



HAL
open science

Amélioration de la connaissance et de la prévision des vents de vallée en conditions stables : expérimentation et modélisation statistique avec réseau de neurones artificiels

Florian Dupuy

► To cite this version:

Florian Dupuy. Amélioration de la connaissance et de la prévision des vents de vallée en conditions stables : expérimentation et modélisation statistique avec réseau de neurones artificiels. Océan, Atmosphère. Université Toulouse 3 – Paul Sabatier, 2018. Français. NNT : . tel-02293020

HAL Id: tel-02293020

<https://theses.hal.science/tel-02293020v1>

Submitted on 20 Sep 2019

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.



THÈSE

En vue de l'obtention du DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par l'Université Toulouse 3 - Paul Sabatier

Présentée et soutenue par

Florian DUPUY

Le 27 novembre 2018

Amélioration de la connaissance et de la prévision des vents de vallée en conditions stables : expérimentation et modélisation statistique avec réseau de neurones artificiels

Ecole doctorale : **SDU2E - Sciences de l'Univers, de l'Environnement et de
l'Espace**

Spécialité : **Océan, Atmosphère, Climat**

Unité de recherche :

LA - Laboratoire d'Aérodynamique

Thèse dirigée par

Pierre DURAND et Thierry HEDDE

Jury

Mme Chantal STAQUET, Rapporteur

Mme Sophie BASTIN, Rapporteur

M. Joan CUXART RODAMILANS, Rapporteur

M. Eric PARDYJAK, Examineur

M. Frank ROUX, Examineur

M. Pierre DURAND, Directeur de thèse

M. Thierry HEDDE, Co-directeur de thèse

M. Damien PIGA, Examineur (invité)

M. Pierre ROUBIN, Co-directeur de thèse (invité)

Remerciements

Je souhaite tout d'abord remercier mes encadrants Pierre Durand, Thierry Hedde et Pierre Roubin pour m'avoir accepté en thèse et pour le super encadrement tout au long de ces trois années. Merci également pour vos multiples conseils et relectures du manuscrit de thèse. Je remercie également Gert-Jan Duine puisque c'est avec lui que j'ai commencé à échanger sur ce qui allait devenir mon travail de thèse. Le travail qu'il a produit pendant sa thèse m'a également facilité la vie dans la prise en main du sujet.

Je remercie également les membres de mon jury en commençant par Sophie Bastin, Joan Cuxart et Chantal Staquet qui ont accepté d'évaluer mon travail en tant que rapporteur. Merci à Frank Roux d'avoir accepté de présider la soutenance, à Eric Pardyjak d'avoir fait un très long voyage pour assister à la soutenance ainsi qu'à Damien Piga.

Je tiens également à remercier les membres des deux comités de suivi de thèse lors desquels ont émergé de nombreuses idées et conseils. Merci à Alexandre Paci, Eric Pardyjak, Damien Piga et Jonathan Virga.

Un grand merci également à Eric pour m'avoir accueilli à Salt Lake City. C'était très enrichissant de découvrir un nouveau laboratoire et un nouveau pays. And thank you Nipun for your help during my stay in Utah and for making me discover the american way of life.

Je remercie finalement mes anciens collègues en commençant par le LMTE pour la super ambiance pendant ces trois années. Je dois également remercier Julien pour ses très nombreux conseils et son aide avec mes problèmes de codage. T'as gagné, j'ai fini par me mettre au python! Je remercie également le groupe des contractuels Toulousains avec qui j'ai partagé les trois derniers mois de thèse au LA.

Résumé

L'objet de cette thèse est l'étude des écoulements en conditions stables dans une vallée dans le but de participer à l'amélioration des calculs de transport et dispersion atmosphérique dans ces conditions. Les stratifications stables sont en effet critiques car elles limitent les mouvements verticaux, ce qui réduit le mélange et la dispersion des composés émis près de la surface. De plus, dans ces conditions, les écoulements créés sont fortement dépendants du relief local, et donc difficiles à calculer précisément avec des modèles de prévision opérationnels, car la résolution requise devient prohibitive. Le site étudié est la vallée de Cadarache, située dans le sud-est de la France, et dans laquelle est installé le centre CEA de Cadarache. Cette vallée de largeur 1 km, avec des versants de 100 m de hauteur, est le siège de courants de densité dès que les conditions stables s'établissent. Ces courants ont des interactions complexes avec les écoulements d'échelle supérieure.

Le premier objectif a été d'améliorer les connaissances sur la stratification et les écoulements à l'échelle de la vallée grâce aux observations de la campagne KASCADE 2017. Une profondeur adimensionnelle de la vallée permet de distinguer les situations de couplage ou découplage entre les vents dans et au-dessus de la vallée. Ces situations ont été reliées à l'intensité de l'hétérogénéité des températures et des vents. En situation de découplage, la stratification dans la vallée est de type « cold pool marginale ». Les vents de pente sur les versants et dans le fond de la vallée, associés aux cold pools marginales, ont été analysés.

Le second objectif a été d'améliorer la détermination des vents près de la surface dans la vallée, à partir d'informations d'échelle supérieure. Une prévision immédiate statistique de ce vent, à partir d'observations du vent au-dessus de la vallée et de la stratification dans cette dernière, a été réalisée à l'aide d'un réseau de neurones artificiels (RNA). La méthode s'est révélée performante et le RNA a donc également été appliqué aux prévisions opérationnelles réalisées avec une résolution de 3 km sur la région avec le modèle météorologique WRF. Cette mise en œuvre consiste donc à effectuer une descente d'échelle statistique à partir de simulations dans lesquelles ni le relief ni l'occupation des sols ne sont correctement représentés à l'échelle de la vallée. Les résultats produits permettent d'améliorer significativement la prévision du vent près de la surface dans la vallée.

Abstract

The aim of this thesis is to study the valley flows under stable stratification to participate to the atmospheric dispersion calculations improvement under such conditions. These conditions are studied because vertical motions are reduced under stable conditions, which diminishes mixing and dispersion of compounds emitted close to the surface. Furthermore, the winds are strongly dependent on topography, which makes their calculation very hard with mesoscale NWP models. The region studied is the Cadarache valley, in south-east of France, which hosts the Cadarache CEA center.

The first objective was to improve our knowledge on stratification conditions and local winds at the valley scale thanks to the KASCADE 2017 experiment. A non-dimensional valley depth allowed us to determine whether above and inside valley winds are coupled or not. Each of the two situations has been related to the temperature and wind heterogeneity in the valley. When winds are not coupled, the stratification in the valley behaves like a marginal cold pool. Slope and valley winds associated to this stratification regime have been analysed.

The second objective was to improve the wind definition, inside the valley and close to the surface, from a larger scale information. A statistical nowcasting of this wind was performed, based on above valley wind observations and the knowledge of the stratification at the valley scale, with an artificial neural network (ANN). Since the method has proved to be efficient, it was thus applied to operational forecasts done on the area with the WRF meteorological model at a 3-km horizontal resolution. This constitutes a statistical downscaling process, applied to simulations where neither the topography nor the land use are adequately resolved at the local scale of the valley. This allows to improve significantly the valley wind forecast.

Table des matières

Résumé	iv
Abstract	vi
Introduction	1
1 La couche limite atmosphérique stable sur terrain complexe – théorie et modélisation	5
1.1 Concepts physiques de base définissant la CLA	6
1.1.1 Équations de base	6
1.1.2 Caractérisation de la turbulence	9
1.1.3 Les flux d'énergie à la surface	11
1.1.3.a Flux radiatifs	11
1.1.3.b Flux turbulents	11
1.1.4 La structure verticale de la CLA	12
1.2 La CLA sur terrain complexe	14
1.2.1 Écoulements sur les zones au relief complexe	14
1.2.1.a Écoulements forcés par le relief	14
1.2.1.b Écoulements forcés par des effets thermiques	15
1.2.1.c Origine des vents dans une vallée	18
1.2.1.d Distinction du couplage/découplage des vents	19
1.2.2 Stratification dans une vallée	20
1.2.3 Sources d'hétérogénéités liées aux terrains complexes	22
1.2.3.a Modulation du rayonnement net par le relief	22
1.2.3.b Modulation des flux turbulents par les hétérogénéités de la surface	23
1.3 Modélisation numérique de la CLA stable sur terrain complexe	23
1.3.1 Représentation de la turbulence par les méthodes RANS et LES	24
1.3.2 Domaines d'application des méthodes RANS et LES	25
1.3.3 Simulation de cas réels de CLA stable sur terrain complexe	26
1.3.3.a Représentation des hétérogénéités de la surface	26
1.3.3.b Des domaines RANS vers les domaines LES	27
1.3.3.c Autres limitations liées aux cas réels	27
1.3.3.d Exemples d'applications	28
1.4 Modélisation statistique	29
1.4.1 Descente d'échelle statistique	29
1.4.2 Réseaux de neurones artificiels	29

2	Recherches préliminaires et passées sur le site d'étude	31
2.1	Description du site	32
2.1.1	Terrain complexe	32
2.1.2	Météorologie régionale	33
2.2	Connaissances sur la météorologie locale du site	35
2.2.1	Observations continues	35
2.2.2	KASCADE 2013	36
2.2.3	Climatologie des vents sur une période d'un an	38
2.3	Prévisions WRF	40
2.3.1	Configuration des simulations Mistrau	40
2.3.2	Représentation du terrain dans les simulations	42
2.3.3	Critères statistiques pour l'évaluation de la prévision des vents	43
2.3.4	Comparaison aux observations	44
2.3.4.a	Calcul des vents WRF à 2 m	44
2.3.4.b	Vents à 2 m	47
2.3.4.c	Vents à 110 m	48
2.3.4.d	Relation entre les vents à différentes hauteurs	49
2.3.4.e	Impact du forçage par les vents d'altitude	51
2.3.4.f	Stratification entre 110 m et 2 m	52
2.3.5	Simulations à 1 km de résolution horizontale	54
2.4	Prévision immédiate du vent local par méthode statistique	54
2.5	Conclusion	55
3	La campagne de mesures KASCADE 2017	57
3.1	Présentation de l'expérience	58
3.1.1	La stratégie expérimentale	58
3.1.2	Les observations continues	59
3.1.3	Les périodes d'observations intensives (POI)	62
3.2	Aperçu des conditions pendant l'expérience	64
3.2.1	Conditions météorologiques sur la période décembre 2016 – mars 2017	64
3.2.2	Oscillations	68
3.2.2.a	Précédentes observations et analyses d'oscillations pour un vent de vallée	69
3.2.2.b	Oscillations observées pendant KASCADE 2017	71
3.3	La CLA stable dans la vallée de Cadarache	79
3.3.1	Mesures de la stratification dans la vallée	79
3.3.2	Extension verticale de la CLA stable	79
3.3.3	Stratification dans la vallée de Cadarache	81
3.3.3.a	Stratification et hétérogénéité	81
3.3.3.b	Caractéristiques de la couche stratifiée dans la vallée	87
3.3.3.c	Stratification dans le fond de la vallée	91
3.4	Écoulements dans la vallée de Cadarache	96
3.4.1	Couplage/découplage des écoulements dans et au-dessus de la vallée	96
3.4.2	Hétérogénéité des directions des vents	98
3.4.2.a	Climatologie des vents	98
3.4.3	Évolution temporelle du vent de vallée de Cadarache	104
3.5	Conclusion	107

4	Prévision immédiate statistique du vent dans la vallée de Cadarache	109
4.1	Fonctionnement des RNA	110
4.2	Description de la structure des réseaux utilisés	111
4.3	Conclusion	112
4.4	Article	113
5	Descente d'échelle statistique de simulations méso-échelle pour la prévision d'un vent local	131
5.1	Construction du réseau de neurones artificiels	132
5.1.1	La stratégie	132
5.1.2	Sélection des variables d'entrée	132
5.2	Performance de prévision du vent avec le RNA	135
5.2.1	Comparaison aux résultats du RNA–Observations	135
5.2.2	Apport du RNA par rapport aux simulations WRF	135
5.2.3	Analyse des erreurs	140
5.2.3.a	Erreur de forçage au-dessus de la vallée	141
5.2.3.b	Erreur de prévision de la différence de température potentielle	142
5.3	Conclusion	145
	Conclusion et perspectives	147
	Bibliographie	153
A	Étude de sensibilité sur le calcul de la stratification	167
A.1	Théorie	167
A.1.1	Calcul théorique de la stratification	168
A.1.2	Exemples	169
A.2	Calculs à partir du ballon captif (KASCADE 2013)	171
B	Inter-comparaison	173
B.1	Température à 2 m	173
B.2	Humidité relative à 2 m	175
B.3	Pression à 2 m	178
B.4	Rayonnement SW descendant à 2 m	181
B.5	Module du vent à 2 m	183
B.6	Direction du vent à 2 m	185
B.7	Température de surface	186
B.8	Ballon captif et radiosondages	189
C	Composition des jeux de données d'entrée pour les RNA	193
C.1	Sélection des variables	193
C.1.1	Observations sur la période de un an	193
C.1.1.a	Première étape – relations inter-variables	193
C.1.1.b	Seconde étape – choix des variables les plus pertinentes	194
C.1.2	Données WRF	197
C.1.2.a	Première étape – relations inter-variables	197
C.1.2.b	Seconde étape – choix des variables les plus pertinentes	202
C.2	Longueur du jeu de données	205

Introduction

La couche limite atmosphérique (en abrégé CLA) est la partie de l'atmosphère qui se situe le plus près du sol. Nous respirons donc l'air de la CLA, d'où l'importance de l'étude de sa composition étant donné que certaines espèces chimiques présentes peuvent être problématiques d'un point de vue sanitaire. À la surface, les sources de composés, qu'ils soient naturels ou anthropiques, sont nombreuses et variées comme par exemple les pollens, l'érosion des sols ou les sels marins pour les particules fines (ou PM10 pour celles dont le diamètre est inférieur à 10 μm) d'origine naturelle, ou encore la combustion pour les particules fines ou les oxydes d'azote (NO_x) d'origine anthropique.

Les composés émis à la surface se retrouvent ensuite dans la CLA où ils sont mélangés avec une efficacité dépendante des conditions thermiques et dynamiques de l'atmosphère. Ainsi, en journée, la convection et la turbulence permettent en général de les mélanger et de les transporter en altitude impliquant une dilution importante. Au contraire, lorsque les conditions sont stables, c'est à dire que la CLA est marquée par une couche d'inversion près du sol et que la turbulence est bien moindre qu'en journée, le mélange est beaucoup moins efficace et les composés émis se retrouvent piégés près de la surface. Ils peuvent alors atteindre de hautes concentrations faisant des conditions stables les conditions les plus problématiques du point de vue sanitaire. Les épisodes anticycloniques hivernaux conduisent d'ailleurs fréquemment au dépassement des seuils réglementaires (fixés par les articles [R221-1](#) à [R221-3](#) du code de l'environnement en ce qui concerne la France).

Dans les régions au relief important, avec alternance de vallées et montagnes, la stabilité rencontrée pendant les conditions anticycloniques hivernales peut persister plusieurs jours et les concentrations en polluants augmenter de manière importante dans les vallées. La vallée de l'Arve et la ville de Grenoble dans les Alpes françaises sont des exemples fréquemment utilisés pour illustrer les problèmes de qualité de l'air rencontrés dans ces conditions ([Paci et al., 2017](#); [Largeron and Staquet, 2016](#)).

Cette thèse se place dans le contexte du centre du commissariat à l'énergie atomique et aux énergies alternatives (CEA) de Cadarache, situé dans les Bouches-du-Rhône dans le sud-est de la France. Bien que l'intérêt pour le CEA ne porte pas exactement sur les mêmes composés que ceux cités précédemment, les préoccupations relatives à la qualité de l'air sont similaires puisque la région est marquée par un relief complexe et une fréquence importante de conditions stables. Ainsi, une des obligations du CEA est le calcul de l'impact des substances chimiques ou radiologiques émises sur le centre de Cadarache, d'où le besoin d'évaluer leur transport dans les différents milieux naturels (atmosphère, sol, sous-sol, rivières). Ces différents calculs sont effectués par le Laboratoire de Modélisation des Transferts dans l'Environnement (LMTE) du CEA Cadarache.

La représentation des écoulements dans les basses couches de l'atmosphère sur terrain complexe peut être particulièrement difficile, notamment par conditions stables. Ces conditions

gènèrent fréquemment un système de circulation tridimensionnel dans les vallées avec des vents de pentes drainant l'air froid produit sur les versants et alimentant un vent catabatique en fond de vallée. Il est alors important de bien caractériser la circulation en place afin de valider les simulations de ces écoulements. Pour cela, plusieurs campagnes de terrain ont été réalisées, parmi lesquelles on peut mentionner au niveau international : COLPEX ([Price et al., 2011](#)), MAP ([Rotach et al., 2004](#)), VTMX ([Doran et al., 2002](#)) ou MATHERHORN ([Fernando et al., 2015](#)). Sur le site de Cadarache, la campagne KASCADE 2013 (KAtabatic winds and Stability over CADarache for Dispersion of Effluents) s'est déroulée pendant l'hiver 2013 et a permis de documenter la structure verticale de la stratification et des écoulements ([Duine et al., 2017](#)).

Concernant le transport dans l'atmosphère, les calculs effectués au LMTE sont actuellement réalisés à partir de modèles de dispersion atmosphérique de type gaussien. Ces modèles font l'approximation que les panaches sont transportés par le vent moyen et diffusés par la turbulence. Cette dernière étant considérée homogène et isotrope, la concentration de l'espèce émise suit une loi gaussienne dans le plan perpendiculaire à la direction du vent moyen. Les principaux avantages de ces modèles sont leur faible temps de calcul et leur simplicité de mise en œuvre ce qui en fait des outils particulièrement bien adaptés aux situations d'urgence en terrain homogène. Ils possèdent cependant quelques limites, dont la description des conditions météorologiques qui est très simplifiée. En effet, le champ de vent y est souvent supposé uniforme et la stabilité atmosphérique, déterminant l'efficacité du mélange turbulent, y est représentée de façon très simplifiée, par exemple avec des classes discrètes. De plus, ces modèles ne sont généralement valides que dans un rayon limité à quelques kilomètres autour de la source. Enfin, les effets du relief et des bâtiments, qui peuvent être primordiaux, ne sont pas représentés. Les modèles gaussiens de dernière génération, sont eux capables de décrire la stabilité de manière continue ou encore disposent d'une modélisation simplifiée des effets du relief qui restitue un champ de vent non homogène.

Les modèles de type lagrangien, qui discrétisent les substances émises par des particules de type Monte-Carlo, restent tout de même mieux adaptés pour les situations les plus complexes, telles que la dispersion atmosphérique par conditions stables sur une région au relief complexe, grâce notamment à l'utilisation de champs de vents quadridimensionnels permettant la prise en compte de l'effet du relief. Ces modèles sont particulièrement adaptés au site de Cadarache du fait de son relief important et de la fréquence des épisodes de stabilité. Pour cette raison, le LMTE envisage l'utilisation d'un modèle de type lagrangien. Le calcul de la dispersion atmosphérique se déroule en deux étapes : un code de prévision calcule les variables météorologiques utilisées comme forçage pour un code de dispersion atmosphérique. Parmi les différents outils disponibles, le code de prévision météorologique WRF (Weather Research and Forecasting) et le code de dispersion de type lagrangien FLEXPART (FLEXible PARTicle dispersion model) ont été choisis. Ces deux codes sont disponibles en open source, sont tous deux largement utilisés dans la communauté scientifique et ont donc été validés sur de nombreuses situations. De plus, une version couplée des deux codes (nommée FLEXPART-WRF) a été développée ce qui présente un avantage supplémentaire justifiant le choix de cette combinaison d'outils.

Actuellement, les simulations WRF de l'outil baptisé « Mistrau » produisent des prévisions à 4 jours sur un domaine possédant une résolution horizontale de 3 km sur la région autour de Cadarache. Malheureusement, cette résolution n'est pas suffisante pour représenter toute la complexité du relief, en particulier la vallée de Cadarache (large d'environ 1 km et longue d'environ 5 km), dans laquelle sont implantées des stations météorologiques du CEA, n'est pas du tout résolue. Il en résulte une mauvaise prévision des vents de surface, notamment par conditions stables pendant lesquelles ils sont fortement dépendants du relief local.

Ce travail fait suite à la thèse de [Duine \(2015\)](#) pendant laquelle la campagne KASCADE 2013 a été réalisée et analysée. G.-J. Duine a également réalisé des simulations à une résolution horizontale de 1 km avec le code WRF, mais cette résolution, si elle permet de bien représenter la vallée de la Durance, n'est pas encore suffisante pour améliorer la représentation du relief et des vents près de la surface dans la vallée de Cadarache. Il a alors également développé une méthode de prévision instantanée visant à diagnostiquer la présence d'un vent d'origine thermique dans la vallée de Cadarache à partir d'une mesure de la différence verticale de température potentielle (entre 110 m et 2 m) réalisée en continu sur le site de Cadarache ([Duine et al., 2016](#)).

L'objectif du travail présenté ici est double : améliorer la connaissance des écoulements dans la vallée de Cadarache afin d'aider à la validation de futures simulations aérologiques à très haute résolution, et améliorer le modèle de prévision de vent utilisé en opérationnel via l'outil Mistrau à l'aide d'une descente d'échelle statistique.

La campagne de terrain KASCADE 2013 a permis de décrire la structure verticale de la stratification et des vents dans le centre de la vallée de Cadarache. Cependant, la connaissance des écoulements dans le reste de la vallée, apportée par les quelques stations météorologiques placées dans la vallée et opérant en continu, restait limitée. Une deuxième campagne d'observation, KASCADE 2017, a donc été lancée afin de compléter les mesures effectuées en 2013. Elle a visé à étudier la structure de la stratification sur les versants de la vallée, à mesurer les hétérogénéités rencontrées dans celle-ci, notamment celle des directions des vents, et à étudier l'évolution du vent dans la vallée sur toute sa longueur, de son apparition à sa disparition.

La connaissance apportée par ces deux campagnes d'observations révèle la complexité des écoulements locaux. Aucun des systèmes opérationnels, qu'il s'agisse des simulations WRF à 3 km de résolution ou de l'observation permanente sur un seul site à deux niveaux – le fond de la vallée et le sommet de ses flancs – n'est à même de décrire et/ou prévoir les vents de vallée et leur cycle diurne. Il faut donc mettre en œuvre d'autres outils pour pallier ce manque. Une première piste est de réaliser des simulations à très haute résolution (≈ 100 m horizontalement) afin que le relief et l'occupation des sols soient décrits de manière réaliste dans le modèle. Cette piste sera explorée, mais à une échéance qui sort du cadre de cette thèse, car c'est une étude complexe, et qui présentera de plus des limitations pratiques de mise en œuvre (temps de calcul) dans un cadre opérationnel. Une autre possibilité est de mettre en œuvre des techniques de descente d'échelle statistique, soit à partir d'observations disponibles en temps réel et représentatives d'une échelle supérieure au phénomène local qu'on veut décrire (nowcasting), soit à partir des prévisions opérationnelles (forecasting). La première voie a été explorée par [Duine et al. \(2016\)](#) qui à partir d'une simple méthode de seuillage sur un paramètre pertinent ont montré qu'on pouvait, avec un bon taux de succès, déterminer la présence ou l'absence d'un courant de densité dans le fond de la vallée. Les limitations de cette méthode (présence ou absence du vent, mais pas de possibilité de le quantifier), nous ont amené à explorer une technique plus élaborée : les réseaux de neurones artificiels (RNA). Cette méthode de modélisation, qui nécessite une phase d'apprentissage par confrontation de la sortie du RNA avec une réalité (dans notre cas des observations du vent de vallée), a été mise en œuvre à la fois sur les observations opérationnelles, et sur les simulations WRF. L'apprentissage a été rendu possible grâce aux campagnes d'observations KASCADE, et grâce à une station météorologique installée pendant une année complète dans le fond de la vallée.

Le premier chapitre introduit les principes physiques théoriques relatifs aux sujets abordés au cours de cette thèse. Il présente la CLA de manière générale, avec une description détaillée

des écoulements sur les régions à la topographie complexe et notamment par conditions stables. Quelques principes relatifs à la modélisation des écoulements sont présentés, notamment concernant la turbulence afin d'énoncer les principales difficultés anticipées pour la modélisation numérique sur les terrains complexes.

Le deuxième chapitre permet de poser les bases de l'étude menée pendant la thèse. La situation géographique du site ainsi qu'une étude de la météorologie régionale, basée en partie sur le travail de thèse de [Duine \(2015\)](#), y sont présentées. Les simulations Mistrau (WRF) sont ensuite introduites puis analysées par comparaison avec les observations réalisées en continu sur le site de Cadarache afin d'en évaluer les capacités et limites.

Le troisième chapitre est consacré à l'étude de la stratification et des écoulements dans la vallée de Cadarache à partir de la campagne d'observations KASCADE 2017. La méthodologie puis les résultats obtenus avec la campagne y sont présentés. De plus, deux phénomènes particuliers observés de manière régulière par conditions stables sont également présentés : des oscillations des vents de vallée et une distribution particulière des températures près de la surface.

Les deux derniers chapitres sont dédiés au calcul du vent dans la vallée de Cadarache de manière statistique à l'aide d'un RNA. Dans le quatrième chapitre, la méthode est évaluée et validée dans un mode de prévision instantanée par comparaison avec la méthode développée par [Duine et al. \(2016\)](#). Elle est ensuite appliquée avec un objectif de descente d'échelle statistique des simulations Mistrau afin d'en améliorer la prévision locale du vent. Cette méthodologie est relativement originale. Ses performances sont évaluées ici sur un site particulier, mais son potentiel d'application n'est évidemment pas limité à cette zone géographique.

Chapitre 1

La couche limite atmosphérique stable sur terrain complexe – théorie et modélisation

Ce chapitre décrit les concepts physiques généraux qui définissent la couche limite atmosphérique (CLA), notamment le cas des CLA stables sur terrain complexe. Une description détaillée des écoulements sur relief complexe est fournie. Enfin, les difficultés rencontrées par les simulations numériques dans ces conditions sont présentées ainsi que certaines solutions statistiques.

Plan du chapitre

1.1	Concepts physiques de base définissant la CLA	6
1.1.1	Équations de base	6
1.1.2	Caractérisation de la turbulence	9
1.1.3	Les flux d'énergie à la surface	11
1.1.4	La structure verticale de la CLA	12
1.2	La CLA sur terrain complexe	14
1.2.1	Écoulements sur les zones au relief complexe	14
1.2.2	Stratification dans une vallée	20
1.2.3	Sources d'hétérogénéités liées aux terrains complexes	22
1.3	Modélisation numérique de la CLA stable sur terrain complexe	23
1.3.1	Représentation de la turbulence par les méthodes RANS et LES	24
1.3.2	Domaines d'application des méthodes RANS et LES	25
1.3.3	Simulation de cas réels de CLA stable sur terrain complexe	26
1.4	Modélisation statistique	29
1.4.1	Descente d'échelle statistique	29
1.4.2	Réseaux de neurones artificiels	29

1.1 Concepts physiques de base définissant la CLA

La CLA est la couche de l’atmosphère la plus proche du sol, dont l’extension verticale typique est de l’ordre du kilomètre en journée et de la centaine de mètre la nuit. La CLA est directement impactée par la surface sur une échelle de temps courte (minutes/heure). Les interactions entre la surface et la CLA se matérialisent par des flux de matière, d’énergie et de quantité de mouvement qui peuvent avoir des origines thermiques, dynamiques, thermodynamiques, physiologiques, chimiques. . . Ces échanges se matérialisent notamment par du transport à caractère turbulent, des échanges radiatifs ou des changements de phase de l’eau.

1.1.1 Équations de base

La description de l’état et de l’évolution de la CLA sont gouvernés par les équations de la mécanique des fluides et de la thermodynamique : une équation d’état et quatre lois de conservation.

L’équation d’état des gaz s’écrit :

$$p = \rho R_d T_v \quad (1.1)$$

où p est la pression, ρ est la masse volumique de l’air humide, R_d est la constante des gaz pour l’air sec ($R_d = 287 \text{ J K}^{-1} \text{ kg}^{-1}$) et T_v est la température virtuelle qui permet de prendre en compte l’effet de l’humidité (l’humidité spécifique q) sur la densité :

$$T_v \approx (1 + 0,61q) T \quad (1.2)$$

La loi de conservation de la masse, aussi appelée équation de continuité, s’écrit (en utilisant la convention sommatoire) :

$$\frac{\partial \rho}{\partial t} + \frac{\partial (\rho u_j)}{\partial x_j} = 0 \quad (1.3)$$

où u_j est la $j^{\text{ème}}$ composante du vent (u , v et w).

L’équation de conservation de la quantité de mouvement (par unité de masse) s’écrit :

$$\underbrace{\frac{\partial u_i}{\partial t}}_{\text{stockage}} + \underbrace{u_j \frac{\partial u_i}{\partial x_j}}_{\text{advection}} = \underbrace{-\delta_{i3} g}_{\text{force de gravité}} + \underbrace{f_c \epsilon_{ij3} u_j}_{\text{Coriolis}} - \underbrace{\frac{1}{\rho} \frac{\partial p}{\partial x_i}}_{\text{force de pression}} + \underbrace{\nu \frac{\partial^2 u_i}{\partial x_j^2}}_{\text{contraintes visqueuses}} \quad (1.4)$$

avec g la gravité, δ_{ij} le symbole de Kronecker, f_c le paramètre de Coriolis ($f_c = 2\Omega \sin \phi$ où Ω est la vitesse de rotation de la Terre et ϕ est la latitude), ϵ_{ijk} le symbole de Levi-Civita et ν la viscosité cinématique moléculaire.

L’équation de conservation de l’humidité (pour la partie gazeuse) s’écrit :

$$\underbrace{\frac{\partial q}{\partial t}}_{\text{stockage}} + \underbrace{u_j \frac{\partial q}{\partial x_j}}_{\text{advection}} = \underbrace{\frac{S_q}{\rho}}_{\text{source}} + \underbrace{\nu_q \frac{\partial^2 q}{\partial x_j^2}}_{\text{contraintes visqueuses}} \quad (1.5)$$

où ν_q est la diffusivité moléculaire pour la vapeur d’eau dans l’air et S_q représente une source ou un puits d’humidité lié à un changement de phase.

Il est nécessaire d'introduire la notion de température potentielle (θ) avant de formuler l'équation de conservation de la chaleur. La température absolue de l'air (T) varie en fonction de la pression, ce qui engendre de forts gradients verticaux. La valeur théorique de ce gradient pour une atmosphère sèche en stratification neutre et dans les basses couches vaut $-9,8\text{ °C km}^{-1}$ alors que les valeurs pour une atmosphère standard (par exemple en troposphère libre) sont proches de -6 °C km^{-1} . Il est donc courant d'avoir recours au calcul d'une température potentielle pour comparer les températures de deux masses d'air ayant des pressions différentes. La température potentielle représente la température absolue qu'aurait une masse d'air (considéré comme un gaz parfait) amenée adiabatiquement (c'est à dire en suivant une transformation isentrope) à la pression $p_0 = 1000\text{ hPa}$:

$$\theta = T \left(\frac{p_0}{p} \right)^{\frac{R}{C_p}} \quad (1.6)$$

avec C_p la capacité calorifique de l'air humide et R la constante des gaz pour l'air humide. Une autre variable couramment utilisée pour comparer deux masses d'air aux contenus en eau différents est la température potentielle virtuelle, notée θ_v . Cette variable permet de prendre en compte les différences de densité liées au contenu en eau de l'air. Elle représente la température potentielle qu'aurait une particule d'air sec de même masse volumique et à la même pression :

$$\theta_v = \theta (1 + 0.61q) \quad (1.7)$$

L'équation de conservation de la chaleur s'écrit donc :

$$\underbrace{\frac{\partial \theta}{\partial t}}_{\text{stockage}} + \underbrace{u_j \frac{\partial \theta}{\partial x_j}}_{\text{advection}} = \underbrace{\nu_\theta \frac{\partial^2 \theta}{\partial x_j^2}}_{\text{contraintes visqueuses}} - \underbrace{\frac{1}{\rho C_p} \frac{\partial R n_j^*}{\partial x_j}}_{\text{source rayonnement}} - \underbrace{\frac{S_\theta}{\rho}}_{\text{source chaleur latente}} \quad (1.8)$$

où ν_θ est la diffusivité thermique moléculaire, $R n_j$ est la composante du rayonnement net dans la $j^{\text{ème}}$ direction et S_θ est un terme de source ou puits représentant les échanges d'énergie issus d'un changement de phase de l'eau.

Représentation de la turbulence dans les équations de Navier-Stokes

Dans un écoulement turbulent, les paramètres atmosphériques sont caractérisés par des fluctuations rapides et de petite échelle spatiale (typiquement de l'ordre du millimètre à quelques centaines de mètres). Il existe en permanence un ensemble continu d'échelles où l'énergie du mouvement est transmise de proche en proche des grandes échelles vers les plus petites. L'énergie cinétique du mouvement turbulent est finalement, à l'issue de cette cascade énergétique, transformée en énergie interne (chaleur) par dissipation aux échelles des plus petits tourbillons (Kolmogorov, 1941).

Une description complète de la turbulence jusqu'aux petites échelles est inaccessible à la simulation aux grands nombres de Reynolds caractérisant la CLA (cf. partie 1.1.2). Une description simplifiée consiste à décomposer les échelles du mouvement en une composante grande échelle (le mouvement moyen) et une composante petite échelle (les fluctuations autour du mouvement moyen : la turbulence) : c'est la décomposition de Reynolds qui pour une variable α s'écrit :

$$\alpha = \bar{\alpha} + \alpha' \quad (1.9)$$

où $\bar{\alpha}$ représente la composante grande échelle (moyenne d'ensemble) et α' représente la composante petite échelle.

L'application de la moyenne de Reynolds aux équations de Navier-Stokes permet d'écrire les variables sous la forme de moments statistiques à une échelle où ils peuvent être représentés dans les modèles. De plus, les termes de diffusion moléculaire sont négligeables devant les autres termes, notamment devant ceux d'advection et de transport turbulent. Le système d'équations s'écrit alors :

$$\bar{p} = \bar{\rho} R_d \bar{T}_v \quad (1.10)$$

$$\frac{\partial \bar{p}}{\partial t} + \frac{\partial \bar{p} \bar{u}_j}{\partial x_j} + \frac{\partial \overline{\rho' u'_j}}{\partial x_j} = 0 \quad (1.11)$$

$$\frac{\partial \bar{u}_i}{\partial t} + \bar{u}_j \frac{\partial \bar{u}_i}{\partial x_j} = -\delta_{i3} g + f_c \epsilon_{ij3} \bar{u}_j - \frac{1}{\rho} \frac{\partial \bar{p}}{\partial x_i} - \frac{\partial (\overline{u'_i u'_j})}{\partial x_j} \quad (1.12)$$

$$\frac{\partial \bar{q}}{\partial t} + \bar{u}_j \frac{\partial \bar{q}}{\partial x_j} = \frac{S_q}{\rho} - \frac{\partial (\overline{u'_i q'})}{\partial x_j} \quad (1.13)$$

$$\frac{\partial \bar{\theta}}{\partial t} + \bar{u}_j \frac{\partial \bar{\theta}}{\partial x_j} = -\frac{1}{\rho C_p} \frac{\partial \overline{R n_j}}{\partial x_j} - \frac{S_\theta}{\rho} - \frac{\partial (\overline{u'_i \theta'})}{\partial x_j} \quad (1.14)$$

Ce système d'équations fait intervenir de nouveaux termes sous la forme de divergence de flux turbulents (les termes qui se trouvent à droite dans les équations 1.11, 1.12, 1.13 et 1.14). Ces termes montrent une relation non-linéaire entre les champs moyens et fluctuants, ce qui signifie que la solution moyenne grande échelle est perturbée par les mouvements turbulents.

Les termes $\overline{u'_i \alpha'}$ sont les flux turbulents qui représentent le transport des fluctuations de α par les fluctuations de vitesse dans la direction i . Le transport turbulent réalise un mélange du fluide et donc favorise son homogénéisation.

Les flux turbulents apportent de nouvelles inconnues dans les équations ce qui pose un problème de fermeture. La fermeture du système d'équations vise à supprimer des inconnues en les exprimant en fonction de variables connues. Une approche courante consiste à admettre que les flux turbulents sont proportionnels au gradient des quantités moyennes (Lilly, 1967) :

$$\overline{u'_i \alpha'} = -K_\alpha \frac{\partial \bar{\alpha}}{\partial x_i} \quad (1.15)$$

avec K_α un coefficient de viscosité turbulente qui est une caractéristique de l'écoulement. Ce choix constitue une fermeture d'ordre 1 parce que les moments d'ordre 2 (les flux turbulents) sont exprimés en fonction des valeurs moyennes. Il existe d'autres méthodes de fermeture d'ordres supérieurs avec des équations pronostiques pour tout ou partie des flux turbulents et des paramétrisations pour les éventuels moments d'ordre supérieur créés.

1.1.2 Caractérisation de la turbulence

Le nombre de Reynolds

Il existe deux régimes d'écoulements possibles permettant de caractériser l'écoulement d'un fluide : laminaire ou turbulent. Un écoulement laminaire est caractérisé par de faibles variations spatiales et temporelles de directions de sorte qu'il paraît organisé. Au contraire, un écoulement turbulent est caractérisé par un comportement chaotique et présente des structures tourbillonnaires d'échelles spatiales et temporelles très variables. Le régime de l'écoulement est déterminé par la compétition entre les forces d'inertie et visqueuses qui s'exprime au travers du nombre de Reynolds :

$$Re = \frac{Uh}{\nu} \quad (1.16)$$

avec U une vitesse caractéristique de l'écoulement et h une longueur caractéristique. Il existe une valeur critique de ce nombre qui permet de distinguer les écoulements laminaires et turbulents. Au-delà d'une valeur de $Re \approx 10^3-10^5$, les écoulements sont systématiquement considérés comme turbulents. Les nombres de Reynolds associés aux écoulements dans la CLA sont largement supérieurs à 10^5 (par exemple, pour $U = 15 \text{ m s}^{-1}$, une hauteur de CLA $h = 1000 \text{ m}$ et avec $\nu \approx 1,45 \times 10^{-5} \text{ m}^2 \text{ s}^{-1}$, $Re \approx 10^9$) ce qui en fait des écoulements turbulents. Cette classification des écoulements par le nombre de Reynolds est purement mécanique, or, dans la CLA, la flottabilité joue également un rôle important (voir les paragraphes suivants).

L'énergie cinétique turbulente

L'énergie cinétique turbulente (ECT) par unité de masse est définie à partir des variances des trois composantes du vent :

$$e = \frac{1}{2} (\overline{u'^2} + \overline{v'^2} + \overline{w'^2}) \quad (1.17)$$

L'ECT est d'autant plus grande que les fluctuations de vitesse le sont ce qui en fait une variable directement représentative de la turbulence. Son évolution s'écrit :

$$\underbrace{\frac{\partial e}{\partial t}}_{\text{stockage}} + \underbrace{\overline{u_j} \frac{\partial e}{\partial x_j}}_{\text{advection}} = \underbrace{-\overline{u'_i u'_j} \frac{\partial \overline{u_i}}{\partial x_j}}_{\text{production mécanique}} + \underbrace{\frac{g}{\theta_v} \overline{w' \theta'_v}}_{\text{flottabilité}} - \underbrace{\frac{\partial \overline{u'_j e}}{\partial x_j}}_{\text{transport turbulent}} - \underbrace{\frac{1}{\rho} \frac{\partial \overline{u'_j p'}}{\partial x_j}}_{\text{redistribution}} - \underbrace{\epsilon}_{\text{dissipation}} \quad (1.18)$$

L'ECT peut être produite par le cisaillement du vent, avec des valeurs plus importantes par vent fort, ou par la flottabilité lorsqu'il y a création de mouvements convectifs. En revanche, la flottabilité peut également détruire l'ECT dans le cas d'une couche stable. Elle est également détruite par dissipation dans les plus petits tourbillons (transformation en chaleur, voir partie 1.3.1).

Turbulence liée à la flottabilité

Une masse d'air peut être caractérisée, entre autres, par sa stabilité. Dans une atmosphère stratifiée en densité, dans le cas d'une augmentation de la température potentielle virtuelle θ_v avec l'altitude, la force de flottabilité aura tendance à contrer chaque petite perturbation de la position verticale qui s'applique à une particule fluide en la ramenant vers son état initial :

c'est un équilibre stable. Au contraire, une petite perturbation appliquée à une particule fluide dans une atmosphère où θ_v diminue avec l'altitude sera amplifiée par la flottabilité : c'est un équilibre instable. Une atmosphère instable est propice à la formation de mouvements turbulents au contraire d'une atmosphère stable. Le gradient vertical de θ_v peut être utilisé pour caractériser la stabilité de l'atmosphère :

- si $\frac{\partial\theta_v}{\partial z} < 0$, la couche d'atmosphère est instable,
- si $\frac{\partial\theta_v}{\partial z} > 0$, la couche d'atmosphère est stable,
- si $\frac{\partial\theta_v}{\partial z} = 0$, la couche d'atmosphère est neutre.

Dans le cas d'un équilibre stable, la fréquence de Brunt-Väisälä caractérise l'oscillation d'une particule fluide autour d'une position d'équilibre, sous l'effet de la gravité et de la poussée d'Archimède. La pulsation correspondante, exprimée en rad s^{-1} , s'écrit :

$$N = \sqrt{\frac{g}{\theta_v} \frac{\partial\theta_v}{\partial z}} \quad (1.19)$$

avec g l'accélération de la gravité et θ_v une température potentielle virtuelle de référence de la couche. Ces oscillations ne sont pas présentes dans un environnement instable parce qu'une particule fluide qui subit une perturbation ne sera pas rappelée vers sa position initiale.

Le nombre de Richardson

Le nombre de Richardson (de flux, noté Ri_f) permet de caractériser différents régimes de turbulence à partir des termes de production thermique (flottabilité) et dynamique (cisaillement) d'ECT (cf. équation 1.18). Il s'écrit :

$$Ri_f = \frac{\frac{g}{\theta_v} \overline{w'\theta'_v}}{\overline{u'w' \frac{\partial u}{\partial z}} + \overline{v'w' \frac{\partial v}{\partial z}}} \quad (1.20)$$

Le terme de production dynamique étant toujours positif, le signe de Ri_f est opposé au signe du terme de production par la flottabilité. Si la flottabilité est positive, les mouvements convectifs créent de la turbulence et le régime est instable : $Ri_f < 0$. Dans une atmosphère neutre, la flottabilité est nulle tout comme Ri_f . Lorsque l'atmosphère est stable, la flottabilité est négative ce qui signifie que le mélange vertical est contré par la stratification. Cela représente un puits de turbulence et Ri_f est positif.

Quand les flux ne sont pas accessibles, Ri peut être calculé en fonction des gradients (nombre de Richardson de gradient noté Ri_g , équation 1.21) ou en approximant les gradients par des différences Δ entre deux hauteurs (nombre global de Richardson ou bulk Richardson number en anglais noté Ri_B , équation 1.22).

$$Ri_g = \frac{\frac{g}{\theta_v} \frac{\partial\theta_v}{\partial z}}{\left(\frac{\partial\bar{u}}{\partial z}\right)^2 + \left(\frac{\partial\bar{v}}{\partial z}\right)^2} \quad (1.21)$$

$$Ri_B = \frac{\frac{g}{\theta_v} \Delta\bar{\theta}_v \Delta z}{\bar{\theta}_v \left((\Delta\bar{u})^2 + (\Delta\bar{v})^2 \right)} \quad (1.22)$$

Le nombre de Richardson permet également de caractériser en conditions stables les transitions entre la génération et l'amortissement de la turbulence et inversement, à partir de valeurs critiques. Il n'y a pas de consensus sur les valeurs précises de ces seuils, mais il est accepté qu'un écoulement stable génère de la turbulence quand $Ri_g < 0,25$ et à l'inverse qu'il l'amortit pour $Ri_g > 1$. En revanche, ces valeurs critiques ne sont adaptées qu'aux gradients locaux, et non aux différences finies comme utilisées dans le calcul du Ri_B . Cependant, plus l'épaisseur de la couche sur laquelle est calculé Ri_B est faible et plus les valeurs critiques se rapprochent des seuils trouvés pour le Ri_g .

1.1.3 Les flux d'énergie à la surface

1.1.3.a Flux radiatifs

Le rayonnement solaire représente une source d'énergie pour l'atmosphère terrestre. En traversant l'atmosphère, ce rayonnement peut être réfléchi (par exemple par les nuages), diffusé ou absorbé. Le rayonnement solaire atteignant la surface terrestre est majoritairement constitué de longueurs d'ondes du spectre visible et il est noté SW_{\downarrow} (pour Short Wave descendant). SW_{\downarrow} peut ensuite être réfléchi (SW_{\uparrow}) ou absorbé par la surface. La part de rayonnement réfléchi dépend de l'albédo qui est une caractéristique de la surface.

Comme n'importe quel corps, la surface terrestre émet un rayonnement dépendant de sa température (LW_{\uparrow}). La gamme de températures de la surface terrestre permet l'émission de rayonnement principalement dans les longueurs d'ondes dites de l'infra-rouge thermique. De même, l'atmosphère peut également être considérée comme un corps qui émet un rayonnement dépendant de sa température (LW_{\downarrow}). Le bilan de ces quatre flux radiatifs est appelée le rayonnement net Rn :

$$Rn = SW_{\downarrow} - SW_{\uparrow} + LW_{\downarrow} - LW_{\uparrow} \quad (1.23)$$

1.1.3.b Flux turbulents

Les mouvements turbulents permettent des échanges au sein de l'atmosphère ainsi qu'entre le sol et l'atmosphère qui se matérialisent par des flux turbulents. Le flux de chaleur sensible H correspond à l'échange de chaleur entre la surface et l'atmosphère et peut être exprimé comme :

$$H = \rho C_p \overline{w'T'} \quad (1.24)$$

H est positif dans le cas où le transfert de chaleur est dirigé vers le haut. Le flux de chaleur latente $L_v E$ correspond aux échanges de chaleur liés aux changements d'état de l'eau :

$$L_v E = \rho L_v \overline{w'q'} \quad (1.25)$$

avec L_v l'enthalpie de chaleur latente. $L_v E$ est positif dans le cas où l'humidité est transportée vers le haut (évaporation). Le bilan des échanges d'énergie à la surface s'exprime en fonction des flux radiatifs et turbulents :

$$Rn = H + L_v E + G \quad (1.26)$$

où G est le flux de chaleur transmis dans le sol par conduction. Enfin, le flux de quantité de mouvement τ s'écrit :

$$\tau = -\rho \left(\overline{u'w'} + \overline{v'w'} \right)^{1/2} \quad (1.27)$$

1.1.4 La structure verticale de la CLA

L'évolution des échanges de chaleur entre le sol et l'atmosphère impacte la structure de la CLA (figure 1.1).

