



HAL
open science

Utilisation des données de capteurs météorologiques connectés pour la prévision numérique à courte échéance de phénomènes de méso-échelle

Alan Demortier

► **To cite this version:**

Alan Demortier. Utilisation des données de capteurs météorologiques connectés pour la prévision numérique à courte échéance de phénomènes de méso-échelle. Sciences de la Terre. Université de Toulouse, 2024. Français. NNT : 2024TLSEP072 . tel-04721974

HAL Id: tel-04721974

<https://theses.hal.science/tel-04721974v1>

Submitted on 4 Oct 2024

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.

Doctorat de l'Université de Toulouse

préparé à Toulouse INP

Utilisation des données de capteurs météorologiques
connectés pour la prévision numérique à courte échéance de
phénomènes de méso-échelle

Thèse présentée et soutenue, le 4 septembre 2024 par

Alan DEMORTIER

École doctorale

SDU2E - Sciences de l'Univers, de l'Environnement et de l'Espace

Spécialité

Océan, Atmosphère, Climat

Unité de recherche

CNRM - Centre National de Recherches Météorologiques

Thèse dirigée par

Olivier CAUMONT et Vivien POURRET

Composition du jury

M. Jean-Pierre CHABOUREAU, Président, Observatoire Midi-Pyrénées

M. Remko UIJLENHOET, Rapporteur, Technische Universiteit Delft

M. Emmanuel COSME, Rapporteur, Université Grenoble Alpes

Mme Jacqueline SUGIER, Examinatrice, Met Office

M. Roger RANDRIAMAMPINANINA, Examineur, Norwegian Meteorological Institute

M. Olivier CAUMONT, Directeur de thèse, Toulouse INP

M. Vivien POURRET, Co-directeur de thèse, CNRS Occitanie Ouest

Membres invités

M. Marc Mandement, Météo-France

M. Patrick Moll, Centre National de Recherches Météorologiques

Trois ans pour s'appropriier un sujet de recherche et produire des résultats scientifiques pertinents. C'est court, c'est long, c'est parfois lent, parfois intense. J'espère avoir mené un projet assez robuste ; qui résiste à minima aux rafales du temps. Construire un projet, c'est essayer, rater, recommencer. Et parfois, à force de recommencer, on se demande si tout cela a bien un sens. Alors on tourne le raisonnement pour qu'il coule dans le *bon* sens. Et pourtant il y a des faits qui restent : des calculs qui sont trop gourmands et une quête à la résolution qui ne va certainement pas y remédier. Alors, vous qui avez l'esprit neuf, vous vous forgerez votre propre avis.

Et à la fin, une question reste : à quoi ressemblera le système de prévision numérique du temps de méso-échelle de demain ? Qui-sait, peut-être que les observations d'opportunités constitueront un des maillons de la longue chaîne que représente le système de prévision.

Remerciements

Cette thèse n'aurait jamais pu voir le jour sans un grand nombre de personnes très différentes qui m'ont portées, inspirées et qui m'entourent de près ou de loin. La force d'une foule (*crowd*) provient avant tout de sa diversité. Et, soit dis en passant, je crois qu'imposer une manière de penser à l'échelle d'une foule est risqué et ne peut mener qu'à une fuite en avant. Un météorologue le sait, une prévision d'ensemble à faible dispersion, n'a que peu d'intérêt. Alors, merci à celles et ceux qui osent réaliser de beaux projets.

Je crois que l'on sous-estime souvent l'importance et l'influence que des proches ou des pairs peuvent avoir et la force que cela représente, alors je tiens sincèrement à les remercier.

D'abord mes encadrants, ce trio de choc : Olivier Caumont qui a monté le sujet et qui m'a inspiré la rigueur et l'exigence nécessaire pour mener à bien ce projet scientifique ; Vivien Pourret pour sa bienveillance, son aide technique et les longues discussions à me répéter les bases de l'assimilation de données, parce qu'on y revient toujours ; et enfin Marc Mandement, toujours présent, toujours là pour répondre à mes milliers de questions, avec son énergie-enthousiasme presque infinie. C'était un vrai plaisir de travailler avec vous !

Mes collègues du couloir Precip/Phy-NH pour les nombreuses pauses-café et repas partagés. Quelle chance de vous avoir rencontrés.

Au pôle d'appui à la recherche pour votre aide pour les missions.

Au support et aides sur Olive : Véronique Mathiot, Maud Martet.

Aux autres doctorants qui permettent une belle entraide. Une pensée spéciale à ceux qui sont partis de manière anticipée sur d'autres chemins Maëva, Malo, Arthur, Morgane. . .

Mes colocs de cœur : Arnaud, Arthur, Gabriel et Isabell. Que de bons moments rue Rappas ! Merci Arnaud pour toutes ces pauses à/dorer au soleil pour redonner de la valeur à ces petits instants précieux.

Aux copain · es Octave, Emma, Célést, Tanguy, Alice, Thomas. . . et les autres à qui je pense tout autant. Aux copains de toujours de l'ENM.

Romane qui m'inspire au quotidien. Merci pour ton soutien inconditionnel. Merci pour tes conseils précieux surtout lors des derniers mois.

Ma famille élargie, pilier bienveillant et intemporel.

Et à tous mes compagnons d'aventure du quotidien, infiniment merci !

*Anticiper une tornade,
s'émerveiller de papillons,
un sens unique.*

Résumé

L'assimilation de données à méso-échelle permet d'initialiser les modèles régionaux de prévision numérique du temps, en combinant une ébauche du modèle avec les observations disponibles notamment à cette échelle. L'émergence de stations météorologiques connectées fournit une source d'observations d'opportunité permettant de densifier le réseau d'observation de surface. Ces nouvelles stations permettent d'observer à fine échelle la convection profonde près de la surface (Mandement, 2020). Cette thèse propose ainsi d'explorer la capacité de ces observations d'opportunité de température, d'humidité relative et de pression à améliorer la prévision de phénomènes météorologiques de méso-échelle, en les assimilant dans le modèle régional AROME-France.

L'étape de pré-traitement de ces observations est nécessaire pour corriger et filtrer les observations considérées anormales vis-à-vis de l'ébauche du modèle, et d'écrêter les observations spatialement proches. Ce pré-traitement permet d'apporter à l'assimilation près de 129 fois plus d'observations au sol de pression, 3 fois plus d'observations de température et 6 fois plus d'observations d'humidité relative que dans le cas où seules les observations standards de surface sont utilisées.

Des expériences avec et sans observations d'opportunité ont été menées (méthode OSE) sur une période d'un mois. Ces OSE ont d'abord été réalisées en assimilant de manière indépendante les trois types d'observations d'opportunité susmentionnés, puis de manière conjointe. Les résultats de l'assimilation des observations d'opportunité de pression montrent une diminution de l'écart quadratique moyen entre la prévision à 1 h d'échéance et les observations standards de pression de 10.3 % en France, avec le schéma d'assimilation 3DEnVar. Il en est de même pour l'assimilation des observations d'opportunité de température et d'humidité relative, qui permettent une diminution de 0.9 % et de 1 % respectivement par rapport aux observations standards de la même grandeur. Avec le réglage actuel du schéma 3DEnVar, il apparaît que les observations de surface disposent d'un poids moindre à l'analyse par rapport aux autres observations et par rapport à l'ébauche. L'assimilation de nombreuses observations d'opportunité de surface contribue à compenser cet effet. Le type de schéma d'assimilation et son réglage jouent un rôle important dans l'assimilation de ces nouvelles observations. Le schéma d'assimilation 3DVar actuellement opérationnel n'est ainsi pas capable de tirer profit des observations d'opportunité, du fait de la création d'incrémentes à la fois isotropes et de grandes dimensions horizontales.

Des études de cas ont été menées, dans un premier temps sur des situations comportant de forts gradients de pression de surface — comme dans le cas d'un système convectif de méso-échelle, d'un front, d'une dépression de méso-échelle ou d'un mascaret atmosphérique ; puis dans des cas comportant de forts gradients de température ou d'humidité relative près de la surface — dans des cas de brise de mer, de brouillard, d'îlot de chaleur urbain ou de ligne orageuse. Ces études de cas ont permis de mettre en évidence la capacité des observations d'opportunité à améliorer la représentation de ces phénomènes, en contraignant l'analyse près de la surface.

Abstract

Mesoscale data assimilation is used to initialise regional numerical weather prediction models, by combining a first guess with the available observations particularly at this scale. The emergence of personal weather stations (PWS) is a source of observations that can be used to densify the surface observation network. These new observations have demonstrated their ability to observe deep convection near the surface at fine scales. This thesis proposes to explore the capacity of these temperature, relative humidity and pressure observations to improve the forecasting of mesoscale meteorological phenomena, by assimilating them in the AROME-France regional model.

These observations need to be pre-processed in order to correct and filter out observations that are considered anomalous in relation to the first model guess, and to be thinned out to remove close observations in space. This pre-processing makes it possible to assimilate almost 129 times more pressure observations, 3 times more temperature observations and 6 times more relative humidity observations than in the case where only standard surface observations are used.

Experiments with and without PWS observations were carried out (OSE method) over a period of one month. These OSEs were first carried out by assimilating the three types of PWS observations mentioned above independently, and then jointly. The results of the assimilation of PWS pressure observations show a reduction in the mean square deviation between the 1 h forecast and the standard pressure observations of 10.3 % in France, using the 3DEnVar scheme. The same applies to the assimilation of PWS temperature and relative humidity observations, which result in a reduction of 0.9 % and 1 % respectively compared to standard observations of the same parameter. With the current setting of the 3DEnVar scheme, it appears that surface observations are given less weight in the analysis than both other observations and the first guess. The assimilation of many PWS surface observations helps to compensate for this effect. The type of assimilation scheme and its adjustment play an important role in the assimilation of these new observations. The use of the 3DVar assimilation scheme is thus unable to take advantage of dense PWS observations, due to the creation of increments that are both isotropic and of large horizontal dimensions.

Case studies were carried out, firstly on situations involving strong gradients in surface pressure - as in the case of a mesoscale convective system, a front, a mesoscale low pressure system or an atmospheric bore; then on cases involving strong gradients in surface temperature or relative humidity - in the case of a sea breeze, a fog, an urban heat island or a squall line. These case studies highlighted the ability of PWS observations to improve the representation of these phenomena, by constraining the surface analysis.

Table des matières

Introduction générale	1
Chapitre 1 Phénomènes météorologiques de méso-échelle : des observations au sol aux modèles de prévision numérique du temps	5
1.1 Observations au sol	6
1.1.1 Un réseau mondial d'observations au sol	6
1.1.2 Stations météorologiques standards	7
1.1.3 Réseau d'observations non conventionnelles	10
1.2 Modèle de prévision numérique du temps régional	19
1.2.1 Équilibre atmosphérique et résolution	19
1.2.2 Le modèle de prévision AROME-France	20
1.3 Traces au sol des phénomènes de méso-échelle : de l'observation au modèle	25
1.3.1 Minimum dépressionnaire et front	25
1.3.2 Système convectif de méso-échelle	29
1.3.3 Phénomènes de basses couches	33
1.4 Synthèse	36
Chapitre 2 Assimilation d'observations au sol dans AROME-France	39
2.1 Assimilation de données	40
2.1.1 Formalisme général du problème	40
2.1.2 Schéma variationnel	40
2.1.3 Schéma d'interpolation optimale	42
2.1.4 Résolution séquentielle	43
2.1.5 Application pour différents modèles à aire limitée	43
2.2 Cycle d'assimilation d'AROME	44
2.2.1 Contrôle de qualité du modèle	44
2.2.2 Opérateurs d'observation au sol	45
2.2.3 Espace des observations	48
2.2.4 Réglage des covariances d'erreur d'ébauche	50
2.2.5 Schéma d'assimilation de surface	52
2.2.6 Chaîne de prévision opérationnelle	53
2.3 Assimilation des observations de surface	55
2.3.1 Quel impact ont les observations de surface?	55
2.3.2 Pourquoi l'assimilation d'observations standards de surface est-elle un challenge?	56
2.3.3 Quels défis pour l'assimilation d'observations d'objets connectés?	58
2.3.4 Quels défis pour l'assimilation de données à échelle convective?	59
2.4 Synthèse	60

Chapitre 3 Assimilation d’observations de pression de surface provenant des stations personnelles dans AROME-France 61

3.1	Résumé de l’article	62
3.2	Article : <i>Assimilation of surface pressure observations from personal weather stations in AROME-France</i>	63
3.2.1	Introduction	64
3.2.2	Data, tools, and methodology	67
3.2.3	Pre-processing algorithm	69
3.2.4	Results of 1 month assimilation experiments	71
3.2.5	Results on the 14 September 2021 heavy precipitation event	74
3.2.6	Discussion and conclusion	81
3.3	Compléments de l’article	85
3.3.1	Supervision a posteriori des observations SWS assimilées	85
3.3.2	Gain spatial des observations en pression des PWS	85
3.3.3	Vérifications du système d’assimilation	87
3.3.4	Tests de sensibilité	88
3.4	Phénomènes à fort gradient de pression	91
3.4.1	Front - 5 octobre 2021	92
3.4.2	Dépression secondaire - 3 novembre 2023	93
3.4.3	Mascaret atmosphérique - 18 juin 2022	95
3.5	Synthèse et perspectives	98

Chapitre 4 Assimilation d’observations de température et d’humidité relative provenant des stations personnelles dans AROME-France 101

4.1	Résumé de l’article	101
4.2	Article : <i>Assimilation of temperature and relative humidity observations from personal weather stations in AROME-France</i>	102
4.2.1	Introduction	103
4.2.2	Datasets and methods	105
4.2.3	Assimilation experiments design	113
4.2.4	Results of the assimilation experiments	118
4.2.5	Results on case studies : heatwave and fog events	121
4.2.6	Discussion and conclusion	125
4.3	Complément de l’article	135
4.3.1	Filtres des observations de surface	135
4.3.2	Cas d’étude : ICU du 17 juin 2023	137
4.3.3	Cas d’étude : ligne orageuse le 8 septembre 2021	137
4.4	Synthèse et perspectives	141

Chapitre 5 Assimilation conjointe de plusieurs types d’observations 143

5.1	Assimilation conjointe de plusieurs types d’observations	143
5.1.1	Impact en surface	143
5.1.2	Impact en altitude	146
5.1.3	Score indicateur	147
5.1.4	Augmentation du coût de calcul de l’assimilation des observations PWS	148
5.2	Synthèse	149

Conclusions générales et perspectives	151
Annexes	156
A Cas d'études réalisés	156
B Coût énergétique des expériences réalisées	158
B.1 Contexte	158
B.2 Méthode de Lannelongue <i>et al.</i> (2021)	159
B.3 AROME-France : expériences sur un mois	159
B.4 Statistiques sur l'étude complète	160
B.5 Conclusion	161
Sigles et acronymes	163
Liste des figures	165
Liste des tableaux	170
Références bibliographiques	173

Introduction générale

Les phénomènes météorologiques impactent un grand nombre de domaines, de l'agriculture à l'industrie, en passant par le transport et la sécurité des personnes et des biens (SPB). S'intéresser à la prévision du temps permet de comprendre le fonctionnement de notre environnement et de mieux anticiper des phénomènes météorologiques parfois extrêmes.

La compréhension des lois de la physique de l'atmosphère et le développement des moyens de calcul au cours du siècle dernier ont permis l'émergence de la prévision numérique du temps (PNT). La PNT est fondée sur la résolution des équations de la dynamique des fluides calculant l'évolution de variables atmosphériques. Celle-ci ne cesse de s'améliorer grâce à l'augmentation de nos connaissances scientifiques, et au prix d'une puissance de calcul de plus en plus importante.

La PNT alimente directement bon nombre de produits à destination du grand public ou de contrats commerciaux. Les champs modèles issus de la PNT sont complexes et l'expertise de prévisionnistes constitue le maillon fondamental pour la prise de décisions importantes, telles que le passage en vigilance et l'avertissement des centres opérationnels de zone (COZ), qui coordonnent les moyens de la sécurité civile en cas d'événements météorologiques dangereux.

L'anticipation de la prévision d'événements météorologiques dangereux est aujourd'hui possible du fait de l'amélioration des modèles de PNT. Les modèles de PNT régionaux, aussi nommés modèles à aire limitée, ont une maille plus fine que les modèles globaux, et sont plus à même de prévoir des phénomènes de plus fine échelle. Parmi les nombreux moyens d'améliorer la prévision numérique du temps se trouve (i) l'amélioration de la représentation des processus, en particulier avec des schémas de paramétrisation physique plus complets, (ii) l'amélioration des conditions initiales et (iii) l'amélioration des forçages aux bords, que ce soit pour les conditions de surface ou les conditions latérales des modèles régionaux. Dans cette thèse, on s'attachera au deuxième point, concernant l'amélioration de l'état initial, aussi appelé analyse.

Deux stratégies sont possibles pour former l'analyse d'un modèle régional de PNT. Une première stratégie consiste à interpoler la prévision d'un modèle global à la résolution du modèle régional. Cette méthode n'intègre pas de nouvelles observations, et garde la cohérence des phénomènes physiques décrits par le modèle global. Elle permet ainsi une stabilité entre les réseaux successifs de prévision. Une seconde stratégie consiste à réaliser une analyse à méso-échelle, c'est-à-dire à intégrer de nouvelles observations dans le modèle, ce qu'on appelle « assimilation de données ». Pour les modèles régionaux, on parle d'assimilation de données de méso-échelle ou assimilation à échelle convective. Sur certains cas d'événements précipitants, ce n'est qu'en introduisant dans l'instant initial des structures de méso-échelle que le modèle parvient à simuler ces événements avec réalisme (Ducrocq *et al.*, 2000), montrant ainsi l'importance de l'initialisation à fine échelle pour la prévision d'événements à enjeux (Ducrocq *et al.*, 2002; Ricard, 2002).

Les méthodes d'initialisation ont été adaptées pour des échelles spatiales et tempo-

relles plus fines, et des réseaux d'observations plus denses. L'interpolation optimale a laissé la place aux méthodes variationnelles (Talagrand et Courtier, 1987), permettant l'assimilation des observations de manière directe et conjointe. L'évolution récente vers des méthodes d'assimilation ensemblistes (EnVar) permet une meilleure prise en compte de la situation météorologique au moment de l'assimilation.

De nombreux centres météorologiques développent leurs modèles à aire limitée. AROME est le modèle développé par Météo-France et ses partenaires. Les modèles à aire limitée, comme AROME, se sont dotés au fil du temps de l'assimilation d'observations issues de réseaux d'observations spatialement denses, tels que les réseaux d'observations RADAR (Caumont *et al.*, 2010). Cependant, malgré l'augmentation du nombre d'observations assimilées dans les modèles régionaux, ces dernières restent insuffisantes pour contraindre l'ensemble des degrés de liberté du modèle à l'instant initial. L'assimilation de nouveaux systèmes d'observations est à l'étude, tels que les observations issues d'avions (Pourret *et al.*, 2021), les observations d'éclairs (Combarrous, 2023) ou les observations infrarouges MTG-IRS (Coopmann *et al.*, 2023).

Parallèlement, l'arrivée de nouvelles observations dites d'opportunité, issues d'objets connectés, apparaît comme un nouveau potentiel (Muller *et al.*, 2015). En particulier, les stations météorologiques personnelles (PWS) achetées par le grand public fournissant des observations disponibles en temps réel se font remarquer. En effet, les PWS atteignent une quantité jusqu'alors jamais atteinte par des réseaux d'observations de surface conventionnelles (Bell *et al.*, 2015; Meier *et al.*, 2017; Bell *et al.*, 2022). Les PWS peuvent mesurer la température, l'humidité relative, la pression, le vent ou bien le cumul de précipitations.

Les travaux de Mandement (2020) ont montré que les observations de PWS permettent de représenter des signatures de phénomènes convectifs comme des surpressions à l'avant de systèmes convectifs de méso-échelle. Plus généralement, les observations issues de PWS permettent l'étude de phénomènes locaux, tels qu'un orage de grêle (Clark *et al.*, 2018), un épisode fortement précipitant (Caumont *et al.*, 2021), ou des îlots de chaleur urbains (Madelin et Dupuis, 2020). Les données de PWS représentent une opportunité pour améliorer l'analyse et, in fine, la prévision de phénomènes à forts enjeux tels que les vagues de chaleur, le brouillard, les orages ou les tempêtes.

Cette thèse propose des premières expérimentations d'assimilation d'observations issues de stations météorologiques personnelles dans AROME-France. Elle se focalisera sur les observations de pression, de température et d'humidité relative, qui sont des variables de stations conventionnelles déjà assimilées dans le modèle AROME et dont les capacités à représenter des phénomènes de méso-échelles sont déjà connues.

Les défis posés par les précédents auteurs sont considérables. Les défis qui suivent se posent pour chaque type d'observation de PWS. Quels sont les pré-traitements nécessaires à leur assimilation tout en gardant l'information de petite échelle contenue dans les données disponibles? En effet, il existe de nombreuses techniques de correction et de contrôle de qualité permettant de traiter les observations des stations personnelles (Coney *et al.*, 2022). Une comparaison des traitements serait nécessaire pour en déterminer les avantages et inconvénients, tout en composant avec les contraintes du temps réel imposées par l'assimilation de données.

Un défi supplémentaire réside dans la qualité du schéma d'assimilation de données utilisé. Les systèmes d'assimilation de données actuels permettent-ils l'assimilation d'informations de petite échelle fournies par ces nouvelles observations? Ridal *et al.* (2019) ont montré les difficultés de l'assimilation des observations issues de PWS dans le système d'assimilation variationnelle 3D (3DVar) avec les réglages actuels. L'émergence de

systèmes d'assimilation EnVar pour les modèles à aire limitée pourrait permettre une meilleure utilisation d'un tel système d'observations spatialement denses.

Ainsi, si le système d'assimilation de données est capable de tirer parti de ces nouvelles observations, est-ce qu'il y a un gain sur la prévision de phénomènes de méso-échelle, tels que des systèmes convectifs de méso-échelle ? Est-ce qu'en moyenne les prévisions issues de ces analyses sont meilleures (c'est-à-dire plus proches des observations conventionnelles) ?

Enfin, si un gain est apporté par ces nouvelles observations, est-ce qu'il pourrait être envisageable d'assimiler ces nouvelles observations d'opportunités dans le modèle AROME-France opérationnel ?

Dans le chapitre 1, nous décrivons les systèmes d'observations de surface disponibles ainsi que le modèle de PNT AROME, et nous montrerons les capacités qu'ont ces deux systèmes à observer et à simuler des événements de méso-échelle. Dans le chapitre 2, nous décrivons les différents systèmes d'assimilation de données utilisés pour les modèles à aire limitée opérationnels, qui permettent d'initialiser les modèles en tirant parti de la dernière prévision et des observations disponibles. Un intérêt particulier sera porté sur le système d'assimilation de données d'AROME. Nous nous attarderons particulièrement sur l'assimilation des observations de surface et les défis posés par l'assimilation des observations des PWS. Dans le chapitre 3, nous présenterons les résultats de l'assimilation des observations de pression provenant des PWS dans AROME. Nous nous baserons conjointement sur des scores à long terme et sur des études de cas pour comprendre l'influence d'un réseau dense d'observations au sol dans le système de PNT. Dans le chapitre 4, nous verrons les premiers résultats de l'assimilation des observations de température et d'humidité relative provenant des PWS dans AROME. Nous verrons quelle influence ces observations peuvent avoir sur la prévision de la couche limite atmosphérique. Enfin, dans le chapitre 5, nous testerons l'assimilation combinée des observations de pression, température et humidité relative des PWS, pour connaître les gains de prévisions qu'elles permettent et le coût numérique qu'elles induisent. Nous concluons sur les perspectives qu'offrent les observations des stations météorologiques personnelles pour l'assimilation de données, et plus largement pour la communauté de recherche météorologique.

Ce travail s'inscrit dans le contrat d'objectif et de performances (COP) 2022–2026 de Météo-France (objectif n°6), et à l'échelle européenne via le consortium ACCORD¹ qui rassemble les partenaires euro-méditerranéens pour la recherche et le développement de modèles de PNT dont le modèle AROME (objectif n°4, RWP-ACCORD).

1. A Consortium for COnvection-scale modelling Research and Development

Chapitre 1

Phénomènes météorologiques de méso-échelle : des observations au sol aux modèles de prévision numérique du temps

Sommaire

1.1	Observations au sol	6
1.1.1	Un réseau mondial d'observations au sol	6
1.1.2	Stations météorologiques standards	7
1.1.3	Réseau d'observations non conventionnelles	10
1.2	Modèle de prévision numérique du temps régional	19
1.2.1	Équilibre atmosphérique et résolution	19
1.2.2	Le modèle de prévision AROME-France	20
1.3	Traces au sol des phénomènes de méso-échelle : de l'observation au modèle	25
1.3.1	Minimum dépressionnaire et front	25
1.3.2	Système convectif de méso-échelle	29
1.3.3	Phénomènes de basses couches	33
1.4	Synthèse	36

Ce premier chapitre vise à montrer les capacités des observations de surface et du modèle de prévisions numériques AROME, à respectivement visualiser et simuler des phénomènes de méso-échelle.

Le terme de méso-échelle est utilisé pour désigner une échelle intermédiaire entre l'échelle synoptique, qui comprend la circulation planétaire décrivant la position des dépressions et des anticyclones, et les échelles aérologiques inférieures à 2 km (Orlanski, 1975b). La méso-échelle comprend des phénomènes d'échelle méso- α , allant de 200 à 2000 km, des phénomènes méso- β de 20 à 200 km et méso- γ de 2 à 20 km. Ce sont principalement des phénomènes méso- β qui vont nous intéresser ici, désignant des phénomènes tels que des supercellules, des lignes de grains, des ondes orographiques ou des brises de mer (figure 1.1). La méso-échelle joue un rôle important dans le transfert de l'énergie des grandes échelles vers les petites échelles, répartissant par exemple le surplus d'énergie reçu à l'équateur vers les moyennes latitudes. De même, les phénomènes de méso-échelle permettent de répartir l'énergie des plus petites échelles vers les grandes, par exemple sans elle l'humidité et la chaleur sensible serait piégée à la surface. Autrement dit, il existe une continuité entre grandes et petites échelles, caractérisés par des échanges d'énergies ou cascade d'énergie (Kolmogorov (1941)). Pour appréhender un phénomène de méso-échelle, il convient de comprendre et de représenter dans les modèles régionaux les processus physiques de l'échelle synoptique jusqu'aux plus petites échelles de la turbulence.

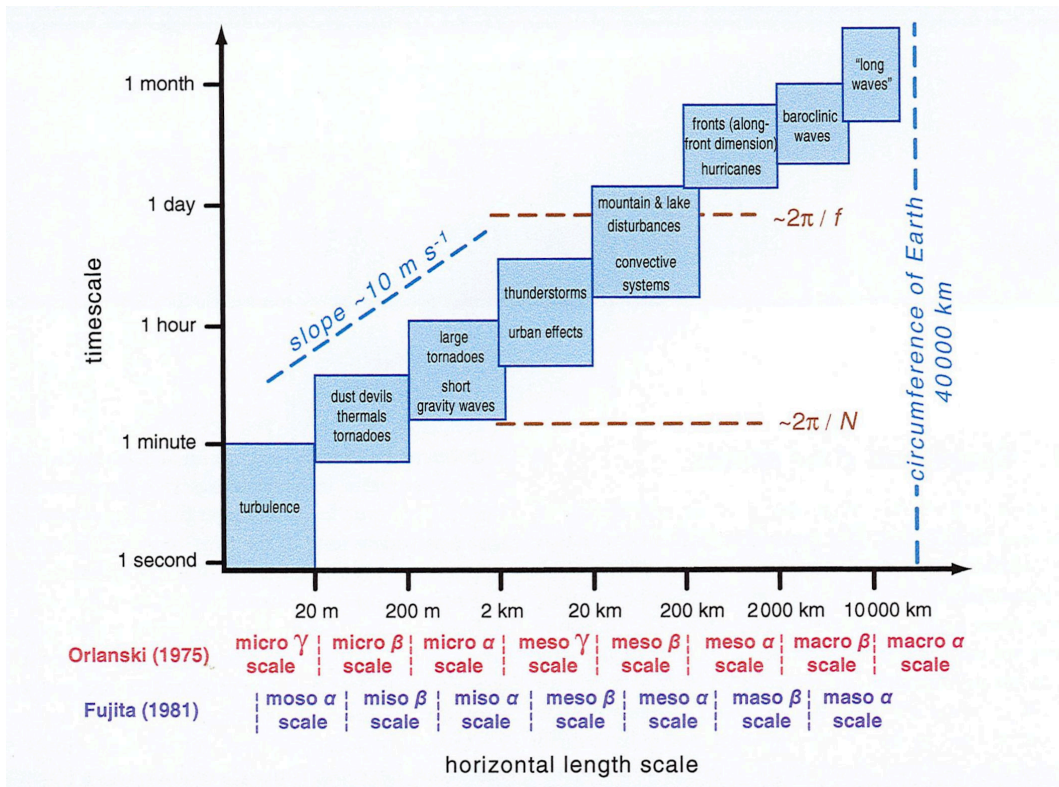


FIGURE 1.1 – Définition des échelles spatiales et temporelles caractéristiques de phénomènes météorologiques variés. Classification de Orlanski (1975a) et Fujita (1981). Figure issue de Markowski et Richardson (2011).

1.1 Observations au sol

1.1.1 Un réseau mondial d'observations au sol

Une observation météorologique est une mesure quantitative d'une variable physique, tel que la pression atmosphérique. Elle est accompagnée de sa position géographique et d'une référence temporelle. Plusieurs observations météorologiques peuvent être réalisées au sein d'une station de mesures. Ces stations sont rassemblées sous forme de réseaux de stations à différentes échelles.

La Veille Météorologique Mondiale (VMM) est un des principaux programmes de l'Organisation Météorologique Mondiale (OMM) qui regroupe, pour chaque pays membre, les moyens dédiés aux systèmes d'observations, aux télécommunications et aux traitements des données. Ce programme comprend une branche « Instruments et méthodes d'observation » dont l'objectif est la standardisation et le développement des observations météorologiques. Le réseau ainsi créé forme le système mondial d'observations de l'OMM (WIGOS, *WMO Integrated Global Observing System*). Le réseau WIGOS, comprend à la fois les observations satellitaires et le réseau mondial d'observations de surface nommé GBON (*Global Basic Observing Network*). Le réseau mondial d'observations GBON est fondé sur un accord international qui répond au besoin de partage des données les plus essentielles. GBON repose sur le principe d'un partage libre et gratuit des observations et sur la collaboration entre les nations.

L'intérêt de maintenir et de développer un tel système est multiple. Il permet à la fois, comme on va le voir, d'améliorer les prévisions météorologiques, mais aussi de fournir des avis sur des phénomènes météorologiques dangereux et d'alimenter le réseau d'obser-

vations climatologique permettant d’avoir de longues séries continues d’observations. En pratique, cette mission a un coût, qui est lié à l’installation des stations et à la maintenance du réseau d’observations. Pour aider au développement du réseau, l’OMM a mis en place une assistance financière et technique de long terme, via le système SOFF (*Systematic Observations Financing Facility*). Le SOFF permet d’aider particulièrement les pays pour lesquels le réseau d’observations est le moins étendu, priorisant ainsi les nations et états insulaires en développement.

L’ensemble des métadonnées des observations, c’est-à-dire les données associées aux observations telles que leurs coordonnées, est constamment mise à jour mondialement via l’interface OSCAR. La figure 1.2 illustre l’ensemble des stations de surface actuellement disponible. On peut constater que certaines régions restent peu denses en observations de surface, telles que les régions désertiques et montagneuses.

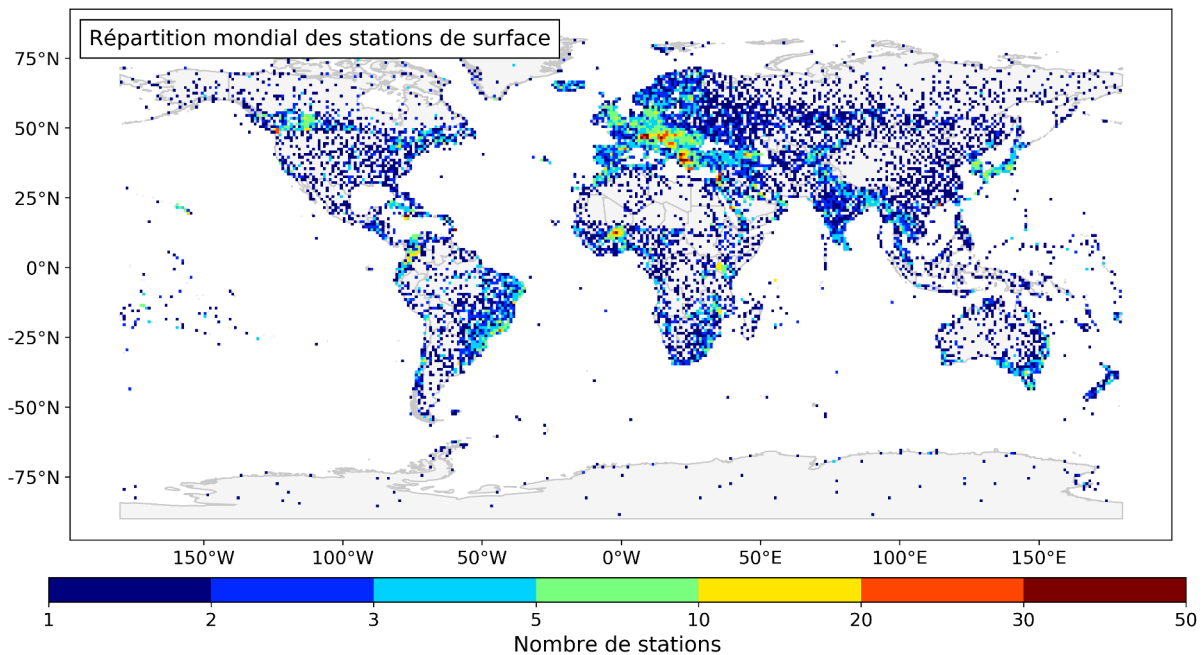


FIGURE 1.2 – Nombre de stations de surface échangeant des observations via l’OMM. Le nombre de stations est divisé en mailles d’environ $1^\circ \times 0.5^\circ$. À la date du 27 février 2024, on dénombre 16823 stations de surface, comprenant les stations GBON de surface, ainsi que des réseaux locaux de stations automatiques de surface (AWS), stations SYNOP et climatologiques (<https://oscar.wmo.int/surface/>).

1.1.2 Stations météorologiques standards

En France, le réseau d’observations est d’abord manuel, avec le réseau climatologique d’état (RCE), organisé depuis 1945 par la Météorologie Nationale (future Météo-France). La Météorologie Nationale se charge alors de l’installation, l’inspection et l’entretien des capteurs, et de la collecte des données pluviométriques (hauteur de précipitations en 24 h) et de données de températures extrêmes journalières. À l’issue de la Seconde Guerre mondiale, la première station automatique est expérimentée à Trappes en 1948 (Météo-France, 2020). Puis, le premier réseau opérationnel de stations automatiques arrivera à la fin des années 1960. Ce réseau est composé de stations SATIN qui sont les premières stations numériques qui transmettent toutes les heures les données sous forme binaire à un

calculateur central. Elles sont remplacées par les stations SIMOUN en 1974, les premières à fonctionner avec le système des concentrateurs².

Il faudra attendre janvier 1996, pour que Météo-France lance le projet RADOME (Réseau d'acquisition de données et d'observations météorologiques étendues) qui initie le réseau de stations de surface que l'on connaît aujourd'hui. Il sera complété par le réseau régional automatisé (RRA) axé sur des données plus spécifiques à l'aéronautique et aux armées. Le réseau RADOME réunit actuellement 600 stations, auxquelles s'ajoute le réseau NIVOSE (30 stations) spécifiques aux zones montagneuses. Initialement, les stations, de type Xaria, permettaient de mesurer la température et l'humidité relative sous abri, le vent à 10 mètres, la pression et les précipitations, ainsi que le rayonnement solaire global et la température au-dessus et dans le sol. Par la suite, les stations ont progressivement été remplacées par des stations Mercury, permettant entre autres une version compacte embarquée à bord de navires et de bouées ancrées (Guillou, 2018).

Parmi les stations complémentaires collectées par Météo-France, on peut noter (i) les stations institutionnelles (par exemple provenant des sites nucléaires), (ii) les stations issues de prestations de services (par exemple provenant de contrats aéronautiques ou routiers) et (iii) les stations issues de réseaux d'intérêt commun (RIC), tels que les stations provenant des services de prévision des crues, du réseau EDF³ ou de l'INRAE⁴). Pour plus d'informations sur le développement et le nombre de stations de surface, nous renvoyons le lecteur vers la thèse de Mandement (2020).

Les réseaux de stations dites au sol ou à la surface réalisent la plupart de leurs observations au-dessus du sol. Afin de mesurer les propriétés de l'air au-dessus de la surface de façon homogène partout dans le monde, les capteurs de température et d'humidité relative doivent être placés entre 1.25 et 2 m de hauteur au-dessus du sol (Organisation Météorologique Mondiale, 2023). Les observations de température et d'humidité relative à ces hauteurs-là sont souvent désignées par les termes de température à 2 m (t2m) et d'humidité relative à 2 m (rh2m). Les capteurs de vent doivent, quant à eux, être placés à 10 m de hauteur au-dessus du sol (Organisation Météorologique Mondiale, 2023).

De plus, les mesures de température et d'humidité relative sont réalisées dans un abri météorologique. Un abri ou un écran anti-rayonnement est conçu de manière à constituer une enceinte dont la température intérieure est uniforme et identique à la température extérieure. Il ne laisse pénétrer ni les rayons incidents, ni les précipitations ou autres phénomènes susceptibles d'influencer le résultat de la mesure, et garantit une ventilation suffisante. Il existe une grande diversité d'abris, de tailles et de ventilations diverses. L'abri est un élément nécessaire pour garantir la qualité des observations de température et d'humidité relative, mais n'est pas suffisant.

1.1.2.1 Qualité des observations des réseaux de stations de surface

Chaque observation de surface est entachée d'erreurs, rassemblées sous le terme d'**erreur d'observation**. L'erreur d'observation peut se subdiviser en une **erreur instrumentale**, liée à la qualité de la mesure, et en une **erreur de représentativité** liée à la capacité de l'observation à représenter une large étendue spatiale. Pour chaque type d'erreur, il existe

2. Le concentrateur transcode les messages de la station automatique en code normalisé BUFR. Les messages BUFR sont transmis par le réseau interne de Météo-France vers le concentrateur de messages, situé à Toulouse, puis rediffusés vers les sites internes et externes (Tardieu et Leroy, 2003).

3. Électricité de France

4. Institut national de recherche pour l'agriculture, l'alimentation et l'environnement

TABLEAU 1.1 – Critères d’homologation d’une classe de performance maintenue pour les observations de température, d’humidité relative et de pression (Leroy, 2010). (*) Quand l’incertitude requise n’est pas actuellement atteignable, il est indiqué l’incertitude atteignable. La maintenance doit être réalisée dans un laboratoire accrédité.

		A	B	C	D	E
Température	Incertitude de la mesure	*0.2 °C	0.5 °C	1 °C	>1 °C	∅
	Maintenance	1 an	5 ans			
Humidité relative	Incertitude de la mesure	*3 %	6 %	10 %	>10 %	∅
	Maintenance	6 mois	1 an	2 ans		
Pression	Incertitude de la mesure	*0.3 hPa	0.5 hPa	1 hPa	>1 hPa	∅
	Maintenance	1 an	2 ans	Adaptée		

une classification spécifique : « la classification de performance maintenue » pour l’erreur instrumentale et « la classification du site » pour l’erreur de représentativité.

La classification de performance maintenue est définie dans une note de Météo-France (Leroy, 2010). Elle couvre à la fois les aspects de performance inhérents à l’instrument de mesure et les moyens mis en place pour la maintenance et l’étalonnage des capteurs. Cinq classes sont ainsi définies, allant de A à E. Les classifications des mesures sont effectuées pour chaque variable. Les critères pour les observations de température, d’humidité relative et de pression sont indiqués dans le tableau 1.1.

La classification d’un site de mesure, définie dans le Guide des instruments et des méthodes d’observation de l’OMM (Organisation Météorologique Mondiale, 2023), est quant à elle, appliquée au sein de chaque service météorologique. Différentes classes ont été créées, allant de 1 (classe de référence) à 5, et permettent de définir la représentativité de la mesure à petite échelle. Plus la mesure est proche de la classe 1 et plus elle sera représentative d’une zone étendue. Si le site possède différentes mesures, il est possible de caractériser la classe du site, en prenant la classe la plus élevée présente sur le site. Par convention, pour un paramètre donné, on mettra en classe 1 tout site respectant les critères de l’OMM. L’attribution d’une classe doit être revue périodiquement, car son environnement peut évoluer au fil du temps. Il est nécessaire de réaliser une vérification visuelle annuelle, et une actualisation complète des classes attribuées à un site tous les 5 ans.

Prenons l’exemple de la mesure de la température et de l’humidité relative sous abri. Deux critères sont vérifiés, liés (i) à l’éloignement aux obstacles susceptibles de générer une ombre portée et (ii) au type de surface au-dessus duquel il se trouve afin d’éviter les sources de chaleur ou autres biais contextuels.

La classe 1 pour une mesure de température ou d’humidité relative est définie suivant les critères suivants :

- Terrain plat et horizontal, entouré d’une surface de dégagement dont la pente est inférieure à 1/3 (19°) ;
- Sol recouvert de végétation naturelle basse (< 10 cm) représentative de la région ;
- À plus de 100 m de sources de chaleur artificielles ou de surfaces réfléchissantes (bâtiments, aires bétonnées, parcs de stationnement, etc.), d’étendues d’eau (sauf si elles sont significatives dans la région), et à l’écart de toute ombre portée lorsque la hauteur du soleil est supérieure à 5° au-dessus de l’horizon.

En pratique, pour définir si une source de chaleur peut avoir un impact significatif, différents critères de superficie de la source et de sa distance au capteur sont définis et résumés

dans la figure 1.3 pour la classe 1.

TABLEAU 1.2 – Pourcentage des 600 stations du réseau RADOME pour chaque classe environnementale (le 1 février 2024). (*) 5 et les stations sans classe environnementale.

Classe	Température	Humidité relative	Vent
1	29.7 %	29.4 %	12.8 %
2	33.7 %	33.7 %	31.0 %
3	19.8 %	20.2 %	38.9 %
4	15.5 %	15.4 %	15.1 %
5*	1.3 %	1.3 %	2.2 %

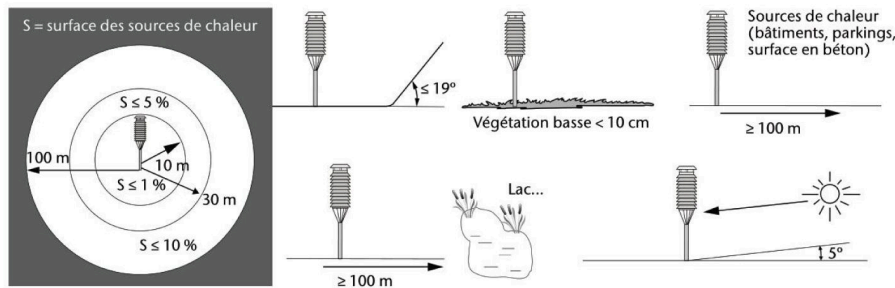


FIGURE 1.3 – Critères d’homologation en classe 1 d’une mesure de température et d’humidité relative de surface (Extrait de WMO, 2021). Une source de chaleur doit couvrir moins de 5 % de la surface si elle est à moins de 10 m, moins de 10 % de la surface si elle est à plus de 30 m et moins de 1 % de la surface si elle est à moins de 100 m.

Pour les stations RADOME en France, un décompte par classe de mesure a été réalisée (Tableau 1.2). Toutes les observations de température ne se valent pas, et seulement 30 % d’entre elles respectent les conditions de l’OMM. Ces classes peuvent ainsi intervenir pour qualifier l’incertitude de la mesure lors d’un record climatique.

Les observations de pression ne sont quant à elles pas sujettes à une erreur de représentativité, car elles ne sont pas directement influencées par la surface.

Les stations standards de surface (SWS, *standard weather stations*) ne sont pas les seules stations météorologiques. Comme présenté ci-après, des réseaux de stations tierces se développent conjointement.

1.1.3 Réseau d’observations non conventionnelles

L’engouement pour l’Internet des Objets (IdO), qui désigne l’interconnexion entre l’Internet et les environnements physiques, a permis la croissance du nombre d’objets connectés, créant de nouvelles possibilités pour développer les réseaux d’observations météorologiques. Les observations météorologiques issues d’objets connectés sont alors qualifiées **d’observations d’opportunité**. Comme nous allons le constater, elles diffèrent grandement en quantité et en qualité par rapport au réseau développé jusqu’à présent.

Les observations d’opportunité peuvent être décomposées en trois catégories. En premier lieu, elles proviennent de tous les objets connectés mesurant des variables météorologiques d’intérêt, comme ceux de capteurs embarqués dans les véhicules, de stations météorologiques agricoles connectées ou de stations météorologiques personnelles connectées. Ces dernières sont rassemblées dans la littérature sous le nom de *personal weather stations* (PWS) ou *citizen weather stations* (CWS). En second lieu, les observations d’op-

portunité proviennent d'objets connectés fournissant indirectement une information météorologique, comme les capteurs embarqués par des smartphones, des tablettes ou des caméras, dont il est possible d'en restituer une observation météorologique telle que la pression atmosphérique. Enfin, un dernier type d'observations d'opportunité provient des réseaux sociaux. On y trouve les observations participatives, par exemple déposées sur l'application mobile de Météo-France. À titre d'exemple, 46 729 observations humaines du temps présent ont été réalisées en 24 h sur l'application le 7 août 2018 durant une dégradation orageuse de grande ampleur (Bernède, 2022). S'ajoutent aux observations participatives tous types d'observations postés sur les réseaux sociaux tels que les observations de grêle sur Twitter.

1.1.3.1 Les PWS : une panoplie de réseaux amateurs à semi-professionnels

Les PWS correspondent à des stations météorologiques vendues pour les particuliers par divers constructeurs et connectées à Internet. Par extension, les PWS incluent toutes les stations de mesures non conventionnelles dont les métadonnées ne permettent pas de qualifier la qualité de la station. A fortiori, la qualité de ces stations est variable, allant de stations basiques de faible qualité à des stations semi-professionnelles mesurant plusieurs dizaines de variables (Bell *et al.*, 2015). Certains fournisseurs de données permettent à quiconque de télécharger directement les observations émises par ces stations, gratuitement et en temps quasi réel via une API (*application programming interface*). Parmi les différents réseaux de stations d'opportunité, on retrouve :

- Des réseaux de stations liés à un fournisseur de stations privées. Les principaux réseaux, avec leurs nombres de stations vendues dans le monde, sont : Netatmo (> 300 000), Davis Instruments ($\approx 42\,800$) ou WeatherFlow (≈ 5000). Ils sont complétés par des réseaux agricoles, dont la liste non exhaustive contient DEMETER (≈ 1000), DTN ($\approx 26\,000$), Pessl Instruments ($\approx 60\,000$), Sencrop ($\approx 13\,000$), ou Weenat (≈ 4000).
- Des réseaux de stations déployés au sein d'une métropole, comme le réseau CAPI-TOUL (Masson *et al.*, 2008) à Toulouse, ou le réseau MUSTARD à Dijon.
- Des réseaux de stations déployés par des associations, ou regroupant des stations de réseaux amateurs. On retrouve parmi eux (i) le réseau WOW (*weather observations website* du Met Office (<https://wow.metoffice.gov.uk/>) qui rassemble différents réseaux de stations, dont certaines proviennent de particuliers, principalement basés au Royaume-Uni; (ii) le réseau StatIC (<https://www.infoclimat.fr/stations/static.php>) géré par l'association Infoclimat, rassemblant des stations de réseaux amateurs et de passionnés depuis 2007. Ce dernier est composé de stations dites semi-professionnelles, dont la majorité sont des Davis Vantage Pro2. Depuis 2021, l'association met en place un portail d'accès aux données, permettant une utilisation gratuite; et (iii) les stations développées par des associations sportives, telles que la Fédération Française de Vol Libre (<https://www.balisemeteo.com/>), les sports nautiques avec plus de 150 anémomètres (<https://www.winds-up.com/>), jusqu'aux associations territoriales telles que le réseau Météo-Bretagne (<https://www.asso-mb.fr/>), ou le réseau d'observations météorologiques du massif alpin (ROMMA, <https://www.romma.fr/>).

Le réseau d'observations Netatmo s'affirme comme étant le plus grand réseau mondial de stations météorologiques connectées, avec plus de 50 000 stations en France. Netatmo

est une entreprise française, spécialisée dans les objets connectés, fondée en 2011. La station météorologique connectée de Netatmo possède par défaut deux modules, un extérieur (*outdoor*) mesurant la température et l’humidité relative et un module intérieur (*indoor*), mesurant la température, l’humidité relative, la pression, le bruit et la concentration en CO_2 . Concernant le module intérieur, seule la mesure de pression est accessible via l’API Netatmo. Deux modules complémentaires sont vendus par l’entreprise : un pluviomètre et un anémomètre. Contrairement aux mesures conventionnelles, la fréquence de mesure est de 5 min pour tous les capteurs, avec un cycle de démarrage qui dépend de chaque station. L’acquisition d’une mesure aux heures pleines n’est donc pas garantie. Les caractéristiques techniques des capteurs embarqués dans les stations Netatmo sont définies dans le Tableau 1.3.

TABLEAU 1.3 – Caractéristiques des capteurs composant la station météo intelligente Netatmo. (*) Le prix est celui identifié sur le site du constructeur au 1^{er} février 2024.

	Plage de mesure	Précision	Résolution	Prix*
Température	-40 à 65 °C	± 0.3 °C	0.1 °C	
Humidité relative	0 à 100 %	± 3 %	1 %	190 €
Pression	260 à 1260 hPa	± 1 hPa	0.1 hPa	
Vent	0 à 160 km h ⁻¹	1.8 km h ⁻¹	1 km h ⁻¹	110 €
Précipitation	0.2 à 150 mm h ⁻¹	1 mm h ⁻¹	0.1 mm h ⁻¹	80 €

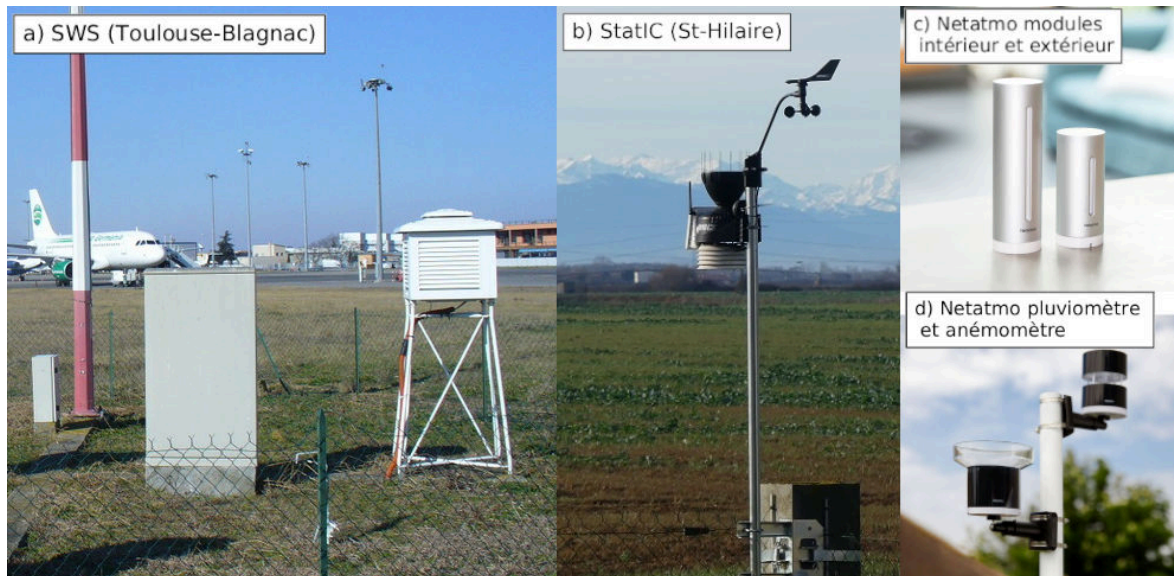


FIGURE 1.4 – Exemples de stations météorologiques : (a) station RADOME avec abri Stevenson à l’intérieur de l’aéroport Toulouse-Blagnac, (b) station StatIC à St-Hilaire équipé d’une Davis Instruments Vantage Pro2, (c) station Netatmo avec module intérieur (pression) et extérieur (température et humidité relative), et (d) pluviomètre et anémomètre.

Depuis 2022, l’entreprise Netatmo propose un abri pour le module extérieur. Malheureusement, cet abri, comparé à un abri Davis 7714 de référence, ne garantit pas un équilibre entre l’air intérieur et extérieur du module, ce qui pourrait être due à son étroitesse, le type de matériaux, son nombre d’ouvertures, ou simplement sa conception. Certains propriétaires mettent leurs modules extérieurs dans des abris homologués, mais cela semble rare d’après les photos d’installations auxquelles nous avons eu accès, et aucune

métadonnée ne permet de le savoir.

La transmission des données entre les modules se fait par ondes radio. Pour la transmission des données, le module intérieur, doit être relié par Wi-Fi à un point d'accès Internet. La visualisation des données se fait exclusivement par l'application mobile ou le site Internet du constructeur. Lors de l'installation, l'application mobile positionne la station dans l'espace grâce au GPS intégré au smartphone, fournissant ainsi la latitude, la longitude et l'altitude de la station. Le propriétaire peut, à tout moment, changer la position de la station. Si l'application mobile n'a pas accès au positionnement GPS, la position est attribuée d'après l'adresse IP sur le réseau Wi-Fi, ainsi plusieurs stations peuvent être placées aux mêmes coordonnées. L'information sur l'altitude de la station influence le calcul de la pression : si elle est fournie, l'observation de pression émise correspond à une pression réduite au niveau de la mer calculée à partir de l'équation hydrostatique, sinon, la réduction n'est pas faite. La qualité de l'observation de pression dépend de la fiabilité de son altitude. Un décalage d'1 m entre l'altitude fournie et l'altitude réelle entraîne un décalage systématique d'environ 0.1 hPa avec la pression qui est observée à l'altitude fournie. Par ailleurs, seul l'utilisateur est en mesure de réaliser un étalonnage du capteur de pression, ce qui peut expliquer une dérive de la mesure dans le temps par rapport à une mesure d'un baromètre de référence.

Pour se rendre compte de la forte proportion de stations Netatmo installées par rapport au réseau conventionnel, la figure 1.5 illustre l'ensemble des observations de température dans la région toulousaine (sur un domaine de 36 km × 45 km), provenant de différents réseaux. On y dénombre 4 stations conventionnelles et 886 stations Netatmo. Malgré leur densité, la qualité des observations Netatmo doit être évaluée, par exemple en les comparant à des stations de référence. En zone urbaine, peu de sites de mesure peuvent être de classe 1, du fait de la densité du bâti. À Toulouse, le réseau CAPITOUL, bien que réalisant des mesures à 5 m, fait office de stations de référence urbaine. L'ensemble du réseau CAPITOUL s'accorde sur une valeur de température homogène autour de 28 à 30 °C, alors que les observations Netatmo sont plus disparates avec 90 % des valeurs entre 28.0 et 32.4 °C et des extremums à 15.7 °C et 40.8 °C (figure 1.5). Il convient alors de chercher à comprendre d'où proviennent ces différences.

1.1.3.2 Qualité des observations de température et d'humidité relative des PWS

La qualité des capteurs des stations Netatmo a été évaluée en laboratoire (Météo-France, 2017 ; Büchau, 2018 ; Varentsov *et al.*, 2020). Des tests réalisés par Météo-France constatent des erreurs de température de 0.1 à 0.4 °C vis-à-vis d'un capteur de référence, incertitude proche de l'intervalle annoncé par le constructeur (Tableau 1.3). Pour l'humidité relative, la précision des capteurs annoncée par le constructeur est atteinte hormis à très basse humidité. Néanmoins, les observations Netatmo du module extérieur sont sujettes à des erreurs d'observation qui sont propres à ces stations. La compréhension de ces erreurs permet ensuite de réaliser le traitement le plus adéquat.

1.1.3.2.1 Erreur d'instrumentation et de représentativité

À cause de leur coque et leur ventilation de faible qualité, les capteurs sont sujets à des biais en température. Pour comprendre physiquement ces biais, Büchau (2018) a réalisé un bilan d'énergie idéalisé du module de mesure du capteur de température. Ce bilan

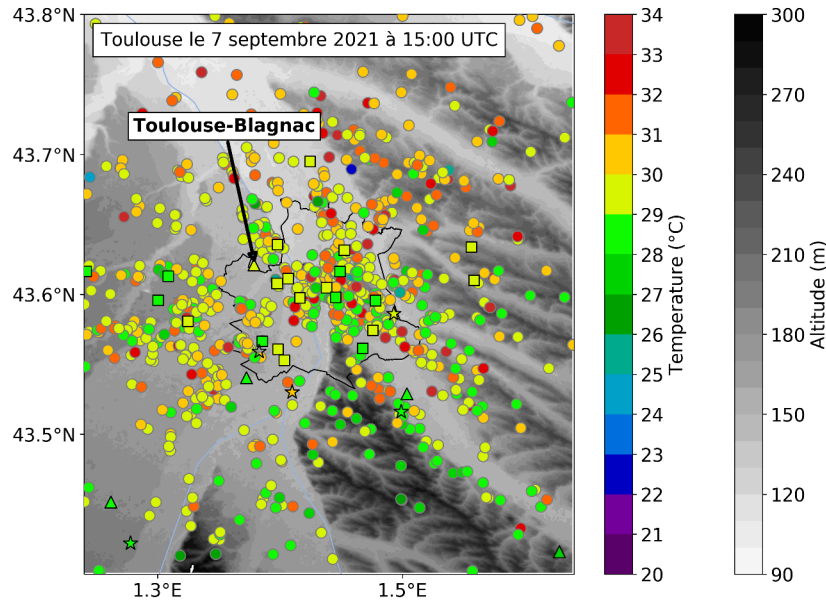


FIGURE 1.5 – Réseau d’observations de température à Toulouse le 7 septembre 2021 à 15:00 UTC. Avec 886 observations Netatmo brutes (en rond), six observations conventionnelles (en triangle), quatre observations du réseau StatIC (en étoile) et 21 observations du réseau Toulouse-Métropole (en rectangle). Le fond de carte représente le relief de la région.

d’énergie prend en compte les différents rayonnements de grandes et petites longueurs d’onde entrant et sortant et les flux de chaleur sensible et latent. Il estime l’albédo du module extérieur à 0.50, alors qu’un abri conventionnel aurait un albédo de 0.75 à 0.85 suivant son état de vieillissement (Perrin de Brichambaut et Leroy, 1995). Il conclut que le rayonnement solaire entrant est le facteur dominant qui génère la surchauffe en cas d’exposition importante. De plus, il estime que la température mesurée par une station Netatmo se rapproche d’une moyenne entre la température à 2 m de hauteur et la température au sol. En complément, Varentsov *et al.* (2020) ont réalisé une étude idéalisée en plaçant dix stations Netatmo dans des configurations particulières que l’on retrouve sur l’ensemble du réseau installé : situées proches d’un mur, sous un arbre ou en hauteur. L’ensemble des configurations sont décrites dans la figure 1.6. Les biais de température dépendent de la position du capteur : l’exposition contre un mur surexpose les masses d’airs qui surchauffent par rapport à l’air environnant, spécialement le matin pour la station exposée à l’est et l’après-midi pour celle exposée ouest. La station exposée sud possède un biais montant jusqu’à 6 °C. Seules les stations respectivement à 15 m, ombragée, protégée dans l’abri Stevenson ou sous l’arbre ont des biais inférieurs à 1 °C. Dans cette dernière configuration, la température de nuit a tendance à être sous-estimée. L’erreur d’observation pour une station ombragée ne provient pas d’une erreur instrumentale, mais de représentativité par rapport à une station à 2 m et sous abri Stevenson, configuration considérée comme la référence. Autrement dit, la station ombragée n’est pas représentative de la masse d’air sur une zone géographique étendue. Varentsov *et al.* (2020) montrent ainsi que l’emplacement d’une station Netatmo influence la qualité de l’observation de température. Des comportements similaires sont attendus pour les observations d’humidité relative des stations Netatmo.

1.1.3.2.2 Décalage temporel de la mesure

À cause de son assemblage, le capteur de température et d’humidité relative prend un

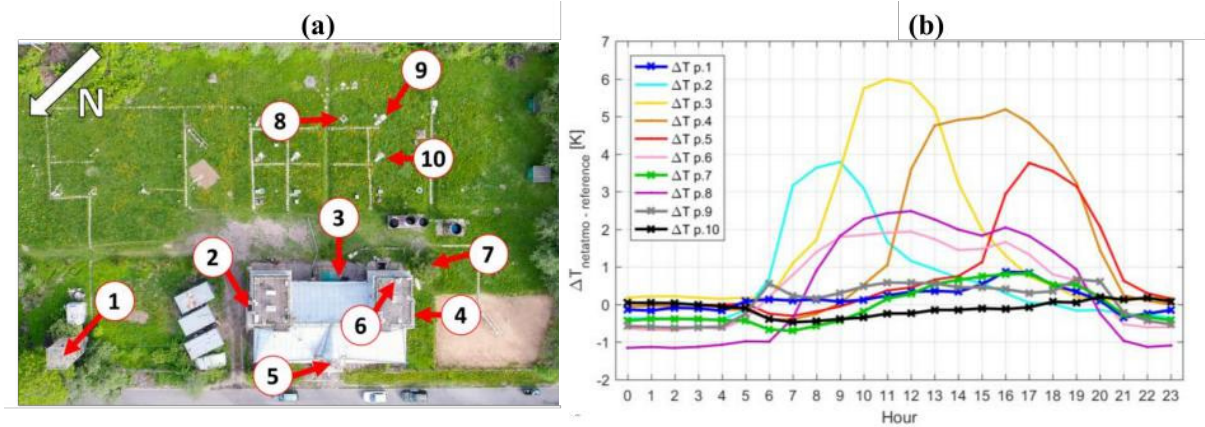


FIGURE 1.6 – Tests des positionnements des stations Netatmo. (a) Schéma de l'emplacement des stations Netatmo et (b) variation diurne du biais de la température moyenne (ΔT Netatmo-Reference) pour les dix modules extérieurs. Les stations sont positionnées dans les configurations suivantes : (1) à 15 m de hauteur, (2–5) aux quatre expositions (nord, sud, est, ouest) contre un mur, (6) sur un toit, (7) sous un arbre, (8) sur un pieu métallique à 2 m, (9) côté ombragé et (10) à l'intérieur de l'abri Stevenson. Figures issues de (Varentsov *et al.*, 2020).

certain temps à mesurer les brusques changements d'air environnant (Bell *et al.*, 2015; Meier *et al.*, 2017; Büchau, 2018). Büchau (2018) montre qu'il y a en moyenne 15 à 20 minutes de décalage temporel entre la température de l'air et sa mesure effective.

Il est donc obligatoire pour utiliser l'information en température et en humidité relative de passer par une technique de correction ou de contrôle qualité (Coney *et al.*, 2022).

1.1.3.3 Traitement des observations d'opportunité

Il existe une panoplie de traitements possibles qui s'adapte aux types de données et à l'utilisation souhaitée. Le tout dépend de l'échelle spatiale qu'on étudie et des contraintes temporelles (c'est-à-dire en temps réel ou pour une étude a posteriori). Deux stratégies de traitement des observations sont réalisées : une correction des observations ou un contrôle de qualité des observations.

Concernant les méthodes de correction, elles concernent notamment les observations de température et d'humidité relative (Clark *et al.*, 2018; Cornes *et al.*, 2019; Mandement et Caumont, 2020; Beele *et al.*, 2022; Sgoff *et al.*, 2022). Nous reviendrons dans le chapitre 4 sur la correction de biais des observations d'opportunité en température et en humidité relative.

L'objectif du contrôle de qualité est de rejeter les observations considérées anormales. Pour les qualifier d'anormales, il est nécessaire de comparer les observations à une référence. Plusieurs types de références sont possibles :

- Des seuils de références climatologiques ;
- Des observations conventionnelles voisines ;
- Des observations voisines du même réseau, par vérifications mutuelles.
- Par une ébauche ou une analyse ;

Ainsi, pour chaque observation PWS notée \mathcal{O}_{PWS} , on peut définir un seuil de rejet α , tel que :

$$|\mathcal{O}_{PWS} - \mathcal{O}_{Ref}| > \alpha \tilde{\sigma}_{Ref}, \quad (1.1)$$

Avec $\tilde{\sigma}_{Ref}$ une estimation de l'écart type des observations voisines et \mathcal{O}_{Ref} qui peut être défini comme la moyenne ou la médiane des observations voisines. Pour plus de précision, \mathcal{O}_{Ref} peut aussi être défini comme l'interpolation des observations voisines à la position de la PWS. Ce seuil de rejet revient à réaliser un test de Student, en prenant pour hypothèse la normalité de la distribution des observations.

Les premières procédures de contrôle de qualité (Chapman *et al.*, 2017; De Vos *et al.*, 2017) sont conçues pour les mesures de température et d'humidité relative, et sont faites pour fonctionner sans l'aide de données supplémentaires de référence. Meier *et al.* (2017) développent par la suite un contrôle de qualité plus complet pour les observations de température, qui se décompose en quatre étapes :

- Contrôle des métadonnées et du nombre de mesures sur le dernier jour et le dernier mois;
- Filtre des stations qui mesurent une température constante (par ex. les stations d'intérieur), suivant la moyenne et l'écart-type de leur température minimale journalière;
- Filtre lié au rayonnement solaire incident. S'il y a une corrélation positive et significative entre la température et le rayonnement incident, on rejette l'observation;
- Filtre sur la différence avec un réseau de référence. Si l'écart dépasse de trois fois l'écart-type des stations de référence voisines, on rejette l'observation.

Napoly *et al.* (2018) reprendront les travaux de Meier *et al.* (2017), en créant une procédure de contrôle de qualité, sans réseau de référence, mise en libre accès sous le nom de CrowdQC (Grassmann *et al.*, 2018). Cette dernière est reprise dans un grand nombre d'études (par ex., Feichtinger *et al.*, 2020; Fenner *et al.*, 2021; Potgieter *et al.*, 2021; Zumwald *et al.*, 2021, et à grande échelle par Venter *et al.*, 2020).

Néanmoins, après l'utilisation de CrowdQC des observations comportant des erreurs liées au rayonnement solaire incident subsistent (Feichtinger *et al.*, 2020). D'autres auteurs par la suite ont essayé d'améliorer l'étape de vérification mutuelle (Nipen *et al.*, 2020) et du choix du seuil de rejet des observations (Mandement et Caumont, 2020).

Nipen *et al.* (2020) proposent une procédure TITAN en libre accès (<https://github.com/metno/TITAN/wiki/SCT>), composé de sous fonctions utilisables séparément. La nouveauté provient d'un contrôle de cohérence spatiale (SCT, *spatial consistency check*). Le SCT est une procédure itérative. Pour chaque observation PWS, une estimation de l'observation par interpolation optimale des observations voisines est réalisée et un rejet par seuil est appliqué.

Mandement et Caumont (2020) ont quant à eux développé une méthode utilisant un seuil de rejet adaptatif. D'abord, une erreur associée à chaque observation PWS est calculée, définie par l'écart de l'observation aux stations de référence. Une liste de seuils à tester est alors définie. Pour chaque seuil, les observations PWS comportant une erreur supérieure au seuil sont rejetées. Puis, pour chaque observation de référence, son estimation est réalisée, d'une part avec les observations de référence, et d'autre part avec les observations de référence et les observations PWS filtrées simultanément. La différence entre l'observation et son estimation, avec et sans les PWS, correspond au gain apporté par les observations PWS pour chaque seuil testé. Le seuil retenu est celui maximisant le gain apporté par les observations PWS. Avec un seuil bas, proche de zéro, les observations de référence suffisent et la majorité des observations PWS sont rejetées. Avec un seuil élevé, les observations PWS apportent une plus-value et plus de stations PWS sont

TABLEAU 1.4 – Présentation des quatre principaux contrôles de qualité en température et humidité des observations PWS

		Meier <i>et al.</i> (2017)	TITAN (Ni-pen <i>et al.</i> , 2020)	Mandement et Caumont (2020)	CrowdQC+ (Fenner <i>et al.</i> , 2021)
	Descriptif	Référence	Cohérence spatiale	Seuil adaptatif	CrowdQC avancé
Tests	Metadonnées	✓	✓	✓	✓
	Doublons	Lon/Lat	Lon/Lat	Lon/Lat	Lon/Lat
	Interpolation temporelle	✓		✓	✓
	Test d'altitude		✓	✓ et correction	✓
Correction			Décalage temporel	Sur 24 h par rapport au réseau conventionnel	Décalage temporel
QC	Base de l'interpolation	Linéaire	Optimale	Linéaire	Linéaire
	Réseau de référence	SWS puis PWS	PWS	SWS	PWS
	Seuil de rejet	Fixe	Fixe	Adaptif	Fixe
Suppl.	Stations d'intérieur	✓ avec l'analyse de la TN	✓ avec <i>Buddy check</i>		✓ avec <i>Buddy check</i>
	Surchauffe diurne	✓ si erreur corrélée au rayonnement			✓ avec un test de Student

conservées. Ce seuil optimal minimise l'erreur de l'ajout des observations et maximise la plus-value en même temps.

Le projet *iOBS*⁵ a permis de tester de nouvelles procédures de contrôle de qualité. Une première concerne l'utilisation d'une méthode de krigeage pour interpoler les observations. Une méthode de classification des stations est testée, consistant à regrouper les stations par classes, et permet de filtrer des groupes de stations incohérents. Une méthode similaire est proposée par Marques (2023) en utilisant l'information temporelle et en réalisant une classification de courbes (*curve clustering*).

Fenner *et al.*, 2021 proposent d'améliorer CrowdQC, nommée CrowdQC+, avec notamment l'ajout de filtres sur les métadonnées et un contrôle de vérification mutuelle. S'y ajoute une correction pour prendre en compte le retard temporel des observations, lié à la mauvaise ventilation du module de mesure. Les quatre principaux contrôles de qualité sont résumés dans le Tableau 1.4.

La différence principale entre ces méthodes réside dans le type de référence utilisée. Pour disqualifier une observation, les algorithmes TITAN et CrowdQC+ se servent des observations PWS voisines, alors que MC2020 utilise des observations conventionnelles

5. <https://wiki.neic.no/wiki/IOBS>

voisines. L'enjeu est donc le suivant : le réseau de référence est-il capable d'observer un phénomène météorologique de fine échelle, détecté par les observations PWS ? Dans le cas où la référence n'y parvient pas, le risque est qu'elle puisse rejeter à tort les observations PWS. Un réseau de référence le plus dense possible est donc souhaitable pour contrôler la qualité des PWS, en utilisant les stations voisines du réseau PWS lui-même, comme le font TITAN et CrowdQC+. Enfin, il est aussi possible de prendre le modèle comme référence, en utilisant l'analyse ou une prévision à courte échéance, ce dernier étant capable de résoudre des phénomènes de fine échelle. Une comparaison des différentes stratégies est formulée dans le chapitre 4.

Les observations de pression comportent, elles aussi, un biais, qui correspond au premier ordre à un mauvais étalonnage de l'altitude de la station. Il est donc nécessaire de soustraire une erreur systématique (*offset*), c'est-à-dire une moyenne de la mesure de stations de référence sur six heures pour Clark *et al.*, 2018, et sur 24 heures pour Mandement et Caumont, 2020. De plus, certaines observations de pression restent incohérentes, et des traitements ont été développés pour rejeter ces observations anormales, comme précédemment, à l'aide d'un écart à une référence. McNicholas et Mass, 2018a créent une référence par interpolation de stations conventionnelles et rejettent les stations s'écartant trop de la référence à hauteur de quatre écart-types. Cette idée est reprise dans l'algorithme TITAN (Nipen *et al.*, 2020) et celui de Mandement et Caumont, 2020 décrit précédemment.

Synthèse

Les réseaux d'observation au sol ont été développés par différents acteurs, du niveau mondial au niveau local. Au niveau mondial, il existe une méthodologie, régie par l'OMM, garantissant un réseau d'observations homogène et travaillant à une meilleure répartition des observations sur tous les continents. Les réseaux d'observation en France sont de diverses origines, et un classement de leur qualité est réalisé pour les stations météorologiques standards. Au niveau local, des stations météorologiques personnelles complètent les réseaux standards, d'une qualité hétérogène, mais d'une plus grande densité. Différents traitements ont été proposés pour en tirer profit. Se pose alors la question : le réseau d'observations au sol actuel permet-il de détecter des phénomènes de méso-échelle ? Cette question sera traitée dans la section 1.3. Intéressons-nous avant aux possibilités qu'offrent les modèles de prévision numérique du temps régionaux.

1.2 Modèle de prévision numérique du temps régional

1.2.1 Équilibre atmosphérique et résolution

Les mouvements dans l'atmosphère sont régis par un ensemble d'équations, connues sous le nom d'équations de Navier-Stokes. Les équations de Navier-Stokes correspondent à des équations aux dérivées partielles non linéaires qui décrivent le mouvement des fluides newtoniens. Elles sont basées sur les équations d'Euler (EE) et sont les équations de base de la mécanique des fluides pour les fluides compressibles.

Ce système d'équations n'a pas de solution analytique dans le cas général. Seules les méthodes numériques discrètes permettent de calculer des solutions approchées. L'ensemble des méthodes numériques reposent donc sur la notion de discrétisation spatio-temporelle des champs décrivant l'état du fluide (Malardel, 2009). De par cette discrétisation, les équations d'évolutions de l'atmosphère peuvent être écrites sous la forme :

$$\frac{dX}{dt} = \mathcal{A} + \mathcal{F} \quad (1.2)$$

Avec X une variable pronostique, dont l'on veut suivre l'évolution ; \mathcal{A} représentant les effets qui sont explicitement représentés, souvent appelés termes adiabatiques ou noyau dynamique du modèle ; et \mathcal{F} , représentant les effets sous-mailles, appelés termes diabatiques ou physiqe du modèle.

Il existe différents types de modèles atmosphériques suivant leur couverture spatiale. Si le modèle résout le système d'équations sur tout le globe, il est dit global. S'il modélise seulement une partie restreinte de l'atmosphère à une résolution plus fine, on parle de modèle à aire limitée (LAM, *Limited Area Model*) ou modèle régional. Les modèles à aire limitée ont besoin de conditions aux limites latérales (LBC, *Lateral Boundary Conditions*) qui sont obtenues grâce à un modèle global.

Certaines hypothèses, admises pour simplifier le système d'équations pour les modèles globaux, ne sont plus valides pour les modèles à aire limitée. En effet, les LAM sont généralement "non-hydrostatiques" (NH), c'est-à-dire qu'ils ne font pas l'approximation hydrostatique qui consiste à négliger le terme d'accélération verticale devant les autres termes de l'équation du mouvement vertical, et "élastiques", pour une meilleure représentation des grandes échelles, comme des ondes de Rossby. Des modèles NH à aire limitée permettent de résoudre la convection profonde sans passer par une paramétrisation de celle-ci.

En Europe, les modèles opérationnels à aire limitée sont principalement développés par des consortiums de services météorologiques nationaux, tels que HIRLAM, ALADIN et LACE. En 2020, ces consortiums se sont unis pour former le consortium ACCORD (*A Consortium for COnvection-scale modelling Research and Development*), rassemblant plus de 26 services météorologiques nationaux euro-méditerranéens. L'intention de ce consortium est d'unir les capacités des pays membres, leurs expériences, leurs connaissances et leurs ressources pour exploiter des modèles communs de prévision numérique du temps (PNT). Le consortium ACCORD maintient et développe trois principaux modèles, que sont AROME, HARMONIE-AROME, et ALARO. Il existe plus de 60 configurations opérationnelles de ces trois modèles, déclinées par les pays membres du consortium.

D'autres modèles à aire limitée opérationnels sont développés au sein des services météorologiques mondiaux. Ces modèles peuvent être caractérisés par leur résolution ho-

horizontale et leur nombre de niveaux verticaux. Parmi eux, on note :

- COSMO, avec une résolution de 2.8 km et 50 niveaux, utilisé par le service météorologique allemand (DWD). Le modèle COSMO est aussi développé à Météo-Suisse avec une résolution de 2.2 km et 60 niveaux (Baldauf *et al.*, 2011) ;
- HRRR, avec une résolution de 3 km et 51 niveaux, développé par le centre américain (NOAA) utilisant le modèle WRF en accès libre (Benjamin *et al.*, 2016) ;
- UKV, avec une résolution de 1.5 km et 70 niveaux, développé par le Met Office (Tang *et al.*, 2013) ;
- le MSM, avec une résolution de 5 km et 76 niveaux, et LFM, avec une résolution de 2 km et 58 niveaux, développés par l'agence météorologique japonaise (JMA).

1.2.2 Le modèle de prévision AROME-France

AROME-France est le modèle de prévision numérique du temps (NWP) à aire limitée développé par Météo-France (Fig. 1.7) et opérationnel depuis décembre 2008 (Seity *et al.*, 2011 ; Brousseau *et al.*, 2016). Il est couplé au modèle ARPEGE pour les conditions aux bords. Bien qu'AROME existe sous différentes versions, telles que AROME-OM pour les territoires d'outre-mer, AROME-PI pour la prévision immédiate, ainsi que des versions commerciales et militaires, nous nous référerons à AROME comme étant la version métropolitaine.

Le modèle AROME est discrétisé sur une maille dont les points de grille sont espacés horizontalement d'1.3 km et possède 90 niveaux verticaux partant de 5 m au-dessus du sol, jusqu'au niveau de pression 10 hPa. Les **variables pronostiques** correspondent aux variables directement calculées par le modèle. Elles décrivent l'état de l'atmosphère et sont régies par une équation différentielle impliquant leur dérivée temporelle. AROME est un modèle spectral, où la plupart des variables pronostiques sont basées sur une décomposition dans l'espace de Fourier. Les variables pronostiques d'AROME peuvent être séparées en deux catégories :

- Les variables dont les évolutions sont explicitement résolues par la dynamique du modèle : les deux composantes du vent U et V , la température T , la pression de surface hydrostatique π_s ; ainsi que deux variables non-hydrostatiques liées à la pression et à la divergence verticale de l'écoulement, \hat{q} et d .
- Les variables de la physique du modèle : l'énergie cinétique turbulente TKE et les contenus spécifiques en vapeur d'eau q_v , en pluie q_r , en neige q_s , en graupel q_g , en particules nuageuses q_c et en cristaux de glace q_i .

À partir des variables pronostiques sont calculées des **variables diagnostiques**, comme la température à 2 m.

Enfin, le modèle AROME présente une coordonnée verticale hybride en pression η , laquelle suit la surface sur les premiers niveaux (figure 1.8). Nous examinerons dans le chapitre 2, comment cette particularité influence l'assimilation des observations dans AROME.

1.2.2.1 Processus résolus et paramétrisés

Le noyau dynamique d'AROME est hérité des développements du modèle ALADIN (Bubnová *et al.*, 1995), dont il est issu. Le noyau dynamique se charge de la résolution des phénomènes dont l'échelle caractéristique est supérieure à la taille de la maille, tandis que

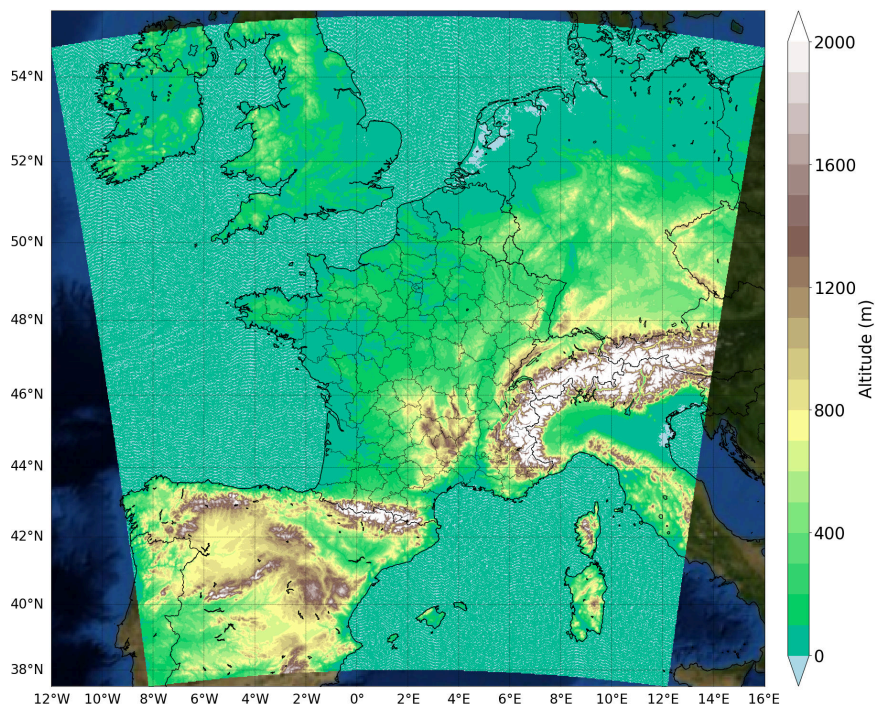


FIGURE 1.7 – Orographie du modèle AROME-France.

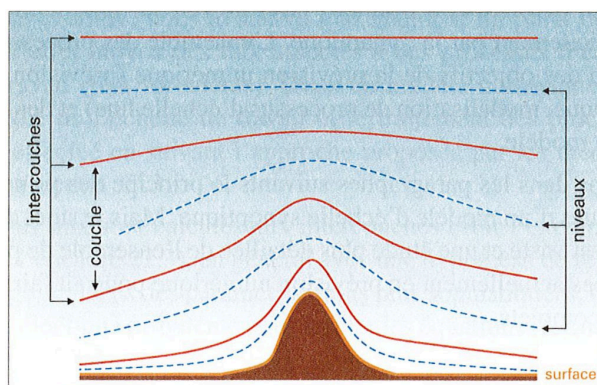


FIGURE 1.8 – Discretisation verticale dans le système de PNT AROME. Figure issue de Malardel (2009).

les autres phénomènes sont paramétrisés. Comme pour les autres modèles opérationnels, il est contraint par son coût (ou temps) de calcul. Ce coût est lui-même dépendant du pas de temps d'intégration du modèle, qui doit notamment respecter la contrainte de convergence de Courant-Friedrichs-Lewy (CFL) pour une résolution spatiale donnée. Pour affaiblir cette contrainte, et intégrer ainsi les équations avec un grand pas de temps, un schéma semi-implicite semi-lagrangien (SI-SL) est utilisé dans AROME. Le schéma semi-implicite se charge du traitement des termes des équations responsables de la propagation des ondes rapides dans l'atmosphère (Bénard *et al.*, 2010), tandis que le schéma semi-lagrangien traite les termes d'advection (Vána *et al.*, 2008). La combinaison de ces deux schémas permet ainsi d'utiliser un pas de temps de 50 s pour une résolution de 1.3 km. Les modèles n'utilisant pas ces schémas sont contraints d'utiliser des pas de temps environ dix fois plus faibles (Burgot, 2021).

Comme évoqué précédemment, AROME résout explicitement l'équation des mouvements verticaux sans faire l'hypothèse hydrostatique :

$$\gamma \frac{dw}{dt} = -\frac{\partial P}{\partial z} - \rho g, \quad (1.3)$$

avec w est la vitesse verticale, P est la pression, ρ est la masse volumique, et g est la constante de gravitation. La vitesse verticale est ainsi une variable pronostique du modèle. L'absence d'approximation hydrostatique est donc particulièrement pertinente lors de forts mouvements verticaux, comme on en trouve dans le cas de convection profonde. L'approximation hydrostatique est valide pour des phénomènes respectant :

$$\frac{H}{L} \ll 1, \quad (1.4)$$

c'est-à-dire tant que la taille horizontale du phénomène L est nettement supérieure à la taille verticale H . Si l'on prend la hauteur de la troposphère autour de 10 km pour H , la limite revient donc à L de l'ordre de 10 km. D'après Malardel (2022), la limite ne doit pas être comparée à la résolution du modèle, mais à sa **résolution effective**, à savoir la taille minimale des processus effectivement résolus. Dans AROME et IFS, cette résolution effective est autour de huit à dix fois la maille. La différence entre le modèle hydrostatique et non-hydrostatique est donc notable à partir de résolution avoisinant l'ordre de 1 km. Ce constat justifie l'intérêt de résoudre les équations sans faire l'approximation hydrostatique pour le modèle AROME dont la résolution horizontale est de 1.3 km. Il convient par ailleurs de souligner que la résolution du modèle dépend non seulement de la taille de sa maille, mais également de la troncature spectrale des variables.

