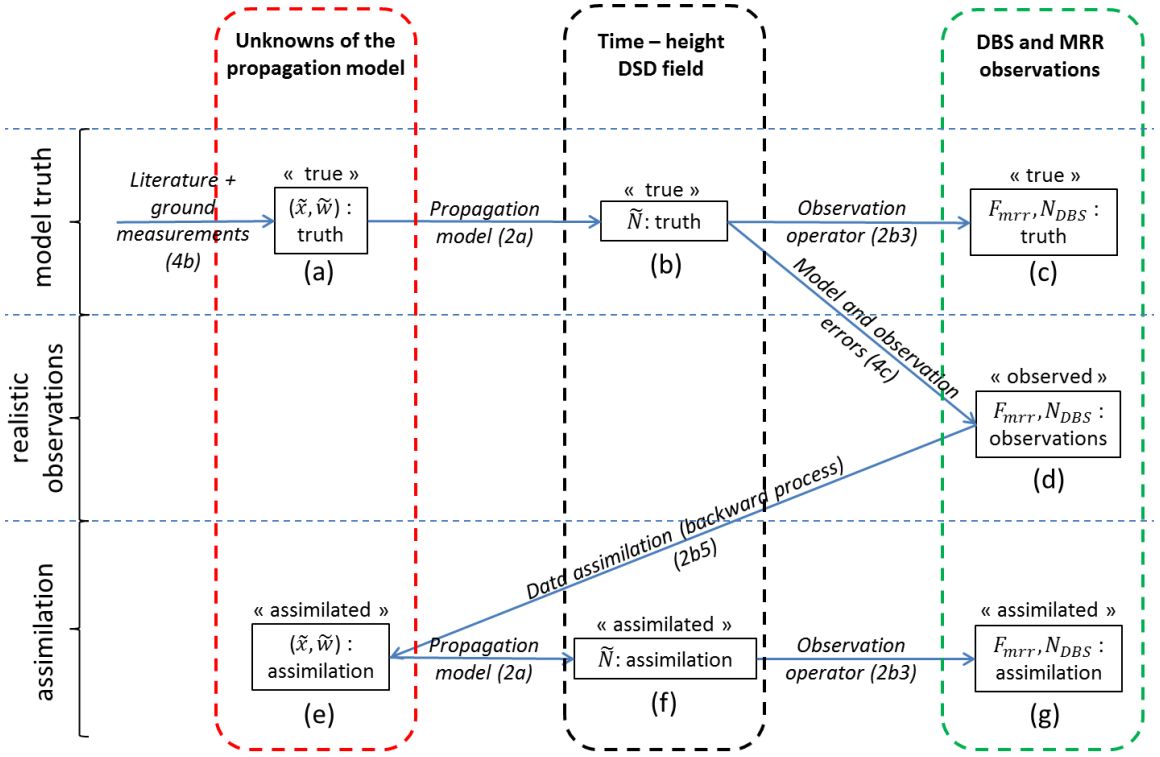

Figure 5. Diagram summarizing the twin experiment process (encompassing the observation simulation). See Sect. 4.1 for the description of the scheme. The sections of the paper in which the different simulation processes are detailed are indicated in brackets below the arrows.