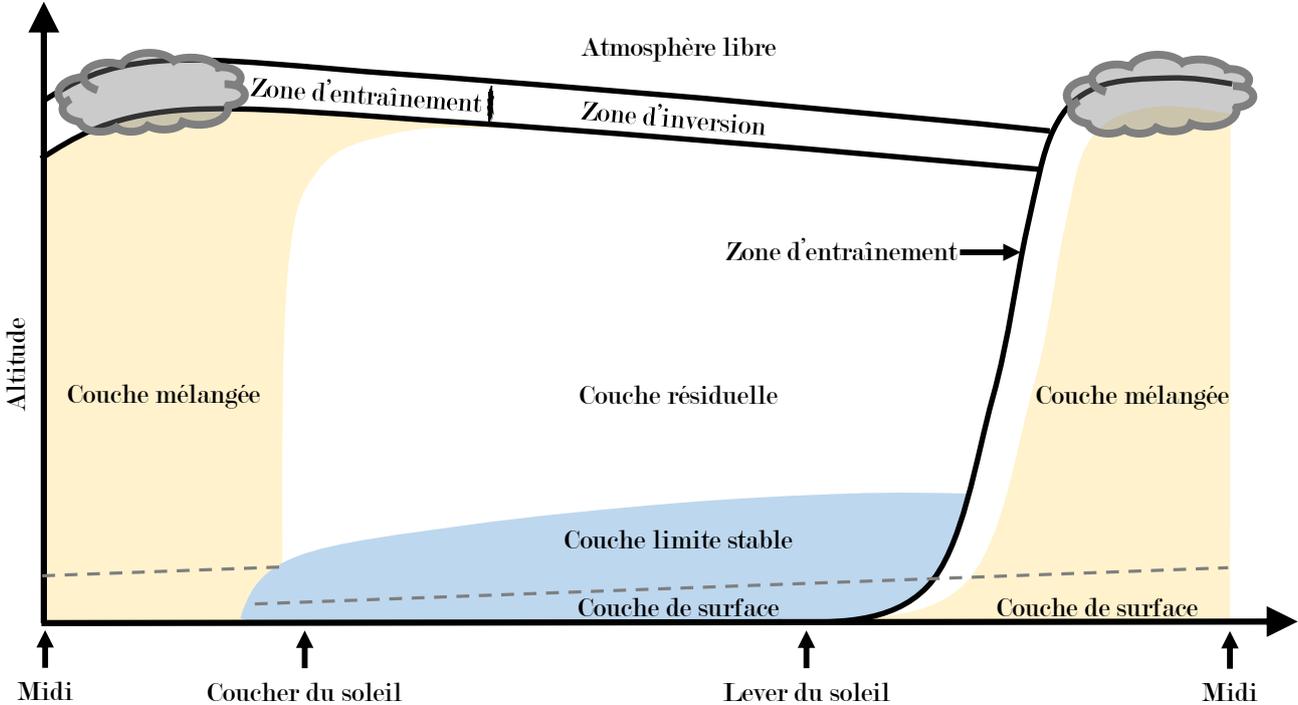


FIGURE 1.1 : Schéma de l'évolution de la couche limite au cours d'une journée où la couverture nuageuse est peu importante et le vent est faible. Adapté de [Stull \(2012\)](#).

En journée, le rayonnement net à la surface est positif grâce au flux solaire, ce qui signifie que le sol se réchauffe. Ce réchauffement est transmis à l'air proche du sol (par les flux turbulents de chaleur) qui devient plus chaud que l'air situé au-dessus : $\partial\theta_v/\partial z < 0$. La CLA est donc instable, ce qui crée des mouvements convectifs et par conséquent la CLA est mélangée. Dans certains cas, ces mouvements convectifs peuvent former des nuages au sommet de la CLA. La composante d'origine thermique de la turbulence dans cette couche est donc importante, et même prépondérante par vent faible quand le cisaillement de vent est faible.

En fin d'après-midi, le flux solaire diminuant, la quantité d'énergie reçue par la surface décroît également. De plus, dans le cas où le ciel est dégagé, le rayonnement LW_{\uparrow} est supérieur au rayonnement LW_{\downarrow} parce que la température de la surface est plus élevée que la température équivalente de l'atmosphère (considérée comme un corps unique rayonnant). Il résulte alors un refroidissement de la surface, et consécutivement de la couche d'air à son contact ($H < 0$) : $\partial\theta_v/\partial z > 0$. Le refroidissement de l'air peut éventuellement se trouver atténué s'il s'accompagne de condensation ($L_v E < 0$). Le refroidissement des couches proches du sol conduit à un puits pour la turbulence puisque la flottabilité est négative. Si de plus le vent est faible, la turbulence créée par le cisaillement n'est pas assez importante pour mélanger la couche qui devient alors une couche stable. Au-dessus de la couche stable se trouve la couche résiduelle qui correspond au reliquat de la couche mélangée de la journée passée.

La couche de surface

La partie la plus basse de la couche limite, caractérisée par de faibles variations de flux turbulents et où la force de Coriolis est négligeable, est appelée la couche de surface. Son extension verticale atteint quelques mètres à plusieurs dizaines de mètres et représente environ 10 % de l'extension verticale de la CLA. Le profil vertical de la vitesse du vent y est décrit comme suivant une loi logarithmique selon la relation suivante :

$$u(z) = \frac{u_*}{k} \ln \left(\frac{z}{z_0} \right) \quad (1.28)$$

avec u le module du vent, k la constante de von Karman (qui vaut environ 0,4), z la hauteur au-dessus du sol ou de la canopée, z_0 la longueur de rugosité en mètres (définie comme la hauteur à laquelle le vent devient nul) et u_* la vitesse de frottement qui s'exprime par :

$$u_* = \left(\overline{u'w'^2} + \overline{v'w'^2} \right)^{\frac{1}{4}} \quad (1.29)$$

La relation entre la vitesse du vent et l'altitude traduit que l'écoulement près de la surface subit une force de friction qui le modifie. Ce modèle convient tel quel aux conditions neutres (Malardel, 2005) et doit être complété pour les conditions stables ou instables mais il permet d'illustrer le cisaillement de vent générateur de turbulence sous l'effet de la rugosité de la surface.

La théorie de similitude de Monin-Obukhov (ou MOST pour Monin-Obukhov similarity theory) a permis de développer des relations empiriques décrivant l'état de l'atmosphère dans la couche de surface. La longueur d'Obukhov (L) est alors un paramètre utile pour la description de l'état de la couche de surface :

$$L = \frac{-\overline{\theta}_v u_*^3}{kg \left(\overline{w'\theta'_v} \right)} \quad (1.30)$$

La théorie de la similitude fait ressortir le rapport z/L (avec z la hauteur de mesure), qui est une grandeur caractéristique de la stabilité statique dans la couche de surface :

- si $z/L > 0$: la couche est statiquement stable ;
- si $z/L = 0$: la couche est neutre ;
- si $z/L < 0$: la couche est statiquement instable.

Pour les conditions proches de la neutralité, z/L est en première approximation le rapport du terme de production par la flottabilité au terme de production dynamique de l'équation d'évolution de l'ECT (équation 1.18). Pour ces conditions, $z/L \approx Ri$ (Businger et al., 1971) et L caractérise donc la hauteur à laquelle la flottabilité commence à devenir le processus dominant, par rapport au cisaillement, dans la génération de turbulence.

1.2 La CLA sur terrain complexe

1.2.1 Écoulements sur les zones au relief complexe

Sur des zones au relief complexe, les écoulements sont générés par la combinaison des effets thermiques et de la modification des vents méso-échelle par le relief. Ces effets sont décrits par [Whiteman \(2000\)](#).

1.2.1.a Écoulements forcés par le relief

Écoulements autour d'un relief

Les écoulements forcés par le relief résultent d'un couplage entre les vents en altitude, représentatifs de l'écoulement libre, et ceux près de la surface, modifiés par le relief. Ces modifications peuvent intervenir à la fois sur la direction et la vitesse du vent qui peut être forcé à passer au-dessus d'un obstacle ou à le contourner en fonction de trois facteurs ([Whiteman, 2000](#)) :

- la stabilité de la masse d'air qui approche le relief ;
- la vitesse de l'écoulement ;
- les caractéristiques géométriques du relief à franchir.

Dans le cas où la masse d'air est instable ou neutre, l'air n'a aucun mal à s'élever et à franchir le relief. En revanche, le comportement d'une masse d'air stable est plus compliqué à anticiper. Il peut néanmoins être caractérisé par le nombre de Froude, Fr , qui est un nombre adimensionnel prenant en compte les 3 facteurs listés ci-dessus :

$$Fr = \frac{U}{NH} \quad (1.31)$$

où U est la vitesse horizontale du vent, N la fréquence de Brunt-Väisälä et H une longueur caractéristique de la hauteur de l'obstacle à franchir. Plus le relief est haut, plus l'énergie nécessaire à son franchissement est importante. De même pour un contournement, plus l'obstacle est large et plus l'énergie nécessaire est importante ([Whiteman, 2000](#)). [Stull \(2012\)](#) décrit les différences de comportement de l'écoulement en fonction de la valeur du nombre de Froude.

Les nombres de Froude faibles ($Fr \ll 1$) sont caractéristiques d'un écoulement lent dans une masse d'air stable. Dans ce cas, l'air a généralement tendance à contourner le relief. Cependant, il est également possible que l'écoulement soit différencié en fonction de l'altitude, l'air le plus près du sol contournant le relief quand la partie se trouvant en altitude passe au-dessus. Dans ce dernier cas, le flux de l'air ayant franchi le relief par le haut forme des ondes orographiques à l'aval de l'obstacle.

Dans le cas où la masse d'air est faiblement stable, le vent fort ou l'obstacle petit, le nombre de Froude est élevé ($Fr > 1$) et l'air a tendance à passer par dessus l'obstacle. Ce cas de figure se présente également lorsque l'obstacle est très large, rendant le contournement de ce dernier difficile.

Canalisation par le relief

Le contournement des reliefs par les écoulements peut conduire à une canalisation entre deux reliefs (figure 1.2). Dans ce cas, les vents longent les reliefs de façon à ce que leur direction

soit alignée avec l'orientation du relief. C'est le cas de certains vents régionaux comme par exemple le Mistral qui est un vent canalisé entre les Alpes et le Massif Central, ou la Tramontane et l'Autan qui sont canalisés entre les Pyrénées et le Massif Central.

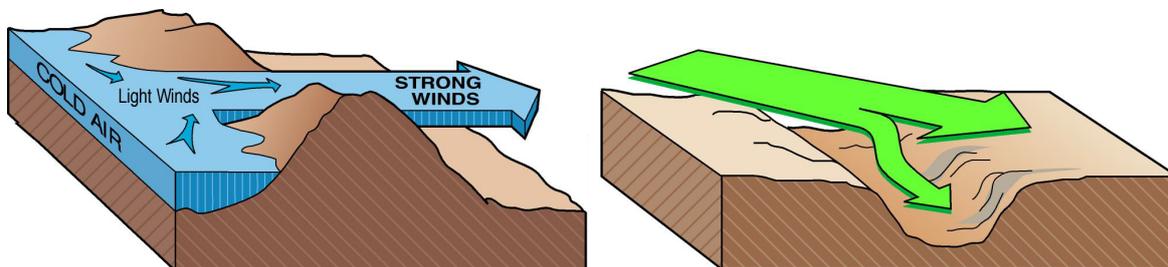


FIGURE 1.2 : Schématisation de la canalisation et de l'accélération de la vitesse de l'écoulement au passage d'un col (à gauche) et de la canalisation des vents d'altitude dans une vallée (à droite), extrait de Whiteman (2000).

À une échelle spatiale plus locale, au passage au-dessus d'une vallée, le vent peut être canalisé de manière à ce que le vent dans la vallée soit aligné avec l'axe de la vallée (figure 1.2). Dans ce cas, le changement de direction entre le vent au-dessus et celui dans la vallée sera le plus faible possible, déterminant la direction montante ou descendante du vent dans la vallée.

1.2.1.b Écoulements forcés par des effets thermiques

Vents de pentes thermiques

Le relief peut également impacter les écoulements sans nécessairement qu'il y ait d'interaction avec le flux d'air circulant au-dessus des reliefs. Sous l'effet de différences thermiques spatiales, l'inclinaison du sol, même faible avec des pentes de l'ordre de 0,1 % (Brost and Wyngaard, 1978; Mahrt, 1981), peut suffire à générer des vents de pentes.

Le mécanisme de formation de ces vents (figure 1.3) est basé sur les gradients horizontaux de température. Les variations de températures de l'air sont fortement dépendantes de la distance de la masse d'air au sol, l'air près du sol étant plus rapidement impacté par les échanges de chaleur avec le sol que l'air situé au-dessus. Ainsi par ciel clair, en journée, l'air près du sol est plus chaud que l'air au-dessus et inversement pendant la nuit. Sur un terrain pentu, ces gradients verticaux de température engendrent également des gradients horizontaux de température puisque deux masses d'air situées à la même altitude à deux endroits différents de la pente sont situées à des distances différentes de la surface. Ces gradients horizontaux de température sont également associés à des gradients de pression qui sont responsables de la formation de vents orientés des hautes vers les basses pressions. De plus, la flottabilité rajoute une composante verticale au vent créé par le gradient horizontal de température.

En journée, l'air au contact du sol est chauffé; il est donc plus léger que l'air au-dessus et que l'air à la même altitude situé en aval donc plus loin de la surface. Le vent a donc une composante horizontale orientée de façon à ce que l'air se rapproche de la surface et une composante verticale orientée vers le haut, ce qui produit un vent qui remonte la pente : c'est un vent anabatique. Pendant la nuit, l'effet inverse se produit avec un refroidissement de l'air près de la surface. Le vent produit est alors orienté de façon à ce qu'il descende la pente : c'est un vent catabatique. Le profil vertical typique du vent de pente présente des vitesses faibles près du sol

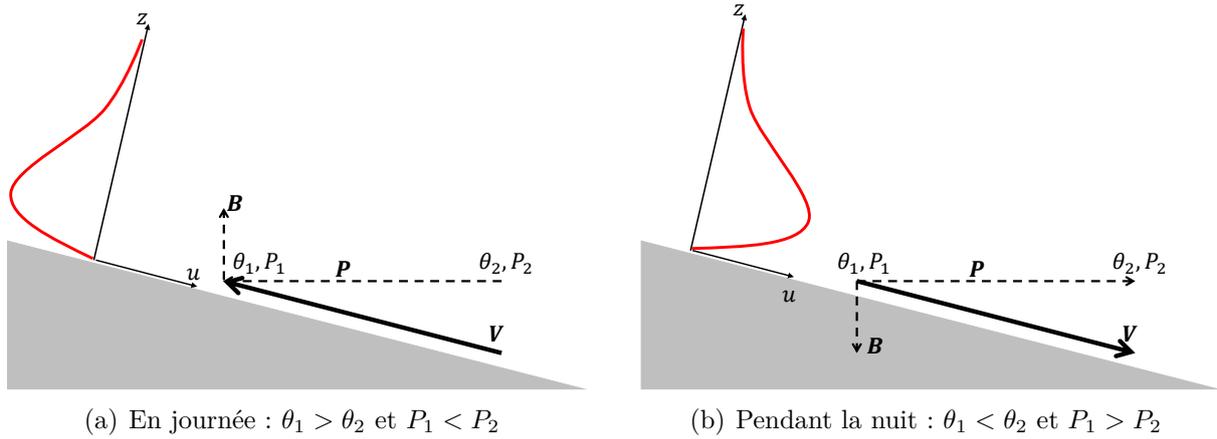


FIGURE 1.3 : Schématisation des forces responsables de la formation des vents de pente \mathbf{V} . \mathbf{B} représente la flottabilité et \mathbf{P} la force de pression. Le profil vertical de vent type est tracé en rouge.

à cause des frottements avec le sol. Les frottements diminuent avec l'altitude ce qui permet au vent d'atteindre un maximum de vitesse à une hauteur variable en fonction du site.

Ces vents ne peuvent être formés que si le forçage exercé par les vents en altitude est faible. Dans le cas contraire, la turbulence générée par ces vents peut créer suffisamment de mélange pour empêcher la formation d'un vent de pente.

Vents de vallées

À l'échelle d'une vallée, ces vents d'origine thermique engendrent une circulation tridimensionnelle qui peut être décomposée en une composante suivant l'axe de la vallée et une composante transverse à la vallée sur les pentes qui constituent ses parois. Cette circulation suit une évolution diurne.

Le long de l'axe de la vallée, le vent remonte la vallée en journée et la descend pendant la nuit. Ce vent s'établit sur une épaisseur pouvant varier de 10 à 400 m avec des vitesses maximales comprises entre 1 et 8 m s^{-1} en fonction de la profondeur de la vallée (Neff and King, 1987). Par exemple, Duine et al. (2017) ont observé un vent de vallée d'une épaisseur d'environ 50–100 m avec des vitesses maximales de $2\text{--}3 \text{ m s}^{-1}$ à environ 30 m au-dessus du sol dans une vallée de 100 m de profondeur. Dans une vallée plus profonde (600 m), le vent de vallée observé par Clements et al. (1989) s'établit sur une épaisseur de 400–600 m avec des vitesses maximales de $5\text{--}6 \text{ m s}^{-1}$ à 80–100 m au-dessus du sol. Une branche de retour associée à cette circulation est parfois présente au-dessus de ce vent de vallée, dans la direction opposée (Stull, 2012).

La circulation qui s'établit dans le plan transverse à la vallée a également une origine thermique. L'évolution de cette circulation est représentée sur la figure 1.4. En fin de journée, la CLA est bien mélangée et la diminution de la turbulence au coucher du soleil crée une couche résiduelle neutre dans la vallée. Au cours de la nuit, le refroidissement radiatif de la surface entraîne un refroidissement de l'air près du sol dans la vallée, générant une couche d'inversion et des vents catabatiques qui descendent le long des flancs de la vallée. La couche d'inversion est d'abord confinée à l'air proche de la surface puis elle s'épaissit au cours de la nuit au point d'occuper toute la vallée. Au lever du soleil, l'air se réchauffe d'abord par le bas, détruisant la stratification près du sol. Des vents de pente anabatiques apparaissent et créent une circulation

avec des branches montant le long des pentes et une branche descendant dans le centre de la vallée. La stratification est alors peu à peu détruite par le bas sous l'effet de la turbulence créée par le réchauffement de l'air près du sol et par le haut sous l'effet du mélange créé par la branche de la circulation qui redescend au centre de la vallée.

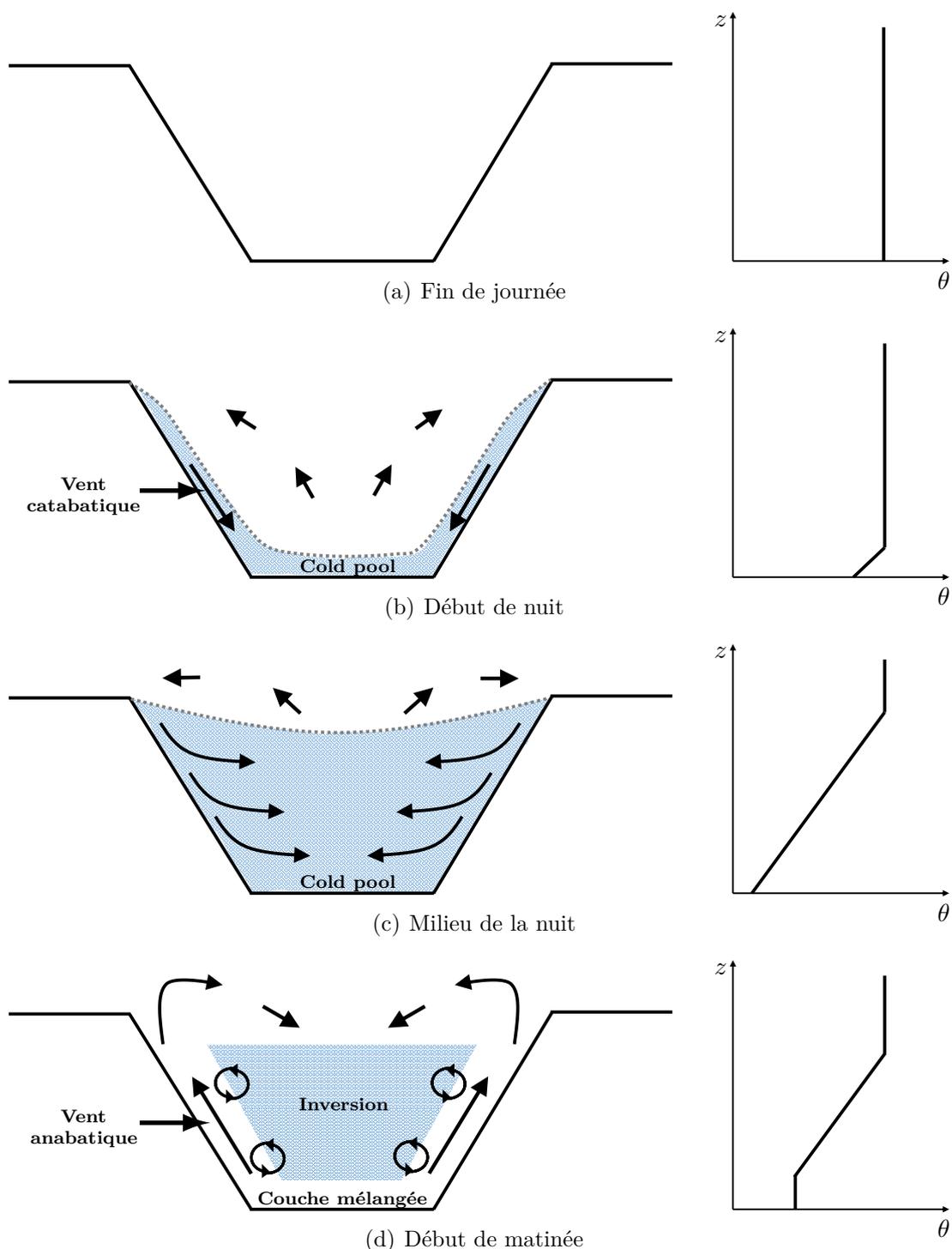


FIGURE 1.4 : Schématisation de l'évolution de la stratification et des vents créés sur un plan transverse à l'axe d'une vallée. Les profils verticaux de température potentielle représentent la stratification dans le centre de la vallée. Adapté de [Stull \(2012\)](#).

1.2.1.c Origine des vents dans une vallée

Dans la réalité, les vents dans une vallée sont générés par une combinaison des deux phénomènes décrits précédemment. [Whiteman and Doran \(1993\)](#) ont décrit une méthodologie permettant d'estimer le phénomène dominant la formation du vent dans une vallée en étudiant la relation entre les vents dans et au-dessus de la vallée. Ils proposent 4 mécanismes d'interaction permettant d'expliquer la formation des vents dans la vallée.

« Downward momentum transport » (flux descendant de quantité de mouvement)

Le premier mécanisme proposé décrit un couplage fort du vent dans la vallée avec le vent au-dessus. Dans ce cas, le vent en altitude transmet sa direction aux vents des couches inférieures et ce jusqu'au sol. Les vents aux différentes altitudes ont donc tous des directions identiques. Il est à noter toutefois que dans le cas d'un forçage par le vent géostrophique, l'augmentation des frottements à l'approche de la surface peut provoquer une rotation entre le vent géostrophique et le vent près de la surface. Ce phénomène est le plus fréquemment observé dans une atmosphère instable ou neutre et dans des vallées larges et peu profondes. La figure 1.5.(a) montre la relation entre les directions des vents dans et au-dessus de la vallée dans le cas où le « downward momentum transport » est le mécanisme prépondérant.

« Thermally driven » (forçage thermique)

Dans ce cas de figure, les vents dans la vallée suivent le mécanisme de formation des vents de pentes. Ils sont donc générés par des gradients locaux de pression le long de la vallée. Ces vents sont alors complètement découplés de ceux présents au-dessus de la vallée et il n'y a alors pas d'interaction entre les deux vents. Ainsi, quelle que soit la direction du vent au-dessus, le vent dans la vallée montera la vallée en journée et la descendra pendant la nuit. Les conditions nécessaires à la formation de ces vents sont un faible forçage par les vents en altitude et des conditions stables la nuit. La relation entre les directions des deux vents est représentée sur la figure 1.5.(b) pour une vallée orientée du sud-est vers le nord-ouest (comme la vallée de Cadarache qui est la vallée étudiée dans cette thèse).

« Forced channeling » (canalisation)

Le troisième mécanisme décrit un phénomène de canalisation des vents par les flancs de la vallée. Les vents d'altitude sont alors redirigés en descendant dans la vallée afin qu'ils s'alignent avec l'axe de la vallée. Le vent dans la vallée n'a alors que deux directions possibles, montant ou descendant la vallée en fonction de la direction du vent au-dessus de la vallée. La canalisation est plus efficace dans les vallées profondes et étroites ([Whiteman and Doran, 1993](#); [Weber and Kaufmann, 1998](#)). La relation entre ces vents est représentée sur la figure 1.5.(c).

« Pressure driven channeling » (forçage par gradient de pression)

Enfin, le dernier mécanisme d'interaction est basé sur les gradients de pression. Il est représenté sur la figure 1.5.(d). Dans ce cas, le vent au-dessus de la vallée est aligné avec les isobares et le champ de pression crée un gradient de pression le long de l'axe de la vallée. Dans la vallée, le vent s'écoule des hautes vers les basses pressions (la force de Coriolis étant négligeable) tout en étant canalisé. La particularité de ce mécanisme est que le vent géostrophique, qui est dirigé par le gradient de pression, peut avoir une direction opposée à celle du vent dans la vallée ce

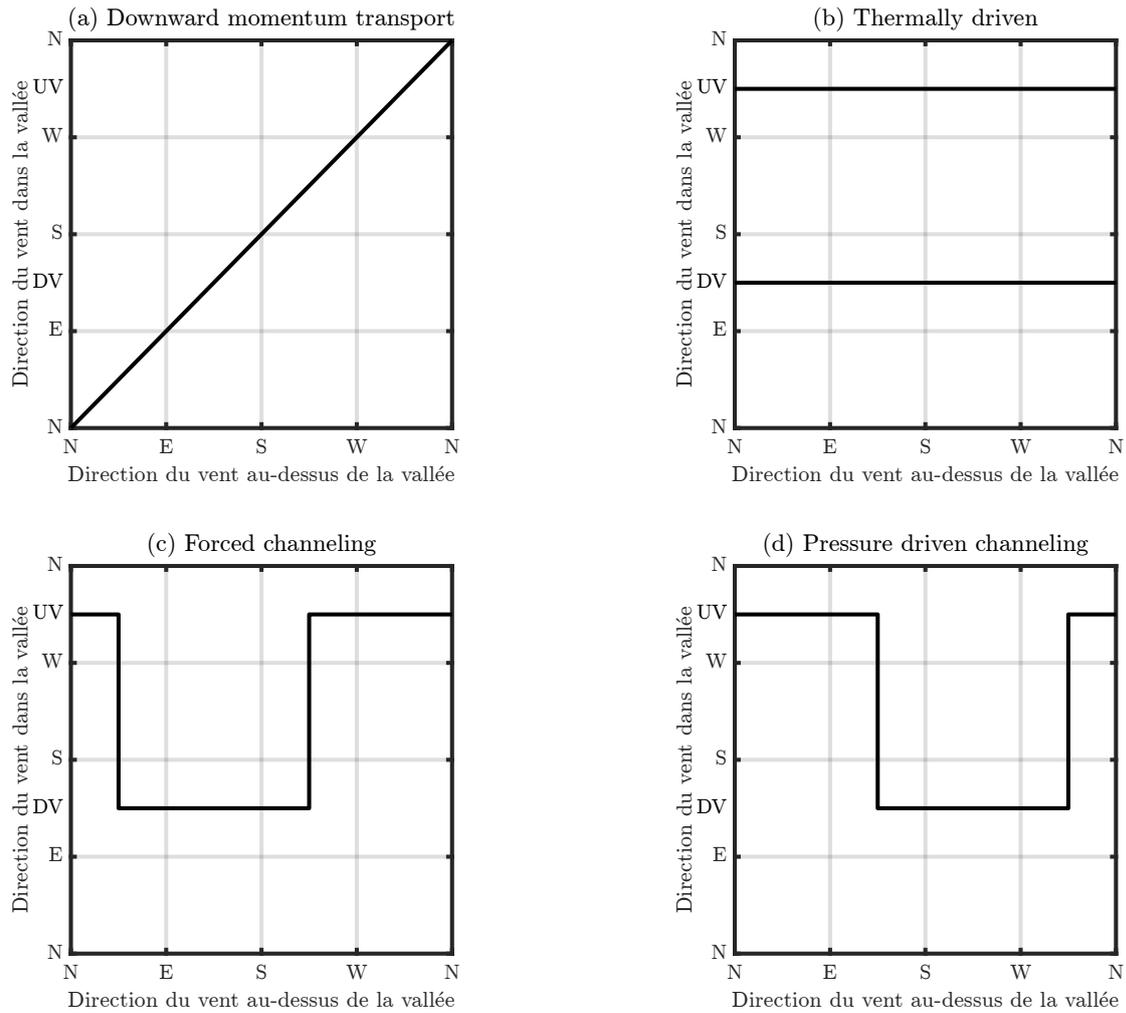


FIGURE 1.5 : Relation entre les directions des vents au-dessus d’une vallée et dans la vallée pour quatre forçages différents. Dans cet exemple, la vallée est orientée du sud-est vers le nord-ouest comme la vallée de Cadarache qui est étudiée dans cette thèse. UV (pour upvalley) et DV (pour downvalley) repèrent les directions des vents montant et descendant la vallée en suivant son axe. Adapté de [Whiteman and Doran \(1993\)](#).

qui crée donc un contre-courant en altitude par rapport au vent dans la vallée ([Wippermann, 1981, 1984](#); [Weber and Kaufmann, 1998](#)).

1.2.1.d Distinction du couplage/découplage des vents

Plusieurs études ont tenté de distinguer les conditions menant à un couplage/découplage des vents dans et au-dessus d’une vallée sur la base d’une adaptation du nombre de Froude Fr (équation 1.31) pour lequel la longueur caractéristique est la profondeur de la vallée.

Une valeur critique de Fr permet de distinguer les conditions pour lesquelles le vent présent au-dessus de la vallée est capable d’imposer ou non ses caractéristiques aux vents dans la vallée. Dans une vallée d’environ 150 m de profondeur, [LeMone et al. \(2003\)](#) ont trouvé que des valeurs de Fr supérieures à 3,3 correspondent à des situations de couplage entre les vents dans et au-dessus de la vallée. [Mason \(1987\)](#) a trouvé une valeur critique de 1 à partir d’observations et [Bell and Thompson \(1980\)](#) ont trouvé une valeur critique de 1,3 à partir d’expérimentations numériques et en laboratoire.

Dans d'autres cas, $1/Fr$ est utilisé pour cette caractérisation. Ce paramètre est appelé profondeur adimensionnelle de la vallée et est noté \hat{H} . À partir de simulations numériques de vallées peu profondes (inférieures à 150 m) en 2 dimensions, [Vosper and Brown \(2008\)](#) ont trouvé une valeur critique de \hat{H} comprise entre 0,35 et 0,64 (Fr entre 2,86 et 1,56) en fonction de la profondeur de la vallée simulée. [Holden et al. \(2000\)](#) ont trouvé une valeur critique de \hat{H} valant 0,5 à partir d'observations réalisées dans une vallée de 400 m de profondeur.

Dans ces études, les écoulements présents au-dessus de la vallée ont des directions transverses à l'axe de la vallée. Pour des vents en altitude dont les directions sont alignées avec l'axe de la vallée, [Largerion \(2010\)](#) a trouvé des valeurs de Fr critique comprises entre 0,06 et 0,30. Ces valeurs sont bien plus petites que dans le cas de vents transverses ce qui montre que, dans le cas où les vents en altitude sont alignés avec l'axe de la vallée, la stabilité doit être plus forte ou les vents en altitude plus faibles pour que les écoulements soient découplés.

Ces nombreux exemples montrent que la caractérisation des écoulements dans une vallée est possible à partir de grandeurs telles que le nombre de Froude ou la profondeur adimensionnelle de la vallée. Cependant, les différences de valeurs critiques permettant la distinction des situations de couplage et découplage semblent indiquer que cette valeur critique est propre à chaque vallée.

1.2.2 Stratification dans une vallée

Dans les vallées stratifiées, la circulation générée par les vents de pente peut être responsable d'une accumulation d'air froid dans le fond de la vallée qui peut engendrer des différences de température très importantes entre l'air froid près de la surface et l'air plus chaud au-dessus. Dans certains cas, ces différences peuvent largement dépasser la dizaine de degrés si le relief contrarie le drainage de l'air froid ([Clements et al., 2003](#); [Zängl, 2005](#)).

[Mahrt \(2017\)](#) a présenté une classification de régimes de stratification dans une vallée afin d'en décrire la structure. Pour cela il se base sur 3 indicateurs caractérisant la stratification.

- Le CPI (pour cold-pool intensity) représente l'intensité de la stratification sur le côté de la vallée, c'est donc la différence de température potentielle entre la crête et le fond de la vallée.
- Le $D\theta$ représente le refroidissement que subit une parcelle d'air lors de sa descente de la crête jusqu'au fond de la vallée.
- Le $-\theta^*$ représente le déficit de température potentielle au niveau de la crête de la vallée, c'est à dire la différence de température potentielle entre l'air situé près du sol au niveau de la crête et l'air situé à la même altitude au centre de la vallée.

Mature cold pool

Le régime appelé « mature cold pool » est défini par des surfaces isentropes horizontales (figure 1.6.a). Cette stratification est favorisée par un refroidissement radiatif important correspondant à des valeurs de CPI élevées, de faibles valeurs de $-\theta^*$ et un drainage faible le long de la vallée ([Sheridan et al., 2014](#)). Dans ce cas, les mesures de température potentielle effectuées sur les côtés de la vallée sont un bon indicateur de la distribution verticale des températures potentielles au centre de la vallée ([Whiteman et al., 2004a](#)). Cette cold pool mature ne couvre pas obligatoirement la profondeur de la vallée dans sa totalité.

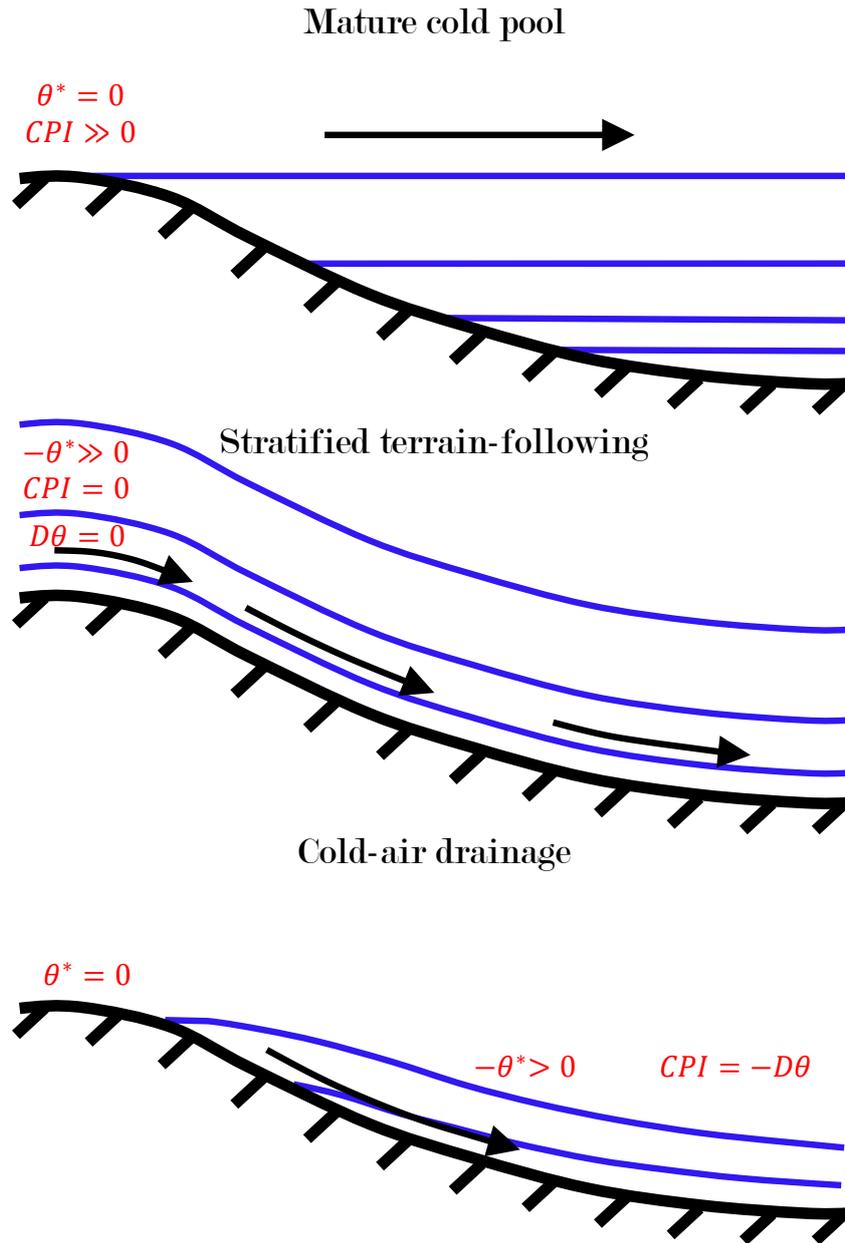


FIGURE 1.6 : Schématisation de différents régimes de stratification qui peuvent se produire dans une vallée d'après Mahrt (2017).

Les vents présents au-dessus de la vallée ne pénètrent pas dans la cold pool ce qui crée un découplage avec les écoulements dans la vallée. Ce régime est associé à des vallées profondes marquées par une forte stratification et un faible écoulement de grande échelle (\hat{H} élevé ou Fr bas).

Stratified terrain-following

Le deuxième régime est appelé « stratified terrain-following » (figure 1.6.b). Dans ce cas, les surfaces isentropes sont parallèles au sol (CPI faible et $-\theta^*$ élevé). Le vent présent en dehors de la vallée descend jusqu'au fond de celle-ci en suivant les isentropes et doit être suffisamment rapide pour que l'air ne se refroidisse pas au contact du sol en descendant dans le fond de la vallée ($D\theta$ faible). Ce régime empêche l'accumulation d'air froid dans le fond de la vallée.

Ce régime a été rencontré par [Mahrt \(2017\)](#) dans une petite vallée de 270 m de large et 12 m de profondeur. De manière générale, ce genre de stratification est associé à des vallées peu profondes, à une faible stratification et à un écoulement de grande échelle fort (\hat{H} faible ou Fr élevé).

Cold-air drainage

Ce régime est dirigé par un drainage d'air froid le long des pentes de la vallée avec un refroidissement de l'air en descendant vers le fond de la vallée ($D\theta$ positif, cf. figure 1.6.c). L'air traverse donc les isentropes en descendant (au contraire des deux régimes décrits plus haut). Le refroidissement produit lors de la descente est équivalent à l'intensité de la stratification sur le côté de la vallée ($CPI \approx -D\theta$). Ce régime de stratification est généralement rencontré en fin d'après-midi sur les pentes ombragées. Le gradient horizontal de température potentielle au sommet de la vallée est faible comparé à l'intensité de la stratification le long de la vallée ($-\theta^* \ll CPI$).

Marginal cold pool

Ce régime est utilisé pour définir les cold pools dont les caractéristiques sont situées à mi-chemin entre celles des cold pools matures et des stratified terrain following. Dans ce cas, l'intensité de la stratification est plus faible que dans le cas d'une cold pool mature mais il existe également une différence horizontale de température au niveau du sommet de la vallée ($-\theta^* > 0$).

1.2.3 Sources d'hétérogénéités liées aux terrains complexes

Les flux d'énergie échangés à la surface pilotent un grand nombre de processus atmosphériques, notamment dans la CLA. Sur les terrains complexes, la variabilité de la topographie et des propriétés de la surface génèrent des évolutions spatiales et temporelles particulières. En conditions stables, la faible intensité de la turbulence ne permet pas de bien homogénéiser les hétérogénéités spatiales ainsi causées.

1.2.3.a Modulation du rayonnement net par le relief

Le relief impacte significativement les quatre composantes du rayonnement net à la surface (équation 1.23). Une pente ainsi que son orientation modifient la quantité d'énergie reçue sous forme de rayonnement solaire ainsi que la distribution de cette énergie au cours de la journée ([Whiteman et al., 1989](#); [Whiteman, 2000](#); [Barry, 1992](#)). Les ombres générées par la topographie peuvent modifier l'heure apparente du lever et coucher du soleil ce qui a également un impact sur la quantité de rayonnement solaire reçue ([Whiteman, 1990](#)).

Concernant le rayonnement dans les grandes longueurs d'ondes, le relief modifie la part de ciel vue par la surface ce qui a pour conséquence de créer une importante variabilité spatiale de rayonnement reçu ([Whiteman, 1990](#); [Dubayah and Loechel, 1997](#)). La diminution de la part de ciel vue par la surface (au profit de la surface des flancs de la vallée) entraîne une augmentation de l'énergie reçue par rayonnement LW_{\downarrow} en journée dans le fond de la vallée par rapport aux flancs de la vallée ([Matzinger et al., 2003](#)). Dans le cas de cuvettes orographiques, [Whiteman et al. \(2004b\)](#) suggèrent même que la diminution de la surface de ciel vue est le paramètre

topographique principal contrôlant le refroidissement par conditions de vent calme et ciel clair. Enfin, la création de cold pool par conditions stables a pour effet de réduire le rayonnement LW_{\uparrow} dans le fond de la cold pool par rapport aux surfaces situées plus haut (Hoch and Whiteman, 2010).

Des observations de variabilité du rayonnement ont été faites dans un cratère de météorite en Arizona (Hoch and Whiteman, 2010). La géométrie régulière du site a facilité l'interprétation des différences observées. La plus grande hétérogénéité est causée par les différences de rayonnement solaire reçu liées aux pentes et à leur orientation par rapport à l'angle d'incidence du rayonnement direct. Les différences de rayonnement SW_{\downarrow} journalier liées à l'exposition de la surface (combinaison de l'orientation et de la pente) peuvent atteindre 6% d'un site à un autre. L'ombre produite par les parois sur le fond du cratère génère une diminution de 6% du rayonnement SW_{\downarrow} journalier. Le rayonnement SW_{\downarrow} diffus est plus important au fond du cratère que sur sa crête à cause de la réflexion par les parois du cratère. Enfin, les parois impactent également le rayonnement LW net, tout comme la diminution du rayonnement LW_{\uparrow} émis dans le fond du cratère par conditions stables du fait de la formation d'une cold pool (la perte d'énergie au fond du cratère ne représente alors que 72% de la perte en haut du cratère).

1.2.3.b Modulation des flux turbulents par les hétérogénéités de la surface

L'état de la CLA dépend grandement de la transformation de l'énergie disponible à la surface en chaleur sensible et chaleur latente. Ces transformations sont entre autres modulées par l'humidité dans le sol, l'occupation et le type de sol (Entekhabi et al., 1996). Les hétérogénéités de nature du sol font apparaître des différences d'albédo qui modulent le rayonnement SW_{\uparrow} amenant à une absorption plus ou moins importante d'énergie par le sol (flux dans le sol G , cf. équation 1.26). L'énergie absorbée est ensuite restituée à l'atmosphère essentiellement sous forme de flux turbulents de chaleur (H et $L_v E$). Par exemple, Beyrich et al. (2002) ont observé des différences importantes de flux turbulents liées à la couverture du sol. Les hétérogénéités spatiales d'occupation du sol conduisent ainsi à des différences spatiales de température qui peuvent être importantes.

L'évaporation et la transpiration par la végétation permettent de transférer indirectement la chaleur du sol vers l'atmosphère via le flux de chaleur latente. Les hétérogénéités d'humidité dans le sol ainsi que le type de végétation peuvent ainsi conduire à des modifications dans le bilan des flux turbulents de chaleur (Taha, 1997; Arnfield, 2003; Ronda et al., 2002) responsables d'hétérogénéités de températures.

1.3 Modélisation numérique de la CLA stable sur terrain complexe

Une simulation numérique de l'état de l'atmosphère est composée d'un ensemble de modèles couplés visant à représenter différents types de phénomènes. Par exemple, le modèle de transfert radiatif sert au calcul des flux radiatifs, le modèle de transfert dans le sol permet de calculer des flux de surface etc. Dans cette partie, l'objectif est d'explorer les possibilités actuelles de modélisation numérique de la turbulence atmosphérique afin de comprendre les difficultés rencontrées dans la modélisation de la couche limite en terrain complexe, notamment par conditions stables.

Trois approches basées sur la résolution des équations pronostiques permettent de simuler les écoulements de fluides turbulents, chacune d’elles se différenciant par la méthode utilisée pour représenter la turbulence : la simulation numérique directe (DNS en anglais pour Direct Numerical Simulation), la modélisation statistique faisant intervenir la moyenne de Reynolds dans les équations de Navier-Stokes (RANS en anglais pour Reynolds-Averaged Navier-Stokes, voir partie 1.1.1) et la simulation des grandes échelles (LES en anglais pour Large Eddy Simulation).

La DNS consiste à résoudre les équations de Navier-Stokes (système d’équations formé par les équations 1.1, 1.3, 1.4, 1.5 et 1.8) ce qui impose de résoudre explicitement toutes les échelles spatiales et temporelles de la turbulence. Cette approche est la plus précise du point de vue de la description de l’écoulement turbulent. Cependant, le coût de calcul étant proportionnel à Re^3 (Pope, 2001), la DNS n’est pas adaptée à la simulation des écoulements atmosphériques qui sont fortement turbulents, donc caractérisés par des Re élevés, et couvrent de vastes domaines spatio-temporels.

La simulation RANS a pour but de décrire l’écoulement de manière statistique. Pour cela, la moyenne de Reynolds est appliquée aux équations de Navier-Stokes. Seule la moyenne statistique des variations est représentée explicitement alors que l’effet de la partie fluctuante sur le mouvement moyen est modélisé par une paramétrisation. Cette approche a l’avantage d’être beaucoup moins consommatrice de temps de calcul que l’approche DNS, mais la turbulence n’est plus représentée que par des moments d’ordre 1 ou 2 (modélisation). Pour cela, la résolution du domaine dans une simulation RANS doit être choisie pour que toutes les échelles de la turbulence restent des phénomènes sous-maille, ce qui limite la réduction des tailles de maille.

Du point de vue de la représentation de la turbulence, la LES est à mi-chemin entre la DNS et la simulation RANS puisqu’elle ne vise pas à représenter la totalité du spectre turbulent, mais la partie correspondant aux échelles pertinentes vis à vis d’un besoin. Elle est également plus coûteuse en temps de calcul que les simulations appliquant la méthode RANS.

Pour les prévisions opérationnelles, le besoin d’effectuer les calculs sur des domaines étendus et la limitation des capacités de calcul actuelles imposent le choix des simulations méso-échelle par résolution des équations RANS qui sont moins coûteuses en temps de calcul. Les résolutions kilométriques actuelles des modèles opérationnels (1,1 km pour le modèle opérationnel suisse COSMO, 1,3 km pour le modèle opérationnel de Météo-France AROME) engendrent donc quelques limitations quand les hétérogénéités affectant la CLA concernent des échelles sub-kilométriques.