1.2.2.2 Processus physiques représentés

Dans la partie physique du modèle sont paramétrés les processus qui jouent un rôle significatif pour la prévision, mais qui ne sont pas représentés dans la partie dynamique du modèle. Le but des paramétrisations n'est pas d'augmenter la résolution du modèle, mais d'évaluer statistiquement l'effet moyen du non résolu (le sous-maille), pour assurer la continuité entre les échelles. La partie physique traite alors conjointement des échanges avec les autres composantes du système Terre (par ex. avec le sol, l'océan, la biosphère), mais aussi du bilan radiatif, des changements d'état de l'eau. En d'autres termes, il s'agit des processus moyens qui ne sont pas directement résolus par la dynamique à l'échelle de la maille. La physique du modèle AROME est principalement issue du modèle de recherche MESO-NH.

Certains schémas sont particulièrement importants, tels que le schéma radiatif, le schéma de surface, le schéma de convection peu profonde, le schéma microphysique et le schéma de turbulence. Leurs caractéristiques sont résumées dans le Tableau 1.5.

A la création d'AROME en 2008, le premier niveau du modèle était à 17 m, ce qui l'éloignait des champs de surface. Masson et Seity (2009) ont ainsi développé un algorithme spécifique, nommé *Surface Boundary Layer* (SBL), permettant le diagnostic de la température à 2 m, l'humidité relative à 2 m et le vent à 10 m avec un schéma de turbulence 1D sur six niveaux entre la surface et le niveau le plus bas du modèle. Lors du passage à une résolution plus fine de 1.3 km et ses 90 niveaux dont le premier niveau à 5 m, il a été décidé de désactiver ce module (Brousseau *et al.*, 2016). Un opérateur est défini pour interpoler la température à 2 m, en utilisant le niveau le plus bas du modèle et la surface (Chapitre 2).

TABLEAU 1.5 – Principales paramétrisations physiques d’un modèle de PNT à aire limitée et transcription dans le cas d’AROME-France (Fischer *et al.*, 2018)

Paramétrisation	Schéma ou modèle	Référence
Rayonnement	RRTMG_LW_SW6 (schémas de rayonnement solaire et thermique, à respectivement six et seize bandes spectrales)	Mlawer <i>et al.</i> (1997) et Iacono <i>et al.</i> (2008)
Turbulence	CBR (schéma en énergie cinétique turbulente pronostique)	Bougeault et Lacarrère (1989) et Cuxart <i>et al.</i> (2000)
Microphysique	ICE3 (microphysique à un moment dont les variables sont les masses des gouttelettes nuageuses, de la pluie, des cristaux de glace primaire, des graupels et de la neige)	Pinty et Jabouille (1998)
Convection peu profonde	PMMC09 (schéma en flux de masse)	Pergaud <i>et al.</i> (2009)
Convection profonde	-	
Nuages	Schéma basé sur une représentation statistique de l’écart local à la saturation dans la maille	Pergaud <i>et al.</i> (2009)
Sédimentation des variables micro-physiques (pluie, neige, graupel, eau nuageuse)	Schéma de sédimentation statistique	Bouteloup <i>et al.</i> (2011)
Surface	SURFEX	Masson <i>et al.</i> (2013)

1.2.2.3 Couplage atmosphère/surface

Les échanges entre l’atmosphère et la surface se font via le modèle SURFEX. Le modèle reçoit les forçages atmosphériques, tels que la température, l’humidité, le vent, la pression, le rayonnement à courte et grande longueur d’onde, ainsi que les précipitations, et retourne depuis la surface les flux de chaleur sensible et latente, ainsi que les caractéristiques radiatives de surface. SURFEX représente la surface avec quatre tuiles, chacune dédiée à un type de surface et associée à un modèle de surface :

- une tuile "nature" associée au modèle Interaction Sol-Biosphère-Atmosphère, ISBA (Noilhan et Planton, 1989) ;
- une tuile "ville" avec le modèle Town Energy Balance (TEB) pour les surfaces anthropisées (Masson, 2000) ;
- une tuile "lac" avec le modèle FLake (Mironov *et al.*, 2010) ;
- et une tuile "mer/ocean" avec le modèle COARE 3.0 (Fairall *et al.*, 2003).

Un pourcentage de chacune des quatre tuiles est associé à chaque maille de la grille. Cette décomposition de la grille en tuiles est réalisée grâce à une base de données physiographiques de terrain appelée ECOCLIMAP. Le système SURFEX est connecté au modèle atmosphérique AROME. En entrée, chaque tuile reçoit le même forçage atmosphérique, et en retour, c’est la moyenne pondérée des flux des quatre tuiles qui est transférée au modèle atmosphérique. Il convient de noter que la tuile "nature" est elle-même décomposée en 12 sous-tuiles, dont le flux correspond à la moyenne des flux des sous-tuiles. Concrètement, la base ECOCLIMAP, dans sa version 1, a une résolution horizontale de 1 km et est composée de 215 types de couverts, correspondant à des surfaces dans lesquelles le couvert est considéré homogène. Ce sont ces couverts, associés à chaque maille, qui nous renseignent sur le pourcentage de tuiles. Cette décomposition est importante pour comprendre comment est calculer une t2m dans le modèle. Nous y reviendrons dans le chapitre 2.

Synthèse

Les modèles de PNT se déclinent de l’échelle globale à l’échelle régionale. AROME est le modèle à aire limitée développé et maintenu par Météo-France et ses partenaires. Les champs atmosphériques sont représentés par une grille d’1.3 km de résolution horizontale, où les équations de mécanique des fluides sont décomposés en une partie explicitement résolue (« dynamique du modèle ») et une partie intégrant des processus sous maille paramétrés (« physique du modèle »). En pratique, la complexité des modèles à aire limitée rend l’interprétation des champs délicate (Malardel, 2022). Prenons l’exemple de l’interprétation d’une valeur d’un champ de température du modèle AROME. La valeur d’un point de la grille représente la moyenne de la maille, plus la moyenne des processus non résolus explicitement à l’échelle de la maille et estimés par les paramétrisations du modèle. Enfin, il s’agira de garder en mémoire que s’ajoutent des hypothèses liées à la résolution des équations dues à la résolution en mode spectral, et avec l’utilisation d’un schéma SI-SL. Ainsi, malgré tous ces efforts, les champs du modèle comportent des erreurs de modélisation. Des post-traitements sont nécessaires pour corriger ponctuellement les champs du modèle au niveau des points d’observations au sol. Au regard de ces

éléments, une question se pose : quelles sont les capacités des modèles à aire limitée à simuler des phénomènes de méso-échelle ?

1.3 Traces au sol des phénomènes de méso-échelle : de l'observation au modèle

Les phénomènes de méso-échelle peuvent être décrits à la fois par des modèles de fine échelle comme AROME et grâce aux observations combinées de SWS et de PWS au sol. Plusieurs types de phénomènes seront étudiés. D'abord des phénomènes liés à des mouvements verticaux importants, tels que des fronts ou des dépressions de méso-échelle et des systèmes convectifs de méso-échelle. Puis suivront des phénomènes de basses couches au sein de la couche limite atmosphérique.

Des exemples seront présentés, comparant les performances du modèle AROME et les observations combinées des SWS et PWS, nommées SPWS. Dans le cas des prévisions AROME, on prendra l'échéance disponible la plus proche, généralement la prévision à 1 h considérée décorrélée de l'analyse. Pour ce qui est des observations personnelles, le traitement utilisé provient de Mandement et Caumont (2020). La fenêtre temporelle pour le traitement a été adaptée pour fonctionner en temps réel, c.-à-d. avec seulement des données antérieures⁶ à la mesure. Pour chaque phénomène décrit, une description des mécanismes physiques expliquant leur cycle de vie sera faite.

1.3.1 Minimum dépressionnaire et front

Cette première partie vise à décrire les processus mis en jeu dans l'atmosphère lors du cycle de vie des dépressions et des fronts, que nous visualiserons après avec les observations et le modèle.

1.3.1.1 Cyclogénèse et frontogénèse

La naissance d'une dépression répond à un déséquilibre atmosphérique et se forme à nos latitudes au niveau d'une zone **barocline** (Malardel, 2009). Une zone barocline de grande échelle est caractérisée par un fort gradient méridien de température, entre l'équateur et le pôle. La relation du vent thermique indique que la zone barocline est associée à un fort cisaillement vertical de vent zonal, c'est-à-dire une augmentation du vent d'est en ouest avec l'altitude. En pratique, la zone barocline donne naissance à un fort courant de vent zonal au niveau de la tropopause⁷ à moyenne latitude, nommée **courant jet**. C'est le courant jet qui favorise le passage des dépressions, aussi nommées perturbations baroclines aux moyennes latitudes, formant le rail des dépressions. Les perturbations baroclines contribuent au mélange méridien entre l'équateur et les pôles, atténuant le gradient méridien moyen de température. Une perturbation barocline est généralement le résultat de l'interaction entre un précurseur d'altitude et un précurseur de basses couches dans une configuration favorable. Ce que l'on appelle précurseur de

6. Pour les observations de température et d'humidité relative, la correction de biais est basée sur les données des 24 dernières heures, avec un pas de temps de dix minutes, et le contrôle de qualité est effectué au dernier pas de temps. Pour les observations de pression, la correction de biais et le contrôle de qualité sont quant à eux basés sur les trois dernières heures à pas de temps dix minutes.

7. Limite supérieure de la troposphère, entre 10 et 12 km d'altitude.

basses couches correspond à une zone dépressionnaire, associée à une anomalie chaude (par ex. le courant océanique chaud ou la libération de chaleur latente depuis un système précipitant). Parallèlement, on appelle précurseur d'altitude une zone de basse tropopause liée à une anomalie froide et un tourbillon positif. Le précurseur d'altitude peut venir, par exemple, d'une ancienne perturbation barocline. Une anomalie de surface associée à un positionnement par rapport au jet d'altitude, appelé « entrée droite, sortie gauche de jet », peut créer les conditions nécessaires à l'apparition de précurseur d'altitude.

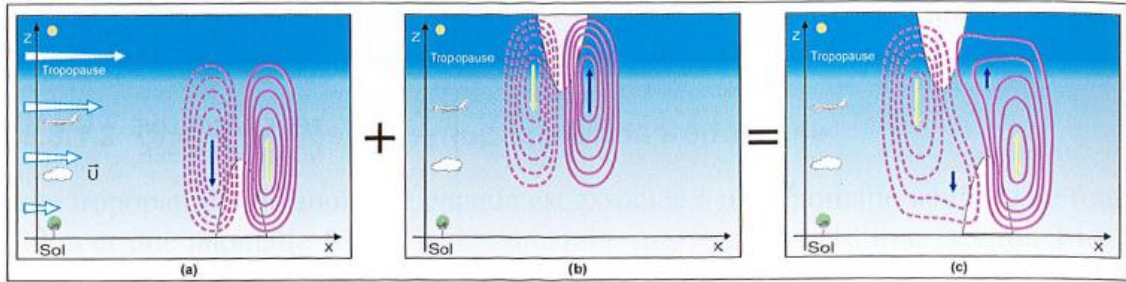


FIGURE 1.9 – Coupe verticale dans un plan ouest-est de la vitesse verticale associée à : (a) une anomalie de tourbillon de basses couches dans un jet d'ouest, (b) une anomalie de tourbillon d'altitude dans un jet d'ouest, (c) la superposition d'une anomalie de basses couches et d'une anomalie de tourbillon d'altitude décalée vers l'ouest par rapport à l'anomalie de basses couches. Figure issue de Malardel (2009).

Si l'anomalie de tourbillon positif en altitude est décalée en amont de celle de surface, alors les ascendances qu'elle induit augmentent le tourbillon au sol par étirement (figure 1.9). Le corollaire à la surface est également valable : l'anomalie de tourbillon positif au sol induit de la subsidence en amont, amplifiant l'anomalie en altitude par étirement. Cet effet de rétroaction positif est appelé développement barocline. Il est accompagné de l'amplification des anomalies de température et de pression en surface. Le creusement d'une dépression induit une amplification des vents qui tournent autour de la dépression selon la loi de Buys-Ballot, cette phase est appelée **cyclogénèse**. On peut classer les dépressions suivant leur vitesse de creusement, allant de dépressions modérées (moins de 5 hPa en 24 h), aux dépressions actives (de l'ordre de 10 hPa en 24 h), jusqu'aux bombes (plus de 24 hPa en 24 h).

Pour décrire les mécanismes de formation d'un front, on parle de **frontogénèse**. La frontogénèse est initiée par une déformation du vent géostrophique⁸ de surface. Cette déformation s'accompagne de la convergence du vent agéostrophique⁹ qui forme un fort gradient de température à la surface. Le fort gradient de température est à l'origine du terme de front, correspondant à une notion de discontinuité entre deux masses d'air. Cette zone de fort gradient de température de surface est associée à de forts tourbillons de basses couches, c'est-à-dire à une forte rotation du vent en surface, formant une anomalie de basses pressions. Cette continuité de basses pressions le long du front est associée à un **talweg**. Le fort tourbillon de surface se traduit par un important cisaillement de vent parallèle au front, que l'on identifie par un jet de basses couches parallèles au front en surface. La frontogénèse est le résultat d'une boucle de rétroaction positive entre l'augmentation du gradient de température et l'augmentation de la convergence de vent agéostrophique en surface.

8. Le vent géostrophique est défini comme la composante du vent dont la force de pression horizontale équilibre l'accélération d'inertie de Coriolis.

9. Le vent agéostrophique est l'écart entre le vent réel et le vent géostrophique.

Le comportement de méso-échelle d'un front est moins bien identifié. Il correspond à des structures complexes de couches nuageuses et d'irrégularité spatiale des précipitations au sein du système. Comme il n'existe pas de modèle conceptuel simple pour les expliquer, les observations de fine échelle sont les meilleurs moyens pour les décrire, grâce aux observations radar et à des avions instrumentés. Les observations d'opportunité représentent un moyen supplémentaire pour identifier les structures d'un front, et de son bon positionnement géographique dans le modèle.

1.3.1.2 Front du 5 octobre 2021

Le 5 octobre 2021, un flux synoptique zonal de moyenne troposphère se met en place et une dépression se positionne sur l'Angleterre. Un front froid s'étend au passage de la dépression avec un axe méridien sur le nord de la France. La figure 1.10 illustre le champ de pression réduit au niveau de la mer pour la prévision 1 h d'AROME, et pour les observations conventionnelles et personnelles à 07:00 UTC. Le minimum dépressionnaire simulé par AROME est en mer, et les valeurs observées sont conformes (par exemple, à Boulogne-sur-Mer avec 999.6 hPa contre 999.7 hPa dans AROME). Le front simulé par AROME, visible par une baisse de pression à son passage, s'étend de Boulogne-sur-mer au Mans. Il est visible par les observations standards de vent, où l'on observe la bascule du flux de sud au flux d'ouest à son passage. Par contre, le réseau d'observations standards de pression de surface ne permet pas de visualiser la baisse de pression. A contrario, le réseau d'observations personnelles, avec sa forte densité de stations, permet de vérifier le positionnement du front simulé par le modèle. Le modèle est légèrement en retard d'une dizaine de kilomètres, avec une valeur observée au Mans de 1005.2 hPa contre 1004.5 hPa pour AROME. Le réseau de PWS permet d'identifier le talweg qui s'étend juste en amont de la station du Mans. Cette anomalie de pression est cohérente avec la rotation de vent, du sud à l'est du Mans, passant sud-ouest au Mans. De plus, on remarque avec le réseau de PWS une anomalie de pression autour d'Évreux ville qui se détache dans le front. Cette anomalie de pression est aussi simulée par AROME, mais plus fine et plus méridionale. La synergie entre les observations est importante pour mieux appréhender les structures de fine échelle. Ici, l'apport des observations personnelles est net par rapport aux observations standards et informe sur les capacités d'AROME à bien simuler la méso-échelle du front. Nous verrons dans le chapitre 3 quel gain peut avoir l'assimilation des observations de pression des PWS pour la prévision du modèle AROME de ce front froid.

1.3.1.3 Dépression secondaire du 2 novembre 2023

Prenons maintenant l'exemple d'une dépression de méso-échelle, aussi appelée dans notre cas dépression secondaire. À la suite de la tempête Ciarán qui a traversé le nord de la France les 1^{er} et 2 novembre 2023, une dépression secondaire se creuse dans le golfe de Gascogne avant d'atteindre la côte le 3 novembre à 02:00 UTC (Fig. 1.11). Les observations standards de pression au sol ont une répartition spatiale irrégulière sur le territoire, mais quatre stations au sud du bassin d'Arcachon permettent de capturer des valeurs proches du minimum dépressionnaire à 982.5 hPa (Fig. 1.11b). On y voit la rotation du vent, avec des vitesses instantanées du vent atteignant les 156 kmh^{-1} à Biscarosse à 02:31 UTC. La direction du vent observé autour du minimum pointe vers l'intérieur de la dépression. Cet effet est dû aux effets de friction au sol, déviant le vent légèrement de la direction tangente aux isobares. Le réseau d'observations personnelles (Fig. 1.11c) en pression est plus dense, mais ne couvre pas une zone à l'est du minimum dépressionnaire

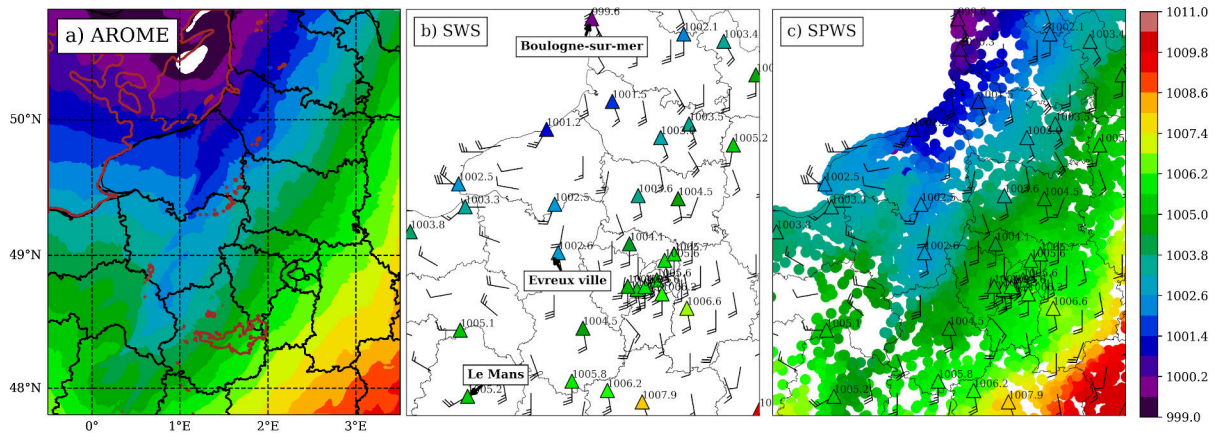


FIGURE 1.10 – Champ de pression réduite au niveau de la mer (en hPa) (a) prévu à 1 h d’échéance par AROME, (b) observé par le réseau d’observations standards et (c) observé par le réseau d’observations standards et personnelles pré-traitées, le 5 octobre 2021 à 07:00 UTC. Les triangles indiquent les observations SWS et les ronds indiquent les observations PWS. Les iso-lignes marron représentent le module de vent de surface à 10 m s^{-1} simulé par AROME. Les barbules (kt) représentent les observations de vent moyen sur 10 min du réseau d’observations standards.

correspondant à la forêt des Landes. Grâce au réseau de PWS, le creux dépressionnaire est établi à une échelle spatiale plus fine. Par ailleurs, on remarque que AROME simule un minimum dépressionnaire en retard et moins creux (+ 6 hPa). Dans le chapitre 3, nous nous demanderons si les observations de pression PWS peuvent permettre d’intensifier le creusement dépressionnaire dans le modèle AROME.

1.3.1.4 Mascaret atmosphérique du 18 juin 2022

Le 18 juin 2022 sur la côte fleurie en Normandie, un vent violent se lève à l’arrivée d’un stratocumulus volutus, faisant 1 mort, 5 blessés et 52 personnes choquées. Il s’agit d’un phénomène de propagation d’onde de gravité interne au sein d’une masse d’air stable et stratifiée (Mandement *et al.*, 2023). Ces ondes sont associées à la propagation d’un mascaret atmosphérique. Il s’agit d’un phénomène similaire au mascaret hydraulique que l’on retrouve dans les systèmes fluviaux. Durant le passage de cette onde, la température chute de 7°C et l’humidité relative monte de 20 % en 30 minutes à Saint-Gatien-des-Bois. La pression de surface quant à elle augmente de 2 hPa puis oscille (Fig. 1.12). Si l’on considère un seuil de 0.1 hPa pour la détection de l’onde, on remarque six périodes, avec une vitesse de phase de l’ordre de 15 m s^{-1} . Le réseau d’observations personnelles réalise des observations toutes les 10 min et ne peut donc pas capturer l’évolution temporelle de l’onde, contrairement aux stations d’observation standard procurant des données toutes les minutes. En revanche, la densité importante des stations personnelles permet de caractériser l’étendue spatiale de l’onde observée avec une hausse de 1 hPa sur 20 km (Fig. 1.13c). Cette onde est bien simulée à 1 h d’échéance par AROME (Fig. 1.13a) avec un décalage spatial de quelques kilomètres, une intensité plus marquée (hausse de plus de 2 hPa en son centre) et plus resserrée (s’étendant sur une dizaine de kilomètres). Cet exemple montre les capacités, indépendamment, du modèle AROME à simuler un phénomène d’échelle méso- γ (de 2 à 20 km), et du réseau d’observations personnelles à le visualiser.

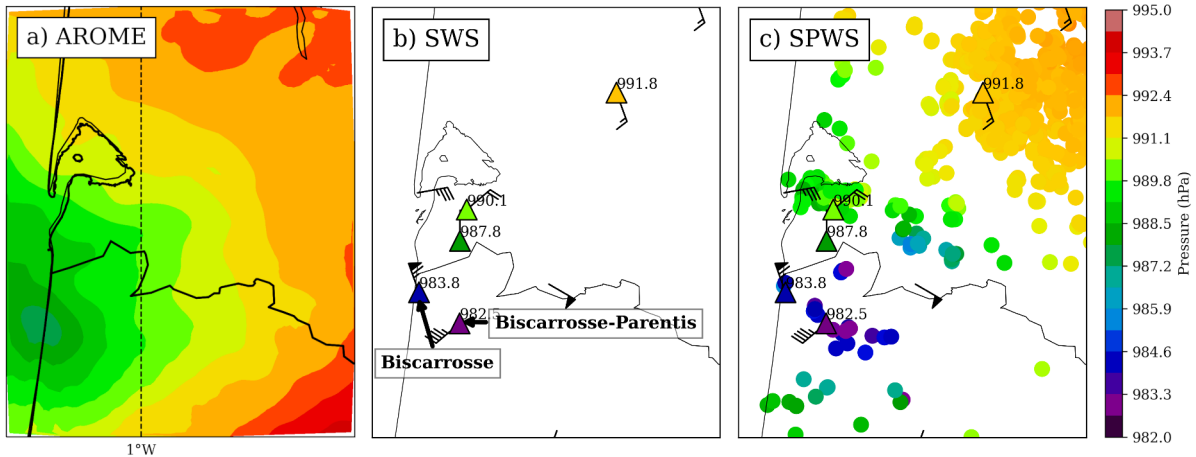
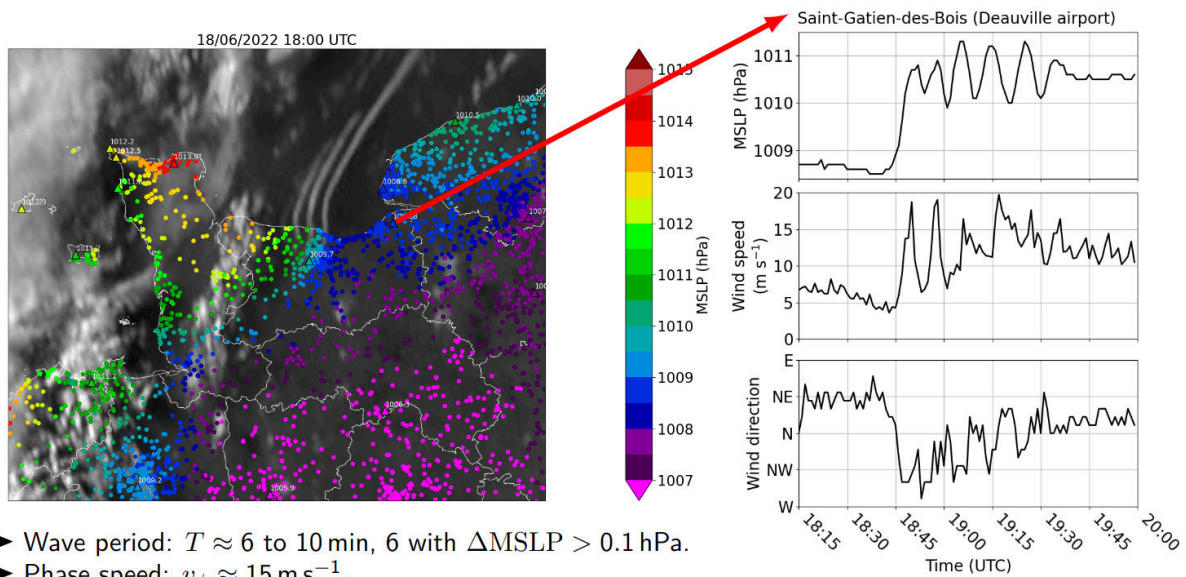


FIGURE 1.11 – Champ de pression réduite au niveau de la mer (a) simulé par la prévision 30 min d'AROME, (b) observé par le réseau d'observations standards et (c) observé par le réseau d'observations personnelles pré-traitées, le 3 novembre 2023 à 02:30 UTC. Les barbules représentent les observations de vent du réseau d'observations standards.



- Wave period: $T \approx 6$ to 10 min, 6 with $\Delta\text{MSLP} > 0.1$ hPa.
- Phase speed: $v_{\phi} \approx 15$ m s⁻¹.

FIGURE 1.12 – Passage du mascaret atmosphérique le 18 juin 2022 à 18:00 UTC. Le fond de l'image représente le canal haute résolution visible du satellite MSG-3. Les couleurs représentent les observations de pression réduite au niveau de la mer (MSLP), des stations personnelles (cercle) et standards (triangle). À droite, les séries temporelles de MSLP, vitesse et direction du vent à Saint-Gatien-des-Bois. Figure issue de Mandement *et al.* (2023).

1.3.2 Système convectif de méso-échelle

1.3.2.1 Formation des systèmes orageux

La convection est l'ensemble des processus caractérisant une masse d'air instable vis-à-vis de la flottabilité¹⁰. La convection est dite peu profonde lorsqu'elle n'est pas précipitante. En été, elle peut correspondre à des cumulus sans impacts au sol, qui sont qualifiés de « cumulus de beau temps ». La convection est dite profonde dès que les mouvements verticaux s'étendent sur toute la hauteur de la troposphère. La convection profonde se ca-

10. La force de flottabilité est la résultante de la force de gravité et de la force d'Archimède.

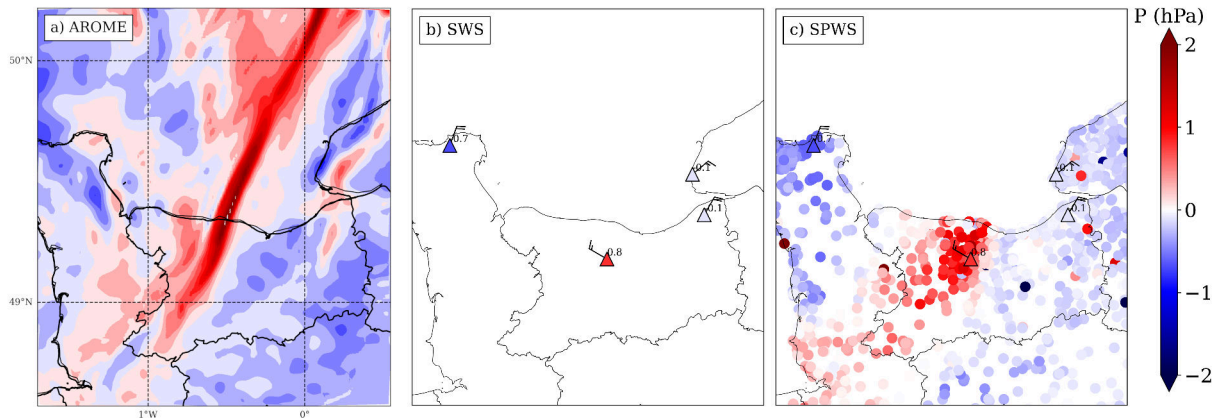


FIGURE 1.13 – Différence de pression réduite au niveau de la mer de 10 min (a) simulée par la prévision 1 h d'AROME, (b) observée par le réseau d'observations standards et (c) observée par le réseau d'observations personnelles pré-traitées, le 18 juin 2022 à 18:00 UTC. Les barbules représentent les observations de vent du réseau d'observations standards.

ractérisé par des nuages de type cumulonimbus, seuls nuages associés aux orages. L'impact des orages peut être multiple, pouvant provoquer des éclairs et de la grêle. La convection profonde peut donner lieu à d'autres types de vigilance tels que des avertissements de vent violent ou de pluies-inondation dans le cas d'épisode fortement précipitant (HPE).

L'initiation de la convection profonde nécessite trois conditions. La première condition consiste en la présence de **zones d'instabilités**, qui correspondent à un profil thermodynamique vertical où les particules d'air se retrouvent dans un état de flottabilité négative. De surcroît, le profil vertical doit être **humide** sur une large épaisseur en moyenne troposphère. Enfin, même si le potentiel convectif est présent, il est nécessaire que les particules d'air atteignent l'état instable. Un **forçage** est généralement nécessaire pour initier les mouvements verticaux.

Une fois la convection en place, on parle de phase mature. De fortes ascendances apparaissent pouvant atteindre des vitesses de l'ordre du m s^{-1} . Dès les premiers mouvements ascendants, une compensation de grande échelle ou locale se met en place, par des mouvements subsidents. Lorsqu'une particule d'air subsidente atteint le sol, elle est plus froide que l'air environnant. Étant plus dense, ces particules se propagent à la surface et provoquent un **courant de densité**. Ces courants de densité peuvent être à l'origine d'importantes rafales et de dégâts. Les courants de densité influencent l'environnement autour des orages ; ils peuvent alors aussi bien être à l'origine d'autres orages en servant de forçage (on parlera d'initiation secondaire), qu'annihiler la convection en séparant une cellule de sa réserve d'instabilité.

L'organisation de la convection profonde permet de comprendre le passage d'une simple cellule convective ordinaire à des orages multicellulaires ou supercellulaires. L'organisation d'un système orageux dépend du **cisaillement vertical** du vent et de la stabilité de l'environnement. Le cisaillement renseigne à la fois sur le changement du module et de la direction du vent sur la verticale.

Une cellule convective ordinaire est constituée d'un cumulonimbus isolé. Ce dernier s'étend généralement sur 10 km^2 à 100 km^2 à l'horizontale, sur une grande partie de la troposphère sur la verticale et son cycle de vie est court, de 30 à 50 min. Lors de l'ascension rapide des particules d'air, la vapeur d'eau se condense, formant des particules d'eau liquides et de glace, qui lorsqu'elles atteignent un poids suffisant, précipitent. Les courants

subsidents associés aux précipitations forment un courant de densité arrivant à la surface, inhibant la convection et provoquant la disparition de l'orage.

Lorsque plusieurs cellules orageuses sont agrégées au sein d'une même masse nuageuse, on parle d'orages multicellulaires. Avec un cisaillement de vent important, le courant de densité d'une cellule convective initie une nouvelle cellule, donnant un axe préférentiel à l'orage multicellulaire. La vitesse de déplacement de l'orage correspond à la somme vectorielle du vent synoptique, et de la vitesse de propagation discrète liée à l'initiation de nouvelles cellules. Si ces vitesses sont opposées, l'orage se maintient sur place. On dit que le système multicellulaire est stationnaire. La stationnarité du système orageux peut être due à d'autres phénomènes, comme la stationnarité du forçage à l'origine de l'orage (par exemple une convergence de vent stationnaire).

Les systèmes orageux peuvent adopter différentes configurations. Ceux composés d'un alignement de cellules convectives sont appelés lignes de grains. Les systèmes nuageux produisant une zone de précipitation continue sur une longueur supérieure à 100 km dans au moins une direction (Houze Jr, 2014, 2018) sont appelés systèmes convectif de méso-échelle (MCS). Un MCS est composé de cellules convectives principales et de cellules affaiblies qui constituent une région stratiforme à l'arrière du système. Les MCS développent des circulations de méso-échelle, dont font partie l'ascendance de méso-échelle dans la partie haute du système et la subsidence de méso-échelle dans la partie basse de la région stratiforme. Contrairement aux cellules convectives isolées à faible durée de vie, les MCS durent de l'ordre de quelques heures, jusqu'à 10 h. Les MCS quasi stationnaires à régénération rétrograde sont à l'origine d'épisodes fortement précipitants. Les HPE adviennent notamment sur le bassin méditerranéen, et sont alors qualifiés d'épisodes méditerranéens ou cévenols suivant leur localisation.

Les systèmes convectifs se découpent temporellement suivant la phase d'initiation, la phase mature et la phase de dissipation. La signature au sol de chaque phase a été étudiée grâce à des simulations idéalisées (Madaus et Hakim, 2016). Durant la phase d'initialisation, les anomalies de surface de pression restent faibles (< 0.1 hPa), jusqu'à 10 min avant l'occurrence des précipitations. Durant cette phase, les anomalies les plus importantes au sol concernent la température causée par les ombres portées des nuages cumuliformes 40 à 60 min avant les premières précipitations. Concernant la phase mature, les signatures au sol sont pilotées par les subsidences. Le courant de densité et la plage froide associée provoquent une baisse de température, et une hausse de l'humidité relative, de la pression et du module de vent. L'étalement et l'amplitude de la plage froide varient suivant la taille, le nombre de cellules et le moment du cycle de vie du système orageux (Ducrocq *et al.*, 2002 ; Engerer *et al.*, 2008).

Les variations de pression de surface sont observées lors du passage d'un système orageux (figure 1.14). Ces structures sont positionnées par rapport au déplacement du système, avant et après son passage. Une surpression de méso-échelle se crée à l'arrière de la ligne convective, associée aux forts courants de subsidences, et grossit durant la phase de développement de l'orage. À l'arrière des précipitations se forme une dépression de sillage, générée par le réchauffement du jet d'afflux arrière (Haertel et Johnson, 2000 ; Johnson et Mapes, 2001). Cette anomalie négative de pression atteint son maximum de développement pendant la phase de dissipation, avant le retour à l'équilibre du système.

La prévisibilité d'un système convectif reste faible. Avec un cas idéalisé, Cintineo et Stensrud (2013) montrent que la prévisibilité de l'occurrence d'une supercellule est de 2 h ; la position du mésocyclone au sein de la supercellule, elle, se réduit à 40 min en avance, et enfin la position exacte des plus forts cumuls reste imprévisible.

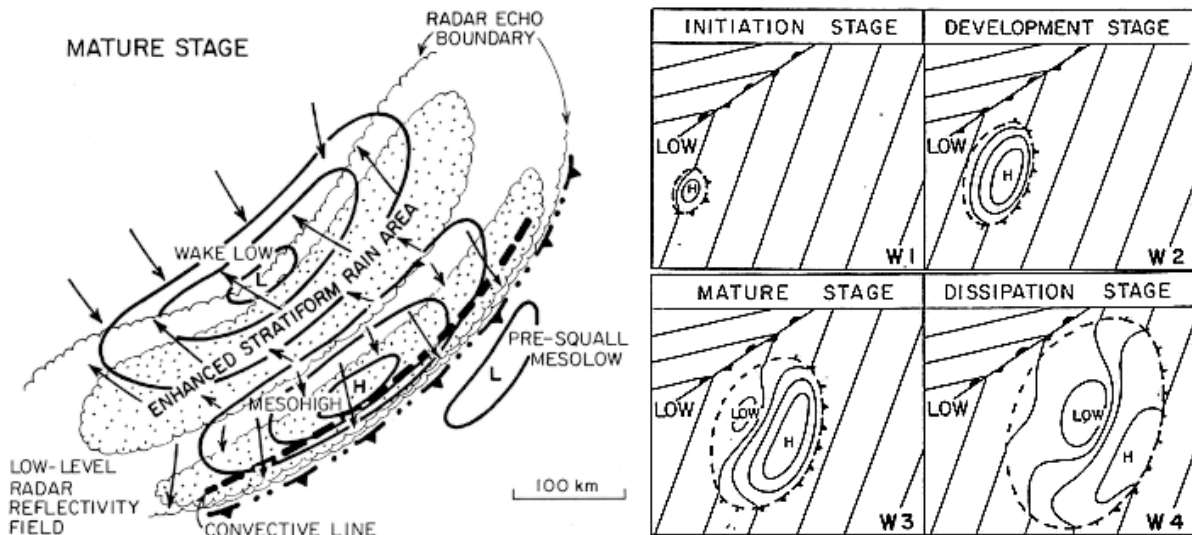


FIGURE 1.14 – Structure horizontale de lignes de grains matures. La pression de surface est indiquée en traits pleins, les vecteurs représentent le vent près de la surface et les régions où les précipitations sont significatives sont pointillées ou grisées. Adapté de Haertel et Johnson (2000).

1.3.2.2 Description de phénomènes convectifs grâce aux observations de PWS

La densité d’observations de PWS est d’une aide précieuse pour caractériser des phénomènes convectifs de méso-échelle. Clark *et al.* (2018) montrent la capacité d’un réseau de 600 observations de PWS en température, pression et vitesse de vent à décrire les caractéristiques locales d’une supercellule.

Par la suite, les travaux de Mandement (2020) ont largement contribué à la description des phénomènes convectifs à l’aide des observations de PWS. Dans Mandement et Caumont (2020), des situations orageuses sont décrites à l’aide d’analyses combinant les stations conventionnelles et personnelles (analyses SPWS). Bien que le modèle soit capable de simuler des structures d’orages, ceux-ci restent éloignés dans le positionnement et l’intensité des cellules réelles. Les observations de pression des stations PWS permettent de représenter une surpression accompagnant une ligne de grains, ainsi que des dépressions de sillage à l’arrière de ces lignes.

Mandement et Caumont (2021) étudient l’épisode méditerranéen du 15 octobre 2018 dans l’Aude, comportant un MCS quasi stationnaire. L’utilisation d’analyses SPWS de pression et de température potentielle virtuelle a permis de mettre en évidence un front de méso-échelle quasi stationnaire à l’est des précipitations les plus intenses, durant près de 6 h. Une comparaison du positionnement du front entre les analyses SPWS et les analyses du modèle AROME a pu être réalisée. Les analyses AROME sont globalement proches des analyses SPWS concernant le positionnement du front, mais sous-estiment la durée de stationnarité du gradient de pression mise en évidence par l’analyse SPWS. On voit dans cet exemple la faculté d’AROME à décrire des phénomènes de fine échelle, mais un manque de réalisme dans le positionnement spatial et temporel des éléments qui sont décrits finement avec l’aide des observations des PWS. Nous verrons dans le chapitre 3, comment les observations de pression des PWS, permettent de décrire l’évolution de la pression de surface durant l’événement fortement précipitant du 14 septembre 2021 dans le Gard, ainsi que le gain apporté par leur assimilation dans le modèle AROME.

1.3.3 Phénomènes de basses couches

1.3.3.1 Théorie de la couche limite atmosphérique

La couche limite atmosphérique (CLA) est la couche de l'atmosphère dans laquelle l'écoulement du fluide atmosphérique est en interaction directe avec la surface terrestre et avec un temps de réponse court, inférieur à une heure (Malardel, 2009). L'épaisseur verticale de la CLA est de l'ordre de 1 km, tout en variant en fonction des conditions atmosphériques. L'ensemble de l'atmosphère au-dessus de la CLA, appelé atmosphère libre, est aussi en interaction avec la surface, mais à des échelles de temps plus longues. La CLA est en équilibre entre les flux d'énergie et de rayonnement entrant et sortant entre la surface et l'atmosphère libre, suivant :

$$R_n = K \downarrow + K \uparrow + I \downarrow + I \uparrow = H + LE - G \quad (1.5)$$

Avec $K \downarrow$, $K \uparrow$ les flux de rayonnements solaires incident et réfléchi, $I \downarrow$, $I \uparrow$ les flux infrarouges reçu et émis, H le flux de chaleur sensible, LE le flux de chaleur latent et G le flux de chaleur échangé avec le sous-sol.

Lors de conditions calmes, la CLA évolue durant la journée selon un cycle diurne caractéristique. Durant la journée, la CLA est chauffée par le sol, créant une production thermique de turbulence : on parle de couche de mélange. Quand arrive la nuit, le sol n'est plus chauffé et une couche limite stable se forme à partir de la surface. Elle s'élève sous le résidu de la couche mélangée, qui devient elle-même une couche résiduelle neutre, peu turbulente. Au sommet de la CLA, il n'est pas rare en conditions calmes de voir se former une couche stable caractérisée par une inversion, où la température augmente avec l'altitude. Il faut attendre le lever du soleil pour que le cycle recommence. Ce cycle dépend ainsi de la durée relative du jour et de la nuit en fonction du jour de l'année, ainsi que du type de sol. Par la suite, nous prendrons trois exemples liés à des discontinuités de surface qui interviennent au sein de la CLA : l'îlot de chaleur urbain, l'effet de brise de mer, et le brouillard de rayonnement.

L'îlot de chaleur urbain (ICU) correspond à une différence de température entre une zone urbaine et une zone péri-urbaine ou rurale adjacente. Il est particulièrement intense la nuit et favorisé par des situations de ciel clair et de vents faibles, lors de vagues de chaleur notamment. La différence de température s'explique par différents facteurs. D'abord, (i) le stockage de chaleur dans les surfaces artificialisées est plus important en ville. Les matériaux urbains (par ex. l'asphalte, le béton, les tuiles), par leur capacité thermique, ont tendance à stocker plus d'énergie en journée et ne la restituent que lentement la nuit en comparaison à des surfaces naturelles (Marques, 2023). Par ailleurs, (ii) le flux anthropique en ville vient s'ajouter aux autres sources d'énergie. Cette source de chaleur provient principalement de l'industrie, du transport et de la climatisation (Pigeon *et al.*, 2007). Enfin, (iii) le flux de chaleur latent dégagé par l'évaporation de l'eau de la végétation est moins important en ville, inhibant le caractère refroidissant de la végétation, qui a lieu en zone rurale. Une fois l'ICU défini, il convient d'en trouver un indicateur. L'indicateur dépend du réseau d'observations dont on dispose. Quand on définit l'ICU par rapport à la température de l'air à la hauteur d'un piéton, on parle d'ICU de canopée urbaine.

La brise de mer est quant à elle due à la différence thermique entre la côte et la mer. Ce phénomène s'explique par la différence de capacité calorifique entre l'eau et la terre, mais aussi par la différence d'absorption de chaleur : une épaisseur d'eau est capable de stocker plus d'énergie par convection et donc s'échauffe moins que le sol terrestre (Malardel, 2009).

En recevant plus de chaleur sensible sur terre que sur mer, l'épaisseur entre les isobares sur terre devient plus importante. Il se crée un gradient de pression au sommet de la couche limite, entraînant une divergence du vent de la terre vers la mer. Ce premier mouvement au sommet de la couche limite correspond à la branche haute initiant la cellule de brise. Cette divergence en hauteur entraîne une baisse de pression en surface sur terre, servant d'appel d'air à la branche basse de la cellule de brise, de la mer vers la terre. C'est ce vent de surface qui est appelé brise de mer. Si l'humidité est suffisante, les vitesses verticales ascendantes sur la côte forment des cumulus qui informent sur le positionnement du front de brise. D'autres brises existent, tout autant causés par un chauffage différentiel, comme la brise de terre, la brise de pente ou la brise de vallée.

Le brouillard de rayonnement est, quant à lui, causé par le refroidissement radiatif à la surface. Il prédomine la nuit sans nuages, quand le rayonnement radiatif infrarouge sortant domine le bilan de surface, induisant une baisse de la température. Il est nécessaire d'avoir un faible vent, permettant à l'air proche de la surface d'atteindre le point de rosée provoquant la condensation. Petit à petit, le brouillard s'intensifie et devient opaque au rayonnement infrarouge, limitant ainsi la baisse de température à la surface. Dès le lever de soleil, si le rayonnement solaire est suffisant, le brouillard se dissipe. La dissipation peut aussi être causée par un changement de situation synoptique, comme un renforcement du vent, mélangeant la couche limite avec de l'air plus sec. Il existe d'autres types de brouillard, comme le brouillard par affaissement de stratus ou le brouillard d'advection.

1.3.3.2 Description de phénomènes de basses couches grâce aux observations de PWS

Les observations issues d'observations de PWS permettent de suivre les phénomènes d'îlot de chaleur urbain, identifiés dans de nombreuses métropoles comme à Londres (Chapman *et al.*, 2017), à Berlin (Fenner *et al.*, 2017; Meier *et al.*, 2017; Napoly *et al.*, 2018), à Paris (Napoly *et al.*, 2018; Madelin et Dupuis, 2020), à Oslo (Venter *et al.*, 2020) et à Moscou (Varentsov *et al.*, 2020). D'autres types d'observations d'opportunité permettent de caractériser l'ICU, comme les thermomètres embarqués au sein des voitures connectées à Rennes et à Dijon (Marquès *et al.*, 2022).

Parmi ces études, Madelin et Dupuis (2020) étudient l'îlot de chaleur urbain à Paris grâce aux observations de PWS, sur une longue période de 2017 à 2020. Madelin et Dupuis (2020) montrent ainsi que, quantitativement, les observations de PWS sont en accord avec les observations conventionnelles, tout en apportant une résolution spatiale bien supérieure. Peu d'études se sont encore intéressées à la description d'autres phénomènes de basses couches avec des observations de PWS.

1.3.3.3 Cas d'étude de l'ICU du 17 juin 2023

Au matin du 17 juin 2023, à 8 h locale, une différence notable de température apparaît après la nuit entre le centre urbain de Paris et sa zone périphérique (figure 1.15). Alors que dans le centre urbain on retrouve une valeur de 21.8 °C à Paris-Montsouris (station de référence de classe 3) et 22.8 °C à Longchamp (classe 2), en zone péri-urbaine on retrouve des valeurs entre 18 et 19 °C comme à Dourdan avec 18.3 °C (classe 1); donnant une différence de 4.5 °C à une distance de 55 km.

La prévision 1 h d'AROME est capable de simuler un îlot de chaleur urbain (figure 1.15a). La valeur à 5 m indique 18.6 °C à Dourdan (0.3 °C par rapport à l'observation), 21.0 °C à Paris-Monsouris (-0.8 °C par rapport à l'observation) et 22.8 °C à Longchamp

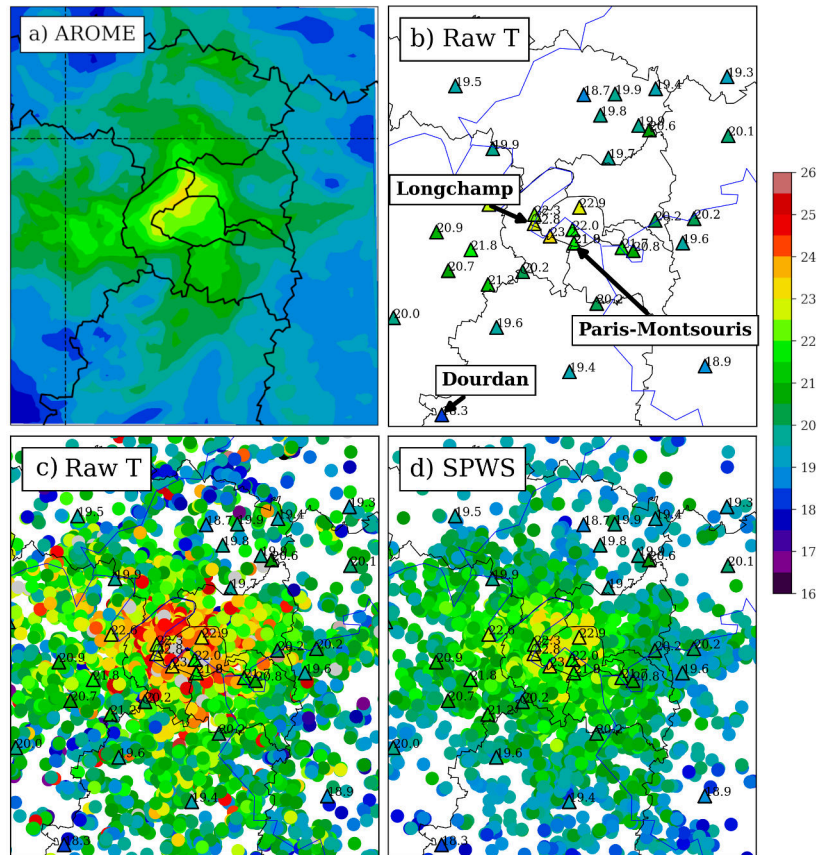


FIGURE 1.15 – Température (en °C) le 17 juin 2023 à 06:00 UTC. (a) Prévision d'AROME à 1 h d'échéance, (b) observations du réseau de stations standards et (c) observations (c) brutes et (d) pré-traitées du réseau de PWS.

(égale à l'observation). Le modèle diminue la différence entre Longchamp et Dourdan à 5 m, à 4.2 °C.

Le réseau d'observations PWS est très dense dans les zones urbanisées. Sur la zone décrite se trouvent 36 observations conventionnelles, pour 9753 observations PWS brutes, soit 270 fois plus d'observations. Cette densité se réduit à 5005 observations PWS une fois pré-traitées, soit 139 fois plus d'observations. Les observations PWS indiquent des valeurs très élevées jusqu'à 34.6 °C. Le calcul de l'ICU peut être réalisé avec des températures de l'air, correspondant aux observations PWS, et même si le rayonnement solaire est faible à cette heure (Varentsov *et al.*, 2020), les observations restent entachées d'erreurs de représentativité. Une fois traitées, les observations PWS présentent un maximum à 24.3 °C. Le traitement permet de les corriger et ne conserver que des observations proches des températures à 2 m conventionnelles, très éloignées des mesures des PWS. Les stations traitées sont cohérentes entre elles (figure 1.15.d). Elles correspondent au premier ordre à l'interpolation des observations conventionnelles, ce qui diminue le gain apporté par le réseau PWS. Il est difficile de comparer directement le dernier niveau du modèle aux observations à 2 m. Néanmoins, en comparant les observations PWS traitées et le modèle, on remarque une sous-estimation de la plage chaude dans le nord-ouest de Paris dans le modèle.

1.3.3.4 Cas de brouillard le 23 septembre 2021

Dans la nuit du 22 au 23 septembre 2021, un brouillard de rayonnement se développe dans le val de Saône. Nous reviendrons dans le chapitre 4 sur les circonstances de son développement, au fil des échéances. Ce qui nous intéresse ici est la différence entre le modèle et les observations de surface, qui nous donnent une information différente sur la plage humide avec des valeurs à plus de 90 % d'humidité relative. Les observations de surface conventionnelles permettent de donner une information ponctuelle de la plage humide, mais ne permettent pas de délimiter son étendue. Le modèle AROME, lui, simule une plage humide méridienne avec un maximum à 103.2 % (tout en sachant que les observations ne peuvent dépasser 100 %). Le réseau PWS d'observations brutes permet d'identifier une plage humide méridienne, mais possède une diversité d'observations intermédiaires entre 60 et 90 % d'humidité relative. En fin de nuit, le rayonnement solaire joue un faible rôle sur le biais des stations ; on peut supposer que l'emplacement de la station est principalement responsable de cette hétérogénéité. L'humidité relative est un paramètre qui met en exergue les différences spatiales très locales de la surface. Après le traitement des observations personnelles, la plage humide méridienne est nettement discernable. On remarque que les observations proches de la station de référence St-Georges-Reneins se sont nettement rapprochées de sa valeur avec une correction de plus de 30 %. Le contrôle de qualité conserve les stations proches de l'interpolation linéaire des stations conventionnelles. Le gain apporté par les observations d'opportunité est alors faible.

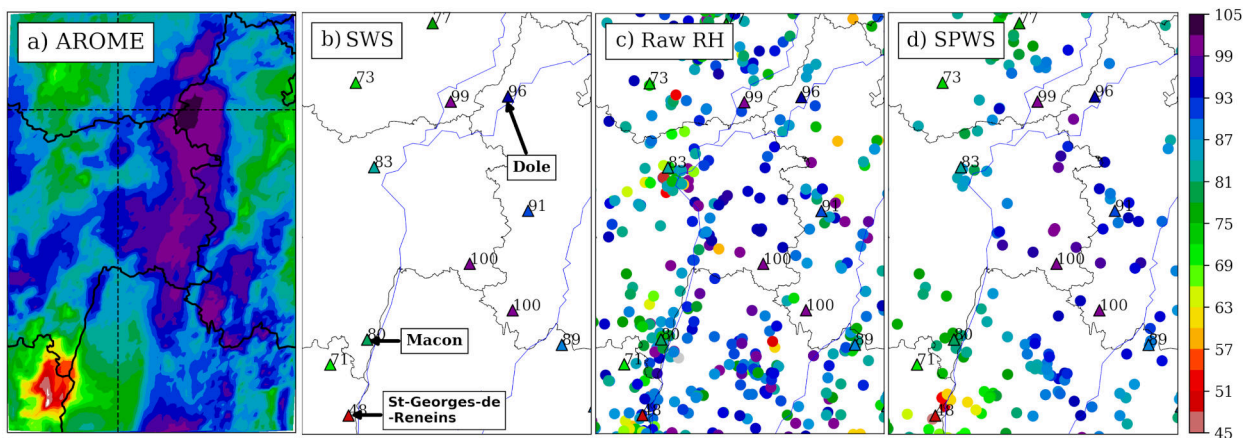


FIGURE 1.16 – Humidité relative (en %) le 23 septembre 2021 à 09:00 UTC. Voir Fig. 1.15 pour plus de détails.

Ces deux études de cas permettent de voir le potentiel et les difficultés que peuvent apporter les observations d'opportunité de température et d'humidité relative.

1.4 Synthèse

Les observations météorologiques conventionnelles de surface (SWS) couvrent l'ensemble des zones terrestres, avec une répartition spatiale irrégulière suivant les pays. Ces observations suivent les conventions dictées par l'OMM. Les SWS possèdent des métadonnées permettant de fournir des informations sur l'observation et de caractériser sa qualité. Un classement des observations est réalisé en fonction de leurs erreurs d'instrumentation et de représentativité. En plus des observations conventionnelles, les stations

d'observations personnelles (PWS) représentent une aubaine par leur densité spatiale en France métropolitaine. Les observations PWS nous renseignent sur les échelles fines de l'atmosphère, allant de l'échelle méso- β à méso- γ , c'est-à-dire de 2 à 200 km. Du fait de la construction des capteurs de température et d'humidité relative des PWS, leur mesure sont sujette à des biais importants. Il en est de même avec les observations de pression du fait de l'incertitude sur l'altitude de ces stations. Le traitement des observations PWS est primordial, et à cause de leur manque de métadonnées, leur exploitation est un défi à part entière. Il est important de noter que le terme « d'observations de surface » (voir « observations au sol ») réfère par la suite à toutes les observations reliées au sol et placées entre le sol et 2 m de hauteur.

En parallèle sont développés des modèles de prévision numérique du temps à aire limitée (LAM) de résolution proche du kilomètre. Les LAM sont capables de simuler des phénomènes locaux, tels que la convection profonde. Le modèle AROME-France est le LAM développé par Météo-France à résolution de 1.3 km sur l'horizontale.

Enfin, une description de phénomènes de méso-échelle, pour lesquels les observations PWS apportent un gain par rapport aux observations SWS, a été menée. Tout d'abord, les phénomènes concernés sont ceux comprenant de forts mouvements verticaux, tels qu'une dépression de méso-échelle, un front, un mascaret atmosphérique ou un MCS, qu'il est possible de décrire grâce aux observations de pression de fine échelle. Ensuite, des phénomènes de couches limite peuvent être décrits à l'aide d'observations de température et d'humidité relative provenant d'observations d'opportunité. En parallèle, le modèle de prévision AROME a montré sa capacité à représenter ces phénomènes de méso-échelle. Néanmoins, un décalage spatial, temporel ou une mauvaise intensité est parfois constatée. L'assimilation de données doit permettre de tirer profit des observations d'opportunité pour améliorer la représentation de ces phénomènes de méso-échelle au sein du modèle de PNT.

Chapitre 2

Assimilation d'observations au sol dans AROME-France

Sommaire

2.1	Assimilation de données	40
2.1.1	Formalisme général du problème	40
2.1.2	Schéma variationnel	40
2.1.3	Schéma d'interpolation optimale	42
2.1.4	Résolution séquentielle	43
2.1.5	Application pour différents modèles à aire limitée	43
2.2	Cycle d'assimilation d'AROME	44
2.2.1	Contrôle de qualité du modèle	44
2.2.2	Opérateurs d'observation au sol	45
2.2.3	Espace des observations	48
2.2.4	Réglage des covariances d'erreur d'ébauche	50
2.2.5	Schéma d'assimilation de surface	52
2.2.6	Chaîne de prévision opérationnelle	53
2.3	Assimilation des observations de surface	55
2.3.1	Quel impact ont les observations de surface ?	55
2.3.2	Pourquoi l'assimilation d'observations standards de surface est-elle un challenge ?	56
2.3.3	Quels défis pour l'assimilation d'observations d'objets connectés ?	58
2.3.4	Quels défis pour l'assimilation de données à échelle convective ?	59
2.4	Synthèse	60

La prévision numérique du temps à courte échéance dépend notamment de la qualité des variables d'état du modèle à l'instant initial de la prévision. Pour réaliser cet état initial de l'atmosphère, nous avons recours au processus d'**assimilation de données**. Ce processus combine l'ensemble des **observations** disponibles et la meilleure estimation que l'on a de l'atmosphère, correspondant à la dernière prévision disponible, que l'on nomme **ébauche**. Les variables ainsi initialisées sont nommées **variables de contrôle** et constituent l'**analyse**. Une fois l'analyse obtenue, on réalise la prévision pour la prochaine échéance qui devient l'ébauche du prochain réseau d'assimilation.

Cette section est divisée en trois parties. Tout d'abord, nous présenterons le problème de l'assimilation de données et les schémas de résolution les plus utilisés en météorologie. Nous verrons le principe de l'assimilation de données dans AROME, du traitement des observations de surface, aux réglages du schéma variationnel. Enfin, nous examinerons quels défis pose l'assimilation des observations denses de surface dans un système de PNT régional.

2.1 Assimilation de données

L'assimilation de données repose sur la recherche du meilleur compromis entre l'ébauche et les observations. Les observations disponibles à l'instant de l'analyse seront symbolisées par le vecteur \mathbf{y}^o , de dimension p . L'état de l'atmosphère donné par l'ébauche sera quant à lui symbolisé par le vecteur \mathbf{x}^b de dimension n . L'objectif de l'assimilation est de se rapprocher de l'état vrai $\mathbf{x}^t \in \mathbb{R}^n$ du modèle, état cherchant à représenter au mieux sur la grille du modèle l'état vrai de l'atmosphère. Cet état idéal permet de définir les vecteurs d'erreurs d'ébauche $\boldsymbol{\varepsilon}^b$ et d'observation $\boldsymbol{\varepsilon}^o$, tel que :

$$\mathbf{x}^b = \mathbf{x}^t + \boldsymbol{\varepsilon}^b \quad (2.1)$$

$$\mathbf{y}^o = \mathcal{H}(\mathbf{x}^t) + \boldsymbol{\varepsilon}^o \quad (2.2)$$

où \mathcal{H} est l'opérateur d'observation, qui permet de passer de l'espace du modèle à l'espace des observations : $\mathcal{H} : \mathbb{R}^n \rightarrow \mathbb{R}^p$.

On fera l'hypothèse que les erreurs d'ébauche et d'observation sont de moyenne nulle ($\mathbb{E}[\boldsymbol{\varepsilon}^b] = \mathbb{E}[\boldsymbol{\varepsilon}^o] = 0$), et qu'elles sont distribuées selon une loi normale, avec $\mathbf{B} = \mathbb{E}[\boldsymbol{\varepsilon}^b(\boldsymbol{\varepsilon}^b)^T]$ la matrice de covariances d'erreurs d'ébauche et $\mathbf{R} = \mathbb{E}[\boldsymbol{\varepsilon}^o(\boldsymbol{\varepsilon}^o)^T]$ la matrice de covariances d'erreurs d'observation.

Les matrices de covariances d'erreurs d'ébauche $\mathbf{B} \in \mathbb{R}^{n \times n}$ et d'observation $\mathbf{R} \in \mathbb{R}^{p \times p}$ sont symétriques et définies positives. Nous verrons qu'elles sont cruciales pour définir la confiance relative que l'on accorde à l'ébauche et aux observations.

2.1.1 Formalisme général du problème

On cherche à déterminer la probabilité de la trajectoire de \mathbf{x} connaissant le résultat des observations \mathbf{y}^o . En utilisant la formule de Bayes, il est possible d'estimer cette probabilité à partir de termes connus :

$$p(\mathbf{x}|\mathbf{y}^o) = \frac{p(\mathbf{y}^o|\mathbf{x}) \cdot p(\mathbf{x})}{p(\mathbf{y}^o)} \quad (2.3)$$

Avec $p(\mathbf{y}^o|\mathbf{x})$ la vraisemblance des observations \mathbf{y}^o en connaissant l'état du modèle et $p(\mathbf{x})$ la probabilité a priori de l'état de l'atmosphère. Le dernier terme $p(\mathbf{y}^o)$ ne dépend pas de \mathbf{x} et revient donc à un facteur de normalisation. Avec l'hypothèse que les erreurs d'observation et d'ébauche sont gaussiennes, on peut définir ces termes sous la forme :

$$p(\mathbf{y}^o|\mathbf{x}) \propto \exp\left(-\frac{1}{2} \|\mathcal{H}(\mathbf{x}) - \mathbf{y}^o\|_{\mathbf{R}^{-1}}^2\right) \quad (2.4)$$

$$p(\mathbf{x}) \propto \exp\left(-\frac{1}{2} \|\mathbf{x} - \mathbf{x}^b\|_{\mathbf{B}^{-1}}^2\right) \quad (2.5)$$

2.1.2 Schéma variationnel

L'approche variationnelle revient à résoudre le problème bayésien, en minimisant la fonction de coût $J(\mathbf{x}) = -\ln(p(\mathbf{x}|\mathbf{y}^o))$. Dans le cas d'un vecteur d'état \mathbf{x} en trois dimensions, la méthode d'assimilation variationnelle est appelée **3DVar**. La fonction de coût J

peut alors s'écrire sous la forme :

$$\begin{aligned} J(\mathbf{x}) &= \frac{1}{2} \|\mathbf{x} - \mathbf{x}^b\|_{\mathbf{B}^{-1}}^2 + \frac{1}{2} \|\mathcal{H}(\mathbf{x}) - \mathbf{y}^o\|_{\mathbf{R}^{-1}}^2 \\ &= \underbrace{\frac{1}{2} (\mathbf{x} - \mathbf{x}^b)^T \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}^b)}_{J_b(\mathbf{x})} + \underbrace{\frac{1}{2} (\mathbf{y}^o - \mathcal{H}(\mathbf{x}))^T \mathbf{R}^{-1} (\mathbf{y}^o - \mathcal{H}(\mathbf{x}))}_{J_o(\mathbf{x})} \end{aligned} \quad (2.6)$$

Le premier terme J_b est un terme de rappel à l'ébauche et le deuxième terme J_o est un terme de distance aux observations. L'état \mathbf{x} cherché est un compromis entre la distance à l'ébauche et aux observations.

La minimisation de la fonction de coût peut être résolue par l'**approche incrémentale**, qui est un processus itératif basé sur la méthode de Gauss-Newton. L'approche incrémentale consiste à résoudre le problème variationnel non-linéaire par une suite d'approximations linéaires. On fera donc l'hypothèse de la linéarité de l'opérateur d'observation au voisinage de l'ébauche :

$$\mathcal{H}(\mathbf{x}) - \mathcal{H}(\mathbf{x}_b) \approx \mathbf{H}(\mathbf{x} - \mathbf{x}_b) \quad (2.7)$$

où $\mathbf{H} \in \mathbb{R}^{p \times n}$ est la matrice jacobienne qui définit l'opérateur d'observation tangent linéaire (TL) de \mathcal{H} en x_b . En posant l'incrément $\delta \mathbf{x} = \mathbf{x} - \mathbf{x}_b$, on obtient :

$$J(\delta \mathbf{x}) = \frac{1}{2} \|\delta \mathbf{x}\|_{\mathbf{B}^{-1}}^2 + \frac{1}{2} \|\mathbf{H} \delta \mathbf{x} - \mathbf{d}\|_{\mathbf{R}^{-1}}^2 \quad (2.8)$$

avec $\mathbf{d} = \mathbf{y}^o - \mathcal{H}(\mathbf{x}_b)$ le vecteur des innovations.

Le gradient de la fonction de coût est donné par :

$$\nabla J = \mathbf{B}^{-1} \delta \mathbf{x} + \mathbf{H}^T \mathbf{R}^{-1} (\mathbf{H} \delta \mathbf{x} - \mathbf{d}) \quad (2.9)$$

où \mathbf{H}^T est l'opérateur d'observation adjoint (AD), qui permet de passer de l'espace des observations à celui du modèle, tel que $\mathbf{H}^T \in \mathbb{R}^{n \times p}$.

On cherche $\delta \mathbf{x}$ tel que $\nabla J = 0$, ce qui revient à chercher $\delta \mathbf{x}^*$, tel que :

$$(\mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H}) \delta \mathbf{x}^* = \mathbf{H}^T \mathbf{R}^{-1} \mathbf{d} \quad (2.10)$$

Le terme de gauche étant symétrique et défini positif, on retrouve $\delta \mathbf{x}^* = \mathbf{K} \mathbf{d}$ avec $\mathbf{K} = (\mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{R}^{-1}$ la formulation primale du Best Linear Unbiased estimator (BLUE). On peut montrer que la formulation primale peut se réécrire sous la formulation duale, avec :

$$\mathbf{x}^a = \mathbf{x}^b + \mathbf{K}(\mathbf{y}^o - \mathbf{H} \mathbf{x}^b) \quad (2.11)$$

$$\mathbf{K} = \mathbf{B} \mathbf{H}^T (\mathbf{H} \mathbf{B} \mathbf{H}^T + \mathbf{R})^{-1} \quad (2.12)$$

La matrice \mathbf{K} est nommée **gain de Kalman**. Le BLUE est la solution générale du problème en prenant pour hypothèse la linéarité de l'opérateur d'observation. La différence $\mathbf{x}^a - \mathbf{x}^b$ est appelée **incrément d'analyse**.

En pratique, on ne peut pas inverser les matrices \mathbf{B} et \mathbf{R} , ni même les stocker du fait de leur taille. À la place, on résout directement le système de l'eq. 2.12 grâce à la méthode

du gradient conjugué (méthode de Krylov)¹¹.

Dans l'approche incrémentale, il est courant de réaliser un pré-conditionnement des variables de contrôle. Cela permet conjointement de découpler les variables d'état, de simplifier son codage, et d'améliorer le conditionnement des problèmes de minimisation. En pratique on réalise un pré-conditionnement en $\mathbf{B}^{1/2}$ dans AROME. Cela revient à chercher $\mathbf{U} \in \mathbb{R}^{n \times q}$ vérifiant $\mathbf{B} = \mathbf{U}\mathbf{U}^T$. Autrement dit, cela revient à poser $\delta\mathbf{x} = \mathbf{U}\boldsymbol{\chi}$. Ce changement de variable rend la fonction de coût moins elliptique ou plus circulaire, ce qui la rend plus rapide à minimiser.

Une faiblesse du schéma 3DVar est de ne pas prendre en compte la composante temporelle des observations. Toutes les observations actives comprises dans la fenêtre d'assimilation sont considérées à la même heure (au milieu de la fenêtre) et comparées à une unique ébauche. Le schéma *First Guess at Appropriate Time* (3DVar-FGAT) permet de comparer l'ébauche au temps réel de l'observation. On peut alors obtenir des innovations différentes pour des observations très proches spatialement, tout en considérant les innovations à la même heure d'analyse lors de la minimisation, pour ne réaliser qu'une minimisation. D'autres extensions du schéma d'assimilation variationnel existent : (i) le schéma 4DVar, qui prend en compte la composante temporelle en ramenant le modèle au temps de l'observation, (ii) le schéma 3DEnVar, qui actualise la matrice d'erreur d'ébauche à chaque analyse à l'aide d'un ensemble de prévision, et (iii) le schéma 4DEnVar, combinant les deux approches.

Enfin, dans le cas où l'on considère le modèle parfait, le système d'assimilation est dit à « contrainte forte ». Si le modèle est considéré imparfait, un terme facultatif J_c s'ajoute à la fonction de coût, prenant en compte les erreurs du modèle ; le système d'assimilation est dit à « contrainte faible ».

2.1.3 Schéma d'interpolation optimale

L'interpolation optimale (IO) consiste à calculer explicitement la formule de l'analyse du BLUE. Afin de réduire le coût de calcul, on sélectionne un faible nombre d'observations N autour de chaque point de l'ébauche. En pratique il s'agit d'observations localisées dans un certain voisinage du point considéré. On peut réécrire la formulation duale du BLUE (eq. 2.11 et eq. 2.12), où pour chaque composante j du vecteur d'état, on cherche l'incrément d'analyse $\Delta(\mathbf{x}^a)_j$, tel que :

$$\Delta(\mathbf{x}^a)_j = \sum_{i=1}^N W_i \Delta \mathbf{x}_i \quad (2.13)$$

avec $\Delta \mathbf{x}_i$ l'écart à l'ébauche de l'observation i , et W_i son poids optimal donné par :

$$(\mathbf{B}_{loc} + \mathbf{R}_{loc})\mathbf{W} = b \quad (2.14)$$

avec $\mathbf{B}_{loc} \in \mathbb{R}^{N \times N}$ et $\mathbf{R}_{loc} \in \mathbb{R}^{N \times N}$ les restrictions de $\mathbf{H}\mathbf{B}\mathbf{H}^T$ et \mathbf{R} , aux N observations sélectionnées ; et b les N coefficients correspondants de la ligne i de $\mathbf{B}\mathbf{H}^T$.

Dans le cas de l'interpolation optimale, il est courant de choisir \mathbf{B} sous la forme de

11. Le gradient conjugué consiste en un processus itératif, qui est qualifié de boucle interne ; la boucle externe faisant référence aux approximations linéaires successives de l'approche incrémentale. En pratique, AROME ne réalise qu'une seule linéarisation de \mathbf{H} , réduisant la boucle externe à une seule itération.

coefficients $b_{ll'}$ avec l et l' des points de la grille :

$$b_{ll'} = \sigma_b^2 \times \alpha(r_{ll'})\beta(z_{ll'}) = \sigma_b^2 \times \exp\left(-\frac{1}{2}\left[\frac{r_{ll'}}{d}\right]^2\right) \times \exp\left(-\left[\frac{z_{ll'}}{h}\right]^2\right) \quad (2.15)$$

où $r_{ll'}$ et $z_{ll'}$ sont les distances horizontales et verticales entre deux points de grille l et l' , $\alpha(r_{ll'})$ et $\beta(z_{ll'})$ les fonctions de structures horizontales et verticales. d et h sont des paramètres de distance horizontale et verticale.

2.1.4 Résolution séquentielle

D'autres méthodes existent pour résoudre le problème de l'assimilation de données. Le filtre de Kalman (KF, *Kalman Filter*) en fait partie, correspondant à une résolution séquentielle. On sépare les différentes dates d'observation ou paquets d'observation en séquences de t_0, t_k, \dots, t_K . Chaque pas d'itération est composé de deux étapes successives, nommées prédiction et correction. L'étape de prédiction utilise un opérateur d'évolution incrémentale pour établir une estimation de l'état courant à partir de l'état estimé au pas précédent. L'étape de correction utilise les observations courantes pour améliorer l'estimation en corrigeant l'état prédit. Autrement dit, on assimile séquentiellement les observations en mettant à jour les matrices de covariances d'erreur d'ébauche et d'analyse au cours des cycles de correction-prédiction successifs.

Parmi les extensions possibles du filtre de Kalman, on peut citer (i) le filtre de Kalman étendue (*Extended KF*, EKF) qui prend en compte des fonctions non-linéaires, (ii) le filtre de Kalman d'Ensemble (EnKF) qui estime les matrices de covariances d'erreur d'ébauche et d'analyse à partir d'un ensemble de prévision valide au même instant, et (iii) le LETKF (*Local Ensemble Transform Kalman Filter*), qui sépare l'analyse en sous-domaines et en sous-ensembles d'observations. Le LETKF permet de paralléliser l'exécution numérique de façon efficace. Le lecteur pourra trouver plus d'information concernant le filtre de Kalman et ses extensions dans Houtekamer et Zhang (2016).

2.1.5 Application pour différents modèles à aire limitée

En pratique, chaque centre météorologique utilise le schéma d'assimilation qu'il a historiquement développé, en lien avec ses moyens et ses partenaires.

TABLEAU 2.1 – Vue d'ensemble des schémas d'assimilation utilisés opérationnellement pour différents modèles à aire limitée. Pour plus de détails, voir Gustafsson *et al.* (2018).

Centre	Modèle	Méthode
Météo-France	AROME-France	3DVar incrémental ; En préparation : 3DEnVar incrémental
Consortium ALADIN	HARMONIE- AROME	3DVar incrémental
DWD/MétéoSwiss	COSMO	LETKF
Met Office	UKV	3DVar incrémental et 4DVar
NOAA	HRRR	3DEnVar hybride incrémental et analyse de nuages non variationnel
JMA	MA	4DVar incrémental

D'autres paramètres s'ajoutent aux schémas d'assimilation, tels que le type et le

nombre d'observations assimilées, la fenêtre temporelle des observations assimilées, le nombre de variables de contrôle, le pré-conditionnement, l'estimation de B, de sa localisation pour réduire le bruit d'échantillonnage et de son inflation.

L'assimilation de surface est généralement effectuée indépendamment de l'assimilation atmosphérique. L'IO est la technique la plus couramment utilisée pour l'assimilation opérationnelle de surface (Giard et Bazile, 2000 ; Drusch, 2007). Néanmoins, l'EKF remplace désormais l'IO pour l'analyse de l'humidité du sol dans le système du DWD (Hess, 2001) et au Met Office (Dharssi, 2012). Une autre solution explorée par le JMA consiste à inclure les variables de surface dans les variables de contrôle du schéma d'assimilation atmosphérique (Gustafsson *et al.*, 2018). Nous examinerons par la suite les réglages du schéma d'assimilation utilisés dans le modèle AROME.

2.2 Cycle d'assimilation d'AROME

Le système d'assimilation d'AROME est inspiré de celui d'ALADIN-France (Fischer *et al.*, 2005) et du système ARPEGE-IFS (Courtier *et al.*, 1998), dont il tire sa formulation incrémentale, ses opérateurs d'observation, sa méthode de minimisation et sa gestion des flux de données. Le modèle AROME utilise le schéma d'assimilation 3DVar pour initialiser les variables atmosphériques et le schéma d'assimilation d'interpolation optimale pour initialiser les variables de surface.

Dans cette section, nous allons voir les différentes étapes de l'assimilation de données dans AROME : le traitement des observations, les réglages des matrices de covariances d'erreur d'observation et d'ébauche et la description des opérateurs d'observations de surface. Nous concluons sur la mise en œuvre pratique de la chaîne d'assimilation-prévision opérationnelle d'AROME-France.

2.2.1 Contrôle de qualité du modèle

Afin de filtrer les observations anormales, une étape de traitement des observations est réalisée à l'entrée de l'assimilation de données, étape communément appelée *screening*. Nous traiterons ici uniquement du cas des observations de surface. Pour plus d'information concernant l'étape de *screening*, nous renvoyons le lecteur vers la documentation du modèle global IFS (ECMWF, 2023).

L'ensemble des observations est classé selon leur utilisation dans l'analyse, à partir de quatre statuts : actif, passif, rejeté ou blacklisté. Les observations actives sont utilisées pour la minimisation alors que les observations passives sont gardées pour certains usages spécifiques. Pour déterminer leur statut, différents tests de vérification, qu'on appelle « événements », sont réalisés pour chaque observation. Il existe 26 événements dans les modèles IFS/ARPEGE/AROME.

L'événement le plus important, nommé *too big first guess departure*, consiste à rejeter les observations ayant un écart à l'ébauche supérieur à un seuil fixé.

$$\|OmB\| < \alpha \sqrt{\sigma_o^2 + \sigma_b^2} / \sigma_{coef}, \quad (2.16)$$

avec OmB l'écart de l'observation à l'ébauche (*Observation minus Background*), σ_o l'erreur d'observation, σ_b l'erreur d'ébauche et α le seuil de rejet. $\sigma_{coef} = 0.9$ est un coefficient additionnel pour le *screening* d'altitude, augmentant le seuil de rejet, afin de garder plus d'observations.

TABLEAU 2.2 – Paramètres σ_o prescrits et σ_b estimés le 6 septembre 2021 à 01:00 UTC, pour différentes observations de surface.

		σ_o	σ_b	α
t2m	Atmosphère	1.4 °C	0.3-0.5 °C	4
	Surface	1.3 °C	1.6 °C	
rh2m	Atmosphère	10 %	10 %	
	Surface			
z	Atmosphère	78.4532 m ² s ⁻²	40-150 m ² s ⁻²	5

Les valeurs de σ_o et σ_b pour le *screening* du modèle atmosphérique et celui de surface sont rassemblées dans le tableau 2.2. Pour donner un ordre de grandeur du seuil de rejet lors du *screening*, le 1 octobre 2021 à 01:00 UTC sont rejetées les observations dont l'écart à l'ébauche était supérieur à 6.5 °C en température, 70 % en humidité relative et 530 m² s⁻² en géopotential (≈ 5.3 hPa) ; valeurs représentatives sur la période.

D'autres événements existent pour filtrer les observations de surface :

- Redondance des observations aux mêmes temps/latitude/longitude, qui font office de doublons.
- Le rejet d'une observation de température à 2 m entraîne le rejet de l'observation d'humidité relative à 2 m de la même station, car la qualité d'une observation d'humidité relative dépend de la température.

Le *screening* montre ses limites lors d'événements météorologiques mal représentées par le modèle. Il existe des événements de méso-échelle qui entraînent le rejet inapproprié de stations au sol. Si le cycle d'assimilation-prévision n'est pas assez fréquent, il suffit d'un décalage temporel dans la prévision d'un phénomène pour rejeter des stations trop éloignées du modèle. Ceci est identifié lors du derecho survenu en Corse le 18 août 2022, rejetant 39 observations au sol. Ces rejets proviennent de 21 stations de surface, principalement d'observation de t2m mais aussi de rh2m, du vent et du géopotential devenus trop éloignés de l'ébauche.

2.2.2 Opérateurs d'observation au sol

Les observations provenant des stations au sol présentent l'avantage de mesurer directement des paramètres liés aux variables prévues par AROME. Néanmoins, les observations n'étant pas positionnées sur un point de la grille, une interpolation horizontale et verticale est nécessaire. L'opérateur d'observation comprend (i) les diagnostics des variables à 2 m à partir du modèle, (ii) des ajustements pour les différences entre la hauteur de la station et la hauteur du modèle, et (iii) l'interpolation horizontale et temporelle (Ingleby (2015), pour IFS).

2.2.2.1 Géopotential

Le modèle AROME a une coordonnée verticale hybride en pression (chapitre 1). Une observation de pression n'est pas assimilée directement sous forme de pression dans le modèle, mais indirectement grâce au géopotential. La valeur de géopotential est calculée grâce à l'altitude fixe du baromètre de la station. Cette valeur de géopotential est alors assimilée au niveau de pression observé¹².

12. En anglais : *Geopotential at the Level of the Observed station Pressure* (GLOP)

L'opérateur d'observation consiste à calculer le géopotentiel modèle Φ_E au niveau pression observé p_O . Le calcul est effectué via la coordonnée de pression hydrostatique π . L'opérateur d'observation utilise la modélisation de l'atmosphère selon l'atmosphère type ISA¹³ reprise par l'OACI¹⁴, qui permet de définir une pression et une température idéalisée, indépendante du temps et de l'espace. Le modèle ISA divise l'atmosphère en différentes couches avec une distribution linéaire de la température. Par exemple, au niveau de la mer, la norme donne une pression de 1013.25 hPa, une température de 15 °C et un gradient thermique adiabatique de -6.5 °C km^{-1} .

$\Phi_E(p_O)$ aux coordonnées (x_O, y_O) est calculé par les étapes suivantes d'après Courtier *et al.* (1998) :

1. On se place à la position horizontale de l'observation x_O, y_O et on dispose du profil vertical des variables du modèle en ce point par interpolation horizontale.
2. À cette position horizontale, on réalise alors une interpolation verticale, telle que :

$$\Phi_E(p_O) = \Phi_E(\pi_S) + \Phi_{\text{OACI}}(p_O) - \Phi_{\text{OACI}}(\pi_S) + \Delta\Phi(p_O) \quad (2.17)$$

avec π_S la pression hydrostatique de surface et $\Delta\Phi(p_O)$ obtenu par l'interpolation verticale du géopotentiel entre les niveaux du modèle.

3. Cas particuliers : p_O entre le niveau modèle le plus proche du sol et le sol ou sous le relief modèle. On modélise la partie sous le dernier niveau modèle par une atmosphère standard. On vient alors calculer : (i) la température de surface OACI T_S , (ii) la température de surface modèle T^* , en l'extrapolant à partir des deux derniers niveaux du modèle, et (iii) la température équivalente au niveau de la mer T_0 . On a ainsi :

$$\Phi_E(p_O) = \Phi_E(\pi_S) - \frac{R_{\text{dry}} T^*}{\gamma} \left[\left(\frac{p_O}{\pi_S} \right)^\gamma - 1 \right] \quad (2.18)$$

avec $\gamma = \frac{R_{\text{dry}}}{\Phi_S} (T_0 - T_S)$ et R_{dry} la constante des gaz parfaits pour l'air sec.

2.2.2.2 Température à 2 m

L'opérateur d'observation pour la température à 2 m réalise une interpolation verticale entre le premier niveau du modèle et la surface, qui dépend de la stabilité de la couche limite. L'ensemble de la démonstration est décrit dans Geleyn (1988).

On utilise l'énergie statique sèche $s = c_p T + \phi$ et la théorie de Monin–Obukhov, pour définir le gradient vertical de s près du sol :

$$\frac{\partial s}{\partial z} = \frac{s_*}{\kappa(z + z_0)} \phi_H \left(\frac{z + z_0}{L} \right) \quad (2.19)$$

Avec L la longueur de Obukhov :

$$L = \frac{c_p T u_*^2}{g \kappa s_*} \quad (2.20)$$

où z la hauteur, z_0 est la longueur de rugosité, u le vent horizontal, u_* et s_* le vent et l'énergie statique de frottement (c.-à-d., la composante du second ordre), $\kappa = 0.4$ la constante de von Kármán, c_p la capacité calorifique et ϕ_H une fonction analytique

13. *International Standard Atmosphere*

14. Organisation de l'aviation civile internationale

prescrite.

Il ne reste qu'à intégrer deux fois l'eq. 2.19 entre le sol et le niveau de l'observation z , ainsi qu'entre le sol et le niveau le plus bas du modèle z_1 . On posera B_N le coefficient neutre et B_H le coefficient de chauffage.

Dans le cas stable (le nombre de Richardson $Ri \geq 0$) :

$$s(z) = \tilde{s} + \frac{s(z_1) - \tilde{s}}{B_H} \left[\ln \left(1 + \frac{z}{z_1} (e^{B_N} - 1) \right) - \frac{z}{z_1} (B_N - B_H) \right] \quad (2.21)$$

avec \tilde{s} l'énergie statique sèche à la surface.