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

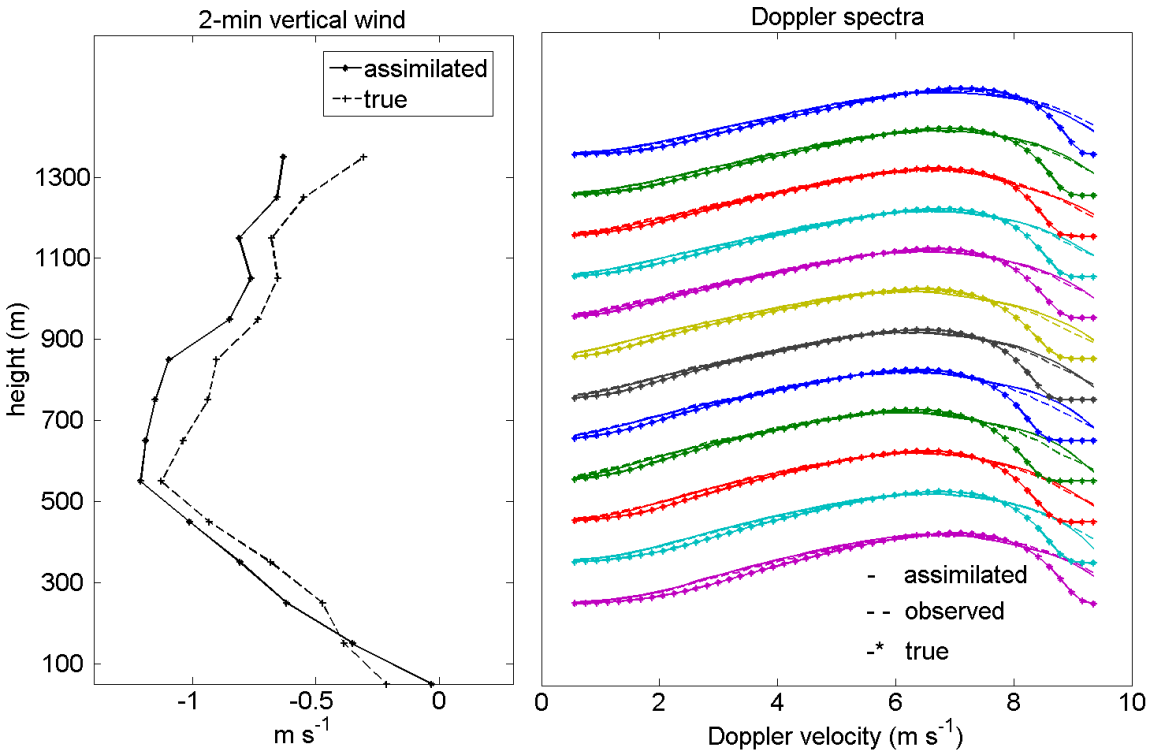

Figure 6. For the twin experiment: vertical wind and Doppler spectra vertical profiles for a given time. (left): vertical wind (positive downward) according to the height. Dashed line is the true wind; solid line is the assimilated wind. (right): Doppler spectra for the successive radar gates from gate 4 (bottom) up to gate 15 (up). Marked lines are the true spectra, dashed lines the observed spectra (with turbulence and instrumental noise) and solid lines the assimilated spectra.

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

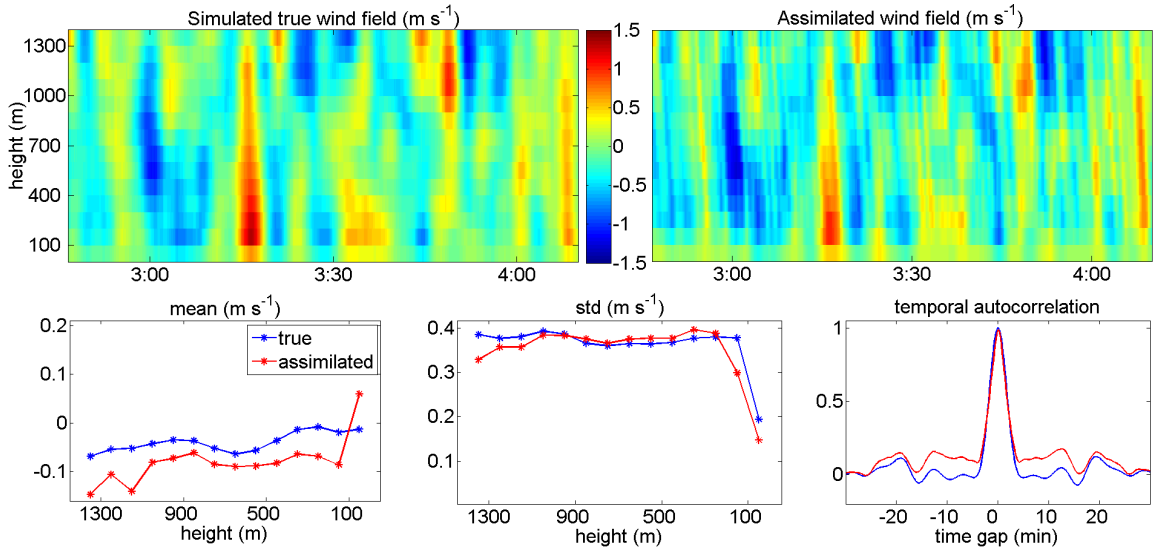

Figure 7. For the twin experiment: ability of the assimilation algorithm to retrieve the wind field characteristics. (up, left): true wind field; (up, right): assimilated wind field; (bottom, left): mean values of these two fields according to the height: blue: truth; red: assimilation; (bottom, middle): the same for the standard deviation; (bottom, right): temporal normalized autocorrelation function of the two wind fields. The wind is positive downward.