1.3.1 Représentation de la turbulence par les méthodes RANS et LES

La méthode RANS n’a pas pour objectif d’étudier explicitement le détail du spectre turbulent dans la CLA. Dans ces simulations, la turbulence est donc paramétrée dans un schéma de couche limite planétaire (schéma PBL pour Planetary boundary layer en anglais). Cette paramétrisation suppose une CLA homogène horizontalement à l’intérieur d’une maille. Elle permet de représenter les effets du mélange turbulent vertical dans l’ensemble du domaine et notamment dans la CLA en y prenant en compte les effets des échanges turbulents de chaleur, humidité et quantité de mouvement sur les quantités moyennes. Les flux horizontaux, autres que ceux turbulents qui sont ici négligeables, sont quant à eux majoritairement résolus.

Dans la simulation LES, les mouvements turbulents ne sont représentés qu'en partie. Pour cela, un filtre est appliqué aux équations de Navier-Stokes ce qui permet de résoudre explicitement les structures turbulentes les plus grandes alors que l'effet des petites échelles sur la structure des grandes échelles est modélisé. Dans cette approche, on considère que la description des grandes échelles est suffisante pour décrire l'état d'un écoulement, puisque cet écoulement est gouverné par les structures de grande échelle. Cela suppose donc que la résolution du domaine soit suffisamment importante pour que les grandes échelles de la turbulence ne soient pas des phénomènes sous-maille. Généralement, une LES est considérée satisfaisante quand au moins 80 % de l'énergie cinétique turbulente est représentée explicitement (Pope, 2001).

1.3.2 Domaines d'application des méthodes RANS et LES

Dans un écoulement turbulent, l'énergie cinétique associée à la turbulence est portée par des tourbillons de tailles très variables. Le spectre de la figure 1.7 représente l'énergie cinétique turbulente en fonction de la taille des tourbillons. L'énergie y est injectée par prélèvement sur l'écoulement moyen et se traduit par la création de grands tourbillons. Ces grands tourbillons se fractionnent en tourbillons plus petits (dans la zone inertielle) qui se fractionnent eux mêmes en tourbillons encore plus petits (cascade de Richardson) jusqu'à une taille (échelle de Kolmogorov) où leur énergie se dissipe en chaleur.

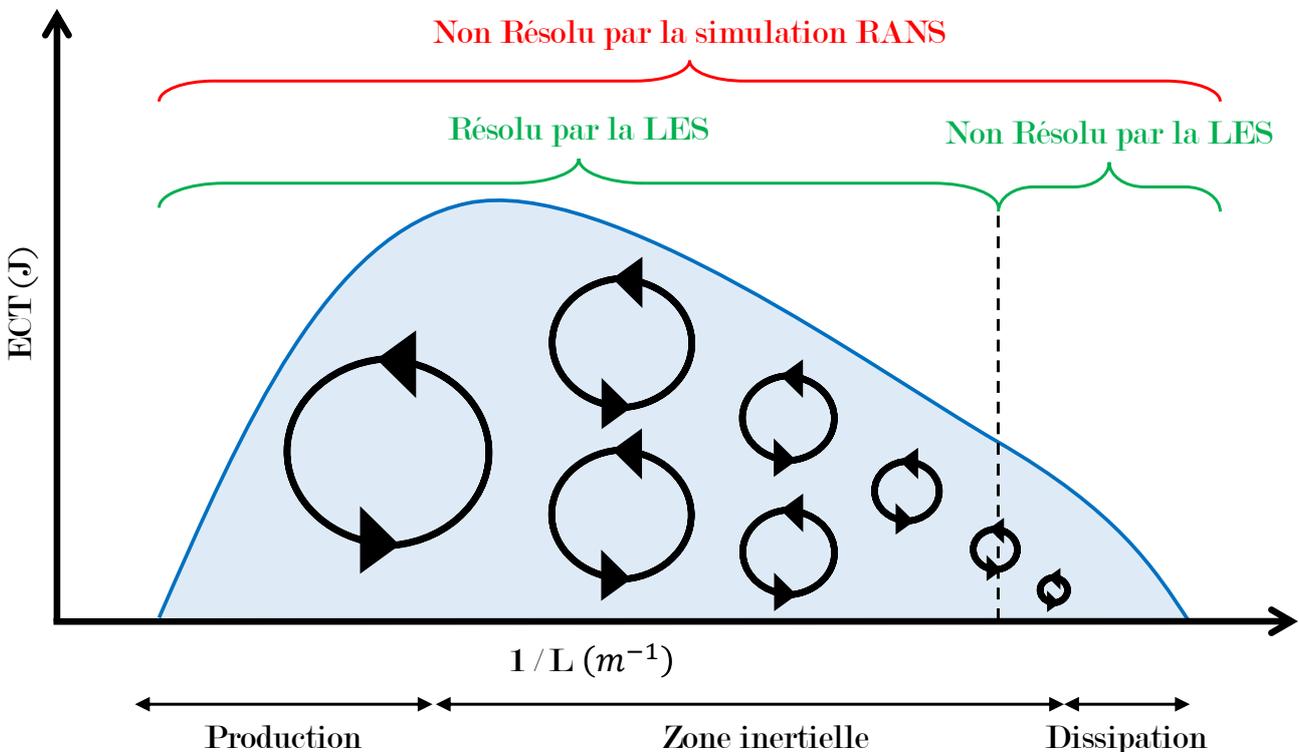


FIGURE 1.7 : Schéma du spectre de l'énergie en fonction de la taille des structures turbulentes (L).

La méthode RANS suppose que toute la turbulence est sous-maille, autrement dit la taille de la maille est plus grande que la taille caractéristique des tourbillons les plus grands. Idéalement, une LES devrait pouvoir représenter explicitement les tourbillons jusqu'à des tailles

correspondant aux plus petits tourbillons de la zone inertielle.

L'application des méthodes RANS et LES étant conditionnée par des limitations dans le choix des résolutions horizontales, [Wyngaard \(2004\)](#) a introduit la « zone grise » (ou « terra incognita »). Elle représente une gamme de résolutions horizontales (environ de 1 km à 100 m) pour lesquelles il n'est ni possible d'utiliser une paramétrisation de type PBL, parce qu'une partie de la turbulence est résolue, ni possible d'utiliser une LES parce qu'une trop grande partie de la turbulence est sous-maille.

[Cuxart \(2015\)](#) donne des ordres de grandeur de résolutions horizontales nécessaires à la représentation de certains phénomènes rencontrés en terrains complexes par une simulation LES. Par exemple, il suggère l'utilisation de résolutions d'environ 20 m pour représenter un jet de vallée et de 5 à 10 m (voire moins) pour une atmosphère stratifiée.

1.3.3 Simulation de cas réels de CLA stable sur terrain complexe

1.3.3.a Représentation des hétérogénéités de la surface

La simulation numérique, pour être précise, nécessite que ses forçages soient réalistes. Cela suppose entre autres que les hétérogénéités de la surface (relief, occupation du sol, ...) soient bien représentées. La résolution horizontale des mailles du domaine simulé est alors déterminante dans l'optique d'une bonne représentation des caractéristiques de la surface puisque chaque maille possède dans le modèle des caractéristiques uniformes. Dans le cas d'une simulation sur un terrain complexe, les hétérogénéités de surface sont nombreuses ce qui signifie qu'une bonne représentation n'est possible qu'avec une haute résolution horizontale.

Plusieurs études ont montré l'importance de la représentation du relief dans une simulation. Des biais systématiques ont été observés dans les vitesses de vent près de la surface causés par une mauvaise représentation du relief ([Milton and Wilson, 1996](#); [Georgelin et al., 2000](#); [Rontu, 2006](#); [Schmidli et al., 2018](#)) produisant des vitesses trop fortes dans le fond des vallées et trop faibles au niveau des crêtes. [Jiménez and Dudhia \(2012\)](#) ont proposé une paramétrisation visant à améliorer la représentation des vitesses de vent de surface dans WRF (le modèle météorologique Weather Research and Forecasting) pour les simulations en terrain complexe. Pour cela, ils utilisent un écart-type du relief sous-maille et le Laplacien de la topographie qui permet de prendre en compte le relief sous-maille. Leur paramétrisation permet d'augmenter la rugosité liée à la sous représentation du relief (où WRF a généralement un biais positif de vent) et d'augmenter les vitesses sur les crêtes des montagnes (où WRF a généralement un biais négatif). Cette méthode permet d'améliorer les calculs de vitesse notamment pendant la nuit mais il subsiste un biais négatif en journée qui peut être corrigé en ajoutant l'effet de la stabilité à la paramétrisation ([Lorente-Plazas et al., 2016](#)). [Santos-Alamillos et al. \(2015\)](#) ont montré que l'amélioration de la couverture du sol peut également permettre d'améliorer le calcul des vitesses et directions du vent. Cependant, [Jiménez and Dudhia \(2013\)](#) ont montré les limites du code WRF dans la reproduction des directions de vents de surface sur un terrain complexe où le relief n'est pas bien représenté.

[Jiménez-Esteve et al. \(2018\)](#) ont montré que l'augmentation de la résolution horizontale et l'amélioration de la représentation de l'occupation du sol donnent en terrain complexe de meilleurs résultats sur la température et l'humidité relative à 2 m. C'est notamment le cas pendant les nuits stables où la température minimale observée dans le fond des vallées est mieux

représentée ce qui conduit donc à une meilleure représentation de l'inversion thermique.

1.3.3.b Des domaines RANS vers les domaines LES

Lorsque l'objectif d'une simulation est de représenter une situation réelle, un grand nombre d'échelles spatiales doivent être prises en compte et chacune d'elles nécessite des moyens de calcul différents, notamment en terme de résolution. La technique de modélisation visant à reproduire le plus finement possible les caractéristiques d'une région particulière, tout en gardant un temps de calcul relativement faible, consiste à imbriquer plusieurs domaines les uns dans les autres. On commence alors par un grand domaine (appelé domaine père) avec une résolution horizontale faible et on finit par un petit domaine (appelé domaine fils) centré sur la région étudiée avec une résolution horizontale élevée. Le domaine père, qui permet de simuler la circulation et les champs de grande échelle, fournit les conditions initiales et aux limites du domaine fils, dont l'objectif est de simuler les caractéristiques locales plus finement. En retour, le domaine fils peut avoir une rétroaction sur le domaine père. Une règle communément utilisée (pour WRF) pour le choix des résolutions des différents domaines consiste à augmenter la résolution horizontale d'un facteur 3 ou 5 du domaine père vers le domaine fils. Cette méthode semble donc être adaptée aux simulations en terrain complexe, puisqu'elle permet d'augmenter la résolution sur une zone précise, mais elle pose toutefois quelques problèmes pour les résolutions avoisinant la zone grise.

Il est donc possible dans une même simulation d'avoir un domaine avec une résolution de plusieurs kilomètres permettant de simuler les mouvements de grande échelle et un domaine avec une résolution kilométrique voire plus fine permettant de simuler les caractéristiques locales. Les phénomènes représentés à ces deux résolutions ne permettent pas toujours l'utilisation des mêmes paramétrisations, notamment pour la représentation de la turbulence pour les raisons évoquées précédemment.

La solution serait donc d'imbriquer un domaine LES dans un domaine RANS. Cependant, il n'est pas possible d'éviter complètement la zone grise en utilisant un facteur 3 ou 5 dans l'augmentation de la résolution horizontale. De plus, les différences dans la représentation de la turbulence (1D pour le domaine RANS et 3D pour le domaine LES) posent des problèmes dans l'initialisation de la turbulence dans le domaine LES. Le domaine père ne fournit pas assez de turbulence au domaine fils pour qu'il produise à son tour suffisamment de turbulence (Muñoz-Esparza et al., 2014). Mirocha et al. (2014) ont tenté d'accélérer la génération de la turbulence en ajoutant des perturbations dans les champs de température potentielle et de vent aux bords du domaine fils mais cela a conduit à des erreurs dans le calcul des champs moyens.

De plus, l'évolution de la taille des plus grands tourbillons de la CLA au cours du cycle diurne complique les simulations de cas réels. Les plus grands tourbillons en journée ont une taille de l'ordre du kilomètre ce qui empêche la paramétrisation complète de la turbulence pour des résolutions inférieures au kilomètre, alors qu'une simulation LES d'une CLA nocturne stable nécessite une résolution d'environ 5 à 10 m au plus (Cuxart, 2015).

1.3.3.c Autres limitations liées aux cas réels

Les paramétrisations physiques utilisées dans les simulations utilisant la méthode RANS ont été développées sur la base d'hypothèses qui considèrent que le terrain est plat et homogène horizontalement (schéma de couche limite planétaire et schéma de rayonnement). La turbulence n'est alors traitée qu'en considérant les échanges verticaux alors que les processus

turbulents dans la CLA en terrain complexe sont pleinement tridimensionnels (Goger et al., 2018). Par exemple, les vents de vallée et vents de pente contribuent aux échanges horizontaux (Wagner et al., 2015; Leukauf et al., 2016). Concernant le schéma de rayonnement, l’inclinaison des pentes modifie le rayonnement solaire incident ce qui doit être pris en compte (Buzzi, 2008).

Cependant, la modélisation dans la zone grise de la turbulence (environ de 1 km à 100 m) peut avoir des avantages pour des simulations en terrain complexe parce qu’il y a une amélioration importante des calculs grâce à une meilleure représentation du relief et de l’occupation du sol (Zhou et al., 2014). Ce résultat est cohérent avec les résultats de Wagner et al. (2014) qui ont montré, à partir de simulations idéalisées d’une vallée représentée avec différentes résolutions horizontales, qu’une bonne représentation de la topographie est plus importante que le choix de représentation de la turbulence.

1.3.3.d Exemples d’applications

Quelques études présentent des applications de simulation à haute résolution en terrain complexe par conditions stables nocturnes.

Jiménez and Cuxart (2014), Udina et al. (2017) et Conangla et al. (2018) ont simulé les vents dans la vallée de La Cerdagne dans les Pyrénées avec WRF à une résolution de 400 m en portant une attention particulière aux phénomènes observés par conditions stables. Cette résolution permet de bien représenter la vallée qui a une largeur de plusieurs kilomètres. Ces simulations à haute résolution permettent notamment de représenter les jets de vallées, les vents catabatiques qui descendent les pentes et l’évolution de la stratification dans la vallée. Vionnet et al. (2015) ont également effectué des simulations à très haute résolution (250 m) sur le Canada avec une amélioration du calcul des champs de température (notamment en altitude) et de vent.

Pour ces simulations, la turbulence est représentée par un schéma 1D. Ce choix est cohérent avec les recommandations de Cuxart (2015) qui préconise un choix de représentation de la turbulence adapté aux phénomènes étudiés pour les domaines qui ont des résolutions qui tombent dans la zone grise. Pour les études portant sur les phénomènes de CLA stable, une résolution de quelques centaines de mètres ne permet pas de représenter les grands tourbillons et un schéma 1D est alors mieux approprié (bien qu’il ne soit pas approprié en journée). Au contraire, si l’étude porte sur des phénomènes convectifs, pour la même résolution horizontale, un schéma 3D représentera mieux les phénomènes turbulents.

Schmidli et al. (2018) ont montré que l’augmentation de la résolution, de 2,2 km à 1,1 km permet de mieux représenter les caractéristiques des vents dans des vallées larges ou moyennement larges, mais n’améliore pas les calculs dans les petites vallées où une résolution kilométrique n’est pas suffisante pour représenter avec assez de précision le relief. Pour les petites vallées comme la vallée de Cadarache, une résolution de 1 km ne permet pas de représenter correctement la topographie ce qui conduit à des erreurs de prévision notamment pour le vent (Duine, 2015). Il est alors important de bien comprendre les phénomènes en jeu afin de trouver des techniques alternatives pour la prévision ou la descente d’échelle (Vosper and Brown, 2008).

1.4 Modélisation statistique

Certaines situations, comme les conditions stables en terrain complexe, sont hors des limites du champ d'application des simulations numériques opérationnelles, il en résulte alors des erreurs de prévision. Dans ce cas, des techniques statistiques peuvent permettre d'améliorer la qualité des prévisions a posteriori. La combinaison de méthodes déterministes et statistiques permet alors de fournir des prévisions pour des phénomènes non représentés dans le modèle de prévision, ce qui correspond donc à une descente d'échelle statistique.

1.4.1 Descente d'échelle statistique

[Sheridan et al. \(2010\)](#) ont corrigé les erreurs de prévision de la température à 2 m liées à la mauvaise représentation de la topographie dans le modèle. Pour cela, ils calculent un gradient thermique vertical qui leur permet d'appliquer une correction en fonction de la différence entre l'altitude du relief dans la maille de la simulation et l'altitude réelle de la station.

[Focken et al. \(2002\)](#) ont corrigé les erreurs de prévision de vitesse du vent en utilisant un modèle de corrélation spatiale, ce qui consiste à calculer un vent en un point à partir d'observations du vent en plusieurs autres points. Ils ont ainsi corrigé les biais plus efficacement qu'en ne prenant les données que d'un seul site parce que la combinaison des sites permet un lissage spatial qui réduit les erreurs.

[Lei et al. \(2009\)](#) fournissent une vue d'ensemble des différentes méthodes permettant le calcul de vitesses de vent et les classent en quatre catégories : les modèles physiques (simulations numériques basées sur les équations de Navier-Stokes), les modèles de corrélation spatiale, les modèles statistiques et les autres méthodes dont fait partie le réseau de neurones artificiels (RNA). Une des conclusions de la comparaison de ces différentes méthodes est que les modèles basés sur les RNA produisent moins d'erreurs que les autres méthodes. Ils ont donc déjà été largement utilisés dans le domaine de la recherche météorologique.

1.4.2 Réseaux de neurones artificiels

Les RNA sont des outils mathématiques qui permettent de trouver des relations non-linéaires entre des variables. Cette faculté n'est pas spécifique aux RNA. En revanche, ils ont la particularité d'être des « approximateurs parcimonieux », c'est à dire qu'à précision égale, le nombre de paramètres ajustables nécessaires est plus faible dans le cas du RNA que dans le cas des autres approximateurs couramment utilisés (régression linéaire, régression polynomiale, ...). Les RNA combinent des fonctions dont la forme elle-même dépend de paramètres ajustables. Cela donne un nombre de degrés de liberté important qui permet de reproduire une fonction avec un nombre limité de ces paramètres. Au contraire, les autres modèles combinent des fonctions dont les seuls paramètres ajustables sont les poids de chaque fonction, d'où un nombre de degrés de liberté moins grand que dans le cas d'un RNA et donc un nombre de paramètres ajustables nécessaire plus grand à précision égale ([Dreyfus et al., 2002](#)).

Les paramètres sont ajustés au cours d'une étape dite d'apprentissage visant à obtenir des résultats les plus proches possibles d'une variable observée. La propriété parcimonieuse est alors fondamentale parce que le nombre de données utilisées pendant la phase d'apprentissage doit être d'autant plus grand que le nombre de paramètres ajustables est élevé afin que l'estimation

de ces paramètres soit la plus significative possible (Dreyfus et al., 2002). Donc, à précision égale, la taille du jeu de données utilisé pendant l'apprentissage est plus petite pour une méthode parcimonieuse comme le RNA. Une description plus approfondie de cette méthode est fournie dans la section 4.1.

Gardner and Dorling (1998) donnent un aperçu des possibles utilisations de RNA dans le domaine des sciences de l'atmosphère. Ils ont déjà été utilisés pour calculer un champ de vent interpolé à partir des mesures de quelques stations météorologiques (Philippopoulos and Deligiorgi, 2012). Les RNA peuvent également déduire une variable à partir de variables de natures différentes comme dans l'étude de Delon et al. (2007) qui ont calculé les émissions d'oxydes d'azote du sol à partir de différentes observations environnementales. Les RNA peuvent également être utilisés pour faire de la prévision à court terme (quelques heures) comme l'ont montré Cadenas and Rivera (2009) et More and Deo (2003). Ils ont majoritairement été utilisés pour déduire des variables à partir d'observations, mais il y a également quelques exemples d'utilisation dans un but de descente d'échelle à partir de simulations. Ahn et al. (2012) ont utilisé des données issues de simulations méso-échelle pour faire une descente d'échelle permettant de reproduire des températures locales sur une région marquée par un relief complexe, mais cette étude ne porte pas sur des prévisions. De même, Liu and Coulibaly (2011) ont amélioré des prévisions hydrologiques à partir de modèles globaux. D'autres études font état de l'utilisation de RNA pour faire de la descente d'échelle à partir de simulations globales dans un but de prévision à court terme (Kumar et al., 2012; Coulibaly et al., 2005).

Les RNA ont donc déjà été utilisés avec succès pour le calcul en prévision instantanée ou à court terme des vitesses de vent (Cadenas and Rivera, 2009; Salcedo-Sanz et al., 2009; Kani and Ardehali, 2011) mais il n'y a pas d'études concernant la direction du vent, probablement à cause de la difficulté supplémentaire liée au fait que c'est une variable cyclique. Ce point particulier a été exploré au cours de cette thèse et il sera présenté par la suite.

1.5 Conclusion

Les phénomènes en jeu dans la couche limite atmosphérique sont complexes, notamment quand la surface présente d'importantes hétérogénéités (relief, type de sol, couverture). Les nuits sont souvent marquées par des conditions stables à cause du refroidissement radiatif de la surface. Dans les vallées, les différences spatiales de température alimentent une circulation tridimensionnelle (à l'échelle de la vallée) qui accompagne la stratification. La complexité du terrain est alors primordiale dans l'évolution de la couche limite et dans le développement des écoulements.

Le travail de cette thèse est relatif aux écoulements aérologiques dans la vallée de Cadarache. Le site d'étude est présenté dans le chapitre suivant, ainsi que les résultats de recherches effectuées sur la région, antérieures ou préliminaires au travail réalisé pendant cette thèse. Cela permettra de choisir une orientation adaptée à l'objectif plus large qui est d'effectuer des prévisions météorologiques fiables sur Cadarache. Si le travail de thèse est appliqué au site de Cadarache, la méthodologie est néanmoins généralisable à la descente d'échelle pour n'importe quelle vallée non résolue.

Chapitre 2

Recherches préliminaires et passées sur le site d'étude

Cette partie présente les caractéristiques du site d'étude qui sont déterminantes pour l'analyse des écoulements dans une vallée (comme vu dans le chapitre 1) : hétérogénéités du sol, vents d'altitude typiques qui peuvent créer un forçage pour les vents dans la vallée. Un bilan des connaissances sur les écoulements locaux du site étudié est présenté sur la base d'observations continues et d'observations complémentaires effectuées lors d'une campagne de mesure à l'hiver 2013. Enfin, les simulations numériques effectuées quotidiennement sur la région sont présentées et leur capacité à reproduire les écoulements locaux est évaluée sur la base des observations effectuées en un seul point de la vallée. Cette partie pose donc les bases du travail effectué dans cette thèse puisqu'elle justifie les voies explorées pour l'amélioration de la prévision des vents en vue d'un calcul de dispersion atmosphérique.

Plan du chapitre

2.1	Description du site	32
2.1.1	Terrain complexe	32
2.1.2	Météorologie régionale	33
2.2	Connaissances sur la météorologie locale du site	35
2.2.1	Observations continues	35
2.2.2	KASCADE 2013	36
2.2.3	Climatologie des vents sur une période d'un an	38
2.3	Prévisions WRF	40
2.3.1	Configuration des simulations Mistrau	40
2.3.2	Représentation du terrain dans les simulations	42
2.3.3	Critères statistiques pour l'évaluation de la prévision des vents	43
2.3.4	Comparaison aux observations	44
2.3.5	Simulations à 1 km de résolution horizontale	54
2.4	Prévision immédiate du vent local par méthode statistique	54
2.5	Conclusion	55

2.1 Description du site

2.1.1 Terrain complexe

Le site étudié englobe le centre CEA de Cadarache situé dans le sud-est de la France, à environ 50 km au nord de la Méditerranée dans une région pré-alpine caractérisée par un relief complexe de vallées et collines. Des éléments de reliefs notables sont situés entre 20 et 50 km autour de Cadarache. À 50 km à l'ouest, la vallée du Rhône orientée du nord vers le sud, à 20 km le Luberon et la Sainte Victoire sont deux massifs orientés ouest-est qui culminent respectivement à 1125 m et 1101 m et encadrent le site, le Luberon au nord-ouest et la Sainte Victoire au sud. Enfin, à 20 km au nord-est, il y a le plateau de Valensole, plateau en pente douce descendant en direction du sud-ouest à une altitude d'environ 600 m. Plus proche du site, à 7 km, la montagne de Vautubière culmine à 658 m (voir la figure 2.1 pour une vue de la topographie de la région).

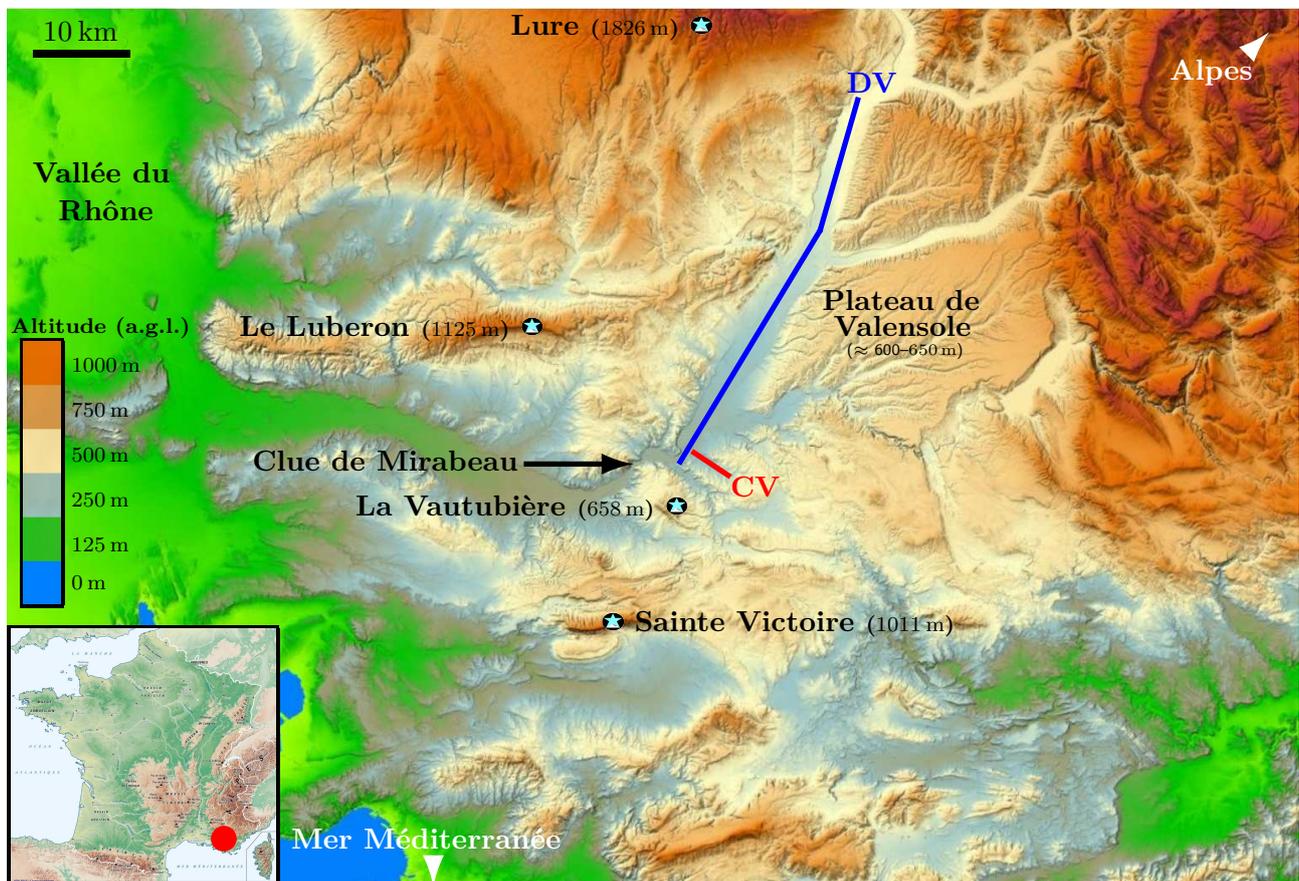


FIGURE 2.1 : Carte du relief de la région autour de la zone d'étude. CV représente la vallée de Cadarache et DV la vallée de Durance. Source : Institut national de l'information géographique et forestière (IGN) : <https://www.geoportail.gouv.fr/>.

Cadarache est situé à la confluence de deux vallées (voir la figure 2.2.(a) pour une vue de la topographie locale). La première est la moyenne vallée de la Durance (notée DV), orientée du nord-nord-est vers le sud-sud-ouest, longue de 67 km, large d'environ 5 km et profonde d'environ 200 m (son fond est situé à une altitude de 250 à 350 m dans sa partie médiane, représentée en bleu sur la figure 2.1). La seconde est la vallée de Cadarache (CV, en rouge sur les figures 2.1 et 2.2.a), orientée du sud-est vers le nord-ouest, longue de 5 km, large d'environ 1 km et profonde de 100 m (son fond est situé à une altitude comprise entre 250 et 300 m et les crêtes alentours

culminent à environ 400 m). La vallée de Cadarache présente une géométrie complexe, et ses versants sont entaillés d'une multitude de petites vallées.

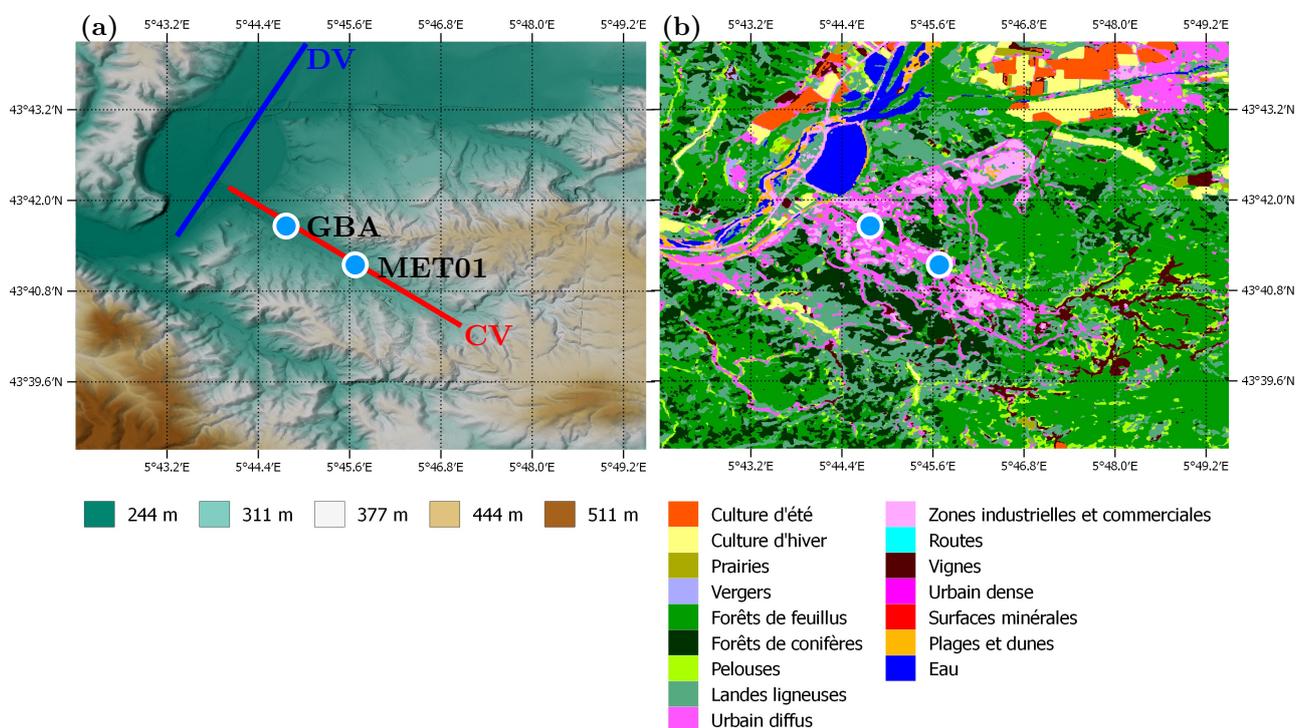


FIGURE 2.2 : Carte de la zone autour de Cadarache montrant la topographie (a) et l'occupation du sol (b) issue de la base de données fournie par le pôle Theia : <https://www.theia-land.fr/>. Les points bleus matérialisent les positions des deux stations permanentes du CEA (appelées GBA et MET01), dont l'instrumentation est détaillée dans le tableau 2.1.

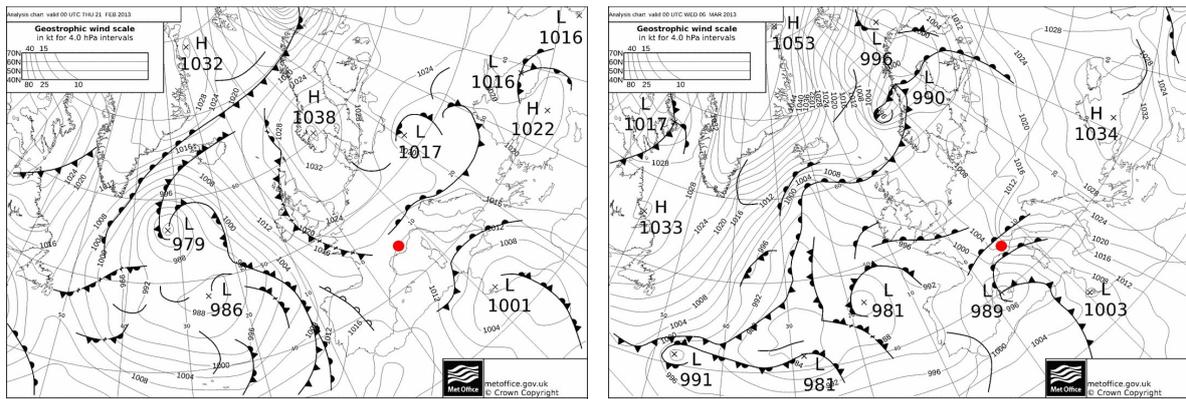
La figure 2.2.b montre l'hétérogénéité d'occupation du sol existant sur le site et ses environs. Le fond de la vallée de Cadarache est occupé par de nombreux bâtiments dispersés ainsi que par quelques prairies. Les flancs et les crêtes présentent majoritairement des surfaces arborées, avec une densité variable. La vallée de la Durance présente beaucoup moins de surfaces artificielles. La Durance y alimente un bassin d'une surface de 1,3 km² situé quelques centaines de mètres au nord de la confluence des vallées de Durance et Cadarache. Le fond de la vallée de Durance est occupé également par des terrains agricoles cultivés et quelques forêts de feuillus plus ou moins denses. Enfin les alentours des deux vallées sont majoritairement couverts de forêts et de quelques surfaces cultivées.

2.1.2 Météorologie régionale

Les vents typiques rencontrés sur la région sont la brise de mer (Cros et al., 2004; Bastin et al., 2005), les vents d'ouest, les vents de sud-est associés au « mauvais temps » (pluie ou nuages) et le Mistral.

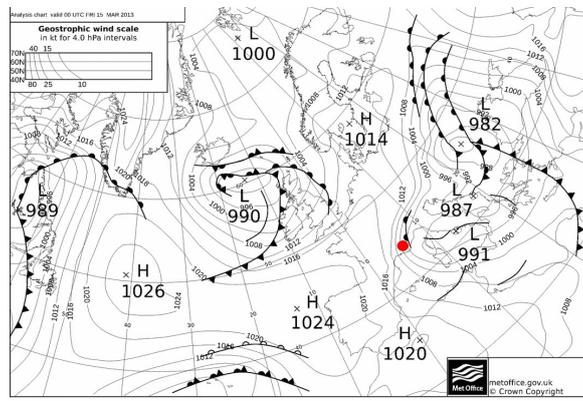
Le Mistral est un fort vent régional résultant de la présence d'un anticyclone à l'ouest-sud-ouest de la France et d'une dépression aux alentours du golfe de Gênes (cf. figure 2.3.(c) pour la situation du 15 mars 2013) produisant à l'échelle synoptique un vent venant du nord sur le sud-est de la France. Il est alors canalisé près de la surface par la vallée de Rhône, entre les Alpes à l'est et le Massif Central à l'ouest provoquant une augmentation des vitesses. Il

est le plus souvent associé à un ciel dégagé et prend une direction nord-ouest sur Cadarache du fait d'un étalement des directions au débouché de la vallée du Rhône et à l'approche de la Méditerranée. Une situation typique du Mistral sur Cadarache est décrite sur la figure 2.4 (colonne de droite, graphiques c, f et i) pour la période du 14 au 16 mars 2013. Le ciel dégagé produit des valeurs élevées de rayonnement solaire incident (SW_{\downarrow}) pendant le jour et un bilan radiatif fortement déficitaire pendant la nuit. Malgré cela, les fortes vitesses de vent limitent la stratification. Les vents près de la surface sont alors forcés par les vents forts d'altitude et prennent donc globalement une même direction nord-ouest (qui est aussi la direction montante de la vallée).



(a) Forçage faible : 21/02/2013 00h00 UTC

(b) Mauvais temps : 06/03/2013 00h00 UTC



(c) Mistral : 15/03/2013 00h00 UTC

FIGURE 2.3 : Cartes montrant la situation synoptique sur l'Europe et l'Atlantique nord. Les courbes isobares au niveau de la mer sont représentées en gris et les centres des dépressions et anticyclones sont représentés par les lettres L et H. Les fronts sont également représentés en suivant la représentation météorologique classique. Les points rouges matérialisent la position de la vallée de Cadarache. Les cartes sont produites par le Met Office (<https://www.metoffice.gov.uk/>).

Les conditions de mauvais temps sont généralement apportées par les vents de sud-est (Berthou et al., 2016), notamment pendant le printemps et l'automne. Ce cas résulte de la présence d'une dépression au-dessus de la Méditerranée ouest, au niveau des îles Baléares (cf. figure 2.3.(b)). Dans ce cas, qui est également illustré sur la figure 2.4 (colonne du centre, graphiques b, e et h) pour la période du 5 au 7 mars 2013, les forts vents de sud-est en altitude forcent les directions des vents près du sol qui prennent alors la même direction, alignée avec la direction descendante de la vallée.

Le troisième cas typique correspond à des conditions pendant lesquelles le forçage synoptique est faible (pas d'influence marquée de dépression, cf. 2.3.(a) : vents d'altitude faibles et ciel clair), comme pour la période du 20 au 22 février 2013 présentée sur la figure 2.4 (colonne de gauche, graphiques a, d et g). Le refroidissement radiatif nocturne crée alors une forte stratification stable et favorise la production de vents d'origine thermique descendant les vallées de Cadarache et de la Durance. Ces conditions ont fait l'objet d'observations pendant la campagne de mesures KASCADE 2013 qui s'est déroulée pendant l'hiver 2013 (voir partie 2.2.2).

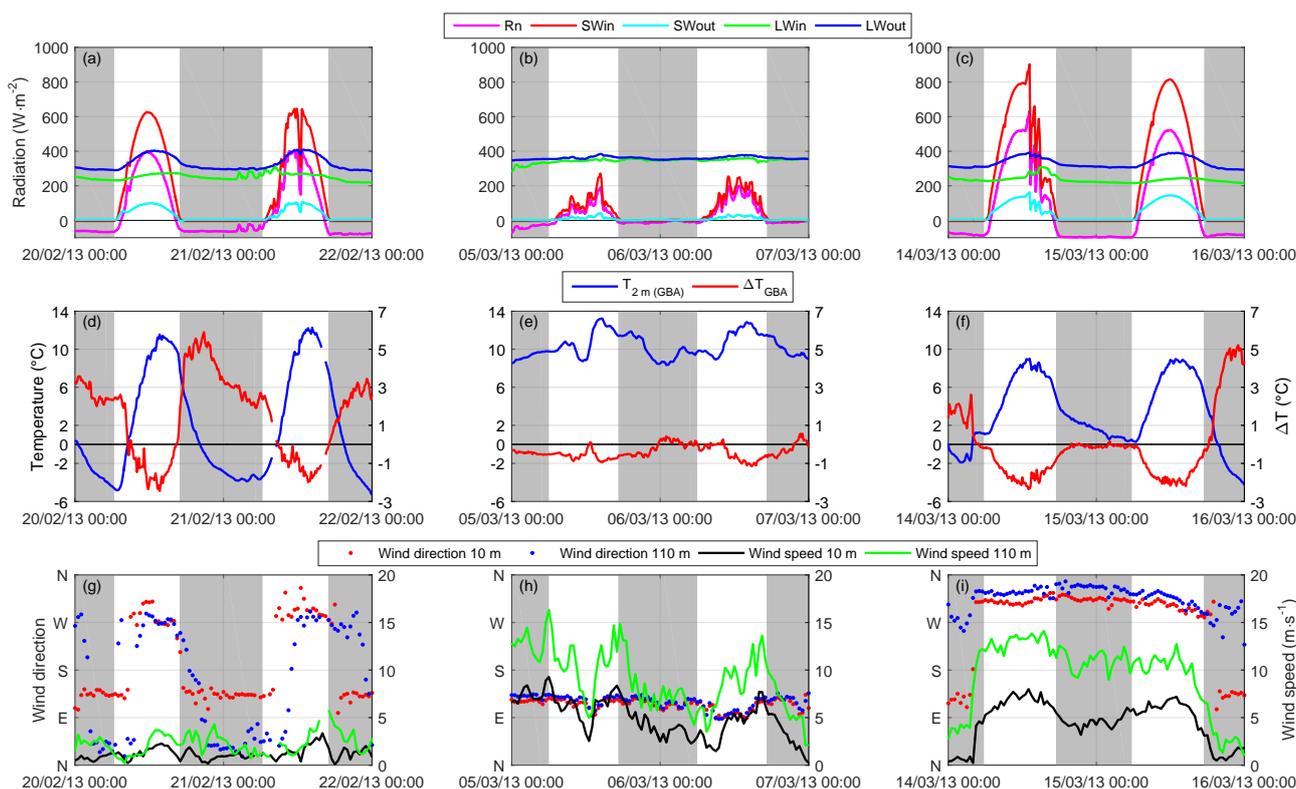


FIGURE 2.4 : Conditions météorologiques observées dans la vallée de Cadarache pendant 3 périodes de 2 jours consécutifs présentant des conditions météorologiques typiques de la région : beau temps et vent faible (colonne de gauche), nuageux ou pluvieux (au centre) et Mistral (à droite). Les figures (a), (b) et (c) montrent les composantes incidentes (in) et ascendantes (out) des rayonnements de courtes (SW) et grandes longueurs d'ondes (LW) ainsi que le rayonnement net Rn . Les figures (d), (e) et (f) montrent la température à 2 m et la différence de température entre 110 m et 2 m mesurées à la station GBA (les stations sont décrites ci-dessous). Les figures (g), (h) et (i) montrent les vitesses et directions des vents mesurés à 110 m (station GBA) et à 10 m (station MET01) dans la vallée de Cadarache. Les zones grisées matérialisent les périodes nocturnes.

2.2 Connaissances sur la météorologie locale du site

2.2.1 Observations continues

La vallée de Cadarache est équipée de deux stations météorologiques permanentes localisées sur les sites GBA et MET01 (figure 2.2). Ces deux stations sont situées dans le fond de la vallée, environ à mi-longueur de la vallée pour la station MET01 et à son extrémité aval, près de la jonction entre les vallées de Cadarache et de Durance, pour la station GBA.

La station GBA est équipée d'un mât de 110 m de hauteur qui permet entre autres de mesurer le vent à 110 m, c'est à dire juste au-dessus des crêtes de la vallée, permettant ainsi une observation continue de vents caractéristiques d'une échelle spatiale supérieure à celle de la vallée elle-même. Cette station permet également de mesurer la stratification dans la vallée grâce à des mesures de température à 110 m et 2 m. La station MET01 permet quant à elle d'observer le vent au fond de la vallée. Un récapitulatif des observations effectuées sur ces deux stations est donné dans le tableau 2.1.

Station	Hauteur	Paramètres mesurés	Date	Instrument
MET01	2 m	SW_{in}	Depuis le 17/02/2015	Pyranomètre Hukseflux LP02
		P	Depuis le 17/02/2015	Baromètre Setra CS100
		T et Rh	Depuis le 17/02/2015	Thermo-hygromètre Campbell Sci. HC2S3
		Pluie	Depuis le 17/02/2015	Pluviomètre EML ARG100
	10 m	WS, WD	17/02/2015 – 17/02/2016	Anémomètre à hélice Young 05103
MET01	2, 10 et 30 m	u, v, w	13/12/2012 – 16/03/2013	Anémomètres soniques Young 81000 et Campbell CSAT3
	GBA	T	En continu	Thermomètre Rotronic PT100
Rh		En continu	Hygromètre Rotonic	
P		En continu	Baromètre Vaisälä PTB1001C	
WS, WD		En continu	Anémomètre sonique Metek	
GBA	110 m	T	En continu	Thermomètre Rotronic PT100

Tableau 2.1 : Récapitulatif des mesures effectuées aux stations GBA et MET01. Rh pour l'humidité relative et WS et WD pour la vitesse et la direction du vent.

2.2.2 KASCADE 2013

La campagne de mesures KASCADE (KAtabatic winds and Stability over CAdarache for Dispersion of Effluents) 2013 avait pour objectif d'observer les vents dans les vallées de Cadarache et de la Durance, de décrire leur structure verticale ainsi que leur occurrence par conditions stables. Bien que ces conditions soient présentes tout au long de l'année, les longues nuits d'hiver sont plus propices à leur observation. Les mesures continues ont ainsi été effectuées durant l'hiver 2013, entre le 13 décembre 2012 et le 18 mars 2013. L'instrumentation permanente du CEA (station GBA) a été complétée par un mât de 30 m (positionné sur le site MET01, avec entre autres des mesures de vent à 2, 10 et 30 m) et un sodar profileur de vent (positionné dans la vallée de la Durance) sur cette période. Ces mesures continues ont été complétées par 23 périodes d'observation intensive (POI) qui ont permis d'améliorer l'observation sur la verticale, durant certaines nuits propices à la formation de conditions stables, en effectuant des radiosondages troposphériques ainsi que des mesures entre la surface et environ 400 m à l'aide d'un ballon captif à l'emplacement de MET01. Un récapitulatif des mesures effectuées pendant cette expérience peut être trouvé dans [Duine et al. \(2017\)](#) et [Duine \(2015\)](#).

[Duine et al. \(2017\)](#) ont représenté l'évolution diurne du profil vertical de vent typiquement rencontré pendant les POI (figure 2.5). En journée, les vents viennent majoritairement de l'ouest au nord-ouest. Après le coucher du soleil, quand le forçage synoptique est suffisamment faible (ciel clair et vents faibles en altitude), des conditions stables s'installent et un vent descendant

la vallée de Cadarache apparaît (la direction des vents descendant la vallée de Cadarache est notée CDV pour Cadarache down-valley). Son extension verticale peut varier avec les conditions de vents au-dessus de la vallée mais atteint généralement une hauteur d'environ 50 m avec un maximum de vitesse (1 à 4 m s^{-1}) vers 25–30 m. La situation redevient rapidement instable après le lever du soleil et les vents d'ouest réapparaissent à leur tour environ une à deux heures après le lever du soleil.