Dans le cas instable ($Ri < 0$) :

$$s(z) = \tilde{s} + \frac{s(z_1) - \tilde{s}}{B_H} \left[\ln \left(1 + \frac{z}{z_1} (e^{B_N} - 1) \right) - \ln \left(1 + \frac{z}{z_1} (e^{B_N - B_H} - 1) \right) \right] \quad (2.22)$$

On obtient alors, au niveau $z = 2$ m, l'énergie statique sèche, puis la température T .

La figure 2.1 illustre les fonctions d'interpolation verticale entre le premier niveau du modèle et la surface, pour le modèle global IFS. L'interpolation verticale décrite précédemment pour AROME correspond au cy48 d'IFS. En pratique, l'interpolation s'éloigne de la théorie pour des fortes valeurs de stabilité, montrant les limites de l'opérateur d'observation de température. Cet écart est corrigé dans la dernière version mise en place du modèle IFS (cycle 49), en introduisant une fonction d'interpolation modifiée, plus proche de la théorie et plus lisse, évitant le saut abrupt vers la température du niveau le plus bas du modèle à une valeur fixe de stabilité.

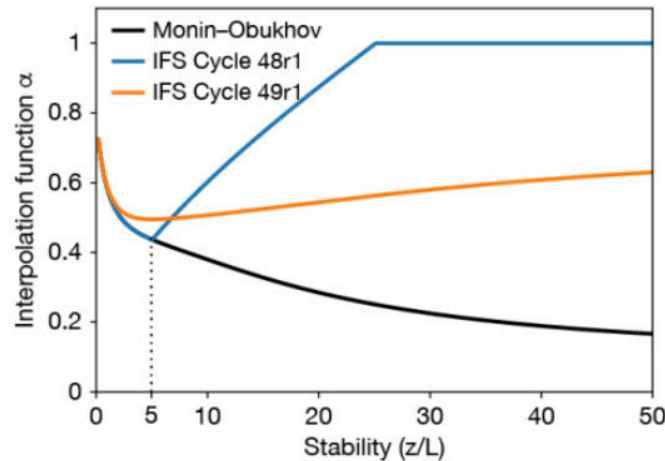


FIGURE 2.1 – Schéma de la fonction d'interpolation α de la température à 2 m (t2m) sur la verticale, en fonction de la stabilité (Ingleby *et al.*, 2024). $\alpha = 1$ indique que la t2m est égale à la température au niveau le plus bas du modèle et $\alpha = 0$ quand la t2m est égale à la température de surface. En noir la fonction basée sur la théorie de Monin-Obukhov, en bleu la formule d'interpolation actuelle (cycle 48r1) et en orange la fonction qui sera introduite dans le prochain cycle d'IFS (cycle 49r1).

2.2.2.3 Humidité relative à 2 m

L'humidité relative est calculée selon :

$$RH = \frac{100Pq}{e_{\text{sat}}\left(\frac{R_{\text{dry}}}{R_{\text{vap}}} + q\left(1 - \frac{R_{\text{dry}}}{R_{\text{vap}}}\right)\right)}, \quad (2.23)$$

avec e_{sat} la pression de vapeur saturante en hPa, P la pression de surface en hPa, q l'humidité spécifique sans dimension, $R_{\text{dry}} = 287.0597 \text{ J kg}^{-1} \text{ K}^{-1}$ et $R_{\text{vap}} = 461.5250 \text{ J kg}^{-1} \text{ K}^{-1}$ sont les constantes des gaz pour l'air sec et la vapeur d'eau, respectivement, au niveau le plus bas du modèle. Il est à noter que le calcul de l'humidité relative à 2 m du modèle ne prend pas en compte l'humidité spécifique de surface, mais seulement celui du niveau le plus bas du modèle.

2.2.3 Espace des observations

Le terme de distance aux observations de la fonction de coût J_o peut être ajusté en fonctions des observations utilisées via la matrice de covariances d'erreurs d'observation \mathbf{R} . Deux méthodes supplémentaires seront présentées, et peuvent être activées indépendamment pour chaque système d'observations : la correction de biais variationnelle (VarBC) et le contrôle de qualité variationnel (VarQC) des observations.

2.2.3.1 Réglage des covariances d'erreur d'observations

La matrice de covariances d'erreurs d'observation \mathbf{R} est prescrite dans le système d'assimilation d'AROME. Afin de simplifier la résolution du problème de l'assimilation de données, la matrice \mathbf{R} utilisée dans AROME est diagonale. Les termes diagonaux indiquent le niveau de confiance que l'on accorde à chaque système d'observation. Cette hypothèse de diagonalité est lourde, car elle suppose que les erreurs d'observation sont décorréelées. Cette hypothèse peut se révéler fautive pour des observations spatialement proches comportant un biais similaire, causé par les observations ou par l'opérateur d'observation. Dans le modèle AROME, cette dernière contrainte conduit à l'écrémage des données à haute définition avec une distance minimale entre chaque observation assimilée : de 15 km pour les observations des radars, 25 km pour les observations d'avions et 70 km pour les observations de SEVIRI¹⁵ (Brousseau *et al.*, 2014).

De nombreux centres de météorologie s'intéressent à modéliser une matrice \mathbf{R} non diagonale, permettant d'assimiler plus de données à plus fine échelle. Les techniques de modélisation d'une matrice \mathbf{R} non diagonale sont encore nouvelles en assimilation de données. Une première piste est donnée par Guillet *et al.* (2019). La matrice \mathbf{R} étant trop lourde pour être stockée (jusqu'à 10^{14} coefficients non nuls) et comme il n'est pas possible d'effectuer les calculs dans l'espace spectral (les observations étant positionnées sur une grille irrégulière) on remplace alors le coût en mémoire par un coût en calcul. On fait l'hypothèse que \mathbf{R} n'est plus diagonale, mais diagonale par bloc de système d'observations, ce qui revient à faire l'hypothèse que les erreurs de différents systèmes d'observations sont décorréelées entre elles. On utilise alors une méthode de diffusion qui remplace l'inverse de la matrice de covariances d'erreurs d'observation \mathbf{R}^{-1} par un produit de la forme \mathbf{L}^m , où \mathbf{L} est une matrice creuse facile à stocker et m un exposant. En pratique, \mathbf{L} provient

15. *Meteosat Spinning Enhanced Visible and Infrared Imager*

de la discrétisation d'un opérateur de diffusion dont les solutions sont des fonctions de Matérn qui ont les propriétés adéquates (par exemple, ressemblent à des gaussiennes, décroissent en fonction de la distance). La modélisation de \mathbf{R}^{-1} , testée pour l'assimilation d'observations SEVIRI admet encore des erreurs trop importantes, de l'ordre de 10 à 30 %, pour être utilisable opérationnellement (Guillet *et al.*, 2019).

2.2.3.2 Diagnostic de Desroziers pour l'estimation d'erreur d'observation

Desroziers *et al.* (2005) démontre, de manière statistique et géométrique, qu'il est possible d'estimer a posteriori la matrice de covariances d'erreurs d'observation. Le diagnostic de Desroziers est tel que :

$$\mathbb{E}[\mathbf{d}_a^o(\mathbf{d}_b^o)^T] = \mathbf{R} \quad (2.24)$$

où les vecteurs d'innovation \mathbf{d}_a^o et \mathbf{d}_b^o correspondent respectivement aux observations moins l'analyse et aux observations moins l'ébauche. Il est important de noter que la méthode Desroziers ne fournit qu'une estimation des erreurs. La méthode suppose que les covariances des erreurs d'ébauche et d'observation sont spécifiées avec précision et que l'erreur d'ébauche ne soit pas corrélée aux erreurs d'observation. En pratique, les séries temporelles d'observation étant parfois biaisées par rapport à l'ébauche, le diagnostic admet des limites qui seront explorées dans le chapitre 3.

2.2.3.3 Correction variationnelle de biais

La correction de biais variationnelle (VarBC) est une méthode embarquée au sein de la minimisation de la fonction de coût pour corriger les observations. VarBC correspond à une régression linéaire utilisant N prédicteurs liés au biais des observations. Chaque prédicteur est associé à un coefficient, qui sont rassemblés dans un vecteur nommé $\boldsymbol{\beta}$. L'opérateur d'observation est ainsi modifié, tel que :

$$\tilde{\mathbf{H}}(\mathbf{x}, \boldsymbol{\beta}) = \mathbf{H}(\mathbf{x}) + \sum_{i=1}^N \beta_i p_i(\mathbf{x}) \quad (2.25)$$

où $p_i(\mathbf{x})$ un prédicteur (une variable modèle, une autre observation, une métadonnée ou une combinaison des trois) et β_i son coefficient associé. L'entraînement de la correction de biais consiste à trouver le vecteur $\boldsymbol{\beta}$ qui s'ajuste le mieux entre le modèle et les observations. Avec cette formulation, la fonction de coût s'adapte :

$$\begin{aligned} J(\mathbf{x}, \boldsymbol{\beta}) = & \frac{1}{2} [\mathbf{y}^o - \tilde{\mathbf{H}}(\mathbf{x}, \boldsymbol{\beta})]^T \mathbf{R}^{-1} [\mathbf{y}^o - \tilde{\mathbf{H}}(\mathbf{x}, \boldsymbol{\beta})] \\ & + (\boldsymbol{\beta} - \boldsymbol{\beta}^b)^T \mathbf{B}_\beta^{-1} (\boldsymbol{\beta} - \boldsymbol{\beta}^b) \\ & + (\mathbf{x} - \mathbf{x}^b)^T \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}^b) \end{aligned} \quad (2.26)$$

où \mathbf{B}_β est la matrice de covariances d'erreurs d'ébauche du paramètre de biais. Le deuxième terme dans l'éq. 2.26 contrôle l'adaptabilité du paramètre de biais, que l'on nomme aussi la contrainte d'inertie. On peut voir le vecteur $\boldsymbol{\beta}$ comme une extension de la variable de contrôle qui sera définie de façon optimale pour minimiser la fonction coût. À chaque cycle d'assimilation, on initialise $\boldsymbol{\beta}$ avec les valeurs déterminées au réseau précédent, en commençant par une valeur arbitraire (dépendant de la connaissance a priori des biais). Il est courant d'utiliser des analyses cyclées du modèle pour entraîner la correction de biais durant une période de l'ordre d'un mois (Auligné *et al.*, 2007). Si l'on compare cette

méthode à une correction de biais *offline*, c'est-à-dire réalisée avant la minimisation de la fonction de coût, alors la correction n'est pas capable de faire la différence entre un biais modèle ou un biais d'observations. Une correction de biais *offline* va corriger les observations pour s'adapter au modèle. Le grand avantage de la méthode de correction de biais variationnelle est de pouvoir distinguer le biais du modèle de celui des observations. Ceci peut se faire à condition de maintenir un système d'observation d'ancrage qui évite à la correction de biais d'être contaminée par le biais du modèle ou même de dériver. Toutes les observations non corrigées par le modèle peuvent servir d'ancrage, grâce à la matrice de covariance d'erreur d'ébauche. Néanmoins, les observations d'ancrage les plus efficaces pour réduire l'effet du biais du modèle sont celles qui observent la même variable d'état que les observations corrigées. Elles servent d'autant plus qu'elles sont fiables et localisées proches des observations à corriger (Francis *et al.*, 2023).

À titre d'exemple, la méthode VarBC s'est montrée pertinente pour corriger les observations de vent provenant d'avions, en utilisant comme prédicteurs la direction et la vitesse de l'avion (Pourret *et al.*, 2021).

2.2.3.4 Contrôle de qualité variationnel

Le contrôle de qualité variationnel (VarQC) correspond à une modification de la fonction de coût des observations, en prenant en compte la nature non gaussienne des erreurs d'observation (Anderson et Järvinen, 1999). Un poids W est défini pour chaque observation, suivant le rapport :

$$W = \frac{J_o^{QC}}{J_o} \quad (2.27)$$

avec J_o la précédente fonction de coût des observations faisant l'hypothèse d'erreur gaussienne et J_o^{QC} prenant en compte la non-gaussianité des erreurs d'observations.

Une adaptation de la méthode VarQC, avec la norme de Huber, est opérationnelle pour le modèle global IFS (Tavolato et Isaksen, 2015). Cette méthode permet de réduire le poids d'une observation éloignée de l'ébauche lors de la minimisation de la fonction de coût. En parallèle, cette méthode permet de prescrire une erreur d'observation σ_o plus faible et un seuil de rejet lors du *screening* moins contraignant afin d'utiliser des observations plus éloignées de l'ébauche.

Néanmoins, cette méthode étant basée sur un formalisme bayésien, elle nécessite de pouvoir ajuster le poids des observations à chaque itération du schéma variationnel. Ceci est possible pour des schémas réalisant plusieurs linéarisations de l'opérateur d'observation (boucle externe), ce qui n'est pas le cas pour AROME, et est moins efficace dans les itérations de la minimisation (boucle interne) (Tavolato et Isaksen, 2015).

2.2.4 Réglage des covariances d'erreur d'ébauche

La matrice de covariances des erreurs d'ébauche \mathbf{B} joue un rôle central dans le processus d'assimilation de données. Les variables de contrôle, mises à jour par le processus d'assimilation et intégrées dans \mathbf{B} , sont au nombre de cinq : la température, l'humidité spécifique, la pression de surface, le tourbillon et la divergence du vent. Il y a moins de variables de contrôle que de variables pronostiques dans AROME. Autrement dit, certaines variables pronostiques nécessaires au modèle de prévision ne sont pas incluses dans \mathbf{B} , et sont dites *cyclées*, c'est-à-dire qu'à l'analyse elles sont copiées de l'ébauche précédente. On y retrouve les hydrométéores, la pression non-hydrostatique, la divergence verticale

du vent, et l'énergie cinétique turbulente. Les éléments diagonaux de \mathbf{B} , aussi notés σ_b , caractérisent l'incertitude associée à la prévision de l'ébauche. Les éléments non diagonaux, quant à eux, caractérisent la structure des incréments apportés par une observation sur l'horizontale, la verticale et entre les variables de contrôle.

La matrice \mathbf{B} ne peut être stockée en raison de sa très grande dimension. Avec une résolution de 1440×1536 points de grille, à multiplier par les 90 niveaux et le nombre de variables de contrôle, un vecteur d'état du modèle possède 10^9 éléments. La matrice \mathbf{B} possède par conséquent 10^{18} coefficients. La matrice \mathbf{B} doit être modélisée et calculée à chaque analyse. \mathbf{B} est modélisée sous la forme $\mathbf{K}\mathbf{B}_u\mathbf{K}^T$ avec \mathbf{K} contenant des opérateurs d'équilibre verticaux et \mathbf{B}_u une matrice diagonale par bloc contenant les covariances d'erreurs verticales pour chaque variable de contrôle (Berre, 2000).

2.2.4.1 Matrice \mathbf{B} climatologique

Dans le 3DVar, la matrice \mathbf{B} est statique, c'est-à-dire qu'elle est calculée une fois et est valable pour tous les réseaux. De plus, \mathbf{B} est homogène et isotrope, c'est-à-dire que l'incrément d'analyse généré par l'innovation d'une observation sera identique, quelle que soit sa position sur le domaine, et se propage de manière identique dans toutes les directions.

Bien que la matrice \mathbf{B} soit climatologique, elle a évolué au fil des versions d'AROME et de l'évolution de la prévision d'ensemble. Au début d'AROME, fin 2008, \mathbf{B} est estimée grâce à un ensemble AROME-France de six membres décrit par Seity *et al.* (2011), exécuté en temps différé et réalisant des prévisions à 3 h d'échéance, deux fois par jour, pendant 26 jours, précurseur du système actuel d'assimilation d'ensemble AROME. En 2015, la chaîne de prévision AROME passe à une résolution de 1.3 km, avec une mise à jour rapide du cycle d'assimilation (RUC) toutes les heures. Une interpolation de la matrice \mathbf{B} à cette nouvelle résolution est réalisée. L'erreur de l'ébauche à 1 h d'échéance étant plus faible qu'à 3 h d'échéance, les variances d'erreur d'ébauche σ_b ont été réduites, tout en considérant que les covariances (verticales, horizontales et croisées) sont identiques de 1 h à 3 h d'échéance (Brousseau *et al.*, 2016).

2.2.4.2 Matrice \mathbf{B} dépendante du flux

Avec la mise en place en temps réel d'un cycle d'assimilation-prévision ensembliste (EDA¹⁶), la possibilité d'avoir une matrice de covariances d'erreurs d'ébauche \mathbf{B} actualisée à chaque analyse à émergé (Michel et Brousseau, 2021). L'EDA d'AROME est composé de 50 membres à 3.2 km de résolution horizontale et de 90 niveaux verticaux. Le schéma d'assimilation devient alors **variationnel ensembliste (EnVar)**. On notera \mathbf{B}_e la matrice de covariances des erreurs d'ébauche dépendante de la situation météorologique (aussi appelée « dépendante du flux », ou *flow-dependant*), et estimée sous la forme :

$$\mathbf{B}_e = \frac{1}{N-1} \sum_{k=1}^N (\mathbf{x}_k^b - \bar{\mathbf{x}}^b)(\mathbf{x}_k^b - \bar{\mathbf{x}}^b)^T \quad (2.28)$$

où N est le nombre de membres de l'ensemble, \mathbf{x}_k^b l'ébauche du membre k et $\bar{\mathbf{x}}^b$ la moyenne de l'ensemble.

16. *Ensemble Data Assimilation*

La méthode d'assimilation d'ensemble variationnelle 3DEnVar correspond au schéma d'assimilation 3DVar, auquel on injecte la matrice de covariances d'erreurs d'ébauche dépendante du flux \mathbf{B}_e . Le schéma d'assimilation 3DEnVar permet d'avoir des incréments d'analyse non isotropes, contrairement au 3DVar, qui sont liés aux conditions prévues par l'ensemble. Il n'est plus nécessaire de modéliser les matrices d'opérateurs pour définir \mathbf{B} comme c'était le cas en 3DVar. Ainsi, il est plus aisé de rajouter des nouvelles variables de contrôle tant qu'elles sont prévues par le modèle, c'est-à-dire incluses dans les variables pronostiques. C'est le cas des contenus spécifiques en hydrométéores qui sont utiles comme variable de substitution en vue de l'assimilation directe des observations RADAR ou de nouvelles observations comme les éclairs (Combarrous, 2023).

Toutefois, à cause du faible nombre de membres contenu dans l'ensemble (≈ 50) comparé à la taille du vecteur d'état ($\approx 10^9$), du bruit d'échantillonnage apparaît dans l'estimation des corrélations d'erreur d'ébauche. La qualité de l'estimation est également tributaire de la qualité de l'ensemble de prévisions utilisé. Afin de réduire le bruit d'échantillonnage, une localisation spatiale est appliquée (Hamill *et al.*, 2001), ce qui limite la propagation de l'innovation à des distances plus courtes. La matrice localisée $\hat{\mathbf{B}}_e$ est obtenue en faisant un produit élément à élément (connu sous le nom de produit de Hadamard-Schur) entre \mathbf{B}_e et une matrice de corrélation \mathbf{L} de même dimension :

$$\hat{\mathbf{B}}_e = \mathbf{B}_e \circ \mathbf{L}, \quad (2.29)$$

Les fonctions de corrélation dans la matrice \mathbf{L} sont quasi gaussiennes avec des localisations horizontales et verticales (Montmerle *et al.*, 2018). La longueur spécifiée de la localisation correspond à la largeur à mi-maximum de la fonction quasi gaussienne. Le réglage de la localisation horizontale a été choisi de sorte qu'elle parte de 25 km au sol et décroisse linéairement avec l'altitude pour atteindre 150 km au niveau le plus haut du modèle (Michel et Brousseau, 2021). La localisation verticale est fixée à environ 0.15 hPa (Michel et Brousseau, 2021).

Des méthodes plus avancées permettent de s'affranchir d'une localisation fixe à un niveau donné, comme le *Scale Dependent Localisation* (SDL), où l'on applique des échelles de longueur de localisation appropriées, à différentes gammes d'échelles spatiales de covariance d'erreurs d'ébauche (Caron *et al.*, 2019). Cette méthode est déjà opérationnelle pour le modèle global canadien (Caron et Buehner, 2022).

2.2.5 Schéma d'assimilation de surface

AROME dispose d'un schéma d'assimilation dédié pour l'initialisation des variables de surface. Ces variables de surface sont distinctes pour les quatre tuiles de surface (chapitre 1). Pour la tuile nature, les variables de surface correspondent à la température et à un réservoir d'eau, à deux niveaux dans le sol (Giard et Bazile, 2000).

Le schéma d'assimilation de surface se sert des variables de température et d'humidité relative au niveau 2 m, qui ne sont pas des variables de contrôle (figure 2.2). La première étape consiste à les calculer avec l'opérateur d'observation de surface décrit précédemment. Ensuite, le schéma d'interpolation optimale univariée (IO) vient combiner les observations à 2 m avec l'ébauche à 2 m. La distance caractéristique D utilisée dans le schéma IO est fixée à 100 km pour les observations de température et d'humidité relative à 2 m. Cette distance caractéristique est un compromis entre la résolution du modèle, l'échelle des phénomènes météorologiques représentés et la densité des stations de surface. Ensuite, les incréments d'analyse de la température et de l'humidité relative à 2 m sont utilisés pour

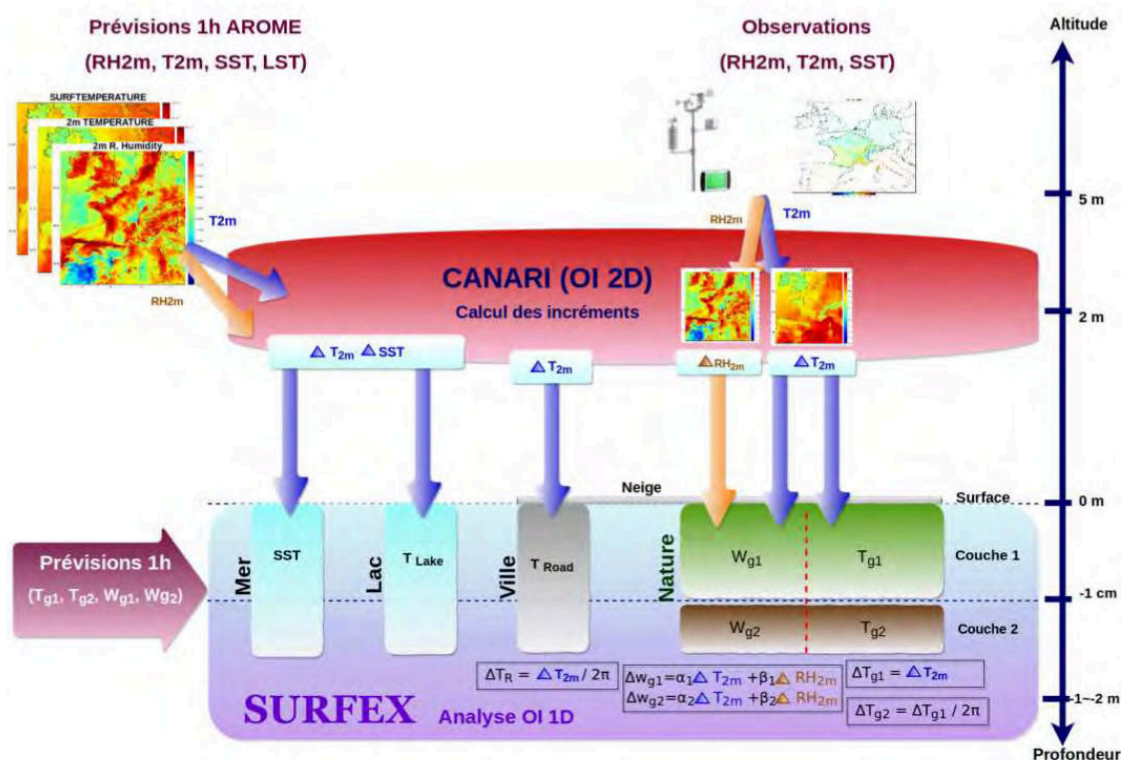


FIGURE 2.2 – Système d’assimilation de surface dans le modèle AROME de type interpolation optimale. Figure extraite de Sassi *et al.* (2019).

corriger les variables du sol avec des interpolations linéaires. Plus de détails sur le schéma d’assimilation de surface sont donnés par Sassi *et al.* (2019).

2.2.6 Chaîne de prévision opérationnelle

Un cycle d’assimilation et de prévision opérationnel est continuellement exécuté sur les super-calculateurs de Météo-France. La figure 2.5 schématise son fonctionnement. Le modèle régional AROME-France a besoin de conditions aux limites (LBC) qui correspondent au modèle global ARPEGE opérationnel à Météo-France. Le cycle actuel d’assimilation d’AROME-France est horaire : l’ébauche provient d’une prévision à $h - 1$ pour réaliser l’analyse à h . La fenêtre d’assimilation est de 1 h centrée sur l’heure d’analyse : toutes les observations disponibles entre $[h - 30 \text{ min}, h + 30 \text{ min}[$ sont récupérées pour l’analyse de l’heure h . Toutes les 3 h une analyse de surface est réalisée. À partir de ces analyses, toutes les 3 h, de longues prévisions jusqu’à 51 h d’échéance sont lancées. On dénombre huit longues prévisions par jour à 00:00 UTC, 03:00 UTC, 06:00 UTC, 09:00 UTC, 12:00 UTC, 15:00 UTC, 18:00 UTC et 21:00 UTC. Une longue prévision lancée à l’heure h utilise l’analyse à l’heure h , mais aussi l’analyse à l’heure $h + 1$ déjà disponible : c’est ce qu’on appelle l’*Analysis Incremental Update* (IAU, Brousseau *et al.*, 2016). Le terme IAU désigne aussi une autre méthode qui consiste à introduire progressivement l’incrément d’analyse durant la première heure de prévision, en l’ajoutant petit à petit à chaque pas de temps. Cette technique permet de réduire le temps de retour à l’équilibre du modèle après le choc de l’assimilation.

Une panoplie d’observations est assimilée dans le modèle AROME. La figure 2.3 représente la proportion des observations actuellement assimilées en moyenne sur le mois

d'étude du 6 septembre au 5 octobre 2021.

Sur ce mois, 49 253 observations sont assimilées chaque heure, se répartissant selon :

- Les observations RADAR (56.5 %) françaises et européennes, sous la forme de vitesse radiale (dopp) et d'humidité relative (rh). On notera que le nombre d'observations RADAR assimilées varie, passant de 3000 au minimum en temps calme à plus de 78 000 observations en temps pluvieux.
- Les observations de surface issues des messages SYNOP (17 %) décrites dans le chapitre 1. Les observations de surface incluent les observations GPS-sol (apdss), donnant une mesure indirecte et intégrée sur toute la colonne d'atmosphère de l'humidité relative.
- Les observations de satellites SATEM (9.8 %), donnant des températures de brillance (rawbt). Ces observations proviennent de satellites à orbites basses, présentant un pic d'observations à 10:00 UTC et 21:00 UTC, et du radiomètre SEVIRI de la mission Météosat 2^e génération.
- Les observations de radiosondages TEMP (9.6 %) lancés à 00:00 UTC et 12:00 UTC à Trappes, Nîmes, Ajaccio, Bordeaux et Brest, ainsi que dans les pays voisins, notamment en Allemagne.
- Les observations d'avions AIREP (6.1 %). Leur nombre est variable au cours de la journée : constant autour de 5000 en moyenne horaire entre 05:00 UTC et 19:00 UTC, avant de décroître en soirée pour devenir très faible pendant la nuit (figure 2.4b).
- Les observations de profileurs de vent PILOT (1.0 %).
- D'autres observations sont assimilées telles que les observations de bouées dérivantes et ancrées.

En parallèle de la chaîne opérationnelle, la chaîne en double d'AROME-France utilise le schéma d'assimilation 3DEnVar. Celui-ci se repose sur un cycle d'assimilation ensembliste (EDA d'AROME) produisant toutes les 3 h des prévisions ensemblistes permettant

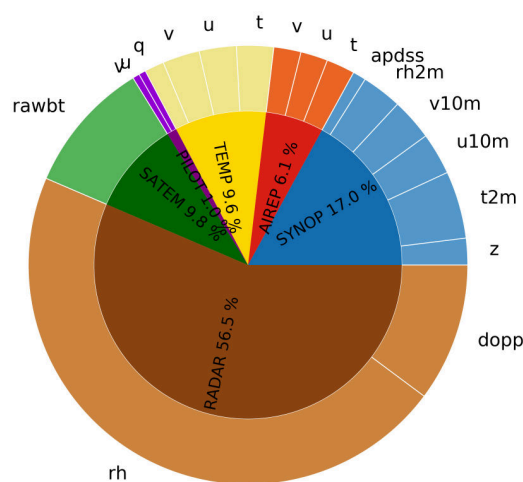


FIGURE 2.3 – Proportion moyenne du nombre d'observations actives (qui ont passé le *screening*) assimilées par réseau entre le 6 septembre et le 5 octobre 2021. Seules les observations de bouées ne sont pas intégrées, représentant moins de 1 % des observations.

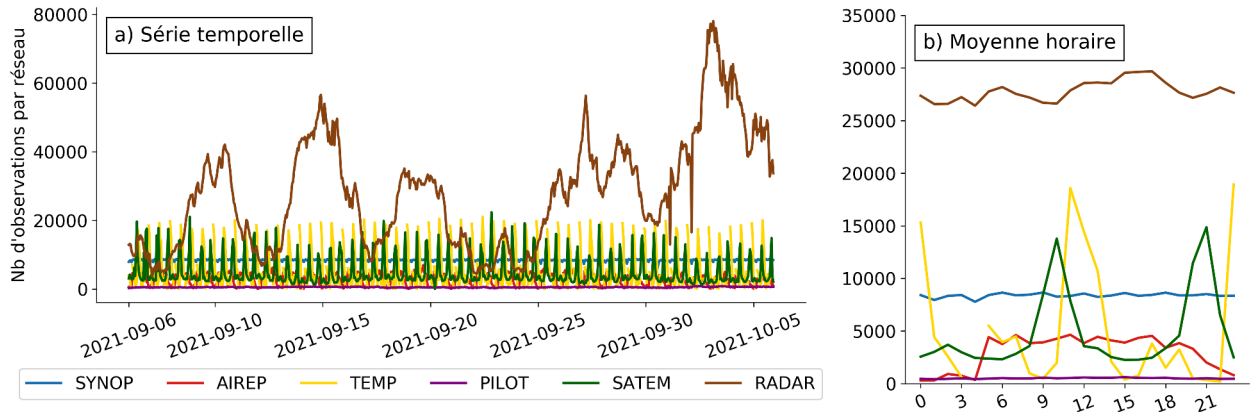


FIGURE 2.4 – Nombre d'observations par type d'observation, (a) par réseau successif et (b) en moyenne horaire, pendant la période du 6 septembre au 5 octobre 2021. Voir le texte pour une description des types d'observations.

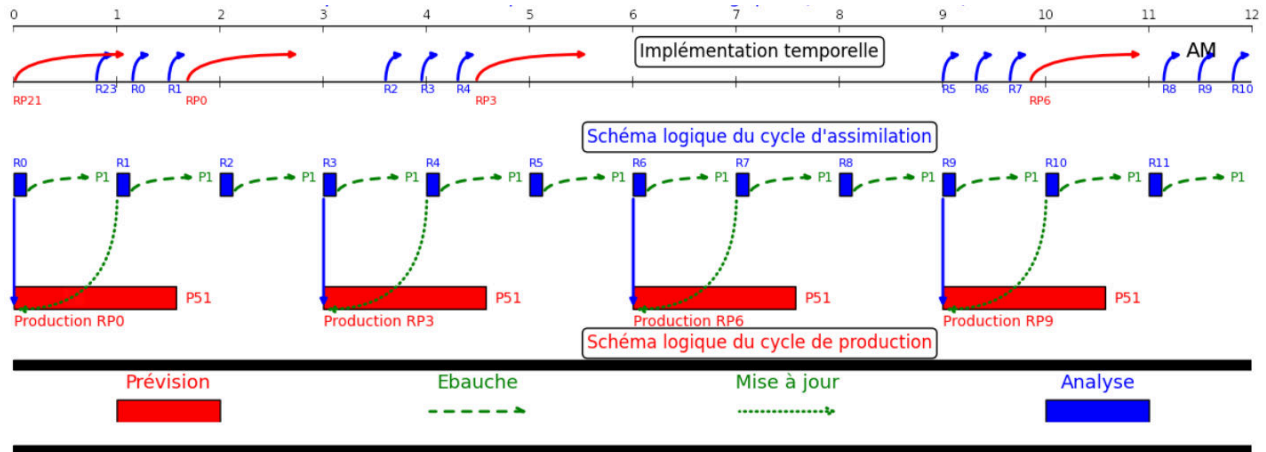


FIGURE 2.5 – Chaîne de prévision opérationnelle pour AROME-France de 00:00 à 12:00 UTC (source : <http://controle.meteo.fr>).

le calcul de la matrice d'erreur d'ébauche \mathbf{B}_e . On se sert des prévisions à 5, 6 et 7 h d'échéance, car les réseaux plus récents ne sont pas encore disponibles. Par exemple, pour produire l'analyse AROME de 07:00 UTC, on utilise la prévision à 7 h d'échéance du réseau de 00:00 UTC de l'EDA d'AROME.

2.3 Assimilation des observations de surface

2.3.1 Quel impact ont les observations de surface ?

Dans la section précédente, nous avons décrit le nombre d'observations disponibles à chaque analyse. Cependant, du fait de leur positionnement dans l'espace, des erreurs d'observation prescrites, ainsi que de la matrice de covariances d'erreurs d'ébauche, les observations n'ont pas toutes le même poids dans l'analyse.

Le calcul de l'importance relative des observations a été effectué avec une version d'AROME antérieure à 2015 (Brousseau *et al.*, 2014) et sert encore de référence pour le cycle opérationnel. Brousseau *et al.*, 2014 reprennent la formulation de Desroziers *et al.*, 2005 pour estimer l'importance relative de chaque système d'observations lors d'une journée pluvieuse. Cette méthode permet d'estimer la réduction de variance de l'ébauche

à l'analyse fournit par un système d'observation. Il y montre que les observations à 2 m sont les observations qui permettent la principale réduction de variance dans la couche limite atmosphérique¹⁷. Pour les champs de température et en dessous de 900 hPa, la réduction de la variance est principalement assurée par les mesures de température à 2 m. De plus, l'information fournie par les observations de température à 2 m est propagée de façon décroissante de la surface jusqu'à 900 hPa par les corrélations verticales de la matrice de covariances d'erreur d'ébauche. Pour les champs d'analyse d'humidité spécifique, les observations d'humidité relative à 2 m sont les observations qui permettent la principale réduction de variance du sol jusqu'à 800 hPa. Ainsi, les corrélations verticales sont plus importantes pour les observations d'humidité relative à 2 m que pour celles de température à 2 m. Par ailleurs, la réduction de variance du champ de température atteint un maximum de 80 % à la surface. Autrement dit, même en 3DVar, une part des champs de la couche limite atmosphérique en température et en humidité spécifique ne sont pas contraints du fait d'un manque d'observations de surface.

Une quantification de l'influence des observations dans AROME a aussi été effectuée par Fischer *et al.* (2018), utilisant la méthode DFS (*Degrees of Freedom for Signal*). Le DFS est une mesure de l'importance relative de chaque système d'observation dans l'incrément d'analyse. Ce dernier a montré que lors d'une journée pluvieuse de février 2022, le pourcentage d'observations de surface et leur importance relative donné par le DFS, étaient identiques, soit égal à 9%. Pendant les journées peu pluvieuses, où peu d'observations radar sont assimilées, le DFS des observations de surface monte jusqu'à 38%.

Dans un système 3DVar, il est important d'avoir des observations à différents niveaux du modèle. En effet, à l'époque du modèle ALADIN (ancêtre d'AROME) en 2005, les observations à 2 m sont considérées complémentaires aux observations de radiances SEVIRI en milieu de troposphère, évitant une propagation trop importante de ces dernières vers des couches basses du modèle à cause des fonctions de structure de grande dimension du 3DVar sur la verticale (Fischer *et al.*, 2005). Cet exemple permet de comprendre la compétition qui règne entre les différents systèmes d'observation, et l'influence parfois néfaste qu'elles engendrent, liée à la structure de la matrice \mathbf{B} .

De plus, l'impact d'un système d'observation dépend du schéma d'assimilation utilisé. Le réglage de la matrice de covariances d'erreurs d'ébauche peut favoriser un système d'observation par rapport à un autre (Lenfle, 2018). Il n'existe pas à ce jour d'études sur l'influence des observations au sol dans cette version du modèle en utilisant un schéma d'assimilation EnVar.

2.3.2 Pourquoi l'assimilation d'observations standards de surface est-elle un challenge ?

Pour certains modèles de prévision atmosphérique, l'assimilation d'observations de surface ne s'est pas fait immédiatement. Deux exemples en montrent bien les différentes difficultés.

D'abord, prenons l'exemple des modèles globaux et des défis auxquels ils font face. Le modèle global IFS du centre européen (CEPMMT), utilise une assimilation 4DVar. Dans IFS, les observations de t2m et de rh2m sont assimilées dans le modèle de surface,

17. La couche limite atmosphérique est la partie de l'atmosphère qui est influencée par la surface, et est généralement contenue entre le sol et 850 hPa (≈ 1500 m).

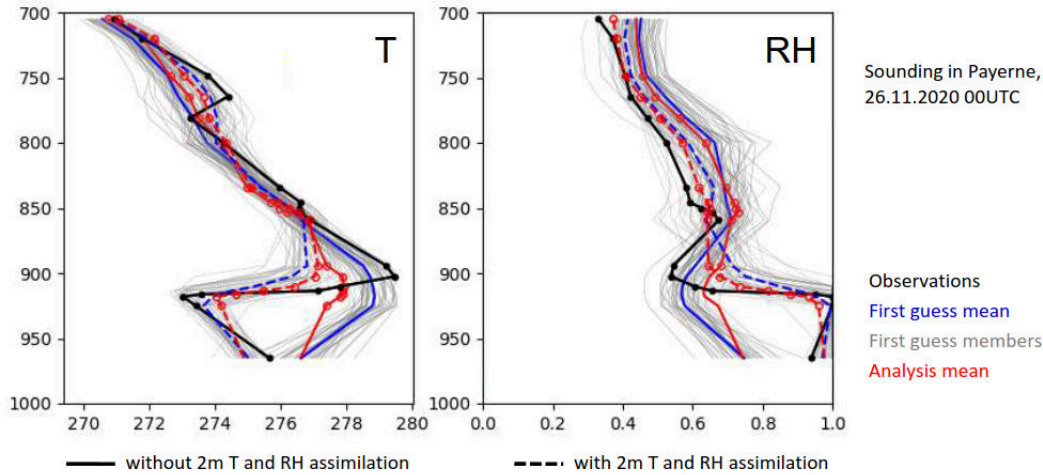


FIGURE 2.6 – Profils de température et d’humidité relative lors d’un épisode d’inversion le 26 novembre 2020 à 01:00 UTC à Payerne (Suisse). Expériences avec et sans l’assimilation d’observations de t_{2m} et de rh_{2m} . Figure issue de (Merker *et al.*, 2023).

mais seules les observations de rh_{2m} sont assimilées dans le schéma d’assimilation atmosphérique 4DVar, et seulement le jour. Ceci est aussi le cas du modèle global ARPEGE opérationnel à Météo-France. Dans IFS, l’assimilation de t_{2m} était responsable de la dégradation de la température à 850 hPa dans des conditions stables hivernales (Ingleby *et al.*, 2024). Parmi les solutions mises en place, on notera l’amélioration de l’opérateur d’observation, plus proche de la formulation théorique, mentionné précédemment, et la réduction du poids donné aux observations de t_{2m} si leur écart à l’ébauche est trop important, par exemple avec VarQC.

Le deuxième exemple concerne le modèle régional suisse COSMO, où l’assimilation des observations de surface ne s’est faite qu’en septembre 2021. COSMO possède un schéma d’assimilation LETKF, nommé KENDA¹⁸. Pour mettre en évidence l’impact des observations de surface, deux expériences avec et sans l’assimilation d’observations de t_{2m} et de rh_{2m} ont été réalisées durant un épisode d’inversion à Payerne (figure 2.6). Ce n’est qu’avec l’assimilation des observations de surface que le modèle COSMO est capable de simuler la couche d’inversion, identifiée par le radiosondage. A contrario, il existe aussi des situations d’inversion de basse couche très marquées, où l’assimilation d’observations de surface dégrade l’analyse en basse couche. En effet, les observations de surface permettent de recalibrer le modèle sur ses premiers niveaux, mais en influençant les niveaux supérieurs avec des covariances d’erreurs d’ébauche verticale trop importantes, elles dégradent les profils de température et d’humidité relative au niveau de la couche d’inversion. Parmi les solutions proposées, on trouve le réglage de la localisation horizontale et verticale ainsi que le rejet de stations si elles sont trop éloignées du modèle.

En résumé, l’impact d’une observation de surface change suivant la situation météorologique et le nombre d’observations disponibles, ce dont les schémas d’assimilation de données peinent à intégrer. Ceci représente un verrou pour l’assimilation d’observations de t_{2m} et de rh_{2m} . Ces deux exemples montrent qu’il est important de s’intéresser simultanément à :

- l’impact spatial autour de l’observation, réglé par la matrice de covariance d’erreur d’ébauche ;

18. *Kilometre-scale ENsemble Data Assimilation*

— l'amplitude de l'incrément, avec le réglage de l'erreur d'observation σ_o .

2.3.3 Quels défis pour l'assimilation d'observations d'objets connectés ?

Les observations provenant d'objets connectés, que ce soit des smartphones ou des PWS, sont une source d'observations importante. Pour les observations des PWS, elles représentent quinze fois plus de stations que le réseau SWS de Météo-France (Mandement et Caumont, 2020). Ces observations sont non seulement soumises aux mêmes difficultés rencontrées par les observations standards, mais présentent également des défis supplémentaires liés à leur densité (notamment en ville), et à leur qualité.

Des premiers tests d'assimilation d'observations de pression de PWS ont été réalisés dans AROME, sans amélioration de scores de prévision ou de représentations de phénomènes précipitants (Hintz, 2019 ; Ridal *et al.*, 2019 ; Scheffknecht *et al.*, 2021). Ces études n'étant que préliminaires, les résultats n'ont pas été soumis au processus d'évaluation par les pairs. Néanmoins, des premières pistes d'enseignement peuvent en être déduits. D'abord les observations doivent faire l'objet d'un traitement spécifique. De plus, il est difficile d'extraire une information des systèmes d'observation de forte densité avec le schéma d'assimilation 3DVar. Celui-ci possède des fonctions de structure de grande échelle peu adaptées à la méso-échelle. A contrario, l'assimilation d'observations de pression provenant de smartphones avec un schéma d'assimilation ensembliste a montré des améliorations de la prévision à 1 h d'échéance sur la température et l'humidité à 2 m ; et sur la trajectoire et l'intensité d'une dépression (McNicholas et Mass, 2018b). L'assimilation des observations de température et d'humidité relative des PWS avec un schéma d'assimilation LETKF dans ICON-D2, a montré une amélioration de la prévision à 1 h d'échéance sur la température et l'humidité à 2 m (Sgoff *et al.*, 2022). Le type de schéma d'assimilation, ainsi que son réglage, tient un rôle important pour prendre en compte la représentativité des observations de surface, surtout pour des systèmes denses d'observations.

Du fait de l'hypothèse de non-corrélation des erreurs d'observation, un écrémage des observations est réalisé pour des systèmes d'observation comportant des corrélations d'erreurs. L'écrémage peut aussi servir à réduire le poids des observations localement, pour assurer un certain niveau de cohérence entre l'échantillonnage des systèmes d'observation et la résolution du modèle, pour éviter de perdre les équilibres de l'ébauche. Plusieurs types d'écrémage sont possibles. Par exemple, Sgoff *et al.* (2022) sélectionnent 10 % des observations de PWS de manière aléatoire. La recherche d'un seuil d'écrémage optimal permettrait d'augmenter le nombre d'observations assimilé, sans que cela vienne diminuer les scores de longs termes. Pour les observations RADAR, des expériences OSE¹⁹ ont permis de tester différents seuils d'écrémage et un seuil d'écrémage optimal à 15 km a été trouvé (Wattrelot, 2016). Augmenter le nombre d'observations localement peut entraîner une dégradation des scores. Par exemple, l'écrémage des données d'avions Mode-S était réglé *par avion* dans le système 4DVar d'IFS, ce qui n'empêchait pas l'assimilation d'un grand nombre d'observations localement proche des aéroports, et a entraîné une dégradation des prévisions à 1 h du vent. Cet écrémage a dû être revu, en un écrémage par maille permettant de corriger la précédente dégradation (Ingleby, 2024). Ainsi, sans l'écrémage des observations, il est possible de réaliser une analyse qui tire profit des observations en s'en rapprochant fortement, et qui pourtant dégrade les prévisions. L'écrémage des observations est une composante clef pour l'assimilation de systèmes denses d'observation.

19. *Observing System Experiments*

2.3.4 Quels défis pour l'assimilation de données à échelle convective ?

L'assimilation de méso-échelle, ou assimilation à échelle convective (*convective-scale DA system*), se réfère à l'assimilation de données dans les systèmes régionaux de PNT avec une résolution horizontale de 1 à 4 km, où la convection est modélisée explicitement plutôt que paramétrée (Johnson *et al.*, 2017; Huang *et al.*, 2021; Hu *et al.*, 2022).

Malgré une augmentation de la résolution et de l'amélioration des modèles régionaux de PNT, une partie des événements de méso-échelle restent peu prévisibles. On l'a vu des événements de méso-échelle, comme les systèmes convectifs, dépendent à la fois de la grande échelle et de l'échelle aérologique. Par conséquent, même avec une analyse étroitement contrainte en observations SWS, la localisation exacte et l'heure de déclenchement d'un événement convectif reste peu prévisible et peut être sujets à la variabilité de plus petite échelle (Madaus et Hakim, 2016). En effet, si un phénomène de méso-échelle est observé, mais que le modèle ne l'a pas simulé, il serait nécessaire que le schéma d'assimilation de données permettent de générer à l'analyse tous les ingrédients pour simuler le phénomène, simplement en assimilant les observations, ce qui est peu probable avec un seul cycle d'assimilation. Néanmoins, en réalisant des cycles d'assimilation-prévisions, cela laisse le temps au modèle de pouvoir former cycle après cycle les ingrédients pour simuler le phénomène à l'analyse.

Ceci étant dit, le défi reste important. Car, quand bien même le modèle est capable de simuler le phénomène, si celui-ci admet un décalage spatial ou temporel, il se peut que les observations soient rejetées par le *screening*. En effet, les observations deviennent alors trop éloignées de l'ébauche et leur assimilation peuvent entraîner une double peine : d'abord en neutralisant le phénomène mal placé, et en plus sans permettre sa relocalisation. Enfin même si l'ébauche du modèle et les observations positionnent correctement le phénomène, il est nécessaire que la matrice de covariance d'erreur d'ébauche permettent de propager correctement l'information des observations, ce qui nécessite pour un schéma EnVar que la prévision d'ensemble ait correctement prévu le phénomène. La saisie d'informations de fine échelle, par exemple d'humidité ou de pression dans le cas d'un événement convectif, nécessiterait à la fois un réseau dense de stations, un réglage optimal du schéma d'assimilation et des cycles d'assimilation fréquents (Madaus et Hakim, 2016).

Concernant le réglage du schéma d'assimilation, l'évolution de schéma 3DVar vers des schémas prenant en compte des covariances d'erreur d'ébauche du jour, permet de mieux initialiser la convection profonde (Johnson *et al.*, 2017). D'autre part, l'ajout de nouvelles observations spatialement denses à l'échelle convective constitue un autre axe de recherche (Gustafsson *et al.*, 2018). Par exemple, l'ajout des observations d'éclairs, permettant de contraindre les contenus en neige et en graupel, éléments essentiels au développement des orages, pourraient permettre de générer et de déplacer des zones de précipitations (Combarous, 2023). Néanmoins, ajouter des ingrédients permettant la convection n'est pas suffisant pour maintenir de tels systèmes. Ceux-ci peuvent, par exemple, manquer d'énergie au sol à cause d'un manque d'air chaud et humide, induisant une dissipation généralement rapide des systèmes convectifs. Dans ces conditions, l'assimilation de plus d'observations de surface trouve toute sa place (Hu *et al.*, 2022).

2.4 Synthèse

L'assimilation de données cherche à tirer profit de l'information fournie conjointement par l'ébauche et les systèmes d'observation pour réaliser l'analyse, afin d'initialiser les prévisions. Différents schémas d'assimilation de données sont mis en place par les centres de prévisions météorologiques mondiaux pour initialiser les modèles à aire limitée (Gustafsson *et al.*, 2018). Le modèle AROME possède un système d'assimilation 3DVar avec cinq variables de contrôle dont la température, l'humidité spécifique et la pression de surface hydrostatique. Des opérateurs d'observations permettent de définir l'équivalent modèle des observations, pour pouvoir calculer l'écart entre l'ébauche et les observations.

Les observations de surface doivent faire leur place parmi l'ensemble des données assimilées, afin de contraindre les champs du modèle à l'analyse. Les observations de température et d'humidité relative de surface sont les observations qui ont le plus grand impact pour contraindre la couche limite atmosphérique. L'assimilation d'observations, et notamment les observations de surface à méso-échelle, pose malgré tout des défis liés à leur représentativité dans le modèle. Des phénomènes locaux adviennent à la surface, et la propagation de l'information contenue dans une observation ne devrait pas s'étendre au-delà du phénomène. L'utilisation d'une matrice d'erreur d'ébauche actualisée à l'heure de l'analyse, incluse dans le schéma 3DEnVar, permet de propager l'information des observations en fonction des conditions météorologiques simulées.

Des défis supplémentaires s'ajoutent pour l'assimilation des observations au sol provenant de stations météorologiques personnelles, du fait de la quantité et de la qualité de ces observations. L'apport de ces nouvelles observations pour améliorer l'analyse et in fine les prévisions reste à montrer.

Ces défis seront relevés différemment selon le type d'observations assimilées, que ce soit pour des observations de pression (chapitre 3), ou des observations de température et d'humidité relative (chapitre 4).

Chapitre 3

Assimilation d'observations de pression de surface provenant des stations personnelles dans AROME-France

Sommaire

3.1	Résumé de l'article	62
3.2	Article : <i>Assimilation of surface pressure observations from personal weather stations in AROME-France</i>	63
3.2.1	Introduction	64
3.2.2	Data, tools, and methodology	67
3.2.3	Pre-processing algorithm	69
3.2.4	Results of 1 month assimilation experiments	71
3.2.5	Results on the 14 September 2021 heavy precipitation event	74
3.2.6	Discussion and conclusion	81
3.3	Compléments de l'article	85
3.3.1	Supervision a posteriori des observations SWS assimilées	85
3.3.2	Gain spatial des observations en pression des PWS	85
3.3.3	Vérifications du système d'assimilation	87
3.3.4	Tests de sensibilité	88
3.4	Phénomènes à fort gradient de pression	91
3.4.1	Front - 5 octobre 2021	92
3.4.2	Dépression secondaire - 3 novembre 2023	93
3.4.3	Mascaret atmosphérique - 18 juin 2022	95
3.5	Synthèse et perspectives	98

3.1 Résumé de l'article

Des expériences d'assimilation des observations pré-traitées de pression de surface des PWS sont réalisés dans AROME-France. Le bénéfice de l'assimilation de ces nouvelles observations, avec les schémas d'assimilation de données 3DVar et 3DEnVar, est évalué sur une période d'un mois et lors d'un événement fortement précipitant dans le sud de la France. On notera expérience 3DVarP (resp. 3DEnVarP) l'expérience assimilant ces nouvelles observations de pression avec le schéma d'assimilation 3DVar (resp. 3DEnVar).

Les observations PWS de pression comportent une erreur systématique due à un mauvais étalonnage ou une mauvaise connaissance de leur altitude (figure 3.1). Une méthode de correction de biais « offline » est appliquée aux PWS en soustrayant pour chaque station l'écart moyen à l'ébauche des observations sur le dernier mois. Cette correction permet de recalibrer les observations de pression à une altitude conforme au modèle. Les observations ayant un écart-type d'écart à l'ébauche dépassant $60 \text{ m}^2 \text{ s}^{-2}$ sur le dernier mois sont ensuite rejetées, permettant de rejeter 1.4 % des stations supplémentaires considérées anormales. Enfin, les observations sont écrémées en choisissant une observation par maille d'1.3 km, soit à la résolution d'AROME-France. Sur la France, près de la moitié des 55 187 observations de pression des PWS disponibles par heure sont assimilées, soit 129 fois plus que le nombre d'observations de pression des SWS actuellement assimilées.

L'assimilation des observations de pression des PWS, avec le schéma d'assimilation 3DEnVar, montre une amélioration quotidienne systématique de la prévision de géopotential au niveau de pression de la station à 1 h d'échéance et réduit de 10.3 % la racine carrée de l'écart quadratique moyen de la pression de surface entre la prévision à 1 h du modèle et les observations SWS sur la France. Une amélioration significative de la pression moyenne au niveau de la mer est observée au cours des neuf premières heures de prévision. L'assimilation de ces mêmes observations avec le schéma d'assimilation 3DVar montre des résultats mitigés. En venant centrer les observations PWS sur l'ébauche du modèle, avec des écarts moyens des PWS à l'analyse (OmA) proches de zéro (inférieure à $0.2 \text{ m}^2 \text{ s}^{-2}$), le schéma d'assimilation 3DVar répand un biais (par rapport aux observations SWS) de l'ordre de $12 \text{ m}^2 \text{ s}^{-2}$. Ce biais augmente ainsi en moyenne l'écart à l'analyse du géopotential des SWS dans l'expérience 3DVarP, et a fortiori éloigne la prévision des observations SWS de référence. Cet effet est moins remarqué avec le schéma d'assimilation 3DEnVar, car la taille des incréments et le poids relatif des observations de surface y sont plus faibles.

Le diagnostic de Desroziers, permettant l'estimation de l'erreur d'observation, montre ses limites quand l'hypothèse d'observations non biaisées n'est pas respectée, ce qui est ici le cas pour les observations des SWS. Quand les deux systèmes d'observation de pression de SWS et PWS ont été débiaisés par rapport à l'ébauche du modèle, leur erreur d'observation estimée est proche et nettement inférieure à l'erreur d'observation prescrite. Ce diagnostic valide la qualité des données provenant des PWS et justifie leur assimilation avec la même erreur d'observation que celle prescrite pour les observations de pression des SWS.

Enfin, l'assimilation des observations de PWS de pression lors de l'épisode fortement précipitant du 14 septembre 2021 dans le Gard est présentée. Une anomalie de pression de surface générée par un système convectif de méso-échelle (MCS) est observée par les PWS et non visible avec le réseau SWS uniquement. L'assimilation des observations de pression des PWS avec le schéma d'assimilation 3DEnVar permet de retranscrire cette anomalie dans le champ analysé. Dans ce cas d'étude, l'assimilation des nouvelles observations

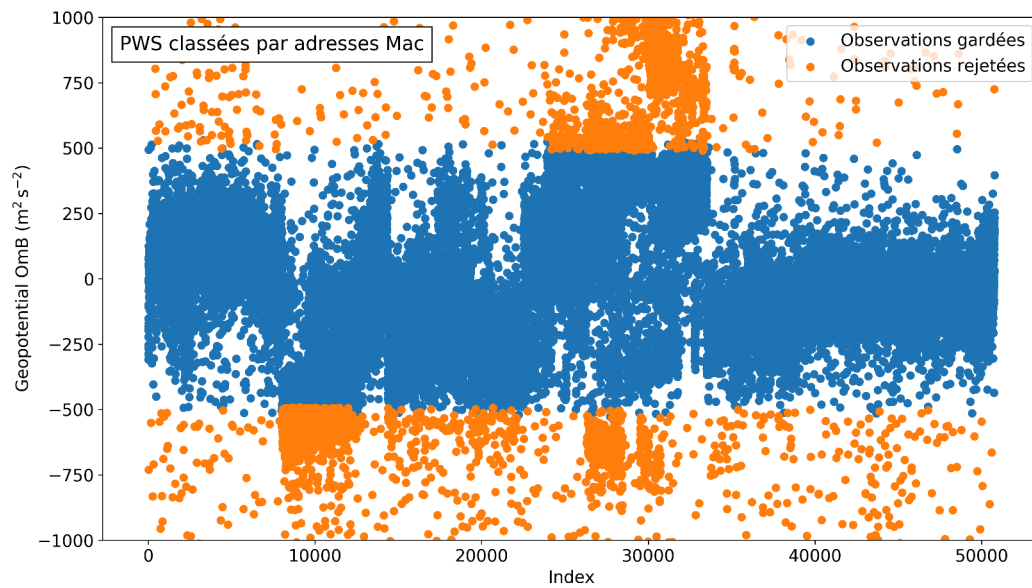


FIGURE 3.1 – Différences de géopotential au niveau de pression de la station entre observations brutes des PWS et prévision à 1 h d’échéance en fonction de leur adresse MAC, le 6 septembre 2021 à 01:00 UTC. L’adresse MAC est un identifiant unique donné à chaque station. L’étape de *screening* rejette les observations considérées anormales (en orange, chapitre 2). Les observations brutes comportent un biais dépendant de leur adresse MAC, dû à une mauvaise connaissance de leur altitude. Il n’y a pas de lien entre l’adresse MAC d’une observation et sa position spatiale.

rend les prévisions de la localisation et de l’évolution temporelle du MCS ainsi que des précipitations associées plus proches des observations de réflectivités radar.

3.2 Article : *Assimilation of surface pressure observations from personal weather stations in AROME-France*



Assimilation of surface pressure observations from personal weather stations in AROME-France

Alan Demortier¹, Marc Mandement¹, Vivien Pourret¹, and Olivier Caumont^{1,2}

¹CNRM, Université de Toulouse, Météo-France, CNRS, Toulouse, France

²Météo-France, Direction des opérations pour la prévision, Toulouse, France

Correspondence: Alan Demortier (alan.demortier@meteo.fr)

Received: 22 June 2023 – Discussion started: 8 August 2023

Revised: 5 February 2024 – Accepted: 12 February 2024 – Published: 19 March 2024

Abstract. Spatially dense surface pressure observations from personal weather stations (PWSs) are able to describe pressure patterns at the surface, such as those associated with convective events, in more detail than with standard weather stations (SWSs) only. In this study, the benefit of assimilating PWS observations with the 3DVar and the 3DEnVar data assimilation schemes of the AROME-France model is evaluated over a 1-month period and during a heavy precipitation event in the South of France. Observations of surface pressure from PWSs are bias-corrected, quality-controlled, and thinned with a spacing equal to the horizontal dimension of an AROME-France grid cell. Over France, almost half of the 55 187 available PWS observations are assimilated, which is 129 times more than the number of assimilated SWS observations. Despite the limited advantages found from their assimilation with the 3DVar assimilation scheme, the 3DEnVar assimilation scheme shows systematic improvement and reduces by 10.3 % the root-mean-square deviation in surface pressure between 1 h model forecasts and SWS observations over France. Significant improvement is observed over the first 9 h of the forecasts in mean sea level pressure. Finally, when PWS observations are assimilated with the 3DEnVar assimilation scheme, a surface pressure anomaly generated by a mesoscale convective system – observed by PWSs and not visible without them – is successfully assimilated. In that case, the forecasts of location and temporal evolution of the mesoscale convective system as well as rainfall are closer to the observations when PWS observations are assimilated.

1 Introduction

Amongst all the challenges to improve data assimilation (DA) for convection-permitting numerical weather prediction (NWP), the use of new observation types has emerged (Hu et al., 2022). Indeed, accurately estimating the initial state (analysis) of a convective-scale model requires the assimilation of spatially dense and temporally frequent observations (Gustafsson et al., 2018). A step forward in improving analyses could come from near-surface crowdsourced observations, e.g. from smartphones or personal weather stations (PWSs) (Hintz et al., 2019). Crowdsourced observations are much more numerous than observations of standard weather stations (SWSs) currently assimilated in NWP models while having a more heterogeneous quality and less metadata (Nipen et al., 2020; Sgoff et al., 2022).

Clark et al. (2018) and Mandement and Caumont (2020, hereafter MC20) showed that PWS observations, when combined with SWS observations, reveal small-scale spatio-temporal variations in screen-level temperature, relative humidity, and surface pressure associated with the life cycles of thunderstorms which are partly visible, or not visible at all, with SWS observations only. Amongst these three observed variables, MC20 showed that use of PWS allows approximately 100 times more surface pressure observations to be assimilated. These observations, when used to produce gridded analyses of mean sea level pressure (MSLP), are the observations that most reduce (of the three observed variables, during 4 stormy days) the root-mean-square error of these analyses compared to analyses using only SWSs. The shape and amplitude of surface pressure perturbations, which vary from one thunderstorm to another, are described by Madaus and Hakim (2016) using idealized simulations for the partic-

ular case of isolated thunderstorms. The interest in assimilating PWS observations into a convection-permitting NWP model has therefore emerged.

As well as varying on a fine scale with the passage of thunderstorms, surface pressure is a crucial variable in NWP models. Surface pressure observations provide information not only near the surface but also over the entire air mass in a column of the atmosphere, and they are less dependent on the surface characteristics (e.g. cities, vegetation; Ingleby, 2015) than other surface observations. Most global and limited-area models assimilate surface pressure observations over land and oceans from the World Meteorological Organization (WMO) Regional/Global Basic Observing Networks (World Meteorological Organization, 2021) exchanged through the WMO Global Telecommunication System (World Meteorological Organization, 2020). In the Met Office global model, Ingleby (2015) showed that surface data contribute to 16.9% of the total observation impact of 24 h forecast (compared to 10.0% for radiosondes and 9.2% for aircraft). In the Météo-France global model ARPEGE, Chambon et al. (2023) showed that surface observations are crucial to reduce forecast errors: over Europe, the Performance Index 18 (IP18) forecast score (based on the arithmetic average of the geopotential, temperature, and wind root-mean-square errors against radiosoundings at 48 and 72 h range and normalized by the value of the year 2008) diminished from 6.51 to -1.63 when the surface observations are removed over a 6-month period (Chambon et al., 2023), i.e. being worse than 2008 score. Surface pressure errors are identified, amongst the other surface observations, as the main source of 24 h forecast errors in terms of a moist energy norm.

These last years, several studies (Madaus et al., 2014; McNicholas and Mass, 2018b; Hintz et al., 2019; Ridal et al., 2019) have reported assimilation experiments of crowdsourced surface pressure observations from dense networks to capture local-scale features. Various techniques were designed depending on the kind of observations (e.g. smartphones or PWSs) and the assimilation system.

Among these studies, Madaus et al. (2014) assimilated between 1300 and 1700 pressure observations from PWSs over the Pacific Northwest in addition to the 140 SWSs commonly assimilated in NWP models. One-month assimilation experiments were done in the WRF model using an ensemble Kalman filter DA system. With appropriate bias correction and quality control, they showed significant improvements of 3 h forecasts of surface pressure, with lower but still significant improvements of 2 m temperature and a 10 m wind meridian component. Furthermore, results showed a better timing of frontal passages (20 to 45 min of improvement) in the forecasts. Then, a few experiments were done using smartphone observations. Hintz et al. (2019) showed the difficulties of the task due to quality control and bias correction issues because the altitudes of the smartphone pressure observations were not accurately known (accentuated by

the moving location of the sensors) and used a 3DVar assimilation system. However, McNicholas and Mass (2018a) developed a machine learning quality control method for smartphone pressure observations. Once applied, quality-controlled observations were assimilated in an ensemble DA system during two specific cases (McNicholas and Mass, 2018b). Results showed an improvement of the 1 h forecast on both 2 m temperature and humidity as well as on the track and intensity of a coastal wind storm. In the same way for PWSs, Ridal et al. (2019) implemented and compared different quality control methods, using in particular a variational bias correction to remove the pressure bias from the observation time series. The results of the 3-week assimilation experiments in HARMONIE-AROME, the limited-area NWP model developed within the ALADIN consortium (Bengtsson et al., 2017), which uses a 3DVar DA scheme, did not show any advantage. In summary, the assimilation of dense surface pressure observations from a few pairs of observation network-DA scheme has been tested, but there is a lack of information on the role that the DA scheme plays in making the most of dense PWS surface pressure observations.

Different DA schemes have been developed for AROME-France. AROME-France is the convection-permitting, limited-area model operational at Météo-France since 2008. AROME-France has a three-dimensional variational (3DVar) DA scheme launched hourly and assimilates, in addition to observations assimilated in ARPEGE, dedicated observations for the mesoscale within a 1 h assimilation window (Seity et al., 2011; Brousseau et al., 2016). The background error covariance matrix determines how the information from the observations is spread in the model space (horizontally and vertically) to make the analysis. In the 3DVar scheme, the background error covariance matrix is climatological, spatially homogeneous, and isotropic. Montmerle et al. (2018) and Michel and Brousseau (2021) developed an ensemble DA scheme, to replace the static matrix with a dynamic background error covariance matrix, improving the spread of the information from the observation. This scheme is called three-dimensional ensemble variational (3DEnVar).

This study intends to answer the following questions: does the assimilation of PWS surface pressure observations in the current AROME-France 3DVar data assimilation system improve weather analyses and forecasts? And what are the differences with the 3DEnVar system currently under development?

The remainder of this paper is organized as follows. Section 2 describes the crowdsourced surface pressure observations used and the different assimilation configurations available for AROME-France. The observation pre-processing and the DA experiments are described in Sect. 3. Section 4 presents the scores of 1-month cycled assimilation experiments, and a case study of a heavy precipitation event (HPE) is analysed (Sect. 5). Finally, the advantages of the 3DEnVar

system for the assimilation of dense surface pressure observations are discussed (Sect. 6).

2 Data, tools, and methodology

A study period of 1 month is chosen, ranging from 6 September 2021 at 00:00 UTC to 5 October 2021 at 23:00 UTC. In total, it represents 717 h for each experiment (excluding two time steps during which PWS observations are missing). This period includes several rainy situations, such as precipitation events in the south-west of France on 8 September, in the south-east of France on 14 September, and over the north and south-east of France between 4 and 5 October.

2.1 Surface pressure observations

Pressure observations are mainly retrieved at the surface with weather stations. Two main surface networks are used: the standard weather stations and the personal weather stations.

2.1.1 Standard weather station observations

Standard weather stations include land surface synoptic stations (sending SYNOP weather reports through the WMO Global Telecommunication System; see World Meteorological Organization, 2020), the French extensive data acquisition and weather observation network (RADOME, Tardieu and Leroy, 2003), and stations located over sea on ships (sending SHIP reports). On average during the 1-month study period, 1124 surface pressure observations per hour are available in the AROME-France domain coming from manual land SYNOP (10.0%), automatic land SYNOP (58.6%), French RADOME (17.9%), manual SHIP (2.0%), and automatic SHIP (11.4%) reports. As France is a fraction of the AROME-France domain, only a part of these stations are located in France. Indeed, among these reports, 214 are located in France over land: 10.1% are automatic land SYNOP and 89.9% are French RADOME. Stations associated with these surface pressure observations are further called SWSs. Most of the scores will focus on SWSs located in France.

2.1.2 Personal weather stations observations

Amongst all types of personal weather station (PWS) networks, the Netatmo network stands out as being at the same time the largest in France and available in near-real time (MC20). The stations are privately owned, which explains why they are not spatially equally distributed (Fig. 1); i.e. they are denser in the cities.

The Netatmo PWS is composed of outdoor and indoor modules. The indoor module measures the surface pressure. Observation times are not exactly regular: the mean time step between two observations is approximately 5 min, and the company updates its servers only every 10 min. The owner of the station can perform on-demand measurements, which

shift the following observation times. Also, the evolution of the surface pressure related to the passage of thunderstorms is of the order of a few minutes. For all these reasons, to limit time lags, the observations are linearly interpolated at fixed times (i.e. round hours) using the two closest measurements around the interpolation time. If the two closest measurements are separated by a period of 700 s or more, the observation is removed.

The absolute accuracy of the pressure sensor of the Netatmo PWS given by the manufacturer is ± 1 hPa (Netatmo, 2023). In 2018, the sensor was an STMicroelectronics LPS25HB, which has the same absolute accuracy indicated by Netatmo and a relative accuracy of ± 0.2 hPa after calibration for standard atmospheric surface conditions (STMicroelectronics, 2016). The exact model of the sensor may change regularly; however the absolute accuracy provided by Netatmo has not changed since 2019 (MC20), suggesting that their characteristics remain similar. Once the PWS is installed and running, pressure sensor calibration is possible but is not performed by default, which can lead to systematic errors. One of the main sources of surface pressure bias is errors in the station altitude, caused in part by uncorrected GPS heights (Pauley and Ingleby, 2022). When PWS MSLP observations were compared to SWS analyses taken as a reference, MC20 showed that the first (Q25) and the third (Q75) quartiles of the systematic errors were equal to -4 and 2 hPa, respectively. The supplementary errors result from lagged-time pressure measurements to pressure change and inaccurate or unavailable metadata. These constraints will be taken into consideration in the pre-processing method.

2.2 AROME-France assimilation system

AROME-France is the Météo-France operational limited-area model (Seity et al., 2011; Brousseau et al., 2016). The model is convection-permitting at 1.3 km horizontal grid spacing with 90 vertical hybrid pressure levels. It does not rely on the hydrostatic assumption and fully resolves elastic Euler equations. Physical parameterizations come from the Meso-NH research model (Lac et al., 2018). At the surface, AROME-France is coupled with the SURFEX model (Masson et al., 2013). The model is regularly updated with new features, creating new development cycles. In the current study, the cycle 48t1 of AROME-France is used. Hereafter, AROME will systematically refer to AROME-France.

The DA scheme combines the active observations (see below for the meaning of active) with the background (for AROME the 1 h forecast starting from the previous analysis), to produce x_a , the analysis state of the atmosphere, obtained by the minimization of the cost function J :

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) + \frac{1}{2}[\mathbf{y}_o - \mathcal{H}(\mathbf{x})]^T \mathbf{R}^{-1}[\mathbf{y}_o - \mathcal{H}(\mathbf{x})], \quad (1)$$

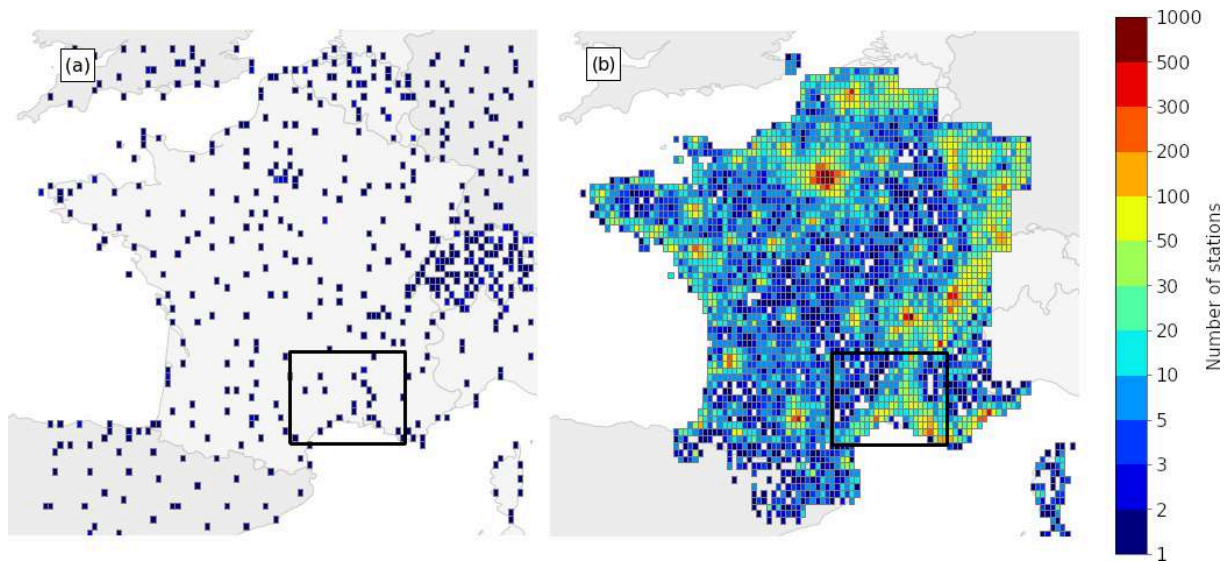


Figure 1. Number of (a) SWSs and (b) PWSs providing at least one observation during the 1-month study period over France. Observation counts are binned into approximately $0.15^\circ \times 0.1^\circ$ bins. The black rectangle delimits the domain used for the case studied in Sect. 5.

in which \mathbf{x}_b is the background state, \mathbf{y}_o the observation vector, \mathcal{H} the nonlinear observation operator, and \mathbf{B} and \mathbf{R} are the background and the observation error covariance matrix, respectively. \mathbf{B} approximates the error of the model at the observation point, propagates the information from the observation in both horizontal and vertical directions, and influences other variables. To reduce the computational cost, the formulation used to find the solution from J is incremental, i.e. using δx the difference between the atmosphere state and the one from the background (Courtier et al., 1998), at full model resolution. If \mathbf{B} is computed at a lower resolution than the model, interpolation operators from low to high resolution are applied.

The difference between an observation and its corresponding value in the background, called observation minus background (OmB), provides a comparison between the observations and the model 1 h forecast. The OmB and the observation minus analysis (OmA) can be calculated using observation operators. However, the OmB is rather preferred in score calculation, due to correlation that analysis has with observations.

AROME uses a 3DVar data assimilation scheme, with an hourly data assimilation cycle, to keep as close as possible to the true state of the atmosphere (Brousseau et al., 2016; Gustafsson et al., 2018). The 3DVar analysis done at time t_0 considers all observations between $t_0 - 30$ min and $t_0 + 30$ min to be valid at t_0 . To avoid redundancy issues, only the observation closest to t_0 is taken into account for stationary stations (SWS except ships) providing multiple observations during the assimilation window. Only PWS observations valid at t_0 are used, following the method described in Sect. 2.1.2. All available observations are submitted to a quality control procedure called screening, which removes

the observations deemed questionable. One of the screening steps compares the observation vector \mathbf{y}_o to the background state \mathbf{x}_b using the observation operator \mathcal{H} and keeps observations such as $\|\mathbf{y}_o - \mathcal{H}(\mathbf{x}_b)\| < 5\sigma$ with $\sigma = \sqrt{\sigma_o^2 + \sigma_b^2 \sigma_{\text{coef}}^2}$, given σ_o and σ_b , the standard deviation of observation and background errors, respectively, and $\sigma_{\text{coef}} = 0.9^{-2}$ a supplementary factor which artificially increases σ , allowing one to keep more observations. Two other screening steps that are important for mass observations (surface pressure and geopotential height) include checks for missing station altitude and compliance with WMO observation standards: if the reported observation level is far from the true station level, σ_o is inflated (more details in Sect. 4.4.1 of ECMWF, 2023). Height differences between stations and the ground level of the model, which can be large in mountainous areas, are not managed by the screening but rather by the observation operator described in Sect. 2.2.1. Observations rejected by the screening are not used further.

After this step, the observations kept are called active observations. This procedure matters, removing suspicious observations, which helps the convergence of the minimization algorithm and limits the spin-up period. However, some observations might be rejected when they should not because of an inaccurate background. Additional quality control is tested in Sect. 3 to reject observations depending on the background departure of their time series rather than the background departure at one specific time.

The 3DVar and 3DEnVar share the same formulation, with the only difference that \mathbf{B} is static in the 3DVar and becomes \mathbf{B}_e , an ensemble error covariance matrix, in the 3DEnVar (Michel and Brousseau, 2021). The matrix \mathbf{B}_e is estimated for each analysis from an ensemble data assimilation (EDA)

system composed of 50 members at 3.2 km horizontal grid spacing and with 90 vertical levels (Brousseau et al., 2016). Despite the fact that the coefficients of the \mathbf{B}_e matrix are given at 1.3 km of resolution, the structure of the increments is still at the spatial resolution of the coarser grid. The advantage of the matrix \mathbf{B}_e is to have a flow-dependent background error statistics updated at each analysis, which allows the model to take better advantage of the observation. However, in order to remove the sampling noise of the model error covariance estimate, a spatial localization is applied (Hamill et al., 2001), which limits the propagation of the OmB to shorter distances. The localized matrix $\hat{\mathbf{B}}_e$ is obtained making an element-to-element product (known as a Hadamard–Schur product) between \mathbf{B}_e and a correlation matrix \mathbf{L} of the same dimension (Eq. 2).

$$\hat{\mathbf{B}}_e = \mathbf{B}_e \circ \mathbf{L}. \quad (2)$$

The correlation functions in the \mathbf{L} matrix are quasi-Gaussian with horizontal and vertical localizations (Montmerle et al., 2018). The specified length of localization corresponds to the half width at half maximum (HWHM) of the quasi-Gaussian function. The tuning of the horizontal localization has been found to be optimal using a linear function from 25 km at the ground to 150 km at the highest model level (Michel and Brousseau, 2021). The vertical localization is set to about 0.15 hPa (Michel and Brousseau, 2021).

2.2.1 Assimilating geopotential at the level of the observed station pressure

In AROME, the geopotential is equal to gz , with $g = 9.80665 \text{ m s}^{-2}$ the standard acceleration of gravity and z the altitude. For each observation of surface pressure (p_s) at the altitude of the station's barometer (z_s), the geopotential equal to gz_s is assimilated at the vertical level corresponding to p_s . For geopotential observations, σ_o is set to $8g$. The model geopotential at the location of the observation is retrieved using an observation operator described by Courtier et al. (1998).

Surface pressure and geopotential variables are both used in this study to refer to the same observations. The surface pressure will be used when studying the observations, whereas the corresponding geopotential at the level of the observed station pressure will be the variable used when looking at assimilation statistics.

The propagation of the OmB from a surface pressure observation differs between the 3DVar and the 3DEnVar DA systems. If we consider that a typical horizontal scale in the surface pressure analysis is given by the distance at which the amplitude of a surface pressure analysis increment divided by its maximum value is below 1 %, then this scale is approximately 200 km for the 3DVar scheme and 40 km for the 3DEnVar scheme (Fig. A1). In 3DEnVar this typical horizontal scale can slightly vary depending on the meteorological situation.

2.3 Evaluation tools

When assimilating new observations, the associated \mathbf{R} matrix has to be specified. The allocated part of the \mathbf{R} matrix for PWSs will have the same characteristics as the one for the SWSs; i.e. it is assumed to be diagonal, and the diagonal values are the same. This assumption requires verifying that observation errors are not correlated and, if they are, using techniques such as thinning to reduce the negative impact of this correlation. The effect of different σ_o between PWS and SWS observations could be investigated in a sensitivity study.

Desroziers et al. (2005) proposed a posterior diagnostic to estimate the observation error covariance matrix \mathbf{R} , which can be estimated by $E(\mathbf{OmA}[\mathbf{OmB}]^T)$. This method has some limitations. It assumes that the background and the observation error covariance matrices are accurately specified and uncorrelated with a sufficiently large and representative sample of forecast error statistics (Pourret et al., 2022). However, the \mathbf{B} matrix is not totally estimated or stored but specified through operators in the 3DVar scheme (Seity et al., 2011; Brousseau et al., 2014). Some limitations will be discussed further.

The estimation of the correlation between observations errors matters. Liu and Rabier (2003) showed that increasing the number of observations with uncorrelated errors generally increases the quality of the analysis and the forecast. However, if the observations have correlated errors greater than a threshold (e.g. 0.2 as used by Liu and Rabier, 2003), it decreases the quality of the analysis and the forecast. To prevent observation errors from being spatially correlated, a thinning technique which removes observations at an optimal spatial resolution can be applied (Benáček et al., 2016). Thinning is done for observing systems providing high spatial resolution observations which have correlated observation errors, such as radar (Wattrelot et al., 2012), satellites (Dando et al., 2007), or aircraft (Pourret et al., 2021).

This method is not used for SWS, but given the high spatial density of the PWS (Fig. 1b), it is tested further. To find this optimal spatial resolution, the Desroziers posterior diagnostic can be extended to a method estimating the spatial correlation of observation error (Mile et al., 2019). First, for each distance bin, from 0 to 30 km at 1 km intervals, it finds equidistant pairs of stations. Then for each pair (i, j) of stations in each bin, the covariate $\text{OmA}_i \cdot \text{OmB}_j$ is calculated. Finally, the average of the covariates is normalized by the average of the first covariate, when the distance is zero, which corresponds to the actual Desroziers diagnostic. This diagnostic gives an estimation of the spatial correlation errors, which can be used to estimate a minimum thinning length where the correlation errors are significantly low. It has the same limitation as the Desroziers diagnostic previously described.

2.4 Design of the experiments

In order to evaluate the benefits of the new observations, the observing system experiment (OSE) framework is used to compare a reference experiment with an experiment where new observations are assimilated (e.g. Pourret et al., 2022). OSEs allow us to compare the background (i.e. 1 h forecast) or the analysis of both experiments to observations. In AROME the examination can be both at the surface with observations from stations and at different levels of the atmosphere with observations from radiosondes, aircraft, and satellites.

Section 3 describes an experiment monitoring the raw pressure observations from the PWSs in a non-cycled 3DVar DA scheme. This monitoring experiment will be used to compute both the bias correction and the quality control of the PWS observations. Then, the Sect. 4 will focus on long cycled experiments with the 3DVar and the 3DEnVar assimilation schemes. Finally, Sect. 5 will use long, cycled experiments and an extra experiment assimilating the new observations only at 07:00 UTC on 14 September 2021. All the experiments are described in Table 1. The experiments without assimilating PWS observations will also be called control experiments.

3 Pre-processing algorithm

All PWS observations are subject to a pre-processing algorithm before assimilation, composed of a bias correction, quality control, and spatial and temporal thinning. The number of observations remaining after each step of the process is provided in Table 2. Starting from raw observations, they are interpolated to round hours, using the method described in Sect. 2.1.2, which removes some observations. Stations with identical coordinates are assumed to have not been properly set up by the owner, leading to automatic placement based on their wireless IP address (Meier et al., 2017). These stations are further removed.

3.1 Bias correction and quality control

Figure 2a shows 1-month time series of surface pressure observations from 30 PWSs close to the city of Toulouse (i.e. less than 0.05° from the city centre), located in the South of France. Most of the time series exhibit similar pressure variations and large differences in the mean value. However, the area being relatively flat (i.e. less than 150 m from lowest to highest point in this area), it is not realistic to have a systematic difference up to 25 hPa (approx. 250 m), i.e. a first glance of the different systematic errors at each station location. Then, it has to be confirmed with long-term statistics, which can be done with the monitoring experiment. In this experiment, which assimilated the raw pressure observations from the PWSs, the observations are compared to the model background (i.e. the 1 h forecast from the operational

AROME model), giving the OmB for the whole set of observations. The 30 PWS time series have systematic biases from -500 to $1200 \text{ m}^2 \text{ s}^{-2}$ in terms of geopotential (Fig. 2b). More globally, Fig. 3 shows the whole set of geopotential OmB during the pre-processing. The raw data exhibit the biases from $\pm 1000 \text{ m}^2 \text{ s}^{-2}$ (looking at the whiskers, representing the percentile range of the data; i.e. from 5th to 95th percentiles). Therefore, a systematic bias correction is crucial.

To implement a geopotential bias correction for PWS data based on AROME, the previously described monitoring experiment with a non-cycled 3DVar DA scheme has been used. PWS individual biases are then obtained by averaging the geopotential OmB for each station over a rolling time window of 1 month to prevent the correction from depending on the meteorological situation and for robust statistics. This bias is then subtracted from the actual observation which, at the end, only changes the altitude of the station in the assimilation scheme. PWSs with missing observations for more than half the rolling time window are eliminated to ensure the robustness of the bias. This bias correction reduces the bias range to $\pm 15 \text{ m}^2 \text{ s}^{-2}$ and removes 4.6 % more stations while keeping numerous ones with heterogeneous quality (Fig. 3b). Contrary to SWSs, some PWSs can suddenly move, and their spatial coordinates or altitude metadata change. Some other stations also show sudden pressure change without an abnormal change in their metadata (Fig. 2a). That is why a supplementary quality control has been tested. First, PWSs that have moved from more than 100 m on the horizontal and 1 m on the vertical are discarded. That represents 0.4 % more stations removed. Also, an interruption in the PWS observation time series is suspected to be an early sign of erroneous observations: PWSs with missing observations for more than a quarter of the last 48 h are eliminated. Finally, PWSs whose geopotential observations deviate too much from AROME background, i.e. with a standard deviation of their last month's geopotential OmB above $60 \text{ m}^2 \text{ s}^{-2}$, are removed. This empirical threshold removes the tail of the OmB distribution (i.e. 1.4 % of the observations), corresponding to unrealistic observations while keeping physical patterns. In the end, 84.8 % of the stations are kept, and the range of the geopotential OmB is under the SWS ones (Figs. 3c, d).