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

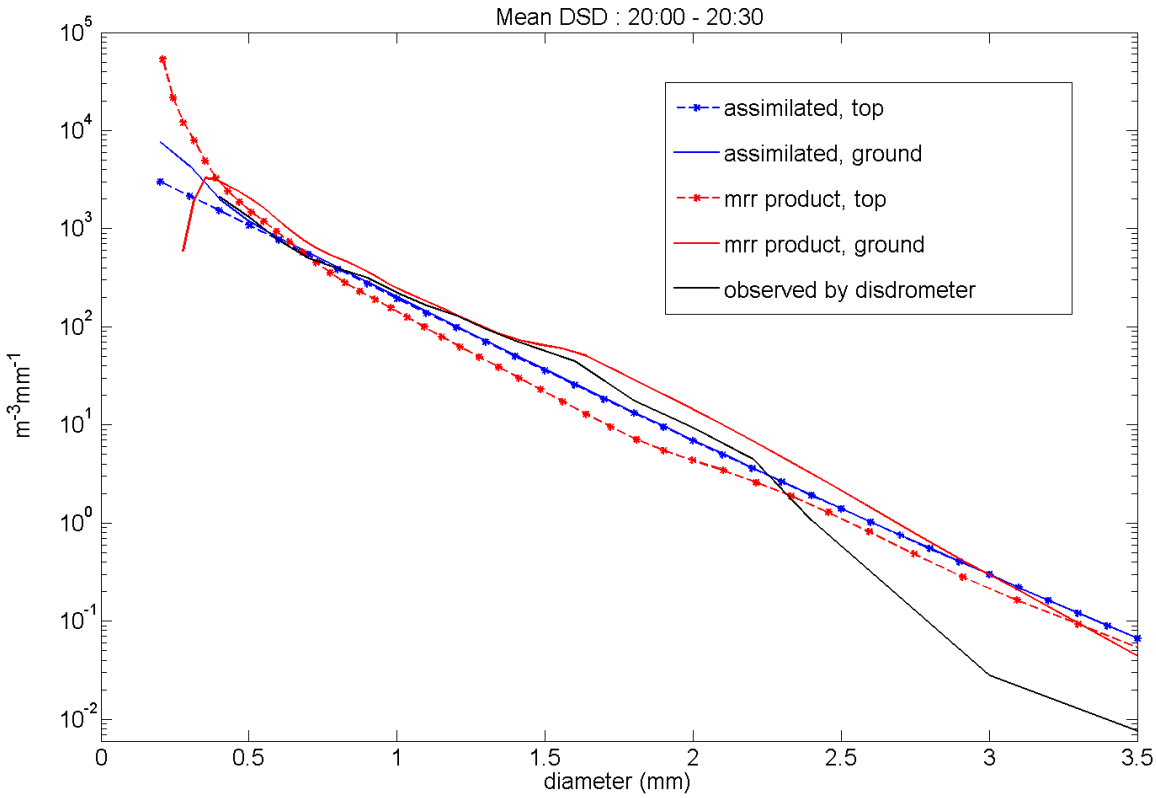

Figure 8. For the real case study: mean DSD between 20:00 and 20:30 evaluated on the ground (continuous lines) from: disdrometer data (black) and assimilation outputs (blue) and from MRR on the lowest radar gate available (the 3rd one) (continuous, red). The same, evaluated on the top (dotted marked lines) from MRR data (red) and assimilation outputs (blue).