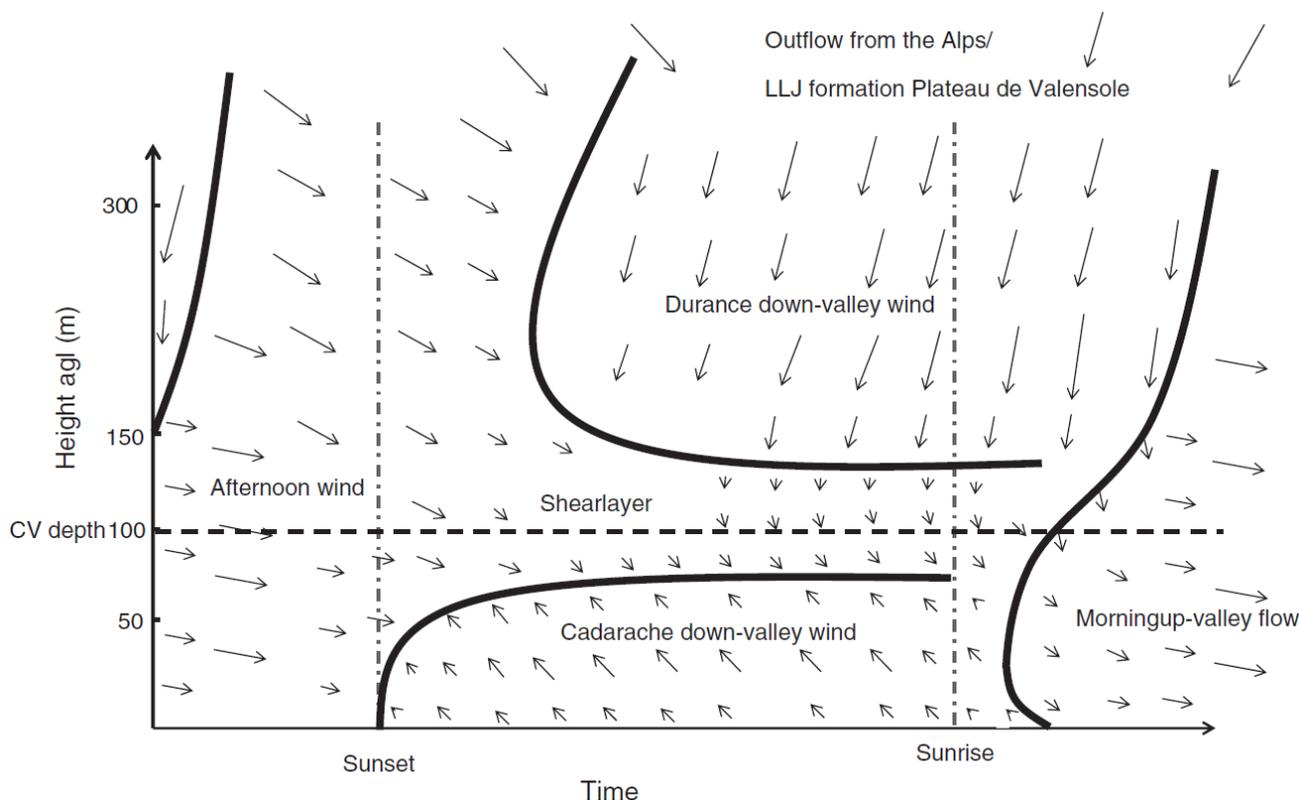


FIGURE 2.5 : Schématisation de l'évolution diurne du profil vertical des directions du vent typiques des conditions observées pendant la campagne KASCADE 2013 (ciel clair et faible forçage synoptique), d'après Duine et al. (2017). Les flèches représentent le vent horizontal à une altitude et un instant de la journée donnés.

Au-dessus du vent de vallée de Cadarache, un autre vent est observé entre 100 m et 500 m (l'extension verticale est très variable) avec une direction nord-nord-est et un maximum de vitesse (4 à 8 m s^{-1}) atteint à une hauteur de 175 à 225 m. Il correspond à un drainage dans la vallée de la Durance de courants prenant naissance sur les pentes des massifs et plateaux encadrant la Durance (Lure, Valensole, ...). La formation de ce vent implique donc des phénomènes qui se produisent à une échelle spatiale régionale, bien plus étendue que ceux générant le vent de vallée de Cadarache, ce qui explique qu'il n'apparaisse au-dessus de Cadarache qu'environ 6 heures après le coucher du soleil et qu'il persiste environ 3 heures après le lever du soleil, avant de laisser place aux vents d'ouest. Plus de détails concernant les écoulements dans la vallée de Durance ont été apportés par des simulations numériques (voir la partie 2.3.5). Enfin, il existe une couche de cisaillement entre les deux vents de vallée dans laquelle la rotation des vents se fait dans le sens horaire en montant (« veering »).

2.2.3 Climatologie des vents sur une période d'un an

Une climatologie des vents dans et au-dessus de la vallée de Cadarache est réalisée sur la base d'un an d'observations aux stations GBA et MET01, comme détaillé dans le tableau 2.1.

La rose des vents établie à 110 m durant cette période (figure 2.6) présente trois directions majoritaires : les vents d'ouest à nord-ouest correspondent à un mélange de Mistral et de vents d'ouest typiquement rencontrés en journée, les vents de nord-nord-est alignés avec la vallée de la Durance et les vents de sud-est souvent associés au « mauvais temps ». Les vents à 2 m sont clairement impactés par la vallée de Cadarache puisqu'une majorité des vents observés sont alignés avec la vallée. La rose des vents à 2 m permet de définir trois classes de direction de vents : une première pour les vents descendant la vallée (zone bleue, entre 105 et 165°), une deuxième pour les vents remontant la vallée (zone verte, de 245 à 335°) et une troisième pour les vents transverses à la vallée (zone rouge, pour toutes les autres directions).

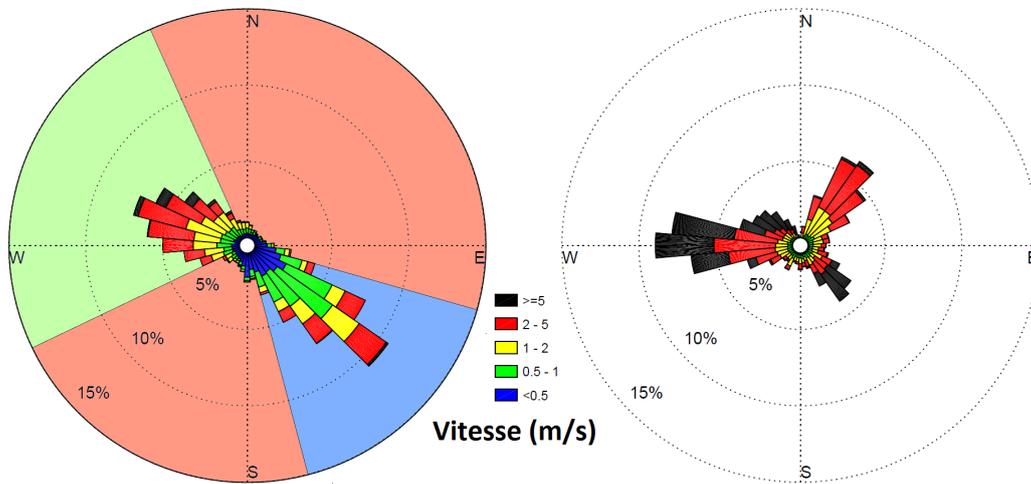


FIGURE 2.6 : Roses des vents établies sur une année (du 17 février 2015 au 17 février 2016) à partir des observations à 2 m à la station MET01 (à gauche) et à 110 m à la station GBA (à droite). À 2 m, la zone bleue matérialise les vents descendant la vallée de Cadarache (entre 105 et 165°), la zone verte matérialise les vents qui remontent la vallée (entre 245 et 335°) et la zone rouge matérialise les directions de vent transverses à la vallée.

À 2 m, les proportions de vents montant (direction CUV pour Cadarache up-valley) et descendant la vallée sont équivalentes avec 39% de vents observés dans chacune des deux classes pour à peine 22% de vents transverses à la vallée. De plus, près de la moitié des vents transverses ont des vitesses inférieures à $0,5 \text{ m s}^{-1}$ (tableau 2.2).

	$< 0,5 \text{ m s}^{-1}$	$> 0,5 \text{ m s}^{-1}$	Total
CDV	10,63 %	27,96 %	38,59 %
CUV	3,47 %	35,67 %	39,14 %
Autres	11,93 %	10,34 %	22,27 %
Total	26,03 %	73,97 %	100 %

Tableau 2.2 : Proportion de vents à 2 m montant, descendant ou transverses à la vallée de Cadarache en fonction des trois secteurs définis sur la figure 2.6. « CDV » pour les vents descendant la vallée, « CUV » pour les vents montant la vallée et « Autres » pour toutes les autres directions.

Le lien entre les vents à 2 m et à 110 m est représenté sur la figure 2.7. Il apparaît que les vents aux deux hauteurs ne sont pas toujours couplés. Les interactions entre les vents dans et au-dessus d'une vallée peuvent alors être décrits d'après [Whiteman and Doran \(1993\)](#) (voir le chapitre 1). Dans le cas des vents dans la vallée de Cadarache, trois phénomènes expliquent la génération des vents à 2 m :

- un forçage direct par les vents d'altitude générant des vents ayant des directions identiques dans et au-dessus de la vallée ;
- une canalisation des vents par la vallée, avec une déviation des vents d'altitude de sorte que les vents dans la vallée soient alignés avec son axe ;
- la génération de vents d'origine thermique qui montent ou descendent la vallée provoquant un découplage des vents dans et au-dessus de la vallée.

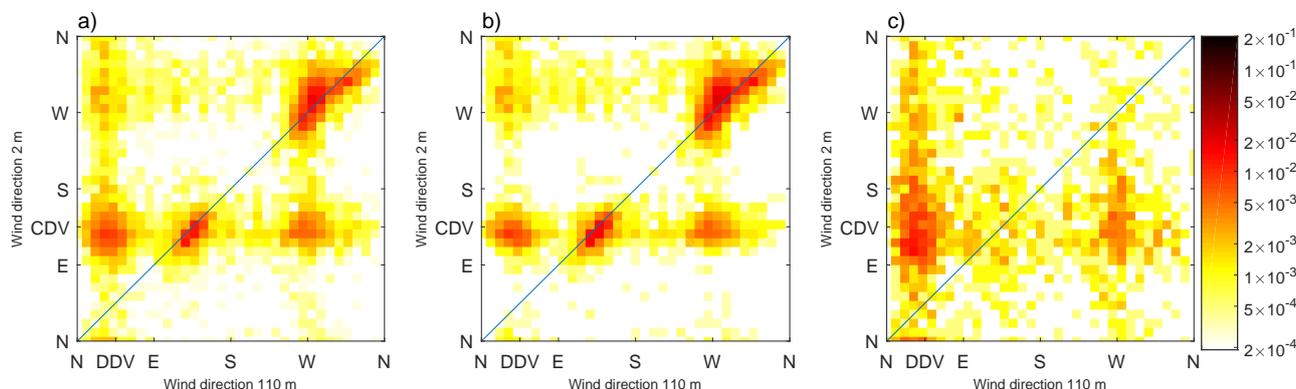


FIGURE 2.7 : Comparaison des directions à 2 m et à 110 m sur un an d'observations représentée en densité de probabilité. (a) pour tout le jeu de données, (b) pour les vents (à 2 m) supérieurs à $0,5 \text{ m s}^{-1}$ (74 % du jeu de données total) et (c) pour les vents inférieurs à $0,5 \text{ m s}^{-1}$ (26 % des données).

La première situation n'est possible qu'avec des vents en altitude suffisamment forts pour forcer les directions près de la surface. On rencontre ce phénomène avec les vents à 110 m qui viennent du sud-est qui sont très majoritairement associés à des vents de sud-est à 2 m, ou encore avec les vents de nord-ouest et une partie des vents d'ouest.

La canalisation des vents par la vallée (deuxième situation) intervient avec une partie des vents d'altitude venant d'ouest qui sont associés à des vents de nord-ouest à 2 m.

Enfin, la troisième situation intervient quand les vents en altitude ne sont pas suffisamment forts pour imprimer leur direction sur les vents près de la surface. On retrouve majoritairement cette situation par conditions stables, avec des vents qui descendent les vallées de Cadarache (visible à 2 m) et de Durance (visible à 110 m). En début de nuit, quand la stratification commence à s'établir, le vent de vallée de Cadarache (sud-est) apparaît ([Duine et al., 2017](#)) alors que le vent à 110 m conserve une direction ouest-nord-ouest. Quelques heures après le coucher du soleil, un vent descendant la vallée de Durance (nord-est) apparaît au niveau de Cadarache et est visible à 110 m. Au lever du soleil, la stratification dans la vallée de Cadarache est rapidement détruite et le vent de vallée de Cadarache cesse alors que le vent de vallée de Durance persiste pendant quelques heures. Le vent dans la vallée de Cadarache garde alors une origine thermique, mais avec une direction nord-ouest (remontant la vallée) jusqu'à ce que le vent de vallée de Durance soit détruit et que le vent d'ouest-nord-ouest revienne à 110 m.

L'objectif du CEA étant le calcul du transport et de la dispersion atmosphérique, la connaissance des caractéristiques des vents rencontrés sur la région de Cadarache est impor-

tante, notamment près du sol. Des simulations numériques sont ainsi effectuées avec l'objectif de reproduire ces vents. Ces simulations et leur évaluation sont présentées dans la partie suivante.

2.3 Prévisions WRF

Le code WRF (Weather Research and Forecast, [Skamarock et al. \(2008\)](#)) est utilisé pour fournir des prévisions météorologiques en mode opérationnel avec l'objectif d'utiliser les champs de vent produits sur la région pour contraindre des outils de calcul de la dispersion atmosphérique des rejets potentiellement émis sur le site. La plateforme numérique produisant ces prévisions a été baptisée « Mistrau ».

WRF a été développé principalement par le National Center for Atmospheric Research (NCAR). Il est utilisé aussi bien pour la prévision opérationnelle que pour la recherche en météorologie et climatologie. Il résout les équations non-hydrostatiques et compressibles ce qui lui permet de simuler les phénomènes convectifs.

La gamme d'applications de WRF est très large, de l'échelle synoptique à la micro-échelle avec la possibilité d'effectuer des LES, des simulations de cas idéalisés, la possibilité de faire de la descente d'échelle en imbriquant différents domaines aux résolutions croissantes ou encore l'utilisation en complément de modules dédiés tels que WRF-Chem ([Grell et al., 2005](#)) pour la prise en compte des effets de la chimie atmosphérique ou FLEXPART ([Brioude et al., 2013](#)) pour le calcul de la dispersion atmosphérique. C'est d'ailleurs dans l'optique de la conduite de calculs FLEXPART que les simulations « Mistrau » sont effectuées.

2.3.1 Configuration des simulations Mistrau

Les prévisions « Mistrau » sont effectuées quotidiennement au LMTE, à l'aide de WRF, de manière opérationnelle depuis la mi-février 2015 (prévisions à +108 h). Elles sont initialisées chaque jour par les analyses GFS (Global Forecast System, [Center \(2003\)](#)) à 0,25° de résolution horizontale. Elles comprennent deux domaines imbriqués centrés sur Cadarache. Le premier de 9 km de résolution horizontale, englobe la France (forcé par les prévisions GFS aux bords du domaine durant toute la simulation). Le deuxième domaine, forcé sur sa bordure par le premier, de 3 km de résolution horizontale, englobe la région Provence-Alpes-Côte d'Azur (figure 2.8).

Les choix de paramétrisations physiques dans les simulations WRF ont fait l'objet de tests de sensibilité lors de simulations visant à reproduire les conditions observées sur certaines POIs de la campagne KASCADE 2013 ([Duine, 2015](#); [Kalverla et al., 2016](#)). Ces simulations comprenaient un domaine supplémentaire, à 1 km de résolution horizontale, imbriqué dans le domaine à 3 km de résolution. Cinq schémas de couche limite planétaire ont été testés (YSU ([Hong et al., 2006](#)), MYJ ([Janjić, 1994](#)), QNSE ([Sukoriansky et al., 2005](#)), MYNN2.5 ([Nakanishi and Niino, 2006](#)) et ACM2 ([Pleim, 2007](#))) sans que le modèle ne montre de sensibilité marquée à ces changements. L'utilisation du schéma de surface terrestre NOAH a permis de mieux représenter les températures de surface ainsi que l'amplitude thermique diurne. Finalement, la combinaison des schémas NOAH pour la paramétrisation de la surface terrestre et QNSE pour la couche limite planétaire a été jugée la plus performante. Le schéma de rayonnement choisi pour les grandes longueurs d'ondes est RRTM. Finalement, le détail des choix de paramétrisations est donné dans le tableau 2.3.

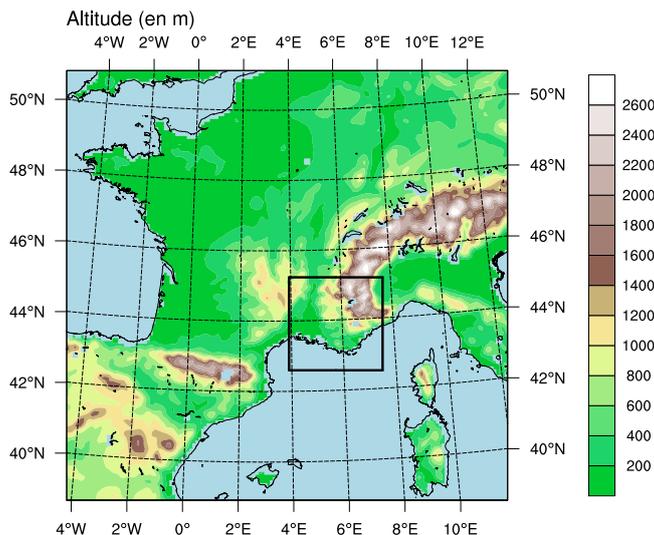


FIGURE 2.8 : Représentation de la topographie à 9 km de résolution horizontale dans les simulations WRF. Le domaine père (résolution horizontale de 9 km) couvre la totalité de la figure alors que le domaine fils (à 3 km de résolution horizontale) est matérialisé par le carré noir situé sur le sud-est de la France.

Forçage météo	prévisions horaires GFS 0,25°
Occupation du sol	CORINE land cover (https://land.copernicus.eu/)
Orographie	SRTM (Farr et al., 2007)
Temps de retour de simulation	10 h
Résolutions horizontales	9 km × 9 km et 3 km × 3 km
Communication entre les domaines	couplage dans les deux sens
Niveaux verticaux	35 niveaux
Rayonnement LW	RRTM (Mlawer et al., 1997)
Rayonnement SW	Goddard (Chou and Suarez, 1994)
Microphysique	WSM6 (Hong and Lim, 2006)
Surface	Noah (Niu et al., 2011)
Couche limite	QNSE (Sukoriansky et al., 2005)
Couche de surface	QNSE (Sukoriansky et al., 2005)

Tableau 2.3 : Récapitulatif des choix de paramétrisation des simulations « Mistrau ».

Après avoir estimé un temps de spin-up¹ à 24 h, les données exploitées des simulations sont les sorties horaires de +24 h à +47 h de chaque prévision, ce qui permet de produire une série horaire continue, sans chevauchement (figure 2.9) depuis le 17 février 2015 12 UTC jusqu'au 17 février 2016 12 UTC (période pendant laquelle les données des simulations et les mesures de vent à 2 m sont disponibles simultanément). Les résultats ont été interpolés (interpolation bilinéaire) vers le point de coordonnées (43,685°N ; 5,760°E) qui correspond à la station MET01 (voir figure 2.2 pour la position). Enfin, les variables 3D ont été calculées à des niveaux altimétriques (10 m, 20 m, 30 m, 50 m, 100 m, 110 m, 150 m, 200 m, 300 m, 400 m, 500 m au-dessus du sol) à partir de la coordonnée en pression normalisée sigma (15 premiers niveaux : 1,0000, 0,9974, 0,9940, 0,9900, 0,9854, 0,9796, 0,9723, 0,9635, 0,9528, 0,9401, 0,9252, 0,9079, 0,8882, 0,8659, 0,8410, avec une valeur de 0 pour une pression de 50 hPa) de discrétisation verticale utilisée par WRF. Les performances de ces prévisions sont évaluées par comparaison avec plu-

1. Temps de stabilisation de la simulation

sieurs mesures effectuées en continu sur le site de Cadarache.



FIGURE 2.9 : Schématisation des périodes simulées pour trois prévisions successives. La partie rouge correspond au spin-up, la partie verte correspond à la période de 24 heures sélectionnée pour l'évaluation des simulations et la partie bleue correspond au reste de la simulation. Seule la partie verte est utilisée (de +24 à +47h).

2.3.2 Représentation du terrain dans les simulations

Dans les simulations WRF, la résolution du terrain dépend directement de la résolution horizontale à laquelle les variables météorologiques sont calculées. Les figures 2.10 et 2.11 montrent une comparaison des caractéristiques du terrain (topographie et occupation du sol) entre leur représentation dans les simulations et une représentation avec une meilleure résolution.

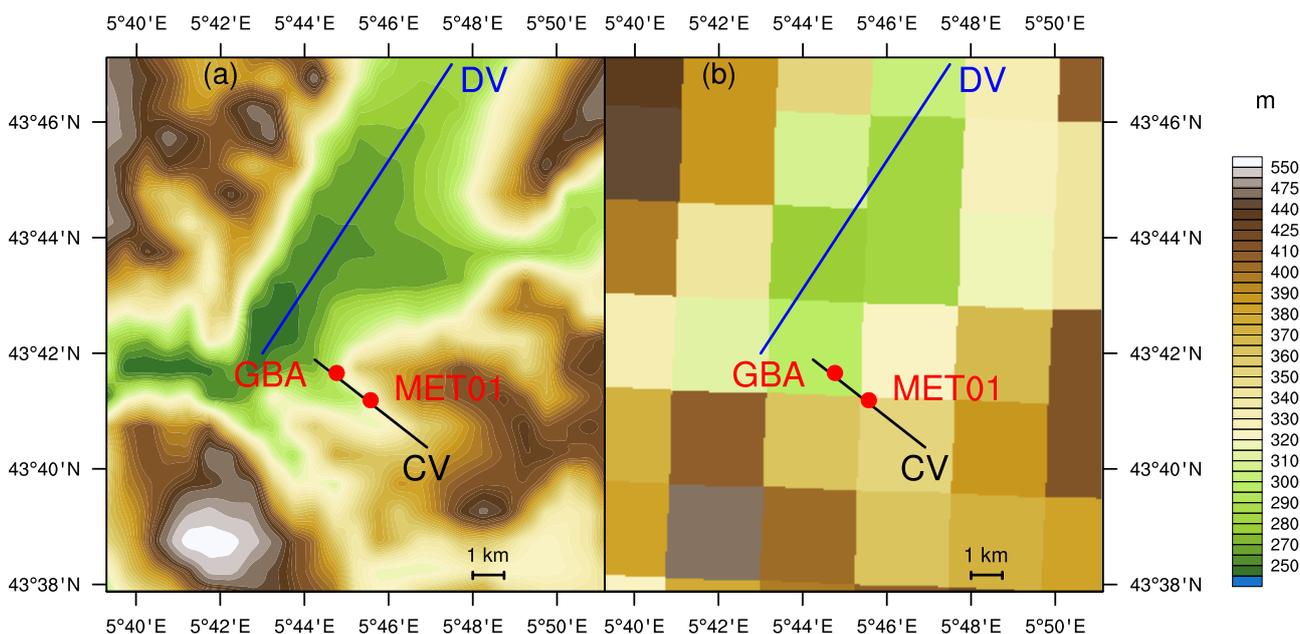


FIGURE 2.10 : Représentation de la topographie avec une résolution horizontale de 110 m (a) et 3 km (b) autour de Cadarache.

À 3 km de résolution horizontale, la description des caractéristiques du terrain est très grossière (figure 2.10). Le relief dans la vallée de la Durance, qui fait environ 5 km de large et 200 m de profondeur, est partiellement résolu. En revanche, la vallée de Cadarache, qui est plus petite (environ 1 km de large et 100 m de profondeur), n'est pas du tout représentée dans WRF.

L'occupation du sol est également représentée de manière très grossière dans WRF (figure 2.11). Les données d'entrée sont issues de la base de données CORINE land cover (résolution maximale de 25 ha avec une largeur minimale de 100 m, <https://land.copernicus.eu/pan-european/corine-land-cover/>). La vallée de Cadarache est caractérisée par la présence de bâtiments dans le fond de la vallée et de végétation plus ou moins dense sur les flancs de la vallée,

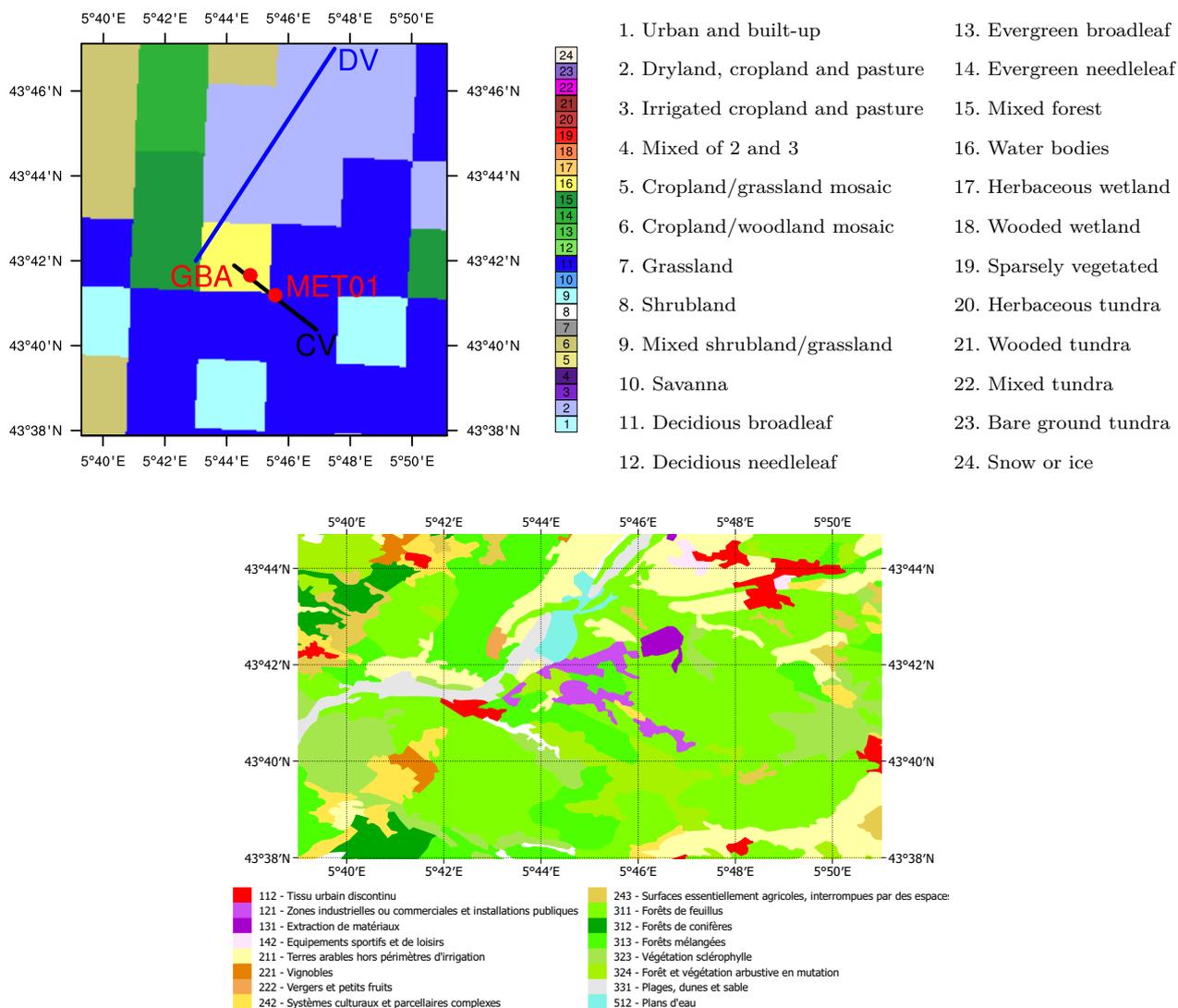


FIGURE 2.11 : Représentation de l'occupation du sol dans les simulations WRF à 3 km de résolution horizontale (en haut), les catégories Corine land cover ayant été réduites aux catégories USGS, et dans la base de données CORINE land cover (en bas) qui est utilisée pour le calcul des classes dans WRF.

ce qui est bien reproduit dans la base de données CORINE land cover. Dans les simulations, la vallée de Cadarache n'est représentée que par 4 mailles, dont 3 affichent la présence d'arbres à feuilles caduques, et une montre un plan d'eau (un bassin d'écluse de 1,3 km² dans la vallée de la Durance). Les zones artificielles présentes dans le fond de la vallée ne sont donc pas représentées dans les simulations. La vallée de la Durance n'est représentée que par des pavés de terres cultivées et de pâturages.

2.3.3 Critères statistiques pour l'évaluation de la prévision des vents

L'évaluation de la représentation des vents dans les simulations WRF se fait par comparaison avec les observations. Concernant les directions du vent, trois indicateurs sont utilisés : le *DACC* (Directional ACCuracy, Santos-Alamillos et al. (2013)) et le *PC* (Proportion Correct) à 2 et 4 classes. Le *DACC* est la fraction de directions de vent horizontal s'écartant de moins de 45° de la direction observée. Le *PC* a été utilisé dans l'étude de Duine et al. (2016). Il représente la fraction de valeurs bien classées en considérant différents secteurs de directions ou

de vitesses de vent. Avec deux secteurs, comme dans l'étude de [Duine et al. \(2016\)](#), l'objectif est de détecter les vents descendant la vallée de Cadarache. Un premier secteur comprend les vents venant du quart sud-est alors que le deuxième regroupe toutes les autres directions. Ce critère est approprié pour l'évaluation de la représentation des vents descendant la vallée de Cadarache. Le *PC4* distingue 4 classes de vents : une première classe pour les vents faibles (inférieurs à $0,5 \text{ m s}^{-1}$) pour lesquels les directions ne sont pas bien définies ([Jarraud, 2008](#)), puis trois catégories de directions choisies en fonction des observations du vent (secteurs de couleur sur la figure 2.6) pour évaluer la représentation de situations météorologiques dominantes. On définit ainsi une classe qui comporte les vents descendant la vallée de Cadarache, une classe pour les vents qui remontent la vallée et la troisième classe pour toutes les autres directions de vent.

L'évaluation de la représentation de la vitesse du vent se fait à partir du biais, de l'erreur absolue moyenne et d'un coefficient de corrélation. Les coefficients de corrélation sont également calculés pour l'évaluation des composantes u et v du vent.

2.3.4 Comparaison aux observations

L'objectif principal des simulations pour le CEA est de produire un champ de vent qui soit le plus réaliste possible, c'est pourquoi une partie importante de notre évaluation des simulations est centrée sur la représentation des vents (à 2 m et 110 m puisque ce sont les deux niveaux auxquels des mesures à long terme sont disponibles). Une autre partie de cette évaluation est dédiée à la représentation de la stabilité puisqu'une grande partie des vents dans la vallée de Cadarache est d'origine thermique.

2.3.4.a Calcul des vents WRF à 2 m

Le niveau le plus bas auquel est calculé le vent dans les simulations WRF est à 10 m. Or les observations du vent dans la vallée de Cadarache étaient auparavant faites à 2 m (elles sont depuis peu effectuées à 10 m). Afin de pouvoir comparer les vents prévus et observés près du sol, les vents des simulations sont extrapolés de 10 m à 2 m.

Un moyen permettant ce calcul consiste à utiliser la loi logarithmique définissant le profil de vent dans la couche de surface atmosphérique pour les conditions de stratification proches de la neutralité (voir chapitre 1). Cette méthode nécessite la connaissance de la longueur de rugosité mécanique z_0 .

Dans les simulations WRF, z_0 est déterminée par le schéma de surface et dépend de l'occupation du sol. La station MET01, où sont effectuées les mesures de vent à 2 m est située dans une maille caractérisée par la présence d'arbres à feuilles caduques (figure 2.11) pour lesquels la longueur de rugosité est fixée à $z_0 = 0,50 \text{ m}$.

Il est également possible de calculer la longueur de rugosité réelle correspondant au site à partir des données de la campagne de mesure KASCADE 2013 pendant laquelle 3 anémomètres soniques étaient placés sur le site MET01 (à 2, 10 et 30 m de hauteur). En effet, les anémomètres soniques permettent de calculer à la fois le vent moyen et la vitesse de frottement u_* à partir des fluctuations de vitesse horizontale et verticale (cf. équation 1.29). Pour la campagne KASCADE 2013, ces données sont disponibles toutes les 30 minutes. La longueur de rugosité

est alors calculée directement avec l'équation 1.28 en sélectionnant les conditions neutres.

Les conditions sont considérées comme étant neutres lorsque les différences de température potentielle entre 110 m et 2 m mesurée à la station GBA sont comprises entre -1 K et +1 K. Les valeurs médianes de z_0 ainsi obtenues valent 0,04 m pour les mesures à 2 m, 0,56 m pour celles à 10 m et 2,96 m pour celles à 30 m.

La loi logarithmique du profil de vent est caractérisée par une longueur de rugosité constante, donc indépendante de l'altitude. Or, la valeur z_0 caractérise la rugosité d'une aire impactant la mesure (le « footprint »). Cette surface dépend de plusieurs paramètres, comme la vitesse du vent et les conditions de stratification, et elle est d'autant plus grande que la hauteur de la mesure est élevée. En cas de terrain non-homogène, les obstacles responsables de la rugosité étant situés dans une aire qui est d'autant plus grande que la hauteur de mesure est élevée, l'augmentation de la surface d'influence permet d'englober de nouveaux obstacles de dimension potentiellement différente, ce qui a pour effet de modifier la valeur de la longueur de rugosité qui n'est alors pas constante sur la verticale.

Pour le site MET01, ces surfaces d'influence sont calculées en utilisant la méthode décrite par Hsieh et al. (2000), qui est basée sur la combinaison de résultats de modèle de dispersion Lagrangien et sur la théorie de similitude. Le rayon d'influence x est alors dérivée de la stabilité atmosphérique, de la hauteur à laquelle est effectuée la mesure z_m et de la longueur de rugosité z_0 :

$$x = \frac{Dz_u^P |L|^{1-P}}{k^2 \ln(a)} \quad \text{avec :} \quad z_u = z_m \left(\ln \left(\frac{z_m}{z_0} \right) - 1 + \frac{z_0}{z_m} \right) \quad (2.1)$$

avec L la longueur d'Obukhov, k la constante de von Karman, a la proportion de flux impactée par la zone déterminée par x (par exemple $a = 0,5$ pour le calcul de la zone impactant 50 % du flux) et D et P des coefficients dont les valeurs sont dépendantes du régime de stabilité :

- $D = 0,28$ et $P = 0,59$ pour les conditions instables ($z_u/L < -0,04$) ;
- $D = 2,44$ et $P = 1,33$ pour les conditions stables ($z_u/L > 0,04$) ;
- $D = 0,97$ et $P = 1$ pour les conditions neutres ou proches de la neutralité.

Seules les données correspondant aux conditions proches de la neutralité ont été utilisées parce que ce sont ces conditions qui permettent de calculer simplement une valeur de z_0 à partir des observations. Une valeur x est calculée pour chaque instant et donne une distance caractérisant la zone d'influence dans la direction d'où vient le vent. Une valeur médiane est ensuite calculée pour chaque secteur de direction de vents de 10° .

Les zones d'influence calculées pour les trois hauteurs d'anémomètres soniques de la station MET01 pendant la période KASCADE 2013 sont représentées sur la figure 2.12. Il y a bien une augmentation de l'étendue du footprint avec la hauteur. À 2 m, 50 % du flux est impacté par des obstacles compris dans un rayon d'environ 50 m, aire dans laquelle la surface est couverte de prairie ce qui est cohérent avec une valeur de rugosité de quelques centimètres. À 10 m, les obstacles impactant la mesure sont présents sur une surface comprenant quelques arbres, d'où une longueur de rugosité de quelques dizaines de centimètres. Enfin, la zone impactant les observations à 30 m englobe des bâtiments voire les premiers reliefs boisés entourant la station, ce qui explique la longueur de rugosité calculée de 2,96 m.

Malgré une définition grossière de l'occupation du sol dans les simulations WRF, la va-

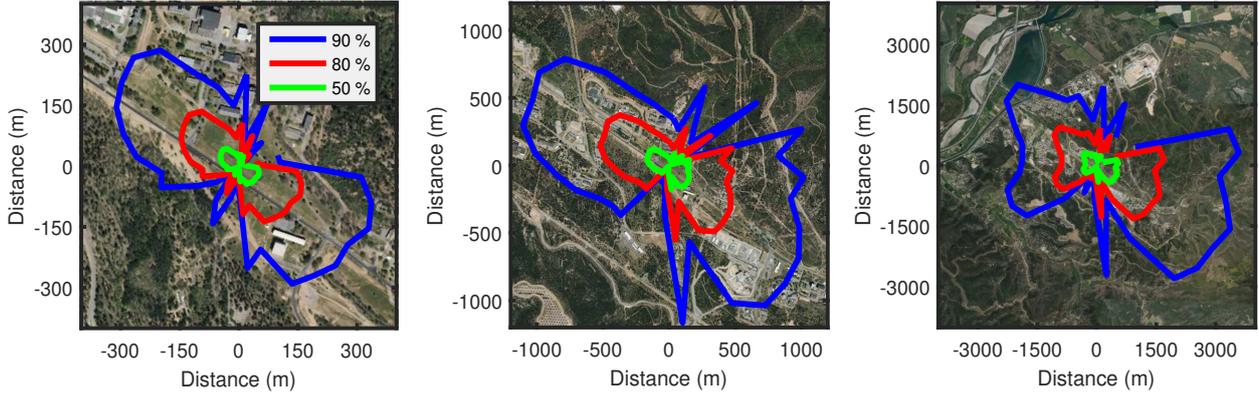


FIGURE 2.12 : Zones d'influence impactant les mesures des anémomètres soniques en conditions neutres (à 2 m à gauche, à 10 m au centre et à 30 m à droite) en fonction de la direction d'où vient le vent, calculées à partir de la méthode décrite par Hsieh et al. (2000). Le contour 50 % (en vert) indique l'aire dans laquelle se trouvent les obstacles impactant 50 % des mesures (de même pour les contour 80 et 90 %). Les images satellites proviennent des données de l'institut national de l'information géographique et forestière (IGN) : <https://www.geoportail.gouv.fr/>.

leur de longueur de rugosité définie à partir de l'occupation du sol est donc proche de la valeur caractérisant le site MET01 à 10 m calculée à partir de mesures in situ (0,50 m pour les simulations contre 0,56 m dans les observations). Cependant, la loi logarithmique définissant le profil de vent dans la couche limite atmosphérique est valide pour des terrains homogènes, avec une longueur de rugosité indépendante de la hauteur. Dans notre cas, la forte variation de z_0 avec la hauteur s'explique par l'hétérogénéité du site et implique que le profil de vent n'est pas logarithmique. Pour cette raison, nous avons considéré que le calcul du vent à 2 m à partir de sa valeur à 10 m ne pouvait se faire avec une loi logarithmique à z_0 constant.

Une simple régression linéaire entre les vents de ces deux niveaux (2 m et 10 m) a alors été testée. Toutes les conditions de vent ont été utilisées pour ce calcul et pas uniquement les conditions neutres. Les résultats de cette régression sont donnés dans l'équation 2.2.

$$U_{(2\text{m})} = 0,8176 \times U_{(10\text{m})} - 0,0728 \quad (\text{m s}^{-1}) \quad (2.2)$$

Les résultats des calculs effectués en utilisant le profil logarithmique avec des longueurs de rugosité soit de 0,56 m soit de 0,04 m (valeurs obtenues à 10 m et 2 m) sont comparés à ceux produits par la régression linéaire sur la figure 2.13. Il est alors clair que la régression linéaire produit les meilleurs résultats avec une erreur absolue moyenne de $0,12 \text{ m s}^{-1}$, soit nettement inférieure aux deux autres méthodes, ce qui représente une incertitude tout à fait acceptable. C'est pourquoi le calcul des vitesses de vent WRF à 2 m s'est fait sur la base d'une régression linéaire entre les vents à 2 et 10 m.

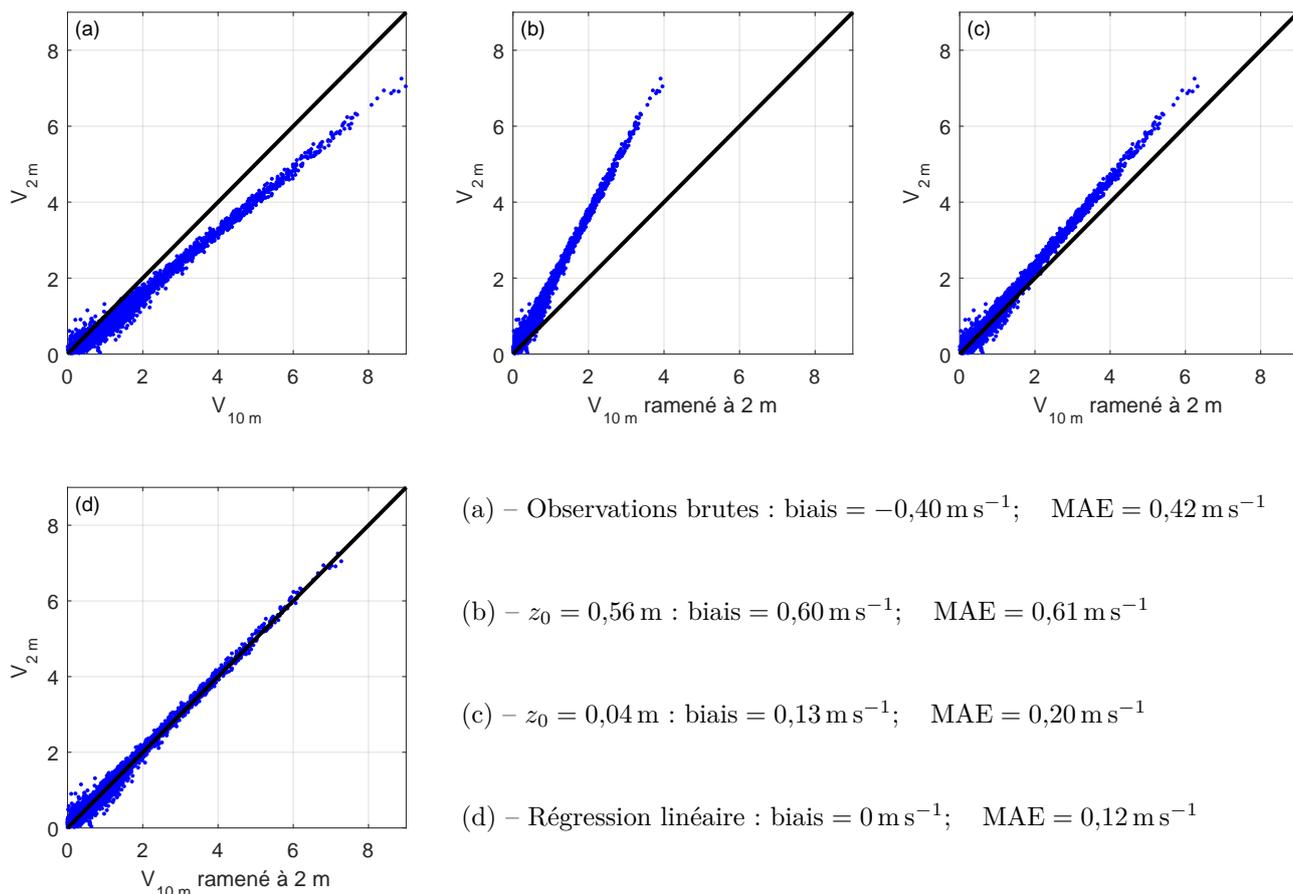


FIGURE 2.13 : Comparaison des vitesses de vent observées à 2 m et 10 m sur la station MET01 pendant KASCADE 2013 (a), puis pour lesquelles les vitesses à 10 m ont été recalculées à une hauteur de 2 m en utilisant la loi logarithmique et une longueur de rugosité $z_0 = 0,56 \text{ m}$ (b) et $z_0 = 0,04 \text{ m}$ (c), et enfin en utilisant une régression linéaire avec l'équation 2.2 (d).

2.3.4.b Vents à 2 m

Comme nous l'avons montré dans la section 2.2.3, les vents observés à 2 m sur la période de un an (tableau 2.4) affichent un comportement assez clair avec deux directions majoritaires, globalement alignées avec l'axe de la vallée. Il y a 38,6 % des vents qui remontent la vallée (entre 245° et 335°), 39,2 % qui descendent la vallée (entre 105° et 165°) et seulement 22,2 % qui ont des directions transverses à la vallée. On ne retrouve pas ces proportions dans les prévisions WRF (ramenées à 2 m avec l'équation 2.2) puisqu'il y a moins de vents descendant la vallée (22,7 %) et plus de vents transverses (38,5 %). Ce surplus de vent transverses s'explique par une proportion trop importante de vents dont les directions sont forcées par des vents en altitude qui ont des directions non alignées avec l'axe de la vallée.

La comparaison des vents prévus par WRF avec les vents observés à 2 m (figure 2.14) permet de se rendre compte que les vents qui remontent la vallée sont globalement mieux représentés que ceux qui la descendent (64 % de vents bien classés pour les vents compris entre 245° et 335° contre 36 % pour ceux compris entre 105° et 165°). Enfin, les simulations produisent des vents beaucoup plus forts que les observations (biais de $+1,01 \text{ m s}^{-1}$). Cela est une conséquence de l'absence de représentation du relief sous-maille dans la simulation. Dans la réalité, la vallée de Cadarache crée un abri qui y réduit la vitesse du vent. Ce biais est également visible sur la figure 2.14 comparant les vitesses prévues à celles observées. Les performances des prévisions de vent par les simulations WRF sont évaluées à l'aide des critères statistiques définis dans la partie 2.3.3 (tableau 2.5).

	$< 0,5 \text{ m s}^{-1}$		$> 0,5 \text{ m s}^{-1}$		Total	
	WRF	Obs	WRF	Obs	WRF	Obs
CDV	2,1 %	10,6 %	20,6 %	28,0 %	22,7 %	38,6 %
CUV	1,8 %	3,5 %	37,0 %	35,7 %	38,8 %	39,2 %
Autres	5,6 %	11,9 %	32,9 %	10,3 %	38,5 %	22,2 %
Total	9,5 %	26,03 %	90,5 %	74,0 %	100 %	100 %

Tableau 2.4 : Proportion de vents à 2 m dans chacune des 3 classes définies à partir de la rose des vents observée sur la période de 1 an (cf. figure 2.6), pour les simulations WRF et pour les observations. Il y a également une distinction suivant la vitesse du vent (seuil à $0,5 \text{ m s}^{-1}$).