3.2 Thinning

After the bias correction and the quality control, the number of PWS observations within a 1.3 km horizontal grid cell of AROME ranges from 0 to 69, 69 being reached in the centre of Paris. In order to keep a minimal weight to the model's background and other observations at such grid points where several PWS geopotential observations are available, it is chosen to thin PWS observations at 1.3 km, the model's horizontal grid spacing, for the whole studied period, independently of the assimilation scheme. Among PWS observations available in a grid point, one is randomly drawn. Such thinning, not usually applied for independent surface observa-

Table 1. Overview of the experiments: 3DVar, 3DVarP, 3DEnVar, and 3DEnVarP experiments were launched on 6 September 2021 at 00:00 UTC. The monitoring experiment was launched on 6 August 2021 at 00:00 UTC (1 month earlier). The 3DEnVarP7UTC experiment was launched on 14 September 2021 at 07:00 UTC. BC means bias correction, and QC means quality control.

Experiments	Duration	Cycling	Data assimilation scheme	Use of PWS surface pressure observations
Monitoring	2 months	No	3DVar	Raw observations monitored
3DVar 3DVarP	1 month	Yes	3DVar	No
3DEnVar 3DEnVarP			3DEnVar	Assimilated after BC, QC, and thinning
3DEnVarP7UTC	1 h	No	3DEnVar	Assimilated after BC, QC, and thinning

Table 2. Average number of PWS surface pressure observations following the pre-processing process between 6 September 2021 at 00:00 UTC and 5 October 2021 at 23:00 UTC (717 time steps). Each step of the process is described in the text. SD refers to standard deviation of the OmB's last rolling month (see text).

	Raw	Preparation		BC		QC	
		interpolation	same lat–long	> 1/2 of the period	station has moved	> 1/4 during the last 48 h	SD > 60 m ² s ⁻²
PWSs remaining	55 187	52 334	50 757	48 275	47 748	47 564	46 800
% of total PWSs	100 %	94.8 %	92.0 %	87.4 %	87.0 %	86.2 %	84.8 %

tions, is generally applied in DA to minimize the risk of having correlated observation errors. For surface pressure, correlated errors could come from instrument systematic bias or correlation of the model altitude error at the subgrid scale. To verify that PWS geopotential observation errors after thinning are not correlated, a spatial Desroziers diagnostic (Sect. 2.3) is computed for two analyses. The analyses have been carried out at midnight and at midafternoon.

Figure 4 presents this diagnostic for pre-processed pressure observations from PWSs. From the first bin (1 km) there is limited spatial correlation of observation error. This diagnostic shows small variability between the two assimilation schemes and the hour of study. It confirms that, if a thinning radius is chosen, there is no need to increase the thinning radius above 1.3 km to mitigate the issue of correlated errors. The effect of assimilation without thinning or shorter thinning lengths is not tested. After thinning, there is an average of 27 889 PWS surface pressure observations per hour assimilated, which represents 51 % of the initial number of PWS observations.

4 Results of 1-month assimilation experiments

When statistical values (STAT) such as standard deviation or root mean square (rms) of time series are compared between experiments, their relative evolution in an experiment (XP)

with respect to another (CTRL) is given by

$$\Delta\text{STAT} = \frac{\text{STAT}_{\text{XP}} - \text{STAT}_{\text{CTRL}}}{\text{STAT}_{\text{CTRL}}}. \quad (3)$$

Special attention will be paid when the CTRL statistical value is close or equal to 0, in which case the relative evolution is not defined. The significance of the values is determined using a bootstrap method, where the rms is computed from the time series sampled with replacement 1000 times. The 1000 rms values form a distribution, from which 95 % confidence intervals are estimated using the “percentile” method (Virtanen et al., 2020).

The OmBs from the four experiments described in Table 1 are compared in this section. When pressure observations from PWSs were added to the AROME assimilation cycles, no convergence problems were reported.

4.1 Impact on geopotential analyses and 1 h forecasts

Statistics of geopotential OmA and OmB for SWSs and PWSs gathered in Table 3 reveal the ability of the assimilation schemes to take advantage of the PWS geopotential observations.

In the 3DVar experiment, the OmA geopotential mean for SWSs is equal to 4.7 m² s⁻², which shows a little bias between AROME analyses and the SWS geopotential observations over France. Over the whole domain, it decreases to 2.0 m² s⁻² (not shown). In the 3DEnVar experiment, the OmA geopotential mean for SWS is higher than in the

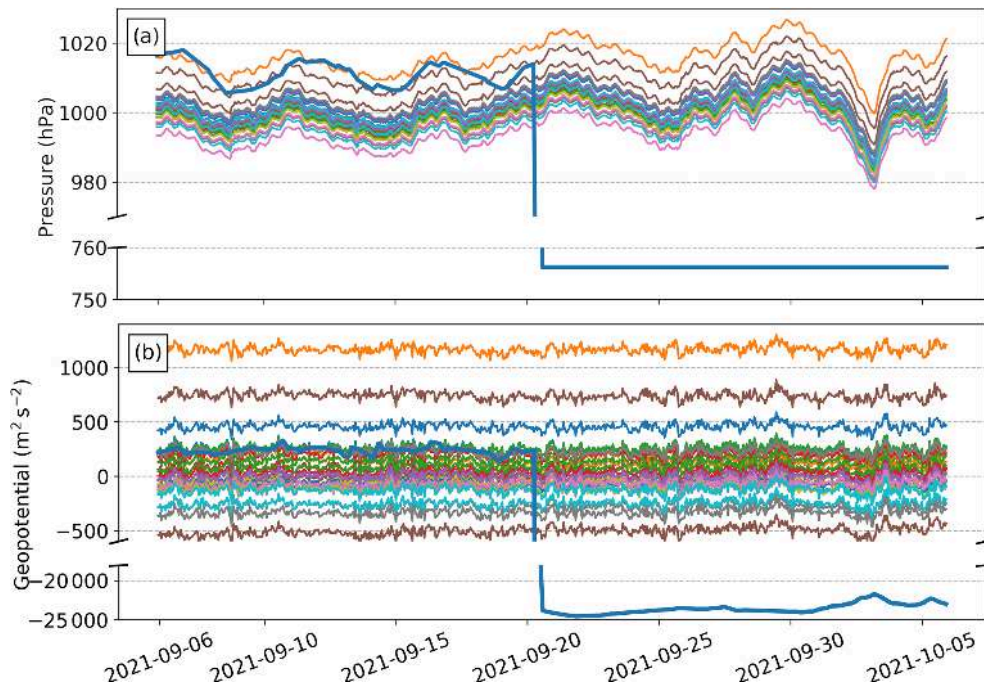


Figure 2. Time series of (a) surface pressure observations and (b) corresponding geopotential OmBs from 30 PWSs around Toulouse and 1 anomalous PWS during the study period from 6 September to 5 October 2021. The background is from the monitoring experiment.

Table 3. Mean and standard deviation of geopotential OmA and OmB ($\text{m}^2 \text{s}^{-2}$) and Desroziers diagnostic ($\tilde{\sigma}_O$ in $\text{m}^2 \text{s}^{-2}$) for SWSs located in France and PWSs over the 1-month study period, which represent respectively 212 and, 27 889 observations per hour on average over the 717 h of the experiments. NA – not available.

Monitoring		3DVar		3DVarP				3DEnVar		3DEnVarP				
PWS		SWS		SWS		PWS		SWS		SWS		PWS		
mean	SD	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD	
OmA	NA	4.7	26.5	11.9	25.4	−0.1	14.6	12.8	38.0	12.2	24.8	0.2	15.9	
OmB	−2.6	36.3	11.4	43.3	16.7	41.7	3.9	48.9	15.6	46.4	15.3	41.3	2.6	49.2
$\tilde{\sigma}_O$	n.a.	31.2		27.7		15.7		43.6		29.3		19.7		

3DVar experiment and equal to $12.8 \text{ m}^2 \text{ s}^{-2}$. The geopotential standard deviation OmA increases by 43 % from 26.5 to $38.0 \text{ m}^2 \text{ s}^{-2}$ between 3DVar and 3DEnVar. This can be due to the finer localization given to geopotential SWS observations in 3DEnVar than in 3DVar. In 3DVarP and 3DEnVarP, experiments in which PWS geopotential observations are assimilated, the OmA geopotential means for PWSs are closer to zero (respectively $−0.1$ and $0.2 \text{ m}^2 \text{ s}^{-2}$). This result was expected given the prior PWS bias correction using AROME’s OmB statistics over a rolling month.

Between 3DVar and 3DVarP, the OmB geopotential mean for SWSs increases, while the geopotential standard deviation OmB for SWSs decreases by 4 %. Time series of geopotential mean and standard deviation OmB shows the same trends every day (Fig. 5a).

Between 3DEnVar and 3DEnVarP, both the OmB geopotential mean and standard deviation for SWSs decrease (i.e. $−12$ % for the standard deviation). Time series of geopotential standard deviation OmB show the same trends every day (Fig. 5b), while 3DEnVarP is not systematically closer to 0 than 3DEnVar for the geopotential mean OmB. Time series also show that improvements in OmB statistics (mean and standard deviation) of 3DEnVarP compared to 3DEnVar are larger on days when 3DEnVar OmB statistics are furthest from zero, which corresponds to days of unsettled weather, such as on 8 and 14 September and 5 October 2021.

In the framework of these experiments, considering SWS observations as the reference, the assimilation of PWS observations in a 3DVar scheme has a mixed impact: it increases OmA and OmB geopotential mean but decreases OmA and OmB geopotential standard deviation. The large increase in

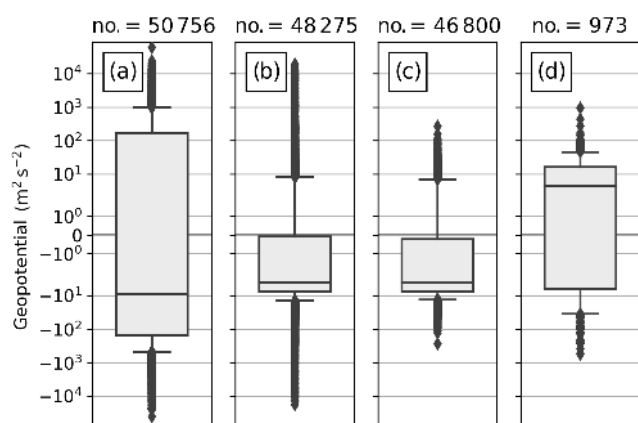


Figure 3. Box plots of geopotential OmB during the 1-month study period. Geopotential observations originate from (a) PWS raw data, (b) PWS bias-corrected data, (c) PWS assimilated (bias-corrected and quality-controlled), and (d) SWS assimilated. The background fields are coming from the non-cycled monitoring experiment. Box-plot titles indicate the hourly average number of stations. The bar across the box represents the median, and its borders represent the 25th and 75th percentiles. The box-plot whiskers indicate the 5th and 95th percentiles of the data. Outliers are plotted as separate dots.

OmA geopotential mean when PWS observations are assimilated could be explained by the PWS bias correction method, which uses OmB statistics and thus contains model or other observation biases (as SWS ones). As 129 times more PWS observations than SWS observations are assimilated, with the same prescribed observation error, the 3DVar long range PWS increments overwhelm the SWS ones, which questions the relative number of observations in each subset and its impact in different analysis methods. On the contrary, the assimilation of PWS observations in a 3DEnVar scheme systematically improves all statistics of both analyses and 1 h forecasts.

The Desroziers diagnostic ($\widetilde{\sigma}_0$), described in Sect. 2.3, calculated from OmA and OmB statistics, encompasses these results. For all experiments, values of this diagnostic estimating the observation error range from 15.7 to 43.6 m² s⁻² (Table 3), which is only between 20 and 56 % of the prescribed value in AROME (78.4532 m² s⁻²). For SWSs, the Desroziers diagnostic increases by 40 % from 31.2 to 43.6 m² s⁻² between 3DVar and 3DEnVar experiments, which shows a high sensitivity to the choice of the assimilation scheme. Also, the Desroziers diagnostic varies for SWSs when PWSs are assimilated: it decreases by 12 % between 3DVar and 3DVarP and by 48 % between 3DEnVar and 3DEnVarP. Finally, the PWS bias correction explains why the Desroziers diagnostic is lower for PWSs than for SWSs in 3DVarP and 3DEnVarP experiments, which could be wrongly interpreted as PWS geopotential observations are of better quality than SWS geopotential observations (see Appendix B). These findings show some limitations of this

diagnostic, which should be used with caution, especially when its assumptions such as that the average OmB and OmA are close to zero are not fulfilled.

4.2 Impact on 1 h forecasts of surface pressure, temperature, relative humidity, and wind

To show how the background (1 h forecast) deviates from SWS observations, the rms OmB of a given variable observed near the ground by SWSs is calculated. This score is presented in Table 4.

The comparison of 3DEnVar with respect to 3DVar shows an 8.0 % increase in rms OmB of surface pressure for SWSs over the AROME domain and 9.3 % over France. It indicates that the 3DEnVar background deviates more than the 3DVar background from SWS observations of surface pressure. For 2 m temperature and 10 m zonal and meridional wind, increases in rms OmB range from 1.1 % to 2.0 %. Only the rms OmB of 2 m relative humidity decreases by 1.1 % for SWSs over France but still increases by 3.6 % over the AROME domain.

The comparison of 3DVarP with respect to 3DVar shows slightly positive Δ rms OmB of surface pressure for SWSs over France (0.7 %) and the AROME domain (0.8 %), showing a slight degradation, but it is not statistically significant. Despite this, significant decreases of -1.6 % and -1.5 % of rms OmB of the zonal and meridional wind for SWSs located in France are found.

The comparison of 3DEnVarP with respect to 3DEnVar shows significantly negative Δ rms OmB of surface pressure for SWSs, reaching -10.3 % over France and -3.0 % over the AROME domain. Small but significant degradation from 0.8 % to 1.3 % is found on 2 m temperature and 2 m relative humidity.

These experiments show that PWS assimilation improves significantly the quality of the surface pressure background when a 3DEnVar scheme is used, and it partly compensates for the significant degradation associated with the transition to a 3DEnVar system from the 3DVar. Even if they are statistically significant, only a little improvement of the background is found in the AROME 3DVar scheme in terms of wind when PWS geopotential observations are assimilated, with a neutral impact on surface pressure.

4.3 Impact on the forecasts up to 30 h range

For 3DVar, 3DVarP, 3DEnVar, and 3DEnVarP experiments, forecasts starting from the analyses at 00:00, 06:00, 12:00, and 18:00 UTC were run each day during the 1-month study period. Statistics of observation minus forecast (OmF) and their evolution between experiments (as indicated in Eq. 3) are computed. By definition, OmF at 0 h time step is equal to OmA, but OmF at 1 h time step is not equal to OmB because of the use of the incremental analysis update technique in AROME forecasts (Brousseau et al., 2016).

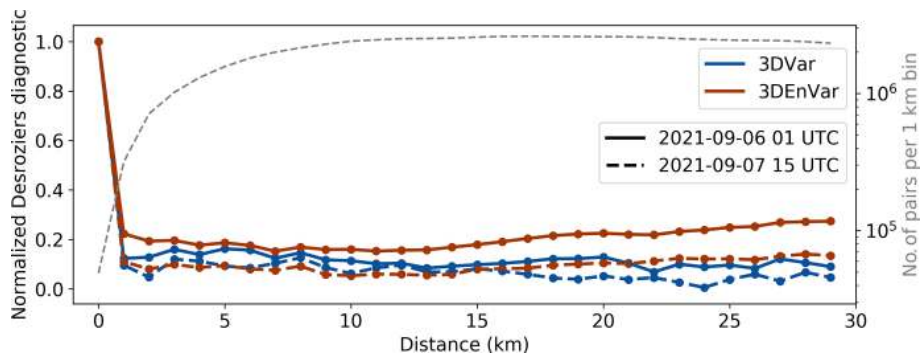


Figure 4. Spatial Desroziers diagnostic of PWS surface pressure observations. The colours represent the assimilation scheme, with the 3DVar, in blue, and 3DEnVar, in brown. The diagnostic is done at (plain lines) 01:00 UTC on 5 September 2021 and (dashed lines) at 15:00 UTC on 6 September 2021. In grey, the number of pairs of equidistant stations (in 1 km bins).

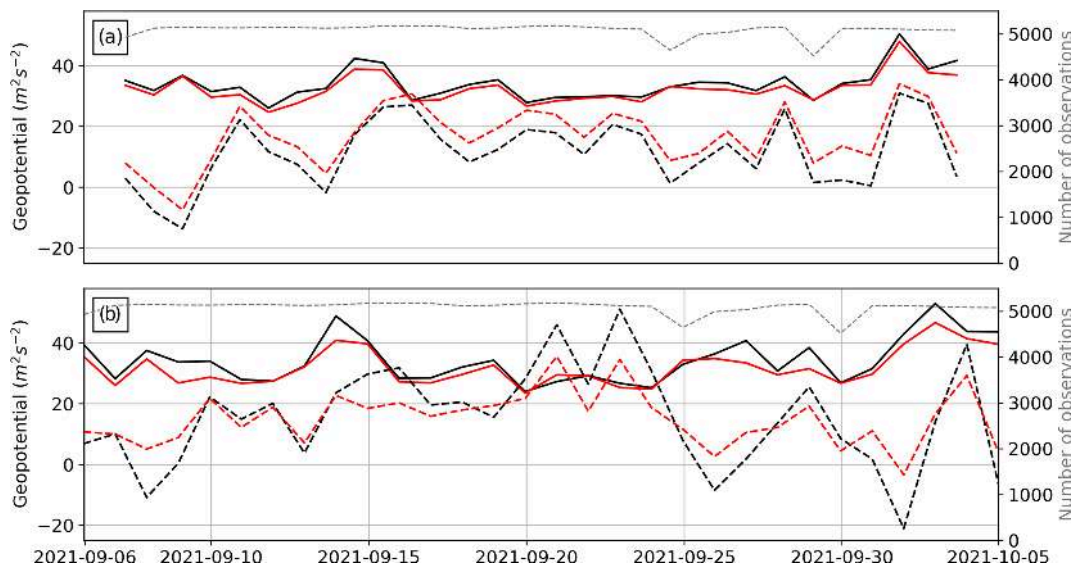


Figure 5. Daily average geopotential (dashed line) mean and (plain line) standard deviation OmB for (a, black) 3DVar, (a, red) 3DVarP, (b, black) 3DEnVar, and (b, red) 3DEnVarP experiments over the 1-month study period. Only observations available for all experiments are taken. The dashed grey line is the daily number of observations.

In the analysis (0h), $\Delta_{rms} OmF$ of MSLP for 3DVarP with respect to 3DVar is negative and equal to -7.5% , and then it rapidly increases during the first hours of the forecast and is positive after only 3 h of forecast (Fig. 6a). For 3DEnVarP with respect to 3DEnVar, $\Delta_{rms} OmF$ of MSLP is negative and equal to -24% in the analysis, also rapidly increases during the first hours of the forecast, but remains negative before 12 h of forecast. This shows that the benefit of the assimilation of PWS observations quickly disappears in 3DVarP, while it remains in the 3DEnVarP experiment. Small but significant degradation is noticeable between 15 to 21 h of forecast for 3DVarP with respect to 3DVar, which is less pronounced and not significant for 3DEnVarP with respect to 3DEnVar. The influence of the assimilation of PWS observations on other variables near the ground is less pronounced, and no significant results are seen (not shown).

Impact at different levels of the atmosphere can be assessed with the help of radiosoundings (Fig. 6b). Because PWS observations are only assimilated over France or near its borders, only five French radiosoundings launched twice a day are taken into consideration. Over the 1-month study period, it represents between 220 (at 1000 hPa) and 272 observations. Negative $\Delta_{rms} OmA$ of geopotential near the ground and for the whole troposphere is seen, reaching at 1000 hPa -11% for 3DVarP with respect to 3DVar and -24% for 3DEnVarP with respect to 3DEnVar. This shows a substantial improvement of geopotential in the analysis throughout the troposphere when PWS observations are assimilated. Regarding other variables, improvement is seen on the wind heterogeneously for the whole troposphere for 3DVarP with respect to the 3DVar experiment (not shown),

Table 4. Δ rms OmB of surface pressure, 2 m temperature, 2 m relative humidity, and 10 m zonal and meridional wind for SWSs over the 1-month study period. Negative values of an experiment (XP) with respect to another (CTRL) indicate that 1 h forecasts of XP are closer than CTRL to SWS observations. Significant values are in bold (see text for details).

	Surface pressure		Temperature		Relative humidity		Zonal wind (U)		Meridional wind (V)	
	AROME domain	France	AROME domain	France	AROME domain	France	AROME domain	France	AROME domain	France
3DEnVar w.r.t. 3DVar	8.0 %	9.3 %	2.2 %	1.2 %	3.6 %	−1.1 %	2.0 %	1.1 %	2.0 %	1.3 %
3DVarP w.r.t. 3DVar	0.8 %	0.7 %	0.0 %	−0.8 %	0.0 %	0.0 %	−1.4 %	−1.6 %	−1.4 %	−1.5 %
3DEnVarP w.r.t. 3DEnVar	−3.0 %	−10.3 %	0.6 %	0.8 %	0.8 %	1.3 %	0.2 %	0.0 %	−0.2 %	−0.2 %

and no significant change could be seen for other radiosounding parameters.

AROME-France being the finer-scale operational model available over France, one of its added values lies in the forecasting of local precipitation events. One common verification score called AROME index (also called IP16) described in Amodei et al. (2015) synthesizes regional Brier skill scores (BSS_NO) of wind gusts and 6 h accumulated rainfall up to 24 h of forecast. It is computed for the different experiments and is shown in Fig. 7. IP16 increases from 0.797 to 0.827 between 3DVar and 3DEnVar, which is a 3.8 % improvement associated with the change of the DA scheme. The assimilation of PWS geopotential observations decreases IP16 by $−0.3 %$ in 3DVarP with respect to 3DVar and increases by 0.7 % in 3DEnVarP with respect to 3DEnVar.

When looking in detail at BSS_NO composing the IP16, the largest improvement of BSS in 3DEnVarP with respect to 3DEnVar is found for the highest threshold of accumulated rainfall: BSS_NO increases by 0.7 % for the 0.5 mm threshold against 1.3 % for the 5 mm threshold. Improvements in BSS_NO are found particularly from 6 to 18 h of the forecast and vanish after 24 h (not shown). Results concerning the 14 September event are shown in the next section.

5 Results on the 14 September 2021 heavy precipitation event

5.1 Presentation of the case

The Mediterranean region is regularly affected by heavy precipitation events (HPEs, Doswell et al., 1998; Nuissier et al., 2020; Caumont et al., 2021). On 14 September 2021, a mesoscale convective system (MCS) remained quasi-stationary for 8 h, giving up to 260 mm of rain in 3 h between 07:15 and 10:15 UTC in the city of Saint-Dionisy, in the Gard department (South of France). The particular organization of mesoscale vertical circulations, clouds, hydrometers, and precipitation inside MCSs can lead to the formation of surface pressure structures which can be observed from the ground.

When PWS surface pressure observations – quality-controlled by MC20 algorithm – and SWS surface pressure observations are combined, surface pressure structures appear near and below the MCS (Fig. 8). At 07:00 UTC a low-pressure anomaly in front of the precipitation system is observed, which could be associated with updraughts, increased relative humidity, increased temperature, and at the rear high-pressure anomaly, potentially due to downdraughts. These surface pressure anomalies could not be seen with SWS observations only; therefore, this section examines the impact of assimilating PWS observations of pressure in the area where these anomalies are observed.

5.2 Impact of PWS observations assimilation on surface pressure increments

Figure 9 shows increments of surface pressure for experiments assimilating PWS observations of geopotential (3DVarP, 3DEnVarP) or not (3DVar, 3DEnVar). Because the four experiments are cycled and starting from much earlier, they have a different background at 07:00 UTC. Since the four experiments start from a different background, this comparison is intended to illustrate the spatial structure of surface pressure increments and, for each observation, the associated change in analysis departure with respect to background departure. Other observations are assimilated in this area at 07:00 UTC, coming from SWSs, from radars and from an aircraft flying towards Marseille Provence airport (not shown).

Increments in experiments using a 3DVar DA scheme (Fig. 9a, b) are smoother than increments in experiments using a 3DEnVar DA scheme (Fig. 9c, d). These differences are due to the differences in background error covariance matrices between both schemes. In the 3DVar experiment (Fig. 9a), all pressure observations from SWSs are consistent with the increment signs, which indicate that with their large analysis increment, they are more predominant than other assimilated observations. This is not the case for the 3DEnVar experiment (Fig. 9c), where anomalies appear close to the centre of the MCS. They are not consistent with the two closest SWS surface pressure observations. To know from which observations these anomalies were coming, a supplementary experiment has been done rejecting the radar data, which is

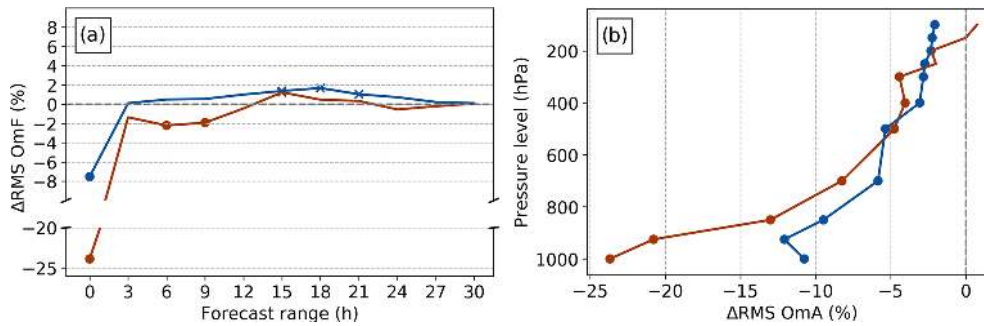


Figure 6. (a) Δ rms OmF of MSLP for SWSs over France between 0 and 30 h forecast range and (b) Δ rms OmA of geopotential for radiosoundings launched over France for (blue) 3DVarP with respect to 3DVar and (brown) 3DEnVarP with respect to 3DEnVar. Points indicate improvement, while crosses indicate degradation of XP with respect to CTRL at 90 % statistical significance level.

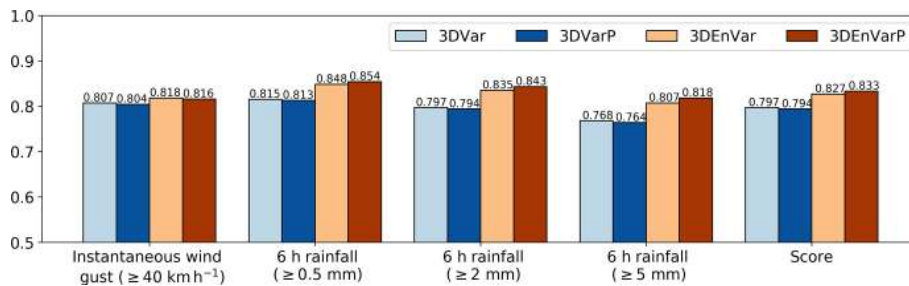


Figure 7. IP16 score of forecasts launched at 00:00 UTC during the 1-month study period except the first day (7 September to 5 October 2021) averaged over four forecast times (6, 12, 18, 24 h) given by horizontal bars for 3DVar (light blue), 3DVarP (navy blue), 3DEnVar (beige), and 3DEnVarP (brown) experiments. All columns correspond to regional Brier skill scores (with a spatial window of 50 km) contributing to IP16: instantaneous wind gusts above 40 km h^{-1} , 6 h accumulated rainfall above or equal to 0.5, 2, and 5 mm. Higher is better.

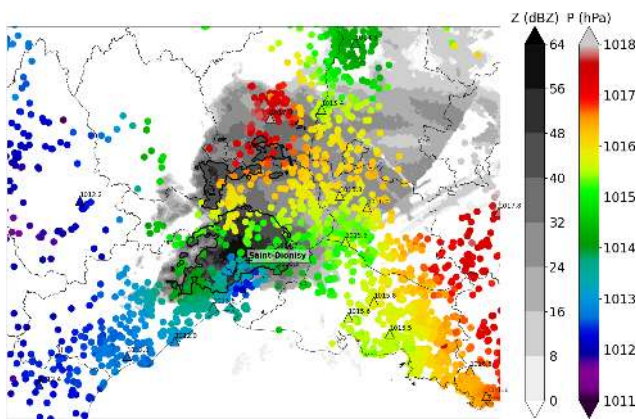


Figure 8. MSLP observations of SWS and PWS networks at 07:00 UTC on 14 September. SWSs are indicated by coloured triangles with black contours and PWSs by coloured circles. The instantaneous (500 or 250 ms) wind gust during the last minute is shown with barbs (in kn). Two values separated by a slash indicate respectively from left to right the instantaneous wind gust during the last minute and during the last 10 min in km h^{-1} . Base reflectivity (Z) in grey colours indicates thunderstorm activity and location. Reflectivities over 40 dBZ are illustrated by bold black edges.

the other type of observation assimilated close to the anomalies; this effectively removes the anomalies (not shown). This confirms that surface pressure observations have a lower impact than other assimilated observations (here the radars) in this specific case with the 3DEnVar scheme.

When assimilating PWS observations, the two assimilation schemes react differently. For the 3DVarP (Fig. 9b), all the positive increments from PWSs add up to form a large positive anomaly. A small group of observations, close to the MCS, are not consistent with this anomaly, which shows the difficulties that the 3DVar DA scheme has to take advantage of the local scale information. However, the 3DEnVarP experiment (Fig. 9d) effectively achieves pushing the model according to the OmBs brought by PWS observations. The number of PWS observations assimilated compensate for the small localization given to the surface observations in the 3DEnVar DA scheme. Only a few stations remain at the boundary between positive and negative anomalies, which do not correspond to the increment. As 3DEnVarP is the only experiment able to reproduce the local anomaly of surface pressure observed by PWSs in the 07:00 UTC analysis, the following section will compare experiments with a 3DEnVar DA scheme.

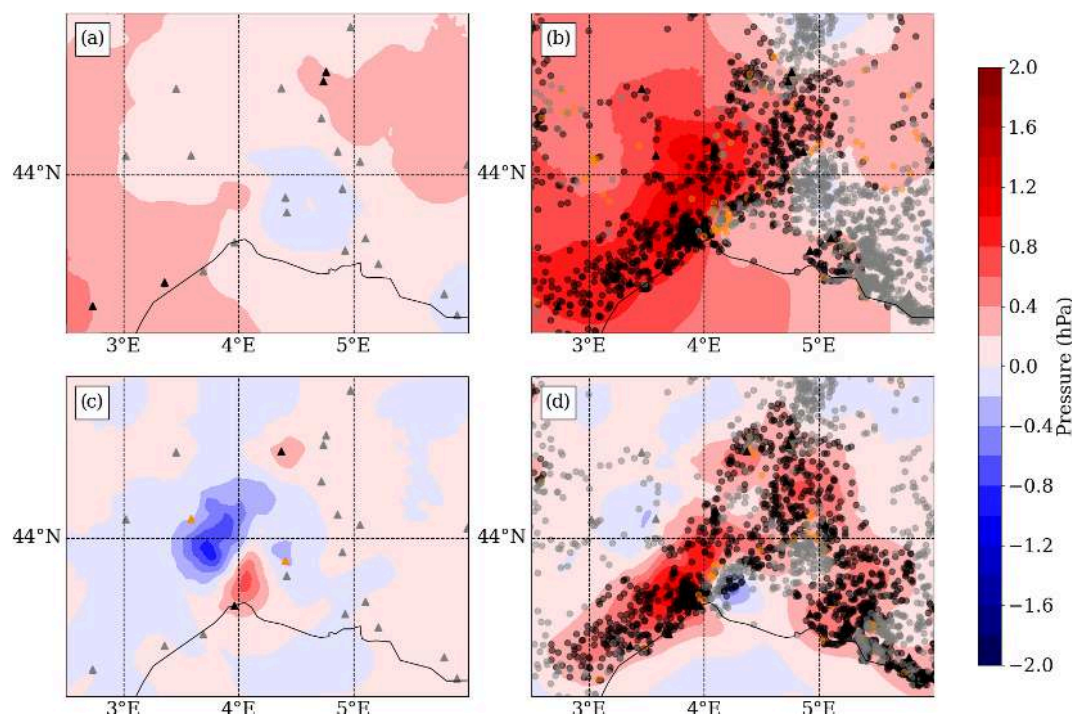


Figure 9. Surface pressure increments (analysis minus background) in (a) 3DVar, (b) 3DVarP, (c) 3DEnVar, and (d) 3DEnVarP experiments on 14 September 2021 at 07:00 UTC. Locations of the stations whose observations of surface pressure are assimilated are indicated by triangles for SWSs and circles for PWSs. Colours of triangles and circles indicate for each surface pressure observation the change in analysis departure (O_{mA}) with respect to background departure (O_{mB}): reduced in black such as $|O_{mA}| - |O_{mB}| < -0.2 \text{ hPa}$, increased in orange such as $|O_{mA}| - |O_{mB}| > 0.2 \text{ hPa}$, or neutral in grey such as $||O_{mA}| - |O_{mB}|| < 0.2 \text{ hPa}$.

5.3 Impact on forecasts from a single-hour PWS assimilation

Given the pressure increments added by the PWSs with the 3DEnVar scheme, the influence on other variables is explored. In that way, an experiment using the 3DEnVar DA scheme and assimilating PWS observations only at 07:00 UTC was realized (Fig. 10). It is called hereafter 3DEnVarP7UTC. In the AROME DA system, the surface pressure, temperature, specific humidity, and the two components of the wind are analysed every hour, while the other prognostic model fields, such as vertical divergence, are only updated during the forecast steps of the assimilation cycles. In consequence, with the help of the 3DEnVar background covariance error matrix, the PWS geopotential observations induce surface pressure, temperature, and relative humidity changes close to the surface. Indeed, above the positive surface pressure increments, between 3 and 4° E, a negative temperature difference implying a denser air mass appears (Fig. 10b, c). Vertical velocity, being a diagnostic field, is not updated in the analysis, and only small anomalies appear due to changes in surface pressure and thus pressure of all grid points on the vertical (Fig. 10d). Regarding the evolution of these anomalies, after a few time steps of forecast, the AROME model balances the information provided by the

surface pressure increments into other prognostic fields and uses them to create new meteorological structures. The surface pressure differences in Fig. 10e have a smaller scale than the analysis increments in Fig. 10a. Near 4° E, after 50 min of forecast, a dipole of positive and negative vertical velocity differences appears (Fig. 10h). At that time, multiple fine-scale differences, up to 1 °C in temperature and up to 20 % in relative humidity, show that the location of convective cells inside the MCS is modified. This particular event shows that changes in surface pressure have an effect on the other variables in the model, which can rapidly extend all the way to the top of the troposphere. As such storms occur rapidly and over a small geographical area compared with the area of the model domain, the modifications of temperature and relative humidity seen in this figure are barely visible on assimilation experiment scores over a time period of 1 month shown in Sect. 4.

5.4 Impact on predicted rainfall

The MCS remained quasi-stationary in the southern part of the Gard department, which caused heavy rainfall during the event. The rainfall hourly time series analysed by ANTILOPE, the Météo-France operational precipitation analysis combining radar and rain gauge observations at

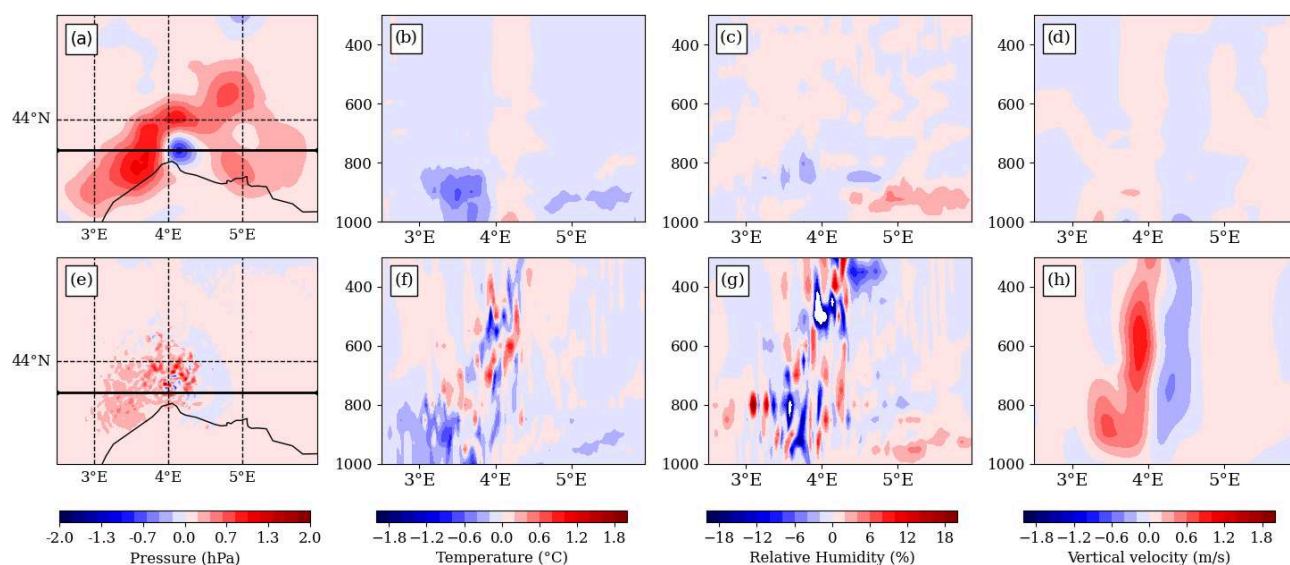


Figure 10. Differences between 3DVarP7UTC and 3DVar experiments of (a, e) surface pressure, (b, f) temperature, (c, g) relative humidity, and (d, h) vertical velocity (a–d) in the analysis of 07:00 UTC and (e–h) at 50 min forecast time. (b–d, f–h) Differences along the vertical cross-section indicated by the thick black line in (a) pressure coordinates (hPa).

1 km grid spacing (Laurantin, 2008), shows the chronology of the event (Fig. 11d). Rainfall started over the area at 00:00 UTC. Maximum hourly rainfall continuously increased until 09:00 UTC and decreased after 09:00 UTC. Between 08:00 and 09:00 UTC, the ANTILOPE analysis (Fig. 11c) shows two convective cells, the southern being the most intense with an hourly maximum rainfall accumulation of 144 mm. The corresponding 9 h forecasts starting from the 3DVar and 3DVarP analyses are shown in Fig. 11a and b. The 3DVar experiment forecasts rainfall over the entire Gard department and even north of it, whereas the 3DVarP experiment is closer to ANTILOPE for both the spatial extension and the location of the rainfall maximum in the south of the Gard department. Looking at the maximum rainfall accumulation, the 3DVarP experiment with 96 mm is closer to the observation (144 mm) than the 3DVar experiment, with 65 mm. It can be assumed that the vertical velocity anomalies previously observed could generate more intense precipitation rates. Fractions skill scores (FSSs; Mittermaier, 2018) show no advantage of 3DVarP with respect to 3DVar for small thresholds of precipitation (e.g. 1 mm). However, over 5 mm thresholds, the FSS is significantly better for the 3DVarP experiment (not shown). This means a better location and intensity of the hourly maximum rainfall accumulation of precipitation at this time. In addition, the time series of the hourly rainfall accumulation of the 3DVarP experiment is closer to ANTILOPE than the 3DVar experiment.

Finally, what forecasters are interested in is the anticipation of the event. Then, the added value of the 12 h forecast launched from the 3DVarP analysis at 00:00 UTC on 14 September 2021 with respect to the 12 h forecast launched

from the 3DVar analysis at the same time is explored. The maximum rainfall accumulation observed during this period is 320 mm. The maximum rainfall accumulation of the 12 h forecast launched from the 3DVarP analysis is 221 mm, which is 22 % higher than the 3DVar one (180 mm). Looking at the FSS, small but significant degradation is seen at the 2 mm threshold, whereas significant improvement can be seen at higher thresholds, especially at 50 and 100 mm (Fig. 12). In this case, the FSS of the 12 h forecast at 100 mm threshold is significantly improved at a window size of 1° in latitude and longitude. This could be explained by more realistic analyses a few hours before the forecast launch, thanks to the assimilation of PWS surface pressure observations.

6 Discussion and conclusion

This study has examined the impact of the assimilation of spatially dense PWS surface pressure observations with the 3DVar and 3DVarP data assimilation schemes of AROME-France in 1-month cycled experiments. A procedure has been presented in order to correct and select the observation dataset for the purpose of assimilation.

The usability of surface pressure observations is related to the accuracy of their altitude metadata. When assimilated in a non-cycled AROME 3DVar experiment (the monitoring experiment), abnormal systematic differences in the OmB statistics were shown in the pressure time series of 30 individual stations around Toulouse. The model topography is different from the true one, and the OmB statistics combine the error from the observation and the error from the model. The choice was made to correct the altitude of the

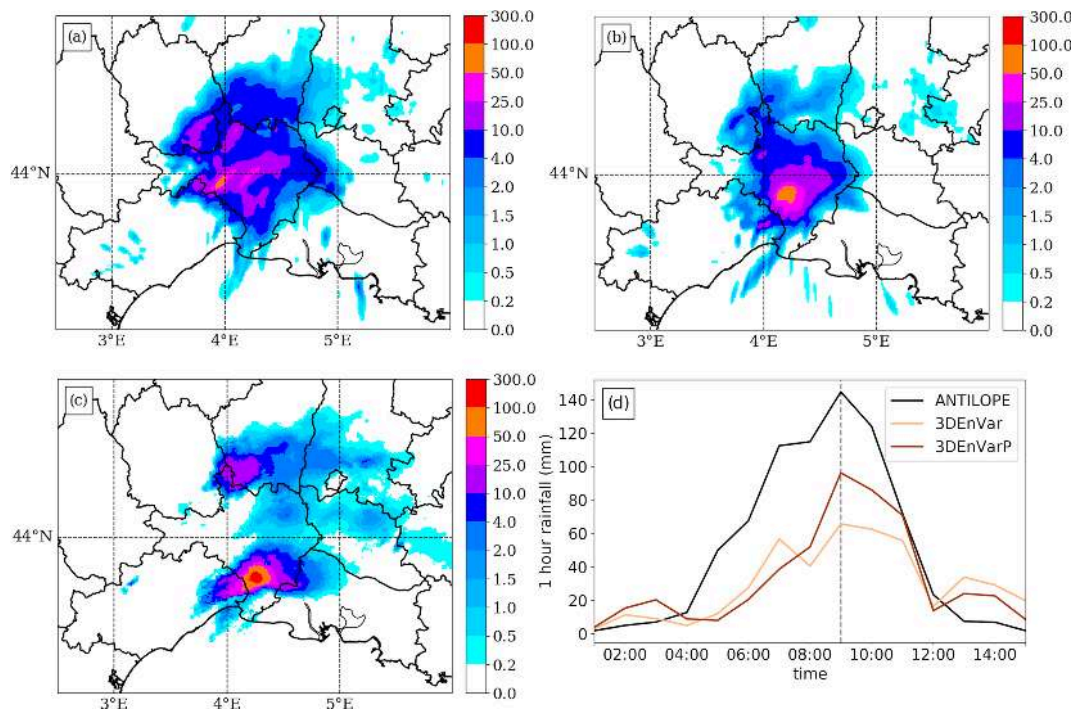


Figure 11. Rainfall accumulation (mm) on 14 September 2021 between 08:00 and 09:00 UTC of forecasts starting from the (a) 3DEnVar and (b) 3DEnVarP analyses at 08:00 UTC on 14 September 2021 and (c) analysed by ANTILOPE. (d) Time series of hourly maximum rainfall over the area of forecasts starting from the (light brown) 3DEnVar and (brown) 3DEnVarP analyses and (black) analysed by ANTILOPE.

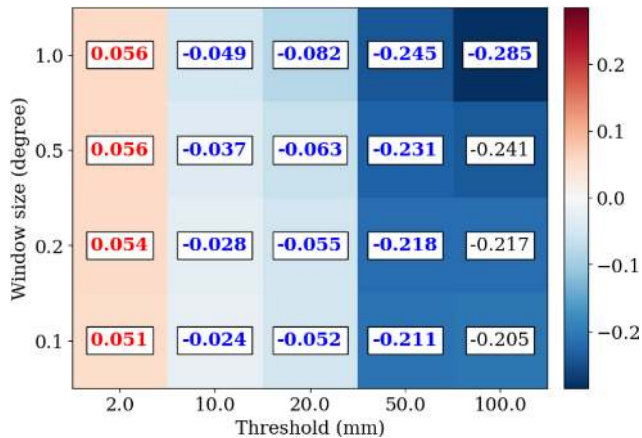


Figure 12. Difference in FSS of 12h precipitation forecasts starting from the 3DEnVar and 3DEnVarP analyses at 00:00 UTC on 14 September 2021. Difference is computed as $FSS(3DEnVarP) - FSS(3DEnVar)$. FSS is computed for 2, 10, 20, 50, and 100 mm thresholds and spatial neighbourhoods of 0.1, 0.2, 0.5, and 1° in latitude and longitude. Significant positive values are in red and negative in blue.

PWS stations to the altitude they would have in the model. The station-dependent bias has been found by averaging the OmBs from the monitoring experiment during one rolling month. The corrected observations have been used for both

3DVar and 3DEnVar experiments. The bias correction substantially reduces the range of individual geopotential bias from ± 1000 to $\pm 15 \text{ m}^2 \text{ s}^{-2}$.

Then, unrealistic pressure observations were filtered. This was done by removing observations when their time series were too sparse and when the standard deviation of the 1-month geopotential departure between the PWS and the model was larger than $60 \text{ m}^2 \text{ s}^{-2}$. This method ensures that observations belonging to time series with abnormal variations, due in particular to sudden changes in altitude or mis-calibration, are removed.

The last pre-processing step was a thinning. With the use of the spatial Desroziers diagnostic, the set of observations was found to have limited spatial correlation of observation error. PWS observations were thinned by keeping one observation every 1.3 km in order to avoid (1) giving too much influence to PWS observations when they are dense compared to conventional observing systems, which can make local minimization of the cost function difficult and ultimately deteriorate the forecast (Eq. 1), and (2) propagating possible correlated observation errors in space, even if their existence has not been proven. After the preparation of the data, the bias correction, the quality control, and the thinning, there are an average of 27 578 pressure observations per hour ready to be assimilated, which represents 51 % of the initial observations.

In order to improve the pre-processing, there are several possibilities. Firstly, with regard to the bias correction, which can induce a slight bias of $-2.6 \text{ m}^2 \text{ s}^{-2}$ on average. The bias correction uses the 1 h forecast of the monitoring experiment, which may be biased. Investigations over a longer period are needed to fully understand this bias. Another limitation of this method may lie in the consistency of assimilating into a 3D EnVar experiment PWS observations that have been bias-corrected thanks to OmB time series from the non-cycled 3DVar experiment. Other results may be expected using a bias correction which uses OmB time series from a non-cycled 3D EnVar experiment. In an operational context, a VarBC method could be a convenient and more consistent solution to routinely correct the bias of PWS observations while being aware of the need to maintain sufficient anchor observations from SWSs (Eyre, 2016).

Concerning the quality control (QC), the method used here sometimes suffers from not removing a wrong PWS surface pressure observation whose time series undergo a small but abnormal pressure variation (e.g. change of spatial position without any change in metadata). If such abnormal pressure variation persists over time, then the monthly rolling standard deviation of the OmBs will keep growing hour by hour, and it will eventually exceed the rejection threshold. The QC will take some time to remove the wrong stations. More sophisticated QC could be done to remove erroneous observations faster and more accurately, with the help of the neighbours' variations. Indeed, a simple buddy check (as used by Vasiljevic et al., 2006) could help while keeping in mind its difficulties to make the difference between erroneous variations and local physical patterns (as the one seen during the 14 September HPE). Finally, a more advanced technique, taking into account a temporal window to remove stations that are unrealistic during a few time steps (as the MC20 method) could be beneficial. In the end, the supplementary QC presented here removed 1.6 % more observations, which is the same order of magnitude as the leave-one-out cross-validation method presented by MC20. More investigation could be done on the utility of such method, moreover, if a superobbing method is tested in place of the thinning one. The necessity of the thinning procedure is not questioned in this study. A dense observational network can be detrimental in the case of correlated observations errors and cause convergence problems and imbalanced fields (Pourret et al., 2021; Brousseau et al., 2016). Therefore, we applied a thinning at 1.3 km scale, which reduced those effects. However, the spatial Desroziers diagnostics have shown small observation correlation errors for the first 1 km distance bin. Experiments without thinning could be done in the future to understand how the assimilation scheme reacts when assimilating pressure observations from PWS in dense observations areas, like in the cities.

Making 1-month cycled experiments has allowed us to see the overall impact of the assimilation of dense PWS surface pressure observations on analyses and forecasts of the

AROME NWP system. When looking at statistics of geopotential OmB for SWS observations, the rms OmB does not diminish when PWS observations are assimilated in the 3DVar DA scheme, due to systematic increases in the mean geopotential OmB for SWS observations. When removing the bias from the geopotential time series of individual SWSs (Appendix B), as it has been done previously for PWSs, the mean and the standard deviation of the OmB for SWSs are not degraded when assimilating PWS geopotential observations. This indicates that new methods to detect and reduce observational biases of SWS geopotential observations, bearing in mind that these anchoring observations are the link to reality, could allow one to take advantage of PWS observations with AROME current 3DVar DA scheme. On the contrary, when assimilating the new observations with the 3D EnVar DA scheme, the OmB rms diminishes up to -10.3% for geopotential observations from SWSs situated in France.

To extend this study, the impact of assimilating PWS surface pressure observations on the quality of analyses could be studied with the degrees of freedom for signal diagnostic (Chapnik et al., 2006). If the added value of assimilating surface pressure observation appears in local areas, further research should be conducted there.

Finally, the impact of the assimilation of a dense network of surface pressure observations related to a high precipitation event has been investigated. The event has been chosen due to its high intensity of 260 mm of rainfall in 3 h observed in the southern part of the Gard department and its associated local pressure variation. While the 3DVar has difficulties in reproducing the local pressure features, the 3D EnVar was found to create reduced length scale surface pressure increments at the location where the PWS network observed the anomalies. This reduction of length scale from 200 km in 3DVar to 40 km in 3D EnVar for a single observation is illustrated in Appendix A. The surface pressure increment is able to modify temperature, relative humidity, and vertical velocity fields 50 min after the analysis. Then, the impact on rainfall has been studied by comparing forecasts to quantitative precipitation estimation products merging observations. The forecast launched from the 3D EnVar experiment assimilating PWS observations approximately 8 h before the most intense rainfall improves both the location and the intensity of the rainfall maximum compared to a forecast launched from the 3D EnVar experiment without assimilating PWS observations. The maximum rainfall accumulation is increased by 22 % in the better forecast, which is closer to the observations. Still, this maximum rainfall accumulation remains 30 % lower from the observed one, underlining the need to further improve both analyses and the AROME model. Further research on other case studies is needed to confirm the positive results.

For operational use, the flows and storage space for these data, as well as the bias correction algorithms and quality control, must be fully implemented in the model's operational environment. Discussions may be held in the mean-

time to pool some of this work with other national meteorological services. AROME-France's 3DVar DA system is currently being tested at Météo-France. The assimilation of PWS observations could therefore be included in a subsequent AROME-France test chain, controlled over several months.

Appendix A: Single-observation assimilation experiments

Single-observations assimilation experiments are performed to describe the ability of a DA scheme to propagate the OmB from an observation. In Fig. A1, one observation of geopotential in Blagnac has been assimilated, and all the other observations have been removed. The increment is quasi-isotropic in 3DVar (Fig. A1a) while slightly northward in 3DVar (Fig. A1b) at this time. The horizontal propagation of the increment on a longitudinal cross-section at the observation latitude is given by Fig. A1c and d. With the 3DVar scheme, the propagation of the increment goes up to 200 km, i.e. where the increment is reduced by 99% of the OmB. With the 3DEnVar scheme, with the localization at 25 km, the propagation goes up to 40 km.

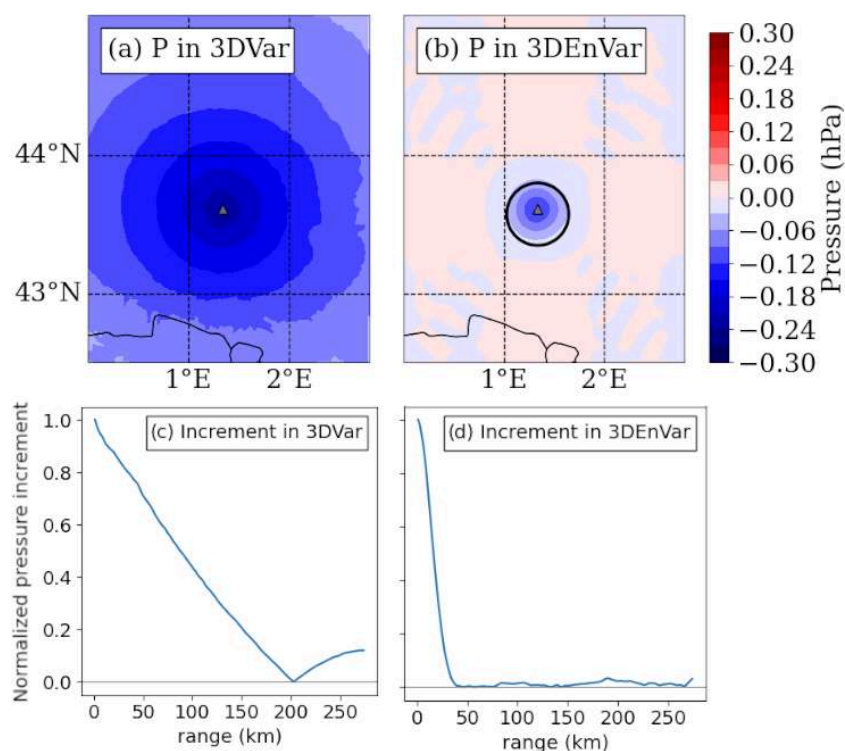


Figure A1. (a, b) Analysis increment of surface pressure resulting from the assimilation of a single observation of pressure at 1 m height in Blagnac, France, and (c, d) amplitude of the longitudinal surface pressure analysis increment divided by its maximum value called normalized pressure increment (a, c) in an idealized 3DVar experiment and (b, d) in an idealized 3DEnVar experiment starting from the same guess. The black circle indicates a 25 km distance around the observation location.

Appendix B: Sensitivity of the Desroziers diagnostic to a geopotential bias correction

In a NWP system, because SWSs observations are, as said by Ingleby (2015), our link to reality, they are not bias-corrected. Thus, the mean geopotential OmA for SWSs ranges between 4.7 and 11.9 m² s⁻² for 3DVar and 3DVarP experiments, respectively. In this article, however, PWS observations are bias-corrected before the assimilation. As a consequence, the mean geopotential OmA for PWSs is -0.1 m² s⁻² in the 3DVarP experiment (Table 3), and the Desroziers diagnostic is lower for PWSs than for SWSs. To study the sensitivity of the bias correction on this diagnostic, an experiment in which SWSs are bias-corrected was realized.

In this experiment called 3DVarP8corr (Table B1), SWS geopotential observations are assimilated after being corrected by subtracting the mean geopotential OmB computed on all time steps of 1 rolling month given by the monitoring experiment, similarly as for PWSs. The 3DVarP8corr experiment was run until the statistics, given in Table B2, were found to be stable over time, which took 8 d.

In 3DVarP8corr, the mean OmA for SWSs become close to zero (i.e. -0.2 m² s⁻²) as wished. Even if the mean OmBs for SWS are slightly diverging from the OmA statistics, the estimated Desroziers diagnostic significantly decreases from 25.0 m² s⁻² for 3DVarP8 to 10.7 m² s⁻² for 3DVarP8corr. The PWS bias correction explains why the Desroziers diagnostic is lower for PWSs than for SWSs in 3DVarP and 3DVarP8 experiments. This result confirms that the diagnostic cannot be used to compare the quality of two observing systems whose observations have not been corrected similarly.

Table B1. Overview of supplementary experiments run during 8 d between 6 September 2021 at 00:00 UTC and 14 September 2021 at 00:00 UTC.

Experiments	Duration	Cycling	DA scheme	Use of PWS surface pressure observations	Use of SWS surface pressure observations
3DVarP8 3DVarP8corr	8 d	Yes	3DVar	Assimilated after BC, QC and thinning	Assimilated without BC Assimilated after BC

Table B2. As Table 3 but for 3DVarP8 and 3DVarP8corr experiments. SWSs provide 182 observations per hour over 215 time steps, and PWSs provide 27997 observations.

	3DVarP8				3DVarP8corr			
	SWS		PWS		SWS		PWS	
	mean	SD	mean	SD	mean	SD	mean	SD
OmA	11.4	21.6	-0.1	13.0	-0.2	8.8	0.0	12.9
OmB	10.4	40.2	-0.6	34.6	-1.9	35.1	-0.8	34.8
$\tilde{\sigma}_O$	25.0		13.9		10.7		13.9	

Code availability. The code used for the assimilation experiments in AROME-France is owned by the members of the ACCORD consortium. This agreement allows each member of the consortium to license the shared ACCORD codes to academic institutions in their home countries for non-commercial research. Access to ACCORD codes, and codes used for the figures, can be obtained by contacting the corresponding author.

Data availability. World Meteorological Organization essential weather station data (SYNOP and SHIP reports), the radar mosaic at a 15 min time step, and AROME analyses and forecasts are available freely, online, in real time at <https://donneespubliques.meteofrance.fr/> (Météo-France, 2024b, c, a, d); archives of RADOME data, radar mosaic data, and AROME analyses and forecasts are available offline by making a request via this website. Personal weather station data are provided by and property of Netatmo and cannot be shared according to their licence: <https://dev.netatmo.com/legal> (Netatmo, 2024). ANTILOPE QPE and the results of data assimilation experiments are available on request from the corresponding author solely for non-commercial research purposes and for reasonable data volumes.

Author contributions. AD: formal investigations, methodology, analysis; resources; writing – original draft; writing – review and editing. MM: formal analysis; resources; writing – review and editing. VP: formal analysis; software; validation; supervision; writing – review and editing. OC: validation; project administration; supervision; validation; writing – review and editing.

Competing interests. The contact author has declared that none of the authors has any competing interests.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors.

Acknowledgements. The authors would like to thank the two anonymous referees for their comments that helped to improve the quality of the article.

Financial support. This research has been supported by the French National programme Les Enveloppes Fluides et l'Environnement (LEFE), projects ASMA and EXDOMO.

Review statement. This paper was edited by Dan Li and reviewed by two anonymous referees.

References

- Amodei, M., Sanchez, I., and Stein, J.: Verification of the French operational high-resolution model AROME with the regional Brier probability score, *Meteorol. Appl.*, 22, 731–745, <https://doi.org/10.1002/met.1510>, 2015.
- Bengtsson, L., Andrae, U., Aspelien, T., Batrak, Y., Calvo, J., de Rooy, W., Gleeson, E., Hansen-Sass, B., Homleid, M., Hortal, M., Ivarsson, K.-I., Lenderink, G., Niemelä, S., Nielsen, K. P., Onvlee, J., Rontu, L., Samuelsson, P., Muñoz, D. S., Subias, A., Tijm, S., Toll, V., Yang, X., and Ødegaard Køltzow, M.: The HARMONIE–AROME Model Configuration in the ALADIN–HIRLAM NWP System, *Mon. Weather Rev.*, 145, 1919–1935, <https://doi.org/10.1175/MWR-D-16-0417.1>, 2017.
- Benáček, P., Brožková, R., Bučánek, A., and Trojáková, A.: Tuning of 3D-Var and ALADIN-CZ system for aircraft data and assimilation, https://www.rclace.eu/media/files/Data_Assimilation/2016/rep_Patrik_ModeS.pdf (last access: 15 June 2023), 2016.
- Brousseau, P., Desroziers, G., Bouttier, F., and Chapnik, B.: A posteriori diagnostics of the impact of observations on the AROME-France convective-scale data assimilation system, *Q. J. Roy. Meteor. Soc.*, 140, 982–994, <https://doi.org/10.1002/qj.2179>, 2014.
- Brousseau, P., Seity, Y., Ricard, D., and Léger, J.: Improvement of the forecast of convective activity from the AROME-France system, *Q. J. Roy. Meteor. Soc.*, 142, 2231–2243, <https://doi.org/10.1002/qj.2822>, 2016.
- Caumont, O., Mandement, M., Bouttier, F., Eeckman, J., Lebeauupin Brossier, C., Lovat, A., Nuissier, O., and Laurantin, O.: The heavy precipitation event of 14–15 October 2018 in the Aude catchment: a meteorological study based on operational numerical weather prediction systems and standard and personal observations, *Nat. Hazards Earth Syst. Sci.*, 21, 1135–1157, <https://doi.org/10.5194/nhess-21-1135-2021>, 2021.
- Chambon, P., Mahfouf, J.-F., Audouin, O., Birman, C., Fourrié, N., Loo, C., Martet, M., Moll, P., Payan, C., Pourret, V., and Raspaud, D.: Global Observing System Experiments within the Météo-France 4D-Var Data Assimilation System, *Mon. Weather Rev.*, 151, 127–143, <https://doi.org/10.1175/MWR-D-22-0087.1>, 2023.
- Chapnik, B., Desroziers, G., Rabier, F., and Talagrand, O.: Diagnosis and tuning of observational error in a quasi-operational data assimilation setting, *Q. J. Roy. Meteor. Soc.*, 132, 543–565, <https://doi.org/10.1256/qj.04.102>, 2006.
- Clark, M. R., Webb, J. D. C., and Kirk, P. J.: Fine-scale analysis of a severe hailstorm using crowd-sourced and conventional observations, *Meteorol. Appl.*, 25, 472–492, <https://doi.org/10.1002/met.1715>, 2018.
- Courtier, P., Andersson, E., Heckley, W., Vasiljevic, D., Hamrud, M., Hollingsworth, A., Rabier, F., Fisher, M., and Pailleux, J.: The ECMWF implementation of three-dimensional variational assimilation (3D-Var). I: Formulation, *Q. J. Roy. Meteor. Soc.*, 124, 1783–1807, <https://doi.org/10.1002/qj.49712455002>, 1998.
- Dando, M. L., Thorpe, A. J., and Eyre, J. R.: The impact of targeted satellite observations on numerical weather prediction, *Q. J. Roy. Meteor. Soc.*, 133, 1945–1957, <https://doi.org/10.1002/qj.176>, 2007.
- Desroziers, G., Berre, L., Chapnik, B., and Poli, P.: Diagnosis of observation, background and analysis-error statistics in

- observation space, *Q. J. Roy. Meteor. Soc.*, 131, 3385–3396, <https://doi.org/10.1256/qj.05.108>, 2005.
- Doswell, III, C. A., Ramis, C., Romero, R., and Alonso, S.: A Diagnostic Study of Three Heavy Precipitation Episodes in the Western Mediterranean Region, *Weather Forecast.*, 13, 102–124, [https://doi.org/10.1175/1520-0434\(1998\)013<0102:ADSOTH>2.0.CO;2](https://doi.org/10.1175/1520-0434(1998)013<0102:ADSOTH>2.0.CO;2), 1998.
- ECMWF: IFS Documentation CY48R1 – Part I: Observations, 1, ECMWF, <https://doi.org/10.21957/0f360ba4ca>, 2023.
- Eyre, J. R.: Observation bias correction schemes in data assimilation systems: a theoretical study of some of their properties, *Q. J. Roy. Meteor. Soc.*, 142, 2284–2291, <https://doi.org/10.1002/qj.2819>, 2016.
- Gustafsson, N., Janjić, T., Schraff, C., Leuenberger, D., Weissmann, M., Reich, H., Brousseau, P., Montmerle, T., Wattrelot, E., Bučánek, A., Mile, M., Hamdi, R., Lindskog, M., Barkmeijer, J., Dahlbom, M., Macpherson, B., Ballard, S., Inverarity, G., Carley, J., Alexander, C., Dowell, D., Liu, S., Ikuta, Y., and Fujita, T.: Survey of data assimilation methods for convective-scale numerical weather prediction at operational centres, *Q. J. Roy. Meteor. Soc.*, 144, 1218–1256, <https://doi.org/10.1002/qj.3179>, 2018.
- Hamill, T. M., Whitaker, J. S., and Snyder, C.: Distance-Dependent Filtering of Background Error Covariance Estimates in an Ensemble Kalman Filter, *Mon. Weather Rev.*, 129, 2776–2790, [https://doi.org/10.1175/1520-0493\(2001\)129<2776:DDFOBE>2.0.CO;2](https://doi.org/10.1175/1520-0493(2001)129<2776:DDFOBE>2.0.CO;2), 2001.
- Hintz, K. S., Vedel, H., and Kaas, E.: Collecting and processing of barometric data from smartphones for potential use in numerical weather prediction data assimilation, *Meteorol. Appl.*, 26, 733–746, <https://doi.org/10.1002/met.1805>, 2019.
- Hu, G., Dance, S. L., Bannister, R. N., Chipilski, H. G., Guillet, O., Macpherson, B., Weissmann, M., and Yussouf, N.: Progress, challenges, and future steps in data assimilation for convection-permitting numerical weather prediction: Report on the virtual meeting held on 10 and 12 November 2021, *Atmos. Sci. Lett.*, 24, e1130, <https://doi.org/10.1002/asl.1130>, 2022.
- Ingleby, B.: Global assimilation of air temperature, humidity, wind and pressure from surface stations, *Q. J. Roy. Meteor. Soc.*, 141, 504–517, <https://doi.org/10.1002/qj.2372>, 2015.
- Lac, C., Chaboureaud, J.-P., Masson, V., Pinty, J.-P., Tulet, P., Escobar, J., Leriche, M., Barthe, C., Aouizerats, B., Augros, C., Aumond, P., Auguste, F., Bechtold, P., Berthet, S., Bielli, S., Bosseur, F., Caumont, O., Cohard, J.-M., Colin, J., Couvreux, F., Cuxart, J., Delautier, G., Dauhut, T., Ducrocq, V., Filippi, J.-B., Gazen, D., Geoffroy, O., Gheusi, F., Honnert, R., Lafore, J.-P., Lebeaupin Brossier, C., Libois, Q., Lunet, T., Mari, C., Maric, T., Mascart, P., Mogé, M., Molinié, G., Nuissier, O., Pantillon, F., Peyrillé, P., Pergaud, J., Perraud, E., Pianezze, J., Redelsperger, J.-L., Ricard, D., Richard, E., Riette, S., Rodier, Q., Schoetter, R., Seyfried, L., Stein, J., Suhre, K., Taufour, M., Thouron, O., Turner, S., Verrelle, A., Vié, B., Visentin, F., Vionnet, V., and Wautelet, P.: Overview of the Meso-NH model version 5.4 and its applications, *Geosci. Model Dev.*, 11, 1929–1969, <https://doi.org/10.5194/gmd-11-1929-2018>, 2018.
- Laurantin, O.: ANTILOPE: Hourly rainfall analysis merging radar and rain gauge data, in: Proceedings of the International Symposium on Weather Radar and Hydrology, Grenoble/Autrans, France, 10–15 March 2008, 2–8, 2008.
- Liu, Z.-Q. and Rabier, F.: The potential of high-density observations for numerical weather prediction: A study with simulated observations, *Q. J. Roy. Meteor. Soc.*, 129, 3013–3035, <https://doi.org/10.1256/qj.02.170>, 2003.
- Madaus, L. E. and Hakim, G. J.: Observable Surface Anomalies Preceding Simulated Isolated Convective Initiation, *Mon. Weather Rev.*, 144, 2265–2284, <https://doi.org/10.1175/MWR-D-15-0332.1>, 2016.
- Madaus, L. E., Hakim, G. J., and Mass, C. F.: Utility of Dense Pressure Observations for Improving Mesoscale Analyses and Forecasts, *Mon. Weather Rev.*, 142, 2398–2413, <https://doi.org/10.1175/MWR-D-13-00269.1>, 2014.
- Mandement, M. and Caumont, O.: Contribution of personal weather stations to the observation of deep-convection features near the ground, *Nat. Hazards Earth Syst. Sci.*, 20, 299–322, <https://doi.org/10.5194/nhess-20-299-2020>, 2020.
- Masson, V., Le Moigne, P., Martin, E., Faroux, S., Alias, A., Alkama, R., Belamari, S., Barbu, A., Boone, A., Bouyssel, F., Brousseau, P., Brun, E., Calvet, J.-C., Carrer, D., Decharme, B., Delire, C., Donier, S., Essaouini, K., Gibelin, A.-L., Giordani, H., Habets, F., Jidane, M., Kerdran, G., Kourzeneva, E., Lafaysse, M., Lafont, S., Lebeaupin Brossier, C., Lemonsu, A., Mahfouf, J.-F., Marguinaud, P., Mokhtari, M., Morin, S., Pigeon, G., Salgado, R., Seity, Y., Taillefer, F., Tanguy, G., Tulet, P., Vincendon, B., Vionnet, V., and Voltaire, A.: The SURFEXv7.2 land and ocean surface platform for coupled or offline simulation of earth surface variables and fluxes, *Geosci. Model Dev.*, 6, 929–960, <https://doi.org/10.5194/gmd-6-929-2013>, 2013.
- McNicholas, C. and Mass, C. F.: Smartphone Pressure Collection and Bias Correction Using Machine Learning, *J. Atmos. Ocean. Tech.*, 35, 523–540, <https://doi.org/10.1175/jtech-d-17-0096.1>, 2018a.
- McNicholas, C. and Mass, C. F.: Impacts of Assimilating Smartphone Pressure Observations on Forecast Skill during Two Case Studies in the Pacific Northwest, *Weather Forecast.*, 33, 1375–1396, <https://doi.org/10.1175/waf-d-18-0085.1>, 2018b.
- Meier, F., Fenner, D., Grassmann, T., Otto, M., and Scherer, D.: Crowdsourcing air temperature from citizen weather stations for urban climate research, *Urban Climate*, 19, 170–191, <https://doi.org/10.1016/j.uclim.2017.01.006>, 2017.
- Météo-France: Animation mosaïque radar France métropole, Météo-France [data set], https://donneespubliques.meteofrance.fr/?fond=produit&id_produit=98&id_rubrique=34 (last access: 15 March 2024), 2024a.
- Météo-France: Observations du réseau sol de France, Météo-France [data set], https://donneespubliques.meteofrance.fr/?fond=produit&id_produit=93&id_rubrique=32 (last access: 15 March 2024), 2024b.
- Météo-France: Observations en mer (Messages SHIP, BUOY), Météo-France [data set], https://donneespubliques.meteofrance.fr/?fond=produit&id_produit=95&id_rubrique=32 (last access: 15 March 2024), 2024c.
- Météo-France: Données de modèle AROME, Météo-France [data set], https://donneespubliques.meteofrance.fr/?fond=produit&id_produit=131&id_rubrique=51 (last access: 15 March 2024), 2024d.
- Michel, Y. and Brousseau, P.: A Square-Root, Dual-Resolution 3DVar for the AROME Model: Formulation and Evaluation

- on a Summertime Convective Period, *Mon. Weather Rev.*, 149, 3135–3153, <https://doi.org/10.1175/MWR-D-21-0026.1>, 2021.
- Mile, M., Benáček, P., and Rózsa, S.: The use of GNSS zenith total delays in operational AROME/Hungary 3D-Var over a central European domain, *Atmos. Meas. Tech.*, 12, 1569–1579, <https://doi.org/10.5194/amt-12-1569-2019>, 2019.
- Mittermaier, M. P.: A “Meta” Analysis of the Fractions Skill Score: The Limiting Case and Implications for Aggregation, *Mon. Weather Rev.*, 149, 3491–3504, <https://doi.org/10.1175/MWR-D-18-0106.1>, 2018.
- Montmerle, T., Michel, Y., Arbogast, E., Ménétrier, B., and Brousseau, P.: A 3D ensemble variational data assimilation scheme for the limited-area AROME model: Formulation and preliminary results, *Q. J. Roy. Meteor. Soc.*, 144, 2196–2215, <https://doi.org/10.1002/qj.3334>, 2018.
- Netatmo: Smart weather station technical specifications, <https://www.netatmo.com/en-gb/smart-weather-station/specifications> (last access: 29 November 2023), 2023.
- Netatmo: Weather API Documentation, Netatmo [data set], <https://dev.netatmo.com/apidocumentation/weather> (last access: 15 March 2024), 2024.
- Nipen, T. N., Seierstad, I. A., Lussana, C., Kristiansen, J., and Hov, Ø.: Adopting Citizen Observations in Operational Weather Prediction, *B. of Am. Meteor. Soc.*, 101, E43–E57, <https://doi.org/10.1175/BAMS-D-18-0237.1>, 2020.
- Nuissier, O., Duffourg, F., Martinet, M., Ducrocq, V., and Lac, C.: Hectometric-scale simulations of a Mediterranean heavy-precipitation event during the Hydrological cycle in the Mediterranean Experiment (HyMeX) first Special Observation Period (SOP1), *Atmos. Chem. Phys.*, 20, 14649–14667, <https://doi.org/10.5194/acp-20-14649-2020>, 2020.
- Pauley, P. M. and Ingleby, B.: *Assimilation of In-Situ Observations*, Springer International Publishing, Cham, 293–371, ISBN 978-3-030-77722-7, https://doi.org/10.1007/978-3-030-77722-7_12, 2022.
- Pourret, V., Mahfouf, J.-F., Guidard, V., Moll, P., Doerenbecher, A., and Pignatelli, B.: Variational bias correction for Mode-S aircraft derived winds, *Tellus A*, 73, 1886808, <https://doi.org/10.1080/16000870.2021.1886808>, 2021.
- Pourret, V., Šavli, M., Mahfouf, J.-F., Raspaud, D., Doerenbecher, A., Bénichou, H., and Payan, C.: Operational assimilation of Aeolus winds in the Météo-France global NWP model ARPEGE, *Q. J. Roy. Meteor. Soc.*, 148, 2652–2671, <https://doi.org/10.1002/qj.4329>, 2022.
- Ridal, M., Bojarova, J., Hagelin, S., Medeiros, P., Gregow, E., and Stappers, R.: Study of forecast quality improvement (magnitude and duration), relative to using a reference set of observations, from assimilation of novel observations, https://www.ecmwf.int/sites/default/files/special_projects/2020/spnolabo-2020-finalreport.pdf (last access: 15 June 2023), 2019.
- Seity, Y., Brousseau, P., Malardel, S., Hello, G., Bénard, P., Bouttier, F., Lac, C., and Masson, V.: The AROME-France Convective-Scale Operational Model, *Mon. Weather Rev.*, 139, 976–991, <https://doi.org/10.1175/2010MWR3425.1>, 2011.
- Sgoff, C., Acevedo, W., Paschalidi, Z., Ulbrich, S., Bauernschubert, E., Kratzsch, T., and Potthast, R.: Assimilation of crowd-sourced surface observations over Germany in a regional weather prediction system, *Q. J. Roy. Meteor. Soc.*, 148, 1752–1767, <https://doi.org/10.1002/qj.4276>, 2022.
- STMicroelectronics: MEMS pressure sensor: 260-1260 hPa absolute digital output barometer, rev. 4, <https://www.st.com/en/mems-and-sensors/lps25hb.html#documentation> (last access: 15 June 2023), 2016.
- Tardieu, J. and Leroy, M.: Radome, le réseau temps réel d’observation au sol de Météo-France, *La Météorologie*, 40–43, <https://doi.org/10.4267/2042/36262>, 2003.
- Vasiljevic, D., Andersson, E., Isaksen, L., and Garcia-Mendez, A.: Surface pressure bias correction in data assimilation, Tech. rep., ECMWF, <https://doi.org/10.21957/UV295RFMX5>, 2006.
- Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat, İ., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., and SciPy 1.0 Contributors: SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python, *Nat. Methods*, 17, 261–272, <https://doi.org/10.1038/s41592-019-0686-2>, 2020.
- Wattrelot, E., Montmerle, T., and Guerrero, C. G.: Evolution of the assimilation of radar data in the AROME model at convective scale, http://www.meteo.fr/cic/meetings/2012/ERAD/extended_abs/NWP_401_ext_abs.pdf (last access: 15 June 2023), 2012.
- World Meteorological Organization: Manual on the Global Telecommunication System: Annex III to the WMO Technical Regulations (WMO-No. 386), https://library.wmo.int/doc_num.php?explnum_id=10469 (last access: 15 June 2023), 2020.
- World Meteorological Organization: Manual on the WMO Integrated Global Observing System: Annex VIII to the WMO Technical Regulations (WMO-No. 1160), <https://community.wmo.int/en/activity-areas/wigos/gbon> (last access: 31 October 2023), 2021.

3.3 Compléments de l'article

3.3.1 Supervision a posteriori des observations SWS assimilées

Le suivi durant l'assimilation des observations PWS de pression, permet une meilleure compréhension du biais et de la qualité de ces données. Le suivi des observations SWS peut être fait en parallèle, aussi appelé *monitoring*, et fait l'objet de cette section. L'expérience 3DEnVar nous renseigne sur les statistiques d'OmB des SWS et l'expérience 3DEnVarP sur celles des PWS.

La figure 3.2 montre le comportement spatial des observations en géopotential des réseaux de SWS et de PWS. La moyenne et l'écart-type des écarts à l'ébauche (OmB) par station SWS (figure 3.2b,e) augmente par rapport à celles des OmA (figure 3.2a,d), notamment sur le relief alpin. Cette augmentation est synonyme d'écartement entre les observations et le modèle à 1 h de prévision par rapport à l'analyse.

Une disparité des statistiques sur les observations est notable suivant les pays européens et met en évidence une disparité de la qualité des données. La moyenne des statistiques par pays est résumée dans le tableau 3.1. L'écart-type renseigne sur la variabilité au second ordre de la différence entre le modèle et l'observation, qui est la somme quadratique de l'erreur du modèle et l'erreur d'observation, synonyme de la qualité des mesures disponibles. Les SWS, individuellement, ont des écart-types homogènes sur le domaine AROME, indiquant une qualité du modèle et des instruments comparables. Les PWS ont, eux aussi, un écart-type comparable aux SWS.

La moyenne des OmB renseigne sur le biais des stations par rapport à l'ébauche (figure 3.2a-c), correspondant à un mauvais étalonnage, une dérive de l'instrument, une mauvaise indication de l'altitude de la station dans le modèle ou au biais de l'ébauche du modèle. Il y a une grande disparité des biais entre les stations et entre les pays. L'Allemagne est le pays qui possède le réseau avec le moins de SWS biaisées par rapport à l'ébauche. Au contraire, l'Espagne ou l'Italie présentent des variations importantes d'une station à l'autre, avec des OmB absolu moyen à $49.7 \text{ m}^2 \text{ s}^{-2}$ et $56.1 \text{ m}^2 \text{ s}^{-2}$ respectivement. Le réseau italien rassemble, en effet, des systèmes d'observations de surface des sources diversifiées et de qualités hétérogènes. Une meilleure prise en compte des observations serait possible avec un étalonnage ou une correction de l'altitude des stations comportant les moyennes d'OmB les plus élevées. Les observations des PWS ayant été débiaisées par rapport aux ébauches du modèle, elles présentent l'OmB absolu moyen le plus bas, avec $9.0 \text{ m}^2 \text{ s}^{-2}$.

3.3.2 Gain spatial des observations en pression des PWS

L'assimilation des observations de pression des PWS avec le schéma d'assimilation 3DEnVar a montré une baisse de l'erreur de prévision 1 h de la pression au niveau des stations de référence plus important en France que sur le domaine AROME. La répartition de cette baisse suivant les SWS est uniforme sur la France et se développe très faiblement sur les stations limitrophes à la France (figure 3.3). Il serait pertinent d'assimiler l'ensemble des observations provenant des PWS disponibles sur tout le domaine AROME. De plus, les SWS biaisées vis-à-vis de l'ébauche (figure 3.2.b) ne montrent pas de baisse de ΔRMS OmB. Autrement dit, en enlevant les SWS de mauvaise qualité, les scores vis-à-vis des SWS seraient encore meilleurs.

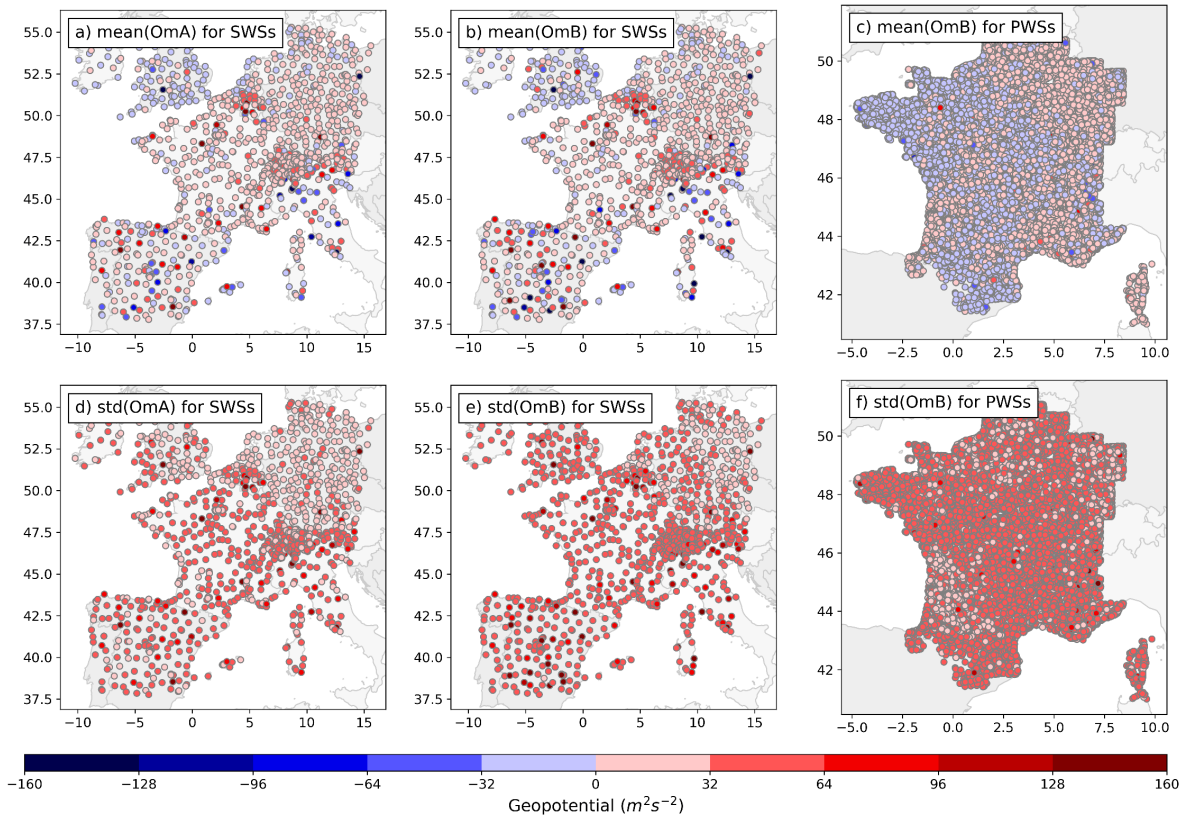


FIGURE 3.2 – Cartes des moyennes (a-c) et écart type (d-f) des écarts entre le modèle et les observations sur le mois d'étude ; (a, d) des écarts OmA des SWS, (b, e) des écarts OmB des SWS de l'expérience 3DEnVar et (c, f) des écarts OmB des PWS après prétraitement de l'expérience 3DEnVarP.

TABLEAU 3.1 – Moyenne par pays comptant plus de 50 stations sur le domaine AROME de la moyenne et de l'écart type (m^2s^{-2}) des séries temporelles d'écart à l'ébauche des observations de géopotential SWS et PWS. Réalisé sur le mois de la période d'étude. Les stations avec un écart-type supérieur à $100 m^2 s^{-2}$ sont enlevées, représentant moins de 2% des stations. (*) Les stations PWS couvrent les zones frontalières des pays limitrophes à la France.