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

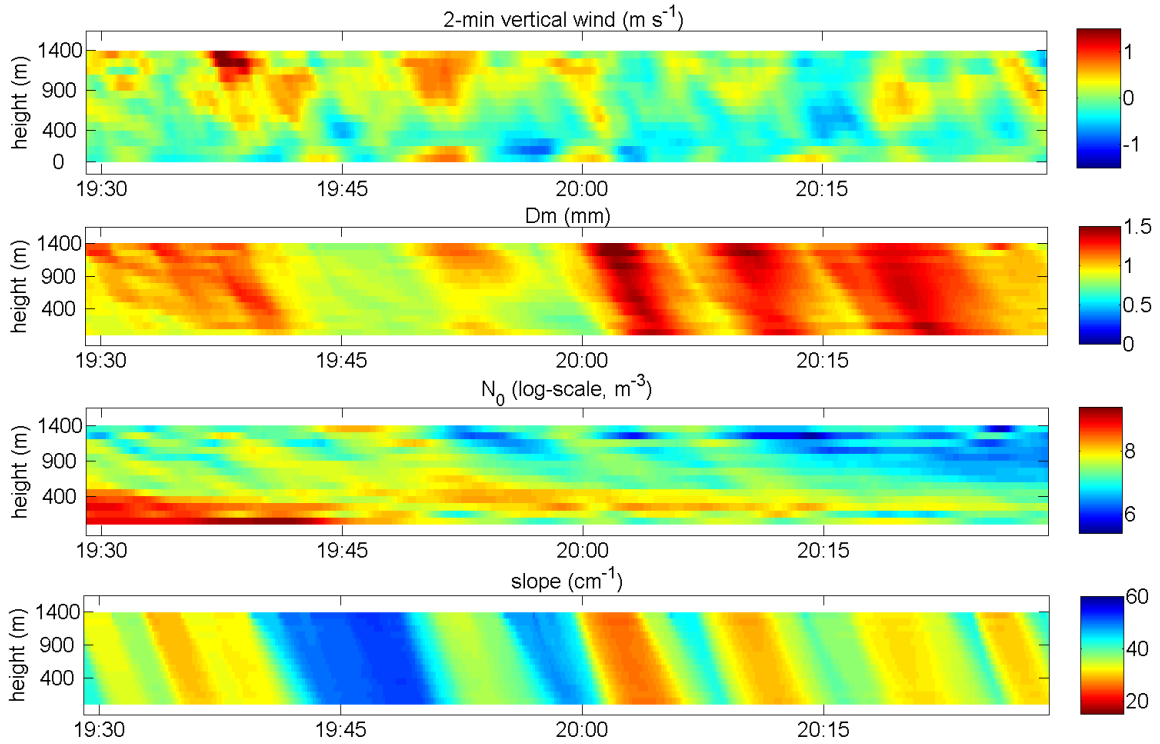

Figure 9. For the real case study: time-height fields of different assimilated parameters induced by DSD field $\tilde{N}$ or wind $\tilde{\mathbf{w}}$. Top: 2 min vertical wind field (positive downward). Middle high: mean volume diameter D_{m}. Middle low: number of drops N_0 (moment of order 0 of the DSD). Bottom: slope of the DSD, evaluated as the slope of the linear regression of the logarithm of the DSD on the diameters.

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

université PARIS-SACLAY

ÉCOLE DOCTORALE
Sciences mécaniques et énergétiques, matériaux et géosciences (SMEMAG)

Titre : Assimilation variationnelle d'observations multi-échelles : Application à la fusion de données hétérogènes pour l'étude de la dynamique micro et macrophysique des systèmes précipitants.

Mots clés : assimilation de données, microphysique de la pluie, télédétection

Résumé : D'une part, les instruments permettant de mesurer les précipitations (pluviomètres, radars, etc.) effectuent des mesures de natures différentes et à différentes échelles. Leurs données sont difficilement comparables. D'autre part, les modèles décrivant l'évolution des précipitations sont eux complexes et difficiles à paramétrer et à valider. Dans cette thèse, nous utilisons l'assimilation de données afin de coupler des observations hétérogènes des précipitations et des modèles, pour étudier les précipitations et leur variabilité spatiotemporelle à différentes échelles (macrophysique, qui s'intéresse aux cellules de pluie, et microphysique, qui s'intéresse à la distribution en taille des gouttes – DSD – qui les composent).

Tout d'abord, nous développons un algorithme permettant de restituer des cartes de précipitations à partir de mesures de l'atténuation causée par la pluie à des ondes provenant de satellites de télévision. Nos restitutions sont validées par rapport à des données radar et pluviomètres sur un cas d'étude dans le sud de la France.
Ensuite, nous restituons, toujours par assimilation de données, des profils verticaux de DSD et de vents verticaux à partir de mesures de flux de gouttes au sol (par disdromètres) et de spectres Doppler en altitude (par radar). Nous utilisons ces restitutions sur 3 cas d'étude pour étudier les phénomènes physiques agissant sur les gouttes de pluie durant leur chute et pour évaluer la paramétrisation de ces phénomènes dans les modèles.

Title : Variationnal assimilation of multi-scale observations: fusion of heterogeneous data for the study of dynamics of rainfall at macro and microscopic scales.

Keywords : data assimilation, rain microphysics, remote sensing

Abstract : On the one hand, the instruments designed to measure rainfall (rain gages, radars, etc.) perform measurements at different scales and of different natures. Their data are hard to compare. On the other hand, models simulating the evolution of rainfall are complex. It is not an easy task to parameterize and to validate them. In this thesis, we use data assimilation in order to couple heterogeneous observations of rainfall and models for studying rain and its spatiotemporal variability at different scales (macrophysical scale, which is interested in rain cells, as well as microphysical scale, which is interested in the drop size distribution – DSD).

First, we develop an algorithm able to retrieve rain maps from measurements of attenuation of waves coming from TV satellites due to rainfall. Our retrievals are validated by comparison with radar and rain gages data for a case study in south of France.
Second, we retrieve – again with data assimilation – vertical profiles of DSD and vertical winds from measurements of rain drop fluxes on the ground (using a disdrometer) and of Doppler spectra aloft (using a radar). We use these retrievals for 3 case studies to study the physical phenomena acting on rain drops during their fall and to evaluate the parameterization of these phenomena in models.