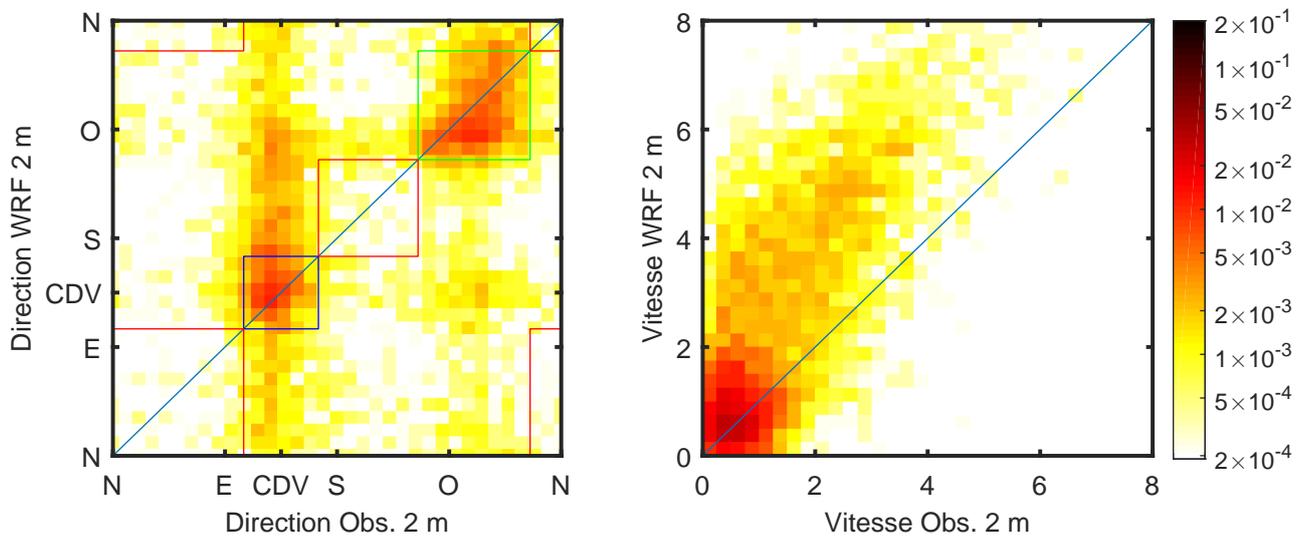


FIGURE 2.14 : Comparaison des vents prévus par les simulations WRF à 2 m avec les observations, représentée en densité de probabilité. (a) pour les directions, (b) pour les vitesses. Les rectangles de couleur correspondent aux trois classes de direction de vent définies à partir de la rose des vents observée à 2 m sur la période de 1 an (cf. figure 2.6).

Cas	DACC	PC2	PC4	Corrélation			Vitesse (m s^{-1})	
				Vitesse	u	v	Biais	MAE
WRF 2 m	0,49	0,67	0,47	0,74	0,69	0,57	+1,01	1,27
WRF-LW 2 m	0,53	0,70	0,44	0,72	0,69	0,58	+1,27	1,44
WRF 110 m	0,63	–	0,66	0,69	0,77	0,60	+1,20	2,06
WRF-LW 110 m	0,63	–	0,66	0,69	0,77	0,60	+1,22	2,07

Tableau 2.5 : Récapitulatif des résultats de performance des prévisions de vent avec WRF sur un an. $PC2$ pour 2 classes et $PC4$ pour 4 classes. Les biais et MAE (erreur absolue moyenne) sont en m s^{-1} . -LW pour les performances calculées après retrait des vents (WRF à 2 m) inférieurs à $0,5 \text{ m s}^{-1}$.

2.3.4.c Vents à 110 m

La comparaison des vents à 110 m (figure 2.15 et tableau 2.5) montre que les vents d'ouest-nord-ouest sont globalement mieux représentés (74 % de vents bien classés pour les vents compris

entre 245° et 335°) que les vents de sud-est (50 % de réussite pour les vents compris entre 115° et 155°) ou les vents de nord-est (52 % de réussite pour les vents compris entre 15° et 75°). Concernant les vents observés de nord-est (vents d'origine thermique qui descendent la vallée de la Durance), ils sont souvent bien prévus, mais pour un nombre significatif de cas le modèle simule un vent de sud-est. Concernant le module, il y a beaucoup de dispersion (coefficient de corrélation de 0,69, plus faible qu'à 2 m où il vaut 0,74).

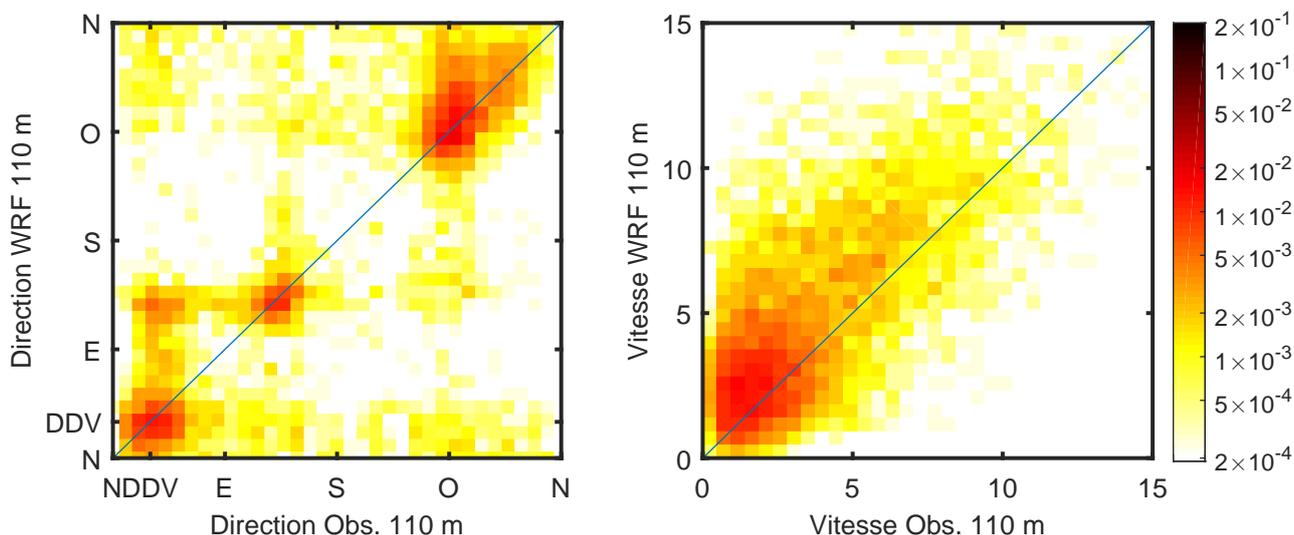


FIGURE 2.15 : Comme la figure 2.14 pour les vents à 110 m. DDV indique la direction du vent descendant la vallée de la Durance (Durance Down-Valley).

2.3.4.d Relation entre les vents à différentes hauteurs

La figure 2.16, qui représente le lien entre les vents à 10 m, 110 m et 300 m dans les simulations, montre un comportement différent des observations (figure 2.7).

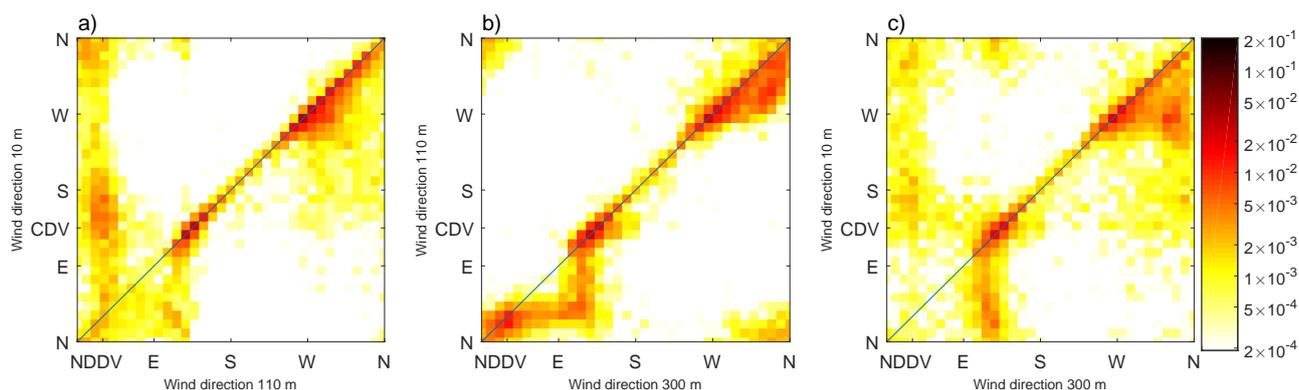


FIGURE 2.16 : Comme la figure 2.7.a pour les vents issus des simulations WRF sur la même période et pour des altitudes de 10, 110 et 300 m.

En effet, les vents d'ouest à nord à 110 m impactent systématiquement les vents à 10 m en forçant leur direction. Les vents à 10 m prennent alors soit exactement la même direction que les vents à 110 m, soit une direction légèrement déviée vers l'ouest. Cette rotation se produit également à des altitudes plus élevées puisqu'elle apparaît entre les vents à 300 m et 110 m. Dans les observations, il est fréquent de retrouver, avec un forçage d'altitude de nord-ouest, des vents de sud-est près de la surface, qui correspondent aux vents de vallée d'origine thermique.

Cette configuration n'est pas du tout représentée dans les simulations.

Les vents d'ouest à sud-est à 110 m forcent systématiquement les directions des vents à 10 m qui prennent la même direction. Dans les observations, au contraire, ces vents sont découplés des vents près de la surface, ces derniers étant des vents d'origine thermique qui montent ou descendent la vallée de Cadarache. Ce phénomène n'est donc jamais représenté dans les simulations.

Les vents du secteur nord-est à 110 m, comprenant les vents qui descendent la vallée de la Durance, sont souvent associés à des vents de sud-sud-est à 10 m, ce qui pourrait correspondre aux vents qui descendent la vallée de Cadarache. Or la vallée n'est pas représentée dans les simulations WRF à 3 km de résolution horizontale. Par contre, les reliefs qui entourent la vallée de Cadarache de l'est au sud sont assez bien représentés dans les simulations, et comme le montre l'exemple de la figure 2.17, les vents de sud-sud-est sont des vents qui descendent ces reliefs pour rejoindre la vallée de la Durance. La figure 2.17.b de droite montre les vents au niveau 960 hPa. La pression au niveau de la station MET01 étant de 973 hPa, ce niveau correspond à une hauteur au-dessus du sol d'environ 110 m (à la position de la station MET01), soit une altitude proche de 420 m asl. Les vents à ce niveau sont alignés avec l'axe de la vallée de la Durance. Il y a alors un découplage entre ces vents et ceux à 10 m au-dessus du sol, ces derniers affichant des directions descendant les pentes locales. Ainsi, cela explique aussi qu'une partie des vents descendant la vallée soient bien représentés dans les simulations en condition calme, donc sans forçage de moyenne échelle, bien que la vallée de Cadarache ne soit pas représentée.

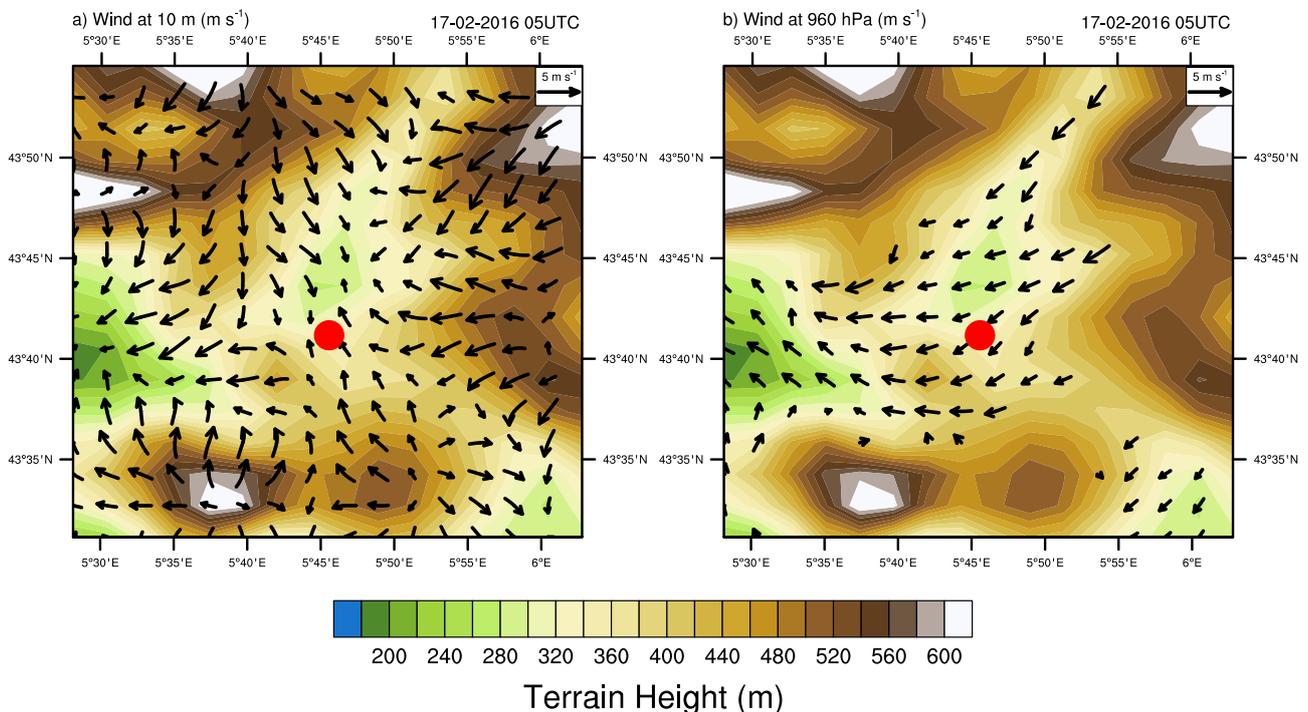


FIGURE 2.17 : Cartes de prévisions du vent de la simulation du 15 février 2016 pour l'échéance du 17 février 2016 05 UTC (+41 h). (a) pour les vents à 10 m au-dessus du sol et (b) pour les vents à 960 hPa (la pression à 2 m au niveau de la station MET01 est de 973 hPa). La topographie dans les simulations WRF est représentée en couleur. Le point rouge matérialise la position de la station MET01 qui est située au milieu de la vallée de Cadarache.

Les fréquents découplages observés entre les vents à 2 m et 110 m ne sont donc pas bien reproduits. À l'inverse, les simulations affichent un forçage important des vents d'altitude sur

les vents près de la surface. Ce forçage est lié à la mauvaise représentation du relief dans les simulations ainsi qu'à la surestimation de la force du vent en altitude (biais positif de $+1,20 \text{ m s}^{-1}$ à 110 m de hauteur).

2.3.4.e Impact du forçage par les vents d'altitude

Il est possible d'identifier deux sources majeures d'erreurs de prévision des vents près de la surface. La première est liée à la mauvaise représentation du relief dans les simulations WRF. La deuxième source d'erreur est liée à une mauvaise prévision des forçages dont les vents de basses couches sont dépendants, notamment du forçage par les vents sus-jacents.

Le tableau 2.6 montre une comparaison de la qualité des prévisions de vent à 2 m en fonction de la qualité du forçage par les vents à 110 m. Le forçage est considéré correct pour une prévision de direction du vent à 110 m qui ne diffère pas de la direction observée de plus de 45° . Il est alors clair que la qualité de la prévision des vents à 2 m dépend grandement de celle des vents à 110 m puisque la plupart des critères statistiques sont meilleurs dans le cas d'une bonne prévision à 110 m. Il est toutefois à noter que le biais et l'erreur absolue moyenne sont plus forts dans le cas d'un bon forçage, ce qui s'explique par le fait que les bons forçages en direction sont associés à des vents forts. Or le biais sur les vitesses est d'autant plus grand que les vents sont forts.

Cas	DACC	PC2	PC4	Corrélation			Vitesse (m s^{-1})	
				Vitesse	u	v	Biais	MAE
Forçage bon	0,61	0,75	0,55	0,76	0,80	0,70	+1,64	1,73
Forçage mauvais	0,33	0,55	0,28	0,52	0,14	0,18	+0,96	1,12

Tableau 2.6 : Récapitulatif des résultats de performance des simulations WRF (vitesses recalculées à 2 m) en fonction de la qualité du forçage par les vents d'altitude (prévision de la direction du vent à 110 m correcte à $\pm 45^\circ$ ou incorrecte).

Cependant, il reste de nombreuses erreurs même avec un bon forçage. La figure 2.18 montre une comparaison des directions de vent observées à 2 m et 110 m dans le cas d'un « bon forçage ». Dans ce cas, les simulations produisent de bonnes prévisions de vents à 2 m (figure 2.18.b) dans le cas où les directions sont forcées par les vents d'altitude. Quelques bonnes prévisions interviennent également quand les vents à 110 m viennent du nord-nord-est (vents qui descendent la vallée de la Durance) et ceux à 2 m viennent du sud-est (vents qui descendent le relief au sud de la vallée de Cadarache).

Cependant, un bon forçage par les vents à 110 m ne suffit pas à bien reproduire la totalité des vents à 2 m lorsqu'ils sont d'origine thermique. Par exemple, les erreurs correspondant aux vents d'ouest en altitude et sud-est en surface (figure 2.18.(c)) correspondent aux débuts de nuits quand le vent de vallée de Cadarache est déjà établi alors que celui de vallée de Durance ne l'est pas encore, ce qui n'est pas reproduit dans les simulations. De même pour les vents de nord-est en altitude et nord-ouest en surface qui correspondent aux débuts de matinée quand les vents dans la vallée de Cadarache remontent la vallée par un effet thermique alors que ceux dans la vallée de la Durance continuent de la descendre pendant encore quelques heures après le lever du soleil.

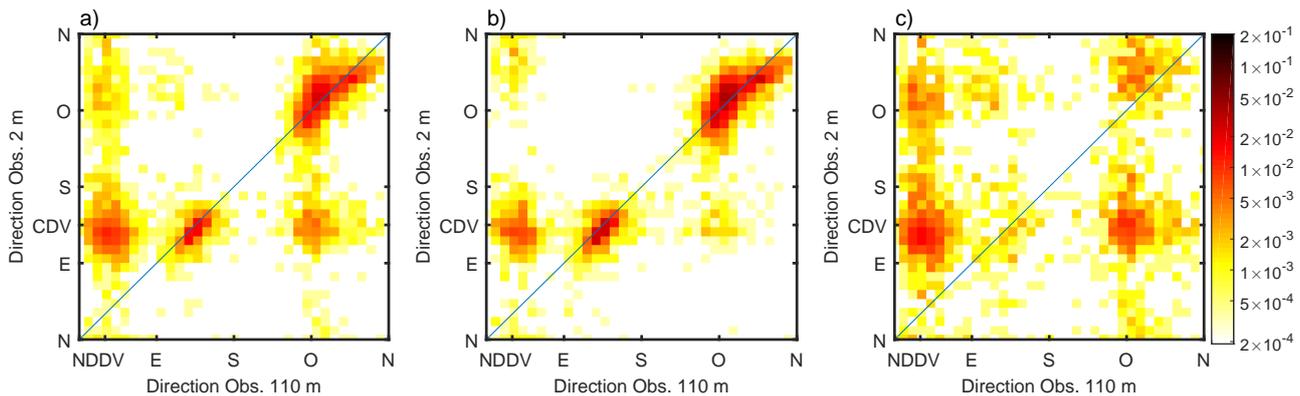


FIGURE 2.18 : Comme la figure 2.7.(a) pour les conditions où le forçage à 110 m est correct. (b) : idem mais pour les prévisions de direction du vent WRF à 2 m correctes (à $\pm 45^\circ$). (c) : idem mais pour les prévisions de direction du vent WRF à 2 m fausses.

2.3.4.f Stratification entre 110 m et 2 m

La stratification simulée est évaluée par comparaison avec les mesures de température effectuées à la station GBA à des hauteurs de 2 m et 110 m. La stratification peut notamment être évaluée à l'aide d'un gradient vertical de température potentielle virtuelle (voir le chapitre 1) mais il manque les mesures de pression et d'humidité à 110 m nécessaires au calcul d'une température potentielle virtuelle.

Des tests de sensibilité ont été effectués afin d'évaluer l'intensité de la stratification à partir des mesures de la station GBA. Les résultats sont détaillés dans l'annexe A. Il en ressort que bien que les θ_v soient très sensibles aux valeurs de pression et humidité, ce sont les gradients de ces variables qui sont déterminants dans le calcul de la différence verticale de θ_v . Plus les gradients verticaux sont faibles et moins la différence verticale est sensible à ces gradients. Les observations réalisées dans la vallée de Cadarache par ballon captif lors de la campagne KASCADE 2013 ont montré que les gradients verticaux d'humidité sont faibles (différence d'humidité spécifique généralement inférieure à 1 g kg^{-1}) ce qui permet d'obtenir une bonne approximation de la différence verticale de température potentielle virtuelle ($\Delta\theta_v$) à partir de la différence verticale de température potentielle ($\Delta\theta$) : biais de $-0,05^\circ\text{C}$ et MAE de $0,07^\circ\text{C}$ sur les 789 profils de ballon captif testés.

Concernant le gradient de pression, l'utilisation de l'approximation hydrostatique ou d'un gradient adiabatique sec standard permettent d'obtenir des résultats proches du $\Delta\theta_v$: biais de $-0,03^\circ\text{C}$ et MAE de $0,07^\circ\text{C}$ avec l'approximation hydrostatique et $-0,73^\circ\text{C}$ et $0,09^\circ\text{C}$ en utilisant le gradient adiabatique sec. La stratification (notée $\Delta\theta_{(110\text{m}-2\text{m})}$) dans les observations est alors approximée en utilisant la méthode du gradient adiabatique sec standard.

La figure 2.19 permet d'évaluer la représentation du $\Delta\theta_{(110\text{m}-2\text{m})}$ dans le modèle par rapport aux observations. Les observations montrent un cycle diurne bien défini avec une stratification stable de l'atmosphère la nuit et des températures plus élevées à la surface le jour. La tendance de ce cycle est globalement bien représentée dans la prévision avec une atmosphère instable pendant la journée ($\Delta\theta_{(110\text{m}-2\text{m})} < 0$) et stable la nuit. Cependant, il y a une nette différence le matin, avec une déstabilisation de l'atmosphère plus rapide dans les prévisions qui est liée au chauffage plus rapide de la surface dans les simulations.

Le changement de stabilité à l'approche du coucher du soleil (figure 2.19.a) est décalé, prévu environ 3 h avant le coucher du soleil contre environ 4 h dans les observations. En re-

vanche, après le coucher du soleil, le maximum de stratification est atteint dans les observations 4 h après le coucher du soleil puis la stabilité diminue progressivement. Au contraire, dans la simulation, la stabilité ne croît que lentement mais le fait tout au long de la nuit. Le matin (figure 2.19.b), la diminution de stratification se produit 1 h après le lever du soleil dans les prévisions alors qu'il faut 1 h de plus dans les observations. De manière générale la stratification est légèrement plus forte dans les observations (biais négatif de $-0,59$ K et erreur absolue moyenne de $1,27$ K) voire franchement plus forte dans la période de transition matinale. Néanmoins, les valeurs simulées restent majoritairement comprises dans l'écart-type des observations, à l'exception de la période de transition matinale. Enfin, le coefficient de corrélation entre les observations et les prévisions sur cette variable est de $0,75$.

Le graphe de densité de fréquence (figure 2.19.c) permet de visualiser le niveau d'accord entre les observations et les prévisions. Le graphe peut être découpé en trois parties. La première correspond aux $\Delta\theta_{OBS} < 0$ K. Dans cette partie, qui correspond à de l'instabilité, les $\Delta\theta_{WRF}$ sont également négatifs et une majorité de points sont regroupés autour de l'axe $x = y$ ce qui montre un bon accord. La deuxième partie du graphe correspond aux $\Delta\theta_{OBS}$ compris entre 0 K et 2 K majoritairement associés à la fin d'après-midi ou au début de matinée en se basant sur les figures 2.19.(a) et (b). Elle affiche un étalement des valeurs qui montre un moins bon accord entre les observations et les prévisions (comprises entre $-1,5$ K et $+4$ K) lié au décalage temporel correspondant aux changements de stratification. Les couleurs majoritairement rouges à gauche de l'axe $x = y$ alors qu'elles sont plutôt orange ou jaunes à droite montrent que la plupart des observations présentent une stratification supérieure aux prévisions. Enfin, la troisième partie correspond aux $\Delta\theta_{OBS} > 2$ K. Dans cette zone, qui correspond aux valeurs nocturnes, les prévisions identifient la stratification stable mais les données sont très dispersées ce qui montre la difficulté des simulations à bien représenter ce gradient de température potentielle en conditions stables.

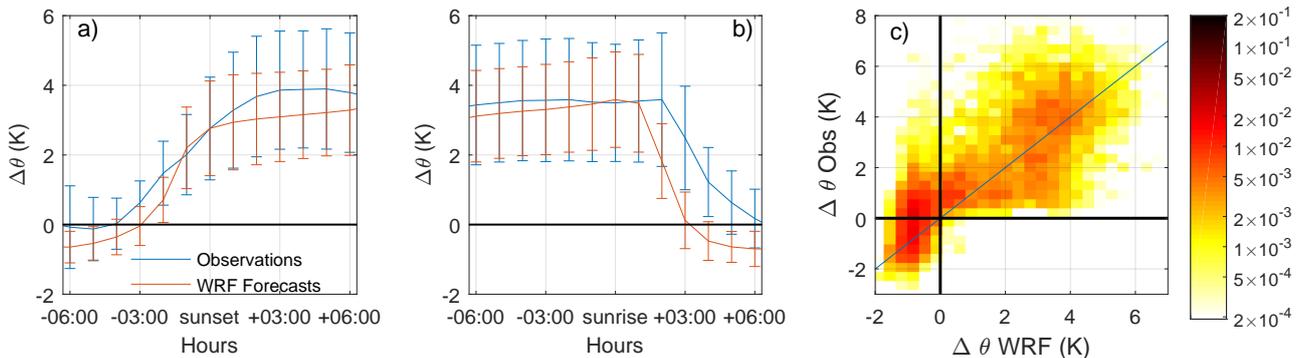


FIGURE 2.19 : (a) : Moyenne horaire de la différence de température potentielle entre 110 m et 2 m présentée relativement à l'heure du coucher de soleil. En bleu pour les observations et en rouge pour la prévision. Les barres verticales représentent l'écart-type ; (b) : idem mais calculé par rapport au lever de soleil ; (c) : densité de fréquence montrant la comparaison entre les observations et les prévisions. Les données horaires de la période de 1 an ont été utilisées pour tracer cette figure.

Étant donné les limites rencontrées dans la représentation déterministe des processus de basse couche avec une résolution horizontale de 3 km, des techniques de descente d'échelle ont alors été explorées (Duine, 2015) : d'abord une méthode déterministe visant à effectuer des simulations WRF avec une résolution horizontale plus importante (voir la partie suivante), puis une méthode statistique (présentée dans la partie 2.4).

2.3.5 Simulations à 1 km de résolution horizontale

Des simulations à 1 km de résolution horizontale ont été effectuées ponctuellement sur la région afin de tenter de mieux représenter les vents dans les vallées de Cadarache et de Durance par conditions stables en période hivernale. Tous les détails concernant ces simulations sont rassemblés dans la publication de [Kalverla et al. \(2016\)](#) et la thèse de [Duine \(2015\)](#). L'objectif de ce paragraphe est de faire une synthèse sur l'apport de ces simulations.

Ces deux études sont basées sur les observations réalisées pendant la campagne de mesure KASCADE 2013. Une résolution de 1 km ne permettant pas de représenter la vallée de Cadarache ni par conséquent le vent de vallée qui lui est associé, ce vent n'a pas été étudié à partir de ces simulations. En revanche la vallée de la Durance est mieux représentée que dans les simulations à 3 km de résolution horizontale et ces simulations ont permis l'étude des écoulements le long de cette vallée (de 60 km en amont jusqu'à quelques kilomètres en aval du site de Cadarache).

Tout d'abord, rappelons que le seul point de la vallée de Durance pour lequel des observations sont disponibles est situé au niveau du site de Cadarache. En ce point, la bonne représentation du relief de la vallée de Durance permet de reproduire les caractéristiques des vents dans cette même vallée. Notamment, la relation entre le vent dans la vallée de Durance et le vent au-dessus de celle-ci, sur la base de la description de [Whiteman and Doran \(1993\)](#), est bien reproduite dans les simulations. Cependant, l'apparition et la disparition du vent de vallée interviennent trop tôt dans les simulations par rapport aux observations.

L'étude le long de la vallée a montré que les origines du vent qui la descend sont similaires en plusieurs points de la vallée et s'expliquent par un mélange de canalisation des vents présents au-dessus de la vallée et de vents d'origine thermique. Finalement, les origines conduisant à la production d'un vent de vallée descendant peuvent être différentes d'une situation à une autre et peuvent conduire à son apparition avec un délai plus ou moins important après le coucher du soleil.

2.4 Prévision immédiate du vent local par méthode statistique

[Duine et al. \(2016\)](#) ont développé une méthode statistique de prévision immédiate permettant de diagnostiquer la présence d'un vent d'origine thermique descendant la vallée de Cadarache. Pour cela, ils se sont basés sur des mesures ponctuelles du vent à 10 m, réalisées pendant la campagne KASCADE 2013 et sur les observations continues de la station GBA.

L'objectif de la méthode est de distinguer, pour les vents à 10 m dans la vallée de Cadarache (à la position de la station MET01), les vents CDV (vents d'origine thermique descendant la vallée) des autres directions, en définissant un seuil sur un paramètre calculé à partir des observations sur la station GBA ($\Delta\theta_{(110m-2m)}$, U_{110m} et Ri_B). Par exemple, pour la vitesse du vent, les vents CDV, qui sont des vents d'origine thermique, sont principalement associés à des vents faibles à 110 m qui ne produisent pas un mélange important. L'hypothèse est alors que les vents CDV sont associés à des vitesses inférieures à une valeur seuil alors que les autres directions sont associées à des vitesses supérieures. Pour différentes valeurs de vitesse testées en tant que seuil, la proportion de directions de vents suivant l'hypothèse énoncée est calculée :

c'est le $PC2$. Une valeur seuil correspondant à la valeur maximale du $PC2$ est alors déterminée pour chacun des trois paramètres testés.

La meilleure prédictibilité ($PC2 = 0,90$) a été obtenue avec la grandeur $\Delta\theta_{(110m-2m)}$ pour la valeur seuil de 2,6 K, les situations supérieures à ce seuil présentant majoritairement des vents thermiques descendant la vallée de Cadarache. Les données utilisées pour arriver à ce résultat sont des moyennes sur 30 minutes auxquelles trois étapes de sélection ont été appliquées :

- suppression des données pour lesquelles une des variables est manquante suite à des problèmes de mesure ;
- suppression des vents inférieurs à $0,5 \text{ m s}^{-1}$ (au niveau 10 m), puisque la direction du vent en-dessous de ce seuil n'est pas bien définie (Jarraud, 2008) ;
- suppression des vents pour lesquels les directions sont comprises dans le quart sud-est à la fois à 10 m et 110 m.

Cette dernière étape de sélection vise à supprimer les vents pour lesquels il n'est pas possible de distinguer avec certitude le mécanisme à l'origine des vents dans la vallée (forçage ou vent de vallée d'origine thermique), puisque les vents aux deux altitudes viennent du sud-est.

La dernière étape de sélection est basée sur les mesures du vent à 10 m. Or, ces mesures étaient temporaires, ce qui fait que cette étape n'est pas applicable à des données indépendantes de cette phase d'apprentissage. Cela intègre donc un biais dans la méthode de calcul quand elle est appliquée à un autre jeu de données puisque l'on ne peut pas appliquer cette troisième étape de sélection. Il en est de même avec la deuxième étape de sélection.

Dans notre étude, nous avons donc adapté cette méthode afin que les deux dernières étapes de sélection ne soient plus nécessaires. Le nouvel objectif est alors de distinguer les vents descendant la vallée de Cadarache (indépendamment du mécanisme menant à leur formation) des autres vents. Ainsi, la méthode peut être appliquée à un jeu de données indépendant de la période utilisée lors de l'apprentissage. La performance de la méthode diminue alors légèrement, le $PC2$ passant de 0,90 à 0,78.

2.5 Conclusion

La région du site d'étude est caractérisée par une grande hétérogénéité de surface qui impacte les écoulements près du sol. L'impact des vallées de Durance et de Cadarache a notamment été observé pendant la campagne de mesures KASCADE 2013. Les conditions stables, fréquemment observées la nuit dans la vallée de Cadarache, facilitent la production de vents d'origine thermique qui descendent les vallées. Cependant, les différences entre les tailles des deux vallées et les échelles spatiales impliquées dans la génération des vents d'origine thermique sont responsables de différences de mode de production, de force, d'étendue verticale et de délais d'apparition et disparition de ces vents au niveau de Cadarache.

Les observations étant concentrées dans la vallée de Cadarache, elles n'ont pas permis une étude approfondie des écoulements dans la vallée de Durance. Cependant, des simulations numériques à 1 km de résolution horizontale (réalisées par Duine (2015)), suffisantes pour modéliser l'impact du relief de cette vallée, ont permis une étude de l'évolution des écoulements le long de la vallée.

Concernant la vallée de Cadarache, KASCADE 2013 a apporté une connaissance impor-

tante sur la structure verticale de la stratification et des vents. Cependant, il subsiste toujours un manque de connaissances sur l'hétérogénéité des conditions dans la vallée du fait de la concentration des observations en un point unique à mi-longueur de la vallée. Cela limite la compréhension des écoulements et la validation des calculs des conditions météorologiques dans la vallée à un seul point. La campagne de mesure KASCADE 2017, qui s'est déroulée durant l'hiver et le printemps 2017, a donc eu pour but d'apporter des connaissances complémentaires à cet égard. La campagne sera présentée dans le chapitre 3. Les résultats concernant l'hétérogénéité et quelques phénomènes particuliers observés sur cette expérience (oscillations et augmentation de la température potentielle en montant dans le fond de la vallée alors qu'il y a bien une stratification sur les flancs de la vallée) seront également présentés dans ce chapitre.

Des prévisions météorologiques sont effectuées quotidiennement à 3 km de résolution horizontale sur la région de Cadarache avec le code WRF. Malheureusement, elles ne sont pas toujours réalistes du point de vue de la représentation des vents, aussi bien en altitude que près de la surface. Ces erreurs de prévision peuvent être attribuées à la fois à un problème de forçage synoptique et à la non représentation du relief de la vallée de Cadarache. En conséquence, à 2 m, les vents de sud-est, qui correspondent majoritairement aux vents d'origine thermique qui descendent la vallée (avec des directions forcées par la vallée), sont moins bien représentés que les vents de nord-ouest qui sont majoritairement observés en journée et dont les directions dépendent plus fortement des vents d'altitude. De même, les directions des vents à 110 m, moins dépendantes du relief que ne peuvent l'être celles des vents à 2 m, sont mieux représentées que celles près de la surface. Enfin, le manque de réalisme dans la représentation du relief conduit également à une surestimation des vitesses des vents en fond de vallée. Une solution visant à mieux représenter le relief serait d'augmenter la résolution dans les simulations WRF afin d'augmenter la qualité des prévisions de vent, mais cela reste encore difficile pour les raisons expliquées au chapitre 1.

Afin de diagnostiquer la présence d'un vent d'origine thermique dans la vallée de Cadarache, une première étape de prévision immédiate exploitant des observations de routine a montré le potentiel des méthodes statistiques (Duine et al., 2016). C'est pourquoi une nouvelle méthode de descente d'échelle statistique basée sur les prévisions WRF à méso-échelle à l'aide d'un réseau de neurones artificiels (RNA) a été développée, dans la continuité de la méthode de prévision immédiate de Duine et al. (2016). Cette méthode sera présentée dans le chapitre 4. Les résultats de son application dans un but de prévision immédiate seront également présentés dans cette partie ce qui permet entre autres d'apprécier l'apport de cette nouvelle méthode par rapport à celle de Duine et al. (2016). Enfin, son application dans un but de descente d'échelle des simulations WRF pour l'amélioration de la prévision des vents dans la vallée sera présentée dans le chapitre 5.

Chapitre 3

La campagne de mesures KASCADE 2017

L'objectif de la campagne expérimentale KASCADE 2017 est de faire suite aux observations effectuées lors d'une première campagne, réalisée durant l'hiver 2013 sur le site de Cadarache. En 2013, les observations visaient principalement à étudier la structure verticale de l'atmosphère par conditions stables. L'objectif de cette nouvelle expérience est de compléter les connaissances acquises en 2013 en étudiant l'hétérogénéité spatiale des conditions météorologiques au niveau du sol en portant une attention particulière au développement spatial et temporel du vent de vallée observé par conditions stables. Le site d'étude (la vallée de Cadarache) a déjà été présenté dans le chapitre 2.

Ce chapitre a donc pour but de présenter la campagne expérimentale KASCADE 2017 réalisée au cours de la thèse. La première partie décrit la méthodologie appliquée durant l'expérience, puis un descriptif des conditions rencontrées sur les trois premiers mois de mesure est donné, et enfin les résultats obtenus à partir de l'analyse de ces observations sont présentés dans les parties suivantes.

Plan du chapitre

3.1	Présentation de l'expérience	58
3.1.1	La stratégie expérimentale	58
3.1.2	Les observations continues	59
3.1.3	Les périodes d'observations intensives (POI)	62
3.2	Aperçu des conditions pendant l'expérience	64
3.2.1	Conditions météorologiques sur la période décembre 2016 – mars 2017	64
3.2.2	Oscillations	68
3.3	La CLA stable dans la vallée de Cadarache	79
3.3.1	Mesures de la stratification dans la vallée	79
3.3.2	Extension verticale de la CLA stable	79
3.3.3	Stratification dans la vallée de Cadarache	81
3.4	Écoulements dans la vallée de Cadarache	96
3.4.1	Couplage/découplage des écoulements dans et au-dessus de la vallée	96
3.4.2	Hétérogénéité des directions des vents	98
3.4.3	Évolution temporelle du vent de vallée de Cadarache	104
3.5	Conclusion	107

3.1 Présentation de l'expérience

3.1.1 La stratégie expérimentale

La vallée de Cadarache présente un terrain complexe du point de vue de la topographie et de l'occupation du sol (figure 2.2), ce qui est une source d'hétérogénéité pour certaines variables atmosphériques. L'objectif de KASCADE 2017 est d'étudier cette hétérogénéité ainsi que le développement du vent d'origine thermique présent dans la vallée de Cadarache par conditions stables. La théorie expliquant la formation d'un vent de vallée, décrite dans la partie 1.2.1, montre que ce vent est alimenté par un drainage le long des pentes. Afin d'observer la création et l'évolution de ce vent, l'instrumentation permettant une observation continue a été positionnée majoritairement suivant deux axes : dans le fond de la vallée en suivant son axe et sur un segment transverse à cet axe (figure 3.1). Il y a donc trois pentes qui sont instrumentées et sur lesquelles il est possible d'étudier le développement des vents de pente : les deux versants de la vallée et la pente de fond de vallée. Cette stratégie permet également de mesurer la stratification le long des versants de la vallée, stratification qui n'est pas toujours semblable à celle observée sur une verticale dans le centre de la vallée.

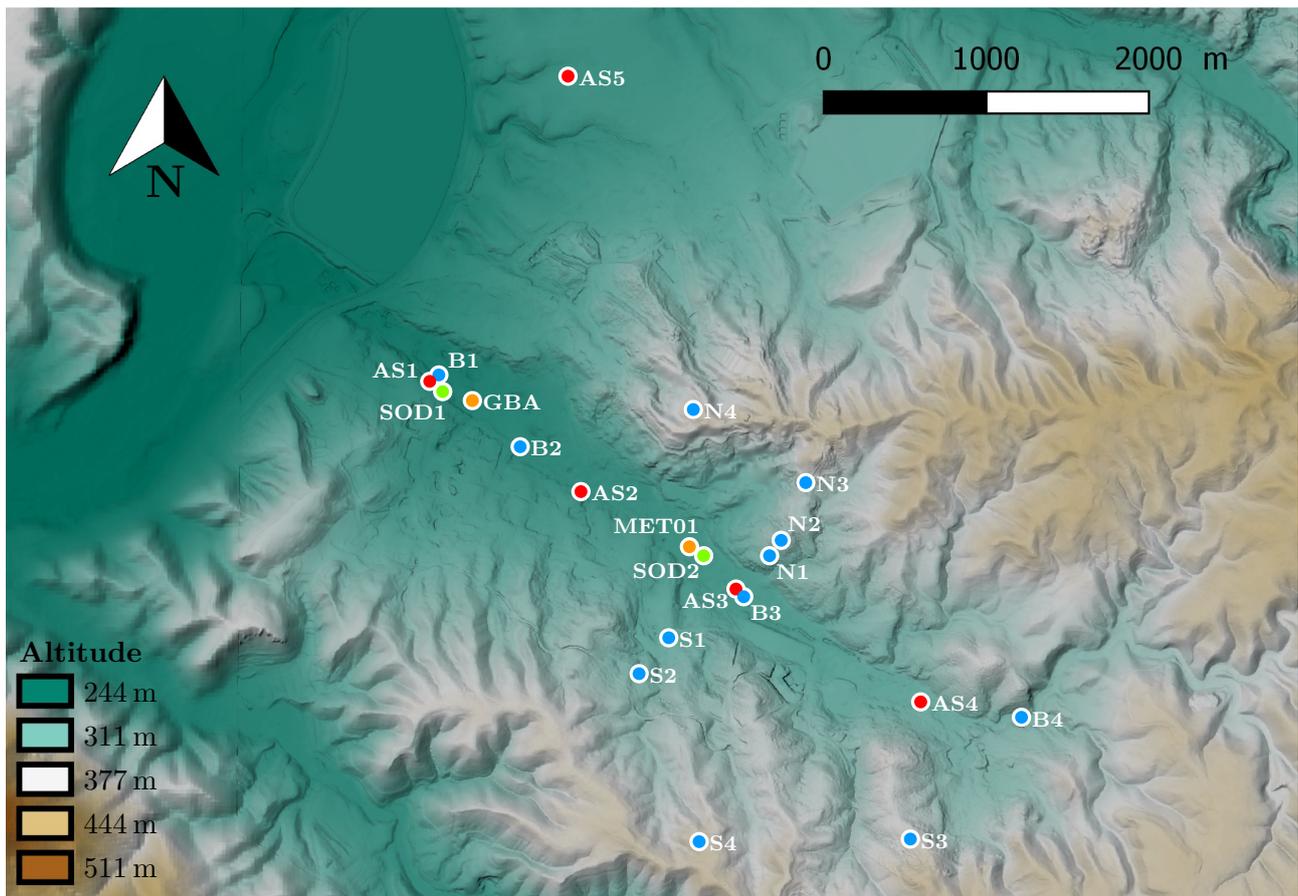


FIGURE 3.1 : Localisation et nom des différentes stations utilisées pendant la campagne KASCADE 2017. L'échelle de couleurs représente la topographie. Les stations dites LEMS sont marquées avec un point bleu et leur nomenclature consiste en un « B » pour les stations situées dans le fond de la vallée (pour bas ou bottom), un « S » pour celles situées sur le versant sud de la vallée et un « N » pour celles situées sur le versant nord. Ils sont numérotés de 1 à 4 avec l'augmentation de l'altitude à laquelle sont positionnées les stations. Les stations AS sont les stations avec anémomètres soniques (en rouge) et les stations SOD sont les sodars (en vert). Les points orange représentent les stations permanentes du CEA Cadarache (GBA et MET01).

Des mesures continues ont été effectuées (stations de la figure 3.1), complétées par 14 périodes d’observations intensives (POI) déclenchées dans le but d’étudier plus particulièrement les nuits les plus propices à la formation de conditions stables. Chaque POI consistait en une période de 24 heures, de 12h00 UTC à 12h00 UTC le lendemain, pendant laquelle des radiosondages étaient lancés toutes les 3 heures. Les radiosondages ont le double avantage d’apporter une connaissance sur la structure verticale de l’atmosphère et de valider les conditions grande échelle des simulations numériques. De plus, une ligne instrumentée avec ballon captif a aussi été installée pendant les nuits de certaines POI afin d’obtenir un profil vertical de température sur les 50 premiers mètres d’atmosphère.

L’instrumentation utilisée pendant les POI est la seule source d’information concernant la structure verticale de l’atmosphère du point de vue de la température et de l’humidité, ce qui est nécessaire en complément des observations continues des conditions stables.

3.1.2 Les observations continues

Un grand nombre de stations météorologiques (21) a été utilisé dans le but d’observer l’hétérogénéité. Ces stations peuvent être classées en quatre catégories : les LEMS (Local Energy-budget Measurement Station), les sodars profileurs de vent, les stations avec anémomètre sonique et les stations permanentes du site. Ces dernières sont les stations GBA et MET01 qui ont été introduites dans le chapitre 2 (elles ne sont donc pas présentées dans cette partie).

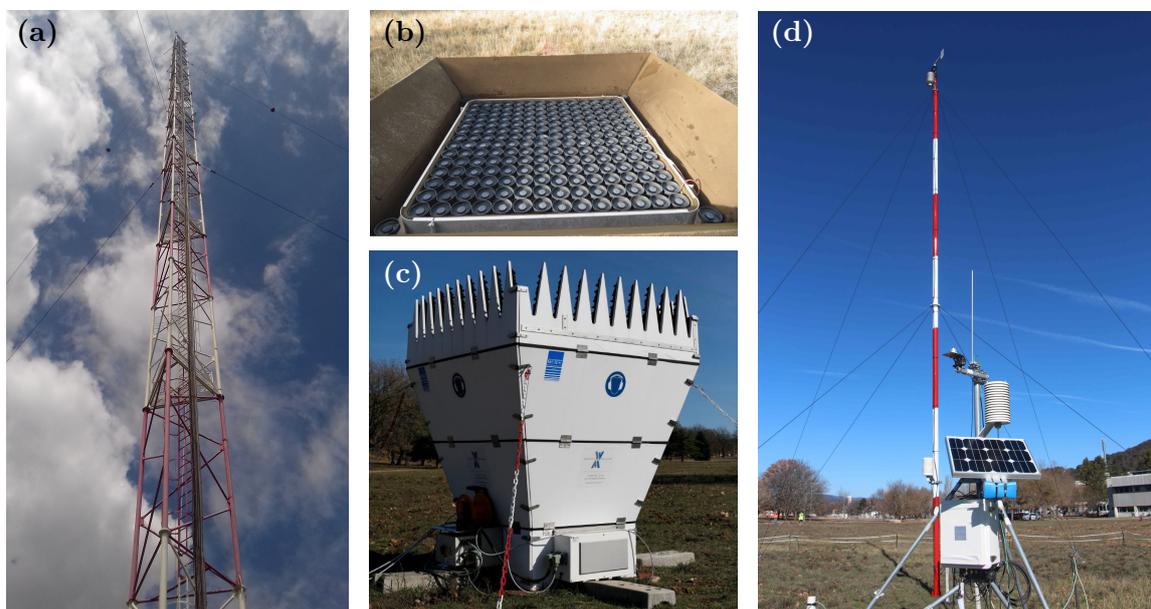


FIGURE 3.2 : Station GBA (a), SOD1 (b), SOD2 (c) et MET01 (d).

Les sodars

Des profils verticaux de vitesse et direction du vent ont été effectués grâce à deux sodars (SONic Detection And Ranging). Le premier (SOD1 sur la figure 3.1, photo 3.2.b), un Remtech PA2 appartenant au laboratoire de physico-chimie de l’atmosphère (LPCA) de l’Université du Littoral Côte d’Opale, était situé près de la sortie de la vallée de Cadarache. Il fournissait des moyennes de direction et vitesse sur des périodes de 15 minutes avec une résolution verticale

de 25 m entre 100 m et 500 m.