Réseau	Pays	Moyenne des $\ OmB\ $	Ecart type des OmB	Nb de SWS
SWS	Allemagne	11.5	30.5	197
	Royaume-Uni	17.7	35.1	96
	France	20.9	38.2	220
	Suisse	30.8	38.2	134
	Espagne	49.7	37.9	134
	Italie	56.1	44.3	54
PWS	France*	9.0	36.1	44251

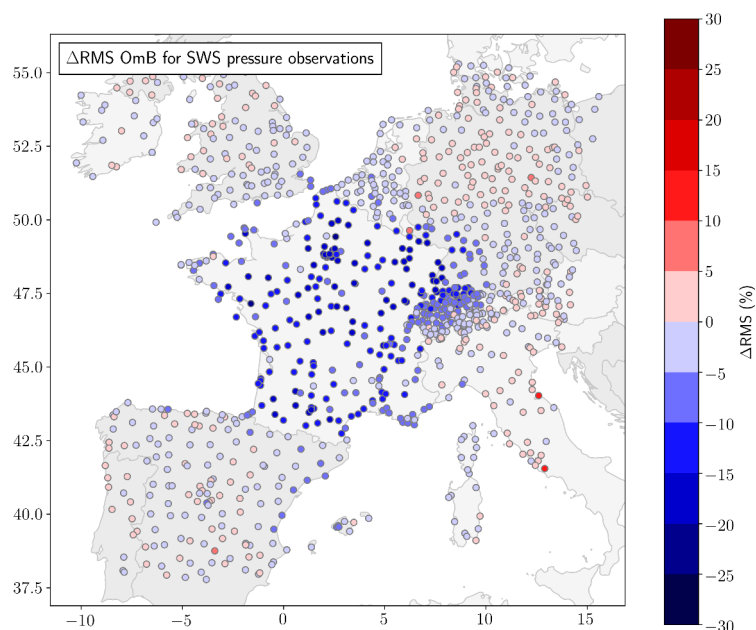


FIGURE 3.3 – $\Delta\text{RMS OmB}$ (%) de l'expérience 3DEnVarP par rapport à 3DEnVar de la pression de surface pour chaque SWS sur le domaine AROME.

3.3.3 Vérifications du système d'assimilation

L'assimilation des observations PWS de pression augmente le nombre d'observations au sol d'un facteur 2.3. Cette augmentation considérable peut poser des problèmes de convergence de la minimisation de la fonction de coût. La minimisation dans AROME-France utilisant le schéma 3DEnVar a une limite de 50 itérations (Michel et Brousseau, 2021). En mode opérationnel, la minimisation va toujours jusqu'au nombre maximum d'itérations, ce qui garantit un temps de calcul quasi identique. Pour expliquer le fonctionnement de la minimisation en mode recherche, nous revenons sur la fonction de coût :

$$J(\mathbf{x}) = J_b(\mathbf{x}) + J_o(\mathbf{x}), \quad (3.1)$$

avec J_o et J_b respectivement les termes d'ébauche et d'observation de la fonction de coût. En mode recherche, la minimisation s'arrête à l'itération i si la norme du gradient du terme d'ébauche de la fonction de coût est inférieure à 10^{-3} fois la norme du gradient du terme d'ébauche de la fonction de coût initiale. La figure 3.4 montre la norme du gradient du terme d'ébauche de la fonction de coût $DJ(i) = \|J_b(\mathbf{x}_i) - J_b(\mathbf{x}_{i-1})\|$ en fonction du nombre d'itérations i . Pour cette minimisation, même si la norme du gradient initial de la fonction de coût est plus élevée dans 3DEnVarP que dans 3DEnVar, la minimisation ne prend qu'une itération supplémentaire (Fig 3.4). Au cours de la période d'étude d'un mois, le nombre moyen d'itérations augmente de 0.6 dans 3DEnVarP par rapport à 3DEnVar. L'ajout des nouvelles observations augmente légèrement le nombre d'itérations, mais reste loin des 50 itérations réalisées systématiquement en opérationnel.

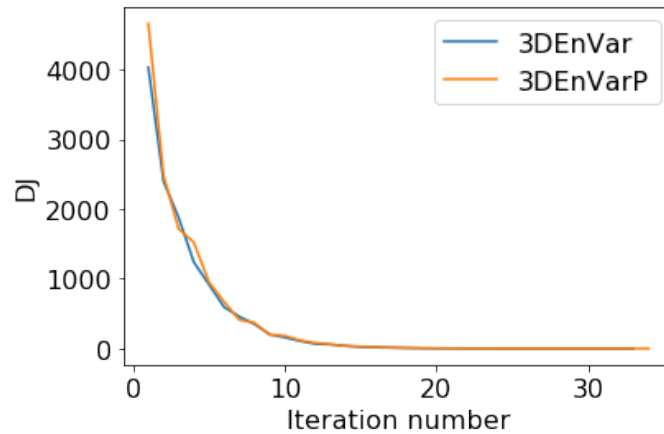


FIGURE 3.4 – Évolution de la norme du gradient du terme d'ébauche de la fonction de coût DJ en fonction du nombre d'itérations lors de la minimisation de la fonction de coût les expériences 3DEnVarP et 3DEnVar le 5 octobre 2021 à 23:00 UTC.

3.3.4 Tests de sensibilité

3.3.4.1 Erreur d'observations et échelle d'écrémage

Différents paramètres peuvent être ajustés pour comprendre l'influence des observations de pression de surface dans le modèle. Deux paramètres ont été modifiés :

- L'écart type de l'erreur d'observation σ_o des observations de pression. Par défaut σ_o est prescrit à $8g$ avec g la constante de gravitation universelle. L'estimation, par le diagnostic de Desroziers, indique une valeur pour les SWS et les PWS proche de $10 \text{ m}^2 \text{ s}^{-2}$ équivalent à $\sigma_o/8$. Deux expériences avec un σ_o divisé par deux ont été produites : 3DEnVar_ $\sigma_o/2$ avec le schéma d'assimilation 3DEnVar en configuration opérationnelle et 3DEnVarP_ $\sigma_o/2$ avec le schéma d'assimilation 3DEnVar et l'assimilation des observations de pression PWS pré-traitées.
- L'échelle de l'écrémage des observations PWS. Par défaut, l'écrémage correspond à sélectionner une PWS par maille de 1.3 km de résolution. Deux expériences ont été réalisées : 3DEnVarP_noTh sans écrémage (*no thinning*) et 3DEnVarP_Th10 avec un écrémage d'une observation par maille de 10 km ; l'un multipliant par 1.7 et l'autre divisant par 5 le nombre d'observations assimilées respectivement.

Ces changements illustrent les réglages faits dans les schémas d'assimilation de données. La figure 3.5 montre l'incrément d'analyse de la première heure d'assimilation des expériences longues. À cette échéance l'ébauche est identique entre les expériences. Au vu des incréments en 3DVar, l'ébauche est éloignée des observations de pression SWS, avec une anomalie négative sur l'ensemble du domaine présentant un minimum de -1.2 hPa (figure 3.5.a). L'expérience en 3DEnVar donne peu de poids aux observations de surface et seulement quelques anomalies localisées sont remarquées (figure 3.5.b). Une option pour redonner plus de poids aux observations de surface, consiste à diminuer leur σ_o . L'expérience 3DEnVar_ $\sigma_o/2$ montre des anomalies plus marquées que précédemment, avec un minimum à -0.8 hPa , mais d'une taille réduite causée par l'échelle de localisation fixée à 25 km pour les observations de surface (figure 3.5.c). Finalement, l'assimilation des observations de pression PWS, couvrant la zone avec une plus forte densité, permet de compenser le faible poids donné aux observations de surface dans le 3DEnVar, avec un incrément d'autant plus fort qu'on donne de l'importance aux PWS soit au travers du σ_o , soit au travers de l'écrémage (figure 3.5.d-f). L'échelle de la localisation pourrait être

modifiée, pour étendre l'influence des observations de surface, mais cela ne résoudrait pas le problème du manque de poids donné aux observations de surface dans cette configuration du 3DEnVar et aurait un impact non maîtrisé sur les autres systèmes d'observation de surface assimilés.

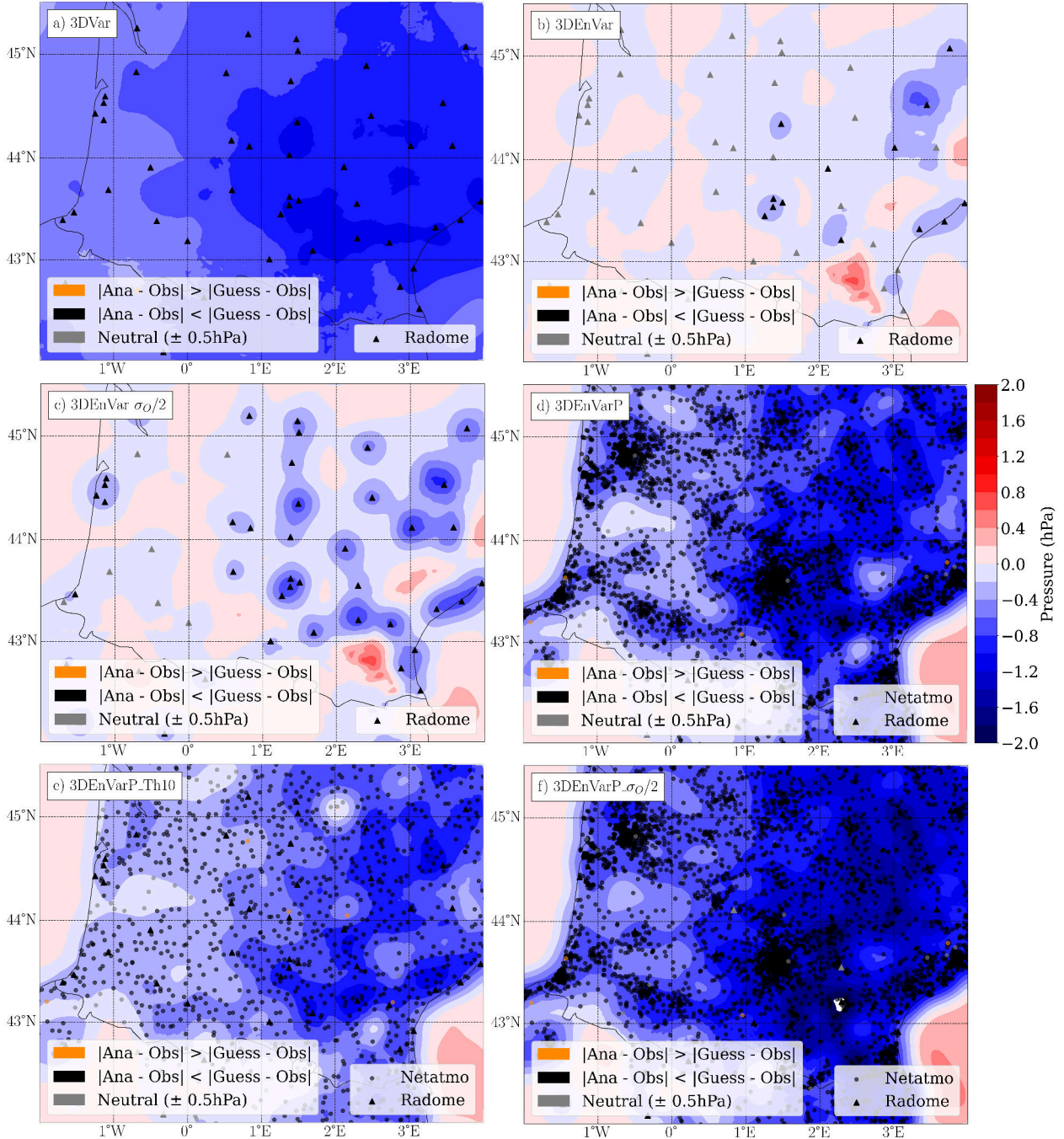


FIGURE 3.5 – Incrément d'analyse de pression de surface le 6 septembre 2021 à 01:00 UTC, partant de la même ébauche. Expériences a) 3DVar, b) 3DEnVar, c) 3DEnVar_σ₀/2, d) 3DEnVarP, e) 3DEnVar_Th10 et f) 3DEnVarP_σ₀/2.

Une dernière question reste en suspens : y a-t-il un avantage systématique à modifier l'erreur d'observation, ou l'échelle spatiale de l'écrémage ? Le Tableau 5.2 présente les scores de prévision à une heure d'échéance à la surface des nouvelles expériences cyclées

TABLEAU 3.2 – Δ RMS OmB (%) de pression de surface, température à 2 m, humidité relative à 2 m et du vent zonal à 10 m provenant des observations SWS sur le mois d'étude. Les valeurs négatives d'une expérience (XP) par rapport à l'expérience 3DEnVar indiquent que les prévisions à 1 h de XP sont plus proches des observations SWS que celles de l'expérience 3DEnVar. On nommera « gain », les valeurs négatives de Δ RMS OmB. Les valeurs significatives sont en gras.

	Surface pressure		Temperature		Relative humidity		Zonal wind (U)	
	France	AROME domain	France	AROME domain	France	AROME domain	France	AROME domain
3DEnVarP	-10.4	-3.0	0.9	0.6	1.4	0.8	0.1	0.2
3DEnVarP_noTh	-10.6	-3.1	0.9	0.7	1.9	1.1	0.1	0.3
3DEnVarP_Th10	-9.0	-2.7	0.5	0.4	0.6	0.4	-0.2	0.0
3DEnVarP_ $\sigma_o/2$	-11.4	-4.9	1.4	1.0	2.4	1.6	0.2	0.3

pendant le mois d'étude. Concernant le réglage de l'échelle d'écrémage, l'augmentation de l'échelle à une maille de 10 km de résolution, filtrant plus d'observations PWS, diminue le gain en pression de surface, passant de -10.4% en 3DEnVarP à -9.0% pour les observations SWS en France, et réduisant, de facto, l'augmentation de l'erreur de prévision à 1 heure d'échéance pour la température et l'humidité relative à 2 m. A contrario, enlever l'écrémage (3DEnVarP_noTh) augmente légèrement le gain passant à -10.6% en pression de surface, mais augmente l'erreur sur la prévision 1 h de l'humidité relative au niveau des observations SWS. Concernant l'expérience avec la diminution de l'erreur d'observation par un facteur 2 pour les observations de pression de surface (3DEnVarP_ $\sigma_o/2$), le gain est plus important encore, passant pour la pression de -10.3% en 3DEnVar à -11.4% pour les observations SWS en France. Comme précédemment, ce gain en pression de surface s'accompagne d'une dégradation plus marquée de l'innovation des observations SWS de température et d'humidité relative.

Rapprocher fortement les champs de pression de surface des observations de pression PWS, peut entraîner une dégradation des autres champs de surface. Cette dégradation peut être en partie due à un mauvais réglage de l'influence des observations de pression sur les autres champs du modèle, provenant des covariances croisées de la matrice d'erreur d'ébauche B en 3DEnVar. La dégradation peut aussi venir des hypothèses faites lors de l'assimilation de données variationnelle, comme celle de la linéarité de l'opérateur d'observation localement, et peut entraîner des déséquilibres de grande échelle dans le modèle. Il est nécessaire de trouver un équilibre dans le réglage du nombre d'observations de PWS et de leur poids relatif par rapport à toutes les autres observations déjà assimilées. Par exemple, en acceptant de diminuer légèrement le gain en pression, avec un écrémage à une échelle de 10 km, la dégradation en humidité relative des innovations de SWS n'est plus significative.

3.3.4.2 Spin-up

En raison des déséquilibres dans les incréments d'analyses, un temps de mise à l'équilibre, permettant le retour à l'équilibre des champs modèle, est nécessaire. Le *spin-up* peut être diagnostiqué grâce à l'évolution de la tendance de pression sur le domaine (Brousseau *et al.*, 2016). Un exemple sur quatre assimilations successives est montré dans la figure 3.6. Partant d'une analyse, la tendance de pression diminue fortement dès le premier pas de temps, puis de moins en moins vite. D'après Brousseau *et al.* (2016), au bout d'une heure de prévision, la tendance de pression a suffisamment baissé pour atteindre un état où les champs sont cohérents entre eux. Sur les quatre réseaux, quand les observations PWS

sont assimilées, une augmentation de la moyenne quadratique de la tendance de pression est notable sur les premières minutes de prévision, mais celle-ci converge à 1 h d'échéance. Sur cet exemple, il n'y a pas d'augmentation systématique du spin-up entre l'expérience assimilant toutes les observations de pression de PWS (3DEnVarP_noTh) et celle réalisant un écrémage à 1.3 km (3DEnVarP). Le calcul moyen du spin-up sur deux journées a été réalisé et ne montre pas d'augmentation significative (non montré). Bien que le calcul de spin-up puisse être réalisée sur une plus grande période, ces résultats préliminaires montrent qu'il n'y a pas d'augmentation notable du spin-up lors de l'ajout des nouvelles observations de surface.

Par ailleurs, un nouveau diagnostic de spin-up a été récemment présenté par Diefenbach *et al.* (2023). Ces derniers suggèrent l'utilisation de l'approximation du gradient de température faible, comme indicateur d'équilibre pour l'assimilation de données à l'échelle convective. Cette métrique pourra être utilisée dans de futures études sur l'assimilation de nouvelles observations et devra être comparée à la précédente métrique.

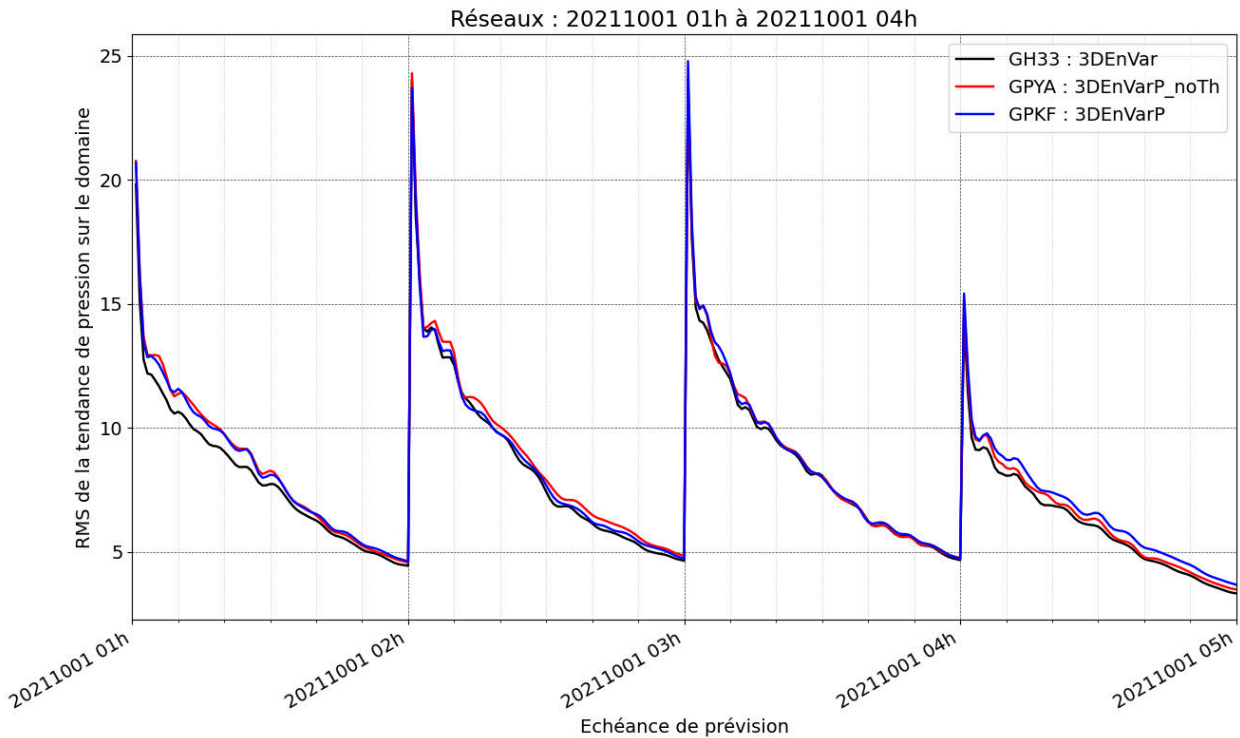


FIGURE 3.6 – Moyenne quadratique de la tendance de pression de surface du cycle d'assimilation-prévision AROME, pendant quatre réseaux successifs au cours des expériences le 1er octobre 2021 à 01:00 UTC, pour l'expérience 3DEnVar (noir), 3DEnVarP (bleu) et 3DEnVarP_noTh (en rouge). 3DEnVarP_noTh correspond à une expérience d'assimilation des observations de pression PWS sans écrémage des observations.

3.4 Phénomènes à fort gradient de pression

L'intérêt de l'assimilation des observations de pression des PWS se porte sur les cas présentant de forts gradients de pression atmosphérique comme pour le cas présenté précédemment concernant le phénomène convectif du 14 septembre 2023 (sect. 3.2.5). D'autres exemples, introduits dans le premier chapitre, sont ici présentés.

3.4.1 Front - 5 octobre 2021

La figure 3.7 présente l'analyse de pression au niveau de la mer (P_{mer}) à l'arrivée d'un front froid sur la Bretagne (nord-ouest de la France). Les quatre expériences 3DVar, 3DEnVar, 3DVarP et 3DEnVarP sont cyclées depuis le 6 septembre 2021, soit depuis 30 jours. Les quatre expériences sont capables de simuler l'arrivée du front, avec un gradient de pression de plus de 8 hPa sur les 250 km du domaine. En 3DVar, le gradient est représenté par des isobares parallèles, alors qu'avec l'expérience 3DEnVar, on observe une anomalie de pression (2.8 °W, 48.1 °N) arrivée lors de l'assimilation de données (non montré). Les deux expériences 3DVarP et 3DEnVarP ont une analyse de P_{mer} comparable, indiquant que les schémas différents d'assimilation utilisés peuvent produire des analyses prenant en compte les structures de pression observées par les PWS.

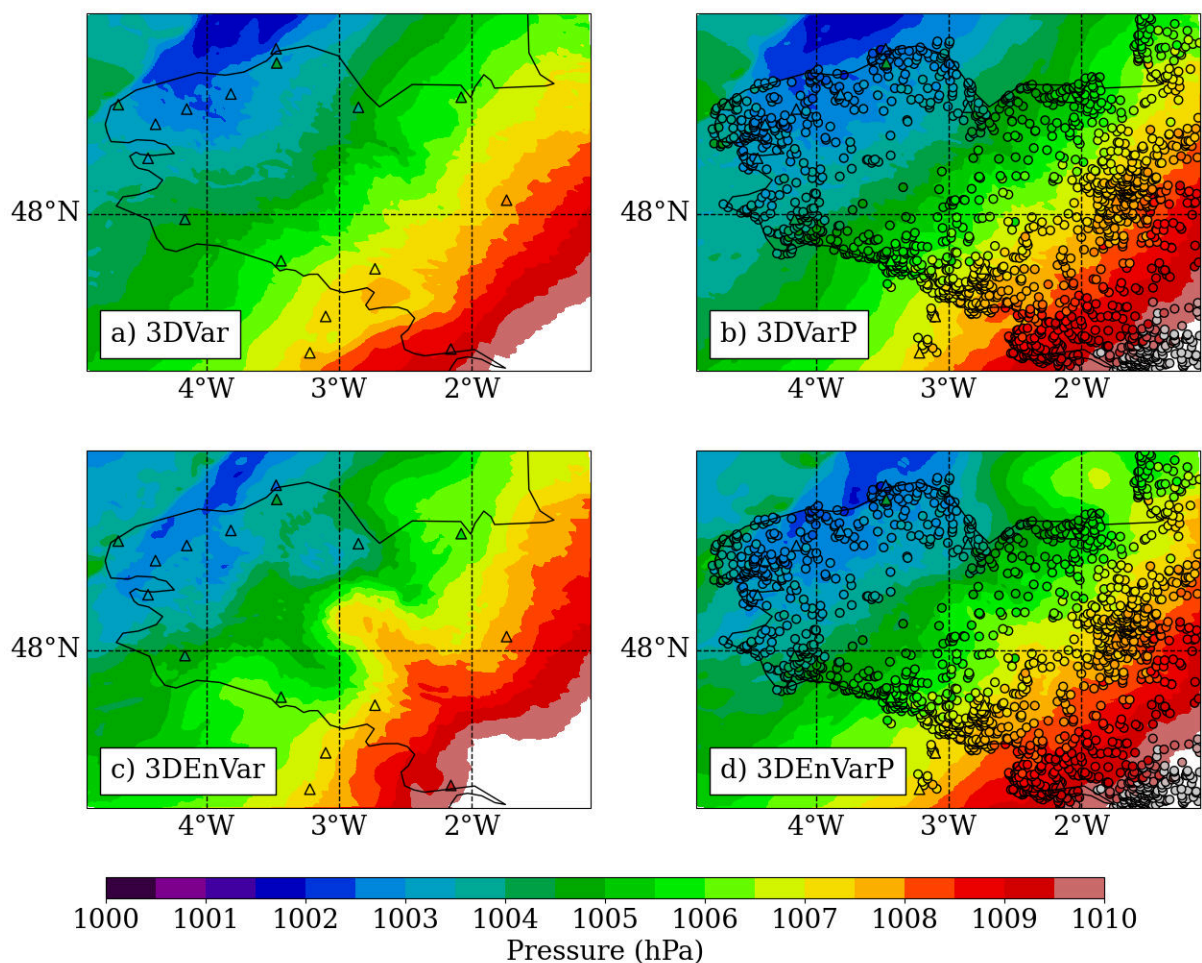


FIGURE 3.7 – Analyse de P_{mer} à 00:00 UTC du front simulé par AROME le 5 octobre 2021. Le fond représente la pression réduite au niveau de la mer des expériences (a) en 3DVar, (b) en 3DVar avec l'assimilation des observations de pression des PWS, (c) en 3DEnVar, et (d) en 3DEnVar avec l'assimilation des observations de pression des PWS. Les expériences sont cyclées depuis le 6 septembre 2021. Le calcul de P_{mer} est effectué suivant Mandement et Caumont (2020).

Le front froid traverse le nord de la France entre 00:00 UTC et 14:00 UTC avant de se combler au nord-est du pays. À partir des analyses de 00:00 UTC, des prévisions jusqu'à

24 h d'échéance ont été produites. Des scores ont été réalisés sur l'écart à la prévision de chaque expérience aux observations de P_{mer} des SWS et PWS (figure 3.8). Les écarts des prévisions aux observations SWS sont plus élevées que pour les observations PWS, même pour les expériences de contrôle 3DVar et 3DEnVar, montrant la qualité des observations PWS (les deux réseaux étant répartis de manière homogène après écrémage des PWS). Les écarts entre les deux analyses 3DVarP et 3DEnVarP, et les observations PWS sont similaires, respectivement de 0.21 et 0.23 hPa. Ces écarts sont inférieurs aux expériences de contrôle 3DVar et 3DEnVar, avec des valeurs respectivement de 0.35 hPa et 0.7 hPa. L'écart de la prévision aux observations PWS provenant de l'expérience 3DVarP (resp. 3DEnVarP) reste inférieure à l'expérience de contrôle 3DVar (resp. 3DEnVar) jusqu'à 9 h d'échéance. Après 09:00 UTC, les écarts de la prévision aux observations des expériences 3DVar et 3DVarP d'une part, et 3DEnVar et 3DEnVarP d'autre part, se rejoignent : l'analyse a toujours une influence importante sur la prévision, mais c'est le schéma d'assimilation qui joue le rôle prépondérant. Enfin après 17:00 UTC, les écarts des prévisions aux observations convergent indépendamment du schéma d'assimilation. Les forçages communs aux expériences deviennent plus importants que l'information issue de l'analyse. Il peut s'agir des conditions initiales sur océan, ou du modèle qui force la prévision sur les bords. Ce résultat est cohérent avec les scores sur la période d'un mois, montrant que l'influence des observations PWS à l'analyse s'étend en moyenne jusqu'à 9 h d'échéance. Pour ce phénomène d'échelle caractéristique méso- α (200 à 2000 km), c'est l'expérience utilisant le schéma d'assimilation 3DVar avec l'assimilation des observations de pression des PWS qui donne les écarts entre la prévision et les observations de P_{mer} les plus bas.

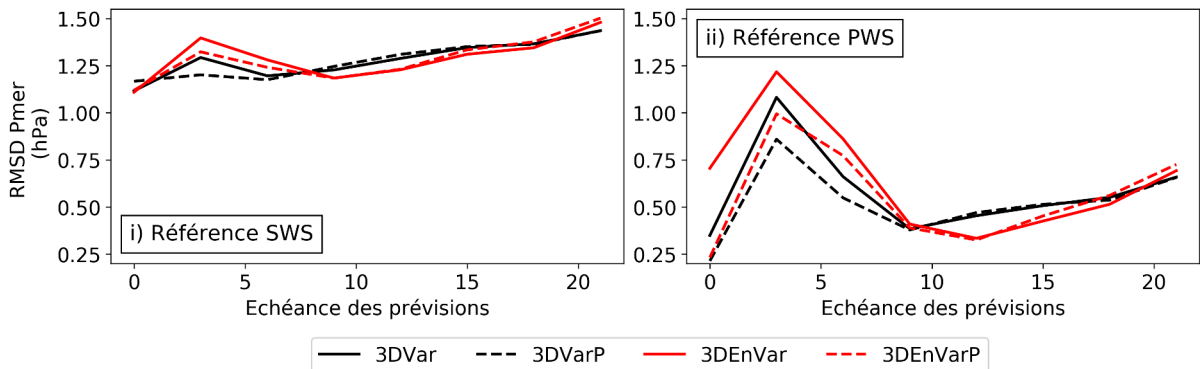


FIGURE 3.8 – Racine carrée de l'écart quadratique moyen des observations a) SWS et b) PWS, à chaque échéance de la prévision partant des analyses à 00:00 UTC des expériences 3DVar, 3DVarP, 3DEnVar et 3DEnVarP (figure 3.7) sur la moitié nord de la France (-5 à 5°W , 46.5 à 51.2°N). Pour une répartition homogène des stations, un écrémage à 10 km de résolution est réalisé pour les PWS, gardant de l'ordre de 2240 stations PWS, pour 128 stations SWS.

3.4.2 Dépression secondaire - 3 novembre 2023

Dans la nuit du 2 au 3 novembre 2023, une dépression secondaire de surface, au sein d'un vaste talweg, se creuse avant d'arriver dans les Landes, dans le sud-ouest de la France. Il n'y a pas de rejet particulier d'observations de PWS, le centre dépressionnaire passe dans une zone peu dense en observations de PWS. À 03:00 UTC, il ne reste que 2 observations PWS proches du centre dépressionnaire avec un minimum observé à 985.1 hPa. Ce phénomène arrivant sur terre à partir de 03:00 UTC, les expériences sont seulement

cyclées à partir de 00:00 UTC.

Cette dépression est bien simulée par les quatre expériences (figure 3.9), néanmoins sa forme et son intensité diffèrent. Avec le schéma 3DVar, l'assimilation des observations PWS (3DVarP) permet d'intensifier le minimum dépressionnaire, passant de 987.7 hPa en 3DVar à 986.5 hPa en 3DVarP, cependant sa forme reste inchangée. Dans l'expérience 3DEnVar, le centre dépressionnaire est plus creux qu'en 3DVar, atteignant 985.7 hPa, proche du minimum observé (soit +0.6 hPa). L'assimilation des observations PWS (3DEnVarP) ne change pas la valeur du minimum. Par contre, dans 3DEnVarP les deux minimums locaux de pression visible dans 3DEnVar, ont été comblés par l'assimilation des observations de pression des PWS. L'analyse de l'expérience 3DEnVarP est la plus proche des observations de pression des SWS et PWS.

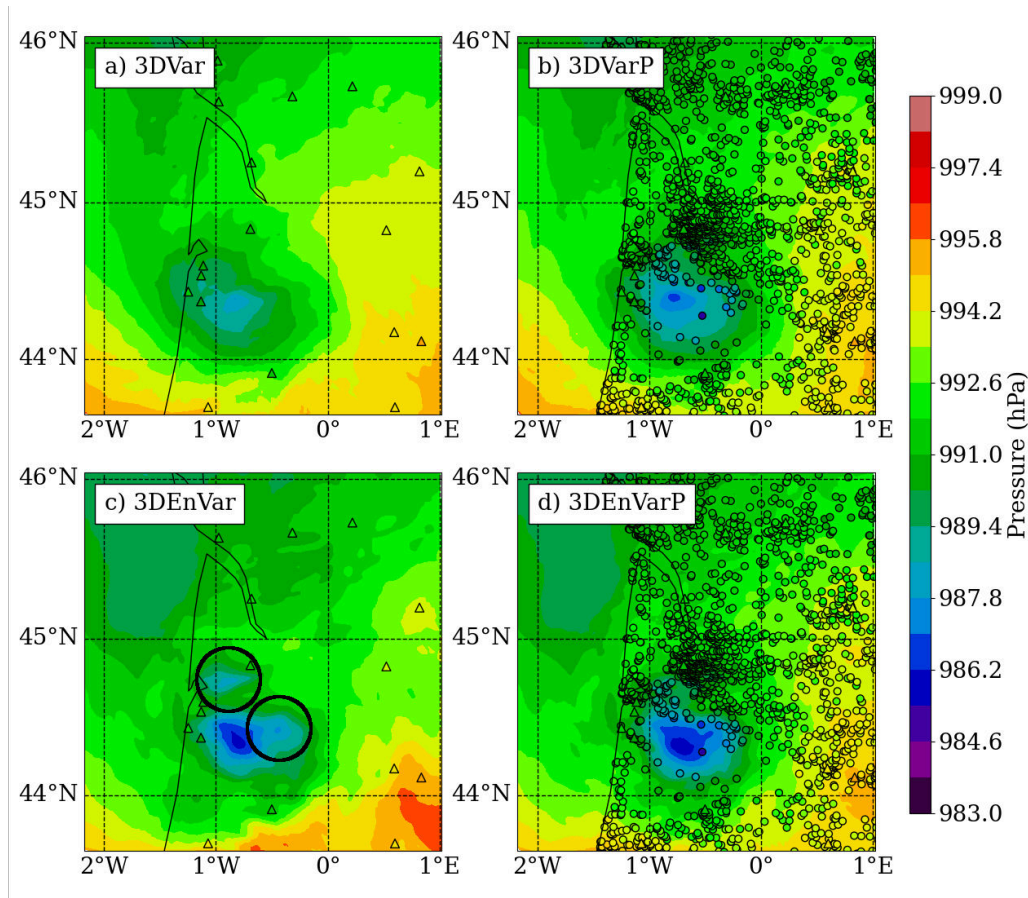


FIGURE 3.9 – Analyse de P_{mer} à 03:00 UTC de la dépression secondaire simulée par AROME le 3 novembre 2023. Les expériences sont cyclées et partent du réseau opérationnel de 00:00 UTC.

La figure 3.10 montre la prévision à 1 h issue des analyses précédentes. La dépression commence à se combler dans les quatre expériences. Dans l'expérience de référence 3DVar, la P_{mer} simulée est cohérente avec les observations SWS, avec un minimum à 990.0 hPa. Cependant, les observations PWS indiquent que le centre dépressionnaire est plus creux, avec une pression de 988.0 hPa, et plus à l'est. Dans l'expérience 3DVarP, le centre dépressionnaire reste plus creux, à 989.6 hPa, comme il l'était à l'analyse comparée à l'expérience 3DVar; ce qui est cohérent avec les observations PWS, même si faiblement en retard. Concernant les expériences 3DEnVar et 3DEnVarP, leur prévision à 1 h d'échéance provenant de l'analyse de 03:00 UTC sont similaires (figure 3.10c-d). Le gain apporté à l'analyse par les PWS n'est plus discernable. Le centre dépressionnaire est en

accord avec les observations PWS, avec un minimum prévu à 988.6 hPa. Mais à l'arrière de la dépression, au niveau de la côte atlantique, la pression ne remonte pas suffisamment par rapport aux observations, avec des différences de 1 à 2 hPa. Faire une analyse cohérente avec un système dense d'observations de surface n'est pas un critère suffisant pour obtenir la prévision la plus réaliste, c.-à-d. la plus cohérente avec les observations. Il convient de se questionner sur l'influence d'autres observations assimilées à des altitudes plus élevées que la surface, qui pourraient avoir influencé la surface aux premières échéances de prévisions en 3DVar. Conjointement, il est possible que les covariances d'erreur d'ébauche du schéma 3DVar ne permettent pas un changement suffisant des autres champs de surface pour maintenir la cohérence entre les champs de pression de surface et les observations de pression des PWS après analyse.

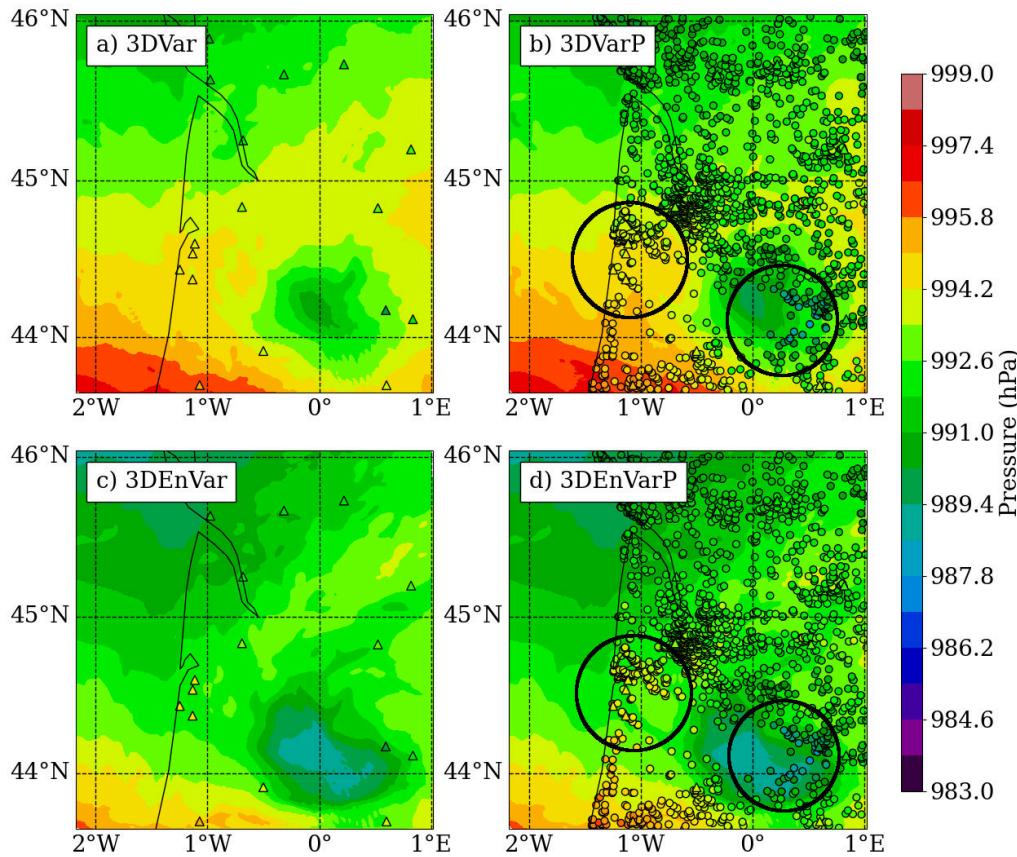


FIGURE 3.10 – Prévision à 1h issue de l'analyse de 03:00 UTC de la dépression secondaire simulée par AROME le 3 novembre 2023 à 04:00 UTC.

3.4.3 Mascaret atmosphérique - 18 juin 2022

Le mascaret atmosphérique est matérialisé par une discontinuité de pression au sol rectiligne coupant la côte normande, au nord-ouest de la France, avec une hausse de plus de 3 hPa en une dizaine de kilomètres (figure 3.12). La zone possède une faible densité d'observations standards de pression avec seulement 3 SWS. Ici, nous présentons l'analyse à 18:00 UTC des quatre expériences partant de la même ébauche à 17:00 UTC du cycle 3DVar opérationnel, afin de décrire l'influence des observations PWS sur l'analyse. Le mascaret est matérialisé par les deux rouleaux nuageux visibles sur l'image satellite (à

0.5 °W).

Les incréments d'analyse (figure 3.11) montrent les réglages et capacités des schémas d'assimilation. En 3DVar les incréments sont de grandes dimensions en comparaison avec ceux du 3DEnVar. On remarque que les incréments ne coïncident pas avec des observations de pression de SWS, ils proviennent d'autres observations grâce aux covariances d'erreur d'ébauche.

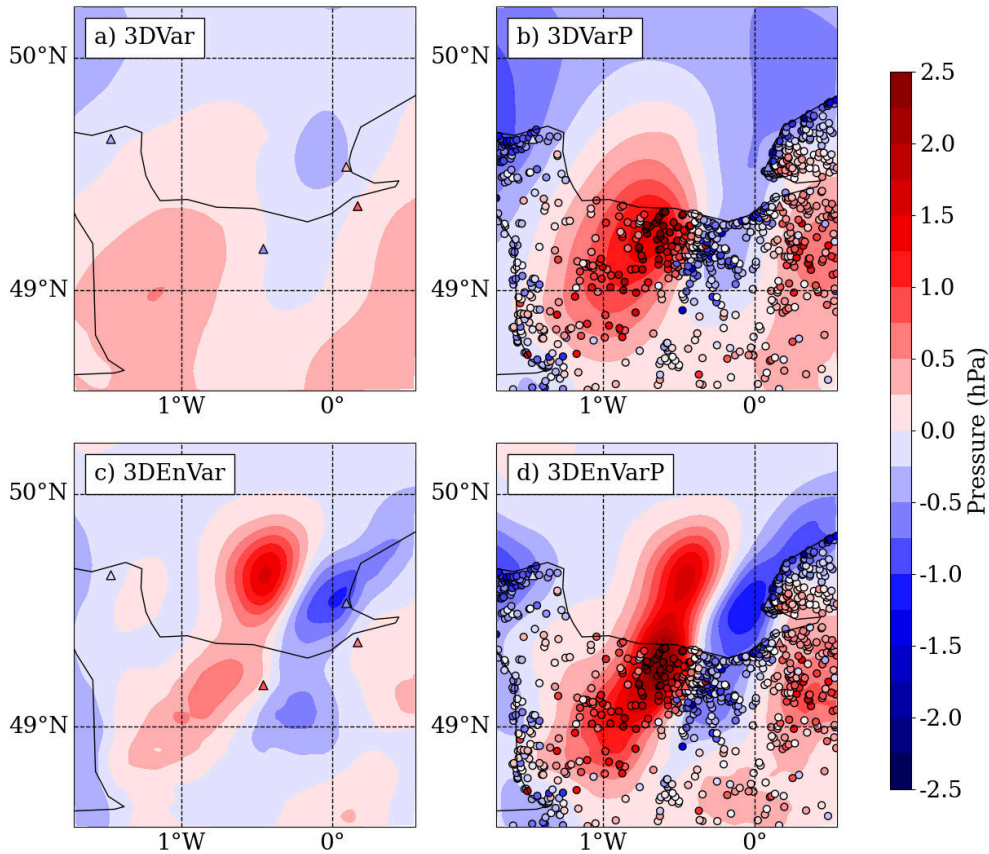


FIGURE 3.11 – Incrément d'analyse à 18:00 UTC du mascaret atmosphérique simulé par AROME le 18 juin 2022. Les quatre expériences partent de la même ébauche AROME à 17:00 UTC.

L'assimilation des observations de pression de PWS (figure 3.11b,d) renforce ces incréments de manière cohérente avec les expériences de contrôle (figure 3.11a, c). Un dipôle d'incréments positifs et négatifs se forme et est de différente taille selon les configurations : dans l'expérience 3DVarP l'incrément positif admet une extension zonale de 64 km, alors qu'en 3DEnVarP il est réduit à 39 km. De plus, en 3DVarP l'incrément positif maximum atteint 1.5 hPa, alors qu'en 3DEnVarP il monte à 2.4 hPa. L'incrément positif en 3DEnVarP se rapproche en taille et intensité de l'innovation des observations de PWS, qui présentent un maximum à 3.4 hPa. On remarque tout de même qu'il reste une différence notable en 3DEnVarP entre l'incrément maximum et l'innovation maximum, avec un écart de 1 hPa, lié aux influences respectives des observations et de l'ébauche dans la minimisation. Néanmoins, comme on l'a vu précédemment, la configuration en 3DEnVar donne peu de portée aux observations de surface, qui est ici compensé par le nombre d'observations PWS assimilées. Ce dipôle d'incrément indique que le mascaret simulé est en retard par rapport aux observations, et que différents systèmes d'observations (dont les PWS)

permettent d'ajuster l'ébauche.

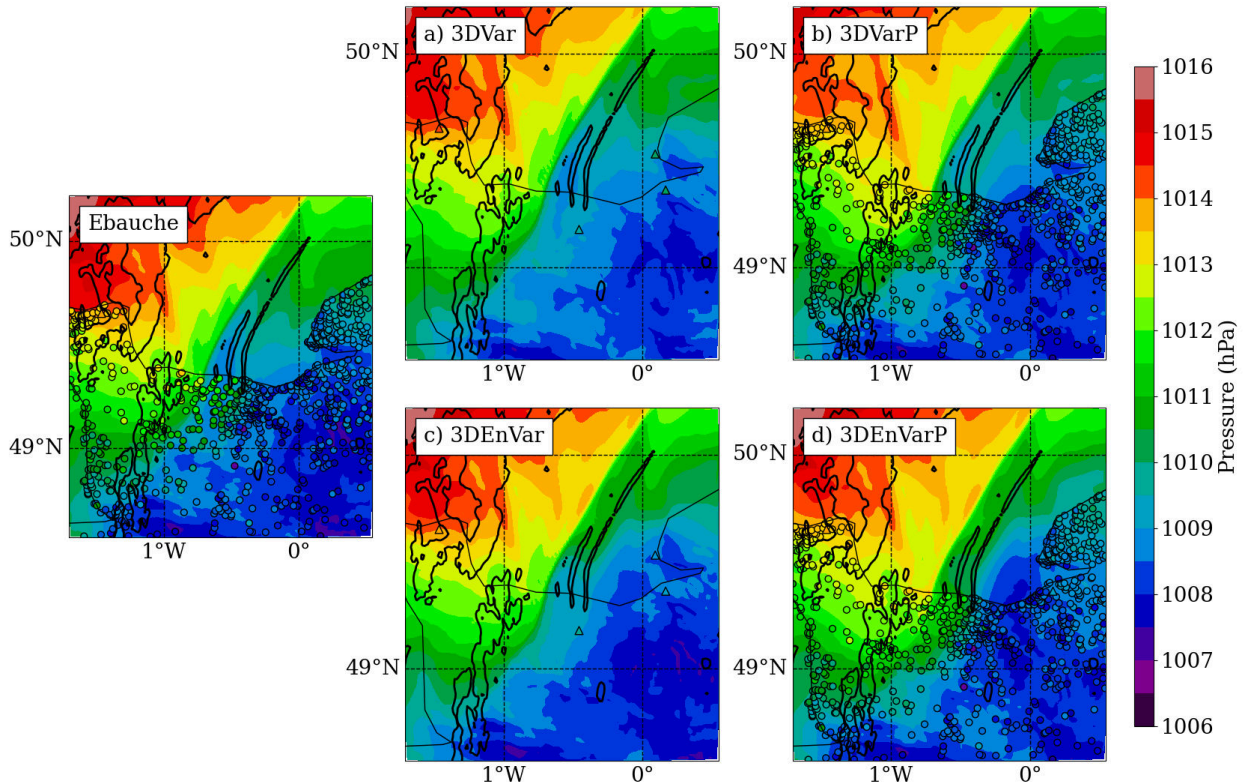


FIGURE 3.12 – Unique ébauche et analyses pour les quatre expériences en P_{mer} à 18:00 UTC du mascaret atmosphérique simulé par AROME le 18 juin 2022. Les réflectances du canal HRV du satellite MSG-3 égales à 43.1 % (110/255) sont représentées par des lignes noires et montrent les trois bandes nuageuses parallèles du mascaret.

La précédente description de l'incrément d'analyse permet de mieux comprendre l'analyse des expériences 3DVarP et 3DEnVarP (figure 3.12b, d). L'analyse en 3DVarP reste éloignée des observations de PWS, tandis que pour l'expérience 3DEnVarP, le schéma d'assimilation arrive à tirer parti de ces observations de fine échelle. On remarque néanmoins que la pression sur la terre, là où sont les observations PWS, augmente plus fortement que la pression en mer. Cette différence de pression entre la mer et la terre peut créer des discontinuités non-physiques qui mériteraient d'être examinées plus en détail. Au final deux gradients de pression sont notables sur l'expérience 3DEnVarP qui est la plus proche des observations PWS : un gradient à l'avant et un autre à l'arrière des trois bandes nuageuses du mascaret. Par la suite, le gradient de pression au sol référera à celui qui est le plus marqué, positionné à l'arrière des bandes nuageuses (le plus à l'est).

Pour comprendre la prévisibilité du phénomène, des prévisions sont produites pour 18:00 UTC à partir des analyses de 12:00 UTC, 15:00 UTC et 17:00 UTC des cycles d'assimilation 3DEnVar et 3DEnVarP (figure 3.13). Avec l'expérience 3DEnVar, le gradient de pression prévu à 18:00 UTC par le réseau de 12:00 UTC est en retard, puis il est mieux positionné avec le réseau de 15:00 UTC et de 17:00 UTC, mais celui-ci est moins marqué. Pour l'expérience 3DEnVarP le positionnement du gradient de pression ne change pas d'une prévision à l'autre, parallèle au mascaret observé par l'image satellitaire. Ici l'assimilation des observations de pression des PWS permet (i) de mieux positionner le gradient de pression 6 h avant l'événement et (ii) de réduire la volatilité d'une prévision à

l'autre. Pour avoir plus d'impact, il faudrait pouvoir assimiler les observations des PWS à une fréquence temporelle plus élevée comme le permettrait le schéma 4DEnVar.

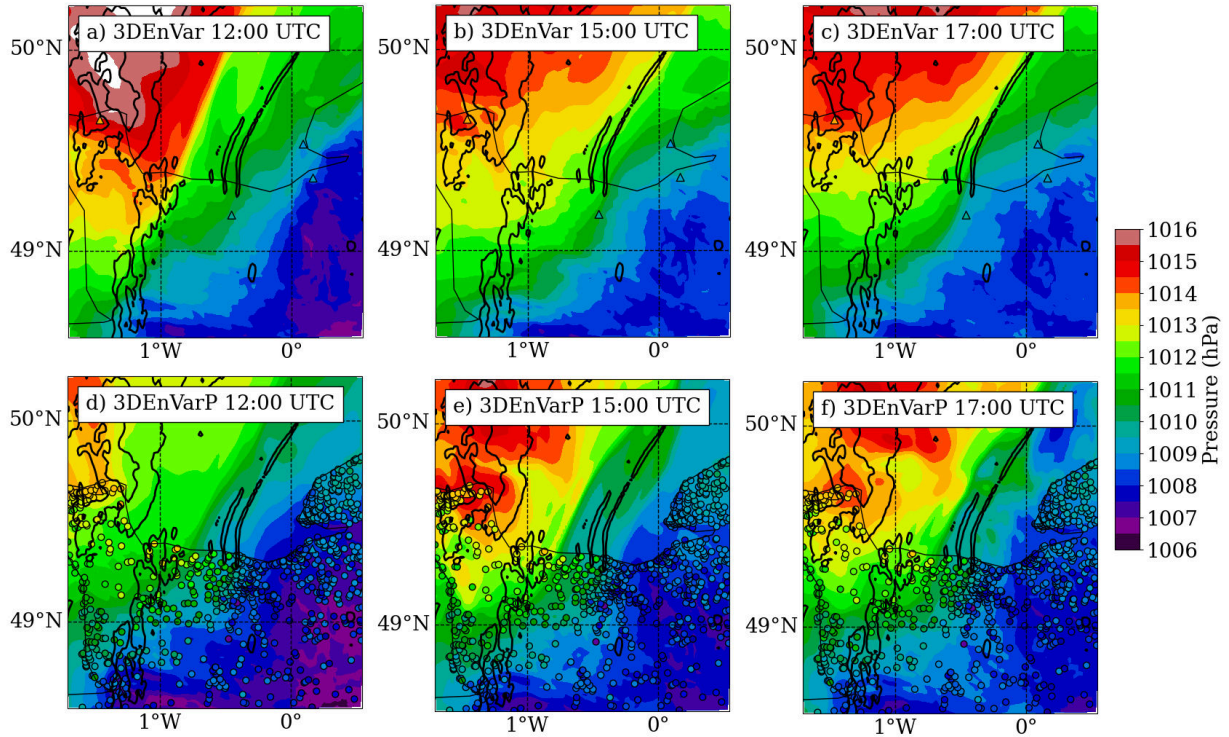


FIGURE 3.13 – Prévisions à 18:00 UTC, des expériences (a-c) 3DEnVar et (d-f) 3DEnVarP partant de l'analyse (a/d) six heures plus tôt à 12:00 UTC, (b/e) trois heures plus tôt à 15:00 UTC et (c/f) une heure plus tôt à 17:00 UTC. Les observations SWS et PWS sont valides à 18:00 UTC.

3.5 Synthèse et perspectives

Dans cette section, un traitement des observations de pression est proposé et leur apport a été étudié de manière objective, sur un mois d'étude, puis via des études de cas.

Une étude de sensibilité à l'erreur d'observation σ_o et à l'échelle d'écrémage a été proposée. Donner plus de poids aux observations de pression de PWS, en diminuant leur σ_o ou l'échelle d'écrémage, rapproche l'ébauche des observations de pression des SWS mais éloigne l'ébauche sur d'autres paramètres comme la température ou l'humidité relative, qui peut être dû à la matrice de covariance d'erreur d'ébauche. Autrement dit, il convient de trouver un juste poids à ces nouvelles observations. L'ajout de ces nouvelles observations n'augmente pas significativement le nombre d'itérations de la minimisation, ni le retour à l'équilibre des champs après l'analyse.

L'apport des observations de pression des PWS pour représenter des structures de fine échelle dans le champ de pression de surface à l'analyse a été montré pour les cas d'un MCS, d'un front, d'une dépression de méso-échelle et d'un mascaret atmosphérique. Cependant, l'influence de ces nouvelles observations sur la prévision à courte échéance est plus disparate. Pour le cas du MCS et du front, avec un cycle d'assimilation long de quelques jours à un mois, on remarque respectivement un bénéfice sur la prévision des précipitations sur 24 h de prévisions et sur la P_{mer} jusqu'à 9 h d'échéance. Pour le cas

d'un gradient de pression associé à un mascaret atmosphérique, l'expérience avec un cycle d'assimilation d'une demi-journée, permet de mieux positionner le gradient de pression de surface associé à l'événement à 6 h de l'échéance. Pour ces cas, assimiler les observations de pression PWS durant un cycle long d'assimilation, permet une meilleure prévision de ces événements. Pour le cas de la dépression de méso-échelle, avec un cycle d'assimilation partant trois heures avant l'événement, le gain concernant la prévision à 1 h d'échéance est moins marquée. Le phénomène se produisant le long de la côte, les observations PWS ne peuvent corriger le phénomène que sur peu de réseaux successifs. Cela pose la question de la prévisibilité d'événements locaux dont l'observation de leurs marqueurs n'intervient que peu de temps avant l'occurrence de l'événement (Madaus et Hakim, 2016).

Assimiler les données provenant d'un nouveau réseau d'observation de surface permet de questionner le réseau d'observation actuel et les réglages des configurations des systèmes d'assimilation développés. Dans la configuration actuelle, des investigations restent à mener concernant :

- Le monitoring des observations de pression des SWS, en fournissant une liste des stations SWS à étalonner, ou à recalculer verticalement.
- La sur-estimation problématique du poids donné à l'ébauche avec la matrice de covariance d'erreur d'ébauche du schéma d'assimilation 3DEnVar sur les niveaux les plus bas du modèle, et a fortiori la faible échelle de localisation horizontale donnée aux observations de surface.

La sous-estimation de la matrice d'erreur d'ébauche au sol, provient d'un défaut de sous-dispersion de l'ensemble de prévision (EDA) au sol. Cette sous-estimation est aussi rencontrée dans le modèle global du centre européen IFS. L'EDA y est sous-dispersif près de la surface, en particulier au-dessus de la neige (Ingleby *et al.*, 2024). Dans le modèle COSMO, la sous-dispersion de l'EDA est aussi relevée par Schraff *et al.* (2016) proposant plusieurs solutions : l'augmentation directe des covariances d'erreur d'ébauche, la relaxation aux perturbations initiales (*Relaxation to prior perturbations*, RTPP) ou la localisation adaptative.

Avant un éventuel passage à l'opérationnel, des tests sont nécessaires :

(i) Ici, une correction de biais *offline* a été proposé. L'utilisation d'une correction de biais variationnelle (VarBC) pourrait rendre le système de débiaisage *online*, autonome et réactif. Néanmoins, son utilisation demande de vérifier qu'il y a suffisamment d'observations d'ancrage et qu'elles sont de bonne qualité. En effet, les observations de pression des SWS, qui pourraient servir d'ancrage, se sont montrées biaisées par rapport aux ébauches du modèle, et nécessiterait a fortiori un contrôle plus restrictif de ces observations.

(ii) Pour améliorer la qualité du système d'observations PWS, un contrôle de qualité plus avancé pourrait être testé, utilisant l'information des observations voisines pour rejeter les observations anormales (Mandement et Caumont, 2020). Le temps de calcul nécessaire pour une telle méthode devra être optimisé pour pouvoir fournir à temps les observations contrôlées au système d'assimilation de données. Enfin, une « liste blanche » de stations dont la qualité a été contrôlée pourra être produite, limitant l'accès aux nouvelles observations installées a posteriori. Ces éléments doivent être discutés pour constituer un système de prétraitement des données efficient et rapide, indispensable pour leur assimilation opérationnelle.

(iii) Le réseau d'observation de pression des PWS, en raison de sa densité en zones urbaines, doit faire sa place parmi les autres réseaux d'observations. Un risque subsiste à donner trop de poids localement aux observations de pression. L'écrémage réalisé en sélectionnant aléatoirement une observation par maille d'1.3 km permet de réduire cet effet.

Néanmoins, une technique de *superobbing* permettrait d'utiliser l'ensemble des observations disponibles. Le *superobbing* consiste à combiner plusieurs observations voisines pour n'en former qu'une, par une moyenne (Waller *et al.*, 2019). Sur le même principe que le *superobbing*, une méthode de calcul des poids adaptative pourrait être testée, en prescrivant une erreur d'observation plus grande aux observations situées dans des zones à forte densité spatiale d'observations. Par exemple, si une maille de taille à définir contient n observations, augmenter leurs erreurs d'observation à $\sigma_o\sqrt{n}$ de manière à ce qu'elles aient le même poids qu'une seule observation dans l'analyse, permettrait de pouvoir garder toutes les observations tout en évitant de choquer le modèle à l'analyse.

D'autres observations provenant des PWS peuvent être assimilées dans AROME, comme celles de température et d'humidité relative, qui sont l'objet du prochain chapitre.

Chapitre 4

Assimilation d'observations de température et d'humidité relative provenant des stations personnelles dans AROME-France

Sommaire

4.1	Résumé de l'article	101
4.2	Article : <i>Assimilation of temperature and relative humidity observations from personal weather stations in AROME-France</i>	102
4.2.1	Introduction	103
4.2.2	Datasets and methods	105
4.2.3	Assimilation experiments design	113
4.2.4	Results of the assimilation experiments	118
4.2.5	Results on case studies : heatwave and fog events	121
4.2.6	Discussion and conclusion	125
4.3	Complément de l'article	135
4.3.1	Filtres des observations de surface	135
4.3.2	Cas d'étude : ICU du 17 juin 2023	137
4.3.3	Cas d'étude : ligne orageuse le 8 septembre 2021	137
4.4	Synthèse et perspectives	141

4.1 Résumé de l'article

Les réseaux de stations météorologiques personnelles (PWS) sont capables de fournir des observations de la température de surface et de l'humidité relative à une densité spatiale auparavant inaccessible aux stations météorologiques standards (SWS), tout comme cela avait été noté pour les observations de pression. L'analyse du modèle AROME-France est sous-contrainte en particulier près de la surface, dans la couche limite atmosphérique, où apparaissent des phénomènes associés à de forts gradients de température et d'humidité relative. Cette étude présente l'impact de l'assimilation d'observations de la température et de l'humidité relative des PWS dans AROME-France pendant des expériences d'assimilation/prévision d'un mois. Les observations étant sujettes à des erreurs d'instrumentation et de représentativité (figure 4.1), un prétraitement particulier leur est appliqué afin d'utiliser des données les plus fiables possibles lors de l'assimilation de données. Le prétraitement des observations PWS comprend la correction du biais, le contrôle de la qualité et l'écrémage. Différentes méthodes sont testées provenant de la littérature : d'une correction diurne, à une correction prenant en compte le rayonnement solaire ; du contrôle de qualité QC-Titan (Nipen *et al.*, 2020), au contrôle de qualité QC-MC (Mandement et Caumont, 2020). Après le prétraitement, près de 4700 observations de température et

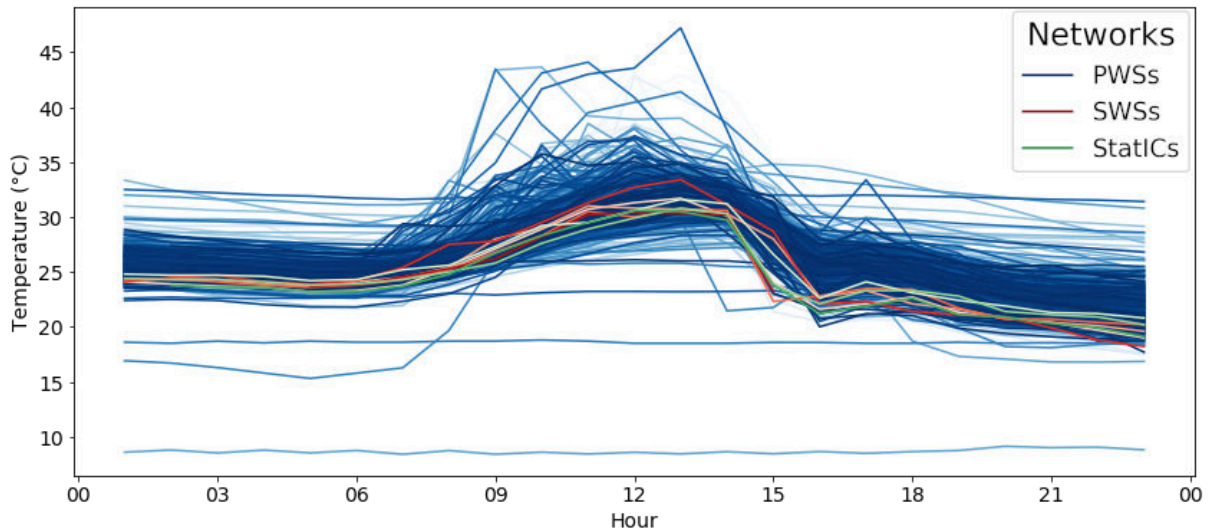


FIGURE 4.1 – Séries temporelles d’observations de température dans le sud-ouest de la France le 12 août 2020. Trois réseaux d’observations sont présents : 121 PWS (palette bleue), 5 SWS (palette rouge) et 3 StatIC (palette verte). Un orage passe à 16:00 UTC sur la zone, avec une baisse de 5°C de la température en 1h. Plusieurs effets sont visibles : la majorité des observations PWS ne correspond pas aux observations de référence, avec une température plus élevée, surtout la nuit ; certaines observations PWS demeurent constantes tout au long de la journée, probablement des températures intérieures ; enfin, certaines observations PWS affichent des températures nettement plus élevées que les références en journée, avec des écarts pouvant atteindre 10 à 15°C par rapport aux SWS.

4200 observations d’humidité relative provenant des PWS sont assimilées par heure, ce qui représente respectivement 3 et 6 fois plus d’observations que celles provenant des SWS. L’assimilation des observations de température pré-traitées des PWS avec le schéma d’assimilation 3DVar réduit significativement l’écart quadratique moyen en température entre les observations des SWS et la prévision à 1 h. Cette réduction est encore observée jusqu’à 3 h de prévision. Des résultats similaires sont obtenus pour l’assimilation des observations d’humidité relative PWS pré-traitées. Une dégradation de la prévision de surface du modèle à 1 h est constatée lorsque les observations PWS sont assimilées à la fois dans l’atmosphère avec le schéma d’assimilation 3DVar et aussi à la surface avec le schéma d’Interpolation Optimale. Ces dégradations montrent l’importance du schéma d’assimilation, et de son réglage, lors de l’assimilation de nouvelles observations. Les études de cas d’une vague de chaleur et d’un épisode de brouillard démontrent la capacité de l’assimilation des observations PWS à mieux représenter les structures météorologiques de méso-échelle à l’analyse.

4.2 Article : *Assimilation of temperature and relative humidity observations from personal weather stations in AROME-France*

Assimilation of temperature and relative humidity observations from personal weather stations in AROME-France

Alan Demortier^a, Marc Mandement^a, Vivien Pourret^a, and Olivier Caumont^{a,b}

^aCNRM, Université de Toulouse, Météo-France, CNRS, Toulouse, France

^bMétéo-France, Direction des opérations pour la prévision, Toulouse, France

Abstract. Personal weather station (PWS) networks owned by citizens now provide near-surface observations at a spatial density unattainable with standard weather stations (SWSs) deployed by national meteorological services. This article aims to assess the benefits of assimilating PWS observations of screen-level temperature and relative humidity in the AROME-France model, in the same framework of experiments carried out to assimilate PWS observations of surface pressure in a previous work. Several methods for pre-processing these observations, in addition to the usual data assimilation (DA) screening, are evaluated and selected. After pre-processing, nearly 4700 temperature and 4200 relative humidity PWS observations are assimilated per hour, representing 3 and 6 times more than SWS observations, respectively. Separate assimilation of each variable in the atmosphere with the 3DVar DA scheme significantly reduces the root-mean-square deviation between SWS observations and forecasts of the assimilated variable at 2 m height above ground level up to 3 h range. Improvements to the near-surface temperature and relative humidity fields analysed are shown for a sea breeze case during a heatwave and a fog episode. However, degradation of short-range forecasts are found when PWS observations are assimilated with the current operational 3DVar DA scheme in the atmosphere or jointly in the atmosphere and at the surface with 3DVar and Optimal interpolation DA schemes. These results demonstrate that the benefit of assimilating temperature and relative humidity PWS observations can be highly dependent on the DA schemes and settings employed.

1 Introduction

The increase in spatial and temporal resolution of regional numerical weather prediction (NWP) models requires their analyses to be initialized by spatially and temporally dense observations, to represent meteorological phenomena on increasingly finer scales (Gustafsson et al., 2018). At Météo-France, the operational regional NWP system designed to forecast up to 51 h lead time such phenomena is AROME-France (Seity et al., 2011; Brousseau et al., 2016). AROME-France currently uses a three-dimensional variational (3DVar) DA scheme. This scheme is planned to be replaced by a three-dimensional ensemble variational (3DVar) DA scheme, with an up-to-date background error covariance matrix, improving the spread of the information from the observations (Montmerle et al., 2018; Michel and Brousseau, 2021).

Phenomena whose representation in the analyses needs to be improved are meso- γ to meso- β scale phenomena (2 to 200 km, Orlandi, 1975) such as thunderstorms, breezes or fog (Stull, 1988) which cause substantial thermodynamic changes, particularly in the atmospheric boundary layer (ABL). New observations are emerging to improve analyses in the ABL, from ground-

based microwave radiometers (Caumont et al., 2016; Bell et al., 2022; Vural et al., 2024), water vapour lidars (Flamant et al., 2021), or aircrafts (Pourret et al., 2022). Even if these instruments provide dense vertical observations, their horizontal density remains low. Work is underway to exploit horizontally dense satellite observations near the surface, such as radiances from the Meteosat Spinning enhanced visible and infrared imager (SEVIRI) (Sassi et al., 2019), which are currently assimilated with a spacing of the order of 30 km and are still difficult to assimilate at full resolution.

The growth of crowdsourced data, i.e. the data obtained from a range of sensors belonging to the public and shared via the Internet, could play a significant role in improving near-surface analyses of NWP models, in particular data from personal weather stations (PWSs) (Hintz et al., 2019; Coney et al., 2022). In order to obtain observations comparable with standard weather station (SWS) observations, PWS observations should be properly pre-processed, i.e. bias-corrected or quality-controlled or both, the bias designating the deviation of a PWS observation time series from a reference time series (other observations or a modelled field). Indeed, PWSs provide observations that may not comply with World Meteorological Organization (WMO) methods: PWS sensors have heterogeneous environments of siting, sometimes unsuitable shelters, can be positioned at various heights above ground level (AGL), and can suffer from time lag compared to sensors following WMO requirements (Bell et al., 2015; Varentsov et al., 2020; Fenner et al., 2021). After pre-processing, observational studies showed that PWS observations, when combined with SWS observations, are able to describe near-surface thermodynamic variations associated with mesoscale phenomena such as thunderstorms partially visible with SWS observations only (Clark et al., 2018; Mandement and Caumont, 2020).

Various bias correction (BC) methods have been developed. The bias can be computed over a short period (e.g. 6 h for Clark et al., 2018 or approximately 24 h for Mandement and Caumont, 2020). To account for the effects of an inappropriate sheltering or siting, the bias can be computed as a linear function of solar incident radiation, or by a more sophisticated decomposition, e.g. with multilinear functions (Sgoff et al., 2022), generalized additive mixed models (GAMMs) (Cornes et al., 2019), or with machine learning methods (Beele et al., 2022; Marquès, 2023).

Regarding quality control (QC) methods, CrowdQC (Meier et al., 2017) and CrowdQC+ (Fenner et al., 2021) were designed to remove PWS temperature observations deemed erroneous based on their deviation from neighbouring observations. CrowdQC was used in numerous studies of urban temperatures (Feichtinger et al., 2020; Venter et al., 2020; Potgieter et al., 2021; Zumwald et al., 2021). Another QC for PWS temperature observations, used operationally in post-processing algorithms by MET Norway, is called Titan (QC-Titan hereafter, Båserud et al., 2020; Nipen et al., 2020). CrowdQC+ and QC-Titan are comparable in a way as they both use the information from their neighbours to remove inconsistent observations, e.g. the m5 procedure of the CrowdQC+ is comparable to the spatial buddy check of QC-Titan (Fenner et al., 2021). In addition, as CrowdQC+ uses local climate zones, it is challenging to use on a scale larger than a city. For both temperature and relative humidity PWS observations, Mandement and Caumont (2020) have developed a QC (QC-MC hereafter) using an adaptive rejection threshold based on a comparison to interpolated SWS observations. Both QC-Titan and QC-MC are tested in this study.

Once PWS observations have been pre-processed, the opportunity of their assimilation arises. Sgoff et al. (2022) carried out experiments assimilating pre-processed (bias-corrected and quality-controlled) PWS temperature and relative humidity

observations – using the flow-dependent local ensemble transform Kalman filter (LETKF) DA scheme of the Icosahedral non-hydrostatic model with 2 km resolution (ICON-D2). They showed that the assimilation of PWS observations is beneficial; however, no experiments assimilating at the same time both SWS and PWS observations were conducted. On the other hand, Demortier et al. (2024, hereafter D24) carried out experiments assimilating simultaneously pre-processed PWS and SWS observations of surface pressure with both 3DVar and 3DEnVar DA schemes in AROME-France. Statistically significant improvements of mean sea level pressure forecasts up to 9 h range were found with the 3DEnVar scheme.

This article extends Sgoff et al. (2022) and D24 works by simultaneously assimilating pre-processed PWS and SWS observations of screen-level temperature and relative humidity, and address the following questions. What are the most effective pre-processing methods for assimilating PWS data? Are AROME-France’s analyses and forecasts improved by assimilating pre-processed PWS temperature and relative humidity observations? What impact does the choice of the DA scheme have?

The remainder of this article is organized as follows. Section 2 describes the AROME-France NWP system and its DA schemes in the atmosphere and at the surface, as well as the observations used and the pre-processing methods of these observations. Section 3 describes the assimilation experiments. Objective results of these experiments are given in Sect. 4 and results for case studies are detailed in Sect. 5. Finally, results are summarized and discussed in Sect. 6.

2 Datasets and methods

The assimilation experiments run from 6 September to 5 October 2021, as in D24. This one-month study period encompasses a diverse range of meteorological events. It includes two periods of anticyclonic conditions, during which a heatwave (from 2 to 8 September) and fog (from 22 to 24 September) were observed. It also includes disturbed weather episodes such as squall lines on 8 September in the south-west of France, a mesoscale convective system in the south-east of France on 14 September, and heavy precipitation events from 2 to 4 October.

2.1 AROME-France NWP system

AROME-France is the limited area NWP model developed by Météo-France, operational since December 2008 (Seity et al., 2011; Brousseau et al., 2016). AROME-France is a spectral model coupled to the global NWP system ARPEGE (Courtier et al., 1998). AROME-France (AROME hereafter) has a 1.3 km grid spacing on the horizontal and has 90 vertical levels in the atmosphere ranging from 5 m height AGL (lowest atmospheric model level) up to 10 hPa.

AROME uses a mass-flux shallow convection scheme (Pergaud et al., 2009). Its associated surface scheme is SURFEX (Masson et al., 2013). AROME and SURFEX exchange flux and near-surface variables such as temperature, humidity, and wind. Surface variables are computed over tiles representing four types of surfaces: nature, town, inland water, and ocean. The surface layers have a dedicated assimilation system, which is described at Sect. 2.1.2.

AROME has a 1 h DA cycle for the atmosphere, and a 3 h DA cycle for the surface. From the analysis, the experiments in this study produce 24 h forecasts at 00:00, 06:00, 12:00 and 18:00 UTC.

2.1.1 AROME atmospheric DA system

The atmospheric DA procedure is composed of two main steps, which are the screening, which is a series of QC checks, and the minimization of the cost function used for variational DA (see D24). During both steps, modelled variables are compared to observed variables by an observation operator. AROME has 13 prognostic variables and 5 control variables, which include temperature and specific humidity. Relative humidity is a diagnostic variable computed from specific humidity, temperature, and pressure. The formula used in AROME is given in Appendix A. The model equivalents of observed screen-level temperature (T) and relative humidity (RH) are 2 m height AGL temperature (T_{2m}) and relative humidity (RH_{2m}), described by Vasiljevic et al. (1992) and ECMWF (2023). Computation of T_{2m} uses the temperature from the lowest atmospheric model level and the surface temperature, based on the Monin–Obukhov theory – with separated solutions for stable and unstable conditions (Cardinali et al., 1994). RH_{2m} is computed using only the lowest level fields of the atmospheric model.

The difference between the observation and the model equivalent from the background (1 h forecast starting from the previous analysis), i.e., the observation minus the background, is further called OmB. In the same way, the difference between the observation and its corresponding value in the analysis is further called OmA.

One of the main QC checks of the screening reject observations verifying:

$$\|\text{OmB}\| > \frac{\alpha}{\sigma_{\text{coef}}} \sqrt{\sigma_o^2 + \sigma_b^2} \quad (1)$$

with $\alpha = 4$ and $\sigma_{\text{coef}} = 0.9$, two coefficients; σ_o is the standard deviation of observation errors set to 1.4 °C and 10%, for screen-level T and RH observations; σ_b is the standard deviation of background errors. As an example, for T_{2m} , during the first hour of the assimilation experiments (i.e., the 6 September 2021 at 01:00 UTC), the average σ_b equals 0.5 °C and the rejection threshold is up to 6.5 °C. For RH_{2m} , the average σ_b equals 10% and the rejection threshold is up to 70%. These rejection thresholds are high because current near-surface observations assimilated (SWS observations) are considered as anchor observations for the model, and screening has been designed to keep as much of these observations as possible.

All active observations, i.e. those that have passed the screening, are combined with the background to produce the analysed state of the atmosphere. AROME currently uses a 3DVar DA scheme to constitute the atmospheric analysis fields. The only difference between the 3DVar and the 3DEnVar DA schemes comes from the background error covariance matrix \mathbf{B} which is static (prescribed) for the 3DVar and dynamic for the 3DEnVar, estimated at each hour from the AROME Ensemble Data Assimilation system (EDA). To illustrate the difference between the two schemes, the increments (i.e. the analysis minus the background) at the lowest atmospheric model level when single screen-level T and RH observations are assimilated in two idealized experiments are shown in Fig. 1. The propagation of the information from an observation is isotropic with the 3DVar DA scheme (Fig. 1a and c), whereas with the 3DEnVar DA scheme it is anisotropic and localized within 25 km around the observation (Fig. 1b and d). At the observation location, the values of the increments differ. From an identical OmB of 3.4 °C in both experiments, the increment at the observation location reaches 1 °C with the 3DVar DA scheme, and is reduced to 0.4 °C with the 3DEnVar DA scheme. Similarly, the assimilation of a single RH_{2m} observation results in an increment of 3.3% (resp. 1.1%) with the 3DVar (resp. 3DEnVar) DA scheme. The relative observation weight is diminished from the 3DVar

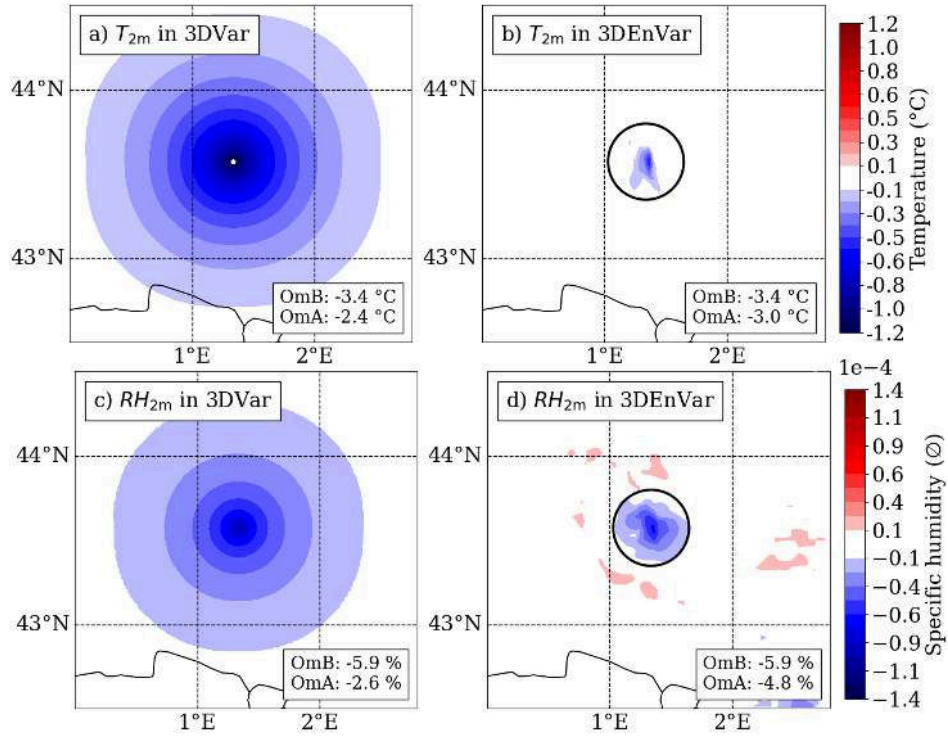


Figure 1. Increments at the lowest atmospheric model level of (a, b) temperature and (c, d) specific humidity, resulting from the assimilation of screen-level T and RH observations, respectively, at Blagnac SWS on 6 September 2021 at 01:00 UTC with the (a, c) 3DVar and (b, d) 3DEnVar DA schemes. The black circle indicates a 25 km distance around the observation location.

125 to the 3DEnVar DA scheme due to lower σ_b on average during the study period for temperature observations. Further details on the 3DVar and the 3DEnVar DA schemes configurations are provided by D24.

2.1.2 AROME surface DA system

AROME has a dedicated surface DA scheme for the initialization of the soil prognostic variables, which consist of two ground temperatures and two ground water contents (Giard and Bazile, 2000). The surface DA scheme has its own screening, which
 130 uses the same formulation as the atmospheric one (Eq. 1) with, for T_{2m} (resp. RH_{2m}), α set to 5 (resp. 2.5), σ_{coef} set to 1, σ_o set to 1.3 $^{\circ}\text{C}$ (resp. 10 %) and σ_b set to 1.6 $^{\circ}\text{C}$ (resp. 10 %).

The surface DA scheme is based on atmospheric increments of T_{2m} and RH_{2m} . First, T_{2m} and RH_{2m} are computed with the observation operator described in Sect. 2.1.1. Then, a univariate optimal interpolation (OI) scheme combines screen-level observations with T_{2m} and RH_{2m} background, using isotropic structure functions. The characteristic distance D used in the OI
 135 scheme is set to 100 km for T_{2m} and RH_{2m} observations, which is a compromise between the resolution of the NWP model,

the scale of the weather phenomenon represented and the density of the weather stations. Then T_{2m} and RH_{2m} increments are used to correct the soil variables with linear interpolations. More details on the surface DA are given by Sassi et al. (2019).

2.1.3 A posteriori diagnostics

140 A number of diagnostic tools exist to understand a posteriori the way in which the DA schemes use the observations. Both the Desroziers diagnostic of observation error and the spatial Desroziers diagnostic (also used in D24) are used (Desroziers et al., 2005). The spatial Desroziers diagnostic is computed to define a minimum thinning length λ , from which the correlated observation errors are found to be significantly low. Then, the observation network is thinned by selecting one random observation per mesh from a horizontally-spaced λ grid (D24).

2.2 Screen-level observations of temperature and relative humidity

145 A distinction is made here between standard weather stations (SWSs) and personal weather stations (PWSs): the former are bought, operated and maintained by national meteorological and hydrological services (NMHS), while the latter are managed by private individuals or third-party organizations. Behind this distinction, it is important to keep in mind that citizens or third-party organizations can acquire stations identical to NMHSs, just as NMHSs can use low-cost stations dedicated to citizens.

2.2.1 Standard weather station (SWS) observations

150 This study uses only SWS observations assimilated in AROME, which constitute a subset of all SWS observations. These observations originate from three types of reports: manual land SYNOP, automatic land SYNOP and the French RADOME (Tardieu and Leroy, 2003). The number of temperature (resp. relative humidity) observations from SWSs is 2440 (resp. 1600) in average per hour in AROME. Only 65 % of SWS temperature observations entering the AROME DA system are located in France. This is only 48 % for relative humidity. Over France, there are more than twice as many observations of temperature
155 as there are observations of relative humidity, and both are fairly evenly distributed (Fig. 2). Regarding the height of observation, air temperature and relative humidity should be and are generally measured between 1.25 and 2 m height AGL (World Meteorological Organization, 2023), the rule being 1.5 m in France, often a little higher up in mountainous areas with heavy snowfall. For SWS wind speed observations used in Sect. 4, wind speed is generally measured at 10 m height AGL.

2.2.2 Personal weather station (PWS) observations

160 Three networks of PWSs are used in this study: Netatmo PWSs, StatIC PWSs and Toulouse Métropole PWSs. In contrast to SWS observations, PWS observations are not currently assimilated operationally in AROME.

The Netatmo PWS network is the largest available in near-real time in France (Mandement and Caumont, 2020). As they are owned by citizens, PWSs are unevenly distributed in France (Fig. 2). The outdoor module of the Netatmo PWS contains temperature and relative humidity sensors. With appropriate sheltering, these sensors have a median and a 95 % range of departures
165 from a reference sensor of about $0^\circ\text{C} \pm 0.9^\circ\text{C}$ in temperature and $3\% \pm 7\%$ in relative humidity, showing their correct

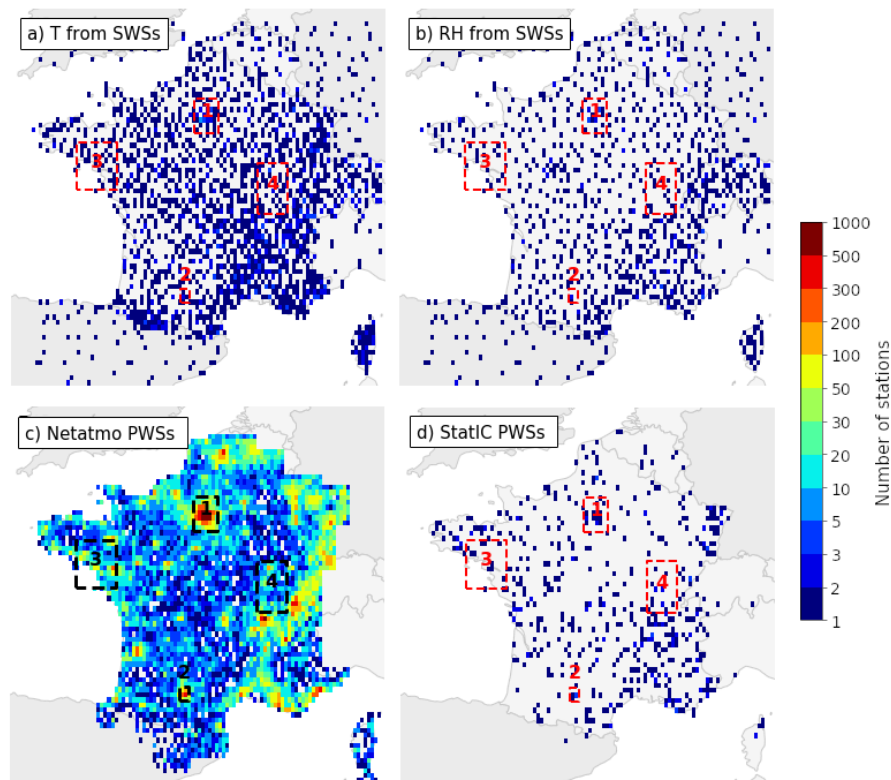


Figure 2. Number of (a) SWSs measuring temperature, (b) SWSs measuring relative humidity, (c) Netatmo PWSs and (d) StatIC PWSs providing at least one observation during the one-month study period over France. Observation counts are binned into approximately $0.15 \times 0.1^\circ$ bins. The dashed rectangles delimit the 4 domains used in the cases studied: (1) Paris, (2) Toulouse, (3) along the Atlantic coast for the heatwave case, and (4) the Saône Valley for the fog case.

intrinsic quality (Mandement and Caumont, 2020). However, the stations as they are sold are not sheltered according to WMO recommendations: short-wave and long-wave radiation affect T and RH , causing departures to sheltered reference sensors (Büchau, 2018). For temperature, these departures are generally overestimations of warm temperatures and underestimations of cold temperatures. These departures are complex depending on the location of the station in relation to its close environment (e.g. wall, balcony, garden; Varentsov et al., 2020), and its actual height AGL.

The StatIC PWS network from the Infoclimat association gathers 880 stations in 2023 (Garcelon et al., 2023), of which almost 600 stations are freely downloadable over the period via their open data portal. It gathers Davis Instruments Vantage Pro2 and Vantage Vue PWSs, and also sensors from Dragino or Talkpool companies in shelters, owned by citizens or by the association itself. To join the network, each station should follow the WMO recommendations adapted by Météo-France (Leroy, 2014), wherever possible: T and RH observation in a normalized shelter, in a clear environment and at 1.5 m height AGL. However, in enclosed urban environments (with low air circulation), the recommended height is 1.5 m above the roof,

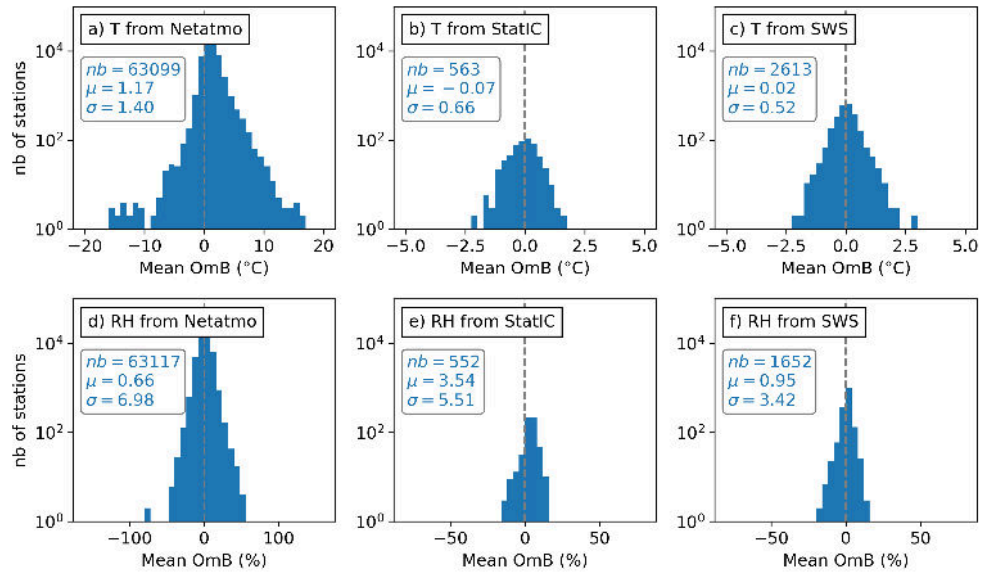


Figure 3. Histograms of average OmB over the one-month study period for (a, d) Netatmo PWSs, (b, e) StatIC PWSs, and (c, f) SWSs observations of (a–c) temperature and (d–f) relative humidity. Boxes indicate nb the number of OmB time series, μ the average OmB for all time series, and σ the standard deviation of average OmB for all time series.

i.e. various heights AGL depending on the height of the roof. StatIC PWS are set up in meteorologically interesting locations such as La Chapelle-en-Vercors which is a cold hole, or areas not well covered by SWS. The accuracy given for the Davis Instruments Vantage Pro2 or Vue PWS is $0.3\text{ }^{\circ}\text{C}$ for temperature and 2% for relative humidity (Garcelon et al., 2023).

180 Toulouse Métropole has deployed Davis Vantage Pro2 PWSs around the city of Toulouse for the study of the urban heat island (Dumas et al., 2021). During the study period, 33 PWSs are available.

Figure 3 shows the average OmB distributions for the two largest PWS networks and the SWS network. For temperature over the study period, there are only one fifth as many StatIC PWS observations as SWS observations, while there are about 25 times as many Netatmo PWS observations. However, the heterogeneous siting of the PWSs, seen by the OmB standard deviation, increases from SWSs to StatIC PWSs to Netatmo PWSs. The results are similar for relative humidity.

In this study, the only PWS observations assimilated are from Netatmo PWS, hereafter referred to as PWS. StatIC and Toulouse Métropole PWS observations are used to evaluate pre-processing methods of PWS observations of temperature. For the evaluation, Netatmo PWS observations are interpolated at the location of StatIC and Toulouse Métropole PWSs, providing an estimate which is compared to the StatIC and Toulouse Métropole PWS observations. The linear interpolation method takes into account the vertical profile of temperature. The use of independent (i.e. non-assimilated) PWS observations allows to verify the absence of bias in the model.

2.3 Satellite observations

Observations of top of the atmosphere bidirectional reflectance from the High resolution visible (HRV) channel from the SEVIRI radiometer, onboard the Meteosat second generation 3 (MSG-3) satellite (also called Meteosat-10) in Rapid Scanning Service (RSS) at 1 km horizontal resolution every 5 min are used in Sect. 5.2 to locate clouds. The HRV channel is a broadband channel which is sensitive to wavelengths between 0.4 and 1.1 μm (Schmetz et al., 2002). Reflectance observations from EUMETSAT are corrected depending on the solar angle following the method of Li and Shibata (2006) applied to visible channels in Météo-France operational products (Derrien et al., 2010).

An estimate of the global solar radiation at each PWS location has been derived from the spatialized global solar radiation product made by Météo-France in near-real time. This product uses both in situ observations and surface solar irradiance from MSG satellites (Eumetsat, 2017). Nighttime values are set to zero. The hourly global solar irradiance is cumulated over each hour and is used as a predictor for one of the bias corrections of the PWS observations.

2.4 Pre-processing methods

In the article, PWS observations used are interpolated to round hours and PWSs with identical coordinates are removed, as in the step preparation of D24. Hereafter, these observations, prepared but not pre-processed, are referred to as raw PWS observations.

2.4.1 Bias correction (BC) methods

The bias correction (BC) ensure that possible biases (i.e. a systematic departure to a reference) in PWS observation time series do not propagate to the model. In this study, raw PWS observations of screen-level temperature and relative humidity have been monitored, i.e. $T_{2\text{m}}$ and $RH_{2\text{m}}$ OmB time series are computed, in the operational AROME 3DVar configuration, which is called the Monitor experiment (as in Sgoff et al., 2022 and D24). Four types of BC methods that can be applied in real-time, because they use past OmB statistics, are tested. The first two use OmB time series from the last hours only, whereas the last two use longer OmB time series. These methods are named as follows:

- BC-D removes the average OmB of the last 24 h (average of 24 hourly OmBs, D for daily), in a manner close to Mandement and Caumont (2020).
- BC-M removes the average OmB of the last 30 d (M for monthly).
- BC-S uses the BC method developed by Sgoff et al. (2022); S for Sgoff. The bias is computed from five predictors composed of trigonometric basis functions from the hour of the measured observations with 2 d inertia.
- BC-R uses a random forest method trained with the hour of the day and cumulative hourly global solar irradiance predictors from the last 30 d, close to the BC used by Beele et al. (2022), but with fewer predictors (R for radiation).

2.4.2 Quality control (QC) methods

Quality control (QC) methods are used in this study to remove PWS observations judged erroneous, because the current screening for near-surface observations has not been designed to deal with observations with OmB statistics so different from those of SWS (Sect. 2.2.2).

225 Titan or titanlib (QC-Titan, Båserud et al., 2020), is an open-source library composed of various checks or tests. In this article, as in Nipen et al. (2020), Titan-QC designates the combination of three main spatial tests which are the buddy check, the spatial consistency test (SCT), and the isolation test. The buddy check consists in removing observations if their deviation from the average is more than twice the standard deviation of the observations in the neighbourhood within a 15 km radius. The SCT consists in an iterative cross-validation procedure. For each observation, the SCT adjusts a vertical profile, and computes
230 an estimation of a true observation, and the standard deviation of the observations in the neighbourhood. The observation is removed if the ratio of the squared deviation from the observation estimation divided by the standard deviation is greater than 4 (or 8 for negative values). The process is repeated until no observations are removed. The isolation test removes observations which have less than 5 stations within a 15 km radius and 200 m elevation difference.

QC-MC (Mandement and Caumont, 2020) has been adapted for use on a larger scale, throughout France, even though it
235 was designed for the scale of a French region: computations are done simultaneously for 11 selected climatological areas. The second main adaptation is the use, for each PWS observation, of OmB statistics from AROME instead of comparisons between this PWS observation and an estimate derived from neighbouring SWS observations. In QC-MC, every hour, in every climatological area, for each PWS, the rms of PWS OmBs are computed on the last 6 h, to separate observations that differ from the background at a given time due to physical reasons (e.g. a phenomenon not well positioned in the background) to the
240 observations that differ all the time from the background. PWS having rms OmB exceeding an adaptive threshold, computed according to (Mandement and Caumont, 2020), are removed.

3 Assimilation experiments design

3.1 Choice of bias correction and quality control

In the literature, some authors carry out a QC without a BC (Nipen et al., 2020). To evaluate this choice, we compare QC-
245 MC and QC-Titan. Note, however, that QC-MC was designed to be used after a BC. Every hour, the observations considered incorrect by the two QCs are removed. The hourly mean OmB for the PWS temperature observations are shown in Fig. 4a. The average OmB exhibits a systematic bias following a diurnal cycle: it varies from about 0.7 °C during the day to 1.3 °C during the night. The systematic warm bias could be due to site effect as the PWSs are placed close to heating environments such as windows or balconies (Bell et al., 2015; Sgoff et al., 2022). For relative humidity (Fig. 5b), the raw PWS observations
250 exhibit a diurnal bias composed of a drier nighttime and a wetter daytime. Its variations are close to what was reported by Sgoff et al. (2022). Regarding the observations after the two QC, the mean OmB is reduced throughout the day when using QC-MC, while it still shows a diurnal bias with QC-Titan (Fig. 4a). QC-MC takes into account a reference, which is here the

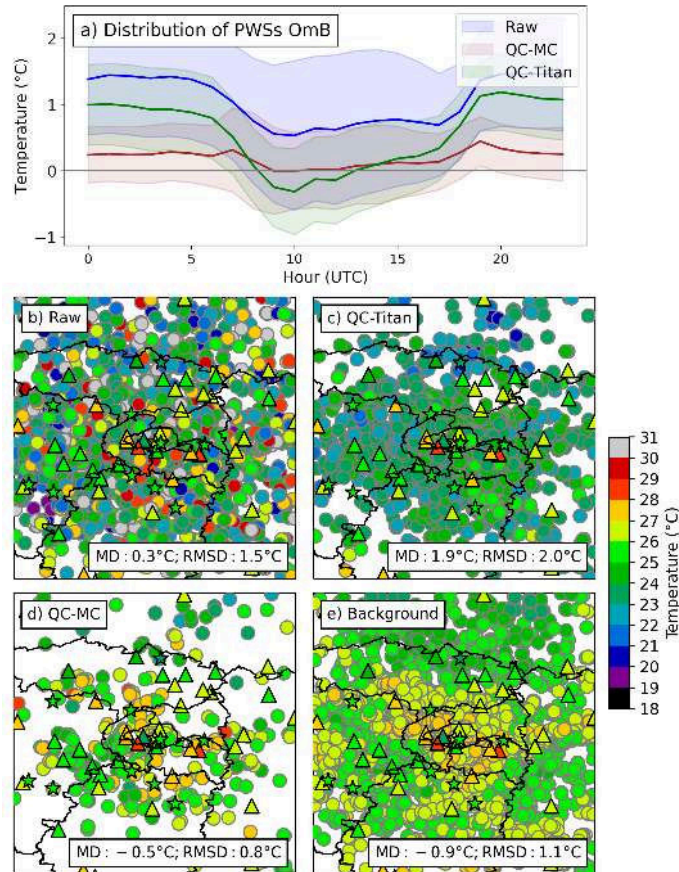


Figure 4. (a) Distributions of PWSs OmB for each hour of the study period: thick lines indicate the mean, thin lines indicate the mean \pm one standard deviation. The area between thick and thin lines is shaded. (b–e) Screen-level temperature observations around Paris on 7 September 2021 at 10:00 UTC from (coloured triangles with black contours) SWSs, (coloured stars) StatIC PWSs. Coloured circles are either (b–d) PWS screen-level temperature observations or (e) T_{2m} AROME background at PWS location. PWS observations are (b) raw or quality-controlled by (c) QC-Titan, (d) QC-MC. Boxes indicate the mean deviation (MD) and the rms deviation (RMSD) of PWS observations compared with StatIC PWS independent observations.

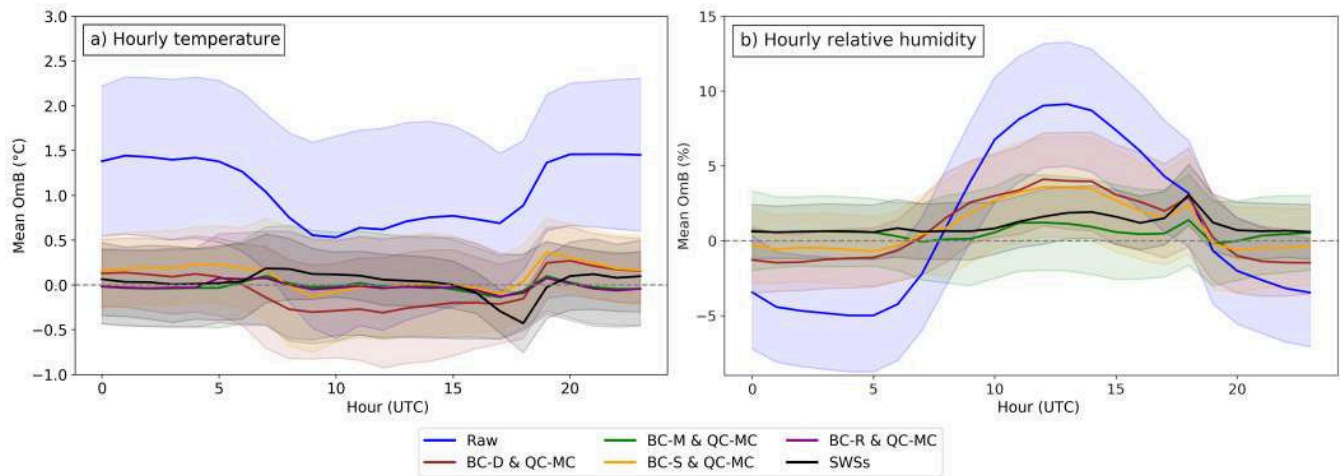


Figure 5. Mean diurnal cycle of OmBs during the 1-month study period of (a) temperature and (b) relative humidity for (blue) raw PWS observations, for PWS observations quality-controlled with QC-MC and bias-corrected with (brown) BC-D, (green) BC-M, (orange) BC-S, (purple, only in a) BC-R, and for (black) SWS observations. Lines and colour shades are as in Fig. 4a.

model background, allowing the QC to remove spatially incoherent observation in relation to the model background. A map of quality-controlled temperature observations near Paris at 12:00 LT is shown in Fig. 4b–e. Without QC, no regional pattern could be seen on the temperature observations from PWSs, explaining a high rms deviation to StatIC PWS observations of 2.4 °C (Fig. 4b). QC-Titan has the ability to remove observations which are too far from their neighbours. After QC-Titan, every single PWS observation is closer to its neighbours, but PWS observations have a mean deviation to StatIC PWS observations of 1.6 °C. When using QC-MC, more stations are removed, but only the observations close to the background are kept (QC-MC PWS OmB threshold is equal to ± 1 °C at the time of the figure). QC-MC will therefore be chosen for the next part of the study. However, QC-MC does not ensure by itself a zero bias of PWS observations: Fig. 4a shows a slight positive bias, and a BC is necessary to remove it.

To remove the remaining bias, 4 BC methods described in Sect. 2.4.1 are tested, in association with the selected QC-MC. Figure 5 shows the diurnal cycle of mean OmBs for both the PWS observations after applying the 4 BC methods, and the SWS observations. For temperature observations (Fig. 5a), mean OmBs for all choices of BC are close to zero and are in the order of magnitude of SWS observations OmBs. By construction, BC-M mean OmBs are near zero throughout the diurnal cycle; BC-R mean OmBs are very close: the addition of hourly global solar irradiance as a predictor does not seem to be making any substantial improvement. For relative humidity observations (Fig. 5b), all BCs tested exhibit a reduced diurnal cycle compared to raw PWS observations. BC-D and BC-S mean OmBs are close, as they both use 24 or 48 h rolling periods to estimate biases, respectively. Still, BC-S OmBs are slightly closer to zero during the night. Once again, the BC-M mean OmBs are near 0 % throughout the diurnal cycle, as SWS OmBs.

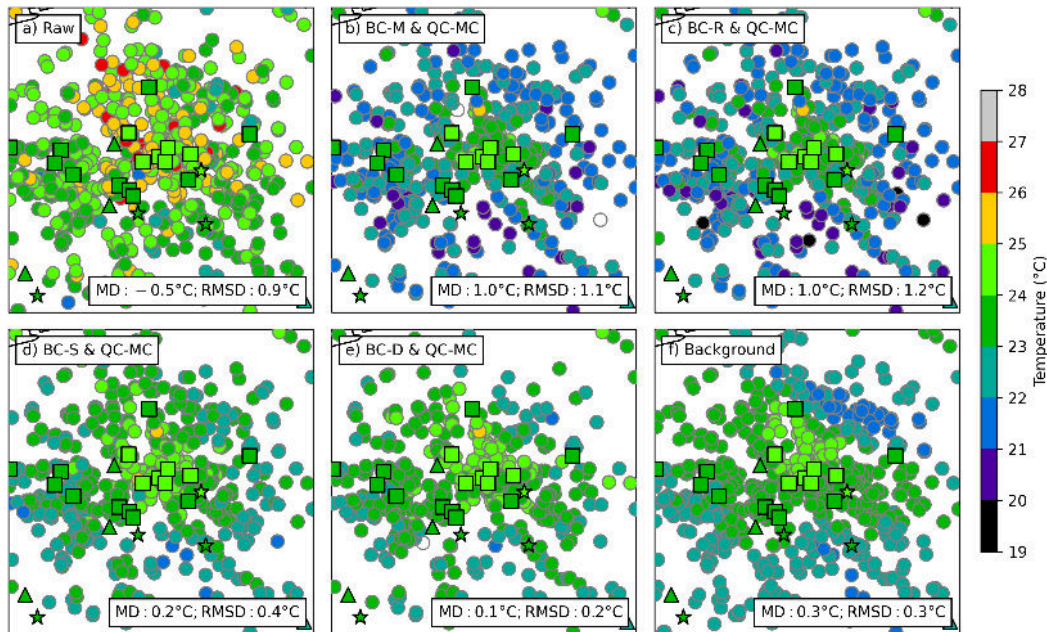


Figure 6. As in Fig. 4b–e, around the city of Toulouse on 7 September 2021 21:00 UTC (23:00 LT). Coloured circles are either (a–e) PWS screen-level temperature observations or (f) T_{2m} AROME background at PWS location. PWS observations are (a) raw, quality-controlled with QC-MC and bias-corrected with (b) BC-M, (c) BC-R, (d) BC-S, (e) BC-D. Coloured squares are Toulouse Métropole PWS screen-level temperature observations. Boxes indicate the mean deviation (MD) and the rms deviation (RMSD) of PWS observations compared with Toulouse Métropole PWS independent observations.

Figure 6 shows how the 4 BC methods modify PWS observations of screen-level temperature on 7 September 2021 23:00 LT around the city of Toulouse (centre of the map). Raw PWS temperature observations are close to observations of SWSs, Toulouse Métropole PWSs or StatIC PWSs networks, even if some PWS observations are warmer north-west of the city. However, BC-M and BC-R which uses monthly rolling periods to estimate biases reduce PWS temperatures in the suburbs of the city, which is inconsistent with other observation networks. In contrast, BC-S and BC-D, which use 24 or 48 h rolling periods to estimate biases, keep observations in agreement with other observations networks: the RMSD of PWS observations compared with Toulouse Métropole PWS independent observations reaches only 0.4°C and 0.2°C , respectively. T_{2m} of the AROME background (Fig. 6f) are close to Toulouse Métropole observations (RMSD of 0.3°C), while unfortunately none are available north-east of the city to support the colder temperatures indicated by the background.

280 3.2 Choice of observation error covariances and thinning

In variational DA schemes, the observation error covariance matrix \mathbf{R} has to be specified. As in D24, the part of \mathbf{R} allocated to PWS observations is prescribed diagonal, as for the part allocated to SWS observations, and the diagonal values (σ_o) are the same as for SWS observations: 1.4°C for T_{2m} and 10% for RH_{2m} . This is done for simplicity's sake, as modelling the

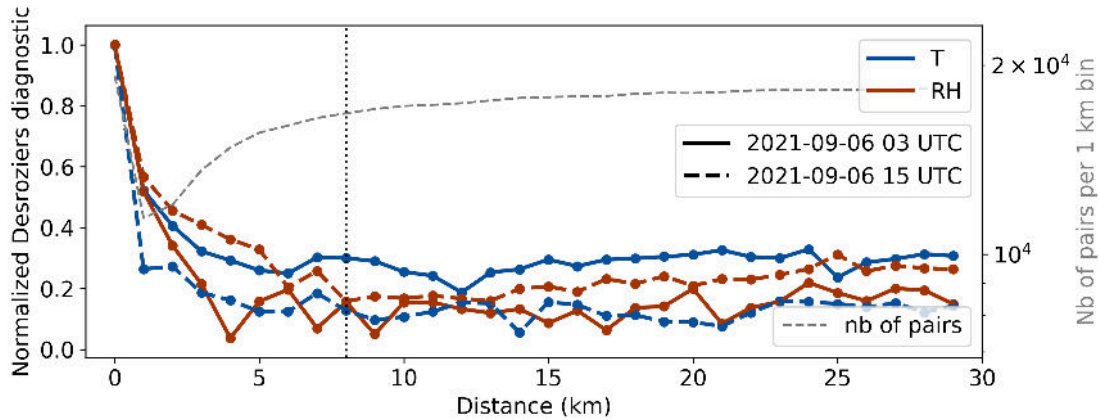


Figure 7. Spatial Desroziers diagnostic of PWS observations of (blue) temperature and (red) relative humidity after BC-M and QC-MC on 6 September 2021 at (plain lines) 03:00 UTC and (dashed lines) 15:00 UTC. In grey, the number of pairs of equidistant stations (in 1 km bins).

non-diagonal terms is an open research question (Guillet et al., 2019), and a study of sensitivity to a change in σ_o is beyond the scope of this article. The diagonal assumption demands to verify that observations errors are not correlated, and if they are, the reduction of the spatial density of the observations, may diminish the correlation between observation errors.

Figure 7 shows the spatial Desroziers diagnostic for PWS observations after BC-M and QC-MC. This diagnostic is computed using OmA and OmB of these PWS observations assimilated without thinning in a cycled 3DEnVar DA experiment launched on 6 September 2021 00:00 UTC. This diagnostic decreases when the distance between the observations increases. Two observations being far from each other have less probability of having correlated observation errors, reducing the influence of their local environment. The spatial Desroziers diagnostic decreases up to distances of 1 to 6 km, depending on the variable and the hour of the day. To reduce the probability of having correlated observation errors, a thinning is applied, selecting one random observation per mesh in an approximately 8 km horizontally spaced regular Gaussian grid (8 km thinning hereafter). One of the limit of the method is the fact that two observations may have correlated errors independently of the distance, for example if they are subject to similar anomalous siting conditions, and these observations are not corrected nor removed by the pre-processing.

Two pre-processing methods of PWS observations are selected for the DA experiments: P-M composed of BC-M, QC-MC and a 8 km thinning, and P-S composed of BC-S, QC-MC and a 8 km thinning. One of the points to bear in mind when using pre-processing methods based on comparisons with the background of the model into which we want to assimilate these observations is that pre-processed PWS observations can overfit the model background (see the triple collocation method shown in Appendix B).

Table 1. Overview of the experiments. The Monitor experiment is launched on 6 August 2021 00:00 UTC. All cycled experiments are launched on 6 September 2021 00:00 UTC. In the name of the experiments, X corresponds to the PWS observation variable assimilated: T for T_{2m} or RH for RH_{2m} .

Experiments	Duration	Cycling	Atmospheric DA scheme	Use of PWS observations
Monitor	2 months	No	3DVar	Raw observations monitored
3DVar	1 month	Yes	3DVar	No
3DVarX	1 month	Yes	3DVar	Variable X assimilated after P-M (BC-M, QC-MC, 8 km thinning)
3DEnVar	1 month	Yes	3DEnVar	No
3DEnVarX	1 month	Yes	3DEnVar	Variable X assimilated after P-M (BC-M, QC-MC, 8 km thinning)
3DEnVarXS	1 month	Yes	3DEnVar	Variable X assimilated after P-S (BC-S, QC-MC, 8 km thinning)
3DEnVarX-surf	1 month	Yes	3DEnVar	Variable X assimilated after P-M (BC-M, QC-MC, 8 km thinning) in both atmospheric (3DEnVar) and surface (OI) DA schemes

3.3 Overview of the experiments

All experiments are described in Table 1. The benefit from assimilating PWS observations of a variable X (T for T_{2m} or RH for RH_{2m}) is evaluated using the observing system experiment (OSE) framework, comparing a reference experiment (3DVar, 3DEnVar) to an experiment where PWS observations are assimilated (Pourret et al., 2022; D24). To explore the role of the pre-processing (or bias correction, since this is the only part that changes), PWS observations after P-M and P-S are assimilated using the 3DEnVar DA scheme (3DEnVarX and 3DEnVarXS, respectively). To explore the role of the DA scheme, PWS observations after P-M are assimilated using the 3DVar DA scheme (3DVarX). Finally, the impact of assimilating simultaneously PWS observations after P-M at the surface using the OI DA scheme and in the atmosphere with the 3DEnVar DA scheme is tested (3DEnVarX-surf). Monitor, 3DVar and 3DEnVar experiments are the same as in D24. Except Monitor, they are cycled, which means that the assimilation of PWS observations at time t influences the background at time $t + 1$ h, used to make the new analysis. During one month, 719 analyses are made, and 119 forecasts are launched.

4 Results of the assimilation experiments

The experiments are evaluated using statistics of one-month OmB and OmF (observation minus forecast) where observations are systematically SWS observations. The relative evolution of the rms OmX (X being B or F) of an experiment (XP) w.r.t. another (CTRL) is given by:

$$\Delta_{\text{rms OmX}} = \frac{\text{rms}(\text{OmX})_{\text{XP}} - \text{rms}(\text{OmX})_{\text{CTRL}}}{\text{rms}(\text{OmX})_{\text{CTRL}}} \quad (2)$$

As in D24, $\Delta_{\text{rms OmX}}$ is considered significant if 0 is not in the 95 % confidence interval around it. This interval is computed by bootstrap with the "percentile" method (scipy.stats.bootstrap function, Virtanen et al., 2020): OmX_{XP} time series is

Table 2. Δ rms OmB (%) of surface pressure, T_{2m} , RH_{2m} and 10 m zonal and meridional wind for SWSs over the one-month study period. Negative values of an experiment (XP) w.r.t. another (CTRL) indicate that the backgrounds (1 h forecasts) of XP are closer than CTRL to SWS observations, i.e. improvement and positive values indicate degradation. Significant values are in bold.

	Surface pressure		T_{2m}		RH_{2m}		Zonal wind		Meridional wind	
	AROME domain	France	AROME domain	France	AROME domain	France	AROME domain	France	AROME domain	France
3DVarT w.r.t. 3DVar	0.2	0.7	1.8	2.4	-0.1	-0.3	-0.3	-0.5	-0.1	-0.2
3DEnVarT w.r.t. 3DEnVar	0.1	1.6	-0.7	-0.9	-0.2	-0.2	-0.1	-0.2	0.0	0.0
3DEnVarTS w.r.t. 3DEnVar	1.3	3.7	-0.6	-0.8	-0.1	-0.2	0.1	0.1	0.0	0.0
3DEnVarT-surf w.r.t. 3DEnVar	0.1	0.9	0.5	0.7	0.3	0.5	0.1	0.4	0.0	0.1
3DVarRH w.r.t. 3DVar	-0.3	-0.8	-0.6	-0.9	2.7	4.5	-0.2	-0.2	-0.1	-0.2
3DEnVarRH w.r.t. 3DEnVar	0.2	0.3	0.2	0.3	-1.0	-1.5	0.0	0.0	0.1	0.2
3DEnVarRHS w.r.t. 3DEnVar	0.5	0.7	0.5	0.8	-0.8	-1.2	0.1	0.2	0.0	0.2
3DEnVarRH-surf w.r.t. 3DEnVar	0.1	0.1	0.3	0.4	-0.9	-1.3	0.1	0.2	0.1	0.2

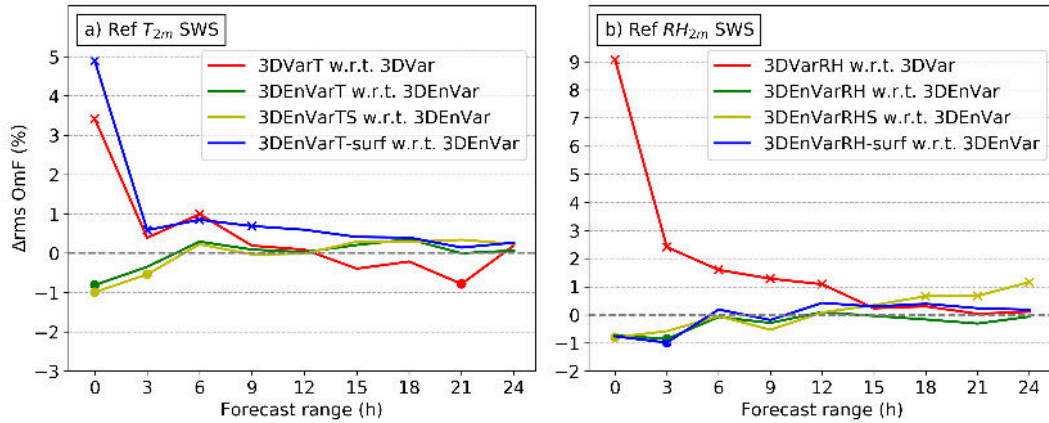


Figure 8. Δ rms OmF of (a) T_{2m} and (b) RH_{2m} for SWSs over France every 3 h between 0 and 24 h forecast range over the one-month study period. Negative values indicate improvement and positive values indicate degradation. Points (respectively crosses) indicate improvement (resp. degradation) of XP w.r.t. CTRL at 90% statistical significance level.

320 randomly sampled with replacement, giving a new time series of the same size. This is done 1000 times, providing 1000 time series, giving 1000 rms values forming a distribution, from which the confidence interval is estimated.

4.1 Impact of the bias correction with the 3DVar DA scheme

When PWS T_{2m} observations are assimilated in 3DVarT or 3DVarTS (Table 2), T_{2m} rms OmBs decrease significantly for both experiments in comparison with 3DVar, which is an improvement. The improvement is slightly larger with 3DVarT (325 Δ rms OmB of -0.7%) than with 3DVarTS (-0.6%) over the AROME domain. The improvement is larger (-0.9% and -0.8% , respectively) over France, as it is the area where PWS observations are assimilated. Surface pressure rms OmB significantly increases over France for 3DVarT ($+1.6\%$) and 3DVarTS ($+3.7\%$). Other variables show no significant evolution.

For PWS RH_{2m} observation assimilation in 3DVarRH or 3DVarRHS, RH_{2m} rms OmBs also decrease significantly (330 over the AROME domain and (Δ rms OmB of -1.0% and -0.8% , respectively). The improvement is larger over France (-1.5% and -1.2% , respectively). However, a slight but significant degradation of T_{2m} rms OmBs is found in the 3DVarRHS experiment both over France and the AROME domain (0.8% and 0.5% , respectively). Other variables show no significant evolution.

For experiments assimilating PWS observations either of T_{2m} or RH_{2m} , the lowest rms OmBs are obtained with the P-M (335 pre-processing for all variables over France.

Scores for forecasts up to 24 h are shown in Fig. 8. When PWS T_{2m} observations are assimilated (3DVarT and 3DVarTS in Fig. 8a), an improvement of T_{2m} rms OmFs is found up to 3 h forecast range; beyond 3 h rms OmFs evolution is stable to slightly degraded. Similar results are found in 3DVarRH and 3DVarRHS when PWS RH_{2m} observations are assimilated. These results are consistent with Sgoff et al. (2022) results showing improvements vanishing or not significant after 5 to 6 h (340 when assimilating bias-corrected PWS observations only.

No significant evolution of rms OmBs or OmFs of other observing systems is noticed (not shown), which could be explained by the small impact of T_{2m} and RH_{2m} above the ABL (Brousseau et al., 2014). Also, for observing systems such as radiosoundings, the low number of observations over France makes it difficult to show significance (D24).

4.2 Impact of the DA scheme

When compared to 3DVar, 3DVarT significantly degrades by $+2.4\%$ over France the T_{2m} rms OmBs (Table 2). For 3DVarRH (345 compared to 3DVar, this degradation reaches $+4.5\%$. The degradation is very large at the analysis and remains significant between 6 to 12 h forecast range for T_{2m} and RH_{2m} , respectively (Fig. 8). As it was found by D24 for PWS surface pressure observation assimilation, the 3DVar DA scheme with its operational settings is not able to take advantage of these PWS observations with the pre-processing (including the thinning) which is selected, mainly due to the large size of the increments.

When PWS T_{2m} observations are concomitantly assimilated by the atmospheric and surface DA systems (3DVarT-surf), (350 the improvement found in 3DVarT turns off to a significant degradation for T_{2m} rms OmBs (Table 2). The degradation is significant up to 9 h forecast range (Fig. 8). When PWS RH_{2m} observations are concomitantly assimilated by the atmospheric and the surface DA system (3DVarRH-surf), RH_{2m} rms OmBs and OmFs are very similar to 3DVarRH, showing an improvement. These large differences between the assimilation of PWS observations of T_{2m} and RH_{2m} in the surface DA

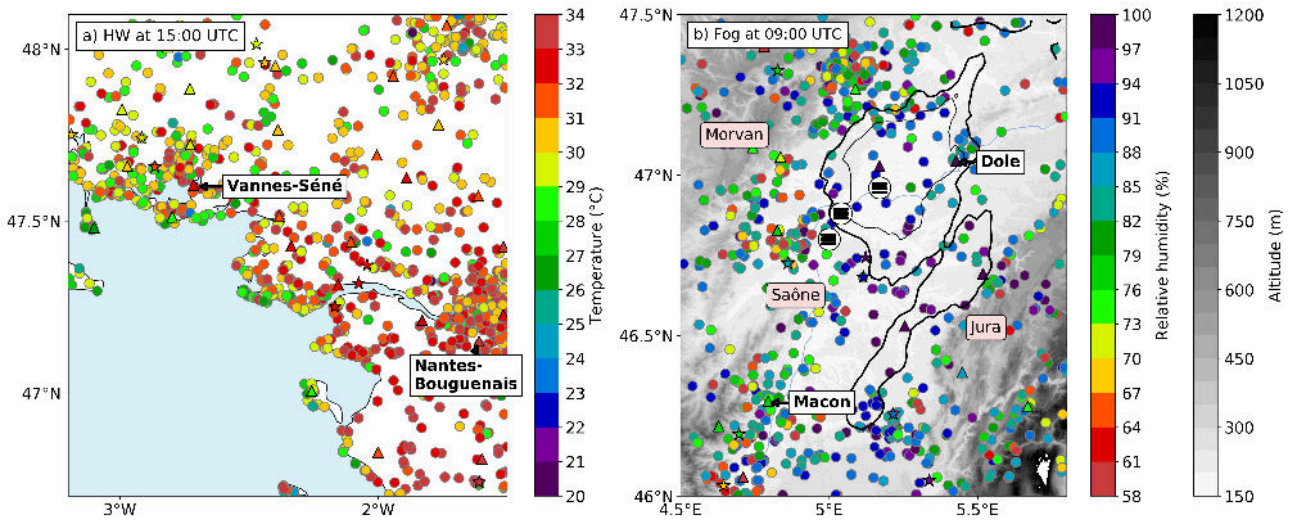


Figure 9. Geographical situation of (a) the heatwave case on 7 September 2021 15:00 UTC with raw screen-level temperature observations and (b) the fog case on 23 September 2021 09:00 UTC with raw screen-level relative humidity observations. SWSs are indicated by coloured triangles with black contours, PWSs by coloured circles, and StatIC PWSs by coloured stars. Human crowdsourced observations of fog between 08:00 and 09:00 UTC are indicated by three encircled horizontal bold lines (fog symbol). MSG-3 satellite HRV channel reflectance observations equal to 24 % and 43 % are shown by bold and thin black edges, respectively. Light blue indicates the sea (the Atlantic Ocean).

355 scheme could be explained by the fact that the RH_{2m} observation operator, contrary to the T_{2m} observation operator, uses only variables from the lowest atmospheric model level, and not surface variables. These large differences could also be attributed to the differential impact of increments in time and space between low level temperature and specific humidity fields.

5 Results on case studies

In case of a strong near-surface gradient of a meteorological variable caused by a meso- β scale meteorological phenomenon, 360 the added value of a dense network of weather stations is likely to be clearly visible. Two cases with substantial temperature and relative humidity gradients in September 2021 are studied: a sea breeze case during a heatwave on 7 September with around 4 °C of screen-level temperature difference between two SWS observations 12 km apart, and a fog case on 23 September with 40 % screen-level relative humidity differences between two SWS observations 25 km apart.

5.1 Sea breeze front during a heatwave on 7 September 2021

365 On 7 September 2021, at the synoptic scale, subtropical warm air came from North Africa, ahead of a low pressure system located over the Atlantic Ocean. Near the surface, the south-eastern to eastern winds carried the warm continental air mass over western France. Maximum daily temperature reached 33.7 °C at Nantes-Bouguenais SWS at 15:07 UTC and 32.2 °C at 15:00 UTC at Vannes-Séné SWS (Fig. 9a). It is 11.3 and 10.3 °C more than the 1991–2020 September normal maximum

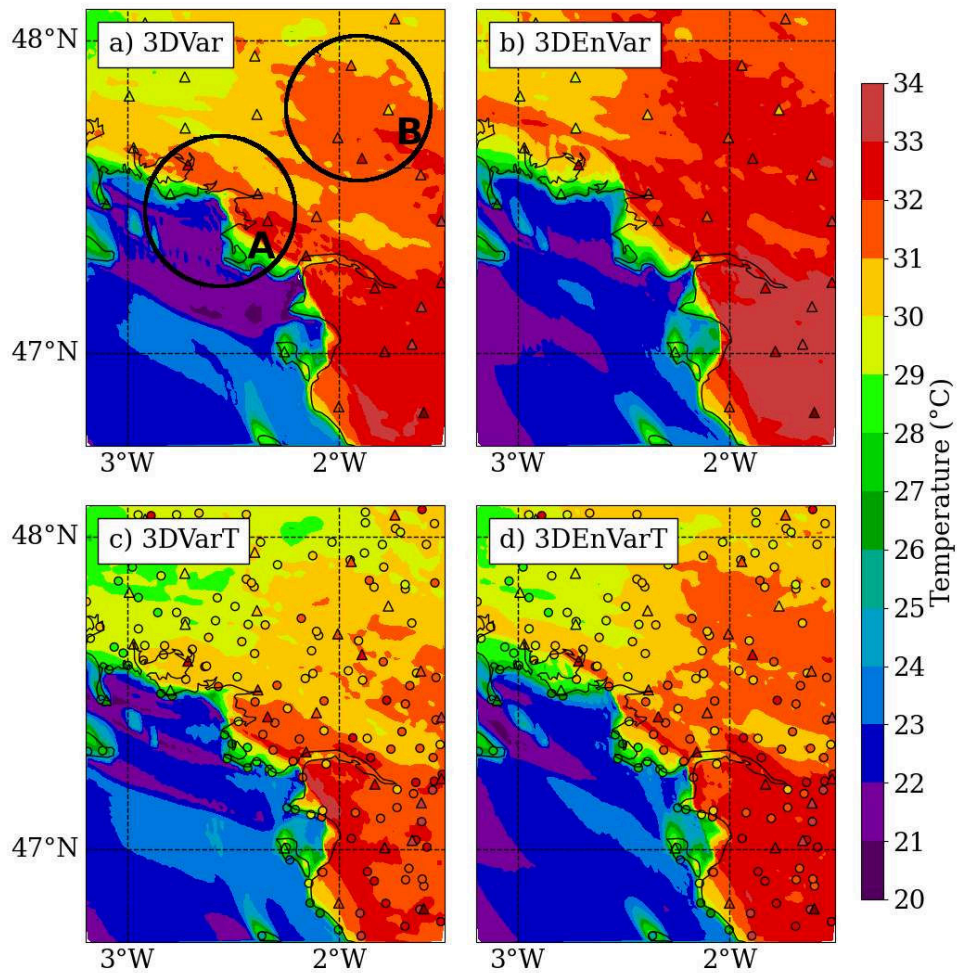


Figure 10. Temperature of lowest atmospheric model level (around 5 m height AGL) analyses for (a) 3DVar, (b) 3DEnVar, (c) 3DVarT, and (d) 3DEnVarT experiments on 7 September 2021 at 15:00 UTC. Locations of the stations whose observations of screen-level temperature are assimilated are indicated by triangles for SWSs and circles for PWSs.

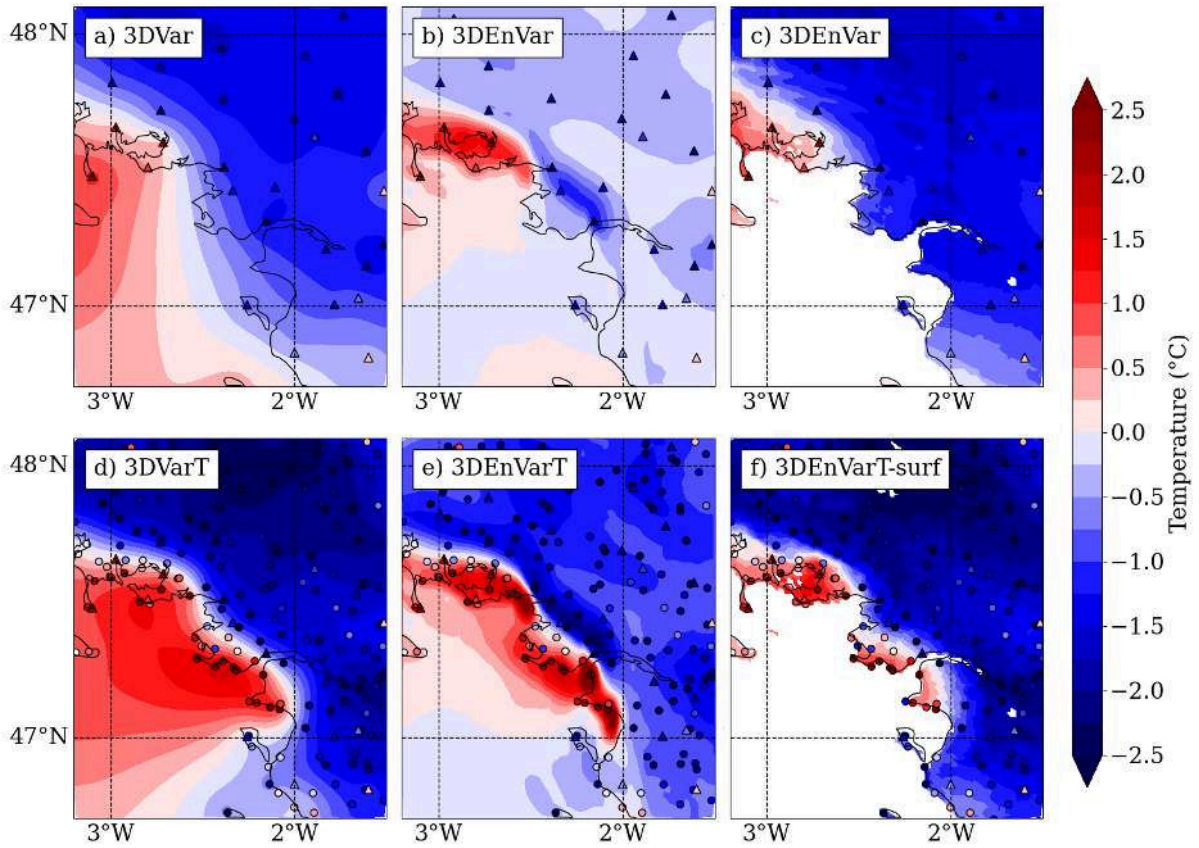


Figure 11. Temperature increments (a–b, d–e) at the lowest atmospheric model level (around 5 m height AGL) and (c, f) at the highest surface model level (ground at 0 m height AGL) on 7 September 2021 at 15:00 UTC. These experiments are designed as (a) 3DVar, (b, c) 3DEnVar, (d) 3DVarT, (e) 3DEnVarT and (f) 3DEnVarT-surf experiments, the only difference being that they all start from the same 3DVar background at 14:00 UTC. The location of stations whose observations of screen-level temperature are assimilated are indicated by triangles for SWSs and circles for PWSs; their colours indicate temperature OmBs.

temperature of these SWSs, respectively. The daily maximum temperature at Nantes-Bouguenais was 5 °C above or equal to
370 the September normal from 2 to 8 September, i.e. for more than 5 consecutive days, which is the WMO criterion for heatwaves.

Figure 10 shows the temperature analysis at around 5 m height AGL, for the different experiments, and both SWS and PWS
screen-level temperature observations. If the height between observations and this analysis field is different, it is shown here
because it is the lowest atmospheric level to which DA increments are applied. The experiments are cycled since 6 September
2021 at 00:00 UTC, which accounts for the significant discrepancy observed between the analysis. The two types of pre-
375 processing show similar results for this case (not shown).

In the afternoon, a sea breeze rises up due to the difference in temperature between the ocean and the land, cooling the coast.
This cooling is observed by three SWSs, and well observed with PWSs, particularly after pre-processing (Figs. 9a and 10).
Analyses from the four DA experiments (Figure 10) are able to reproduce the coastal cooling caused by the breeze, but with
differences in temperature on either side of the breeze front and differences in location of this front.

380 In the 3DVar experiment (Fig. 10a) the temperature analysed is very close to SWS observations of temperature. When PWS
temperature observations are assimilated in 3DVarT (Fig. 10c), the temperature analysed in the area A is lower than in 3DVar,
which increases the deviation from the analysis of the Vannes-Séné SWS observation. In the 3DEnVar experiment (Fig. 10b),
still in the area A near Vannes-Séné SWS, the breeze front is located further inland which is not in agreement with this SWS
observation or PWS observations shown in Fig 10c. Also, in the area B, temperature analysed in 3DEnVar is higher than
385 SWS temperature observations. In this case, 3DEnVarT analysis seems closer than the other analyses to both PWS and SWS
observations, in particular in areas A and B (Fig. 10d).

To illustrate the shape of temperature increments given by each DA scheme depending on the observations assimilated (from
SWSs only or from both SWSs and PWSs), at the lowest atmospheric level for atmospheric DA schemes and at ground level
for surface schemes, experiments starting from the same 3DVar operational background at 14:00 UTC are shown in Fig. 11.

390 Increments at the lowest atmospheric model level have lower horizontal gradients in 3DVar (Fig. 11a) w.r.t. 3DEnVar
(Fig. 11b) and in 3DVarT (Fig. 11d) w.r.t. 3DEnVarT (Fig. 11e). Because the information from observations propagate at
longer distance with the settings of the 3DVar DA scheme than the 3DEnVar DA scheme (as shown in Sect. 2.1.1 mainly due
to the localisation), increments of the same sign as the OmBs of coastal observations propagate over longer distances at sea in
3DVar and 3DVarT than in 3DEnVar or 3DEnVarT experiments.

395 In the area B and generally inland, increments are much lower than the neighbouring SWS OmBs in 3DEnVar whereas
it is not the case in 3DVar. As shown in Sect. 4, the 3DEnVar DA scheme gives less relative weight to surface temperature
observations, which explains why observations have little impact on the temperature analysis. Assimilating PWS observations
(Fig. 11e) increases the amplitude of increments, in particular inland, in 3DVarT w.r.t. 3DVar or in 3DEnVarT w.r.t. 3DEnVar.

400 At ground level, the shape of the ground temperature increments with the OI in 3DEnVar experiment (Fig. 11c) is similar to
the shape of the atmospheric increments in the 3DVar experiment. When assimilating PWS temperature observations with the
OI, the sign of increments at the surface is generally the same as the sign of increments added in the atmosphere in 3DVarT or
3DEnVarT, and their values are closer to 3DVarT than to 3DEnVarT.

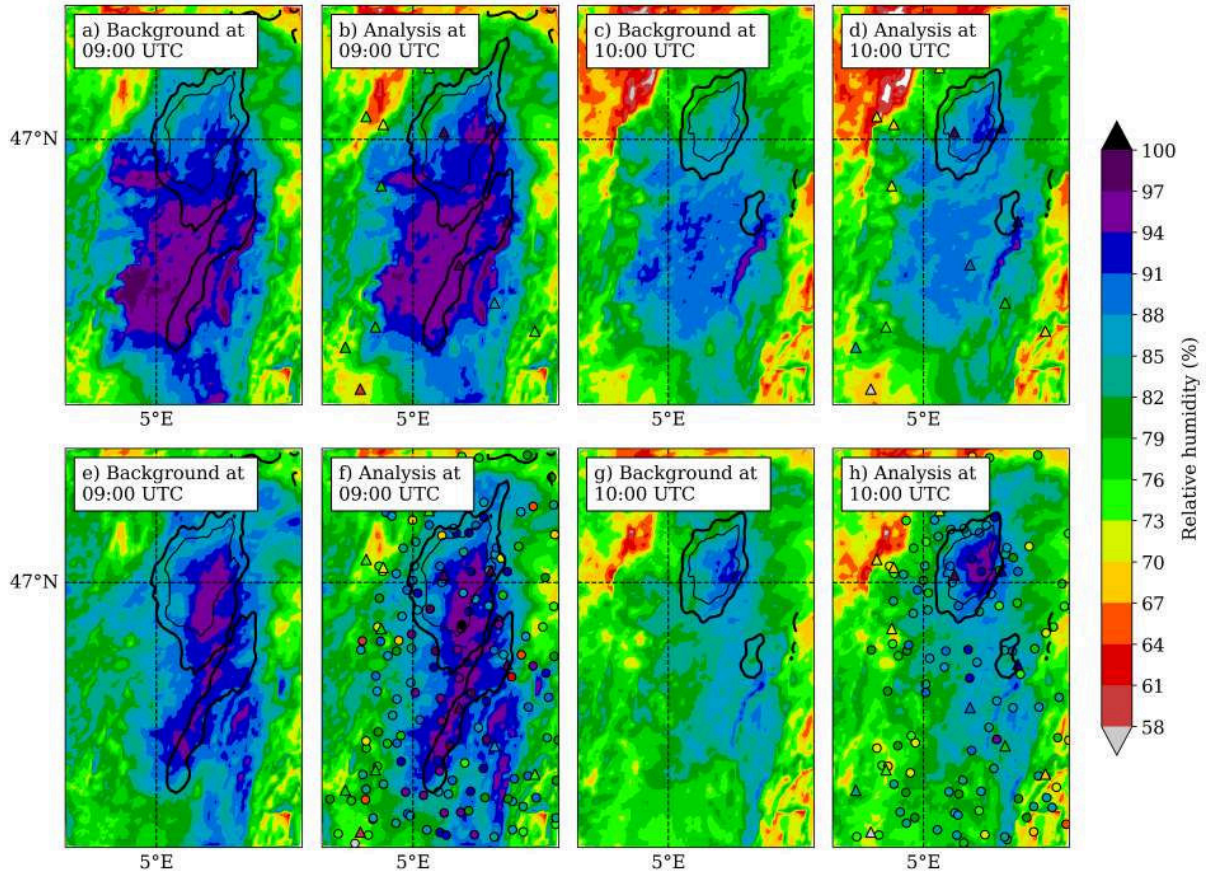


Figure 12. Relative humidity at the lowest atmospheric model level (around 5 m height AGL) for (a–d) 3DEnVar, and (e–h) 3DEnVarRHS experiments on 23 September 2021. (a, c, e, g) Backgrounds and (b, d, e, f) analyses are shown at two consecutive assimilation times: (a, b, e, f) 09:00 UTC and (b, d, f, h) 10:00 UTC. Assimilated observations of relative humidity from (triangles) SWSs and (circles) PWSs are indicated. MSG-3 satellite HRV channel reflectance observations equal to 24 % and 43 % are shown by bold and thin black edges, respectively.

5.2 Fog on 23 September 2021

The Saône valley, in the northwest of France (Fig. 2), is regularly affected by fog during the autumn and winter seasons. Situated between the Morvan and the Jura mountainous massif, it is crossed by a few rivers, including the Saône (Fig. 9b), providing humidity near the surface. On 23 September 2021, high-pressure conditions are associated with stable air masses over France (not shown). At 09:00 UTC (Fig. 9b), MSG-3 satellite HRV observations indicate the presence of low clouds. These clouds are reaching the ground, causing mist and fog: SWS and PWS relative humidity observations reach 90 to 100 %, three crowdsourced human observations indicate fog between 08:00 and 09:00 UTC and a forward scatter sensor at Tavaux

410 SWS, south-west of Dole, observes a meteorological optical range of 1020 m at 09:00 UTC, between mist and fog. Out of the fog, between 4.5 and 5° E, relative humidity observations range from 60 to 90 %.

Figure 12 shows the evolution of the relative humidity in the backgrounds and analyses of two experiments at two assimilation times: 09:00 and 10:00 UTC.

At 09:00 UTC, at the lowest atmospheric level, relative humidity backgrounds of 3DEnVar (Fig. 12a) and 3DEnVarRHS (Fig. 12e) experiments are different, due to the cycled assimilation. High values of relative humidity simulated by 3DEnVarRHS appear closer to the MSG-3 observation of clouds, in its east-west extension, than 3DEnVar. The analysis at 09:00 UTC (Fig. 12b, f) slightly increases the relative humidity under the observed fog, but is not able to substantially modify the shape of the wet area in the background in only one assimilation time step, in particular in 3DEnVar where only SWS observations are assimilated.

420 At 10:00 UTC, the 1 h forecast drastically reduces the wet area for both experiments (Fig. 12c, g). Whereas in 3DEnVar, the area with relative humidity above 91 % is still large in its east-west extension, in 3DEnVarRHS the wettest area is located where clouds are still observed by satellite. The assimilation at 10:00 UTC increases the relative humidity in this area, for both experiments, which exceeds 94 %, as indicated by one SWS, only in 3DEnVarRHS.

This example shows the importance of rapid-update cycles in DA, as one assimilation may have a low impact, but several are gradually changing the simulated relative humidity. The model plays a preponderant role in reducing the simulated fog from 09:00 to 10:00 UTC. Here this situation is illustrated with 3DEnVarRHS because this experiment is closer to the observations than 3DEnVarRH. Regarding the results of 3DEnVarRH-surf, in this case, the assimilation of PWS observations in the OI scheme at the surface does not change the shape and the intensity of the simulated fog (not shown).

6 Discussion and conclusion

430 This study has explored the impact of the assimilation of temperature and relative humidity observations from PWSs in the current operational AROME system. Before the traditional screening stage of a DA system, two pre-processing methods of PWS observations were designed and selected. This was done to obtain pre-processed PWS observations with statistics of differences to the AROME background close to SWS observations currently assimilated in the AROME DA system.

- Of the four bias correction methods evaluated, two have been selected, including the method of Sgoff et al. (2022) designed to reduce the diurnal bias of PWSs. Because these methods are based on statistics of differences between AROME background and observations, there is a risk of transferring model bias to PWS observations, hence the importance of assimilating anchor near-surface observations.
- Two quality control methods have been evaluated. The Titan-QC original design with three checks has been found detrimental when multiple PWS exhibit biases with respect to SWS observations. A QC adapted from Mandement and Caumont (2020) has been used to remove PWS observations judged erroneous based on their departures with the AROME background. Advanced quality control methods could be tested based on background deviation, as First Guess Test from TITAN library (Lussana, 2023).

- A thinning has been applied to reduce the probability of having correlated observation errors, by taking one random observation per mesh in an approximately 8 km horizontally spaced regular Gaussian grid.

445 Then, the OSE framework has been used: the added value of PWS observation assimilation with two selected pre-processing methods and existing DA schemes has been quantified. The assimilation of pre-processed PWS temperature observations with the 3DEnVar DA scheme reduced between 0.7 and 0.9 % the rms SWS temperature OmBs in France, depending on the pre-processing method. Departures between SWS observations and forecasts are reduced up to 3 to 6 h range. Results are similar for PWS relative humidity observations, with a reduction of 1.2 to 1.5 % of the rms SWS relative humidity OmBs in France. These
450 findings are in agreement with benefits found by Sgoff et al. (2022), in experiments where PWS and SWS observations were not jointly assimilated. Assimilating PWS temperature (resp. relative humidity) observations with 3DVar DA scheme shows deterioration of the rms SWS temperature (resp. relative humidity) OmBs, as it was found for PWS pressure observations in D24. Deterioration was also founded when assimilating concomitantly the PWS observations at the surface with the IO scheme.

Two distinct bias corrections were subjected to evaluation. The monthly bias correction (BC-M) demonstrated superior
455 performance in reducing the departure between SWS observations and the model background. In contrast, the two-day temporal window bias correction (BC-S) exhibited adaptive correction capabilities that were shown to be particularly pertinent in case studies. A compromise between the two methods could be achieved by selecting the optimal time window based on the prevailing meteorological conditions, which could be done with third-party observations.

Finally, the assimilation of pre-processed PWS observations with the 3DEnVar DA scheme has been subjectively found to
460 better represent the fine-scale features of a sea breeze during a heatwave event, and the evolution of a fog event. In these case studies, the relative weight given to near-surface observations is found to be lower in 3DEnVar compared to 3DVar, which is compensated by the high amount of PWS observations assimilated.

This study has some limitations. The OmBs used for P-M and P-S pre-processing for all experiments are computed from a monitoring experiment with a 3DVar DA scheme (Monitor). A more rigorous calculation would require a monitoring experi-
465 ment to be carried out for each assimilation scheme tested. To determine the thinning distance, the spatial Desroziers diagnostic (Fig. 7) is computed from an experiment with a 3DEnVar DA scheme. This thinning distance is applied to the experiments with a 3DVar DA scheme as well. Perhaps the scores of the 3DVar experiments are degraded by an inappropriate choice of the thinning length, which could be increased in future studies. Furthermore, no testing has been specifically conducted in complex topographical areas, such as mountainous regions, where the pre-processing may deteriorate the new observations or
470 reject them due to temperature model bias (Préaux et al., 2023).

The joint assimilation of temperature, relative humidity, and surface pressure observations from PWSs in AROME, could now be tested, as these variables describe weather structures more coherently together than separately. Efforts should be made to evaluate such joint assimilation over a wider variety of cases and perhaps over longer study periods.

For operational use, beyond the recommendations of D24 which also apply in this case, this study has not attempted to
475 modify the settings of the assimilation systems used (3DVar, 3DEnVar or OI), and the few new settings have been chosen identically to the operational ones. However, studies of the sensitivity of these settings would be desirable, given their importance in the results obtained and because they are set to obtain the best scores with the observations currently assimilated. In the

atmosphere, while the 3DEnVar system is currently only being tested, there are plans to replace it rapidly with a 4DEnVar system: it would allow more observations to be included in the analysis by increasing the number of assimilation windows. At 480 the surface, the replacement of the OI by a 2DEnVar system is currently under study. This would allow more observations to be assimilated at the surface and ensure greater consistency between atmospheric and surface schemes: it is also a first step before possibly coupling surface and atmospheric DA systems.

Appendix A: Relative humidity computation

The relative humidity (RH) is calculated in AROME as:

$$485 \quad RH = \frac{100Pq}{e_{\text{sat}}\left(\frac{R_{\text{dry}}}{R_{\text{vap}}} + q\left(1 - \frac{R_{\text{dry}}}{R_{\text{vap}}}\right)\right)} \quad (\text{A1})$$

P is the pressure in Pa, q is the specific humidity (dimensionless), $R_{\text{dry}} = 287.0597 \text{ J kg}^{-1} \text{ K}^{-1}$ and $R_{\text{vap}} = 461.5250 \text{ J kg}^{-1} \text{ K}^{-1}$ are the gas constants for dry air and water vapour, respectively.

The saturation vapour pressure e_{sat} in Pa is equal to:

$$e_{\text{sat}} = \exp\left(\alpha_l + (\alpha_s - \alpha_l)\delta - \frac{\beta_l + (\beta_s - \beta_l)\delta}{T} - (\gamma_l + (\gamma_s - \gamma_l)\delta) \ln(T)\right) \quad (\text{A2})$$

490 with $\delta = 0$ if $T \geq 273.16 \text{ K}$ and $\delta = 1$ if $T < 273.16 \text{ K}$ and,

$$\alpha_l = \ln(e_{\text{sat}}(T_t)) + \frac{\beta_l}{T_t} + \gamma_l \ln(T_t) \quad \beta_l = \frac{L_v(T_t)}{R_v} + \gamma_l T_t \quad \gamma_l = \frac{C_l - C_{\text{pv}}}{R_v} \quad (\text{A3})$$

$$\alpha_s = \ln(e_{\text{sat}}(T_t)) + \frac{\beta_s}{T_t} + \gamma_s \ln(T_t) \quad \beta_s = \frac{L_s(T_t)}{R_v} + \gamma_s T_t \quad \gamma_s = \frac{C_s - C_{\text{pv}}}{R_v} \quad (\text{A4})$$

T is the temperature in K, $T_t = 273.16 \text{ K}$ is the triple point temperature, $L_v(T_t) = 2.5008 \times 10^6 \text{ J kg}^{-1}$ is the specific latent heat of vaporization, $L_s(T_t) = 2.8345 \times 10^6 \text{ J kg}^{-1}$ is the specific latent heat of sublimation, $e_{\text{sat}}(T_t) = 611.14 \text{ Pa}$, 495 $C_l = 4218 \text{ J kg}^{-1} \text{ K}^{-1}$ is the specific heat for water in its liquid phase, $C_s = 2106 \text{ J kg}^{-1} \text{ K}^{-1}$ is the specific heat for water in its solid phase and $C_{\text{pv}} = 4R_{\text{vap}}$ is the specific heat at constant pressure and constant volume for water vapour.

Appendix B: Comparison of observations and AROME background by triple collocation

PWS and SWS screen-level observations are considered collocated when separated by less than 1 km horizontally, 50 m vertically and less than 10 min apart in time. When it happens, the AROME background from the Monitor experiment (Table 1) 500 corresponding to this SWS observation, given by the observation operator, provides a third value of the physical variable considered ($T_{2\text{m}}$ or $RH_{2\text{m}}$). This triple collocation method (Stoffelen, 1998) is performed over the 1-month study period, giving a sample of 536000 (respectively 80000) triplets of collocated values for temperature (resp. relative humidity).

For each variable, three triplets with three different pre-processing methods for PWS observations are tested: (i) raw i.e. without pre-processing (ii) P-M without thinning and (iii) P-S without thinning. We will refer to them simply as P-M and

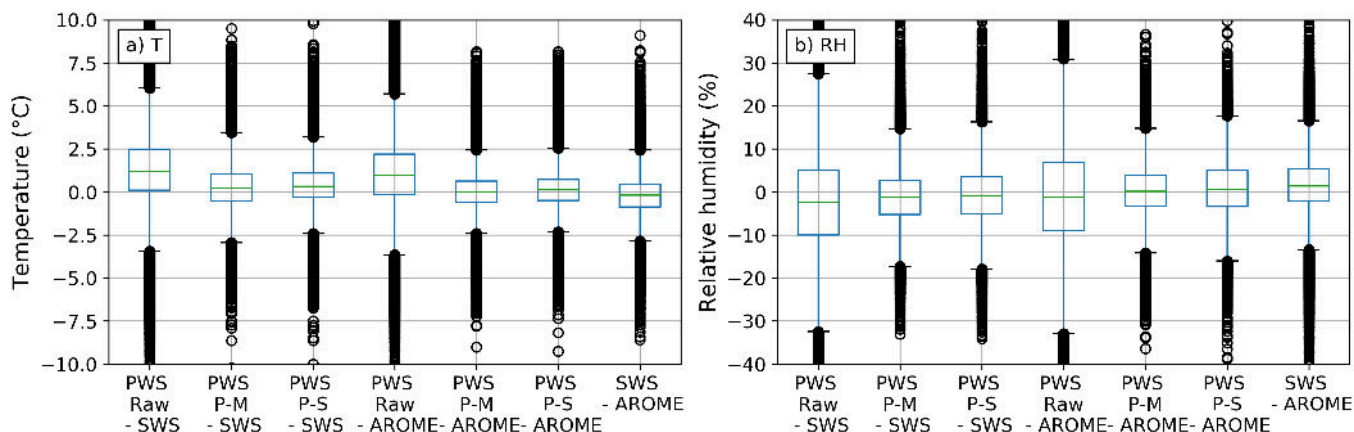


Figure B1. Boxplots showing distributions of pairwise differences of (a) T_{2m} and (b) RH_{2m} in triplets of collocated values (SWS observations, PWS observations, AROME background from Monitor), depending on the pre-processing (raw, P-M without thinning, P-S without thinning) applied to PWS observations.

505 P-S in this paragraph only. Figure B1 shows distributions of pairwise differences between the values in each triplet. Only 7 columns, instead of 9, are shown because 3 pairwise differences (SWS observations minus AROME background from Monitor) are identical.

Regarding the distributions of temperature differences between PWS and SWS collocated observations, their standard deviation is substantially reduced when pre-processing are applied, from 2.3 °C (raw) to 1.3 °C with P-M or 1.3 °C with P-S. Also, 510 mean differences are closer to 0 and the interquartile range is reduced.

Regarding the distributions of temperature differences between the pre-processed PWS observations, with both P-M and P-S, and the AROME background, the standard deviation equals to 1.0 °C which is smaller than 1.3 °C, the standard deviation of differences between the pre-processed PWS observations and SWS observations. This could show that the selected pre-processing methods of PWS observations overfit the AROME background.

515 Similar conclusions can be drawn for relative humidity, although the effects of PWS pre-processing do not reduce interquartile ranges of differences with SWS observations or the AROME background by the same factor as for temperature.

Author contributions. Alan Demortier: Formal investigations, methodology, analysis; resources; writing – original draft. Marc Mandement: Formal analysis; resources. Vivien Pourret: Formal analysis; software; validation; supervision. Olivier Caumont: validation; project administration; supervision; validation.

520 *Competing interests.* The authors declare that they have no conflict of interest.

Code availability. The code used for the assimilation experiments in AROME-France is owned by the members of the ACCORD consortium. This agreement allows each member of the consortium to license the shared ACCORD codes to academic institutions in their home countries for non-commercial research. Access to ACCORD codes, and codes used for the figures, can be obtained by contacting the corresponding author.

525 *Data availability.* WMO essential weather station observations (SYNOP and SHIP reports), French RADOME observation reports, AROME
analyses and forecasts are available freely online (Météo-France, 2024a, b, c, d). PWS observations are provided freely by Infoclimat (2024)
for the StatC network and Toulouse Métropole (2024) for the Toulouse Métropole network; freely in real-time, and archives on request
by Netatmo (2024) for the Netatmo network and they cannot be shared according to their licence: <https://dev.netatmo.com/legal>. Satellite
MSG-3 HRV observations used for this study (corrected by Météo-France) are available on request from the corresponding author, level 1.5
530 observations are provided by Eumetsat (2017). The results of data assimilation experiments are available on request from the corresponding
author solely for non-commercial research purposes and for reasonable data volumes.

Acknowledgements. The authors would like to thank Jean-Baptiste Hernandez from the Centre de Météorologie Spatiale in Lannion for information provided on satellite observations.

Financial support. This work was supported by the French National program Les Enveloppes Fluides et l'Environnement (LEFE), projects
535 ASMA and EXDOMO.

References

- Båserud, L., Lussana, C., Nipen, T. N., Seierstad, I. A., Oram, L., and Aspelién, T.: TITAN automatic spatial quality control of meteorological in-situ observations, *Advances in Science and Research*, 17, 153–163, <https://doi.org/10.5194/asr-17-153-2020>, 2020.
- 540 Beele, E., Reyniers, M., Aerts, R., and Somers, B.: Quality control and correction method for air temperature data from a citizen science weather station network in Leuven, Belgium, *Earth System Science Data*, 14, 4681–4717, <https://doi.org/10.5194/essd-14-4681-2022>, 2022.
- Bell, A., Martinet, P., Caumont, O., Burnet, F., Delanoë, J., Jorquera, S., Seity, Y., and Unger, V.: An optimal estimation algorithm for the retrieval of fog and low cloud thermodynamic and micro-physical properties, *Atmospheric Measurement Techniques*, 15, 5415–5438, <https://doi.org/10.5194/amt-15-5415-2022>, 2022.
- 545 Bell, S., Cornford, D., and Bastin, L.: How good are citizen weather stations? Addressing a biased opinion, *Weather*, 70, 75–84, <https://doi.org/10.1002/wea.2316>, 2015.
- Brousseau, P., Desroziers, G., Bouttier, F., and Chapnik, B.: A posteriori diagnostics of the impact of observations on the AROME-France convective-scale data assimilation system, *Quarterly Journal of the Royal Meteorological Society*, 140, 982–994, <https://doi.org/10.1002/qj.2179>, 2014.
- 550 Brousseau, P., Seity, Y., Ricard, D., and Léger, J.: Improvement of the forecast of convective activity from the AROME-France system, *Quarterly Journal of the Royal Meteorological Society*, 142, 2231–2243, <https://doi.org/10.1002/qj.2822>, 2016.
- Büchau, Y. G.: Modelling Shielded Temperature Sensors: an Assessment of the Netatmo Citizen Weather Station, MSc Thesis, University of Hamburg https://bis-erdsystem.de/fileadmin/user_upload/bise/Texte/MSc_Met_61-20180313.pdf (last access: 15 April 2024), 2018.
- Cardinali, C., Andersson, E., Viterbo, P., Thépaut, J.-N., and Vasiljevic, D.: Use of conventional surface observations in three-dimensional variational assimilation, *ECMWF Technical Memoranda*, p. 31, <https://doi.org/10.21957/p6p470k5x>, 1994.
- 555 Caumont, O., Cimini, D., Löhnert, U., Alados-Arboledas, L., Bleisch, R., Buffa, F., Ferrario, M. E., Haeefe, A., Huet, T., Madonna, F., and Pace, G.: Assimilation of humidity and temperature observations retrieved from ground-based microwave radiometers into a convective-scale NWP model, *Quarterly Journal of the Royal Meteorological Society*, 142, 2692–2704, <https://doi.org/10.1002/qj.2860>, 2016.
- Clark, M. R., Webb, J. D. C., and Kirk, P. J.: Fine-scale analysis of a severe hailstorm using crowd-sourced and conventional observations, *Meteorological Applications*, 25, 472–492, <https://doi.org/10.1002/met.1715>, 2018.
- 560 Coney, J., Pickering, B., Dufton, D., Lukach, M., Brooks, B., and Neely III, R. R.: How useful are crowdsourced air temperature observations? An assessment of Netatmo stations and quality control schemes over the United Kingdom, *Meteorological Applications*, 29, e2075, <https://doi.org/https://doi.org/10.1002/met.2075>, 2022.
- Cornes, R. C., Dirksen, M., and Sluiter, R.: Correcting citizen-science air temperature measurements across the Netherlands for short wave radiation bias, *Meteorological Applications*, 27, <https://doi.org/10.1002/met.1814>, 2019.
- 565 Courtier, P., Andersson, E., Heckley, W., Vasiljevic, D., Hamrud, M., Hollingsworth, A., Rabier, F., Fisher, M., and Pailleux, J.: The ECMWF implementation of three-dimensional variational assimilation (3D-Var). I: Formulation, *Quarterly Journal of the Royal Meteorological Society*, 124, 1783–1807, <https://doi.org/10.1002/qj.49712455002>, 1998.
- Demortier, A., Mandement, M., Pourret, V., and Caumont, O.: Assimilation of surface pressure observations from personal weather stations in AROME-France, *Natural Hazards and Earth System Sciences*, 24, 907–927, <https://doi.org/10.5194/nhess-24-907-2024>, 2024.
- 570 Derrien, M., Le Gleau, H., and Raoul, M.-P.: The use of the high resolution visible in SAFNWC/MSG cloud mask, <https://meteofrance.hal.science/meteo-00604325/>, 2010.

- Desroziers, G., Berre, L., Chapnik, B., and Poli, P.: Diagnosis of observation, background and analysis-error statistics in observation space, *Quarterly Journal of the Royal Meteorological Society*, 131, 3385–3396, <https://doi.org/10.1256/qj.05.108>, 2005.
- 575 Dumas, G., Masson, V., Hidalgo, J., Edouart, V., Hanna, A., and Poujol, G.: Co-construction of climate services based on a weather stations network: Application in Toulouse agglomeration local authority, *Climate Services*, 24, 100 274, <https://doi.org/10.1016/j.cliser.2021.100274>, 2021.
- ECMWF: IFS Documentation CY48R1 - Part I: Observations, chap. 1, ECMWF, <https://doi.org/10.21957/0f360ba4ca>, 2023.
- Eumetsat: Hourly Surface Solar Irradiance (netCDF) - MSG, <https://user.eumetsat.int/catalogue/EO:EUM:DAT:MSG:OSI-304-H/overview> (last access: 31 Mars 2024), 2017.
- 580 Feichtinger, M., de Wit, R., Goldenits, G., Kolejka, T., Hollósi, B., Žuvela-Aloise, M., and Feigl, J.: Case-study of neighborhood-scale summertime urban air temperature for the City of Vienna using crowd-sourced data, *Urban Climate*, 32, 100 597, <https://doi.org/10.1016/j.uclim.2020.100597>, 2020.
- Fenner, D., Bechtel, B., Demuzere, M., Kittner, J., and Meier, F.: CrowdQC+—A Quality-Control for Crowdsourced Air-
 585 Temperature Observations Enabling World-Wide Urban Climate Applications, *Frontiers in Environmental Science*, 9, <https://doi.org/10.3389/fenvs.2021.720747>, 2021.
- Flamant, C., Chazette, P., Caumont, O., Di Girolamo, P., Behrendt, A., Sicard, M., Totems, J., Lange, D., Fourrié, N., Brousseau, P., Augros, C., Baron, A., Cacciani, M., Comerón, A., De Rosa, B., Ducrocq, V., Genau, P., Labatut, L., Muñoz-Porcar, C., Rodríguez-Gómez, A., Summa, D., Thundathil, R., and Wulfmeyer, V.: A network of water vapor Raman lidars for improving heavy precipita-
 590 tion forecasting in southern France: introducing the WaLiNeAs initiative, *Bulletin of Atmospheric Science and Technology*, 2, 1–10, <https://doi.org/10.1007/s42865-021-00037-6>, 2021.
- Garcelon, L., Brana, S., Didez, J., and Ameye, F.: StatIC : un réseau d’observation au service d’une information de proximité et de vulgarisation scientifique., *La Météorologie*, 122, 37–46, <https://doi.org/10.37053/lameteorologie-2023-0071>, 2023.
- Giard, D. and Bazile, E.: Implementation of a New Assimilation Scheme for Soil and Surface Variables in a Global NWP Model, *Monthly
 595 Weather Review*, 128, 997–1015, [https://doi.org/10.1175/1520-0493\(2000\)128<0997:IOANAS>2.0.CO;2](https://doi.org/10.1175/1520-0493(2000)128<0997:IOANAS>2.0.CO;2), 2000.
- Guillet, O., Weaver, A. T., Vasseur, X., Michel, Y., Gratton, S., and Gürol, S.: Modelling spatially correlated observation errors in variational data assimilation using a diffusion operator on an unstructured mesh, *Quarterly Journal of the Royal Meteorological Society*, 145, 1947–1967, 2019.
- Gustafsson, N., Janjić, T., Schraff, C., Leuenberger, D., Weissmann, M., Reich, H., Brousseau, P., Montmerle, T., Wattrelot, E., Bučánek, A., Mile, M., Hamdi, R., Lindskog, M., Barkmeijer, J., Dahlbom, M., Macpherson, B., Ballard, S., Inverarity, G., Carley, J., Alexander, C., Dowell, D., Liu, S., Ikuta, Y., and Fujita, T.: Survey of data assimilation methods for convective-scale numerical weather prediction at
 600 operational centres, *Quarterly Journal of the Royal Meteorological Society*, 144, 1218–1256, <https://doi.org/10.1002/qj.3179>, 2018.
- Hintz, K. S., Vedel, H., and Kaas, E.: Collecting and processing of barometric data from smartphones for potential use in numerical weather prediction data assimilation, *Meteorological Applications*, 26, 733–746, <https://doi.org/10.1002/met.1805>, 2019.
- 605 Infoclimat: Open data - données météo - Infoclimat, <https://www.infoclimat.fr/opendata/> (last access: 16 April 2024), 2024.
- Leroy, M.: Note N° 35B : Classification d’un site, https://www.infoclimat.fr/stations/DSO_35B.pdf (last access: 15 April 2024), 2014.
- Li, J. and Shibata, K.: On the Effective Solar Pathlength, *Journal of the Atmospheric Sciences*, 63, 1365–1373, <https://doi.org/10.1175/JAS3682.1>, 2006.
- Lussana, C.: Titanlib, <https://github.com/metno/titanlib/wiki/First-guess-test>, 2023.

- 610 Mandement, M. and Caumont, O.: Contribution of personal weather stations to the observation of deep-convection features near the ground, *Natural Hazards and Earth System Sciences*, 20, 299–322, <https://doi.org/10.5194/nhess-20-299-2020>, 2020.
- Marquès, E.: Étude à fine échelle de l’îlot de chaleur urbain par modélisation bayésienne à partir de données opportunes, Ph.D. thesis, Institut National Polytechnique de Toulouse, <http://www.theses.fr/2023INPT0016>, 2023.
- Masson, V., Le Moigne, P., Martin, E., Faroux, S., Alias, A., Alkama, R., Belamari, S., Barbu, A., Boone, A., Bouysse, F., Brousseau, P.,
615 Brun, E., Calvet, J.-C., Carrer, D., Decharme, B., Delire, C., Donier, S., Essaouini, K., Gibelin, A.-L., Giordani, H., Habets, F., Jidane, M., Kerdraon, G., Kourzeneva, E., Lafaysse, M., Lafont, S., Lebeaupin Brossier, C., Lemonsu, A., Mahfouf, J.-F., Marguinaud, P., Mokhtari, M., Morin, S., Pigeon, G., Salgado, R., Seity, Y., Taillefer, F., Tanguy, G., Tulet, P., Vincendon, B., Vionnet, V., and Voltaire, A.: The SURFEXv7.2 land and ocean surface platform for coupled or offline simulation of earth surface variables and fluxes, *Geoscientific Model Development*, 6, 929–960, <https://doi.org/10.5194/gmd-6-929-2013>, 2013.
- 620 Meier, F., Fenner, D., Grassmann, T., Otto, M., and Scherer, D.: Crowdsourcing air temperature from citizen weather stations for urban climate research, *Urban Climate*, 19, 170–191, <https://doi.org/10.1016/j.uclim.2017.01.006>, 2017.
- Michel, Y. and Brousseau, P.: A Square-Root, Dual-Resolution 3D-EnVar for the AROME Model: Formulation and Evaluation on a Summer-time Convective Period, *Monthly Weather Review*, 149, 3135–3153, <https://doi.org/10.1175/MWR-D-21-0026.1>, 2021.
- Montmerle, T., Michel, Y., Arbogast, E., Ménétrier, B., and Brousseau, P.: A 3D ensemble variational data assimilation scheme for the
625 limited-area AROME model: Formulation and preliminary results, *Quarterly Journal of the Royal Meteorological Society*, 144, 2196–2215, <https://doi.org/10.1002/qj.3334>, 2018.
- Météo-France: Données de modèle AROME, https://donneespubliques.meteofrance.fr/?fond=produit&id_produit=131&id_rubrique=51 (last access: 5 April 2024), 2024a.
- Météo-France: Observations du réseau sol de France, https://donneespubliques.meteofrance.fr/?fond=produit&id_produit=93&id_rubrique=32 (last access: 5 April 2024), 2024b.
630
- Météo-France: Observations en mer (messages SHIP, BUOY), https://donneespubliques.meteofrance.fr/?fond=produit&id_produit=95&id_rubrique=32 (last access: 5 April 2024), 2024c.
- Météo-France: Données SYNOP essentielles OMM, https://donneespubliques.meteofrance.fr/?fond=produit&id_produit=90&id_rubrique=32 (last access: 5 April 2024), 2024d.
- 635 Netatmo: Weather API Documentation, <https://dev.netatmo.com/apidocumentation/weather> (last access: 5 April 2024), 2024.
- Nipen, T. N., Seierstad, I. A., Lussana, C., Kristiansen, J., and Hov, Ø.: Adopting Citizen Observations in Operational Weather Prediction, *Bulletin of the American Meteorological Society*, 101, E43–E57, <https://doi.org/10.1175/BAMS-D-18-0237.1>, 2020.
- Orlanski, I.: A Rational Subdivision of Scales for Atmospheric Processes, *Bulletin of the American Meteorological Society*, 56, 527–530, <http://www.jstor.org/stable/26216020>, 1975.
- 640 Pergaud, J., Masson, V., Malardel, S., and Couvreux, F.: A Parameterization of Dry Thermals and Shallow Cumuli for Mesoscale Numerical Weather Prediction., *Boundary-Layer Meteorol.*, 132, 83–106, <https://doi.org/10.1007/s10546-009-9388-0>, 2009.
- Potgieter, J., Nazarian, N., Lipson, M. J., Hart, M. A., Ulpiani, G., Morrison, W., and Benjamin, K.: Combining High-Resolution Land Use Data With Crowdsourced Air Temperature to Investigate Intra-Urban Microclimate, *Frontiers in Environmental Science*, 9, <https://doi.org/10.3389/fenvs.2021.720323>, 2021.
- 645 Pourret, V., Šavli, M., Mahfouf, J.-F., Raspaud, D., Doerenbecher, A., Bénichou, H., and Payan, C.: Operational assimilation of Aeolus winds in the Météo-France global NWP model ARPEGE, *Quarterly Journal of the Royal Meteorological Society*, 148, 2652–2671, <https://doi.org/10.1002/qj.4329>, 2022.