Le second (SOD2, figure 3.2.c), un Metek PCS.2000-24 appartenant au CEA, était situé approximativement au centre de la vallée, près de la station MET01. Il fournissait des profils verticaux entre 60 m et 400 m avec une résolution de 20 m et moyennés sur des périodes de 10 minutes.

Les sodars ne permettent pas en général d’observer correctement les vents dans la vallée de Cadarache car leur première fenêtre de mesure est trop haute, mais ils fournissent des informations importantes sur les vents présents au-dessus de la vallée tel que le vent de vallée de Durance.

Station	Hauteur	Mesures	Instrument
LEMS	1,5 m	T et RH	Thermo-hygromètre Sensirion SHT15
	2 m	WD et WS	Anémomètre à coupelle et girouette Davis 7911
	1,8 m	SW_{\downarrow}	Pyranomètre LI-COR LI-200R
	1 m	P	Baromètre Bosch BMP085
	–	T_s	Sensirion Zytemp TN9
	–5 et –25 cm	T et vw	Decagon 5TM
AS1	2 m	u, v, w et T_{son}	Anémomètre sonique Campbell CSAT3
	2 m	H_2O et CO_2	Licor 7500
	2 m	$SW_{\downarrow}, SW_{\uparrow}, LW_{\downarrow}$ et LW_{\uparrow}	Radiomètre Kipp and Zonen CNR4
	1 m	Pluie	Pluviomètre EML ARG100
	–5 cm	Q	Fluxmètre dans le sol Hukseflux HFP01
AS3	2, 5 et 10 m	u, v, w et T_{son}	Anémomètre sonique Young 81000
	2 m	T et RH	Thermo-hygromètre Campbell HMP45C
	1 m	Pluie	Pluviomètre à auget Campbell 52203
AS2–4–5	2 m	u, v, w et T_{son}	Anémomètre sonique Campbell CSAT3
SOD1	100–500 m	WS et WD	Remtech PA2
SOD2	60–400 m	WS et WD	Metek PCS.2000-24
MET01	2 m	SW_{\downarrow}	Pyranomètre Hukseflux LP02
		P	Baromètre Setra CS100
		T et RH	Thermo-hygromètre Campbell Sci. HC2S3
		Pluie	Pluviomètre EML ARG100
10 m	WS, WD	Anémomètre à hélice Young 05103	
	T	Sonde Campbell Sci. HC2S3	
GBA	2 m	T	Thermomètre Rotronic PT100
		RH	Hygromètre Rotronic
		P	Baromètre Vaisala PTB101C
	110 m	WS, WD	Anémomètre sonique Metek
		T	Thermomètre Rotronic PT100

Tableau 3.1 : Récapitulatif des mesures continues effectuées lors de la campagne expérimentale KASCADE 2017. WD et WS pour la direction et la vitesse du vent, T_{son} pour la température dérivée des mesures des anémomètres soniques, T_s pour la température de la surface, vw pour le contenu en eau dans le sol (volumetric water content of soil). Les autres variables ont été introduites précédemment.

Les stations LEMS

Les LEMS (Local Energy-budget Measurement Station, figure 3.3.a) sont des stations météorologiques développées par le Environmental Fluid Dynamics Laboratory (EFDL) de l'université de l'Utah (Gunawardena et al., 2018). Grâce à la collaboration avec le groupe d'Eric Pardyjak (université de l'Utah), et son implication dans la campagne de terrain, douze de ces stations ont été utilisées durant l'expérience. Chacune d'elles permet de mesurer avec une résolution temporelle de 10 secondes la vitesse et la direction du vent à 2 m, la température et l'humidité relative à 1,50 m, la pression atmosphérique, la température de la surface du sol, le rayonnement SW incident ainsi que la température et le contenu en eau dans le sol à deux profondeurs (5 et 25 cm). Le détail des instruments est renseigné dans le tableau 3.1. Un panneau solaire alimente chaque station en énergie. Les positions des stations sont indiquées sur la figure 3.1.

Une période d'intercomparaison de 4 semaines a eu lieu à la fin de l'expérience, durant le mois de juin 2017 (annexe A). Les principales conclusions concernant les LEMS sont que :

- les mesures des 12 LEMS sont très cohérentes entre elles ;
- la comparaison des températures avec les mesures des stations MET01 et AS3 montrent un très bon accord ;
- il existe quelques différences dans les mesures d'humidité relative qui impactent faiblement le calcul de la température virtuelle ($\pm 0,03^\circ\text{C}$ à une température de 0°C) ;
- les mesures de pression sont très cohérentes entre les 12 LEMS mais elles affichent un léger biais négatif (maximum de $-1,32\text{ hPa}$ pour les LEMS B3 et N2) par rapport à la mesure plus précise de la station MET01 sans que cela ne produise des différences importantes lors du calcul de la température potentielle (environ $0,1^\circ\text{C}$) ;
- le montage des capteurs de rayonnement SW_{\downarrow} rend difficile une comparaison des mesures des différentes stations (en effet, il n'est pas garanti que certains capteurs n'aient pas un défaut d'horizontalité) ;
- les modules de vent sont sous-estimés pour les vents faibles (inférieurs à $0,5\text{ m s}^{-1}$) et surestimés pour les vents les plus forts (généralement de moins de $0,5\text{ m s}^{-1}$ pour les vents supérieurs à $0,5\text{ m s}^{-1}$), mais ces mesures ne sont pas corrigées afin de garder une cohérence entre les 12 LEMS (en effet seules 4 des 12 LEMS ont enregistré des mesures de vent pendant la période d'inter-comparaison) ;
- les températures de surface sont globalement proches (biais inférieur à $0,8^\circ\text{C}$ sur l'ensemble des stations), les biais étant les plus faibles pour les faibles températures.

Il n'y a donc eu aucune correction de données à la suite de cette étude d'inter-comparaison entre les LEMS et les stations MET01 et AS3.

Les stations avec anémomètres soniques

Cinq stations avec anémomètres soniques ont été utilisées. Trois de ces stations étaient équipées d'anémomètre soniques Campbell CSAT3 installés à environ 2 m du sol. Une station permettant le calcul des quatre termes de l'équation du bilan d'énergie (cf. équations 1.24, 1.25 et 1.27), dont les flux de chaleur sensible et latente par la méthode d'eddy-covariance, était placée en bas de la vallée sur le même site que celui de la LEMS B1 (figure 3.3.b). Enfin, la station AS3 (figure 3.3.c) avec trois niveaux de mesure, à 2, 5 et 10 m était installée sur le même site que celui de la LEMS B3 (au milieu de la vallée). Le détail des mesures effectuées par ces 5 stations est reporté dans le tableau 3.1. Quatre de ces stations étaient situées dans le fond de

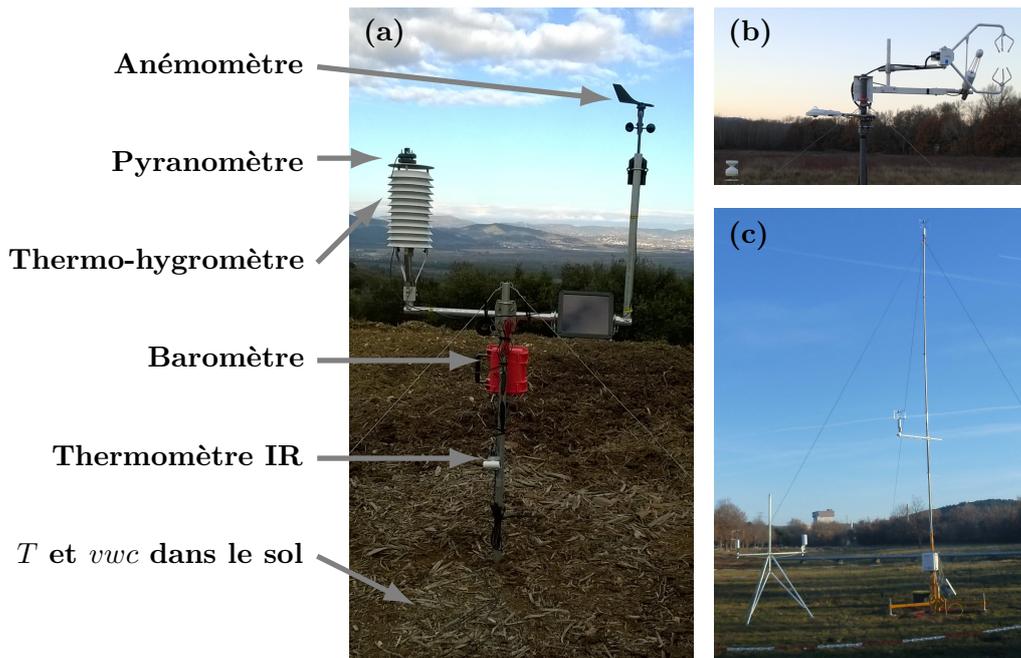


FIGURE 3.3 : Stations LEMS N4 (a), AS1 (b) et AS3 (c).

la vallée de Cadarache et une en dehors, dans la vallée de la Durance (station AS5).

De nombreuses études ont montré que de petites erreurs dans l'alignement des anémomètres soniques peuvent conduire à d'importantes erreurs dans les calculs de flux turbulents (Pond, 1968; Deacon, 1968; Kaimal and Haugen, 1969; Dyer and Hicks, 1972; Dyer, 1981). Ces erreurs proviennent d'une contamination des différentes composantes entre elles provoquée par une inclinaison du capteur qui fait croire à une fluctuation de vitesse horizontale quand c'est en fait une fluctuation sur la verticale et inversement. Il est alors nécessaire de faire pivoter le système de coordonnées pour compenser ces erreurs.

Ce traitement a été effectué avec le logiciel UTESpac qui a été développé à l'EFDL. La rotation du système de coordonnées est effectuée par la méthode du planar fit (Wilczak et al., 2001), ce qui permet d'avoir un système de coordonnées aligné sur l'écoulement moyen (lignes de courant). Cela permet d'avoir des données comparables entre des stations situées sur des sites différents dans les cas où il y a beaucoup d'hétérogénéité, notamment avec le relief. Par exemple, un vent qui suit la pente aura une composante verticale qui dépendra de l'inclinaison de la pente du terrain (les lignes de courants suivent le terrain). Sans correction du système de coordonnées, les calculs de flux verticaux seraient biaisés. Le logiciel UTESpac permet également de calculer les flux turbulents de quantité de mouvement et de flottabilité à partir des données haute fréquence des anémomètres soniques.

3.1.3 Les périodes d'observations intensives (POI)

Les mesures continues permettent d'observer les conditions près du sol dans la vallée mais, à l'exception des profils de vent au-dessus de la vallée de Cadarache fournis par les deux sodars, il n'y a pas d'information sur la structure verticale de l'atmosphère. Des mesures de profils verticaux ont alors été ajoutées pendant les POI.

Le ballon captif

Des profil verticaux de température ont été effectués dans le centre de la vallée à l'aide d'un ballon captif (figure 3.4.a). La ligne portait 5 thermocouples (GEMINI TGP-4020) espacés de 10 m permettant d'obtenir un profil dans les 50 premiers mètres au-dessus du sol, ainsi que d'un thermo-hygromètre (GEMINI TGP-4505) positionné à environ 1 m du sol. Le ballon utilisé pour le maintien de la ligne était un ballon de radiosondage. Sa forme sphérique ne permettait pas de garantir la stabilité de la ligne par vent modéré ou fort. De plus, les capteurs utilisés ne disposent pas d'abri contre le rayonnement ce qui ne permettait pas leur utilisation en journée. La ligne était montée à proximité de la station MET01 ce qui a permis de valider les mesures à 1 et 10 m du ballon captif (biais inférieur à $0,35\text{ }^{\circ}\text{C}$, voir l'annexe A). Par extrapolation, nous considérons que les mesures sur les niveaux supérieurs à 10 m sont également correctes.

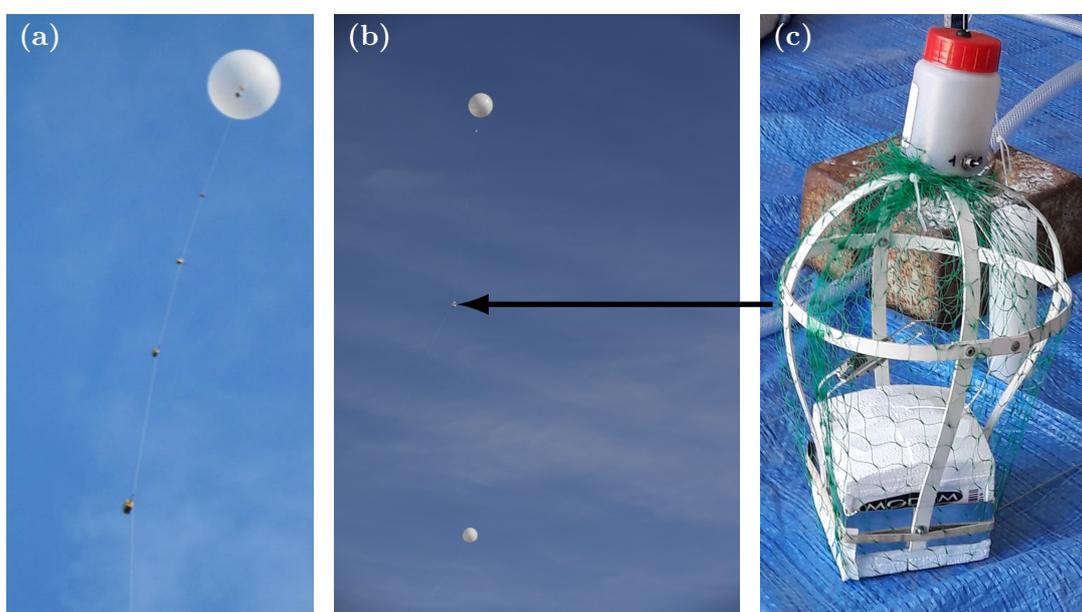


FIGURE 3.4 : Ballon captif (a), radiosondage selon la méthode décrite par [Legain et al. \(2013\)](#) (b) et une radiosonde dans sa cage de protection (c).

Les radiosondages

Des radiosondages ont été réalisés toutes les 3 heures lors de 14 journées d'observations intensives pour un total de 115 radiosondages. Des radiosondages troposphériques classiques, qui ont régulièrement atteint la tropopause, ont été faits à 12h00 UTC et 00h00 UTC. Dans un souci d'économie des radiosondes et parce que l'expérience porte sur l'étude de phénomènes de couche limite, les autres radiosondages ont été effectués en suivant la méthodologie décrite par [Legain et al. \(2013\)](#) (cf. figure 3.4.b et c). Ils avaient une durée d'ascension prédéterminée (typiquement 8 minutes pour une vitesse ascensionnelle d'environ 5 m s^{-1}) de sorte que la couche limite atmosphérique soit sondée sans pour autant aller jusqu'à la tropopause. Le temps de vol étant réduit, les sondes atterrissaient dans une zone proche du lieu de lâcher facilitant leur récupération en vue d'une réutilisation. De plus, la chute de la sonde était freinée par un deuxième ballon permettant l'acquisition d'un nouveau profil à la descente.

Lors de la calibration des sondes, effectuée avant chaque lancer, les conditions météorologiques initiales (température, humidité relative, vitesse et direction du vent près du sol) doivent

être saisies manuellement. Les mesures effectuées à la station MET01 ont été utilisées pour cette initialisation parce que la zone de lancement des radiosondages était proche de cette station.

Les sondes utilisées (modèle M10 du constructeur Modem) mesurent la température et l'humidité relative et embarquent un capteur GPS donnant la position de la sonde, le tout avec une fréquence d'acquisition de 1 seconde. Ces données sont transmises au sol par onde radio, puis le logiciel de post-traitement, fourni avec la station de réception, estime la pression atmosphérique à partir de la position de la sonde. Le vent (direction et module) est calculé à partir de la trajectoire puis lissé pour corriger l'effet de pendulage de la sonde sous le ballon lors de l'ascension.

Une comparaison des profils de température obtenus avec les radiosondages et le ballon captif a montré que l'accord n'est pas toujours bon pour les niveaux les plus près du sol. Il apparaît que les profils convergent souvent à une altitude de 30 m au-dessus du sol. C'est à ce niveau que le biais entre les deux profils est le plus faible ($-0,15$ °C, le radiosondage étant plus froid que les mesures sous ballon captif, voir l'annexe A). Une raison possible de cette différence est que les radiosondages n'étaient pas toujours lancés dans la prairie près de la station MET01, notamment la nuit. Ils étaient plutôt lancés à une centaine de mètres au nord, près d'un hangar à l'abri du vent et depuis une surface goudronnée. Une autre raison possible est le changement de conditions entre le moment du relevé des conditions initiales et le moment du lâcher (10 à 20 minutes). Il a donc été décidé d'ignorer les mesures dans les 30 premiers mètres réalisées par radiosondages.

Un aperçu des observations réalisées lors de cette campagne est présenté dans la partie suivante.

3.2 Aperçu des conditions pendant l'expérience

3.2.1 Conditions météorologiques sur la période décembre 2016 – mars 2017

Les 12 premières semaines de la campagne, du 17 décembre 2016 au 11 mars 2017, sont présentées dans cette partie. C'est la partie de l'expérience qui s'est déroulée au cœur de l'hiver et pendant laquelle les 14 POI de la campagne ont été réalisées. Les météogrammes des figures 3.5, 3.6 et 3.7 permettent d'illustrer les conditions météorologiques rencontrées pendant ces 12 semaines. Leur analyse permet de scinder cette période en trois sous-périodes.

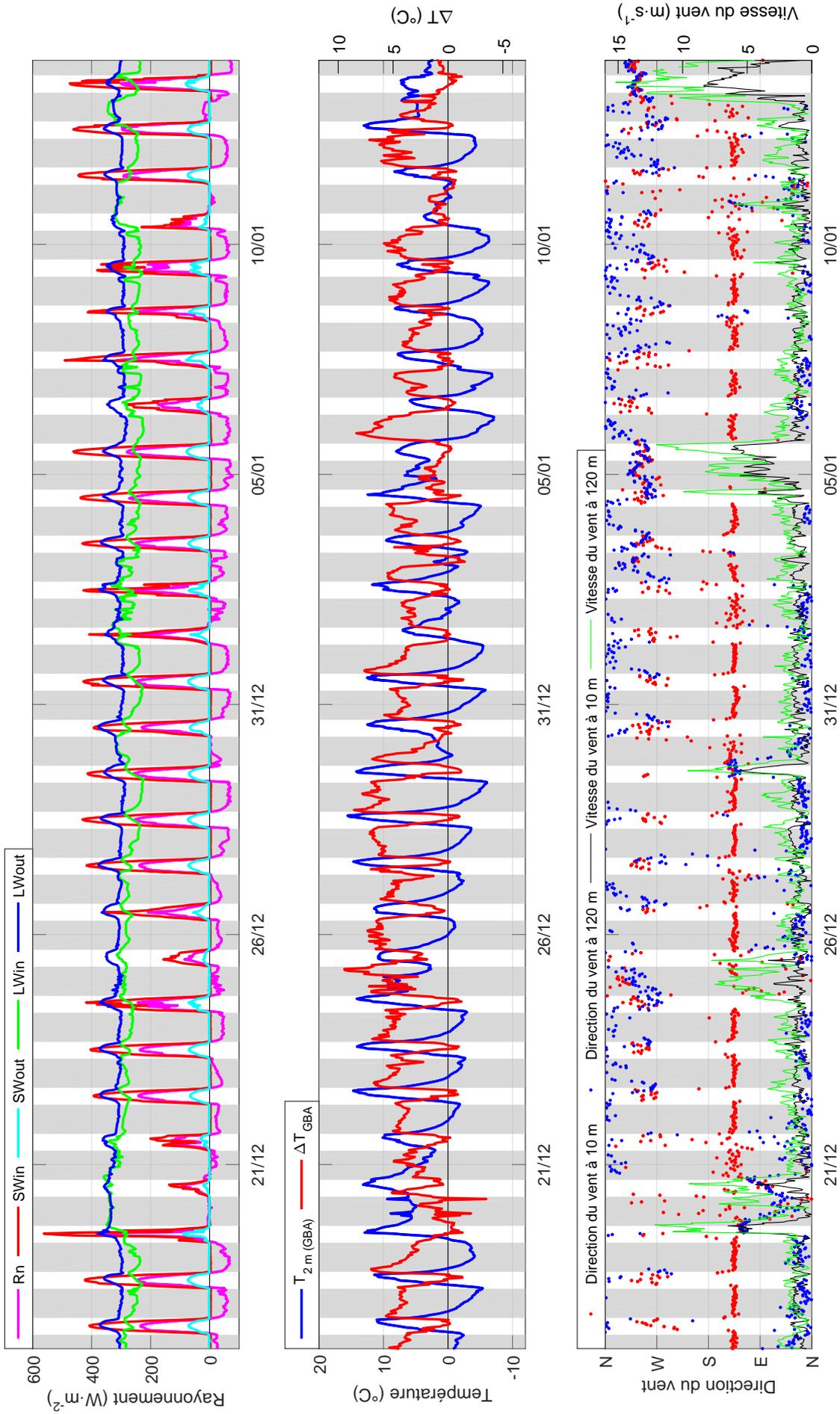


FIGURE 3.5 : Météogrammes des conditions enregistrées pendant les quatre premières semaines de mesure, du 17 décembre 2016 au 14 janvier 2017. Les différents rayonnements observés sur la station AS1 sont représentés en haut, les mesures de température (à 2 m et la différence entre 110 m et 2 m) enregistrées sur la station GBA sont représentées sur le graphe du milieu et les mesures de direction et vitesse du vent à 10 m (station MET01) et 120 m (station SOD2) sont visibles sur le graphe du bas. Toutes les données sont moyennées sur 30 minutes. Les périodes grisées représentent les nuits.

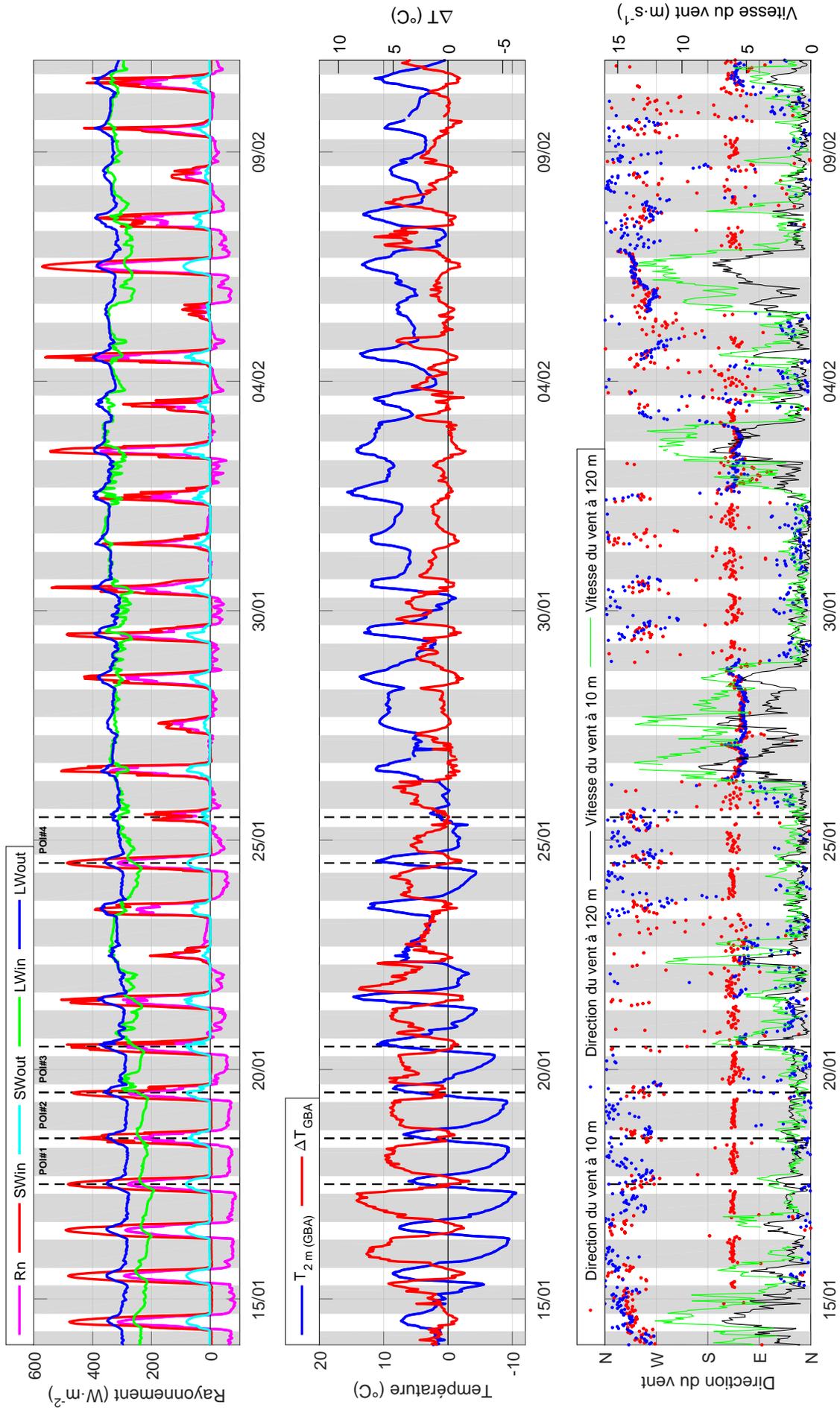


FIGURE 3.6 : Idem figure 3.5 pour la période du 14 janvier 2017 au 11 février 2017. De plus, les périodes des différents POI sont indiquées entre les lignes tiretées noires et leurs numéros sont reportés sur le graphe du haut.

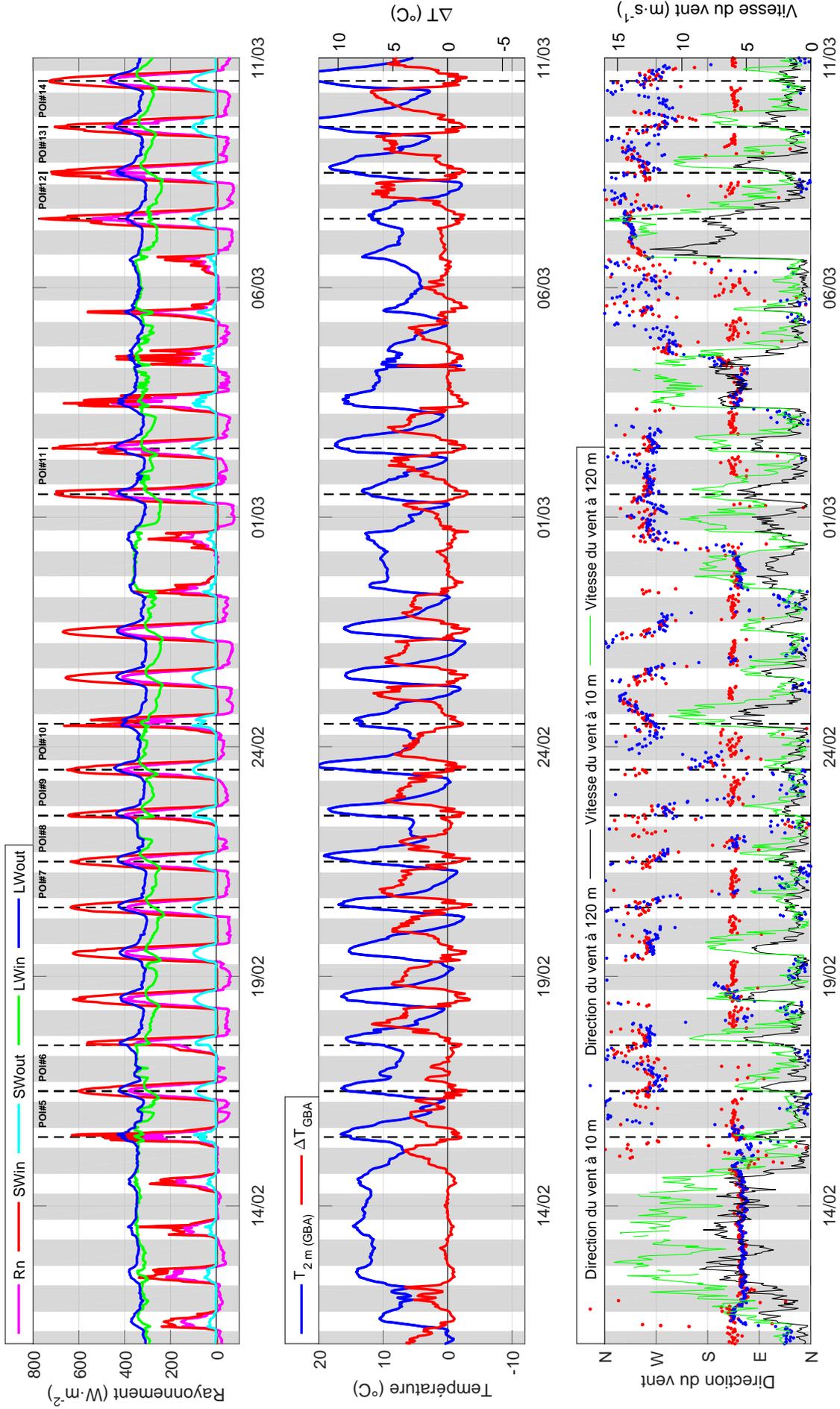


FIGURE 3.7 : Idem figure 3.6 pour la période du 11 février 2017 au 11 mars 2017.

La première sous-période, du 17 décembre 2016 au 22 janvier 2017, a été marquée par une forte occurrence de conditions favorisant le développement de la stabilité pendant la nuit. Les épisodes de vent fort (Mistral d’ouest-nord-ouest ou vent de sud-est) ont été rares et les vents ont été globalement calmes. De plus, les mesures de rayonnement permettent de constater qu’il y a eu peu de couverture nuageuse. En effet, en journée le rayonnement SW_{\downarrow} prend une forme en cloche typique de l’évolution au cours d’une journée avec un ciel dégagé, et la nuit la différence entre les rayonnements LW_{\uparrow} et LW_{\downarrow} permet également de constater l’absence de couverture nuageuse conduisant à la perte d’énergie par le sol et au refroidissement rapide de l’air à son contact. La plupart des journées de cette première sous-période suivent ce schéma (le DTR, pour diurnal temperature range, qui est la différence entre les températures maximales et minimales journalières, est fréquemment supérieur à 15°C) et des conditions stables sont observées la nuit, marquées par une stratification en température dont l’intensité maximale atteint fréquemment 4 à 5°C entre 110 m et 2 m . Des vents de vallée dans les vallées de Cadarache et de la Durance sont alors fréquemment observés la nuit. Il est à noter que les trois premières POI ont été effectuées pendant cette période, du 17 janvier 2017 12h00 UTC au 20 janvier 2017 12h00 UTC.

La deuxième sous-période, du 22 janvier 2017 au 15 février 2017, a été marquée par le passage de plusieurs perturbations entraînant la présence fréquente de nuages. Le refroidissement nocturne a ainsi été beaucoup moins fort que pendant la première sous-période, avec des valeurs de DTR à 2 m fréquemment comprises entre 5 et 10°C . La stratification en température a ainsi également été impactée puisque son intensité est en général bien moindre. Malgré ces conditions, un vent de vallée est fréquemment apparu dans la vallée de Cadarache pendant la nuit par vents faibles. La POI 4 a été réalisée pendant cette période, du 19 au 20 janvier 2017.

Enfin, 10 POI ont été effectuées pendant la troisième sous-période, du 15 février 2017 au 11 mars 2017, qui a été plus favorable aux observations des conditions stables, avec des DTR et intensités de stratification similaires à la première sous-période.

Les 14 POI ont été réalisées entre le 17 janvier 2017 et le 10 mars 2017 (voir le tableau 3.2). Elles ont rencontré différentes conditions météorologiques qui ont impacté le développement de la stratification dans la vallée de Cadarache. Les conditions rencontrées pendant les trois premières POI étaient très favorables pour l’observation des conditions stables. La stratification s’est chaque nuit développée rapidement et s’est maintenue durant toute la nuit. La fin de nuit de la POI 4 a été marquée par un épisode pluvieux qui a détruit la stratification. La stratification a également été partiellement détruite en fin de nuit lors des POI 7 à 10 par l’apparition de brouillard. Les débuts de nuit des POI 11 à 14 étaient marqués par de forts vents qui ont retardé le développement de la stratification.

Dans la partie suivante, les mesures réalisées lors de la campagne KASCADE 2017 sont introduites au travers de la présentation d’un phénomène particulier observé lors de la campagne.

3.2.2 Oscillations

Plusieurs cas d’oscillations de certains paramètres mesurés sur les stations AS ont été identifiés par conditions stables alors que la turbulence était faible. Il est à noter que ce phénomène est observé sur les cinq stations AS, ce qui signifie qu’il ne concerne pas uniquement la vallée de Cadarache (la station AS5 est située dans la vallée de la Durance), et sur différentes variables dont les deux composantes, la vitesse et la direction du vent, ainsi que plus rarement sur la

POI	Date	DTR	$\Delta T_{\text{GBA, max}}$	Remarques
# 01	17–18 janvier 2017	15,5 °C	5,8 °C	Très bonnes conditions
# 02	18–19 janvier 2017	16,1 °C	5,3 °C	Très bonnes conditions
# 03	19–20 janvier 2017	15,8 °C	5,4 °C	Très bonnes conditions
# 04	24–25 janvier 2017	14,2 °C	3,4 °C	Pluie en fin de POI (environ 05h00 UTC)
# 05	15–16 février 2017	16,1 °C	3,4 °C	Très bonnes conditions
# 06	16–17 février 2017	9,6 °C	2,0 °C	Couvert à partir de 21h00 UTC puis brouillard en fin de nuit
# 07	20–21 février 2017	18,4 °C	6,2 °C	Très bonnes conditions
# 08	21–22 février 2017	15,8 °C	4,4 °C	Brouillard de 00h00 UTC à 09h00 UTC
# 09	22–23 février 2017	19,0 °C	5,7 °C	Brouillard lors des RS de 03h00 UTC et 06h00 UTC
# 10	23–24 février 2017	15,7 °C	4,8 °C	Couvert à partir de 21h00 UTC
# 11	1 ^{er} –2 mars 2017	16,3 °C 16,3 °C	5,4 °C 5,4 °C	Vent fort généralisé en début de nuit puis en altitude le reste de la nuit
# 12	7–8 mars 2017	16,5 °C	6,9 °C	Idem POI # 11
# 13	8–9 mars 2017	17,1 °C	6,5 °C	Idem POI # 11
# 14	9–10 mars 2017	18,1 °C	7,0 °C	Idem POI # 11

Tableau 3.2 : Récapitulatif du contexte météorologique des 14 POI réalisées pendant la campagne KASCADE 2017. Chaque POI débute à 12h00 UTC et se termine à 12h00 UTC le lendemain. Le DTR représente le refroidissement observé sur la station GBA à 2 m pendant la POI, il est donc calculé avec la température maximale enregistrée l’après-midi qui précède la nuit de la POI. Le $\Delta T_{\text{GBA, max}}$ représente la stratification maximale en température enregistrée pendant la POI à la station GBA (110 m – 2 m).

température, mais pas sur la vitesse verticale. Un exemple de ces oscillations est présenté sur la figure 3.8 qui sera commentée plus loin.

3.2.2.a Précédentes observations et analyses d’oscillations pour un vent de vallée

De précédentes études menées sur d’autres sites ont rapporté la présence d’oscillations du vent de vallée. [Van Gorsel et al. \(2003\)](#) ont justifié la formation des oscillations observées dans la Riviera Valley (Alpes suisses) par la théorie énoncée par [Fleagle \(1950\)](#). Cette théorie prévoit que le vent drainant l’air froid dans la vallée entraîne un chauffage de cet air par compression pendant la descente, réduisant ainsi la différence de température et le gradient horizontal de pression responsable de la formation du vent de vallée. Il arrive alors un point à partir du quel le chauffage par compression devient dominant par rapport au refroidissement radiatif, créant une inversion du gradient de pression. L’écoulement lié au vent de vallée ralentit alors, de même que la friction, permettant au refroidissement radiatif de dominer de nouveau l’effet de réchauffement par compression, et renforçant alors le gradient de pression et le vent de vallée. Ce cycle peut alors se répéter ainsi. Le modèle théorique proposé par [McNider \(1982\)](#), basé sur la théorie de [Fleagle \(1950\)](#), prévoit que la période d’oscillation du vent catabatique (pour la composante le long de la vallée) vaut $T = 2\pi/N \sin \alpha$, avec N la fréquence angulaire de Brunt-Väisälä et α la pente de la vallée.

[Pinto et al. \(2006\)](#) ont également observé des oscillations dans les vitesses du vent descendant la vallée de Salt Lake (Utah, USA) pendant l’expérience VTMX. Le modèle théorique de

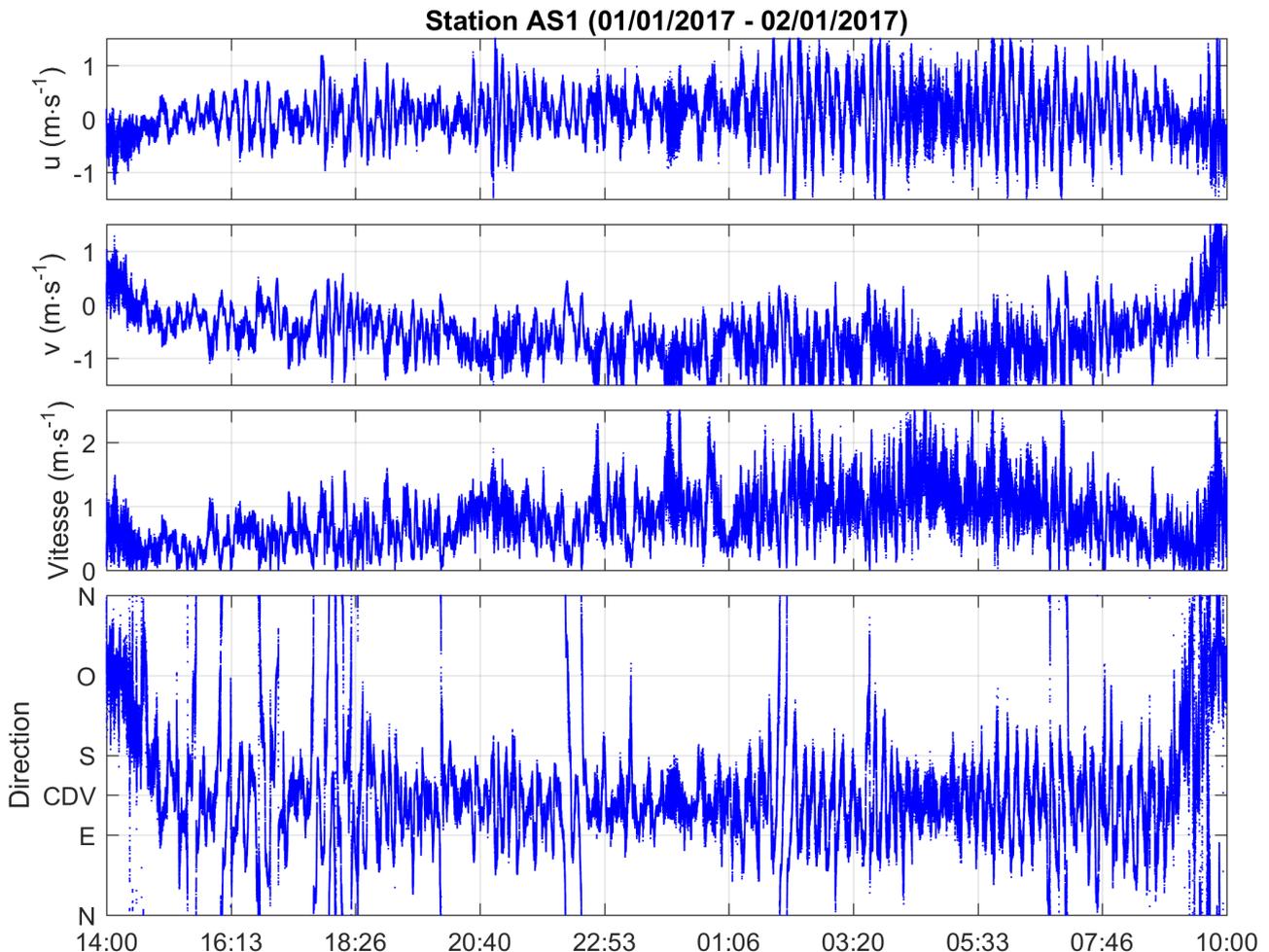


FIGURE 3.8 : De haut en bas, évolutions temporelles des composantes u , v , de la vitesse et de la direction du vent mesurées sur la station AS1 à une fréquence de 10 Hz durant la nuit du 1^{er} au 2 janvier 2017. CDV sur le graphe du bas indique la direction du vent qui descend la vallée de Cadarache.

McNider (1982) permet alors d'en prédire de manière satisfaisante la période d'oscillation (Princevac et al., 2008). Ils ont attribué l'origine de ces oscillations à la présence d'ondes internes de gravité. Gryning et al. (1985) ont aussi observé de telles oscillations caractérisées par une période qui correspond à celle des ondes internes de gravité définie comme $T = \frac{2\pi}{N} \sqrt{1 + (L/H)}$ (Haltiner, 1971), avec L une échelle de longueur horizontale et H une hauteur caractéristique de la profondeur de la vallée. LARGERON et al. (2013) ont simulé à la fois des oscillations dont la période répond au modèle théorique de McNider (1982) et d'autres (sur la composante verticale) produites par les ondes internes de gravité générées par le vent catabatique.

Porch et al. (1991) ont également mis en évidence la présence d'oscillations dans les vitesses d'un vent de vallée descendant dans la vallée Kimball Creek (Colorado, USA). Ces oscillations ont une période de 20 minutes. Dans ce dernier cas, une interaction entre le vent d'une vallée principale et les écoulements dans une de ses vallées affluentes est l'hypothèse avancée pour expliquer ces oscillations. La corrélation négative trouvée entre les vitesses de vent dans les deux vallées est interprétée comme une alternance dans la présence de ces vents. L'écoulement d'air froid dans la vallée principale empêche alors l'air froid de la vallée affluente de s'écouler, ce qui produit une accumulation de cet air dans le bas de celle-ci (vitesses fortes dans la vallée principale et faibles dans la vallée secondaire). L'air froid bloqué finit par s'écouler dans la vallée principale bloquant ainsi à son tour l'écoulement provenant de l'amont dans la vallée principale (vitesses faibles dans la vallée principale et fortes dans la vallée affluente).

3.2.2.b Oscillations observées pendant KASCADE 2017

Dans la vallée de Cadarache, les fluctuations sont étudiées à partir des données haute fréquence des stations équipées d'anémomètres soniques. Les composantes u et v relatives à ces stations, qui sont présentées dans cette partie, ne sont pas orientées suivant la convention météorologique, mais suivant la géométrie de la vallée de Cadarache. Ainsi, u est la composante du vent transverse à l'axe de la vallée (l'axe x associé est orienté positivement vers le sud-ouest) et v est sa composante dans la direction suivant son axe (l'axe y associé est orienté positivement vers le sud-est).

Oscillations versus fluctuations

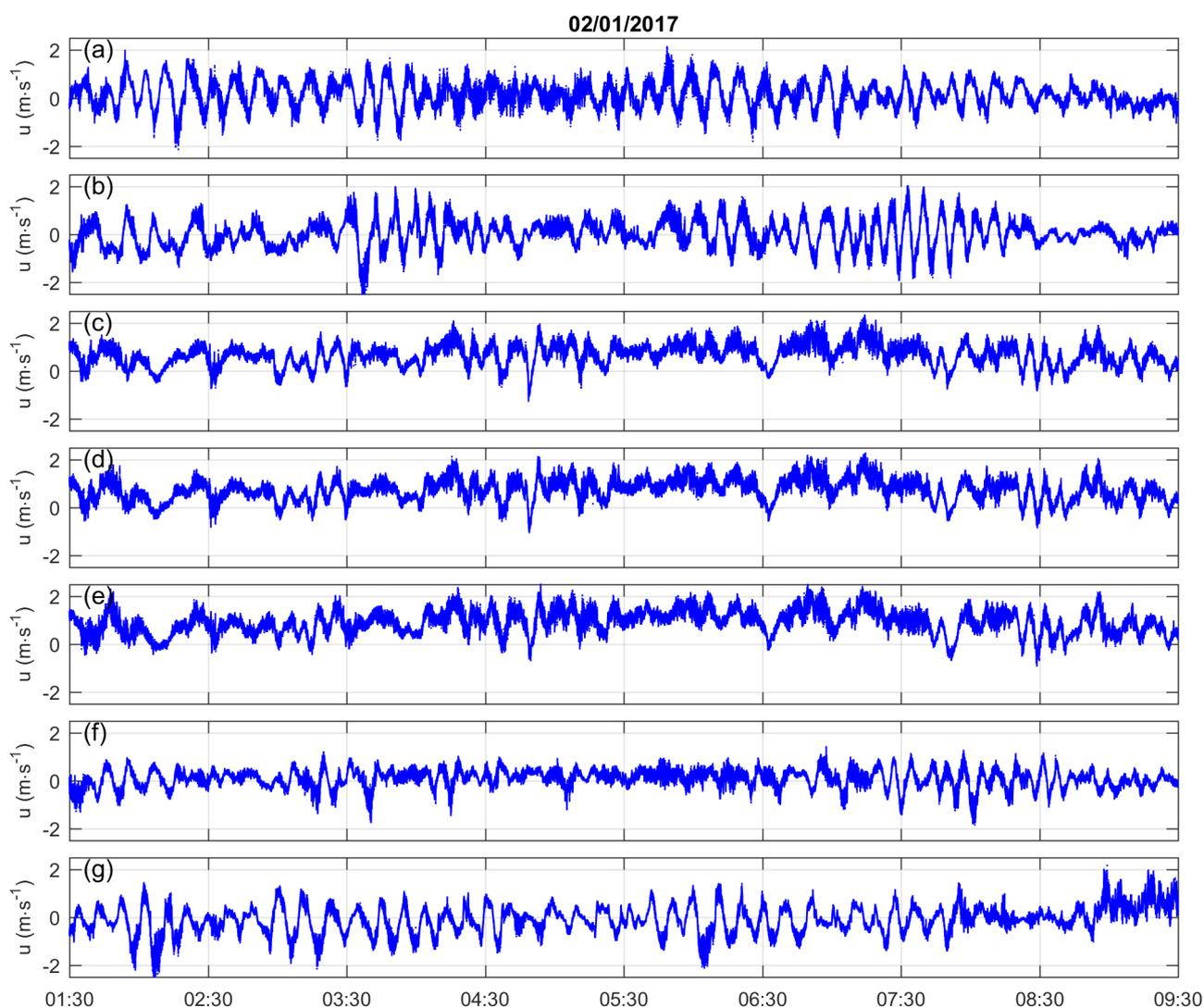


FIGURE 3.9 : Évolution temporelle de la composante u mesurée durant la nuit du 1^{er} au 2 janvier 2017 sur les stations AS1 (a), AS2 (b), AS3 à 2 (c), 5 (d) et 10 m (e), AS4 (f) et AS5 (g), à une fréquence de 10 Hz pour les stations AS1 et AS3 et 20 Hz pour les stations AS2, AS4 et AS5.