- Préaux, D., Dombrowski-Etchevers, I., Gouttevin, I., and Seity, Y.: Study of the 2 m temperature bias of the numerical weather forecasting model Arome over the French Alps, in: EGU General Assembly Conference Abstracts, pp. EGU–7481, <https://doi.org/10.5194/egusphere-egu23-7481>, 2023.
- 650 Sassi, M. Z., Fourrié, N., Guidard, V., and Birman, C.: Use of Infrared Satellite Observations for the Surface Temperature Retrieval over Land in a NWP Context, *Remote Sensing*, 11, 2371, <https://doi.org/10.3390/rs11202371>, 2019.
- Schmetz, J., Pili, P., Tjemkes, S., Just, D., Kerkmann, J., Rota, S., and Ratier, A.: An Introduction to Meteosat Second Generation (MSG), *Bulletin of the American Meteorological Society*, 83, 977–992, [https://doi.org/10.1175/1520-0477\(2002\)083<0977:AITMSG>2.3.CO;2](https://doi.org/10.1175/1520-0477(2002)083<0977:AITMSG>2.3.CO;2), 2002.
- 655 Seity, Y., Brousseau, P., Malardel, S., Hello, G., Bénard, P., Bouttier, F., Lac, C., and Masson, V.: The AROME-France Convective-Scale Operational Model, *Monthly Weather Review*, 139, 976–991, <https://doi.org/10.1175/2010MWR3425.1>, 2011.
- Sgoff, C., Acevedo, W., Paschalidi, Z., Ulbrich, S., Bauernschubert, E., Kratzsch, T., and Potthast, R.: Assimilation of crowd-sourced surface observations over Germany in a regional weather prediction system, *Quarterly Journal of the Royal Meteorological Society*, 148, 1752–1767, <https://doi.org/10.1002/qj.4276>, 2022.
- 660 Stoffelen, A.: Toward the true near-surface wind speed: Error modeling and calibration using triple collocation, *Journal of Geophysical Research: Oceans*, 103, 7755–7766, <https://doi.org/10.1029/97JC03180>, 1998.
- Stull, R. B.: *Mean Boundary Layer Characteristics*, pp. 1–27, Springer Netherlands, Dordrecht, ISBN 978-94-009-3027-8, https://doi.org/10.1007/978-94-009-3027-8_1, 1988.
- 665 Tardieu, J. and Leroy, M.: Radome, le réseau temps réel d’observation au sol de Météo-France, *La Météorologie*, pp. 40–43, <https://doi.org/10.4267/2042/36262>, 2003.
- Toulouse Métropole: Portail d’ouverture des données publiques de Toulouse Métropole — Open Data Toulouse Métropole, <https://data.toulouse-metropole.fr/explore/?sort=modified&refine.keyword=m%C3%A9t%C3%A9o> (last access: 16 April 2024), 2024.
- Varentsov, M. I., Konstantinov, P. I., Shartova, N. V., Samsonov, T. E., Kargashin, P. E., Varentsov, A. I., Fenner, D., and Meier, F.: Urban heat island of the Moscow megacity: the long-term trends and new approaches for monitoring and research based on crowdsourcing data, *IOP Conference Series: Earth and Environmental Science*, 606, 012 063, <https://doi.org/10.1088/1755-1315/606/1/012063>, 2020.
- 670 Vasiljevic, D., Cardinali, C., and Undén, P.: ECMWF 3D variational data assimilation of conventional observations, in: *Workshop on variational assimilation, with special emphasis on three-dimensional Aspects*, 9–12 November 1992, pp. 389–436, ECMWF, Shinfield Park, Reading, https://www.ecmwf.int/sites/default/files/elibrary/1992/76806-ecmwf-3d-variational-data-assimilation-conventional-observations_0.pdf, 1992.
- 675 Venter, Z. S., Brousse, O., Esau, I., and Meier, F.: Hyperlocal mapping of urban air temperature using remote sensing and crowdsourced weather data, *Remote Sensing of Environment*, 242, 111 791, <https://doi.org/10.1016/j.rse.2020.111791>, 2020.
- Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat, 680 Í., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., and SciPy 1.0 Contributors: *SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python*, *Nature Methods*, 17, 261–272, <https://doi.org/10.1038/s41592-019-0686-2>, 2020.
- Vural, J., Merker, C., Löffler, M., Leuenberger, D., Schraff, C., Stiller, O., Schomburg, A., Knist, C., Haeefe, A., and Hervo, M.: Improving the representation of the atmospheric boundary layer by direct assimilation of ground-based microwave radiometer observations, *Quarterly Journal of the Royal Meteorological Society*, 150, 1012–1028, <https://doi.org/10.1002/qj.4634>, 2024.
- 685

World Meteorological Organization: Guide to Instruments and Methods of Observation, Volume I – Measurement of Meteorological Variables (WMO-No. 8), World Meteorological Organization, ISBN 9789263100085, <https://library.wmo.int/idurl/4/68695>, 2023.

Zumwald, M., Knüsel, B., Bresch, D. N., and Knutti, R.: Mapping urban temperature using crowd-sensing data and machine learning, *Urban Climate*, 35, 100739, <https://doi.org/10.1016/j.uclim.2020.100739>, 2021.

4.3 Complément de l'article

Les observations au sol sont cruciales pour contraindre les champs du modèle dans la couche limite atmosphérique. Pour avoir confiance en ces observations, un prétraitement strict est indispensable. Dans ce contexte, nous explorerons le suivi opérationnel du contrôle des observations à Météo-France permettant de *blacklister* les observations anormales. Enfin, deux cas d'étude supplémentaires sont examinés pour comprendre comment les observations PWS peuvent contraindre l'analyse et la prévision d'un îlot de chaleur urbain et d'une ligne orageuse.

4.3.1 Filtres des observations de surface

Le rejet des observations anormales dépend du choix de la référence. Jusqu'à présent, l'ébauche d'AROME a été choisie comme référence, bien qu'elle présente à la fois des avantages et des inconvénients. L'ébauche d'AROME est un champ spatialisé représentant l'état de l'atmosphère à une résolution de 1.3 km. Le modèle est capable de représenter des phénomènes de méso-échelle et s'adapte aux reliefs. Il serait aussi possible de prendre comme référence une analyse produite uniquement avec des observations, n'introduisant pas d'erreur modèle, mais, avec une observation de référence tous les 10 km en moyenne en France, celle-ci possède une résolution horizontale plus grossière que le modèle. Des produits de fusion de données existent, mêlant observations et modèles, comme le produit APROFUS, mais celui-ci peut mêler les désavantages des deux références précédentes (Mandement, 2020). L'ébauche du modèle AROME est choisie par les systèmes de suivi des observations (phase de *monitoring*) pour contrôler les observations en temps réel. Elle permet de créer une "liste noire" rassemblant l'ensemble des mesures de stations, d'un ou de plusieurs paramètres, considérées anormales qui ne sont pas assimilées dans les modèles de PNT. Cette liste noire est basée sur un rejet des observations dépendant d'un seuil sur leur nombre d'erreurs grossières, leur biais ou leur rms (Table 4.1). En guise d'exemple, une liste noire est ici testée a posteriori du travail de l'article précédent, après avoir réalisé le pré-traitement des observations PWS.

TABLEAU 4.1 – Statistiques des écarts à l'ébauche des observations de SWS et PWS. Pour chaque réseau d'observations de température (resp. d'humidité relative), *nb* indique le nombre de stations, *GE* le nombre d'observations dont l'écart à l'ébauche est supérieurs à 15 °C (resp. à 80 %), *biais* le nombre de stations avec un biais mensuel supérieur à 4 °C (resp. à 25 %), et *rms* supérieur à 6 °C (resp. à 40 %), à chaque fois sur l'ensemble du mois d'étude. Ces seuils correspondent aux seuils utilisés en opérationnel. Les PWS pré-traitées correspondent aux observations qui ont passé le pré-traitement P-S (section 4.2).

		<i>nb</i>	<i>GE</i>	<i>biais</i>	<i>rms</i>
Seuils			>15°C	>4°C	>6°C
T	SWS	2603	3	1	0
	PWS brutes	61713	55591	1844	1363
	PWS pré-traitées	20247	114	0	0
Seuils			>80 %	>25 %	>40 %
RH	SWS	1652	3	1	0
	PWS brutes	61713	1030	161	11
	PWS pré-traitées	22795	114	0	0

Sur la période d'étude, seule une SWS présente un biais supérieur à 4 °C, à savoir la

station de l'aiguille du Midi à Chamonix, avec un biais de 4.1 °C, située à 3845 m d'altitude. Pour les observations PWS brutes, 1958 stations seraient considérées anormales, c'est-à-dire rejetée pour au moins un des trois critères (Table 4.1), soit 3 % de stations PWS. Après traitement, toutes les observations seraient admises. Le pré-traitement des observations de PWS (correction de biais et contrôle de qualité) soustrait le biais des observations par rapport au modèle et rejette les observations éloignées de l'ébauche du modèle, mais il peut néanmoins laisser passer ponctuellement des observations très éloignées de l'ébauche, comme en témoignent les 114 observations présentant des erreurs grossières.

Le désavantage de l'utilisation de l'ébauche comme référence pour le contrôle de qualité intervient si le modèle possède des biais systématiques, ce qui peut entraîner une dégradation de l'observation à travers la correction de biais, voir un mauvais rejet des observations PWS. Par exemple, dans des zones topographiques complexes, telles que les régions de montagne, l'ébauche du modèle peut avoir un biais en température par rapport aux observations de référence (Préaux *et al.*, 2023).

Une solution pour ne pas dépendre du modèle peut consister à prendre les observations voisines comme référence, ce qui est fait par le contrôle de qualité Titan. On l'a montré, cette méthode peut aussi être dangereuse si l'ensemble des observations voisines admettent un biais par rapport au modèle. D'autres méthodes explorées par la littérature existent, comme la multi-fidélité (Peherstorfer *et al.*, 2018) ou une classification de courbes (*curves clustering*, Marques, 2023). Cette dernière méthode est présentée dans la figure 4.2.

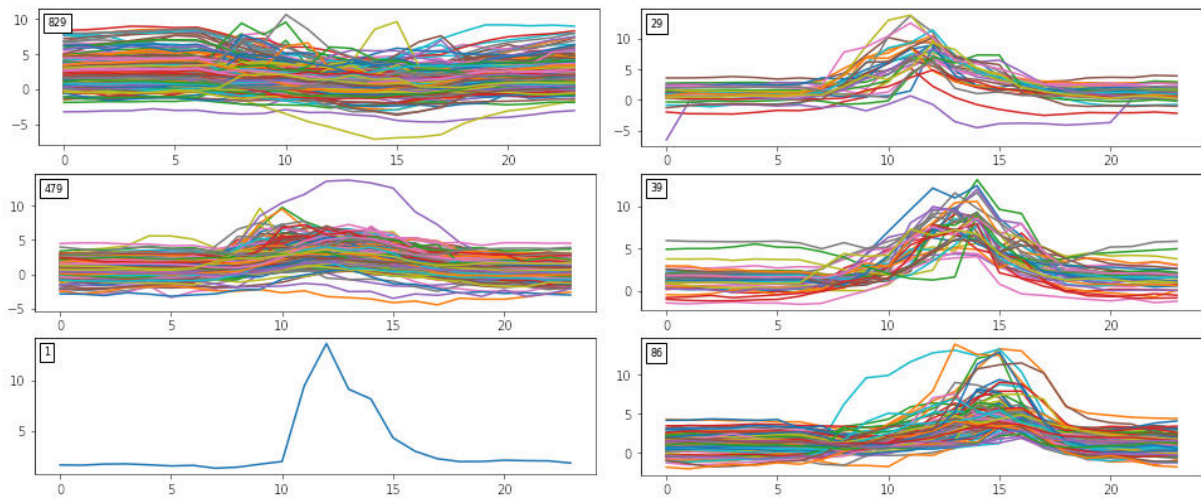


FIGURE 4.2 – Illustration de la méthode de classification de courbes en six classes. Celle-ci consiste en une classification non supervisée par la méthode des *k-means* pour obtenir une discrimination des courbes d'observations de température des PWS. La discrimination est basée sur une décomposition en ondelettes des anomalies de séries temporelles. Les anomalies sont obtenues en soustrayant pour chaque heure et pour chaque observation le quantile 10 % de l'ensemble des observations PWS. Voir Marques (2023) pour plus de détails.

Ce type de classification permet de regrouper des stations qui ont des biais journaliers, par rapport au modèle, comparables entre elles. Ici, une classification est réalisée en un nombre arbitraire de six classes de stations de mesures de température sur un sous-domaine centré sur la ville de Toulouse comprenant 1463 stations. Les deux premières classes regroupent les stations présentant les anomalies journalières les plus constantes. La décomposition forme ensuite trois classes (colonne de droite de la figure 4.2) qui suivent un écart OmB positif, respectivement le matin, le midi, et dans l'après-midi. Il serait

alors possible d'enlever les observations de ces stations aux heures où elles possèdent un biais positif. Cette classification n'est pas parfaite, et peut être considérée comme semi-automatisée, car elle demande une interprétation des classes et le réglage du nombre de classes.

4.3.2 Cas d'étude : ICU du 17 juin 2023

La ville de Paris, en tant que métropole, est fortement exposée aux îlots de chaleur urbains. Des expériences d'assimilation des observations de température des PWS ont été réalisées le 17 juin 2023 à partir des analyses du réseau de 00:00 UTC, avec le schéma 3DVar et 3DEnVar (figure 4.3). L'îlot de chaleur urbain à 06:00 UTC (8 h locale) est bien visible dans les expériences de contrôle 3DVar et 3DEnVar. Cependant, dans l'expérience 3DEnVar la température est plus froide sur l'ensemble du domaine, ce qui résulte du réglage du schéma d'assimilation accordant moins d'impact aux observations de surface, comme observé dans le centre urbain (zone A).

L'assimilation des observations PWS ne modifie pas substantiellement l'intensité de l'ICU, cependant la forme de la plage de température est plus circulaire en 3DVarT que ce qu'elle était en 3DVar : les observations PWS contraignent l'extension spatiale de l'ICU. L'assimilation des observations de température des PWS avec le schéma 3DEnVar permet de relever la température dans le centre urbain (zone A), laquelle reste encore inférieure à celle obtenue avec le schéma 3DVar qui est plus proche des observations SWS. La zone péri-urbaine B est également influencée par les nouvelles observations, où une augmentation de la température de 2 à 3 °C est visible entre les expériences 3DEnVar et 3DEnVarT. On notera que le réseau d'observations PWS pré-traité est conforme au modèle, même si localement certaines observations vraisemblablement anormales subsistent, comme l'observation de PWS à 23.8 °C dans la zone B. Le contrôle de qualité supplémentaire pour les PWS étant basé sur les six dernières heures, si l'observation devient anormale peu de temps avant l'échéance, celle-ci peut être gardée, et ne sera rejetée que si l'erreur persiste suffisamment longtemps sur une période de 6 h. In fine, cette observation ponctuelle étant proche géographiquement d'autres observations de surface, son assimilation n'influence que peu, à elle seule, le champ de température. De plus, des écarts entre les observations d'un réseau indépendant StatIC et l'ébauche à 2 m ont été calculés. L'analyse qui se rapproche le plus du réseau indépendant est l'analyse 3DEnVarT (figure 4.3d). Toutefois, ce réseau ne comprend que neuf stations sur le domaine, et des statistiques sur d'autres cas d'études seraient à réaliser pour vérifier ce résultat. Sur cette situation, diminuer la taille de l'écrémage des observations PWS, permettrait de garder plus d'observations, notamment dans le centre urbain, pour acquérir plus de détails à l'analyse, tout en étant limité par les dégradations que pourraient entraîner la corrélation des erreurs d'observations trop denses.

4.3.3 Cas d'étude : ligne orageuse le 8 septembre 2021

Une ligne orageuse se forme sur le sud-ouest de la France en soirée du 8 septembre 2021, forcée par une convergence de vent en surface, et se déplace vers le nord-est. La figure 4.4 montre les champs à 5 m de température et d'humidité relative à 18:00 UTC, issus des expériences cyclées depuis le 6 septembre à 00:00 UTC, avec et sans l'assimilation des observations PWS. Dans l'expérience de contrôle 3DEnVar, un refroidissement de plus de 7 °C de la température à 5 m est simulé sous le système orageux. La baisse de température

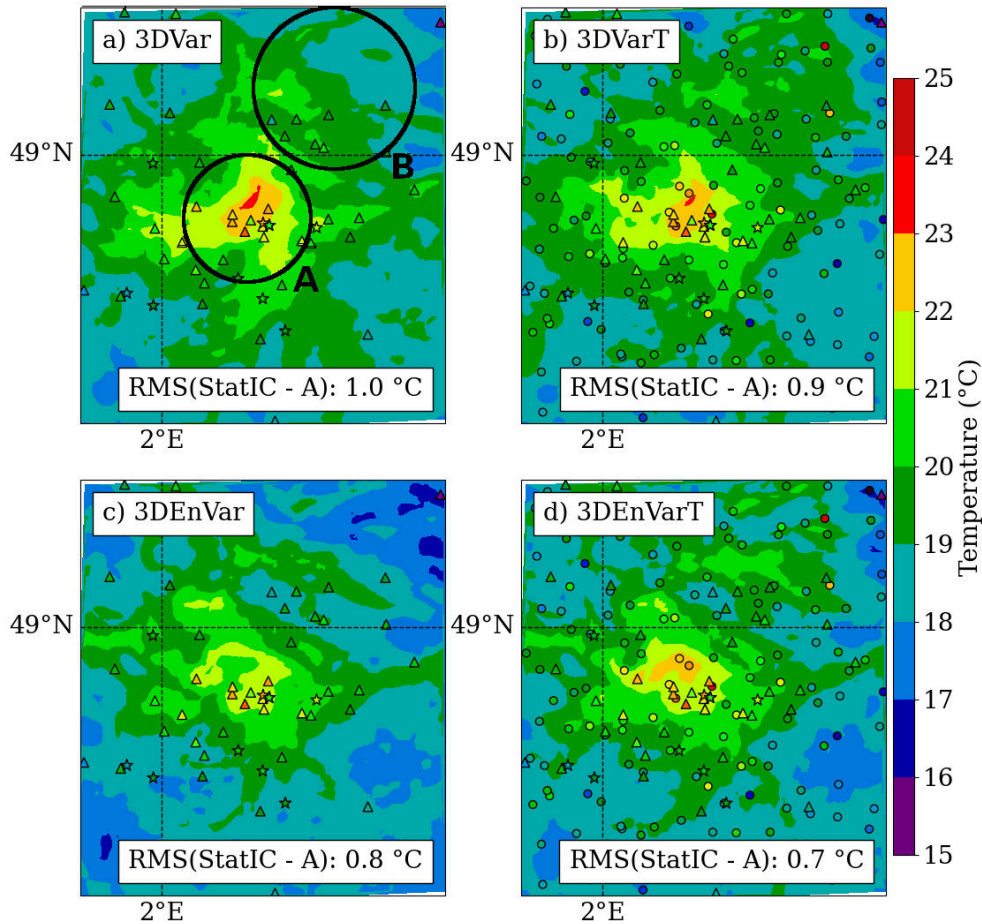


FIGURE 4.3 – Analyse à 06:00 UTC de la température au premier niveau modèle (5 m), de l’îlot de chaleur urbain simulé par AROME le 17 juin 2023, (a) en 3DVar, (b) en 3DVar avec l’assimilation des observations de température des PWS (3DVarT), (c) en 3DEnVar, et (d) en 3DEnVar avec l’assimilation des observations de température des PWS (3DEnVarT). Les expériences sont cyclées et partent d’analyses du réseau de 00:00 UTC. Les observations de températures des PWS ont été pré-traitées suivant P-S comprenant un écrémage à 8 km de résolution. A et B font référence aux deux cercles de (a). Les scores correspondent au rms OmA des neuf stations StatIC indépendantes (étoiles).

est positionnée sous le système observé par le radar. L’analyse de température 3DEnVarTS (figure 4.4b), est quant à elle davantage contrainte avec l’assimilation des observations PWS : le gradient de température est mieux positionné à l’avant et à l’arrière de la ligne orageuse observée. Les observations de température PWS sont cohérentes avec le champ analysé à 5 m, montrant les capacités du schéma d’assimilation 3DEnVar à les prendre en compte.

Il en est de même pour les analyses d’humidité relative (figure 4.4c, d), où la baisse d’humidité relative est mieux positionnée à l’avant et à l’arrière du système orageux dans l’analyse 3DEnVarRHS que dans celle de contrôle 3DEnVar. Cette baisse est notamment visible au sud-ouest de la zone (0.3 °O, 44.7 °N) avec des valeurs baissant de 20 % à 30 % d’humidité relative une fois contraint par les observations PWS. De plus, on remarque que le champ d’humidité relative est nettement réduit sous la ligne orageuse observée. Que ce soit en température ou en humidité relative, les observations PWS sont capables de contraindre les champs associés sur le niveau le plus bas du modèle. D’autant plus

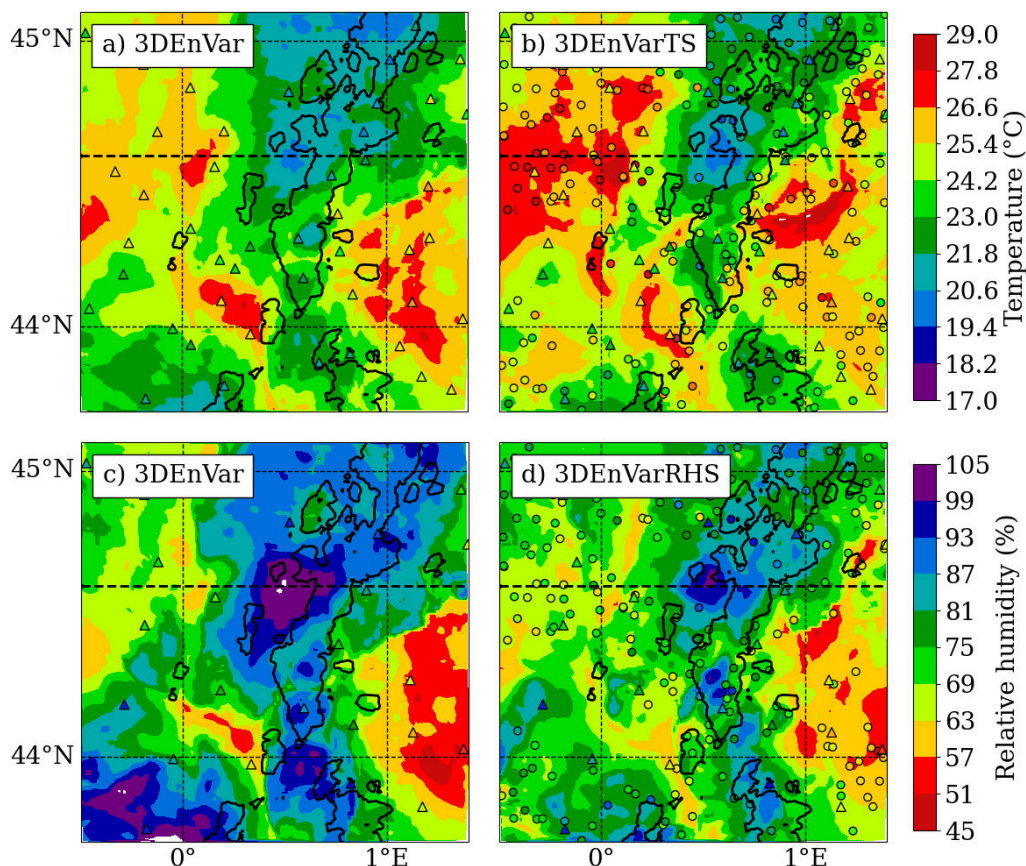


FIGURE 4.4 – Analyses de la ligne orageuse du 8 septembre 2021 à 18:00 UTC (a, b) en température et (c, d) en humidité relative au premier niveau du modèle (5 m). Expériences (a, c) de contrôle (3DEnVar), (b) avec l'assimilation des observations de température PWS (3DEnVarTS), ou (d) d'humidité relative (3DEnVarRHS). Les réflectivités radar égales à 40 dBZ sont représentées par le trait continu. Le trait tireté représente la section verticale réalisée dans la figure 4.5.

qu'il s'agit d'expériences cyclées, et c'est donc grâce à l'assimilation de ces nouvelles observations au fur et à mesure des réseaux successifs qu'il est possible d'arriver à ce résultat.

Une fois les champs de surface décrits, se pose la question du rôle des observations de surface pour contraindre l'analyse sur la verticale. La figure 4.5 montre les coupes verticales d'incrément issues d'expériences additionnelles partant de la même ébauche de référence à 18:00 UTC. Ces coupes correspondent à une section transversale de la ligne orageuse. Les incréments d'analyse sont de l'ordre de 2 °C en température et de 30 % en humidité relative en valeur absolue pour l'expérience de contrôle 3DEnVar (figure 4.5a, d). Si l'on reproduit l'assimilation de données formant l'analyse à 18:00 UTC sans les observations radar (figure 4.5b, e), les incréments sont nettement réduits. Les observations radar sont les observations qui apportent la majeure partie des incréments sur toute la troposphère en matière de température et d'humidité relative, à l'exception de la couche de surface (CLA). Les incréments restants proviennent des observations GPS sol et des observations de surface (non montré), augmentant respectivement l'humidité relative sur toute la troposphère et à la surface, notamment à 0.7 °E, au niveau du système orageux. Enfin, l'assimilation des observations PWS (figure 4.5c, f) apporte des incréments contenus dans la CLA, c'est-à-dire entre 1000 hPa et 900 hPa. En température, les incréments

sont en conformité avec les observations conventionnelles, c'est-à-dire qu'elles réduisent la température à 0.7°E sous le système orageux, et augmentent la température à l'avant et à l'arrière du système orageux. En humidité relative, les observations PWS réduisent faiblement l'humidité relative dessous et à l'avant du système orageux. En définitive, les observations PWS ajoutent des incréments supplémentaires contenus dans la CLA qui sont, pour cette section, cohérents avec le système orageux en place.

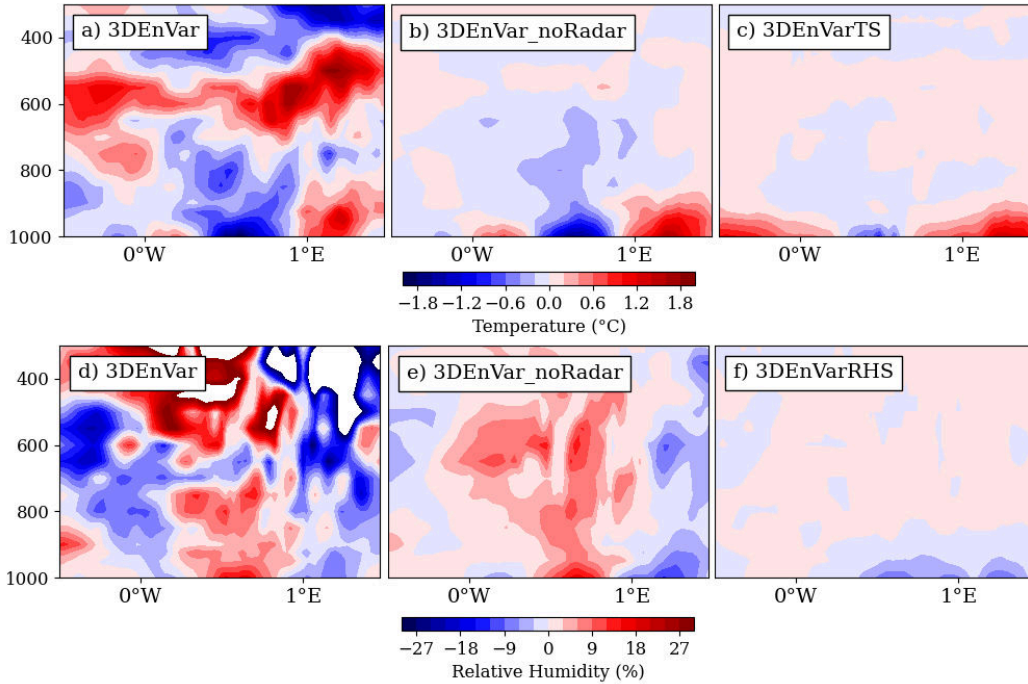


FIGURE 4.5 – Coupes verticales d'incrément (a-c) de température et (d-f) d'humidité relative, partant de la même ébauche. Incréments d'analyse des expériences (a, d) de référence et (b, e) sans les observations radar. Différence d'analyses entre l'expérience assimilant les observations PWS, (c) en température et (f) en humidité relative, et l'analyse de référence. L'axe de la coupe est indiqué en pointillé sur la figure 4.4.

Enfin, des cumuls de précipitations sur 6 h issues des prévisions partant de l'analyse des expériences cyclées de la figure 4.4 sont représentées dans la figure 4.6. Alors que l'expérience 3DEnVar produit une ligne avec un axe méridien de précipitation et un maximum à 97 mm, les expériences assimilant les observations de température et d'humidité relative comportent un maximum plus élevé, de respectivement 113 mm et 110 mm. Ces cumuls sont plus proches des observations de cumul de précipitations qui possède un maximum de précipitations de 131 mm. L'intensité des précipitations est renforcée sur la partie nord du système (à 44.5°N) par l'assimilation des nouvelles observations, au niveau du maximum de précipitation observée. Néanmoins, une cassure dans la partie sud du système est simulée dans les expériences 3DEnVarTS et 3DEnVarRHS (à 43.8°N), en désaccord avec le système orageux observé qui s'étend jusqu'au sud du Gers. L'ajout des nouvelles observations peut à la fois renforcer le système en place, comme c'est le cas dans la partie nord de la ligne orageuse, que déstabiliser les structures existantes dans sa partie sud.

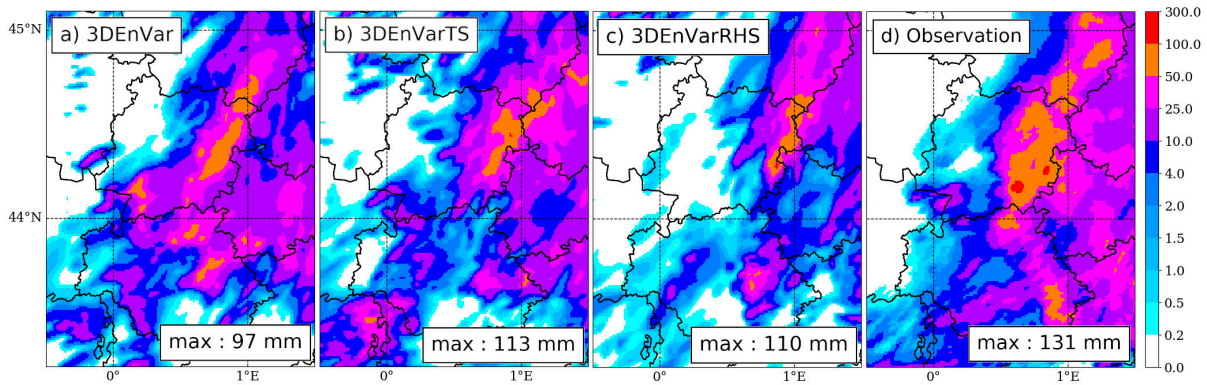


FIGURE 4.6 – Cumul de précipitations sur 6 h (en mm) à partir de l'analyse de 18:00 UTC des expériences cyclées 3DEnVar, 3DEnVarTS et 3DEnVarRHS, ainsi que les observations Antilope.

4.4 Synthèse et perspectives

L'assimilation d'un plus grand nombre d'observations de température et d'humidité relative permet de mieux contraindre les champs de surface à l'analyse. Ces champs sont cruciaux pour décrire les phénomènes de basses couches atmosphériques ainsi que les phénomènes dépendant de la surface, tels que les systèmes convectifs. Des défis restent à relever avant une utilisation opérationnelle. Ces défis concernent (i) le prétraitement des observations, (ii) le réglage des schémas d'assimilation et (iii) l'influence des observations sur les échéances de prévision. Une partie des défis sont similaires avec ceux mentionnés pour les observations PWS de pression ; défis sur lesquelles nous ne reviendrons pas.

(i) Deux prétraitements des observations PWS ont été proposés, P-M et P-S, permettant de réduire le biais diurne des stations et de rejeter les observations considérées anormales. Ces deux prétraitements présentent tout de même des limitations, notamment dû au problème de surapprentissage de l'ébauche du modèle pour corriger les observations. En effet, après correction, les observations PWS peuvent intégrer un biais du modèle par rapport aux observations de référence. Dans la suite le prétraitement P-S a été privilégiée, car avec sa correction de biais sur une courte période de deux jours, c'est celui qui a donné les résultats les plus satisfaisants sur les cas d'études. Pour l'améliorer, deux axes distincts peuvent être envisagés. Pour réduire le problème de surapprentissage, un traitement des observations PWS entièrement basé dans « le monde des observations » peut être envisagé. La correction et le contrôle de qualité se baseraient uniquement sur des analyses réalisées avec les observations standards. Une autre possibilité de prétraitement consiste à rester dans « le monde du modèle », jusque-là privilégié, en corrigeant les observations par une méthode variationnelle (VarBC), comme celle présentée pour les observations de pression des PWS, permettant d'utiliser pleinement l'information contenue par le modèle. Une liste noire de stations présentant des observations anormales pourrait être créée afin d'éliminer d'emblée les observations inappropriées, laquelle pourrait être basée sur une méthode de classification avec une granularité horaire.

Toujours dans le prétraitement, l'étape d'écrémage peut aussi être amenée à être améliorée. Pour certains phénomènes comme l'îlot de chaleur urbain, l'analyse gagnerait à être contrainte par plus d'observations de surface. Un équilibre doit être trouvé, entre réduire l'échelle de l'écrémage pour garder plus d'observations à fine échelle, tout en limitant l'augmentation de la corrélation des erreurs d'observations de ces mêmes observations qui pourrait dégrader la minimisation.

(ii) Le réglage des schémas d'assimilation de données pour l'échelle convective constitue

un axe de recherche à part entière. Des pistes ont été données dans le chapitre précédent pour le schéma 3D EnVar . Le schéma 3D Var s'est montré pertinent pour certains cas d'études, comme celui de l'îlot de chaleur urbain. Mais concernant les scores à l'ébauche du modèle, le schéma 3D Var et l'IO, avec leurs réglages actuels, ne sont pas satisfaisants pour l'assimilation d'observations à méso-échelle. Pour l'IO, un nouveau réglage peut être réalisé, comme cela a été proposé pour l'arrivée de nouvelles observations de surface par Calas (1997), passant du réseau synoptique avec un espacement horizontal moyen de 100 km à un réseau plus dense d'espacement moyen de 40 km. Ce nouveau réglage doit se faire sur trois étapes : (a) en augmentant le nombre d'observations voisines utilisé pour chaque point du modèle, car plus d'observations sont disponibles ; (b) en diminuant la portée horizontale des observations de surface ; (c) en s'assurant un équilibre entre le modèle et les observations avec le réglage des covariances d'erreurs d'observations σ_o et d'ébauche σ_b (c'est-à-dire en gardant un rapport σ_b/σ_o constant). En pratique, le schéma 2D EnVar devrait supplanter le schéma IO dans les années à venir, et l'assimilation des nouvelles observations de surface dans celui-ci devra être testée.

Le réseau d'observations StatIC, en température et en humidité relative, s'est montré de qualité comparable au réseau SWS. Son utilisation en tant que réseau indépendant pourrait servir à évaluer, de manière systématique, la qualité d'une analyse ou d'une prévision sans souffrir de problèmes de corrélations d'erreurs, car celui-ci n'aurait pas été assimilé par le modèle. L'utilisation d'un réseau d'observations indépendant et de qualité suffisante est un problème courant lors de l'évaluation des modèles. À titre de comparaison, les observations de pluviomètres servent d'observations indépendantes, n'étant pas assimilés dans les modèles, pour l'étude des précipitations. Cet usage pourra être repris dans de futures études, et les observations StatIC pourraient, in fine, être assimilées à leur tour.

(iii) Enfin, un dernier axe, se porte sur l'impact des observations de surface au fil des échéances de la prévision. En moyenne, sur la période d'étude d'un mois, les observations PWS amènent une baisse du rms de l'écart des observations de surface à la prévision en température et en humidité relative jusqu'à 3 à 6 h d'échéance. Cette limite d'échéance est plus faible que pour les observations PWS de pression, où il a été montré une baisse jusqu'à 9 h d'échéance de l'écart à la prévision des observations SWS en pression. Pour des cas spécifiques, comme la ligne d'orage du 8 septembre 2021, une meilleure représentation des champs de surface en température et en humidité relative à l'analyse amène une meilleure prévision des structures de précipitations. Des scores sur les prévisions successives précédant un événement de méso-échelle pourront être réalisées dans de futures études. La prévisibilité est un axe crucial pour mieux anticiper les phénomènes à enjeux.

Chapitre 5

Assimilation conjointe de plusieurs types d'observations

Sommaire

5.1	Assimilation conjointe de plusieurs types d'observations	143
5.1.1	Impact en surface	143
5.1.2	Impact en altitude	146
5.1.3	Score indicateur	147
5.1.4	Augmentation du coût de calcul de l'assimilation des observations PWS	148
5.2	Synthèse	149

Après avoir testé l'assimilation des observations des PWS de pression dans le chapitre 3, puis celles de température et d'humidité relative dans le chapitre 4, ce chapitre, qui fait office d'épilogue, démontre les avantages de l'assimilation conjointe de ces observations de surface dans AROME.

5.1 Assimilation conjointe de plusieurs types d'observations

Une expérience assimilant les trois systèmes d'observations de pression, de température et d'humidité relative des PWS, tels que décrits dans les chapitres précédents, a été réalisée pendant le mois d'étude, et sera nommée 3DEnVar_PTRH (Tableau 5.1). Le résultat de l'assimilation conjointe des systèmes d'observation est présenté par des scores à 1 h d'échéance avec des statistiques sur les OmB, jusqu'à 24 h d'échéance avec des statistiques sur les OmF, et grâce à un score indicateur d'AROME qui concerne les précipitations et le vent.

5.1.1 Impact en surface

L'assimilation de chaque type d'observations individuellement, que ce soit de pression, de température ou d'humidité relative, provenant des PWS rapproche la prévision à 1 h d'échéance de la variable correspondant aux observations SWS (Tableau 5.2). L'expérience 3DEnVar_PTRH montre que l'assimilation conjointe des trois types d'observations entraîne une réduction de 7.5 % de l'écart quadratique moyen des OmB des observations de pression des SWS, ainsi qu'une diminution de 0.7 % pour l'humidité relative, sur le domaine France. L'assimilation conjointe correspond au premier ordre à la somme des résultats de l'assimilation des trois types d'observations. Néanmoins, pour chaque paramètre, l'amélioration grâce à l'ajout des nouvelles observations des autres paramètres

TABLEAU 5.1 – Définition et rappel des expériences réalisées sur une période d'un mois du 6 septembre au 5 octobre 2021. Chaque expérience correspond à un cycle d'assimilation horaire et des prévisions à 48 h d'échéance quatre fois par jour.

Observations de PWS	Pression	Température	Humidité relative
Correction de biais	Soustrait la moyenne des OmB sur un mois	Inspiré de Sgoff <i>et al.</i> (2022)	
Contrôle de qualité	Rejet si l'écart-type des OmB sur un mois $> 60 \text{ m}^2 \text{ s}^{-2}$	Inspiré de Mandement et Caumont (2020)	
Écrémage	Grille de 1.3 km de résolution	Grille de 8 km de résolution	
3DEnVarP	✓		
3DEnVarTS		✓	
3DEnVarRHS			✓
3DEnVar_PTRH	✓	✓	✓

TABLEAU 5.2 – ΔRMS OmB (%) de pression de surface, température à 2 m, humidité relative à 2 m et du vent zonal à 10 m provenant des observations SWS sur le mois d'étude. Les valeurs négatives d'une expérience (XP) par rapport à l'expérience 3DEnVar indiquent que les prévisions à 1 h de XP sont plus proches que celles de l'expérience 3DEnVar, des observations SWS. Les valeurs significatives sont en gras.

	Pression de surface		Température		Humidité relative		Vent zonal (U)	
	France	Domaine AROME	France	Domaine AROME	France	Domaine AROME	France	Domaine AROME
3DEnVarP	-10.4	-3.0	0.9	0.6	1.4	0.8	0.1	0.2
3DEnVarTS	3.7	1.3	-0.8	-0.6	-0.2	-0.1	0.1	0.1
3DEnVarRHS	0.7	0.5	0.8	0.5	-1.2	-0.8	0.2	0.1
3DEnVar_PTRH	-7.5	-2.1	0.2	0.2	-0.7	-0.4	0.2	0.4

permet de compenser les dégradations observées dans les simulations mono-paramètre. Par exemple, en humidité relative, la dégradation liée à l'assimilation des observations de pression des PWS, de 1.4% sur la France, est réduite par l'assimilation des observations d'humidité relative des PWS. Autrement dit, pour un paramètre donné du modèle, plus celui-ci est contraint par des observations, moins les covariances croisées de la matrice \mathbf{B} d'observations d'autres paramètres influenceront sur celui-ci. Ce résultat montre l'importance d'assimiler des observations de différents types.

L'assimilation conjointe des observations de PWS a un impact mitigé sur les scores aux prévisions de surface à courte échéance (figure 5.1) :

- L'écart des observations de pression des SWS à la prévision, avec 3DEnVar_PTRH, conserve la diminution significativement déjà démontrée avec 3DEnVarP par rapport à l'expérience de contrôle jusqu'à 9 h d'échéance.
- Le bénéfice sur la prévision à courte échéance de la température est perdu avec 3DEnVar_PTRH par rapport à 3DEnVarTS comme démontrée précédemment par les scores à la prévision à 1 h d'échéance.
- Le bénéfice significatif pour l'humidité relative est maintenu jusqu'à 9 h d'échéance et est plus important avec 3DEnVar_PTRH par rapport à 3DEnVarRHS.
- Une dégradation significative est notée sur la force du vent à l'analyse, et est plus marquée avec 3DEnVar_PTRH. Cette dégradation devient neutre dès la première

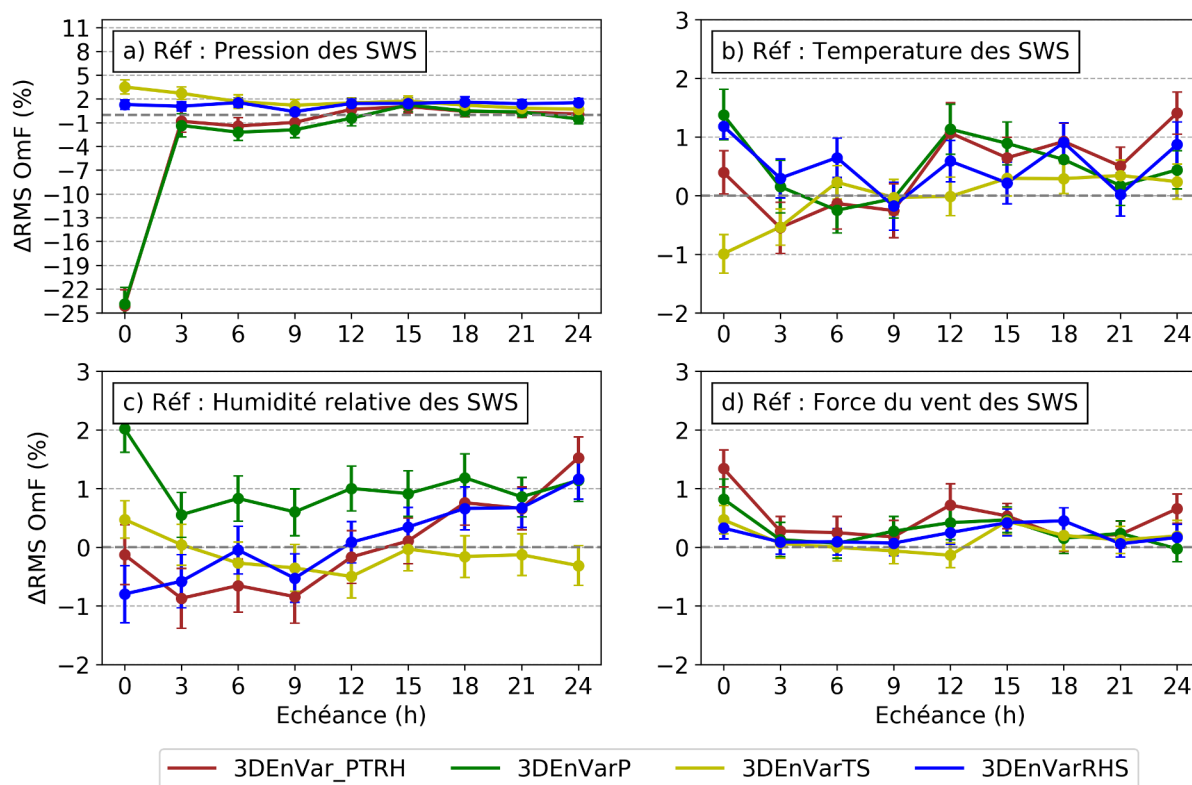


FIGURE 5.1 – RMS de la différence entre les observations de référence (SWS) en France et la prévision pour les expériences 3DEnVar_PTRH, 3DEnVarP, 3DEnVarTS et 3DEnVarRHS, normalisé par l'expérience 3DEnVar. Réalisées, avec 4 prévisions par jour pendant le mois d'étude, pour les observations de (a) pression, (b) de température, (c) d'humidité relative et (d) de force du vent. Les valeurs négatives (resp. positives) indiquent une baisse (resp. une augmentation) de l'écart à la prévision des observations SWS par rapport à l'expérience de contrôle, correspondant à un bénéfice (resp. une dégradation). La barre d'incertitude correspond aux quantiles Q10 et Q90 d'une méthode de bootstrapping.

heure de prévision.

Comme discuté dans les chapitres précédents, l'interprétation des dégradations liées à l'assimilation d'un type d'observation envers les autres paramètres de surface peut provenir de mauvaises covariances croisées des matrices de covariance d'erreur d'ébauche ou d'observations, ces dernières n'étant pas prises en compte.

De plus, des dégradations significatives adviennent après 12 h d'échéance pour l'expérience 3DEnVar_PTRH en température et en humidité relative. Il est possible qu'ajouter les observations à la surface, contraignant l'analyse seulement dans les basses couches, déstabilise les équilibres verticaux du modèle à l'état initial des prévisions et engendre des dégradations au bout de quelques heures de prévisions. Sous cette hypothèse, il conviendrait de réduire le choc à l'analyse de l'assimilation de ces trois types d'observations PWS, en diminuant l'impact de ces observations denses au sol à l'analyse, ce qui peut être réalisé en augmentant leurs erreurs d'observations ou en augmentant l'écrémage.

Enfin, la figure 5.2 présente les scores à la surface, séparés pour les quatre réseaux de prévisions, lors de l'assimilation conjointe des observations PWS. Ces scores de surface admettent une variabilité entre les réseaux. Pour la pression de surface le bénéfice est perceptible uniquement à l'analyse pour le réseau de 06:00 UTC, alors que celui-ci est visible jusqu'à 18 h d'échéance pour le réseau de 18:00 UTC. Globalement, le réseau avec

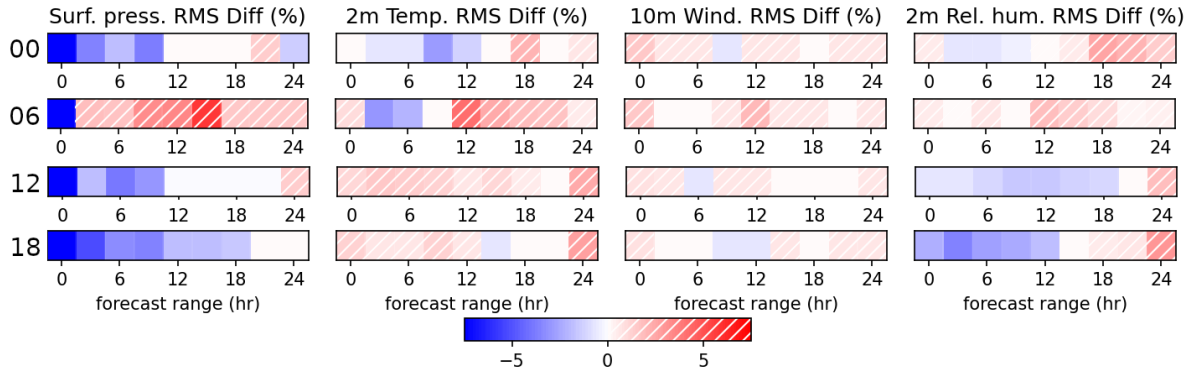


FIGURE 5.2 – RMS de la différence entre les observations de référence (SWS) en France et la prévision issue de l'expérience 3D`EnVar_PTRH` (normalisé par rapport à l'expérience de contrôle), pour les quatre réseaux de prévisions de 00:00 UTC, 06:00 UTC, 12:00 UTC et 18:00 UTC. Pour la pression de surface, la température à 2 m, la force du vent à 10 m et l'humidité spécifique à 2 m. Les valeurs négatives (resp. positives) indiquent une baisse (resp. une augmentation) de l'écart à la prévision de la référence, correspondant à un bénéfice (resp. une dégradation).

le moins d'impact positif est le réseau de 06:00 UTC où l'assimilation conjointe amène une dégradation après 12 h d'échéance pour les quatre paramètres. Alors que les réseaux de 12:00 UTC et 18:00 UTC amènent une amélioration de la pression de surface et de l'humidité spécifique à 2 m jusqu'au moins 9 h d'échéance. Cette disparité entre les réseaux peut s'expliquer par une disparité de qualité entre les analyses liée à des proportions différentes de systèmes d'observations à chaque réseau (par exemple, les observations d'avions sont émises principalement le jour et non la nuit). La disparité entre les réseaux pourrait aussi s'expliquer par une qualité des observations PWS qui peut varier suivant l'heure de la journée. Ces deux hypothèses nécessiteraient des investigations supplémentaires pour être vérifiées.

5.1.2 Impact en altitude

Les scores en altitude sont donnés en comparant la prévision issue de l'expérience assimilant conjointement les observations de PWS à une référence sur la verticale (figure 5.3). Deux références d'altitude sont utilisées : l'analyse de contrôle AROME (3D`Var`), qui correspond à l'état de l'atmosphère le plus proche de la réalité dont on a accès, et les observations de radiosondages. Ces deux références comportent des limites : l'analyse AROME peut avoir des biais dans la couche limite, alors qu'en altitude elle est considérée plus fiable, car liée à la situation synoptique ; les observations de radiosondages sont peu nombreuses, car n'y a que cinq lancés en France qui ont lieu deux fois par jour. Néanmoins, des résultats peuvent être déduits des scores en altitude pour le réseau de 18:00 UTC, là où l'impact est le plus important à la surface :

- En géopotentiel, l'écart entre la prévision de 3D`EnVar_PTRH` et les observations de radiosondages diminue sur toute la verticale, de manière significative au dessus de 500 hPa, jusqu'à 12 h d'échéance. Ce résultat est cohérent avec les scores issus de l'expérience assimilant seulement les observations de pression des PWS (chapitre 3).
- En température, un bénéfice est visible en haute troposphère à 300 hPa jusqu'à 12 h d'échéance avec pour référence les analyses AROME et les radiosondages.
- Pour la vitesse du vent, l'impact est peu discernable sur la verticale vis-à-vis des analyses AROME et disparate vis-à-vis des observations de radiosondages. Ce résultat

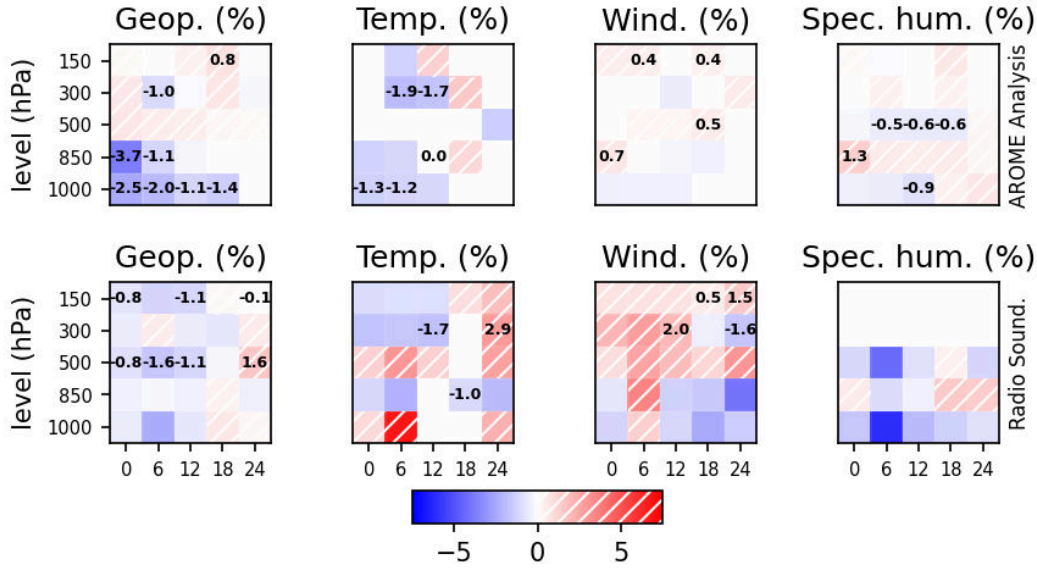


FIGURE 5.3 – RMS de la différence entre la prévision issue de l’expérience 3DEnVar_PTRH et une référence d’altitude (normalisé par rapport à l’expérience de contrôle), pour le réseau de 18:00 UTC. La référence d’altitude correspond (en haut) à l’analyse AROME et (en bas) aux observations de radiosondages. Voir figure 5.2 pour plus de détails.

est cohérent avec ceux obtenus à la surface.

- Pour l’humidité spécifique, un bénéfice est visible sur le bas de la troposphère jusqu’à 500 hPa, et jusqu’à 12 h d’échéance.

En résumé, l’assimilation conjointe des observations de surface des PWS à un impact en altitude pour le géopotential, et plus faible pour la température et l’humidité relative, jusqu’à 12 h d’échéance en moyenne sur le mois d’étude. Ces résultats en altitude restent faibles comparés à l’impact que peuvent avoir des observations d’altitude comme les observations d’avions (Pourret *et al.*, 2022) ou satellitaires (Coopmann *et al.*, 2023).

5.1.3 Score indicateur

L’intérêt particulier du modèle de prévision AROME réside dans sa capacité à anticiper les événements convectifs, notamment ceux caractérisés par des précipitations intenses ou de fortes rafales. Le score de performance de Brier (*Brier skill score*, Amodei *et al.*, 2015) a été calculé pour des seuils de rafales et de cumuls de précipitations (figure 5.4). On l’a vu, l’expérience 3DEnVarP présente des scores supérieurs aux expériences de contrôle 3DEnVar et 3DVar. Bien que l’assimilation conjointe (3DEnVar_PTRH) semble aider pour la prévision de faibles cumuls de précipitations, c.-à-d. supérieurs à 0.5 mm sur 6 h, une légère dégradation est notée pour des cumuls plus élevés, supérieurs à 2 mm et 5 mm sur 6 h. Le score total montre tout de même une amélioration par rapport à l’expérience 3DEnVar de contrôle, passant de 0.827 à 0.829. L’augmentation étant faible, il serait pertinent de prolonger la période d’étude et d’adapter le seuil d’écrémage. Par ailleurs, l’ajout de nouvelles observations dans le cycle d’assimilation-prévision représente un coût numérique supplémentaire qu’il est nécessaire d’évaluer, et qui fera l’objet de la prochaine section.

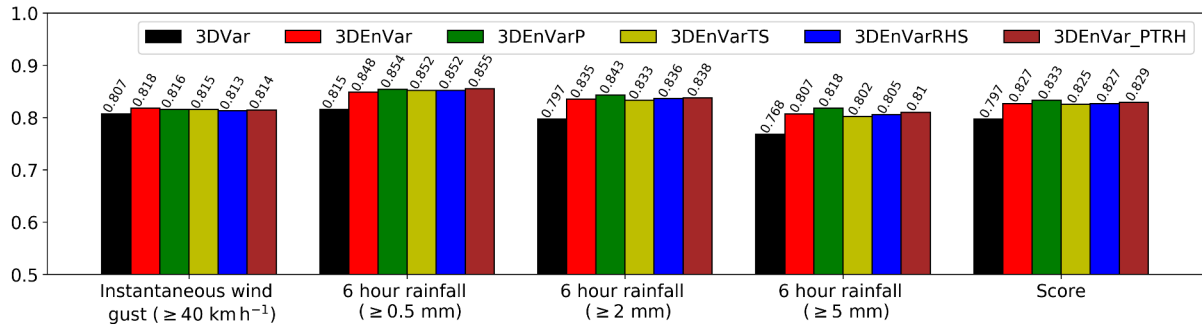


FIGURE 5.4 – Scores de performances de Brier (BSS) pour les prévisions de 00:00 UTC durant le mois d'étude, moyenné sur quatre échéances (6 h, 12 h, 18 h et 24 h), pour les expériences 3DVar, 3DEnVar, 3DEnVarP, 3DEnVarTS, 3DEnVarRHS et 3DEnVar_PTRH. Réalisés avec une échelle spatiale de 50 km, pour les rafales 1 min supérieures à 40 km h^{-1} , le cumul de précipitations sur 6 h supérieure à 0.5 mm, 2 mm et 5 mm. La moyenne des quatre scores forme le score IP16. Le plus élevé, le meilleur est le score de Brier.

5.1.4 Augmentation du coût de calcul de l'assimilation des observations PWS

Pour les expériences réalisées, on a besoin en moyenne de 15 min pour faire toutes les étapes de l'assimilation. Ce temps inclus la préparation de l'ébauche (*fc_init*), des observations (*obsprep*), la minimisation (*analyse*), l'analyse de surface (*surfan*), les post-traitements (*postproc_an*) et la prévision à 1 h d'échéance pour le réseau suivant (*fc_i*). Toutes les étapes de l'assimilation sont réalisées sur 1 noeud de calcul, hormis les deux étapes les plus longues qui sont, la minimisation et la prévision à 1 h d'échéance qui sont faites sur 8 et 13 noeuds respectivement. La prévision à 48 h d'échéance (*fc*) dure en moyenne 2 h 30 min en mode recherche (sur 13 noeuds). En pratique, le temps CPU est 14 fois plus élevé pour une prévision à 48 h d'échéance que pour un cycle d'assimilation (figure 5.5).

Dans AROME-France, en moyenne sur le mois de septembre 2021, 49 250 observa-

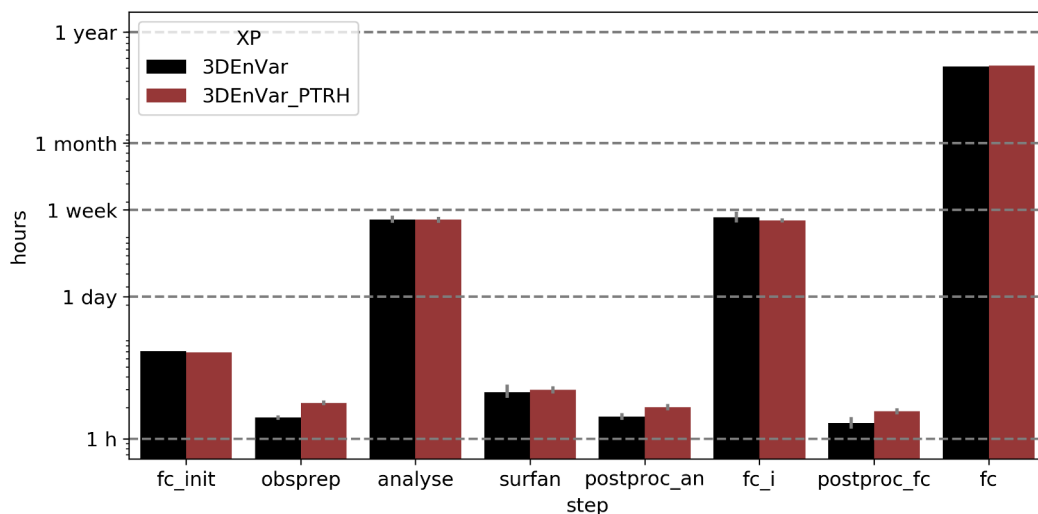


FIGURE 5.5 – Temps CPU moyen pour chaque étape de l'assimilation et de la prévision d'AROME-France. En noir l'expérience de contrôle 3DEnVar et en marron 3DEnVar_PTRH. Réalisé sur 12 jours pour des raisons techniques. Voir texte pour la signification des étapes. Pour plus de détails sur le temps CPU, voir Annexe B.

tions sont assimilées par réseau. L'assimilation d'observations de pression, température et d'humidité relative provenant de PWS dans l'expérience 3DEnVar_PTRH fait passer le nombre moyen d'observations de surface de 14 585 observations par réseau à 52 320 soit une augmentation d'un facteur 3.6.

Sur l'expérience d'une durée d'un mois présentée ici (figure 5.5), les temps de préparation et de post-traitements des observations augmentent significativement, respectivement de 39 % et 24 %. Le temps de préparation reste très faible comparé aux temps de préparation de l'analyse ou de la prévision longue (respectivement 60 et 1900 fois moins long). Il n'y a pas d'augmentation significative du temps pour l'étape de minimisation, ce qui aurait pu en être autrement avec d'autres schémas d'assimilation comme le schéma de filtre de Kalman. Un calcul de significativité avec une méthode de bootstrapping a été réalisé et aucun changement significatif de la moyenne des temps de calcul n'a été observé avec un intervalle de confiance à 95 %. Autrement dit, l'augmentation est inférieure à la variabilité en temps de calcul des différents composants du système d'assimilation-prévision.

5.2 Synthèse

Les observations de pression, de température et d'humidité relative denses de PWS se complètent entre elles afin de réduire les degrés de liberté du modèle près de la surface. On voit ici que l'assimilation des observations produit un choc à l'analyse, et que l'influence réelle des observations après l'analyse est nettement réduite, mais reste bénéfique sur les scores de pression et l'humidité relative en surface jusqu'à 9 h d'échéance. L'assimilation conjointe des observations de PWS améliore les scores de Brier sur la prévision jusqu'à 24 h d'échéance du vent instantané et de cumul de précipitations sur 6 h par rapport à l'expérience de contrôle. L'assimilation des nouvelles observations n'augmente le temps de calcul des cycles d'assimilation-prévision que des phases liées aux traitements des observations. La phase de préparation des nouvelles observations pourra être réduite en optimisant le type de fichier d'entrée utilisé.

Conclusions générales et perspectives

La croissance du nombre de données issues d'objets connectés et leur densité apparaît comme une opportunité pour améliorer l'anticipation de phénomènes convectifs et tout autre type de phénomènes de méso-échelle impliquant la couche limite atmosphérique. L'acquisition des observations d'opportunité étant accessible à tout un chacun, elles représentent une nouvelle source d'observations disponible pour les modèles de prévision numérique du temps régionaux. L'assimilation de nouvelles observations dans les modèles de PNT constitue un axe de recherche en vue d'améliorer l'état initial du modèle, notamment dans les modèles régionaux utilisés pour les prévisions à courtes échéances (Valmassoi *et al.*, 2023).

Dans le cadre de cette thèse, trois types d'observations que sont la pression, la température et l'humidité relative, provenant de stations météorologiques personnelles, ont été assimilés dans le modèle de PNT régional AROME-France. Pour chaque type d'observation un pré-traitement est proposé, des expériences de prévisions en assimilant les nouvelles observations ont été effectuées et l'impact objectif, grâce à des scores de prévisions, et subjectif par des études de cas, ont été réalisés.

D'abord des tests d'assimilation ont été réalisés sur les observations de pression des PWS. Le pré-traitement proposé consiste à (i) corriger leur altitude avec l'ébauche du modèle, (ii) contrôler leur qualité en rejetant les observations éloignées du modèle et (iii) réaliser un écrémage à 1.3 km de résolution. Une fois pré-traitées, les observations de pression des PWS sont d'une qualité comparable aux observations standards. Après pré-traitement 27 889 observations de pression des PWS sont assimilées chaque heure, soit 129 fois plus que celles provenant des SWS. Par ailleurs, des expériences avec les deux schémas d'assimilation 3DVar et 3DEnVar ont pu être menées : tandis que le schéma d'assimilation 3DVar ne permet pas de tirer profit de l'information de méso-échelle des PWS, le schéma 3DEnVar en a les capacités grâce à sa matrice de covariance d'erreur d'ébauche dépendante du flux. En assimilant les observations de pression des PWS, une amélioration significative du champ de pression de surface est remarqué à l'analyse, notamment en France où les observations PWS sont assimilées. Différentes études de cas ont été menées : sur un système convectif de méso-échelle, un front, une dépression de méso-échelle ainsi qu'un mascaret atmosphérique. Pour chacune de ces configurations, le schéma d'assimilation 3DEnVar a permis de prendre en compte les gradients de pression de surface de méso-échelle observés par les PWS pour former l'analyse.

La qualité des observations de température et d'humidité relative apparaît quant à elle plus disparate. Une comparaison de différents traitements a été réalisée, comprenant une correction de biais et un contrôle de qualité des observations. Pour ce faire, il est nécessaire d'avoir une référence la plus précise possible. Deux références ont été testées : avec le réseau PWS lui-même (vérification mutuelle) ou avec le modèle avec la dernière prévision disponible, c.-à-d. à 1 h d'échéance. Il a été mis en évidence que l'utilisation des observations de PWS voisines pour contrôler une PWS peut causer une divergence d'observations entre le réseau de PWS et le réseau de SWS. Le test d'un contrôle de qua-

lité qui a pour référence le modèle s'avère pertinent pour rejeter des stations anormales, mais peut à l'inverse être à l'origine d'un sur-apprentissage du modèle, en ne sélectionnant que les observations corrigées qui sont proches du modèle. Néanmoins, l'assimilation d'observations PWS de température et d'humidité relative pré-traitées diminue l'écart à l'ébauche des observations de SWS. Comme pour les observations de pression, l'assimilation de la température et de l'humidité relative des PWS — avec les réglages actuels du schéma d'assimilation 3DVar pour l'atmosphère et l'IO pour la surface — provoque une dégradation des écarts à l'ébauche des observations SWS. Cette dégradation est principalement due à la taille caractéristique élevée des incréments de surface issue de ces deux schémas d'assimilation de données. Les cas d'étude de brise de mer, de brouillard, d'îlot de chaleur urbain et d'une ligne orageuse ont permis de mettre en évidence l'utilité des observations de PWS avec le schéma 3DEnVar pour contraindre l'analyse à la surface, permettant d'améliorer la représentation de ces phénomènes à la surface.

Enfin, l'assimilation conjointe des trois types d'observations — avec les réglages actuels du schéma 3DEnVar — a montré une baisse significative de l'erreur de la prévision à 1 h d'échéance par rapport aux observations SWS de surface en France : (i) de la pression à hauteur de 7.5 % et (ii) de l'humidité relative à hauteur de 0.7 %. Cette amélioration n'entraîne pas de dégradation des autres champs de température ou de vent de surface. L'assimilation conjointe a montré l'importance d'assimiler des observations de différents types. En effet, pour un champ donné du modèle, plus celui-ci est contraint par des observations, moins les covariances croisées d'autres observations sont susceptibles d'influencer ce champ.

Les limites sur le gain attendu par l'assimilation d'observations des PWS sont résumés ci-après. Ils concernent en premier lieu des problèmes liés à l'analyse :

1. Des observations PWS de mauvaise qualité peuvent dégrader l'analyse, exigeant par conséquent un pré-traitement efficace et un réglage adéquat des erreurs d'observations prescrites. Il convient de faire particulièrement attention aux erreurs d'observation systématiques.
2. Des erreurs d'observation des PWS corrélées spatialement (causées par exemple par l'instrument de mesure, le placement du capteur, ou l'opérateur d'observation), peut induire une dégradation de l'analyse, car ces corrélations d'erreur d'observation ne sont pas prises en compte dans la matrice de covariance d'erreur d'observation \mathbf{R} . Les corrélations d'erreur d'observation croisées entre les systèmes d'observation doivent aussi être prises en compte (par exemple, pour une station donnée, une surchauffe du capteur de température peut impacter la mesure d'humidité relative provoquant une corrélation croisée d'erreur d'observation). Un pré-traitement et un écrémage sont ici réalisés pour réduire les corrélations d'erreur des observations spatialement proches.
3. Même avec des observations de bonne qualité et une matrice \mathbf{R} optimale, l'assimilation de données variationnelle reste limitée par la qualité de la matrice de covariance d'erreur d'ébauche \mathbf{B} et de la localisation, permettant la propagation de l'information des observations. Grâce au passage du schéma 3DVar au schéma 3DEnVar, un progrès a été réalisé en actualisant la matrice \mathbf{B} à chaque analyse, permettant une meilleure prise en compte des observations denses de surface. Néanmoins, des limites restent concernant la faible localisation à la surface de 25 km.

À ces limites à l'analyse s'ajoute des limites concernant la prévision à courte échéance :

- A. Même avec des observations PWS de température et d'humidité relative de bonne

qualité, et des matrices \mathbf{B} et \mathbf{R} permettant de bien prendre en compte les observations à l'analyse, en ajoutant des observations denses seulement à la surface, comme c'est le cas avec le réseau PWS, il y a un risque qu'elles ne contraignent que les champs de basses couches de l'analyse. Il est ainsi possible de créer des discontinuités à l'analyse entre les champs de surface très contraints par les observations et les champs d'altitude peu contraints. Une solution temporaire consiste à réduire l'impact des observations de surface en augmentant leur erreur d'observation, ou en augmentant l'écrémage. Une solution à ce problème est d'assimiler plus d'observations en altitude avec des observations d'avions et satellitaires supplémentaires pour avoir une répartition des observations homogène sur toute la hauteur de la troposphère, ce qui est en cours de recherche.

- B. Contraindre les champs de surface peut s'avérer pertinent pour améliorer la représentation des ingrédients nécessaires à la formation d'un phénomène de méso-échelle (par exemple assimiler de l'humidité au sol pour la formation de brouillard). Néanmoins, certains ingrédients ou marqueurs de méso-échelle n'interviennent que quelques heures à minutes avant la formation du phénomène (par exemple, la plage froide d'un orage n'intervient qu'une fois celui-ci déjà en phase mature). Dans certains cas, les marqueurs d'un phénomène ne sont vus par les observations PWS que trop tard pour améliorer la prévision, comme dans le cas de l'arrivée d'une dépression de méso-échelle sur la côte atlantique, les observations PWS étant restreintes sur terre. Les observations PWS ne peuvent dans ces cas que faiblement améliorer la prévision à longue échéance de ces phénomènes. De plus, s'il y a un décalage spatial ou temporel entre le phénomène simulé et celui observé, le risque de « double peine » est important (par exemple pour le cas d'un système convectif, les observations peuvent initier une plage froide à l'analyse là où il n'y a pas de convection simulée, et réduire la plage froide là où le modèle simule la convection).

Par ailleurs, le coût carbone pour le lancement des expériences d'assimilation-prévision avec un périmètre restreint au calcul et au stockage a été estimé à 6.4 t eqCO₂ (Annexe B). Des possibilités pour réduire l'empreinte de telles études seraient à approfondir, en réduisant le nombre d'expériences ou la taille du domaine d'étude. Des études aussi coûteuses en temps de calcul sont l'opportunité de questionner et de valoriser chaque expérience réalisée. Il incombe toutefois à la communauté scientifique de rester attentive à sa production et à son utilisation croissante de données, associées à des réseaux d'infrastructures numériques intermédiaires (*fog computing*) très consommateurs de métaux et d'énergie. Les objets connectés déjà déployés représentent une source d'observations économique et a priori peu coûteuse pour l'environnement en comparaison avec des observations satellites (Odin *et al.*, 2024). Finalement, l'ajout de nouvelles observations à l'assimilation permet d'améliorer la prévision à courte échéance sans augmenter significativement le temps de calcul du système assimilation-prévision.

Avant l'assimilation opérationnelle des observations PWS dans les modèles régionaux, des travaux importants doivent encore être menés. En particulier, le traitement des observations peut être amélioré en utilisant par exemple la méthode de correction de biais variationnelle qui permet une mise à jour automatique des coefficients de correction par le modèle. Cette méthode n'est possible que grâce au maintien d'un réseau d'observations de référence, servant d'ancrage pour éviter que le cycle d'assimilation-prévision puisse dériver (Vasiljevic *et al.*, 1992). Les observations d'opportunité ne pourront donc en aucun cas se substituer aux observations de référence. Des discussions pourront avoir lieu concernant la mise en place opérationnelle d'une correction de biais et d'un contrôle de qualité

spécifique à ces données d’opportunité. Ce contrôle de qualité pourrait être inspiré des contrôles réalisés pour les observations d’avions Mode-S, complété par la création d’une liste blanche de stations autorisées à être assimilées.

Ces premières expériences d’assimilation d’observations de stations météorologiques personnelles ouvrent la voie au contrôle et à l’assimilation d’autres observations au sein des données d’observations d’opportunité. À court terme, les observations assimilables concernent les réseaux d’observations semi-professionnels et maintenus, provenant du réseau national associatif StatIC ou régional ROMMA. D’autres types d’observations provenant des stations météorologiques personnelles pourraient être testées comme les observations de vent. À plus long terme, les observations d’opportunité assimilables pourraient provenir de véhicules connectés, de smartphones ou même de caméras connectées. Les défis sont multiples tant scientifiquement qu’en termes de coûts et de droits d’accès. Ces questions doivent se discuter à l’échelle de consortiums pour pouvoir en distribuer la charge, initiative déjà lancée par le consortium EUMETNET (*European Meteorological Network*).

Des défis importants restent à mener sur les réglages du schéma d’assimilation 3DEnVar, de l’inflation des covariances d’erreur d’ébauche, aux réglages des erreurs d’observations et de la localisation. La fréquence des observations de PWS à pas de temps 5 min incite la création de systèmes d’assimilation capables de tirer parti d’observations à haute fréquence temporelle. Le schéma d’assimilation 4DEnVar pourrait ainsi permettre de recalculer des systèmes de méso-échelle avant qu’ils ne divergent des observations.

L’assimilation d’observations issues de stations d’opportunité s’inscrit dans les types d’observations aujourd’hui assimilées dans les modèles de PNT. Pendant longtemps, le réseau d’observations au sol déployé et assimilé correspondait à un réseau d’échelle synoptique. On assiste depuis la création de l’assimilation à échelle convective à un basculement vers une assimilation de méso-échelle, avec des réseaux plus denses telles que les observations radars et bientôt d’avions Mode-S supplémentaires. Cette bascule s’accroît avec de nouveaux schémas d’assimilation EnVar, capables de propager l’information des observations de manière cohérente avec les conditions météorologiques du jour. L’assimilation à échelle convective pose de nombreuses questions quant à la capacité des schémas d’assimilation à recalculer des systèmes de méso-échelle déjà en place.

L’assimilation de données permet de passer de simulations météorologiques, que l’on pourrait qualifier d’idéalisées si elles sont éloignées de la réalité, à des simulations météorologiques plus proches de l’état observé de l’atmosphère. Cette amélioration peut être grandement utile pour appréhender le déroulé temporel et spatial d’un phénomène. La qualité de la prévision s’améliore lorsqu’un phénomène météorologique est correctement positionné dans son environnement. Prenons l’exemple d’un phénomène convectif se produisant au-dessus d’une ville : si le modèle rate son développement de quelques kilomètres, il affaiblit l’interaction de l’orage avec la surface urbanisée, accentuant d’autant plus la perte de qualité de la prévision (Forster *et al.*, 2023).

À plus petite échelle, ces observations peuvent remplacer une partie des observations utilisées pour des campagnes de mesures à fine échelle, comme dans l’étude du rôle des surfaces végétalisées en ville (projet PANAME, Haeffelin *et al.*, 2023). Ces observations denses pourraient offrir la possibilité de contraindre les champs de modèles régionaux à résolution plus fine. Elles pourraient ainsi initialiser des modèles de PNT à 750 m de résolution, ou des simulations LES²⁰ à encore plus fine échelle - jusqu’à 100 m de résolution. Pour cela, devront être relevés les défis liés à la modélisation des corrélations d’erreur

20. *Large eddies simulations*

d'observations et à la descente d'échelle de la matrice de covariances d'erreur d'ébauche au sein du schéma d'assimilation de données. Les observations d'opportunité pourraient alors aider à la caractérisation de structures d'écoulement d'échelle aérologique inférieure à 2 km.

La création d'analyses et de réanalyses les plus fidèles possible à la réalité est cruciale. Les réanalyses de modèles à aire limitée pourraient être essentielles pour alimenter en données une base d'apprentissage de futurs émulateurs IA de modèle de prévision de fine échelle. Développer des analyses et des réanalyses qui puissent représenter une diversité de phénomènes de méso-échelle pourrait être une des clefs de réussite pour améliorer ces futurs modèles de PNT. De plus, l'arrivée de nouveaux outils de *machine learning* et d'IA pourraient permettre une meilleure estimation de certains paramètres des schémas d'assimilation de données, voir de remplacer les schémas eux-mêmes. En effet, l'assimilation de donnée et l'IA présentent de nombreuses similitudes : ce sont toutes deux des méthodes inverses qui peuvent être réunies dans un cadre bayésien (Geer, 2021). Des premières expériences d'assimilation-prévision numériques entièrement avec des outils d'IA, nommés *end-to-end AI weather forecasting*, commencent à émerger et pourraient à terme remplacer les modèles physiques (Chen *et al.*, 2023). Ces outils pourraient permettre à terme de produire une prévision à partir des observations elles-mêmes, rendant d'autant plus importants la qualité et la densité des réseaux d'observations, dont les PWS pourraient faire partie.

Enfin, les observations d'opportunité offrent la possibilité à quiconque achetant une station météorologique connectée de pouvoir participer, de façon collaborative, à l'amélioration des modèles de prévision numérique du temps ; une façon d'associer le public à la compréhension et l'anticipation de phénomènes météorologiques à enjeux.

A Cas d'études réalisés

Différents événements de méso-échelle ont été étudiés, qui sont dans l'ordre d'apparition des résultats du manuscrit (Figure A) :

1. Un système convectif de méso-échelle (MCS), se développant dans le Gard le 14 septembre 2021. Théorie, page 29 ; et assimilation, page 78.
2. Un front froid le 5 octobre 2021, passant sur le nord de la France. Théorie, page 27 ; et assimilation, page 92.
3. Une dépression de méso-échelle rentrant sur la côte aquitaine le 3 novembre 2023 dans le sud-ouest de la France. Théorie, page 27 ; et assimilation, page 93.

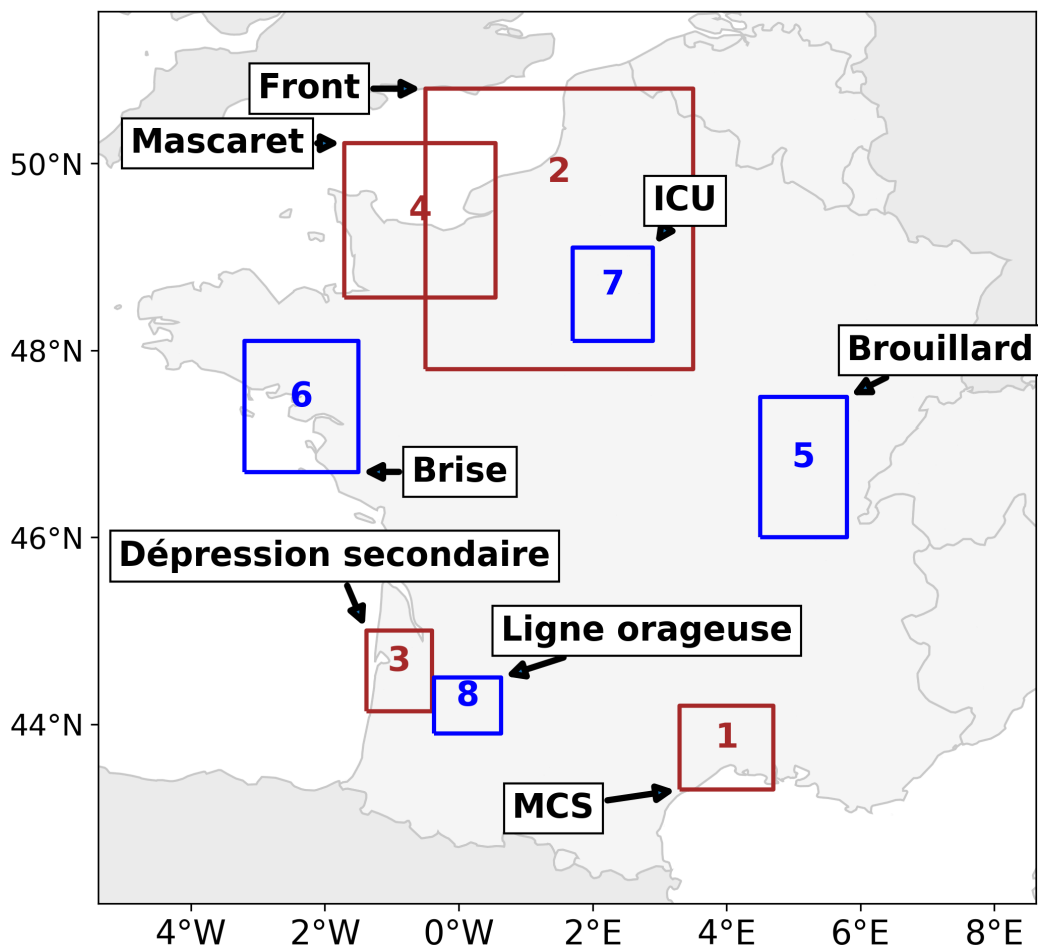


FIGURE A – Carte des différentes situations météorologiques étudiées.

4. Un mascaret atmosphérique rentrant sur la côte normande le 18 juin 2022. Théorie, page 28 ; et assimilation, page 95.
5. Un brouillard de rayonnement le 23 septembre 2021 dans la vallée de la Saône

- (Bourgogne-Franche-comté). Théorie, page 36 ; et assimilation, page 102.
6. Une brise de mer pendant une vague de chaleur le 7 septembre 2021 en Bretagne sud. Théorie et assimilation, page 102.
 7. Un îlot de chaleur urbain sur la métropole de Paris le 17 juin 2023. Théorie, page 34 ; et assimilation, page 137.
 8. Une ligne orageuse se développant le 8 septembre 2021 sur Agen (Lot-et-Garonne). Assimilation, page 137.

B Coût énergétique des expériences réalisées

La recherche pour rester pertinente n'a pas la légitimité de se soustraire aux contraintes de son environnement. Le Comité d'éthique du CNRS (COMETS), dans sa note de 2022, affirme qu'un juste équilibre doit être trouvé entre liberté de la recherche et considérations pour l'environnement (Askenazy *et al.*, 2022). Le chercheur ne peut être neutre dans sa démarche scientifique. À l'idée de neutralité, on préférera donc les notions de quête d'objectivité, de rigueur scientifique et de transparence sur les valeurs (Fragnière *et al.*, 2022).

Pour réaliser une empreinte carbone individuelle il est nécessaire de définir le périmètre qu'elle comprend. Il serait fastidieux et redondant avec une étude en cours, de prendre en compte l'ensemble des coûts carbones nécessaires pour réaliser l'empreinte individuelle totale. On peut néanmoins diviser l'empreinte carbone en trois sous-parties : une partie personnelle (ou privée), une partie inhérente aux conditions de travail, et enfin une partie liée aux expériences de recherche individuelle. C'est cette dernière partie qui nous intéresse ici, souvent oubliée ou négligée par le calcul d'empreinte carbone. Cette section vise ainsi à donner une estimation de la consommation énergétique des expériences d'assimilation des nouvelles observations des PWS dans AROME-France, pour répondre à la question : est-ce que le coût des expériences de prévisions réalisées sont-elles négligeables ou sont-elles à prendre en compte vis-à-vis de l'objectif de réduire l'empreinte carbone individuelle par personne et par an à $2 \pm [1.6, 2.8]$ t eqCO₂²¹ ?

B.1 Contexte

La feuille de route de la France pour lutter contre le changement climatique, à travers la Stratégie Nationale Bas Carbone (SNBC), donne des orientations pour atteindre la neutralité carbone d'ici à 2050. Celle-ci fixe un objectif de réduction des GES de 40 % d'ici à 2030 par rapport à 1990. Météo-France s'engage à réduire de 20 % ses émissions de GES d'ici à 2026, comparé à ses émissions de 2019 (Météo-France, 2023). Pour 2021, les émissions de gaz à effet de serre de Météo-France se sont élevées à 11 241 t eqCO₂ dont 34 % proviennent de la consommation énergétique des bâtiments et 22 % de la consommation énergétique des calculateurs et des radars.