Dans certains cas, les fluctuations se manifestent sous la forme d'oscillations avec une période bien définie, alors que d'autres événements sont marqués par des fluctuations moins

régulières. Par exemple, les oscillations les plus franches ont été observées dans la nuit du 1^{er} au 2 janvier 2017, et ce sur les cinq stations AS (figure 3.9 pour la composante u). Dans ce cas, les oscillations franches observées sur la station AS1 (figure 3.8) affectent principalement la composante u et la direction alors que la composante v et la vitesse affichent des fluctuations moins régulières.

Ces fluctuations ne sont pas toutes ou pas toujours caractérisées par une période bien définie. Par exemple, pour la nuit du 1^{er} au 2 janvier 2017, les stations AS1, AS2 et AS5 affichent des fluctuations (pour la composante u) beaucoup plus franches que les autres stations pour lesquelles il est plus difficile de parler d'oscillations. Les amplitudes ne sont pas égales sur toutes les stations. Pour le cas présenté, elles sont plus grandes aux stations présentant les oscillations les plus franches. Il est à noter que les anémomètres de la station AS3 affichent des variations sur les trois niveaux de mesure (2, 5 et 10 m) avec des caractéristiques très proches les unes des autres.

De même, des oscillations (moins franches) ont été observées sur les composantes u et v pendant la première moitié de la nuit du 12 au 13 janvier 2017, mais les directions résultantes n'affichent pas d'oscillations, mais plutôt des fluctuations qui semblent être aléatoires entre environ 20h00 UTC et 01h00 UTC (figure 3.10). Le même phénomène se produit à la même période sur les quatre autres stations AS.

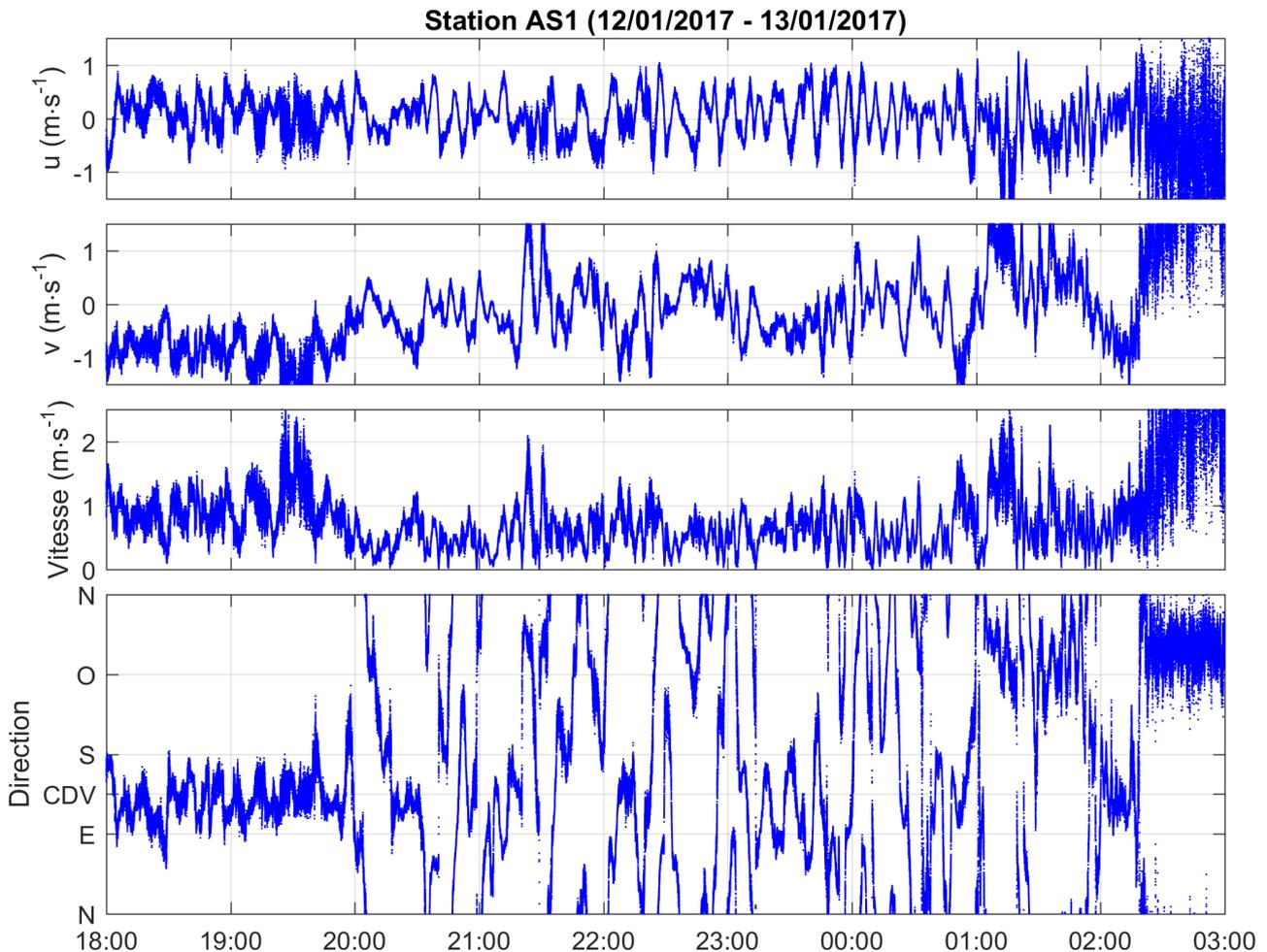


FIGURE 3.10 : Idem figure 3.8 pour la nuit du 12 au 13 janvier 2017.

Les conditions ont beaucoup évolué au cours de cette nuit du 12 au 13 janvier 2017. En

début de nuit et jusqu'à environ 20h00 UTC, les composantes u et v sont marquées par une turbulence modérée et présentent des oscillations avec une période d'environ 10 minutes (de même pour la vitesse du vent). La composante u oscille autour de zéro alors que les valeurs de v restent négatives, ce qui produit des directions oscillant avec une amplitude d'environ 90° autour de la direction moyenne du vent de vallée (vent CDV). Ensuite, jusqu'à environ 01h00 UTC, la turbulence est plus faible et v commence à osciller autour de zéro comme le fait u (la vitesse du vent est également marquée par des oscillations). Ces oscillations ont alors pour effet de produire des vents aux directions très variables sans pour autant que des oscillations apparaissent de manière évidente. La direction est de nouveau marquée par des oscillations entre 01h00 UTC et environ 02h20 UTC puis l'écoulement devient turbulent avec une direction nord-ouest à la faveur d'une augmentation de la vitesse liée à un épisode de Mistral (Mistral également visible sur les figure 3.5 et 3.6).

Oscillations sur les stations LEMS

Les stations LEMS ne permettent généralement pas d'observer ces oscillations de manière aussi précise que les stations AS, probablement parce qu'elles se produisent par vent calme ce qui ne favorise pas la rotation de la girouette des LEMS. Cependant, la LEMS B1 a enregistré quelques oscillations pendant la nuit du 1^{er} au 2 janvier 2017 (figure 3.11). Cette station était située à quelques mètres de l'AS1 (environ 5 m) ce qui permet d'évaluer la restitution du phénomène par la station LEMS.

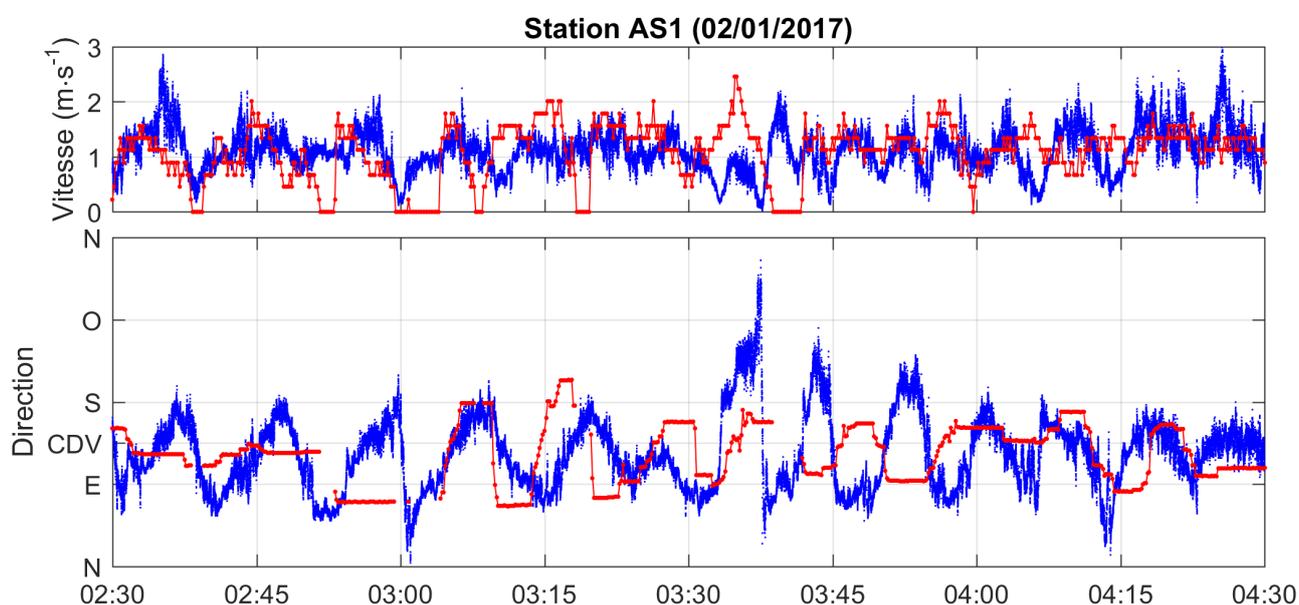


FIGURE 3.11 : Idem figure 3.8 pour la station AS1 (mesures à 10 Hz, en bleu) et la LEMS B1 (fréquence de mesure de 10 secondes, en rouge) pendant la nuit du 1^{er} au 2 janvier 2017.

Pour la période présentée sur la figure 3.11, les oscillations de la vitesse sont globalement bien mesurées par la LEMS à l'exception d'un pic non vu à environ 03h40 UTC et d'un déphasage entre environ 04h05 UTC et 04h15 UTC. Concernant la direction, quelques changements sont visibles sur la première demi heure (mais pas d'oscillations), puis des oscillations apparaissent entre 03h00 UTC et 04h30 UTC. Ces oscillations sont modérément bien corrélées avec les observations de l'anémomètre sonique et présentent même un déphasage d'une demi-période entre environ 03h40 UTC et 04h00 UTC. Ce décalage intervient après une courte période pendant laquelle la vitesse mesurée sur la LEMS est nulle, mais les vitesses sont ensuite supérieures

à 1 m s^{-1} ce qui est du même ordre de grandeur que sur le reste de la période présentée pour laquelle les oscillations sont mieux synchronisées. Ce déphasage ne peut donc pas être attribué à des vitesses de vent faibles.

Oscillations avec une période courte

La période des oscillations est généralement d'environ 10 minutes. Cependant, il y a également quelques épisodes pour lesquels la période est plus courte, comme le montre la comparaison entre les oscillations observées durant les nuits du 1^{er} au 2 janvier 2017 et du 11 au 12 janvier 2017 sur la station AS1 (figure 3.12). La période d'oscillation est alors d'environ 10 minutes lors de la première nuit contre environ 5 minutes pendant la deuxième.

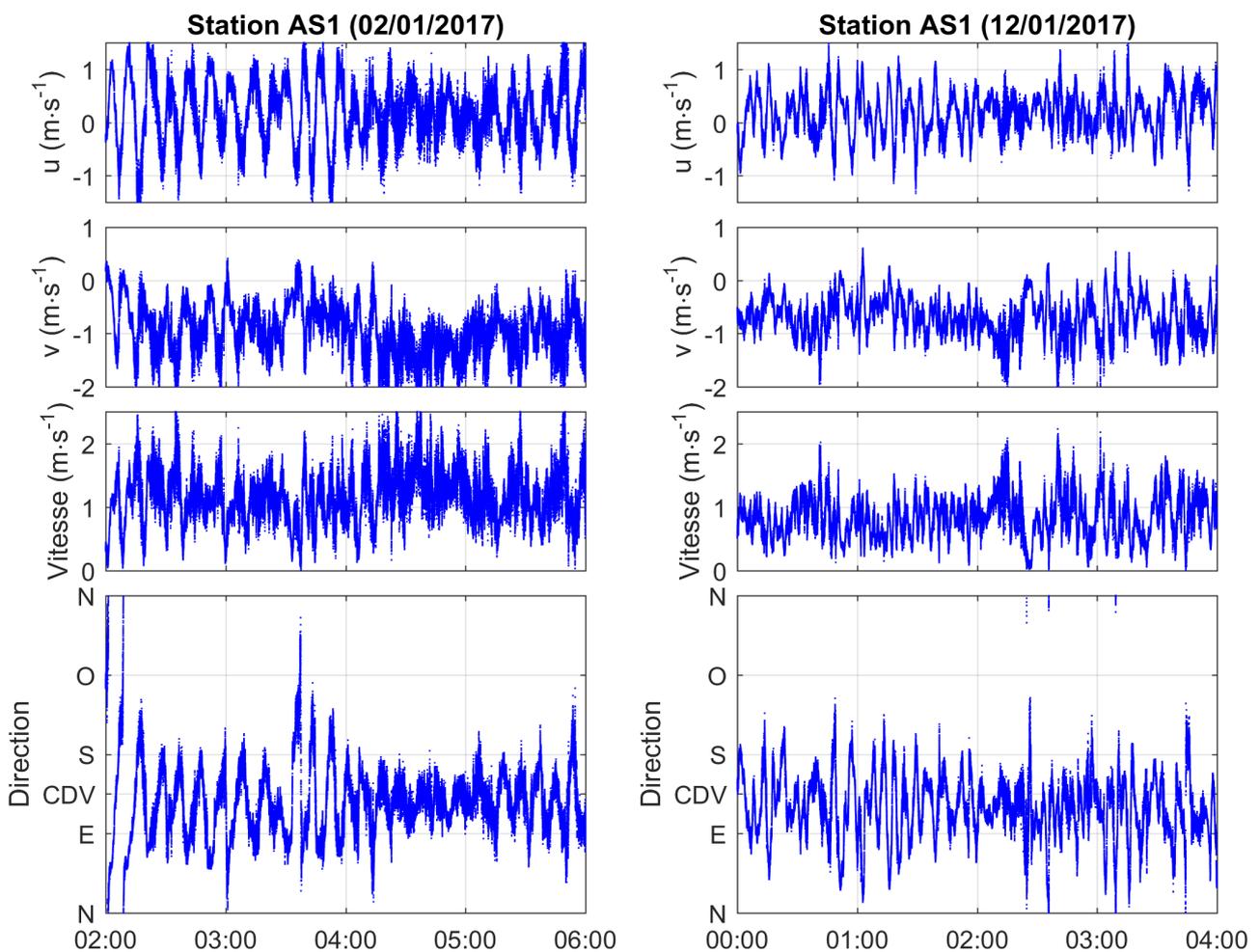


FIGURE 3.12 : Idem figure 3.8 pour la nuit du 1^{er} au 2 janvier 2017 à gauche et du 11 au 12 janvier 2017 à droite.

De la même manière, il est possible d'observer des oscillations relatives aux composantes u et v et de la vitesse ayant des périodes différentes. Par exemple, les oscillations des composantes u et v et de la direction du vent observées sur la station AS5 pendant la nuit du 1^{er} au 2 janvier 2017 (figure 3.13) ont une période d'environ 10 minutes, alors que la vitesse du vent oscille avec une période d'environ 5 minutes, ce qui s'explique par un déphasage d'une demi-période des oscillations des deux composantes qui produit un maximum pour la vitesse à chaque extremum rencontré par les composantes. Sur le cas présenté, les vents les plus forts prennent successivement une direction est ou ouest, et les transitions entre ces deux directions sont marquées par

des vents faibles (figure 3.14). Les transitions entre les deux directions sont généralement brutales, à l'exception de quelques épisodes associés à des vitesses qui diminuent moins fortement (transition vents d'ouest à est entre 03h00 UTC et 03h10 UTC puis transitions vents d'est à ouest entre 03h45 UTC et 03h50 UTC, entre 03h55 UTC et 04h00 UTC et entre 04h10 UTC et 04h20 UTC).

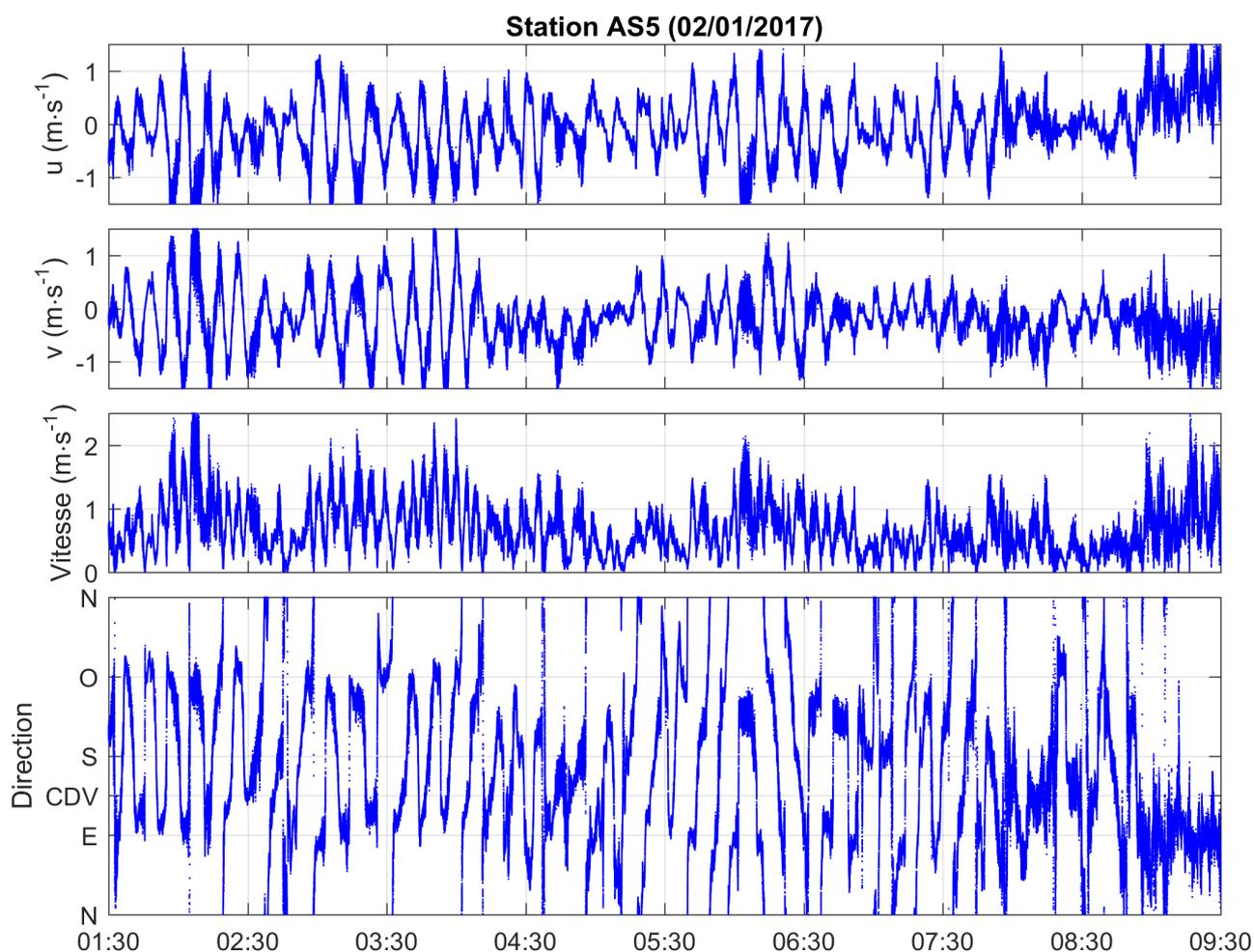


FIGURE 3.13 : Idem figure 3.10 pour la station AS5 pendant la nuit du 1^{er} au 2 janvier 2017 (mesures à 20 Hz).

Les observations réalisées sur les autres stations AS montrent le même comportement pour la nuit en question bien qu'il soit plus difficile de calculer précisément une période d'oscillation pour les vitesses sur ces stations à cause d'une turbulence légèrement plus importante et de l'irrégularité des fluctuations. Les amplitudes des oscillations sur les directions sont également moins importantes (généralement entre 90 et 120° contre environ 180° sur la station AS5) et centrées autour de la direction sud-est (v oscille donc en restant négatif alors que les oscillations de u sont centrées autour de zéro) qui correspond à la direction moyenne du vent descendant la vallée.

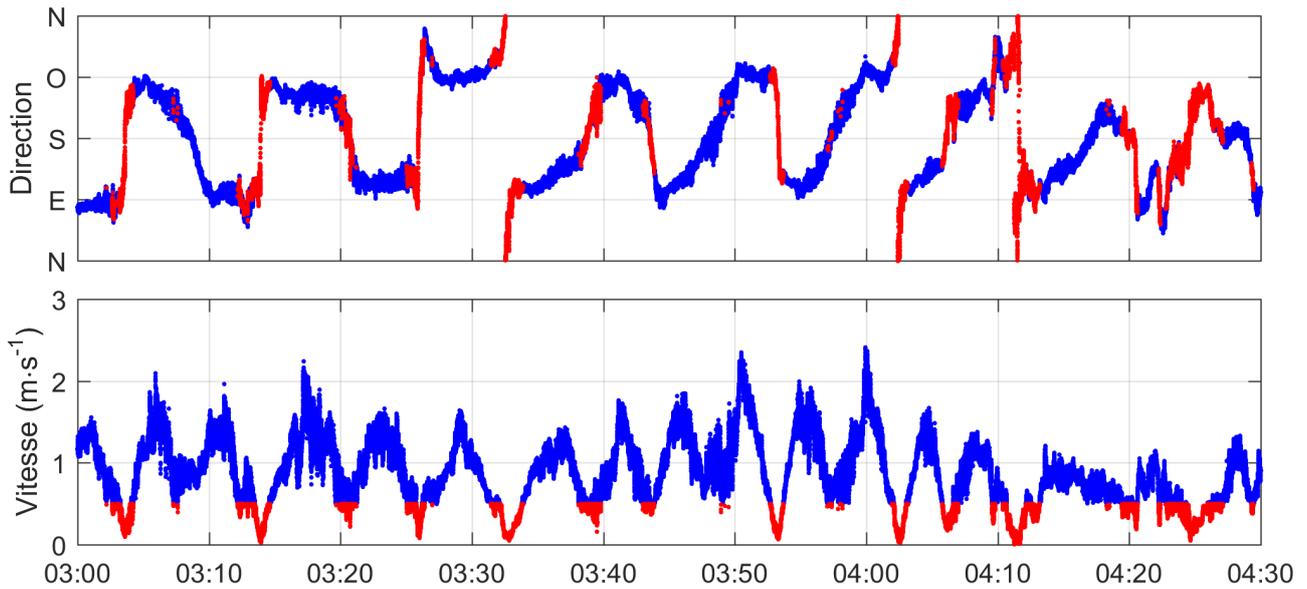


FIGURE 3.14 : Idem figure 3.13 pour la direction (en haut) et la vitesse (en bas). Les points rouges correspondent aux données pour lesquelles les vitesses sont inférieures à $0,5 \text{ m s}^{-1}$.

Oscillations avec une période longue

Un autre type de fluctuations synchrones de la température sonique, de la vitesse et de la direction du vent a été observé, comme le montre la figure 3.15 pour la nuit du 23 au 24 décembre 2016 sur la station AS1. Ces fluctuations correspondent à une intermittence du vent descendant la vallée de Cadarache (zones grisées sur la figure 3.15). Ces intermittences semblent être bien corrélées avec une diminution de la vitesse et de la température ainsi qu’une baisse de la turbulence.

Dans ce cas, la théorie d’un chauffage par compression énoncée par [Fleagle \(1950\)](#) semble pertinente. Dans notre cas, le vent de vallée est associé aux vitesses les plus fortes (observées durant la nuit du 23 au 24 décembre 2016 sur la station AS1) et est plus turbulent que dans les précédents exemples de fluctuations. La présence du vent de vallée semble associée au réchauffement de l’air par compression, puis, le ralentissement de l’écoulement (et de la friction, ce qui est visible avec la baisse de la turbulence pour les trois variables représentées sur la figure 3.15) entraîne un changement de direction du vent et diminue le chauffage de l’air par compression (le refroidissement radiatif redevient prédominant sur le réchauffement par compression d’où la baisse de la température). Ce phénomène est également visible sur la station AS2 et dans une moindre mesure sur les stations AS3 et AS4. Il est à noter que ces fluctuations se répètent avec une période beaucoup plus importante (généralement environ 50 à 60 minutes pour la nuit présentée sur la figure 3.15) que dans les cas d’oscillations présentés précédemment.

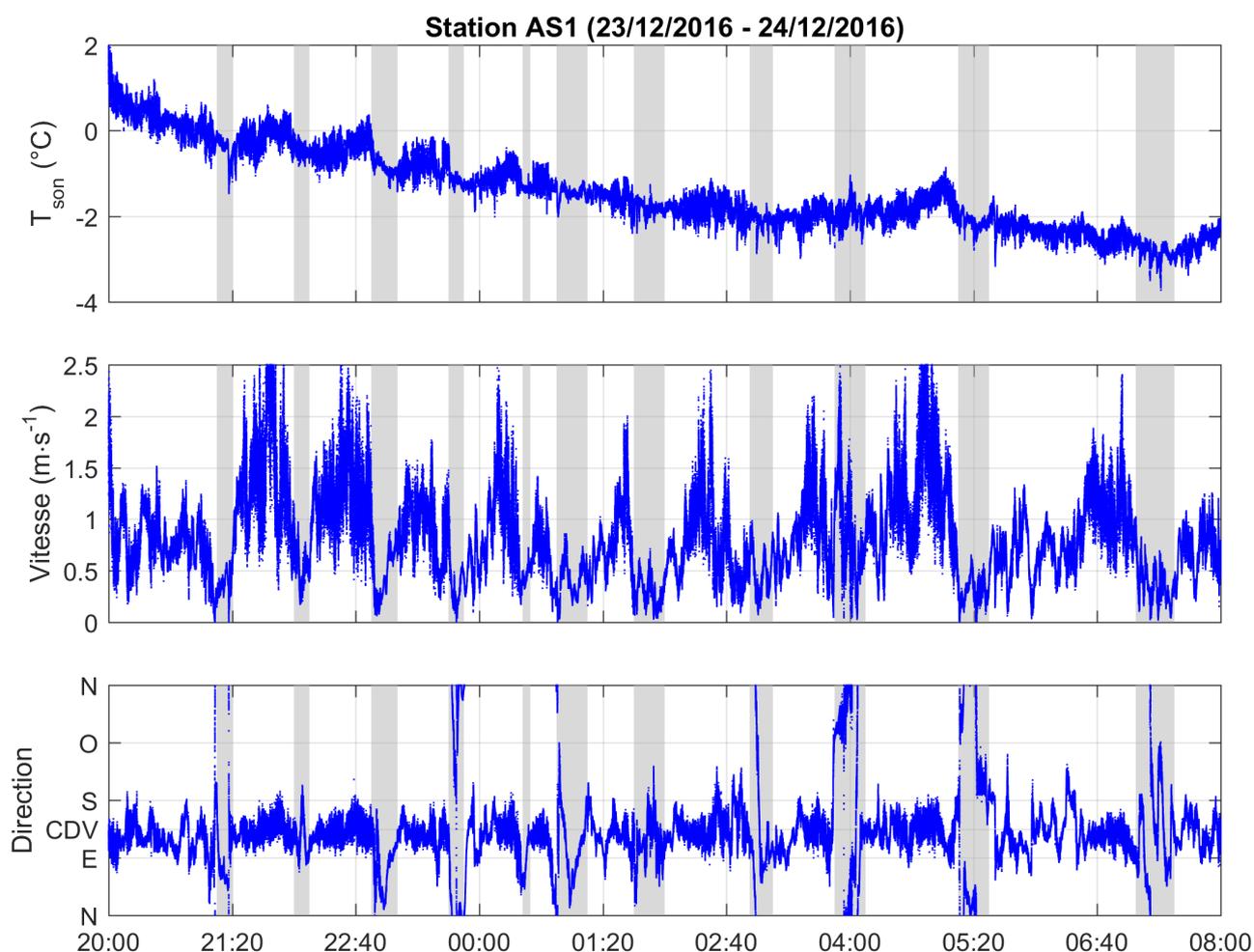


FIGURE 3.15 : Évolutions temporelles de la température sonique (en haut), de la vitesse du vent (au milieu) et de la direction (en bas) mesurées durant la nuit du 23 au 24 décembre 2016 sur la station AS1 à une fréquence de 10 Hz. Les zones grisées matérialisent des intermittences dans la présence du vent CDV.

Les oscillations en résumé

Finalement, des cas d'oscillation ont été observés à de nombreuses reprises au cours de la campagne. Généralement, on observe des oscillations de la direction du vent autour de l'axe descendant de la vallée pour les quatre stations du fond de la vallée. Elles sont observées sur tous les sites instrumentés avec anémomètre sonique bien qu'elles soient plus marquées sur les stations AS1, AS2 et AS5. Le site AS5 qui est situé en dehors de la vallée de Cadarache affiche des fluctuations de directions très importantes mais qui ne peuvent pas toujours être assimilées à des oscillations tant leur évolution peut parfois sembler aléatoire (fluctuations équivalentes à celles présentes entre 20h00 UTC et 01h00 UTC sur la station AS1 la figure 3.10). Ces oscillations ont le plus souvent une période proche de 10 minutes. Elles sont observées par conditions stables, alors que la turbulence est faible ou modérée, et peuvent perdurer sur de longues périodes, voire être présentes durant une nuit complète.

Les oscillations observées ont généralement des amplitudes maximales de 2 à 4 m s⁻¹ pour les composantes u et v , 2 à 3 m s⁻¹ pour les vitesses et de l'ordre de 100° pour les directions des stations AS1 à AS4, la station AS5 pouvant présenter des amplitudes plus importantes pour la direction.

Les trois théories permettant d'expliquer la génération d'oscillations dans une vallée, issues de la bibliographie, ne semblent pas s'appliquer au cas général des oscillations observées sur Cadarache. En effet, les oscillations de la composante verticale du vent relatives aux ondes de gravité n'ont pas été observées. Une alternance entre les courants de densité des deux vallées (Cadarache et Durance) serait principalement marquée par des oscillations de la vitesse ([Porch et al., 1991](#)), or nous avons aussi observé des oscillations sur la direction. Cependant, l'amplitude de ces oscillations (environ 90 à 100° dans la vallée de Cadarache) n'est pas cohérente avec le phénomène de chauffage par compression [Fleagle \(1950\)](#) pour lequel les changements de direction sont proches de 180° . Cette dernière théorie semble toutefois expliquer un cas d'intermittence du vent de vallée.

Cette partie a eu pour objectif de présenter un phénomène intéressant observé pendant la campagne KASCADE 2017. Son analyse reste néanmoins incomplète, notamment vis à vis des processus responsables de sa formation et son étude quantitative, vu la complexité des phénomènes en jeu.

Dans les parties suivantes, les observations de la campagne KASCADE 2017 sont utilisées pour étudier la structure de la stratification et les écoulements dans la vallée de Cadarache.

3.3 La CLA stable dans la vallée de Cadarache

3.3.1 Mesures de la stratification dans la vallée

Différents moyens de mesure étaient à disposition pour évaluer l'intensité de la stratification dans la vallée de Cadarache pendant la campagne KASCADE 2017.

- À la station GBA, la différence verticale de température entre 110 m et 2 m (notée $\Delta T_{110\text{ m}-2\text{ m}}$) est mesurée en sortie de vallée.
- Il est possible de calculer une stratification à partir des différentes stations LEMS placées sur les pentes de la vallée ($\Delta\theta_{v,LEMS}$).
- Enfin, il est également possible de calculer une stratification avec les radiosondages effectués pendant les POI (notée $\Delta\theta_{v,RS}$).

Le calcul exact d'un gradient de température potentielle virtuelle à partir des mesures de la station GBA (noté $\Delta\theta_{v,110\text{ m}-2\text{ m}}$) n'est pas possible du fait du manque de mesure de pression et d'humidité à 110 m. Une étude de sensibilité (voir l'annexe A) a montré que les paramètres les plus importants pour ce calcul (après la température) sont le gradient de pression et le gradient d'humidité. Or, l'analyse des profils de ballon captif réalisés lors de la campagne KASCADE 2013 a permis de montrer que la variation d'humidité dans la vallée de Cadarache sur les 110 premiers mètres est généralement très faible (en général inférieur à 1 g kg^{-1}), limitant ainsi l'influence de ce paramètre dans le calcul de l'intensité de la stratification.

La pression à 110 m peut être calculée en utilisant l'approximation hydrostatique, ce qui permet ensuite de calculer une différence verticale de température potentielle qui dans notre cas est une bonne approximation de $\Delta\theta_{v,110\text{ m}-2\text{ m}}$ (le biais et l'erreur absolue moyenne calculés sur les 789 profils de ballons captifs disponibles sont inférieurs à $0,1\text{ }^\circ\text{C}$).

L'utilisation du gradient adiabatique sec dans l'atmosphère pour la transformation isentrope ($\gamma = dT/dz = -9,8\text{ }^\circ\text{C km}^{-1}$) permet également de réaliser une bonne approximation de la valeur de $\Delta\theta_{v,110\text{ m}-2\text{ m}}$ (le biais et l'erreur absolue moyenne calculés sur les 789 profils de ballons captifs disponibles sont également inférieurs à $0,1\text{ }^\circ\text{C}$).

Ainsi, dans la suite, le gradient adiabatique sec constant est utilisé pour la transformation isentrope, donc :
$$\Delta\theta_{110\text{ m}-2\text{ m}} = \Delta T_{110\text{ m}-2\text{ m}} - \frac{\gamma \times (110 - 2)}{1000}.$$

3.3.2 Extension verticale de la CLA stable

Les radiosondages effectués pendant les POI permettent d'observer la stratification dans et au-dessus de la vallée de Cadarache. Ils atteignent en général au moins une hauteur d'environ 1500 à 2000 m au-dessus du sol ce qui est suffisant pour sonder toute la CLA mélangée pendant une journée d'hiver ainsi que la couche stable la nuit. Les stratifications les plus importantes sont observées lors des sondages effectués à 06h00 UTC, ce qui correspond au dernier sondage de la nuit (pour les premières POI en janvier) ou au premier sondage quelques instants après le lever du soleil (pour les dernières POI de mars).

Les profils de température potentielle virtuelle à 15h00 UTC (en début de POI) et à 06h00 UTC sont tracés sur la figure 3.16 pour six POI qui présentent des conditions particulièrement stables. Le sondage de 15h00 UTC est le dernier effectué avant le coucher du soleil. Il permet donc d'avoir une idée du profil attendu dans la couche résiduelle pendant la nuit suivante puisque les POI sont programmées quand les conditions synoptiques sont celles de

vents calmes et donc de faible advection (pas de passages de perturbations).

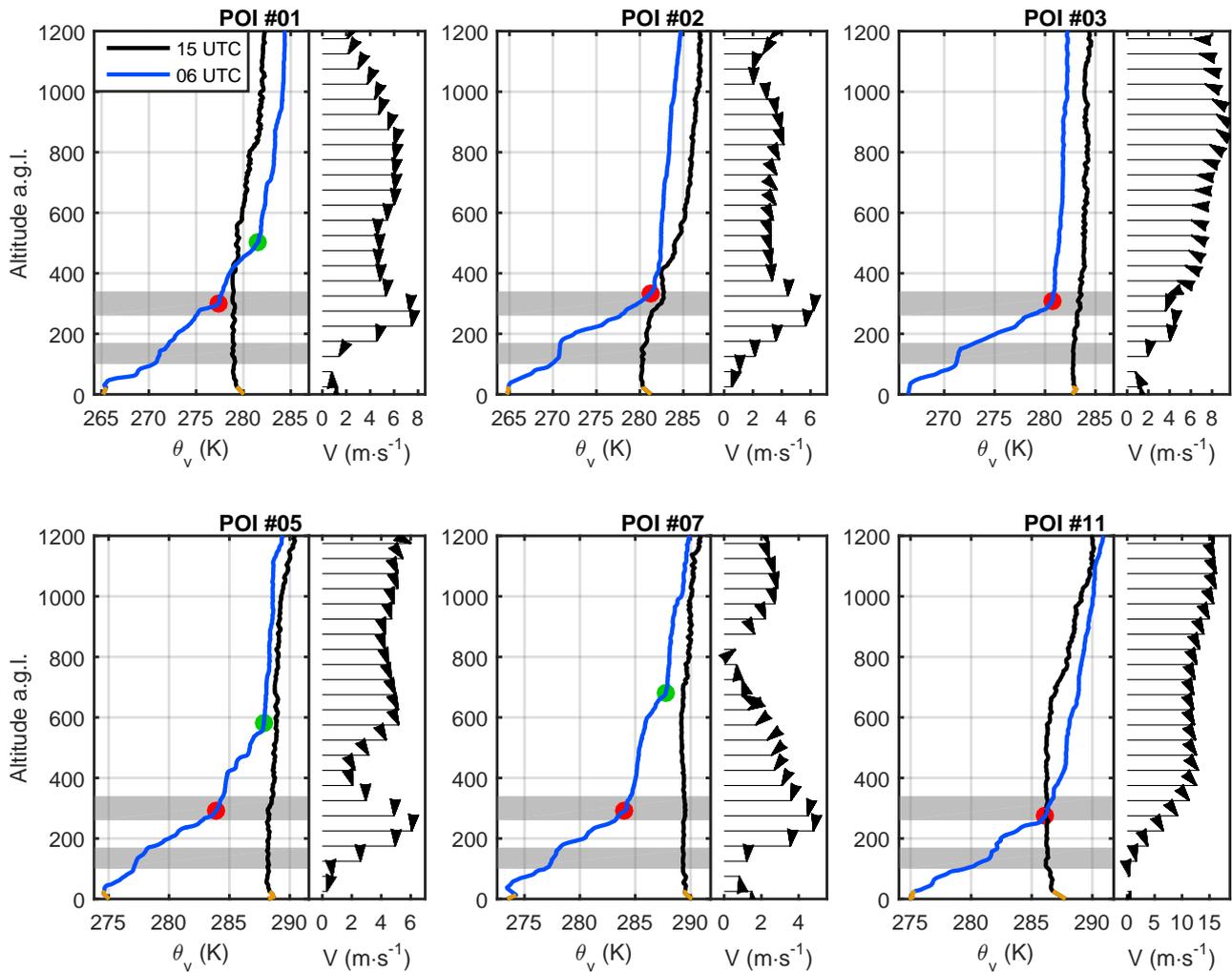


FIGURE 3.16 : Radiosondages typiques effectués lors des POI 1, 2, 3, 5, 7 et 11. Sur les parties gauches des figures sont tracés les profils verticaux de température potentielle virtuelle mesurée par radiosondage à 15h00 UTC (en noir) et à 06h00 UTC (en bleu). Les points rouges correspondent aux altitudes estimées (visuellement) de la hauteur de la couche stable. Les points verts montrent l'altitude à laquelle le sondage de 06h00 UTC rejoint celui de 15h00 UTC ce qui est un indicateur de la base de la couche résiduelle (les conditions synoptiques variant peu). Les bandes grises (entre 100 m et 170 m et entre 260 m et 330 m) matérialisent des couches dans lesquelles la stratification est plus faible. Les portions orange des radiosondages correspondent aux hauteurs inférieures à 30 m qui sont peu fiables. Sur les parties droites des figures, les traits correspondent au module du vent (à 06h00 UTC) et les pointes de flèche indiquent sa direction.

Les hauteurs de couche limite nocturne estimées visuellement en repérant les inversions les plus importantes sont toutes comprises entre 270 m et 330 m pour les 6 nuits présentées (points rouges sur la figure 3.16). Cependant, pour les POI 1, 5 et 7, on observe également une couche stratifiée de moindre intensité entre 300 m et 700 m (points verts sur la figure 3.16) où le profil de 06h00 UTC rejoint celui de la veille à 15h00 UTC, indiquant la base de la couche résiduelle et compliquant l'estimation visuelle de la hauteur de cette couche.

La couche limite nocturne comporte une zone où l'intensité de la stratification est modérée, entre 100 m et 170 m ce qui correspond à la zone de cisaillement entre les vents de vallée de Cadarache et de Durance. Cette diminution de stratification peut donc s'expliquer par une

augmentation du mélange causée par ce cisaillement de vent. Une seconde couche remarquable est située à une altitude comprise entre 260 m et 330 m et elle marque le sommet des couches stratifiées des 6 nuits présentées. Cette couche se situe au-dessus du jet du vent de Durance qui est présent entre 175 et 225 m (Duine et al., 2017). Ce jet est visible sur les sondages des POI 1, 2, 5 et 7 à une altitude comprise entre 200 et 250 m et avec des vitesses de 5–8 m s⁻¹. Le sondage de la POI 3 présente des vitesses trop importantes en altitude pour permettre la mise en place d'un jet associé au vent de vallée de Durance, bien que ce vent soit présent jusqu'à environ 250 m. Les vents forts de la POI 11 empêchent le vent de vallée de s'établir. En revanche, le vent de Durance est présent jusqu'à une altitude d'environ 300-400 m d'après les sondages des POI 2, 3, 5 et 7, ce qui correspond à l'altitude du sommet de la couche fortement stratifiée. Le vent de vallée de Cadarache n'est pas visible avec les radiosondages parce qu'il n'est présent que sur une fine couche (environ les 50 premiers mètres, Duine et al. (2017)) directement en contact avec le sol, là où les radiosondages ne sont pas précis (cf. annexe B).

3.3.3 Stratification dans la vallée de Cadarache

L'intensité de la stratification mesurée à partir des stations LEMS ($\Delta\theta_{v,LEMS}$) est calculée le long d'une pente puisque toutes les mesures de température par ces stations sont effectuées à 1,5 m du sol. Plusieurs combinaisons de stations peuvent alors être utilisées pour calculer cette intensité, parmi lesquelles la différence entre les stations positionnées aux altitudes extrêmes (LEMS B1 pour la plus basse et LEMS S4 pour la plus élevée) ou la différence entre les stations placées sur le segment transverse à l'axe de la vallée (LEMS B3 en bas et S4 ou N4 en haut).

Le profil de température particulier obtenu dans le fond de la vallée par conditions stables (figure 3.17 et partie 3.3.3.c) incite à prendre la LEMS B3 comme référence de mesure de température en fond de vallée parce que ses valeurs sont cohérentes avec le profil vertical attendu, contrairement aux observations de la LEMS B1. Le choix de la station haute (LEMS S4 ou N4) est plus arbitraire car les gradients calculés en utilisant l'une ou l'autre de ces stations sont très proches (figures 3.17). Finalement, la LEMS N4 est choisie parce que fournissant davantage de données, la LEMS S4 ayant été installée plusieurs semaines après le début des mesures. $\Delta\theta_{v,LEMS}$ est donc calculée à partir des LEMS B3 et N4.

3.3.3.a Stratification et hétérogénéité

Il existe deux causes permettant d'expliquer les différences de température potentielle virtuelles observées dans la vallée : la stratification et l'hétérogénéité du site. Lorsque l'atmosphère est stable, les différences d'altitudes des sites où sont installées les stations engendrent des différences de températures alors que dans une atmosphère neutre, bien mélangée, les températures potentielles virtuelles n'ont pas de dépendance à l'altitude. Avant de continuer l'étude de la stratification dans la vallée, cette partie vise à étudier les différences liées aux hétérogénéités indépendamment des effets de la stratification.

LeMone et al. (2003) ont présenté une méthode visant à supprimer les différences de températures liées à la stratification pour ne garder que celles liées aux hétérogénéités. Pour cela, la stratification globale est estimée par la régression linéaire calculée sur les températures potentielles virtuelles des différentes stations en fonction de leur altitude (une régression est calculée pour chaque période moyenne de 30 minutes). L'écart entre les observations (θ_v) et la régression linéaire ($\theta_{v,ze}$ en adoptant la notation de LeMone et al. (2003)) est attribué à l'hétérogénéité du

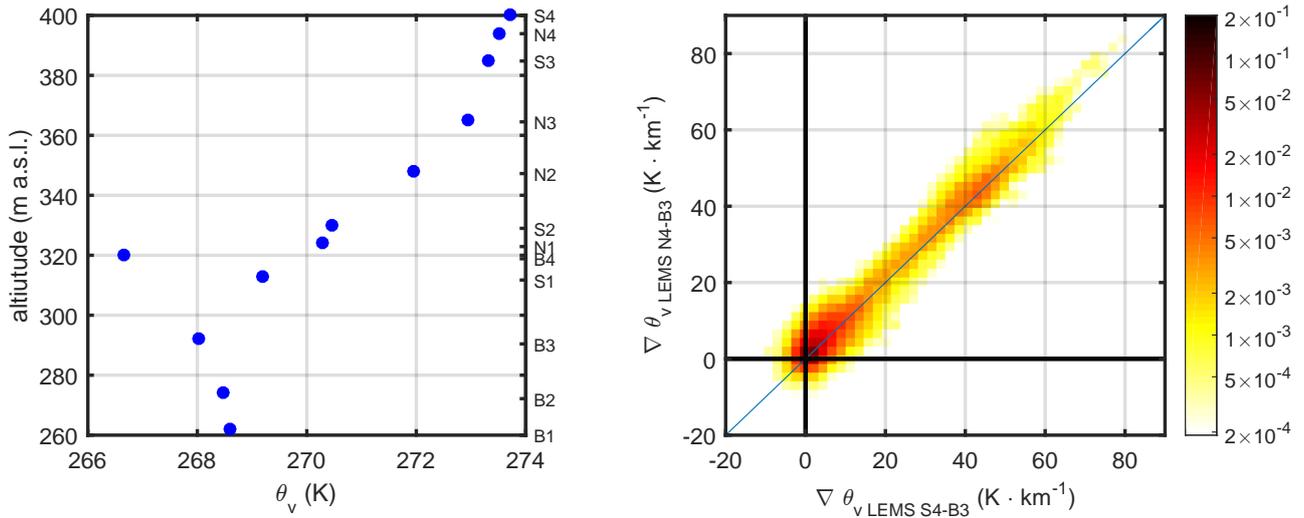


FIGURE 3.17 : À gauche : température potentielle virtuelle moyenne sur les 5 nuits du 15 au 20 janvier 2017 mesurée par les LEMS (les noms des stations sont reportés sur l'ordonnée à droite). À droite : comparaison des gradients mesurés sur les versants sud (LEMS S4 et B3) et nord (LEMS N4 et B3) de la vallée (données de toute la période KASCADE 2017 moyennées sur 30 minutes).

site (figure 3.18). L'hétérogénéité est ensuite quantifiée par l'écart-type de ces différences (σ_{θ_v}).