La puissance de calcul des supercalculateurs (*High performance computing*, HPC) de Météo-France n'a cessé d'augmenter et a permis de franchir des ruptures pour l'augmentation de la qualité de la prévision numérique du temps (PNT) (Lenfle, 2018). Une augmentation de la puissance de calcul est mise en avant dans les COP (contrats d'objectifs et de performances) de Météo-France pour utiliser de manière croissante le nombre d'observations de manières optimisées dans les modèles de PNT (Météo-France, 2023). Météo-France possède et partage deux HPC identiques, nommés Belenos et Taranis. Belenos est composé de racks BullSequana comprenant des processeurs AMD Epyc 7742 (Rome). La chaleur générée par les principaux composants est évacuée par un liquide (DLC) dans un rapport eau/air de 90 %, et est utilisée pour chauffer les bâtiments du site de la Météopôle en hiver. L'efficacité énergétique (PUE) de Belenos est estimée à 1.18, et est, à fortiori, bien plus efficace que la moyenne mondiale des HPC estimés à 1.67 en 2019 (Lannelongue *et al.*, 2021). Le supercalculateur possède 2256 nœuds de calculs. Chaque nœud possède 128 coeurs pour une mémoire RAM de 256 Go et une puissance de 3.5 W

21. Suivi de la Stratégie Nationale Bas-Carbone (SNBC-2) (https://www.ecologie.gouv.fr/sites/default/files/IndicateurscompletsSNBC22023_def.pdf)

par cœur²².

On appelle **temps CPU**, le temps de calcul sur un cœur d'un CPU (*central processing unit*). Il peut être obtenu en multipliant le temps de calcul écoulé (*elapsed*) par le nombre de cœur physique. Cette métrique est couramment utilisée, notamment pour attribuer un temps de calcul par projet. La consommation d'un calcul peut être donnée par (i) une mesure réelle, à l'aide d'un wattmètre directement branché au nœud de calcul, (ii) par une mesure logicielle si le processeur est compatible, ou enfin (iii) à l'aide d'estimateurs liés au temps de calcul. Belenos n'ayant pas à disposition de mesure réelle, les méthodes de mesure logicielle et par estimateur (encadré *Méthode de Lannelongue et al. (2021)*) seront utilisées par la suite.

B.2 Méthode de Lannelongue *et al.* (2021)

La consommation énergétique d'un algorithme est donnée par l'équation 1 :

$$E = t(n_c P_c u_c + n_m P_m) \text{PUE}, \quad (1)$$

où t est le temps de calcul (en heures), n_c le nombre de coeurs, n_m la place mémoire disponible (en GB), u_c le facteur d'utilisation des coeurs (entre 0 et 1). P_c est la consommation électrique d'un coeur de calcul et P_m la consommation électrique de mémoire (en Watt).

L'efficacité d'un data center est donnée par son PUE (Power Usage Effectiveness) avec : $\text{PUE} = \frac{P_{\text{total}}}{P_{\text{compute}}}$, qui correspond au ratio entre la puissance totale utilisée incluant l'ensemble des composants (e.g. le refroidissement, la lumière) et la puissance réellement utilisée pour le calcul. Un data center idéal aurait un PUE de 1.0 où toute la puissance serait utilisée pour les ressources en calcul.

L'empreinte carbone est déduite par l'équation 2 :

$$C = E.CI, \quad (2)$$

Avec CI l'intensité carbone qui correspond à l'empreinte carbone émise par un pays pour produire 1 kWh d'énergie. Pour la France, il est à 59.9 gCO₂e kWh⁻¹ bien en deçà de la moyenne mondiale à 475 gCO₂e kWh⁻¹²³.

Les limitations restent notables. Il n'y a pas de calcul du cycle de vie des composants et de l'utilisation de ces données, pas de prise en compte des incertitudes sur la consommation d'énergie d'un processeur (TDP), ni la consommation liée au stockage qui peut être sous-estimée. Et enfin la fluctuation du mix énergétique du pays peut grandement varier.

B.3 AROME-France : expériences sur un mois

Pour réaliser une prévision de fine échelle comme le permet le modèle de PNT AROME, un certain nombre de modèles intermédiaires sont nécessaires. Une prévision AROME dépend à la fois d'un cycle d'assimilation, d'un cycle d'assimilation d'ensembliste EDA, pour son schéma d'assimilation 3DEnVar, et de la prévision d'un modèle coupleur de

22. <https://www.amd.com/fr/products/cpu/amd-epyc-7742>

23. Facteur d'émission du Labo 1.5 (https://apps.labos1point5.org/static/carbon/FacteursEmission_GES1point5_Juin2021.pdf, dernier accès mars 2024)

grande échelle ARPEGE, qui lui-même dépend d'un cycle d'assimilation.

Les expériences de recherche sur l'assimilation de nouvelles observations sont quant à elles moins coûteuses, car elles nécessitent seulement de reproduire un cycle d'assimilation (une analyse par heure) et de moins nombreuses prévisions (quatre prévisions par jour). Une expérience d'un mois d'AROME-France en 3DEnVar²⁴, dure 7 à 8 jours pour un temps CPU de 28 192 jours. Grâce à l'outil GreenAlgorithms (Lannelongue *et al.*, 2021), la consommation énergétique des CPU est estimée à 2380 kWh. Cette estimation est proche de la valeur obtenue par la méthode de mesure logicielle, c'est-à-dire de 2340 kWh²⁵ (soit 2 % plus faible), ce qui permet de valider l'estimation réalisée par l'outil GreenAlgorithms.

L'outil GreenAlgorithms ajoute à la consommation CPU, la consommation liée à la place mémoire (+21 %) et celle des pertes inhérentes au data center (+21 %). La consommation totale est alors estimée à 3400 kWh avec la méthode GreenAlgorithms. La consommation électrique d'une expérience AROME-France sur un mois équivaut en termes d'empreinte carbone à 195 kg eqCO₂.

B.4 Statistiques sur l'étude complète

Une étude, comme celle présentée ici sur l'assimilation d'un nouveau réseau d'observation, nécessite de faire différents tests de sensibilités et par conséquent de rejouer les expériences plusieurs fois. On nomme *pragmatic scaling factor* (PSF), le nombre de fois où l'expérience a été reproduite. Des dizaines d'expériences ont été produites avant que la méthode de correction des données n'ait été validée à chaque fois pour les observations de pression, de température et d'humidité relative. On estime à 49 le PSF lors de cette étude, correspondant au nombre de mois d'expériences AROME-France réalisées.

Cela correspond à 3800 années de temps CPU et une consommation électrique de 166.6 MWh. L'étude complète accuse en définitive une empreinte carbone de 6.4 t eqCO₂ seulement concernant les expériences numériques. L'incertitude autour de cette estimation est donnée en comparant l'intensité carbone obtenue avec d'autres centres de calcul. À titre indicatif, cette empreinte carbone est comparable à l'émission émise lors d'un déplacement aller/retour en avion d'une mission en Australie.

Intensité carbone

Le temps de calcul écoulé par le modèle est connu avec précision. En revanche, l'**intensité carbone**, qui correspond à la consommation par heure de calcul, est quant à elle plus incertaine. En pratique, pour un même temps de calcul, la consommation d'un algorithme peut varier du simple au quadruple, suivant l'optimisation de l'algorithme. Néanmoins, l'ordre de grandeur de l'intensité carbone, pour un calcul optimal, peut être comparé à d'autres systèmes de calcul. Il est important de noter que l'intensité carbone dépend fortement du **périmètre** choisi, qu'il soit réduit au coût de calcul, ou élargit aux différents composants du HPC (par exemple, aux serveurs ou au stockage), voir au cycle de

24. Cycle 48t1.

25. Cette valeur est calculée en cumulant la consommation électrique de chaque élément du cycle d'assimilation-prévision. Elle a été calculée sur l'expérience 3DEnVar_PTRH sur une période de 17 jours puis extrapolée sur la période complète de 30 jours pour des raisons techniques.

vie des composants et à la maintenance. L'intensité carbone de différents HPC a été calculée par le groupement de recherche Labo1.5, en prenant un périmètre élargi (Benaben *et al.*, 2024). Il est important de comprendre que le périmètre du Labo1.5 comprend non seulement la consommation électrique des HPC, mais aussi l'amortissement des GES émis lors de la production des différents équipements.

Périmètre	Coût de calcul et de stockage	+ Infra-structure	+ ACV complet	
Présente étude	✓			0.32 g eqCO ₂ /hCPU
Curie ¹		✓		0.8 g eqCO ₂ /hCPU
Occigen ¹		✓		1.19 g eqCO ₂ /hCPU
Joliot Curie SKL ¹		✓		1.64 g eqCO ₂ /hCPU
GRICAD ²			✓	2.87 g eqCO ₂ /hCPU

TABLEAU B – Comparaison de l'intensité carbone d'une heure de calcul sur CPU. ⁽¹⁾ D'après le site et la note technique du Labo1.5 ^a avec une incertitude globale estimée de 40 % Benaben *et al.* (2024). ⁽²⁾ Grenoble Alpes Recherche Infrastructure de Calcul Intensif et de Données (GRICAD, Berthoud *et al.*, 2020 ; sans l'empreinte du personnel).

a. <https://apps.labos1point5.org/>

Avec la méthode de Lannelongue *et al.* (2021), l'intensité carbone pour la présente étude atteint une valeur de 0.32 g eqCO₂/hCPU. Cette comparaison permet de valider l'ordre de grandeur estimé par la méthode de Lannelongue *et al.* (2021), et d'étendre le calcul à un périmètre élargi. Belenos étant un HPC avec un PUE bas, il peut être comparé aux meilleurs HPC du labo1.5, comme Curie. L'empreinte pourrait être alors 2 à 3 fois plus élevée en prenant en compte un périmètre élargi à l'infrastructure. Prendre en compte le cycle de vie complet du HPC, peut aller jusqu'à doubler l'intensité carbone, comme pour le GRICAD.

B.5 Conclusion

Concernant l'empreinte énergétique de l'étude, celle-ci a pu être obtenue avec deux méthodes distinctes grâce à une mesure logicielle connectée au ordinateur et une méthode d'estimation donnée par des grandeurs intensives du ordinateur. Les deux méthodes donnent un résultat proche à 2 % près de la consommation électrique des CPU pour l'expérience de référence sur un mois. L'empreinte carbone des expériences de prévisions en prenant en compte seulement l'énergie électrique de calcul et de stockage a été estimé à 6.4 t eqCO₂. Cette valeur, qui peut être divisé par 3 pour être réparti sur les années de l'étude, doit être mise en perspective de l'objectif fixé par les Accords de Paris de limiter « l'élévation de la température moyenne de la planète nettement en dessous de 2 °C par rapport aux niveaux pré-industriels » ce qui correspond à limiter l'empreinte par personne et par an à $2 \pm [1.6, 2.8]$ t eqCO₂ ²⁶. Elle doit être ajoutée à l'empreinte individuelle pour être complète.

Des solutions pour réduire la consommation électrique doivent être recherchées. Des

26. Suivi de la Stratégie Nationale Bas-Carbone (SNBC-2) (https://www.ecologie.gouv.fr/sites/default/files/IndicateurscompletsSNBC2023_def.pdf)

exemples de gain en termes d'énergie peuvent concerner la simplification des codes et l'écriture de moins nombreux fichiers de sorties de plusieurs gigabits. Une autre solution consisterait à réduire la taille du domaine du modèle AROME-France le temps de l'étude (Bakketun *et al.*, [2023](#)).

Sigles et acronymes

3DEnVar Schéma d'assimilation ensembliste variationnel tridimensionnel.

3DVar Schéma d'assimilation variationnel tridimensionnel.

ACCORD Consortium pour la recherche et le développement des modèles à échelle convective (*A Consortium for COvection-scale modelling Research and Development*).

AE-ARO Cycle d'assimilation ensembliste d'AROME.

ANTILOPE ANalyse par spaTIALisation hOraire des PrEcipitations.

AROME Modèle régional de Météo-France (Application de la Recherche à l'Opérationnel à Méso-Échelle).

ARPEGE Modèle global de Météo-France (Action de Recherche Petite Echelle Grande Echelle).

CEPMET Centre européen pour les prévisions météorologiques à moyen terme.

CLA Couche Limite Atmosphérique.

COSMO Consortium pour la modélisation à petite échelle (*Consortium for Small-scale Modeling*).

DWD Service météorologique allemand (*Deutscher Wetterdienst*).

EDA Ensemble d'assimilations (*Ensemble of Data Assimilations*).

EnKF Schéma d'assimilation (*Ensemble Kalman Filter*).

GBON Réseau mondial d'observations de base (*Global Basic Observing Network*).

ICU Îlot de Chaleur Urbain.

IFS Modèle global du Centre européen pour les prévisions météorologiques à moyen terme (*Integrated Forecasting System*).

IO Schéma d'assimilation d'interpolation optimale.

LAM Modèle de prévision numérique du temps à aire limitée (*Limited-Area Model*).

MCS Système convectif de méso-échelle (*Meso-scale Convective System*).

MSG Météosat deuxième génération, une mission de satellites lancés par EUMETSAT (*Meteosat Second Generation*).

OmA Différence entre l'observation et l'analyse au point de l'observation (*Observation minus Analysis*).

OmB Différence entre l'observation et l'ébauche au point de l'observation (*Observation minus Background*).

OMM Organisation Météorologique Mondiale.

OSE Comparaison d'une expérience avec un nouveau système d'observation à une expérience de contrôle (*Observing System Experiments*).

Pmer Pression réduite au niveau de la mer.

PNT Prévision Numérique du Temps.

PWS Stations météorologiques personnelles (*Personal weather stations* ou *Private weather stations*).

RADOME Réseau d'Acquisition de Données et d'Observations Météorologiques Étendues.

rh2m Humidité relative à 2 m de hauteur au-dessus du sol.

SWS Stations météorologiques standards (*Standard weather stations*).

t2m Température à 2 m de hauteur au-dessus du sol.

Liste des figures

- 1.1 Définition des échelles spatiales et temporelles caractéristiques de phénomènes météorologiques variées. Classification de Orlandi (1975a) et Fujita (1981). Figure issue de Markowski et Richardson (2011). 6
- 1.2 Nombre de stations de surface échangeant des observations via l’OMM. Le nombre de stations est divisé en mailles d’environ $1^\circ \times 0.5^\circ$. À la date du 27 février 2024, on dénombre 16823 stations de surface, comprenant les stations GBON de surface, ainsi que des réseaux locaux de stations automatiques de surface (AWS), stations SYNOP et climatologiques (<https://oscar.wmo.int/surface/>). 7
- 1.3 Critères d’homologation en classe 1 d’une mesure de température et d’humidité relative de surface (Extrait de WMO, 2021). Une source de chaleur doit couvrir moins de 5 % de la surface si elle est à moins de 10 m, moins de 10 % de la surface si elle est à plus de 30 m et moins de 1 % de la surface si elle est à moins de 100 m. 10
- 1.4 Exemples de stations météorologiques : (a) station RADOME avec abri Stevenson à l’intérieur de l’aéroport Toulouse-Blagnac, (b) station StatIC à St-Hilaire équipé d’une Davis Instruments Vantage Pro2, (c) station Netatmo avec module intérieur (pression) et extérieur (température et humidité relative), et (d) pluviomètre et anémomètre. 12
- 1.5 Réseau d’observations de température à Toulouse le 7 septembre 2021 à 15:00 UTC. Avec 886 observations Netatmo brutes (en rond), six observations conventionnelles (en triangle), quatre observations du réseau StatIC (en étoile) et 21 observations du réseau Toulouse-Métropole (en rectangle). Le fond de carte représente le relief de la région. 14
- 1.6 Tests des positionnements des stations Netatmo. (a) Schéma de l’emplacement des stations Netatmo et (b) variation diurne du biais de la température moyenne (ΔT Netatmo-Reference) pour les dix modules extérieurs. Les stations sont positionnées dans les configurations suivantes : (1) à 15 m de hauteur, (2–5) aux quatre expositions (nord, sud, est, ouest) contre un mur, (6) sur un toit, (7) sous un arbre, (8) sur un pieu métallique à 2 m, (9) côté ombragé et (10) à l’intérieur de l’abri Stevenson. Figures issues de (Varentsov *et al.*, 2020). 15
- 1.7 Orographie du modèle AROME-France. 21
- 1.8 Discrétisation verticale dans le système de PNT AROME. Figure issue de Malardel (2009). 21
- 1.9 Coupe verticale dans un plan ouest-est de la vitesse verticale associée à : (a) une anomalie de tourbillon de basses couches dans un jet d’ouest, (b) une anomalie de tourbillon d’altitude dans un jet d’ouest, (c) la superposition d’une anomalie de basses couches et d’une anomalie de tourbillon d’altitude décalée vers l’ouest par rapport à l’anomalie de basses couches. Figure issue de Malardel (2009). 26

1.10	Champ de pression réduite au niveau de la mer (en hPa) (a) prévu à 1 h d'échéance par AROME, (b) observé par le réseau d'observations standards et (c) observé par le réseau d'observations standards et personnelles pré-traitées, le 5 octobre 2021 à 07:00 UTC. Les triangles indiquent les observations SWS et les ronds indiquent les observations PWS. Les iso-lignes marron représentent le module de vent de surface à 10 m s^{-1} simulé par AROME. Les barbules (kt) représentent les observations de vent moyen sur 10 min du réseau d'observations standards.	28
1.11	Champ de pression réduite au niveau de la mer (a) simulé par la prévision 30 min d'AROME, (b) observé par le réseau d'observations standards et (c) observé par le réseau d'observations personnelles pré-traitées, le 3 novembre 2023 à 02:30 UTC. Les barbules représentent les observations de vent du réseau d'observations standards.	29
1.12	Passage du mascaret atmosphérique le 18 juin 2022 à 18:00 UTC. Le fond de l'image représente le canal haute résolution visible du satellite MSG-3. Les couleurs représentent les observations de pression réduite au niveau de la mer (MSLP), des stations personnelles (cercle) et standards (triangle). À droite, les séries temporelles de MSLP, vitesse et direction du vent à Saint-Gatien-des-Bois. Figure issue de Mandement <i>et al.</i> (2023).	29
1.13	Différence de pression réduite au niveau de la mer de 10 min (a) simulée par la prévision 1 h d'AROME, (b) observée par le réseau d'observations standards et (c) observée par le réseau d'observations personnelles pré-traitées, le 18 juin 2022 à 18:00 UTC. Les barbules représentent les observations de vent du réseau d'observations standards.	30
1.14	Structure horizontale de lignes de grains matures. La pression de surface est indiquée en traits pleins, les vecteurs représentent le vent près de la surface et les régions où les précipitations sont significatives sont pointillées ou grisées. Adapté de Haertel et Johnson (2000).	32
1.15	Température (en °C) le 17 juin 2023 à 06:00 UTC. (a) Prévision d'AROME à 1 h d'échéance, (b) observations du réseau de stations standards et (c) observations (c) brutes et (d) pré-traitées du réseau de PWS.	35
1.16	Humidité relative (en %) le 23 septembre 2021 à 09:00 UTC. Voir Fig. 1.15 pour plus de détails.	36
2.1	Schéma de la fonction d'interpolation α de la température à 2 m (t2m) sur la verticale, en fonction de la stabilité (Ingleby <i>et al.</i> , 2024). $\alpha = 1$ indique que la t2m est égale à la température au niveau le plus bas du modèle et $\alpha = 0$ quand la t2m est égale à la température de surface. En noir la fonction basée sur la théorie de Monin-Obukhov, en bleu la formule d'interpolation actuelle (cycle 48r1) et en orange la fonction qui sera introduite dans le prochain cycle d'IFS (cycle 49r1).	47
2.2	Système d'assimilation de surface dans le modèle AROME de type interpolation optimale. Figure extraite de Sassi <i>et al.</i> (2019)	53
2.3	Proportion moyenne du nombre d'observations actives (qui ont passé le <i>screening</i>) assimilées par réseau entre le 6 septembre et le 5 octobre 2021. Seules les observations de bouées ne sont pas intégrées, représentant moins de 1 % des observations.	54

2.4	Nombre d'observations par type d'observation, (a) par réseau successif et (b) en moyenne horaire, pendant la période du 6 septembre au 5 octobre 2021. Voir le texte pour une description des types d'observations.	55
2.5	Chaîne de prévision opérationnelle pour AROME-France de 00:00 à 12:00 UTC (source : http://controle.meteo.fr).	55
2.6	Profils de température et d'humidité relative lors d'un épisode d'inversion le 26 novembre 2020 à 01:00 UTC à Payerne (Suisse). Expériences avec et sans l'assimilation d'observations de t2m et de rh2m. Figure issue de (Merker <i>et al.</i> , 2023).	57
3.1	Différences de géopotentiel au niveau de pression de la station entre observations brutes des PWS et prévision à 1 h d'échéance en fonction de leur adresse MAC, le 6 septembre 2021 à 01:00 UTC. L'adresse MAC est un identifiant unique donné à chaque station. L'étape de <i>screening</i> rejette les observations considérées anormales (en orange, chapitre 2). Les observations brutes comportent un biais dépendant de leur adresse MAC, dû à une mauvaise connaissance de leur altitude. Il n'y a pas de lien entre l'adresse MAC d'une observation et sa position spatiale.	63
3.2	Cartes des moyennes (a-c) et écart type (d-f) des écarts entre le modèle et les observations sur le mois d'étude ; (a, d) des écarts OmA des SWS, (b, e) des écarts OmB des SWS de l'expérience 3DEnVar et (c, f) des écarts OmB des PWS après prétraitement de l'expérience 3DEnVarP.	86
3.3	Δ RMS OmB (%) de l'expérience 3DEnVarP par rapport à 3DEnVar de la pression de surface pour chaque SWS sur le domaine AROME.	87
3.4	Évolution de la norme du gradient du terme d'ébauche de la fonction de coût DJ en fonction du nombre d'itérations lors de la minimisation de la fonction de coût les expériences 3DEnVarP et 3DEnVar le 5 octobre 2021 à 23:00 UTC.	88
3.5	Incrément d'analyse de pression de surface le 6 septembre 2021 à 01:00 UTC, partant de la même ébauche. Expériences a) 3DVar, b) 3DEnVar, c) 3DEnVar_ $\sigma_o/2$, d) 3DEnVarP, e) 3DEnVar_th10 et f) 3DEnVarP_ $\sigma_o/2$	89
3.6	Moyenne quadratique de la tendance de pression de surface du cycle d'assimilation-prévision AROME, pendant quatre réseaux successifs au cours des expériences le 1er octobre 2021 à 01:00 UTC, pour l'expérience 3DEnVar (noir), 3DEnVarP (bleu) et 3DEnVarP_noTh (en rouge). 3DEnVarP_noTh correspond à une expérience d'assimilation des observations de pression PWS sans écrémage des observations.	91
3.7	Analyse de P_{mer} à 00:00 UTC du front simulé par AROME le 5 octobre 2021. Le fond représente la pression réduite au niveau de la mer des expériences (a) en 3DVar, (b) en 3DVar avec l'assimilation des observations de pression des PWS, (c) en 3DEnVar, et (d) en 3DEnVar avec l'assimilation des observations de pression des PWS. Les expériences sont cyclées depuis le 6 septembre 2021. Le calcul de P_{mer} est effectué suivant Mandement et Caumont (2020).	92
3.8	Racine carrée de l'écart quadratique moyen des observations a) SWS et b) PWS, à chaque échéance de la prévision partant des analyses à 00:00 UTC des expériences 3DVar, 3DVarP, 3DEnVar et 3DEnVarP (figure 3.7) sur la moitié nord de la France (-5 à 5° W, 46.5 à 51.2° N). Pour une répartition homogène des stations, un écrémage à 10 km de résolution est réalisé pour les PWS, gardant de l'ordre de 2240 stations PWS, pour 128 stations SWS.	93

3.9	Analyse de P_{mer} à 03:00 UTC de la dépression secondaire simulée par AROME le 3 novembre 2023. Les expériences sont cyclées et partent du réseau opérationnel de 00:00 UTC.	94
3.10	Prévision à 1 h issue de l'analyse de 03:00 UTC de la dépression secondaire simulée par AROME le 3 novembre 2023 à 04:00 UTC.	95
3.11	Incrément d'analyse à 18:00 UTC du mascaret atmosphérique simulé par AROME le 18 juin 2022. Les quatre expériences partent de la même ébauche AROME à 17:00 UTC.	96
3.12	Unique ébauche et analyses pour les quatre expériences en P_{mer} à 18:00 UTC du mascaret atmosphérique simulé par AROME le 18 juin 2022. Les réflectances du canal HRV du satellite MSG-3 égales à 43.1 % (110/255) sont représentées par des lignes noires et montrent les trois bandes nuageuses parallèles du mascaret.	97
3.13	Prévisions à 18:00 UTC, des expériences (a-c) 3DEnVar et (d-f) 3DEnVarP partant de l'analyse (a/d) six heures plus tôt à 12:00 UTC, (b/e) trois heures plus tôt à 15:00 UTC et (c/f) une heure plus tôt à 17:00 UTC. Les observations SWS et PWS sont valides à 18:00 UTC.	98
4.1	Séries temporelles d'observations de température dans le sud-ouest de la France le 12 août 2020. Trois réseaux d'observations sont présents : 121 PWS (palette bleue), 5 SWS (palette rouge) et 3 StatIC (palette verte). Un orage passe à 16:00 UTC sur la zone, avec une baisse de 5 °C de la température en 1 h. Plusieurs effets sont visibles : la majorité des observations PWS ne correspond pas aux observations de référence, avec une température plus élevée, surtout la nuit ; certaines observations PWS demeurent constantes tout au long de la journée, probablement des températures intérieures ; enfin, certaines observations PWS affichent des températures nettement plus élevées que les références en journée, avec des écarts pouvant atteindre 10 à 15 °C par rapport aux SWS.	102
4.2	Illustration de la méthode de classification de courbes en six classes. Celle-ci consiste en une classification non supervisée par la méthode des <i>k-means</i> pour obtenir une discrimination des courbes d'observations de température des PWS. La discrimination est basée sur une décomposition en ondelettes des anomalies de séries temporelles. Les anomalies sont obtenues en soustrayant pour chaque heure et pour chaque observation le quantile 10 % de l'ensemble des observations PWS. Voir Marques (2023) pour plus de détails.	136
4.3	Analyse à 06:00 UTC de la température au premier niveau modèle (5 m), de l'îlot de chaleur urbain simulé par AROME le 17 juin 2023, (a) en 3DVar, (b) en 3DVar avec l'assimilation des observations de température des PWS (3DVarT), (c) en 3DEnVar, et (d) en 3DEnVar avec l'assimilation des observations de température des PWS (3DEnVarT). Les expériences sont cyclées et partent d'analyses du réseau de 00:00 UTC. Les observations de températures des PWS ont été pré-traitées suivant P-S comprenant un écrémage à 8 km de résolution. <i>A</i> et <i>B</i> font référence aux deux cercles de (a). Les scores correspondent au rms OmA des neuf stations StatIC indépendantes (étoiles).	138

4.4	Analyses de la ligne orageuse du 8 septembre 2021 à 18:00 UTC (a, b) en température et (c, d) en humidité relative au premier niveau du modèle (5 m). Expériences (a, c) de contrôle (3D <code>EnVar</code>), (b) avec l’assimilation des observations de température PWS (3D <code>EnVarTS</code>), ou (d) d’humidité relative (3D <code>EnVarRHS</code>). Les réflectivités radar égales à 40 dBZ sont représentées par le trait continu. Le trait tireté représente la section verticale réalisée dans la figure 4.5.	139
4.5	Coupes verticales d’incrément (a-c) de température et (d-f) d’humidité relative, partant de la même ébauche. Incrément d’analyse des expériences (a, d) de référence et (b, e) sans les observations radar. Différence d’analyses entre l’expérience assimilant les observations PWS, (c) en température et (f) en humidité relative, et l’analyse de référence. L’axe de la coupe est indiqué en pointillé sur la figure 4.4.	140
4.6	Cumul de précipitations sur 6 h (en mm) à partir de l’analyse de 18:00 UTC des expériences cyclées 3D <code>EnVar</code> , 3D <code>EnVarTS</code> et 3D <code>EnVarRHS</code> , ainsi que les observations Antilope.	141
5.1	RMS de la différence entre les observations de référence (SWS) en France et la prévision pour les expériences 3D <code>EnVar_PTRH</code> , 3D <code>EnVarP</code> , 3D <code>EnVarTS</code> et 3D <code>EnVarRHS</code> , normalisé par l’expérience 3D <code>EnVar</code> . Réalisées, avec 4 prévisions par jour pendant le mois d’étude, pour les observations de (a) pression, (b) de température, (c) d’humidité relative et (d) de force du vent. Les valeurs négatives (resp. positives) indiquent une baisse (resp. une augmentation) de l’écart à la prévision des observations SWS par rapport à l’expérience de contrôle, correspondant à un bénéfice (resp. une dégradation). La barre d’incertitude correspond aux quantiles Q10 et Q90 d’une méthode de bootstrapping.	145
5.2	RMS de la différence entre les observations de référence (SWS) en France et la prévision issue de l’expérience 3D <code>EnVar_PTRH</code> (normalisé par rapport à l’expérience de contrôle), pour les quatre réseaux de prévisions de 00:00 UTC, 06:00 UTC, 12:00 UTC et 18:00 UTC. Pour la pression de surface, la température à 2 m, la force du vent à 10 m et l’humidité spécifique à 2 m. Les valeurs négatives (resp. positives) indiquent une baisse (resp. une augmentation) de l’écart à la prévision de la référence, correspondant à un bénéfice (resp. une dégradation).	146
5.3	RMS de la différence entre la prévision issue de l’expérience 3D <code>EnVar_PTRH</code> et une référence d’altitude (normalisé par rapport à l’expérience de contrôle), pour le réseau de 18:00 UTC. La référence d’altitude correspond (en haut) à l’analyse AROME et (en bas) aux observations de radiosondages. Voir figure 5.2 pour plus de détails.	147
5.4	Scores de performances de Brier (BSS) pour les prévisions de 00:00 UTC durant le mois d’étude, moyenné sur quatre échéances (6 h, 12 h, 18 h et 24 h), pour les expériences 3D <code>Var</code> , 3D <code>EnVar</code> , 3D <code>EnVarP</code> , 3D <code>EnVarTS</code> , 3D <code>EnVarRHS</code> et 3D <code>EnVar_PTRH</code> . Réalisés avec une échelle spatiale de 50 km, pour les rafales 1 min supérieures à 40 km h ⁻¹ , le cumul de précipitations sur 6 h supérieure à 0.5 mm, 2 mm et 5 mm. La moyenne des quatre scores forme le score IP16. Le plus élevé, le meilleur est le score de Brier.	148
5.5	Temps CPU moyen pour chaque étape de l’assimilation et de la prévision d’AROME-France. En noir l’expérience de contrôle 3D <code>EnVar</code> et en marron 3D <code>EnVar_PTRH</code> . Réalisé sur 12 jours pour des raisons techniques. Voir texte pour la signification des étapes. Pour plus de détails sur le temps CPU, voir Annexe B.	148

A	Carte des différentes situations météorologiques étudiées.	156
---	--	-----

Liste des tableaux

1.1	Critères d’homologation d’une classe de performance maintenue pour les observations de température, d’humidité relative et de pression (Leroy, 2010). (*) Quand l’incertitude requise n’est pas actuellement atteignable, il est indiqué l’incertitude atteignable. La maintenance doit être réalisée dans un laboratoire accrédité.	9
1.2	Pourcentage des 600 stations du réseau RADOME pour chaque classe environnementale (le 1 février 2024). (*) 5 et les stations sans classe environnementale.	10
1.3	Caractéristiques des capteurs composant la station météo intelligente Netatmo. (*) Le prix est celui identifié sur le site du constructeur au 1 ^{er} février 2024.	12
1.4	Présentation des quatre principaux contrôles de qualité en température et humidité des observations PWS	17
1.5	Principales paramétrisations physiques d’un modèle de PNT à aire limitée et transcription dans le cas d’AROME-France (Fischer <i>et al.</i> , 2018)	23
2.1	Vue d’ensemble des schémas d’assimilation utilisés opérationnellement pour différents modèles à aire limitée. Pour plus de détails, voir Gustafsson <i>et al.</i> (2018).	43
2.2	Paramètres σ_o prescrits et σ_b estimés le 6 septembre 2021 à 01:00 UTC, pour différentes observations de surface.	45
3.1	Moyenne par pays comptant plus de 50 stations sur le domaine AROME de la moyenne et de l’écart type ($\text{m}^2 \text{s}^{-2}$) des séries temporelles d’écart à l’ébauche des observations de géopotential SWS et PWS. Réalisé sur le mois de la période d’étude. Les stations avec un écart-type supérieur à $100 \text{m}^2 \text{s}^{-2}$ sont enlevées, représentant moins de 2 % des stations. (*) Les stations PWS couvrent les zones frontalières des pays limitrophes à la France.	86
3.2	$\Delta\text{RMS OmB}$ (%) de pression de surface, température à 2 m, humidité relative à 2 m et du vent zonal à 10 m provenant des observations SWS sur le mois d’étude. Les valeurs négatives d’une expérience (XP) par rapport à l’expérience 3DEnVar indiquent que les prévisions à 1 h de XP sont plus proches des observations SWS que celles de l’expérience 3DEnVar. On nommera « gain », les valeurs négatives de $\Delta\text{RMS OmB}$. Les valeurs significatives sont en gras.	90
4.1	Statistiques des écarts à l’ébauche des observations de SWS et PWS. Pour chaque réseau d’observations de température (resp. d’humidité relative), <i>nb</i> indique le nombre de stations, <i>GE</i> le nombre d’observations dont l’écart à l’ébauche est supérieurs à 15 °C (resp. à 80 %), <i>biais</i> le nombre de stations avec un biais mensuel supérieur à 4 °C (resp. à 25 %), et <i>rms</i> supérieur à 6 °C (resp. à 40 %), à chaque fois sur l’ensemble du mois d’étude. Ces seuils correspondent aux seuils utilisés en opérationnel. Les PWS pré-traitées correspondent aux observations qui ont passé le pré-traitement P-S (section 4.2).	135

5.1	Définition et rappel des expériences réalisées sur une période d'un mois du 6 septembre au 5 octobre 2021. Chaque expérience correspond à un cycle d'assimilation horaire et des prévisions à 48 h d'échéance quatre fois par jour. . . .	144
5.2	Δ RMS OmB (%) de pression de surface, température à 2 m , humidité relative à 2 m et du vent zonal à 10 m provenant des observations SWS sur le mois d'étude. Les valeurs négatives d'une expérience (XP) par rapport à l'expérience 3DEnVar indiquent que les prévisions à 1 h de XP sont plus proches que celles de l'expérience 3DEnVar, des observations SWS. Les valeurs significatives sont en gras.	144
B	Comparaison de l'intensité carbone d'une heure de calcul sur CPU. ⁽¹⁾ D'après le site et la note technique du Labo1.5 avec une incertitude globale estimée de 40 % Benaben <i>et al.</i> (2024). ⁽²⁾ Grenoble Alpes Recherche Infrastructure de Calcul Intensif et de Données (GRICAD, Berthoud <i>et al.</i> , 2020 ; sans l'empreinte du personnel).	161

Références bibliographiques

- Amodei, M., Sanchez, I., & Stein, J. (2015). Verification of the French operational high-resolution model AROME with the regional Brier probability score. *Meteorological Applications*, 22(4), 731-745. <https://doi.org/10.1002/met.1510>
- Anderson, E., & Järvinen, H. (1999). Variational quality control. *Quarterly Journal of the Royal Meteorological Society*, 125(554), 697-722. <https://doi.org/10.1002/qj.49712555416>
- Askenazy, P., Bustarret, E., Carayol, H., Guilyardi, E., Jacquier, M., Leclerc, O., Noiville, C., & Thiébaud, S. (2022). Intégrer les enjeux environnementaux à la conduite de la recherche – Une responsabilité éthique. *COMETS AVIS n°2022-43*. <https://hal.science/hal-03907937>
- Auligné, T., McNally, A. P., & Dee, D. P. (2007). Adaptive bias correction for satellite data in a numerical weather prediction system. *Quarterly Journal of the Royal Meteorological Society : A journal of the atmospheric sciences, applied meteorology and physical oceanography*, 133(624), 631-642. <https://doi.org/10.1002/qj.56>
- Bakketun, Å., Blyverket, J., & Müller, M. (2023). Using a Reanalysis-Driven Land Surface Model for Initialization of a Numerical Weather Prediction System. *Weather and Forecasting*, 38(11), 2155-2168. <https://doi.org/10.1175/WAF-D-22-0184.1>
- Baldauf, M., Seifert, A., Förstner, J., Majewski, D., Raschendorfer, M., & Reinhardt, T. (2011). Operational Convective-Scale Numerical Weather Prediction with the COSMO Model : Description and Sensitivities. *Monthly Weather Review*, 139(12), 3887-3905. <https://doi.org/10.1175/MWR-D-10-05013.1>
- Beele, E., Reyniers, M., Aerts, R., & Somers, B. (2022). Quality control and correction method for air temperature data from a citizen science weather station network in Leuven, Belgium. *Earth System Science Data*, 14(10), 4681-4717. <https://doi.org/10.5194/essd-14-4681-2022>
- Bell, A., Martinet, P., Caumont, O., Burnet, F., Delanoë, J., Jorquera, S., Seity, Y., & Unger, V. (2022). An optimal estimation algorithm for the retrieval of fog and low cloud thermodynamic and micro-physical properties. *Atmospheric Measurement Techniques*, 15(18), 5415-5438. <https://doi.org/10.5194/amt-15-5415-2022>
- Bell, S., Cornford, D., & Bastin, L. (2015). How good are citizen weather stations? Addressing a biased opinion. *Weather*, 70(3), 75-84. <https://doi.org/10.1002/wea.2316>
- Benaben, D., Berthoud, F., Guennebaud, G., Ligozat, A.-L., & Valcke, S. (2024). *Estimation de l'empreinte carbone d'une heure de calcul sur un cœur CPU ou sur un GPU* (rapp. tech.). Labo-1.5. <https://labos1point5.org/les-rapports/estimation-empreinte-calcul>
- Bénard, P., Vivoda, J., Mašek, J., Smolíková, P., Yessad, K., Smith, C., Brožková, R., & Geleyn, J.-F. (2010). Dynamical kernel of the Aladin–NH spectral limited-area model : Revised formulation and sensitivity experiments. *Quarterly Journal of the Royal Meteorological Society : A journal of the atmospheric sciences, applied meteorology and physical oceanography*, 136(646), 155-169. <https://doi.org/10.1002/qj.522>

- Benjamin, S. G., Weygandt, S. S., Brown, J. M., Hu, M., Alexander, C. R., Smirnova, T. G., Olson, J. B., James, E. P., Dowell, D. C., Grell, G. A., et al. (2016). A North American hourly assimilation and model forecast cycle : The Rapid Refresh. *Monthly Weather Review*, 144(4), 1669-1694. <https://doi.org/10.1175/MWR-D-15-0242.1>
- Bernède, L. (2022). Observation participative et gestion des épisodes neigeux. *La Météorologie*, 117, 24-32. <https://doi.org/10.37053/lameteorologie-2022-0033>
- Berre, L. (2000). Estimation of synoptic and mesoscale forecast error covariances in a limited-area model. *Monthly weather review*, 128(3), 644-667. [https://doi.org/10.1175/1520-0493\(2000\)128<0644:EOSAMF>2.0.CO;2](https://doi.org/10.1175/1520-0493(2000)128<0644:EOSAMF>2.0.CO;2)
- Berthoud, F., Bzeznik, B., Gibelin, N., Laurens, M., Bonamy, C., Morel, M., & Schwindenhammer, X. (2020). Estimation de l’empreinte carbonée d’une heure de calcul. <https://gricad.univ-grenoble-alpes.fr/documents/GES-h-coeur-GRICAD-2020.pdf>
- Bougeault, P., & Lacarrere, P. (1989). Parameterization of orography-induced turbulence in a mesobeta-scale model. *Monthly weather review*, 117(8), 1872-1890. [https://doi.org/10.1175/1520-0493\(1989\)117<1872:POOITI>2.0.CO;2](https://doi.org/10.1175/1520-0493(1989)117<1872:POOITI>2.0.CO;2)
- Bouteloup, Y., Seity, Y., & Bazile, E. (2011). Description of the sedimentation scheme used operationally in all Météo-France NWP models. *Tellus A : Dynamic Meteorology and Oceanography*, 63(2), 300-311. <https://doi.org/10.1111/j.1600-0870.2010.00484.x>
- Brousseau, P., Desroziers, G., Bouttier, F., & Chapnik, B. (2014). A posteriori diagnostics of the impact of observations on the AROME-France convective-scale data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 140(680), 982-994. <https://doi.org/10.1002/qj.2179>
- Brousseau, P., Seity, Y., Ricard, D., & Léger, J. (2016). Improvement of the forecast of convective activity from the AROME-France system. *Quarterly Journal of the Royal Meteorological Society*, 142(699), 2231-2243. <https://doi.org/10.1002/qj.2822>
- Bubnová, R., Hello, G., Bénard, P., & Geleyn, J.-F. (1995). Integration of the fully elastic equations cast in the hydrostatic pressure terrain-following coordinate in the framework of the ARPEGE/Aladin NWP system. *Monthly weather review*, 123(2), 515-535. [https://doi.org/10.1175/1520-0493\(1995\)123<0515:IOTFEE>2.0.CO;2](https://doi.org/10.1175/1520-0493(1995)123<0515:IOTFEE>2.0.CO;2)
- Büchau, Y. G. (2018). *Modelling Shielded Temperature Sensors : an Assessment of the Netatmo Citizen Weather Station* [MSc Thesis, University of Hamburg] [https://bis-erdsystem.de/fileadmin/user_upload/bise/Texte/MSC_Met_61-20180313.pdf (dernier accès : 4 January 2024)].
- Burgot, T. (2021). *Viabilité et performance des solveurs points de grille dans la résolution du problème implicite du noyau dynamique du modèle AROME* [<http://www.theses.fr/2021INPT0006>]. <http://www.theses.fr/2021INPT0006>
- Calas, C. (1997). *Apport d’une analyse de méso-échelle pour la prévision immédiate des systèmes convectifs* [thèse de doct., Université de Toulouse]. <https://theses.fr/1997TOU30282>
- Caron, J.-F., & Buehner, M. (2022). Implementation of scale-dependent background-error covariance localization in the canadian global deterministic prediction system. *Weather and Forecasting*, 37(9), 1567-1580. <https://doi.org/10.1175/WAF-D-22-0055.1>

- Caron, J.-F., Michel, Y., Montmerle, T., & Arbogast, É. (2019). Improving Background Error Covariances in a 3D Ensemble–Variational Data Assimilation System for Regional NWP. *Monthly Weather Review*, *147*(1), 135-151. <https://doi.org/10.1175/mwr-d-18-0248.1>
- Caumont, O., Mandement, M., Bouttier, F., Eeckman, J., Lebeaupin Brossier, C., Lovat, A., Nuissier, O., & Laurantin, O. (2021). The heavy precipitation event of 14–15 October 2018 in the Aude catchment : a meteorological study based on operational numerical weather prediction systems and standard and personal observations. *Natural Hazards and Earth System Sciences*, *21*(3), 1135-1157. <https://doi.org/10.5194/nhess-21-1135-2021>
- Caumont, O., Ducrocq, V., Wattrelot, É., Jaubert, G., & Pradier-Vabre, S. (2010). 1D3DVar assimilation of radar reflectivity data : a proof of concept. *Tellus A : Dynamic Meteorology and Oceanography*, *62*(2), 173. <https://doi.org/10.1111/j.1600-0870.2009.00430.x>
- Chapman, L., Bell, C., & Bell, S. (2017). Can the crowdsourcing data paradigm take atmospheric science to a new level ? A case study of the urban heat island of London quantified using Netatmo weather stations. *International Journal of Climatology*, *37*(9), 3597-3605. <https://doi.org/10.1002/joc.4940>
- Chen, K., Bai, L., Ling, F., Ye, P., Chen, T., Chen, K., Han, T., & Ouyang, W. (2023). Towards an end-to-end artificial intelligence driven global weather forecasting system. *arXiv preprint*. <https://arxiv.org/abs/2312.12462>
- Cintineo, R. M., & Stensrud, D. J. (2013). On the predictability of supercell thunderstorm evolution. *Journal of the Atmospheric Sciences*, *70*(7), 1993-2011. <https://doi.org/10.1175/JAS-D-12-0166.1>
- Clark, M. R., Webb, J. D. C., & Kirk, P. J. (2018). Fine-scale analysis of a severe hailstorm using crowd-sourced and conventional observations. *Meteorological Applications*, *25*(3), 472-492. <https://doi.org/10.1002/met.1715>
- Combarous, P. (2023). *Assimilation de données de l'imageur d'éclairs de Météosat troisième génération pour la prévision numérique des orages à échelle convective* [thèse de doct., Université Paul Sabatier-Toulouse III]. <https://theses.hal.science/tel-04474838>
- Coney, J., Pickering, B., Dufton, D., Lukach, M., Brooks, B., & Neely III, R. R. (2022). How useful are crowdsourced air temperature observations ? An assessment of Netatmo stations and quality control schemes over the United Kingdom. *Meteorological Applications*, *29*(3), e2075. <https://doi.org/https://doi.org/10.1002/met.2075>
- Coopmann, O., Fourrié, N., Chambon, P., Vidot, J., Brousseau, P., Martet, M., & Birman, C. (2023). Preparing the assimilation of the future MTG-IRS sounder into the mesoscale numerical weather prediction AROME model. *Quarterly Journal of the Royal Meteorological Society*, *149*(757), 3110-3134. <https://doi.org/10.1002/qj.4548>
- Cornes, R. C., Dirksen, M., & Sluiter, R. (2019). Correcting citizen-science air temperature measurements across the Netherlands for short wave radiation bias. *Meteorological Applications*, *27*(1). <https://doi.org/10.1002/met.1814>
- Courtier, P., Andersson, E., Heckley, W., Vasiljevic, D., Hamrud, M., Hollingsworth, A., Rabier, F., Fisher, M., & Pailleux, J. (1998). The ECMWF implementation of three-dimensional variational assimilation (3D-Var). I : Formulation. *Quarterly Journal of the Royal Meteorological Society*, *124*(550), 1783-1807. <https://doi.org/10.1002/qj.49712455002>

- Cuxart, J., Bougeault, P., & Redelsperger, J.-L. (2000). A turbulence scheme allowing for mesoscale and large-eddy simulations. *Quarterly Journal of the Royal Meteorological Society*, *126*(562), 1-30. <https://doi.org/10.1002/qj.49712656202>
- De Vos, L., Leijnse, H., Overeem, A., & Uijlenhoet, R. (2017). The Potential of Urban Rainfall Monitoring with Crowdsourced Automatic Weather Stations in Amsterdam. *Hydrol. Earth Syst. Sci*, *21*(2), 765-777. <https://doi.org/10.5194/hess-21-765-2017>
- Desroziers, G., Berre, L., Chapnik, B., & Poli, P. (2005). Diagnosis of observation, background and analysis-error statistics in observation space. *Quarterly Journal of the Royal Meteorological Society*, *131*(613), 3385-3396. <https://doi.org/10.1256/qj.05.108>
- Dharssi, I. (2012). *Towards a Kalman Filter based land surface data assimilation scheme for ACCESS* (T. 115). https://www.cawcr.gov.au/technical-reports/CTR_054.pdf
- Diefenbach, T., Craig, G., Scheck, L., & Martin, W. (2023). *Measures of imbalance in convective-scale data assimilation* (rapp. tech.). Bologna, Italie. <https://eventi.unibo.it/isda2023/contributions>
- Drusch, M. (2007). Initializing numerical weather prediction models with satellite-derived surface soil moisture : Data assimilation experiments with ECMWF's Integrated Forecast System and the TMI soil moisture data set. *Journal of Geophysical Research : Atmospheres*, *112*(D3). <https://doi.org/10.1029/2006JD007478>
- Ducrocq, V., Lapore, J.-P., Redelsperger, J.-L., & Orain, F. (2000). Initialization of a fine-scale model for convective-system prediction : A case study. *Quarterly Journal of the Royal Meteorological Society*, *126*(570), 3041-3065. <https://doi.org/10.1002/qj.49712657004>
- Ducrocq, V., Ricard, D., Lafore, J.-P., & Orain, F. (2002). Storm-scale numerical rainfall prediction for five precipitating events over France : On the importance of the initial humidity field. *Weather and Forecasting*, *17*(6), 1236-1256. [https://doi.org/10.1175/1520-0434\(2002\)017<1236:SSNRPF>2.0.CO;2](https://doi.org/10.1175/1520-0434(2002)017<1236:SSNRPF>2.0.CO;2)
- ECMWF. (2023). IFS Documentation CY48R1 - Part I : Observations. <https://doi.org/10.21957/0f360ba4ca>
- Engerer, N. A., Stensrud, D. J., & Coniglio, M. C. (2008). Surface characteristics of observed cold pools. *Monthly Weather Review*, *136*(12), 4839-4849. <https://doi.org/10.1175/2008MWR2528.1>
- Fairall, C. W., Bradley, E. F., Hare, J. E., Grachev, A. A., & Edson, J. B. (2003). Bulk Parameterization of Air–Sea Fluxes : Updates and Verification for the COARE Algorithm. *Journal of Climate*, *16*(4), 571-591. [https://doi.org/10.1175/1520-0442\(2003\)016<0571:BPOASF>2.0.CO;2](https://doi.org/10.1175/1520-0442(2003)016<0571:BPOASF>2.0.CO;2)
- Feichtinger, M., de Wit, R., Goldenits, G., Kolejka, T., Hollósi, B., Žuvela-Aloise, M., & Feigl, J. (2020). Case-study of neighborhood-scale summertime urban air temperature for the City of Vienna using crowd-sourced data. *Urban Climate*, *32*, 100597. <https://doi.org/10.1016/j.uclim.2020.100597>
- Fenner, D., Bechtel, B., Demuzere, M., Kittner, J., & Meier, F. (2021). CrowdQC+—A Quality-Control for Crowdsourced Air-Temperature Observations Enabling World-Wide Urban Climate Applications. *Frontiers in Environmental Science*, *9*. <https://doi.org/10.3389/fenvs.2021.720747>
- Fenner, D., Meier, F., Bechtel, B., Otto, M., & Scherer, D. (2017). Intra and inter local climate zone variability of air temperature as observed by crowdsourced citi-

- zen weather stations in Berlin, Germany. *Meteorologische Zeitschrift*, 26, 525-547. <https://doi.org/10.1127/metz/2017/0861>
- Fischer, C., Bouyssel, F., Brousseau, P., El Khatib, R., Pottier, P., Seity, Y., Wattrelot, É., & Joly, A. (2018). Les modèles opérationnels de prévision numérique à aire limitée de Météo-France. *La Météorologie*, 2018(100), 18-28. <https://doi.org/10.4267/2042/65139>
- Fischer, C., Montmerle, T., Berre, L., Auger, L., & Ștefănescu, S. E. (2005). An overview of the variational assimilation in the ALADIN/France numerical weather-prediction system. *Quarterly Journal of the Royal Meteorological Society : A journal of the atmospheric sciences, applied meteorology and physical oceanography*, 131(613), 3477-3492. <https://doi.org/10.1256/qj.05.115>
- Forster, A., Masson, V., & Augros, C. (2023). Study of the urban effect of Paris on several thunderstorm cases in 2022, using hectometric ensemble simulations. *EGU General Assembly Conference Abstracts*, EGU-3249. <https://doi.org/10.5194/egusphere-egu23-3249>
- Fragnière, A., Banwell, N., Suckow, N., Butera, F., Chaperon, D., Chèvre, N., Demay, C., Perrot, P., Eggert, N., Gagnebin, P., Goudet, J., Kaufmann, A., & Palazzo, G. (2022). L'engagement public des universitaires. Entre liberté académique et déontologie professionnelle. *Université de Lausanne*. https://serval.unil.ch/fr/notice/serval:BIB_3053895AFBAF (dernier accès : 1 mai 2024).
- Francis, D. J., Fowler, A. M., Lawless, A. S., Eyre, J., & Migliorini, S. (2023). The effective use of anchor observations in variational bias correction in the presence of model bias. *Quarterly Journal of the Royal Meteorological Society*, 149(754), 1789-1809. <https://doi.org/10.1002/qj.4482>
- Fujita, T. T. (1981). Tornadoes and downbursts in the context of generalized planetary scales. *Journal of Atmospheric Sciences*, 38(8), 1511-1534. [https://doi.org/10.1175/1520-0469\(1981\)038<1511:TADITC>2.0.CO;2](https://doi.org/10.1175/1520-0469(1981)038<1511:TADITC>2.0.CO;2)
- Geer, A. J. (2021). Learning earth system models from observations : machine learning or data assimilation? *Philosophical Transactions of the Royal Society A*, 379(2194), 20200089. <https://doi.org/10.1098/rsta.2020.0089>
- Geleyn, J.-F. (1988). Interpolation of wind, temperature and humidity values from model levels to the height of measurement. *Tellus A*, 40(4), 347-351. <https://doi.org/10.1111/j.1600-0870.1988.tb00352.x>
- Giard, D., & Bazile, E. (2000). Implementation of a New Assimilation Scheme for Soil and Surface Variables in a Global NWP Model. *Monthly Weather Review*, 128(4), 997-1015. [https://doi.org/10.1175/1520-0493\(2000\)128<0997:IOANAS>2.0.CO;2](https://doi.org/10.1175/1520-0493(2000)128<0997:IOANAS>2.0.CO;2)
- Grassmann, T., Napoly, A., Meier, F., & Fenner, D. (2018). Quality Control for Crowdsourced Data from CWS. Accessed November 12, 2021. <https://doi.org/10.14279/depositonce-6740.3>
- Guillet, O., Weaver, A. T., Vasseur, X., Michel, Y., Gratton, S., & Gürol, S. (2019). Modelling spatially correlated observation errors in variational data assimilation using a diffusion operator on an unstructured mesh. *Quarterly Journal of the Royal Meteorological Society*, 145(722), 1947-1967. <https://doi.org/10.1002/qj.3537>
- Guillou, Y. (2018). L'automatisation des observations météorologiques de surface à Météo-France. Une r-évolution. *La Météorologie*, 100(Numéro Spécial Anniversaire 25 ans), 96-99. <https://doi.org/10.4267/2042/65150>
- Gustafsson, N., Janjić, T., Schraff, C., Leuenberger, D., Weissmann, M., Reich, H., Brousseau, P., Montmerle, T., Wattrelot, E., Bučánek, A., Mile, M., Hamdi, R., Lind-

- skog, M., Barkmeijer, J., Dahlbom, M., Macpherson, B., Ballard, S., Inverarity, G., Carley, J., . . . Fujita, T. (2018). Survey of data assimilation methods for convective-scale numerical weather prediction at operational centres. *Quarterly Journal of the Royal Meteorological Society*, *144*(713), 1218-1256. <https://doi.org/10.1002/qj.3179>
- Haeffelin, M., Kotthaus, S., Bastin, S., Bouffies-Cloch e, S., Cantrell, C., Christen, A., Dupont, J.-C., Foret, G., Gros, V., Lemonsu, A., et al. (2023). PANAME-Project synergy of atmospheric research in the Paris region. *EGU General Assembly Conference Abstracts*, EGU-14781. <https://doi.org/10.5194/egusphere-egu23-14781>
- Haertel, P. T., & Johnson, R. H. (2000). The linear dynamics of squall line mesohighs and wake lows. *Journal of the Atmospheric Sciences*, *57*(1), 93-107. [https://doi.org/10.1175/1520-0469\(2000\)057<0093:TLDOSL>2.0.CO;2](https://doi.org/10.1175/1520-0469(2000)057<0093:TLDOSL>2.0.CO;2)
- Hamill, T. M., Whitaker, J. S., & Snyder, C. (2001). Distance-Dependent Filtering of Background Error Covariance Estimates in an Ensemble Kalman Filter. *Monthly Weather Review*, *129*(11), 2776-2790. [https://doi.org/10.1175/1520-0493\(2001\)129<2776:DDFOBE>2.0.CO;2](https://doi.org/10.1175/1520-0493(2001)129<2776:DDFOBE>2.0.CO;2)
- Hess, R. (2001). Assimilation of screen-level observations by variational soil moisture analysis. *Meteorology and Atmospheric Physics*, *77*(1), 145-154. <https://doi.org/10.1007/s007030170023>
- Hintz, K. S. (2019). On The Usage of Crowdsourced Data in Numerical Weather Prediction. https://nbi.ku.dk/english/theses/phd-theses/phd_theses_2019/kasper_stener_hintzs/Kasper_Stener_Hintzs.pdf
- Houtekamer, P. L., & Zhang, F. (2016). Review of the ensemble Kalman filter for atmospheric data assimilation. *Monthly Weather Review*, *144*(12), 4489-4532. <https://doi.org/10.1175/MWR-D-15-0440.1>
- Houze Jr, R. A. (2014). *Cloud dynamics, second edition. International Geophysics Series*. Academic press.
- Houze Jr, R. A. (2018). 100 years of research on mesoscale convective systems. *Meteorological Monographs*, *59*, 17-1. <https://doi.org/10.1175/AMSMONOGRAPHS-D-18-0001.1>
- Hu, G., Dance, S. L., Bannister, R. N., Chipilski, H. G., Guillet, O., Macpherson, B., Weissmann, M., & Yussouf, N. (2022). Progress, challenges, and future steps in data assimilation for convection-permitting numerical weather prediction : Report on the virtual meeting held on 10 and 12 November 2021. *Atmospheric Science Letters*. <https://doi.org/10.1002/asl.1130>
- Huang, B., Wang, X., Kleist, D. T., & Lei, T. (2021). A simultaneous multiscale data assimilation using scale-dependent localization in GSI-based hybrid 4D-EnVar for NCEP FV3-based GFS. *Monthly Weather Review*, *149*(2), 479-501. <https://doi.org/10.1175/MWR-D-20-0166.1>
- Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A., & Collins, W. D. (2008). Radiative forcing by long-lived greenhouse gases : Calculations with the AER radiative transfer models. *Journal of Geophysical Research : Atmospheres*, *113*(D13). <https://doi.org/10.1029/2008JD009944>
- Ingleby, B. (2015). Global assimilation of air temperature, humidity, wind and pressure from surface stations. *Quarterly Journal of the Royal Meteorological Society*, *141*(687), 504-517. <https://doi.org/10.1002/qj.2372>
- Ingleby, B. (2024). ECMWF uses Mode-S aircraft observations again to improve forecasts. *ECMWF*. <https://www.ecmwf.int/en/about/media-centre/news/2024/ecmwf->

- uses-mode-s-aircraft-observations-again-improve-forecasts (dernier accès : 1 mai 2024).
- Ingleby, B., Arduini, G., Balsamo, G., Boussetta, S., Ochi, K., Pinnington, E., & de Rosnay, P. (2024). Improved two-metre temperature forecasts in the 2024 upgrade. *ECMWF Newsletter*. <https://www.ecmwf.int/en/newsletter/178/earth-system-science/improved-two-metre-temperature-forecasts-2024-upgrade> (dernier accès : 1 mai 2024).
- Johnson, A., Wang, X., & Degelia, S. (2017). Design and implementation of a GSI-based convection-allowing ensemble-based data assimilation and forecast system for the PECAN field experiment. Part II : Overview and evaluation of a real-time system. *Weather and Forecasting*, 32(3), 1227-1251. <https://doi.org/10.1175/WAF-D-16-0102.1>
- Johnson, R. H., & Mapes, B. E. (2001). Mesoscale processes and severe convective weather. In *Severe convective storms* (p. 71-122). Springer. [https://doi.org/10.1175/1520-0477\(2001\)082<0013:SMAM>2.3.CO;2](https://doi.org/10.1175/1520-0477(2001)082<0013:SMAM>2.3.CO;2)
- Lannelongue, L., Grealey, J., & Inouye, M. (2021). Green Algorithms : Quantifying the Carbon Footprint of Computation. *Advanced Science*, 8(12), 2100707. <https://doi.org/10.1002/advs.202100707>
- Lenfle, S. (2018). De l'interpolation optimale au 4D-Var : l'émergence d'un nouveau dominant design en assimilation de données météorologiques. *La Météorologie*, (100), 37-46. <https://doi.org/10.4267/2042/65141>
- Leroy, M. (2010). Note technique n°37 : classification de performance maintenue. <http://gedmf.meteo.fr/share/page/document-details?nodeRef=workspace://SpacesStore/9721615a-bd6e-40d9-a909-47c1bb46d055> (rapport interne, dernier accès : 1 mai 2024).
- Madaus, L. E., & Hakim, G. J. (2016). Observable Surface Anomalies Preceding Simulated Isolated Convective Initiation. *Monthly Weather Review*, 144(6), 2265-2284. <https://doi.org/10.1175/MWR-D-15-0332.1>
- Madelin, M., & Dupuis, V. (2020). Intensité et spatialisation de l'îlot de chaleur urbain parisien à partir de données participatives. *Climatologie*, 17, 9. <https://doi.org/10.1051/climat/202017009>
- Malardel, S. (2009). *Fondamentaux de météorologie*. Cépadués Editions.
- Malardel, S. (2022). *A propos de résolution* [Habilitation à diriger des recherches. Non publiée, université de Toulouse, Toulouse].
- Mandement, M., & Caumont, O. (2021). A numerical study to investigate the roles of former Hurricane Leslie, orography and evaporative cooling in the 2018 Aude heavy-precipitation event. *Weather and Climate Dynamics*, 2(3), 795-818. <https://doi.org/10.5194/wcd-2-795-2021>
- Mandement, M. (2020). *Apport des données d'objets connectés pour l'étude de la convection profonde à fine échelle* [thèse de doct., Université de Toulouse]. <https://theses.hal.science/tel-04172418/>
- Mandement, M., & Caumont, O. (2020). Contribution of personal weather stations to the observation of deep-convection features near the ground. *Natural Hazards and Earth System Sciences*, 20(1), 299-322. <https://doi.org/10.5194/nhess-20-299-2020>
- Mandement, M., Delabre, L., & Lac, C. Simulations du coup de vent du 18 juin 2022 sur la Normandie [http://mesonh.aero.obs-mip.fr/mesonh57/TwelfthUsersMeeting?action=AttachFile&do=get&target=20231018_1645_mandement.pdf]. In : In

- 12th Meso-NH User's Meeting, 18–20 October 2023, Observatoire Midi-Pyrénées, Toulouse, France. http://mesonh.aero.obs-mip.fr/mesonh57/TwelfthUsersMeeting?action=AttachFile&do=get&target=20231018_1645_mandement.pdf. 2023, octobre. http://mesonh.aero.obs-mip.fr/mesonh57/TwelfthUsersMeeting?action=AttachFile&do=get&target=20231018_1645_mandement.pdf
- Markowski, P., & Richardson, Y. (2011). *Mesoscale meteorology in midlatitudes*. John Wiley & Sons.
- Marques, E. (2023). *Étude à fine échelle de l'îlot de chaleur urbain par modélisation bayésienne à partir de données opportunes* [thèse de doct., Institut National Polytechnique de Toulouse]. <https://oatao.univ-toulouse.fr/29630/>
- Marquès, E., Masson, V., Naveau, P., Mestre, O., Dubreuil, V., & Richard, Y. (2022). Urban heat island estimation from crowdsensing thermometers embedded in personal cars. *Bulletin of the American Meteorological Society*, 103(4), E1098-E1113. <https://doi.org/10.1175/BAMS-D-21-0174.1>
- Masson, V., Gomes, L., Pigeon, G., et al. (2008). The Canopy and Aerosol Particles Interactions in TOulouse Urban Layer (CAPITOU) experiment. *Meteorol Atmos Phys*, 102, 135-157. <https://doi.org/10.1007/s00703-008-0289-4>
- Masson, V., Le Moigne, P., Martin, E., Faroux, S., Alias, A., Alkama, R., Belamari, S., Barbu, A., Boone, A., Bouyssel, F., Brousseau, P., Brun, E., Calvet, J.-C., Carrer, D., Decharme, B., Delire, C., Donier, S., Essauouini, K., Gibelin, A.-L., . . . Voldoire, A. (2013). The SURFEXv7.2 land and ocean surface platform for coupled or offline simulation of earth surface variables and fluxes. *Geoscientific Model Development*, 6(4), 929-960. <https://doi.org/10.5194/gmd-6-929-2013>
- Masson, V. (2000). A physically-based scheme for the urban energy budget in atmospheric models. *Boundary-layer meteorology*, 94, 357-397. <https://doi.org/10.1023/A:1002463829265>
- Masson, V., & Seity, Y. (2009). Including atmospheric layers in vegetation and urban offline surface schemes. *Journal of Applied Meteorology and Climatology*, 48(7), 1377-1397. <https://doi.org/10.1175/2009JAMC1866.1>
- McNicholas, C., & Mass, C. F. (2018a). Impacts of Assimilating Smartphone Pressure Observations on Forecast Skill during Two Case Studies in the Pacific Northwest. *Weather and Forecasting*, 33(5), 1375-1396. <https://doi.org/10.1175/waf-d-18-0085.1>
- McNicholas, C., & Mass, C. F. (2018b). Smartphone Pressure Collection and Bias Correction Using Machine Learning. *Journal of Atmospheric and Oceanic Technology*, 35(3), 523-540. <https://doi.org/10.1175/jtech-d-17-0096.1>
- Meier, F., Fenner, D., Grassmann, T., Otto, M., & Scherer, D. (2017). Crowdsourcing air temperature from citizen weather stations for urban climate research. *Urban Climate*, 19, 170-191. <https://doi.org/10.1016/j.uclim.2017.01.006>
- Merker, C., Leuenberger, D., Crezee, B., Regenass, D., & Arpagaus, M. (2023). Benefit and challenges in assimilating near-surface temperature and humidity observations in complex terrain. *ISDA*. https://eventi.unibo.it/isda2023/contributions/merker_claire_assimilationofnearsurfacedataatmeteoswiss.pdf/@download/file/Merker_Claire_AssimilationOfNearSurfaceDataAtMeteoswiss.pdf
- Météo-France. (2017). Évaluation des stations Netatmo (rapport interne). <http://gedmf.meteo.fr/share/page/document-details?nodeRef=workspace://SpacesStore/0ae19217-5f6b-46ad-b7f0-13b46d00b81b> (dernier accès : 1 février 2024).

- Météo-France. (2020). Une brève histoire de l'observation. <https://meteofrance.com/magazine/meteo-histoire/observation/une-breve-histoire-de-lobserver>
- Météo-France. (2023). 2022-2026 Contrat d'objectifs et de performance de Météo-France. <https://meteofrance.fr/etablissement/strategie/le-contrat-dobjectifs-et-de-performance> (dernier accès : 1 Mai 2024).
- Michel, Y., & Brousseau, P. (2021). A Square-Root, Dual-Resolution 3DEnVar for the AROME Model : Formulation and Evaluation on a Summertime Convective Period. *Monthly Weather Review*, *149*(9), 3135-3153. <https://doi.org/10.1175/MWR-D-21-0026.1>
- Mironov, D., Heise, E., Kourzeneva, E., Ritter, B., Schneider, N., & Terzhevik, A. (2010). Implementation of the lake parameterisation scheme FLake into the numerical weather prediction model COSMO. *Boreal environment research*, *15*(2). <https://helda.helsinki.fi/server/api/core/bitstreams/d158b23f-f705-4559-aa9d-7dc6508a28e9/content> (dernier accès : 1 mai 2024).
- Mlawer, E. J., Taubman, S. J., Brown, P. D., Iacono, M. J., & Clough, S. A. (1997). Radiative transfer for inhomogeneous atmospheres : RRTM, a validated correlated-k model for the longwave. *Journal of Geophysical Research : Atmospheres*, *102*(D14), 16663-16682. <https://doi.org/10.1029/97JD00237>
- Montmerle, T., Michel, Y., Arbogast, E., Ménétrier, B., & Brousseau, P. (2018). A 3D ensemble variational data assimilation scheme for the limited-area AROME model : Formulation and preliminary results. *Quarterly Journal of the Royal Meteorological Society*, *144*(716), 2196-2215. <https://doi.org/10.1002/qj.3334>
- Muller, C., Chapman, L., Johnston, S., Kidd, C., Illingworth, S., Foody, G., Overeem, A., & Leigh, R. (2015). Crowdsourcing for climate and atmospheric sciences : current status and future potential. *International Journal of Climatology*, *35*(11), 3185-3203. <https://doi.org/10.1002/joc.4210>
- Napoly, A., Grassmann, T., Meier, F., & Fenner, D. (2018). Development and Application of a Statistically-Based Quality Control for Crowdsourced Air Temperature Data. *Frontiers in Earth Science*, *6*. <https://doi.org/10.3389/feart.2018.00118>
- Nipen, T. N., Seierstad, I. A., Lussana, C., Kristiansen, J., & Hov, Ø. (2020). Adopting Citizen Observations in Operational Weather Prediction. *Bulletin of the American Meteorological Society*, *101*(1), E43-E57. <https://doi.org/10.1175/BAMS-D-18-0237.1>
- Noilhan, J., & Planton, S. (1989). A Simple Parameterization of Land Surface Processes for Meteorological Models. *Monthly Weather Review*, *117*(3), 536-549. [https://doi.org/10.1175/1520-0493\(1989\)117<0536:ASPOLS>2.0.CO;2](https://doi.org/10.1175/1520-0493(1989)117<0536:ASPOLS>2.0.CO;2)
- Odin, M., Barret, M., Biancamaria, S., Dassas, K., Firmin, A., Gandois, L., Gheusi, F., Kuppel, S., Maisonobe, M., Mialon, A., Monnier, L., Pantillon, F., & Toubanc, F. (2024). Comprehensive carbon footprint of Earth and environmental science laboratories : implications for sustainable scientific practice. *ESS Open Archive*. <https://doi.org/10.22541/essoar.170965077.73149239/v1>
- Organisation Météorologique Mondiale. (2023). *Guide des instruments et des méthodes d'observation, Volume I – Mesure de variables météorologiques (WMO-No. 8)*. <https://library.wmo.int/idurl/4/68713>
- Orlanski, I. (1975a). A rational subdivision of scales for atmospheric processes. *Bulletin of the American Meteorological Society*, 527-530. <https://www.jstor.org/stable/26216020>

- Orlanski, I. (1975b). A Rational Subdivision of Scales for Atmospheric Processes. *Bulletin of the American Meteorological Society*, 56(5), 527-530. <http://www.jstor.org/stable/26216020> (dernier accès : 1 Mai 2024).
- Peherstorfer, B., Willcox, K., & Gunzburger, M. (2018). Survey of Multifidelity Methods in Uncertainty Propagation, Inference, and Optimization. *SIAM Review*, 60(3), 550-591. <https://doi.org/10.1137/16M1082469>
- Pergaud, J., Masson, V., Malardel, S., & Couvreur, F. (2009). A Parameterization of Dry Thermals and Shallow Cumuli for Mesoscale Numerical Weather Prediction. *Boundary-Layer Meteorol*, 132, 83-106. <https://doi.org/10.1007/s10546-009-9388-0>
- Perrin de Brichambaut, C., & Leroy, M. (1995). La mesure de la température de l'air. *La Météorologie*, 8(12), 14-30. <https://doi.org/10.4267/2042/52002>
- Pigeon, G., Legain, D., Durand, P., & Masson, V. (2007). Anthropogenic heat release in an old European agglomeration (Toulouse, France). *International Journal of Climatology : A Journal of the Royal Meteorological Society*, 27(14), 1969-1981. <https://doi.org/10.1002/joc.1530>
- Pinty, J.-P., & Jabouille, P. (1998). A mixed-phase cloud parameterization for use in mesoscale non-hydrostatic model : simulations of a squall line and of orographic precipitations. *Conf. on cloud physics*, 217-220. http://195.83.22.22/mesonh/dir_publication/pinty_jabouille_ams_ccp1998.pdf (dernier accès : 1 mai 2024).
- Potgieter, J., Nazarian, N., Lipson, M. J., Hart, M. A., Ulpiani, G., Morrison, W., & Benjamin, K. (2021). Combining High-Resolution Land Use Data With Crowdsourced Air Temperature to Investigate Intra-Urban Microclimate. *Frontiers in Environmental Science*, 9. <https://doi.org/10.3389/fenvs.2021.720323>
- Pourret, V., Mahfouf, J.-F., Guidard, V., Moll, P., Doerenbecher, A., & Piguet, B. (2021). Variational bias correction for Mode-S aircraft derived winds. *Tellus A : Dynamic Meteorology and Oceanography*, 73(1), 1886808. <https://doi.org/10.1080/16000870.2021.1886808>
- Pourret, V., Šavli, M., Mahfouf, J.-F., Raspaud, D., Doerenbecher, A., Bénichou, H., & Payan, C. (2022). Operational assimilation of Aeolus winds in the Météo-France global NWP model ARPEGE. *Quarterly Journal of the Royal Meteorological Society*, 148(747), 2652-2671. <https://doi.org/10.1002/qj.4329>
- Préaux, D., Dombrowski-Etchevers, I., Gouttevin, I., & Seity, Y. (2023). Study of the 2 m temperature bias of the numerical weather forecasting model Arome over the French Alps. *EGU General Assembly Conference Abstracts*, EGU-7481. <https://doi.org/10.5194/egusphere-egu23-7481>
- Ricard, D. (2002). *Initialisation et assimilation de données à méso-échelle pour la prévision à haute résolution des pluies intenses de la région Cévennes-Vivarais* [thèse de doct., Toulouse 3].
- Ridal, M., Bojarova, J., Hagelin, S., Medeiros, P., Gregow, E., & Stappers, R. (2019). Study of forecast quality improvement (magnitude and duration), relative to using a reference set of observations, from assimilation of novel observations. https://www.ecmwf.int/sites/default/files/special_projects/2020/spnolabo-2020-finalreport.pdf (last access : 15 June 2023).
- Sassi, M. Z., Fourrié, N., Guidard, V., & Birman, C. (2019). Use of Infrared Satellite Observations for the Surface Temperature Retrieval over Land in a NWP Context. *Remote Sensing*, 11(20), 2371. <https://doi.org/10.3390/rs11202371>

- Scheffknecht, P., Meier, F., & Weidle, F. (2021). Exploring New Observations for Austria's AROME RUC Nowcasting System. https://www.umr-cnrm.fr/accord/IMG/pdf/zamg_assim_2021.pdf
- Schraff, C., Reich, H., Rhodin, A., Schomburg, A., Stephan, K., Periañez, A., & Potthast, R. (2016). Kilometre-scale ensemble data assimilation for the COSMO model (KENDA). *Quarterly Journal of the Royal Meteorological Society*, 142(696), 1453-1472. <https://doi.org/10.1002/qj.2748>
- Seity, Y., Brousseau, P., Malardel, S., Hello, G., Bénard, P., Bouttier, F., Lac, C., & Masson, V. (2011). The AROME-France Convective-Scale Operational Model. *Monthly Weather Review*, 139(3), 976-991. <https://doi.org/10.1175/2010MWR3425.1>
- Sgoff, C., Acevedo, W., Paschalidi, Z., Ulbrich, S., Bauernschubert, E., Kratzsch, T., & Potthast, R. (2022). Assimilation of crowd-sourced surface observations over Germany in a regional weather prediction system. *Quarterly Journal of the Royal Meteorological Society*, 148(745), 1752-1767. <https://doi.org/10.1002/qj.4276>
- Talagrand, O., & Courtier, P. (1987). Variational assimilation of meteorological observations with the adjoint vorticity equation. I : Theory. *Quarterly Journal of the Royal Meteorological Society*, 113(478), 1311-1328. <https://doi.org/10.1002/qj.49711347812>
- Tang, Y., Lean, H. W., & Bornemann, J. (2013). The benefits of the Met Office variable resolution NWP model for forecasting convection. *Meteorological Applications*, 20(4), 417-426. <https://doi.org/10.1002/met.1300>
- Tardieu, J., & Leroy, M. (2003). Radome, le réseau temps réel d'observation au sol de Météo-France. *La Météorologie*, (40), 40-43. <https://doi.org/10.4267/2042/36262>
- Tavolato, C., & Isaksen, L. (2015). On the use of a Huber norm for observation quality control in the ECMWF 4D-Var. *Quarterly Journal of the Royal Meteorological Society*, 141(690), 1514-1527. <https://doi.org/10.1002/qj.2440>
- Valmassoi, A., Keller, J. D., Kleist, D. T., English, S., Ahrens, B., Āurán, I. B., Bauernschubert, E., Bosilovich, M. G., Fujiwara, M., Hersbach, H., et al. (2023). Current challenges and future directions in data assimilation and reanalysis. *Bulletin of the American Meteorological Society*, 104(4), E756-E767. <https://doi.org/10.1175/BAMS-D-21-0331.1>
- Váňa, F., Bénard, P., Geleyn, J.-F., Simon, A., & Seity, Y. (2008). Semi-Lagrangian advection scheme with controlled damping : An alternative to nonlinear horizontal diffusion in a numerical weather prediction model. *Quarterly Journal of the Royal Meteorological Society : A journal of the atmospheric sciences, applied meteorology and physical oceanography*, 134(631), 523-537. <https://doi.org/10.1002/qj.220>
- Varentsov, M. I., Konstantinov, P. I., Shartova, N. V., Samsonov, T. E., Kargashin, P. E., Varentsov, A. I., Fenner, D., & Meier, F. (2020). Urban heat island of the Moscow megacity : the long-term trends and new approaches for monitoring and research based on crowdsourcing data. *IOP Conference Series : Earth and Environmental Science*, 606(1), 012063. <https://doi.org/10.1088/1755-1315/606/1/012063>
- Vasiljevic, D., Cardinali, C., & Undén, P. (1992). ECMWF 3D variational data assimilation of conventional observations. *Workshop on Variational Assimilation, with special emphasis on Three-dimensional Aspects, 9-12 November 1992*, 389-436. <https://www.ecmwf.int/en/elibrary/76806-ecmwf-3d-variational-data-assimilation-conventional-observations> (dernier accès : 1 mai 2024).

- Venter, Z. S., Brousse, O., Esau, I., & Meier, F. (2020). Hyperlocal mapping of urban air temperature using remote sensing and crowdsourced weather data. *Remote Sensing of Environment*, 242, 111791. <https://doi.org/10.1016/j.rse.2020.111791>
- Waller, J. A., Bauernschubert, E., Dance, S. L., Nichols, N. K., Pothast, R., & Simonin, D. (2019). Observation error statistics for Doppler radar radial wind superobservations assimilated into the DWD COSMO-KENDA system. *Monthly Weather Review*, 147(9), 3351-3364. <https://doi.org/10.1175/MWR-D-19-0104.1>
- Wattrelot, E. (2016). Radar assimilation at higher density in the operational AROME model at 1.3 km horizontal resolution. *Combined Newsletter of the HIRLAM and ALADIN Consortia*, (6), 53-63. https://www.researchgate.net/profile/Eric-Wattrelot-2/publication/305758263_Radar_assimilation_at_higher_density_in_the_operation_AROME_model_at_13_km_horizontal_resolution/links/579f5cac08ae80bf6ea7b375/Radar-assimilation-at-higher-density-in-the-operation-AROME-model-at-13-km-horizontal-resolution.pdf (dernier accès : 1 mai 2024).
- WMO. (2021). Manual on the WMO Integrated Global Observing System : Annex VIII to the WMO Technical Regulations (WMO-No. 1160). <https://community.wmo.int/en/activity-areas/wigos/gbon> (dernier accès : 1 mai 2024).
- Zumwald, M., Knüsel, B., Bresch, D. N., & Knutti, R. (2021). Mapping urban temperature using crowd-sensing data and machine learning. *Urban Climate*, 35, 100739. <https://doi.org/10.1016/j.uclim.2020.100739>

Titre : Utilisation des données de capteurs météorologiques connectés pour la prévision numérique à courte échéance de phénomènes de méso-échelle

Mots clés : Stations météorologiques personnels, Contrôle qualité, Assimilation, Convection profonde, Internet des Objets

Résumé : L'assimilation de données à méso-échelle permet d'initialiser les modèles régionaux de prévision numérique du temps, en combinant une ébauche du modèle avec les observations disponibles notamment à cette échelle. L'émergence de stations météorologiques connectées fournit une source d'observations d'opportunité permettant de densifier le réseau d'observation de surface. Ces nouvelles stations permettent d'observer à fine échelle la convection profonde près de la surface (Mandement, 2020). Cette thèse propose ainsi d'explorer la capacité de ces observations d'opportunité de température, d'humidité relative et de pression à améliorer la prévision de phénomènes météorologiques de méso-échelle, en les assimilant dans le modèle régional AROME-France. L'étape de pré-traitement de ces observations est nécessaire pour corriger et filtrer les observations considérées anormales vis-à-vis de l'ébauche du modèle, et d'écrémer les observations spatialement proches. Ce pré-traitement permet d'apporter à l'assimilation près de 129 fois plus d'observations au sol de pression, 3 fois plus d'observations de température et 6 fois plus d'observations d'humidité relative que dans le cas où seules les observations standards de surface sont utilisées. Des expériences avec et sans les observations d'opportunité ont été menées (méthode OSE) sur une période d'un mois. Ces OSE ont d'abord été réalisées en assimilant de manière indépendante les trois types d'observations d'opportunité susmentionnés, puis de manière conjointe. Les résultats de l'assimilation des observations d'opportunité de pression montrent une diminution de l'écart quadratique moyen entre la prévision à 1 h d'échéance et les observations standards de pression de 10.3 % en France, avec le schéma 3DEnVar. Il en est de même pour l'assimilation des observations d'opportunité de température et d'humidité relative, qui permettent une diminution de 0.9 % et de 1 % respectivement par rapport aux observations standards de la même grandeur. Avec le réglage actuel du schéma 3DEnVar, il apparaît que les observations de surface disposent d'un poids moindre à l'analyse par rapport aux autres observations et par rapport à l'ébauche. L'assimilation de nombreuses observations d'opportunité de surface contribue à compenser cet effet. Le type de schéma d'assimilation et son réglage jouent un rôle important dans l'assimilation de ces nouvelles observations. Le schéma d'assimilation 3DVar n'est ainsi pas capable de tirer profit des observations d'opportunité, du fait de la création d'incrémentes à la fois isotropes et de grandes dimensions horizontales. Des études de cas ont été menées, dans un premier temps sur des situations comportant de forts gradients de pression de surface — comme dans le cas d'un système convectif de méso-échelle, d'un front, d'une dépression de méso-échelle ou d'un mascaret atmosphérique ; puis dans des cas comportant de forts gradients de température ou d'humidité relative près de la surface — dans des cas de brise de mer, de brouillard, d'îlot de chaleur urbain ou de ligne orageuse. Ces études de cas ont permis de mettre en évidence la capacité des observations d'opportunité à améliorer la représentation de ces phénomènes, en contraignant l'analyse près de la surface.

Title: Using data from personal weather stations for short range numerical prediction of mesoscale phenomena

Key words: Personal weather station, Quality control, Assimilation, Deep convection, Internet of Things

Abstract: Mesoscale data assimilation is used to initialise regional numerical weather prediction models, by combining a first guess with the available observations particularly at this scale. The emergence of personal weather stations (PWS) is a source of observations that can be used to densify the surface observation network. These new observations have demonstrated their ability to observe deep convection near the surface at fine scales. This thesis proposes to explore the capacity of these temperature, relative humidity and pressure observations to improve the forecasting of mesoscale meteorological phenomena, by assimilating them in the AROME-France regional model. These observations need to be pre-processed in order to correct and filter out observations that are considered anomalous in relation to the first model guess, and to be thinned out to remove close observations in space. This pre-processing makes it possible to assimilate almost 129 times more pressure observations, 3 times more temperature observations and 6 times more relative humidity observations than in the case where only standard surface observations are used. Experiments with and without the PWS observations were carried out (OSE method) over a period of one month. These OSEs were first carried out by assimilating the three types of PWS observations mentioned above independently, and then jointly. The results of the assimilation of PWS pressure observations show a reduction in the mean square deviation between the 1 h forecast and the standard pressure observations of 10.3 % in France, using the 3DEnVar scheme. The same applies to the assimilation of PWS temperature and relative humidity observations, which result in a reduction of 0.9 % and 1 % respectively compared to standard observations of the same parameter. With the current setting of the 3DEnVar scheme, it appears that surface observations are given less weight in the analysis than both other observations and the first guess. The assimilation of many PWS surface observations helps to compensate for this effect. The type of assimilation scheme and its adjustment play an important role in the assimilation of these new observations. The use of the 3DVar assimilation scheme is thus unable to take advantage of dense PWS observations, due to the creation of increments that are both isotropic and of large horizontal dimensions. Case studies were carried out, firstly on situations involving strong gradients in surface pressure - as in the case of a mesoscale convective system, a front, a mesoscale low pressure system or an atmospheric bore; then on cases involving strong gradients in surface temperature or relative humidity - in the case of a sea breeze, a fog, an urban heat island or a squall line. These case studies highlighted the ability of PWS observations to improve the representation of these phenomena, by constraining the surface analysis.