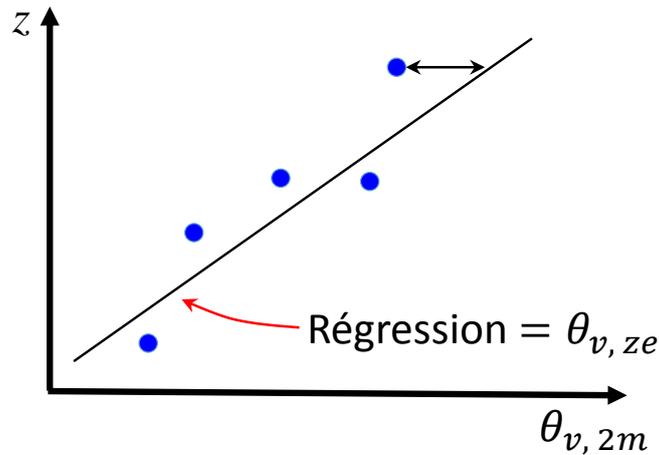


FIGURE 3.18 : Schématisation de la distinction entre la stratification et l'hétérogénéité contenue dans les températures potentielles virtuelles, d'après la méthode de [LeMone et al. \(2003\)](#). La pente de la régression est notée $\theta_{v,ze}$ pour reprendre la notation de [LeMone et al. \(2003\)](#).

La stratification particulière dans le fond de la vallée de Cadarache (figure 3.17 gauche) complique le calcul d'une régression linéaire qui soit représentative de la stratification dans la vallée. Le calcul de la régression a donc été basé uniquement sur les stations placées sur le segment transverse à l'axe de la vallée (les stations B1, B2 et B4 ne sont pas utilisées). En revanche, l'hétérogénéité est bien calculée sur l'ensemble des 12 LEMS. L'évolution de l'hétérogénéité sera étudiée au cours du cycle diurne ainsi qu'en fonction de la valeur de la stratification.

La stratification moyenne estimée par la régression est comparée à la valeur déduite de la différence de température potentielle virtuelle entre les stations N4 et B3 (figure 3.19). La pente de la régression est parfaitement corrélée à la différence de température potentielle virtuelle calculée à partir des LEMS : biais de $-0,6 \text{ K km}^{-1}$ ($\theta_{v,ze}$ plus faible que $\nabla\theta_{v,LEMS}$), erreur absolue moyenne de $3,1 \text{ K km}^{-1}$ et coefficient de corrélation de 0,97. Cela montre que l'intensité

de la stratification à 2 m dans la vallée de Cadarache calculée à partir des stations N4 et B3 est représentative de la stratification observée avec l'ensemble des stations placées sur le segment transverse à la vallée.

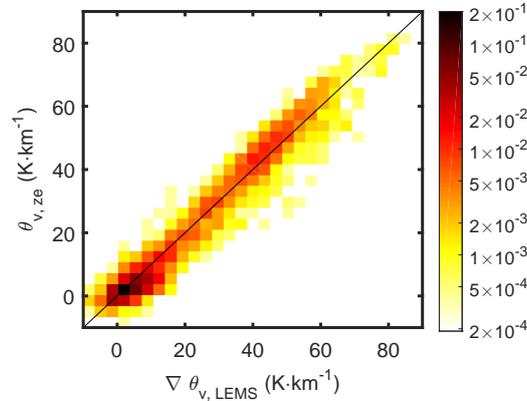


FIGURE 3.19 : Comparaison des intensités de stratification calculées à partir de la différence de température potentielle virtuelle entre les LEMS N4 et B3 ($\nabla\theta_{v,LEMS}$) et estimée à partir de la pente ($\theta_{v,ze}$) de la régression linéaire calculée pour les 9 LEMS placées sur l'axe transverse à la vallée. Les données de toute la période KASCADE 2017 moyennées sur 30 minutes ont été utilisées.

L'évolution de l'hétérogénéité en fonction de l'intensité de la stratification (figure 3.20) montre deux comportements particuliers. Le premier, pour les situations où l'intensité de la stratification est supérieure à 2K, montre une forte corrélation entre les deux paramètres. L'augmentation de la stratification entraîne une augmentation de la stabilité ce qui diminue le mélange non seulement vertical mais aussi horizontal. Il est donc logique d'observer une augmentation de l'hétérogénéité avec le renforcement de la stratification. Pour les stratifications les plus faibles (situations instables, neutres ou faiblement stables, soit pour $\Delta\theta_{v,LEMS} < 1,5$ K), il y a un découplage entre l'intensité de l'hétérogénéité et celle de la stratification.

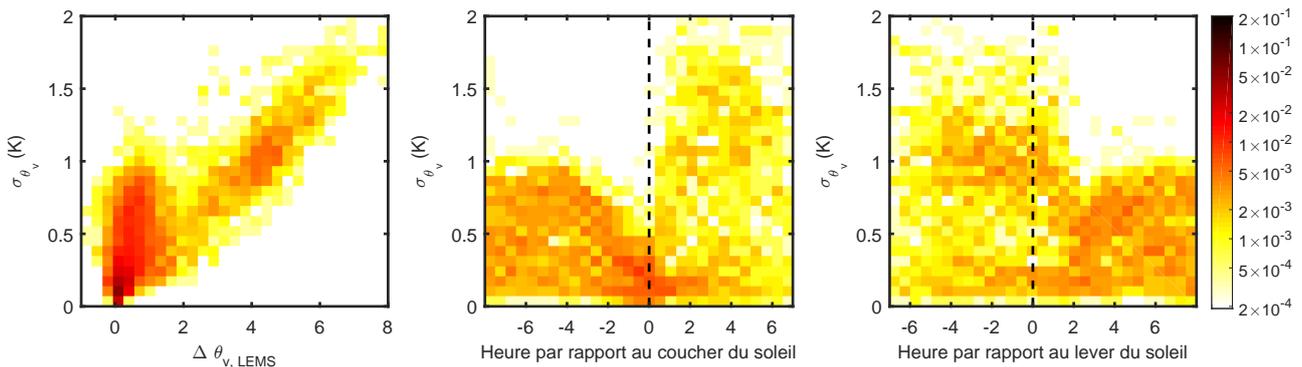


FIGURE 3.20 : Variation de l'hétérogénéité des températures potentielles virtuelles en fonction de l'intensité de la stratification mesurée sur les flancs de la vallée (à gauche), et en fonction du temps référencé soit au coucher du soleil (au centre), soit au lever du soleil (à droite). Les données de toute la période KASCADE 2017 moyennées sur 30 minutes ont été utilisées.

Le cycle diurne de l'hétérogénéité de la stratification est également présenté sur la figure 3.20. En pleine journée, de manière générale, il n'y a pas de lien entre l'intensité de l'hétérogénéité et le moment de la journée. Dès 4 heures avant la fin du jour, et jusqu'à celle-ci, les valeurs maximales diminuent progressivement pour atteindre environ 0,5K. C'est le signe du passage par la « neutralité » en fin de journée. Cela pourrait être lié à la diminution du flux

solaire reçu sur les différents sites en fin de journée (voir le trait bleu sur la figure 3.21) qui fait que la température de l'air près du sol est moins sensible aux hétérogénéités du terrain (pente, albédo, nature et humidité du sol, couvert végétal, ...) qu'en plein ensoleillement. En effet, bien que le relief soit complexe, les vallées sont peu profondes et peu pentues. Les ombres portées n'apparaissent alors qu'en fin de journée (environ 1 heure avant le coucher du soleil, voir le trait noir sur la figure 3.21) alors qu'en milieu de journée, les hétérogénéités du terrain suffisent à générer des différences importantes (les LEMS N1, N2 et N3 sont situées sur les pentes les mieux exposées au soleil et ce sont donc les stations qui enregistrent les températures les plus chaudes en journée).

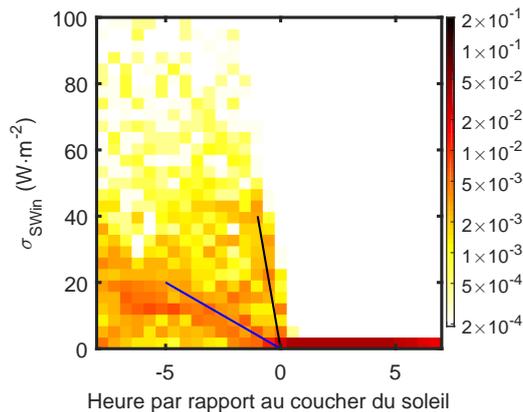


FIGURE 3.21 : Évolution de l'hétérogénéité du rayonnement SW_{\downarrow} autour du coucher du soleil. Le trait bleu repère une première tendance dans l'hétérogénéité spatiale du rayonnement, présente durant la journée, et le trait noir repère une deuxième tendance présente uniquement à l'approche du coucher du soleil. Les données de toute la période KASCADE 2017 moyennées sur 30 minutes ont été utilisées.

Pendant la nuit, les valeurs sont très dispersées ce qui s'explique par des intensités de stratification variables d'une nuit à l'autre. Cependant, la figure 3.20 montre une dépendance entre l'intensité de la stratification et celle de l'hétérogénéité. Par exemple, une CLA stable qui se développe rapidement après le coucher du soleil produit de fortes hétérogénéités dès le coucher du soleil alors que dans le cas d'une CLA stable qui se développe plus lentement, les hétérogénéités apparaissent également progressivement.

Les hétérogénéités les plus fortes sont rencontrées pendant la nuit, ce qui est logique puisque ce sont les conditions stables qui favorisent les fortes hétérogénéités pour les raisons expliquées auparavant. Les fortes stratifications persistent 1 à 2 heures après le lever du soleil tout comme l'hétérogénéité.

La contribution à l'hétérogénéité de chaque site considéré individuellement est représentée sur la figure 3.23 (la position des stations est rappelée sur la figure 3.22). En journée, les hétérogénéités ne dépassent que très rarement $1^{\circ}C$ sur l'ensemble des stations à l'exception des LEMS N1, N2 et N3. Les stations les plus chaudes dans la journée sont justement les LEMS N1, N2 et N3 qui bénéficient d'une meilleure exposition au soleil. Les autres LEMS présentent des températures généralement plus froides que celle prévue par la régression linéaire. La diminution de l'hétérogénéité en fin de journée jusqu'au coucher du soleil, vue sur la figure 3.20, est également clairement visible sur toutes les stations.

Pendant la nuit, les tendances sont moins bien marquées qu'en journée puisque les températures des différentes stations peuvent tour à tour être plus chaudes ou plus froides que les températures prévues par la régression linéaire. Cependant, le comportement particulier de

la stratification dans le fond de la vallée (cf. figure 3.17) explique que les températures soient généralement plus basses que le profil de la stratification sur la LEMS B4 et qu'elles soient plus fortes sur les LEMS B1 et B2. Les autres stations ne font pas apparaître de tendance claire.

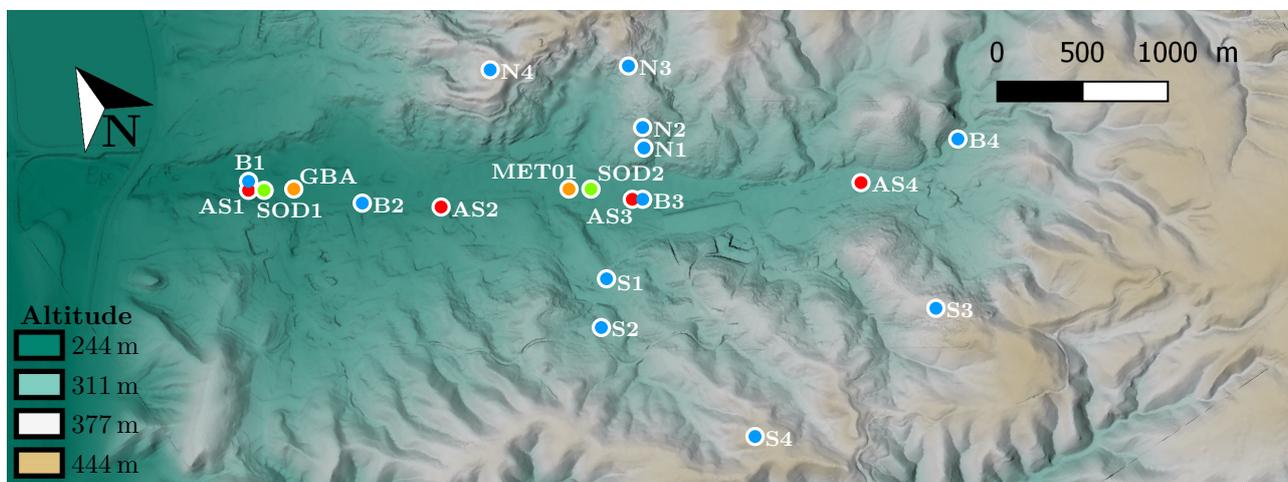


FIGURE 3.22 : Rappel de la position des stations dans la vallée de Cadarache (idem figure 3.1 avec une vue centrée sur la vallée de Cadarache).

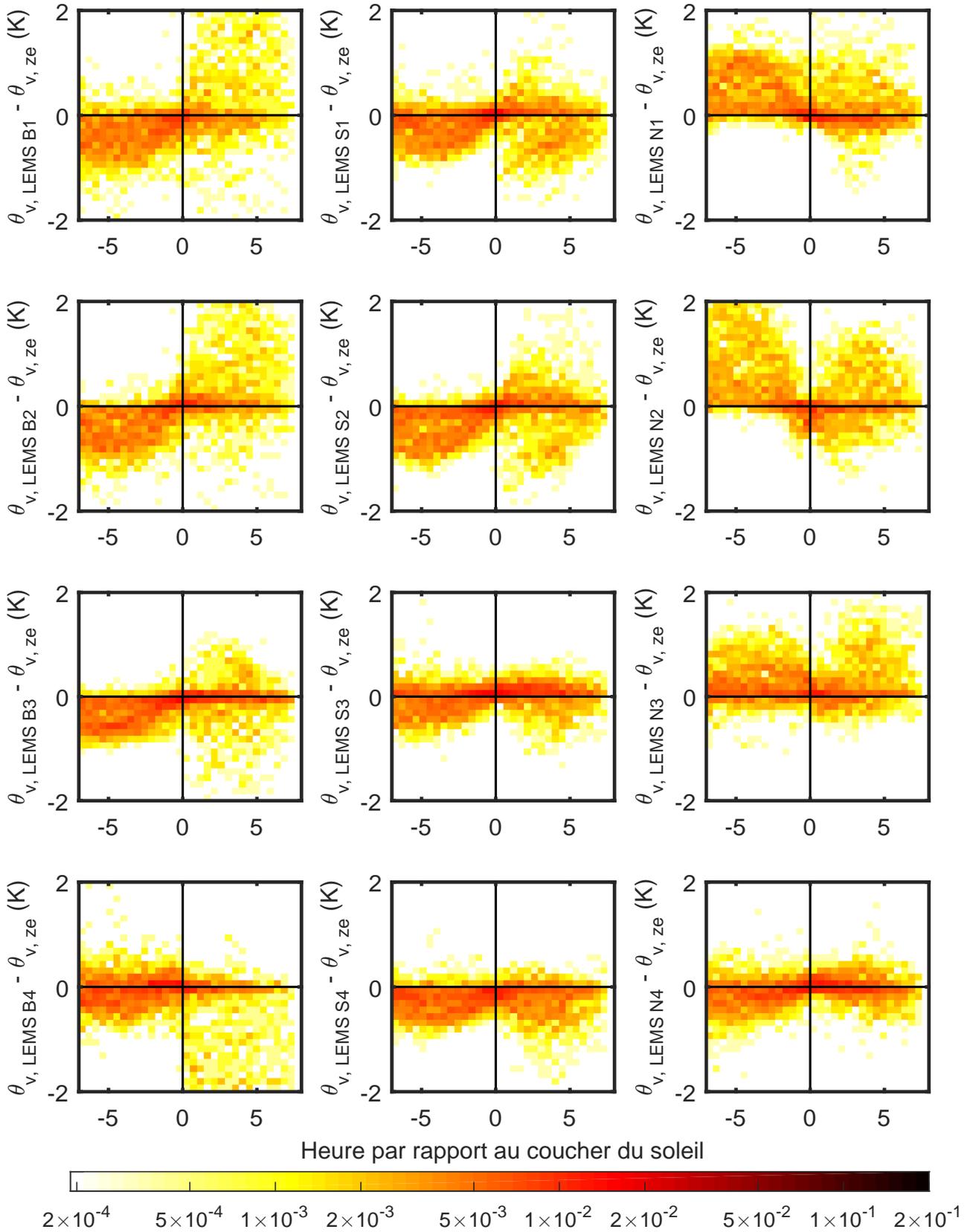


FIGURE 3.23 : Représentation en densité de l'écart $\theta_v - \theta_{v,ze}$ pour chaque LEMS prise individuellement en fonction de l'heure autour du coucher du soleil. Les données de toute la période KASCADE 2017 moyennées sur 30 minutes ont été utilisées. Les positions des différentes LEMS sont rappelées sur la figure 3.22.

3.3.3.b Caractéristiques de la couche stratifiée dans la vallée

La comparaison des profils de température mesurés sur une verticale au centre de la vallée (avec les radiosondages) et près du sol sur les versants de la vallée (avec les LEMS) permet d'aider à caractériser la structure de la couche stratifiée dans la vallée. Les profils observés pendant la deuxième POI sont représentés sur la figure 3.24. Cette POI a été choisie parce que les conditions étaient favorables à la formation et au maintien de la stratification tout au long de la nuit (les résultats sont similaires sur les POI 1 et 3).

À 15h00 UTC le profil vertical formé par les LEMS correspond à une couche de stabilité proche de la neutralité. À cette date, le coucher du soleil intervient 1h30 après, aux alentours de 16h30 UTC. Il est donc cohérent de retrouver un profil neutre d'après les résultats présentés dans la partie 3.3.3.a. Le radiosondage montre que l'atmosphère est également proche de la neutralité dans le centre de la vallée. L'air près de la surface est cependant légèrement plus chaud que l'air aux mêmes altitudes au centre de la vallée.

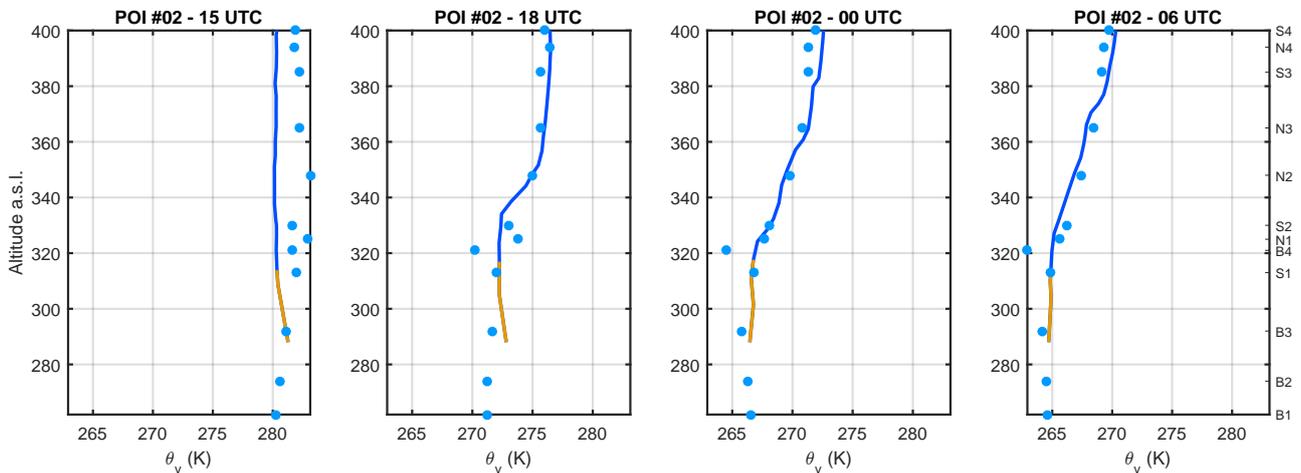


FIGURE 3.24 : Profils verticaux de température potentielle virtuelle mesurés par 4 radiosondages (trait plein, en orange pour les mesures dans les 40 premiers mètres entachées d'incertitude) et par les LEMS (points) aux heures de lâcher pendant la deuxième POI (du 18 au 19 janvier 2017). Le label des stations LEMS est indiqué sur le profil de 06h00 UTC le long de l'ordonnée à droite.

Toutes les stations ont enregistré un net refroidissement de plusieurs degrés entre 15h00 UTC et 18h00 UTC, toutefois plus important pour les stations les plus basses (autour de 10 K pour les 5 stations les plus basses). À 18h00 UTC, il y a donc un début de stabilité sur les versants de la vallée ainsi qu'au milieu de la vallée d'après le radiosondage.

À 00h00 UTC, la couche stable s'est renforcée dans la vallée comme sur ses versants. Cependant, il faut noter que les stations les plus hautes se situent à des altitudes qui correspondent à la zone de cisaillement entre les vents dans et au-dessus de la vallée de Cadarache ce qui peut expliquer que du fait du mélange dynamique, la stabilité y soit moins marquée. Le bon accord entre le sondage et les mesures à 2 m est caractéristique d'une cold pool dans laquelle les isentropes sont horizontales (cf. partie 1.2.2). Seule la LEMS *B4* s'écarte significativement des températures de l'air à la même altitude au centre de la vallée (ce comportement particulier est traité dans la partie 3.3.3.c).

Le refroidissement sur les 6 heures suivantes (jusqu'à 06h00 UTC) est bien moins important que celui produit sur les 6 heures précédant minuit (notamment pour les stations les plus

basses) ce qui limite l'intensité de la stratification ($\Delta\theta_v$ de 6,3 K sur les 100 premiers mètres à 00h00 UTC puis 5,7 K à 06h00 UTC).

Il est possible de décrire la stratification dans une vallée en utilisant la classification de [Mahrt \(2017\)](#) à partir de profils de température dans le centre de la vallée et sur ses versants en suivant la pente (voir la partie 1.2.2). Dans le cas de la vallée de Cadarache, les paramètres permettant cette description sont :

- *CPI* (cold pool intensity) qui représente la stratification sur le versant de la vallée donc $CPI = \Delta\theta_{v,LEMS}$;
- $-\theta^*$ qui représente la différence de température sur l'horizontale au niveau de la crête entre celle-ci et le centre de la vallée, soit $\theta^* = \theta_{v,crête} - \theta_{v,RS}$;
- $D\theta$ qui représente la variation de température potentielle de l'air en descendant la pente du versant.

$D\theta$ n'est pas mesurable avec l'instrumentation disponible pendant la campagne KASCADE 2017 (il faudrait des mesures de vent 3D pour cela). Ce paramètre informe sur le type d'écoulement en place dans la vallée. Un $D\theta$ nul signifie que les écoulements suivent les isentropes sans que l'air ne change de température. Une valeur positive est le signe d'un écoulement qui traverse les isentropes, l'air se refroidissant en descendant dans le fond de la vallée au contact de la surface froide (cf. figure 1.6).

La figure 3.25 montre l'évolution des paramètres pour les six POI présentées sur la figure 3.16. Les situations sont très proches. Ces calculs sont effectués sur deux couches, entre les stations N3 et B3 soit une couche de 72 m, et entre les stations N4 et B3 soit sur la profondeur totale de la vallée (couche de 102 m). La CPI entre les LEMS N3 et B3 montre que la couche est stratifiée entre 18h00 UTC et 09h00 UTC le lendemain matin. Les valeurs de $-\theta^*$ sont proches de zéro pendant les périodes de stratification ce qui indique que les isentropes dans la vallée sont pratiquement horizontales, ce qui est proche du comportement d'une cold pool mature d'après la classification de [Mahrt \(2017\)](#).

La situation est un peu différente entre les stations N4 et B3 avec des valeurs de $-\theta^*$ qui sont nettement positives la nuit ce qui signifie que les isentropes ne sont pas horizontales. Cela est caractéristique d'une cold pool marginale. Cela montre que la situation est proche d'une cold pool mature dans le fond de la vallée (avec des isentropes proches de l'horizontale) et plus de type cold pool marginale au-dessus. Le développement tardif de la stratification pendant la onzième POI s'explique par de fortes vitesses de vent dans la vallée en fin d'après-midi et début de nuit.

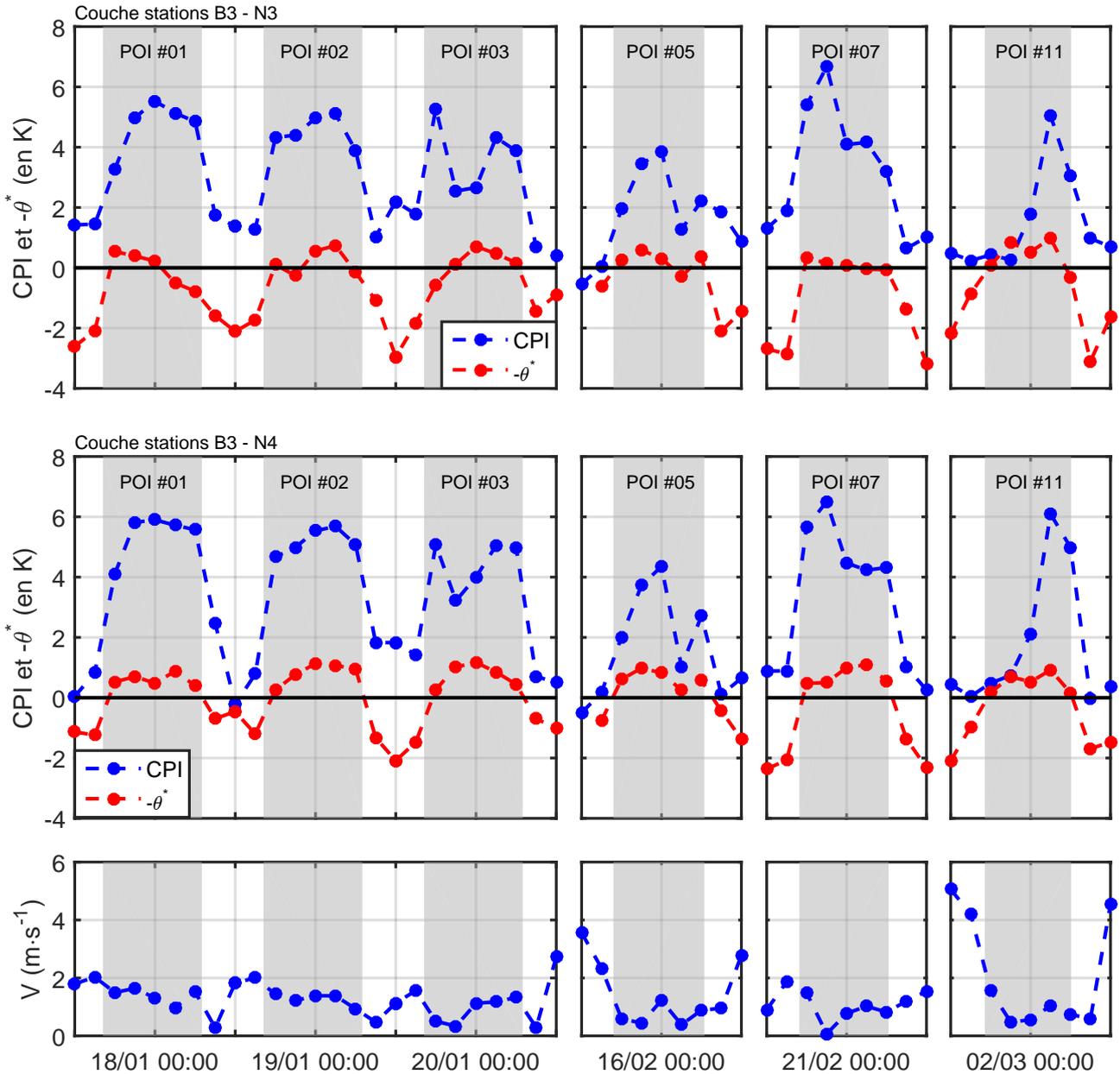


FIGURE 3.25 : Évolution de la CPI et du $-\theta^*$ calculés entre les altitudes des stations B3 et N3 (en haut) et B3 et N4 (au milieu) et vitesse du vent à 10 m mesurée sur la station MET01 (en bas) au cours des 6 POI présentées sur la figure 3.16. Les zones grisés représentent les périodes de nuit.

Mahrt (2017) étudie la relation entre la CPI et le $-\theta^*$ pour mieux déterminer le régime de stratification présent dans la vallée (figure 3.26). Dans la vallée de Cadarache, les valeurs de $-\theta^*$ sont toujours faibles. Concernant la CPI , il existe quelques situations pour lesquelles elle est proche de zéro (ce qui est caractéristique d'une couche neutre observée par vent fort comme en début de nuit lors de la POI #11, figure 3.25), mais elle est généralement forte ce qui est caractéristique d'une cold pool mature. C'est donc ce dernier régime qui semble être dominant dans la vallée de Cadarache, du moins au niveau du transect instrumenté perpendiculaire à l'axe de la vallée.

Enfin, deux couches d'atmosphère, à savoir les 72 premiers mètres entre les LEMS N3 et B3 et sur les 102 m premiers mètres entre les LEMS N4 et B3 (cette dernière prenant donc en compte toute l'extension verticale de la vallée), ont été comparées (figure 3.26). La stratification est légèrement plus forte dans la couche la plus fine. Cela s'explique par le fait que la LEMS N4,

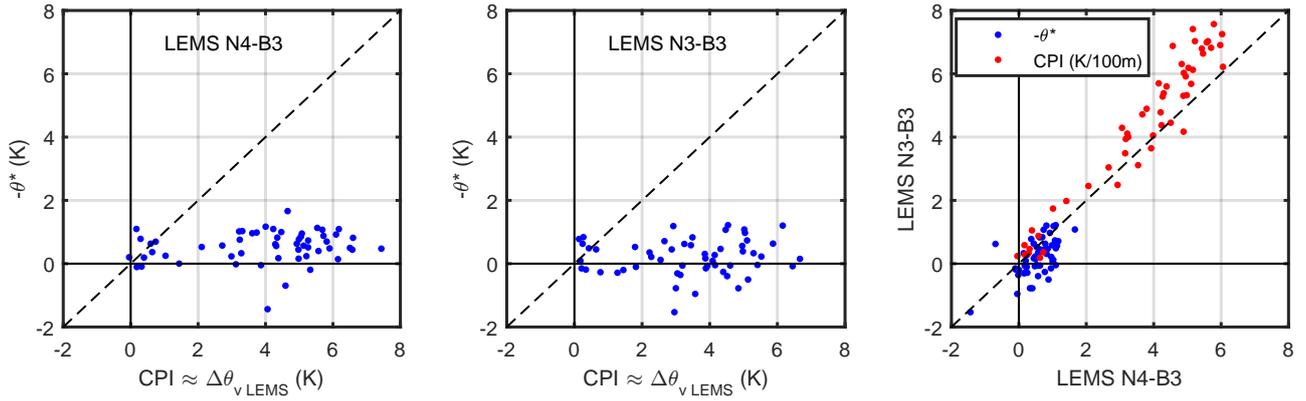


FIGURE 3.26 : $-\theta^*$ en fonction du CPI pour la couche comprise entre les altitudes des LEMS N4 et B3 (à gauche), et N3 et B3 (au centre). Comparaison des critères CPI et $-\theta^*$ des deux couches (à droite). Seules les données correspondant aux heures des radiosondages lancés de nuit sont représentées (pour les 14 POI de KASCADE 2017).

située sur la crête de la vallée, est exposée au cisaillement avec le vent circulant au-dessus de la vallée, dans une zone où l'intensité de la stratification diminue nettement. Les $-\theta^*$ sont proches de zéro dans la couche la plus proche du sol ce qui confirme pour cette couche le caractère de cold pool mature. Le comportement observé sur les 6 POI présentées sur la figure 3.25 semble donc se généraliser pour l'ensemble des 14 POI.

La comparaison des gradients verticaux de température potentielle estimés à partir des LEMS et à la station GBA (figure 3.27) montre que ce comportement se répète sur toute la période de la campagne KASCADE 2017 puisque les deux estimations sont proches, notamment pour les conditions très stables. Pour les stabilités faibles, les gradients sont plus forts sur la station GBA ce qui confirme que dans ces conditions le régime de stratification est de type cold pool marginale. Pour les conditions instables à la station GBA, il y a une anti-corrélation entre les deux estimations ce qui montre que les températures sont plus élevées sur la crête que dans le fond de la vallée. Comme nous avons montré que les gradients mesurés sur les deux versants étaient cohérents (figure 3.17) nous pouvons en conclure que cette tendance est présente sur les deux versants de la vallée.

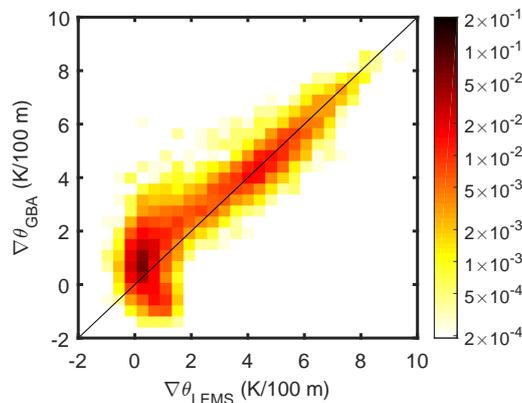


FIGURE 3.27 : Comparaison des gradients de température potentielle mesurés sur les stations GBA et LEMS. Les données de toute la période KASCADE 2017 moyennées sur 30 minutes ont été utilisées.

3.3.3.c Stratification dans le fond de la vallée

Un comportement particulier de la stratification a été observé en suivant le fond de la vallée par conditions stables. On observe une diminution de θ_v en remontant la vallée, alors que ce paramètre augmente lorsqu'on remonte le long des flancs. Ce comportement est illustré sur les figures 3.28 et 3.17 pour la période du 15 au 20 janvier 2017. Les moyennes sur les cinq nuits (figure 3.17) montrent bien des profils opposés entre les stations du fond de la vallée et celles placées sur les versants, ce qui veut dire que l'écoulement qui suit le fond de la vallée (vent CDV) se réchauffe au fur et à mesure de la descente. Seule la LEMS B3 a des températures cohérentes avec le profil que l'on peut attendre par conditions stables.

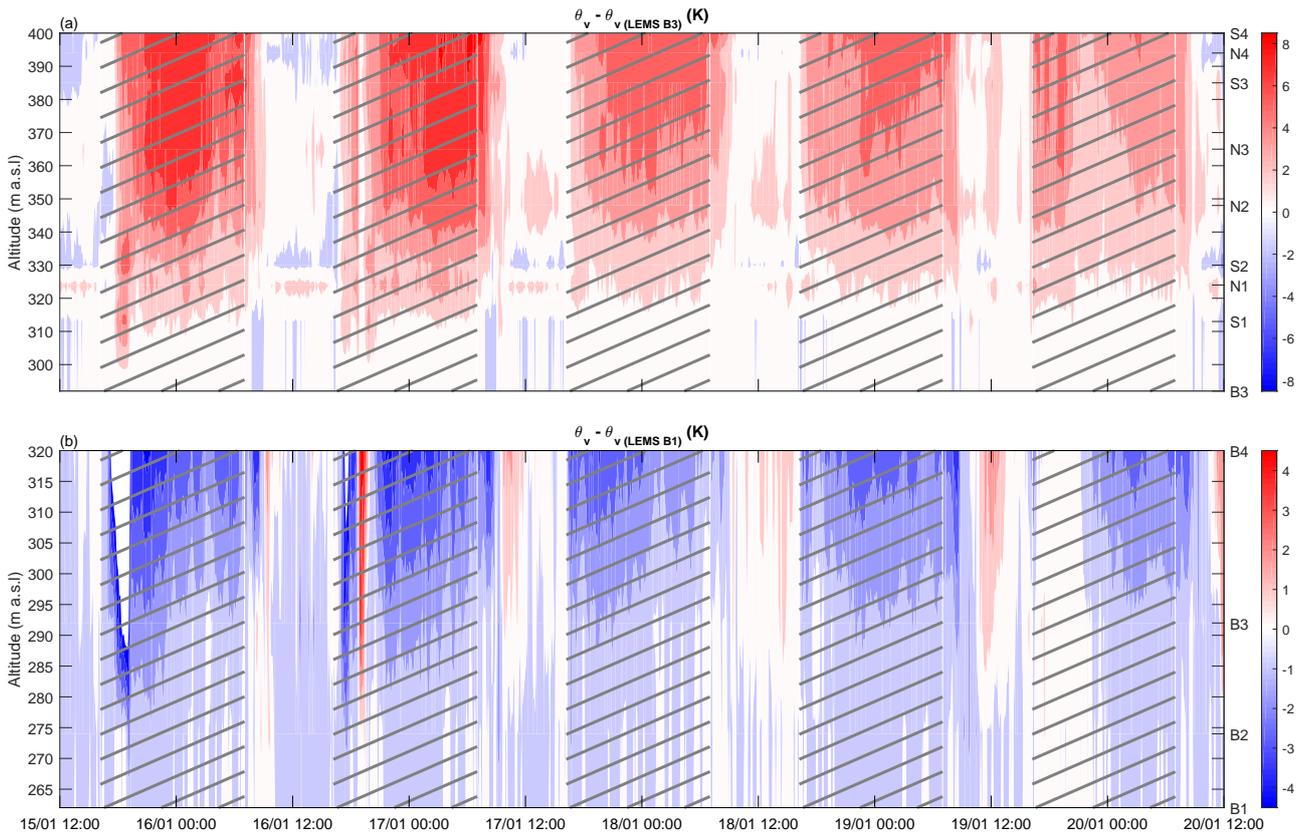


FIGURE 3.28 : Anomalie de températures potentielles virtuelles par rapport aux valeurs mesurées sur la LEMS la plus basse (en altitude) pour les LEMS placées sur les flancs de la vallée (a) et dans le fond de la vallée (b). Les hachures matérialisent les périodes de nuit. Les données sont moyennées sur 10 minutes. Le label des stations LEMS est indiqué le long de l'ordonnée à droite.

La figure 3.28, sur laquelle sont tracées les différences de θ_v entre les LEMS placées dans le fond de la vallée et la LEMS B1, et les différences de θ_v entre les LEMS situées sur les versants de la vallée et la LEMS B3, montre également ce comportement. En journée, les conditions sont proches de la neutralité sur les versants comme dans le fond de la vallée. Ce comportement se répète sur les cinq nuits présentées, avec quelques variations dans les intensités, et il se reproduit également tout au long de l'expérience (figure 3.29).

La comparaison entre les stratifications mesurées sur le versant nord et le long du fond de la vallée montre de nombreuses différences. Les stratifications sont semblables quand les conditions sont neutres sur les versants de la vallée parce que l'atmosphère y est alors bien mélangée. En revanche, il y a une anti-corrélation entre ces stratifications lorsque l'atmosphère devient stable sur les versants de la vallée. Cela est très clair sur la stratification calculée entre les stations B4

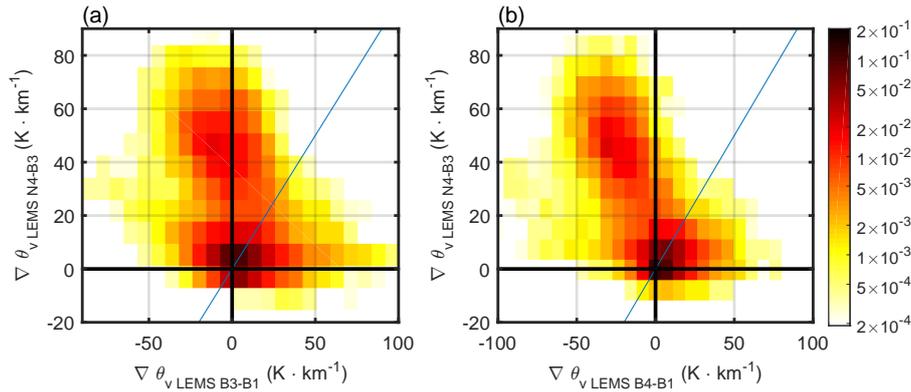


FIGURE 3.29 : Comparaison des gradients de θ_v sur le versant nord de la vallée (entre les LEMS N4 et B3) avec la mesure effectuée dans le fond de la vallée entre les LEMS B3 et B1 (a) et entre B4 et B1 (b). Les données de toute la période KASCADE 2017 moyennées sur 30 minutes ont été utilisées.

et B1 (figure 3.29.(b)), avec une diminution des températures en remontant le fond de la vallée alors que l’atmosphère est stable sur les versants de la vallée. Ce comportement est également bien marqué sur la moitié aval de la vallée, entre les LEMS B3 et B1 (figure 3.29.(a)), mais pas systématique puisque les températures peuvent aussi être parfois constantes ou augmenter légèrement en remontant la vallée (avec toutefois un gradient largement inférieur à celui observé sur les versants de la vallée). Cette différence de comportement entre les stations placées dans la moitié aval de la vallée (B1, B2 et B3) et la station placée le plus en amont (B4) nous incite à penser que des phénomènes distincts sur chaque partie de la vallée sont responsables de ce comportement particulier.

Différentes explications à ce phénomène ont été étudiées, parmi lesquelles :

- une augmentation du mélange dans le bas de la vallée résultant d’une interaction avec les écoulements de la vallée de Durance ou d’un renforcement du vent catabatique de la vallée de Cadarache qui détruirait en partie la stabilité près du sol (Muñoz et al., 2013) ;
- des différences liées aux hétérogénéités de la surface qui signifieraient que le phénomène observé sur quatre stations n’est pas représentatif de la distribution des températures dans le fond de la vallée ;
- une cold pool plus importante en zone amont (moins de drainage en haut qu’en bas) qui expliquerait que les températures en haut de la vallée soient systématiquement très froides.

Ces hypothèses vont maintenant être étudiées en détail.

Augmentation du mélange dans la partie aval de la vallée

Un mélange plus important dans le bas de la vallée permettrait d’expliquer l’augmentation de la température potentielle virtuelle en descendant la vallée par conditions stables. En effet, l’augmentation du mélange provoquerait une diminution de la stratification et donc un réchauffement de l’air près de la surface. Pour vérifier cette hypothèse, le profil moyen d’énergie cinétique turbulente a été calculé avec les anémomètres soniques placés dans le fond de la vallée sur les cinq nuits présentées précédemment (figure 3.30). Les anémomètres soniques AS1 et AS3 sont placés sur les mêmes sites que les LEMS B1 et B3 respectivement, alors que l’AS2 est situé entre les LEMS B2 et B3 et l’AS4 entre les LEMS B3 et B4. Il apparaît que l’énergie

cinétique turbulente augmente en descendant la vallée entre les stations AS3 et AS1, ce qui est cohérent avec l'hypothèse d'un mélange plus important dans le bas de la vallée.

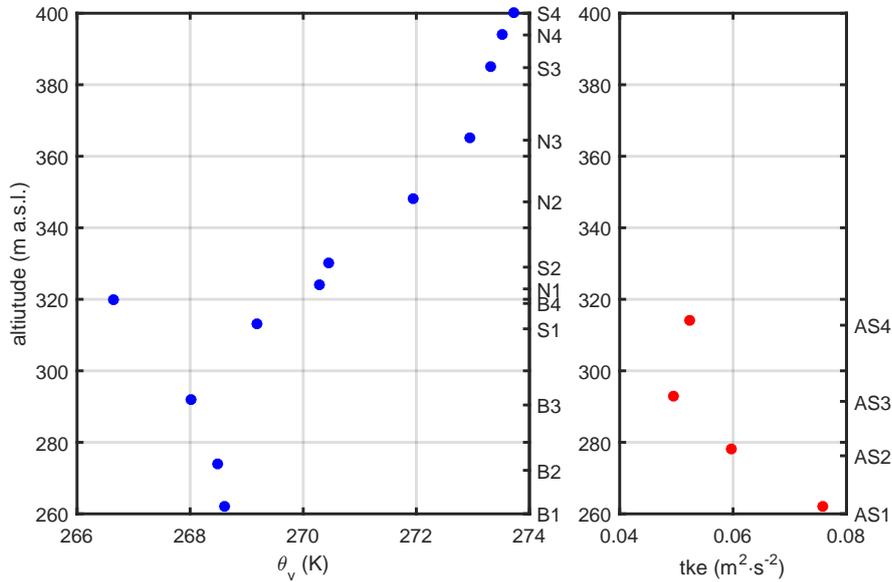


FIGURE 3.30 : Profil moyen sur les 5 nuits du 15 au 20 janvier 2017 de θ_v mesurées sur les LEMS (à gauche, idem figure 3.17) et d'énergie cinétique turbulente mesurée sur les stations AS (à droite).

Le bas de la vallée de Cadarache étant connecté à la vallée de la Durance, ce mélange peut être causé par une interaction avec les écoulements présents dans cette dernière. Les LEMS B1 et B2 sont en effet placées à proximité du débouché de la vallée de Cadarache où ses crêtes ne sont plus aussi élevées et protectrices qu'au niveau des stations B3 et B4, ce qui augmente l'exposition aux écoulements présents dans la vallée de Durance. Une autre explication possible est le phénomène décrit par [Muñoz et al. \(2013\)](#). Ils ont montré que les forts vents de vallée par conditions stables peuvent créer du mélange près de la surface générant une augmentation de la température dans le fond de la vallée en la descendant. Cependant, les vitesses observées par ces auteurs sont bien plus grandes (jusqu'à 20 m s^{-1}) que dans la vallée de Cadarache.

En revanche, l'énergie cinétique turbulente mesurée sur la station AS4 s'écarte de la tendance formée par les trois autres stations. L'hypothèse d'un mélange moins important pour expliquer la diminution des températures en remontant la vallée ne semble donc pas être valable sur la station AS4 située dans la partie amont de la vallée. La situation de la LEMS B4, également située dans la partie amont de la vallée, sera étudiée plus en détail par la suite.

Différences liées aux hétérogénéités du site

Les propriétés du sol modulent les échanges entre la surface et l'atmosphère, entre autres par les flux turbulents de chaleur sensible et latente ainsi que le flux de chaleur dans le sol, de sorte que les hétérogénéités de la surface et du sol sont responsables d'hétérogénéités des variables météorologiques (voir chapitre 1). Les observations effectuées avec les LEMS ne permettent pas de calculer ces flux par les méthodes de référence (eddy-corrélation pour H , LE et τ). En revanche, la différence entre températures de surface et de l'air permet de déterminer le signe du flux de chaleur entre la surface et l'atmosphère.

Les températures potentielles virtuelles à 1,5 m et à la surface ainsi que la différence entre ces deux niveaux sont représentées sur la figure 3.31. À 1,5 m, le comportement décrit plus haut est visible : il y a peu de différences entre les 4 stations en journée puis les différences apparaissent la nuit avec une diminution des températures en remontant la vallée. À la surface, les stations B2 et B3 sont proches durant tout le cycle diurne. La LEMS B1 en est également proche mais uniquement la nuit. En journée, les stations B1 et B4 sont systématiquement plus froides mais restent proches l'une de l'autre. Enfin, la baisse des températures de surface de la LEMS B4 entre fin de journée et début de nuit est singulière puisqu'elle est plus progressive que sur les autres stations. En revanche, alors que le refroidissement sur les autres stations s'atténue fortement quelques heures après le coucher du soleil, il se poursuit sur la LEMS B4 ce qui conduit à des températures de surface plus basses que sur les autres stations en fin de nuit.

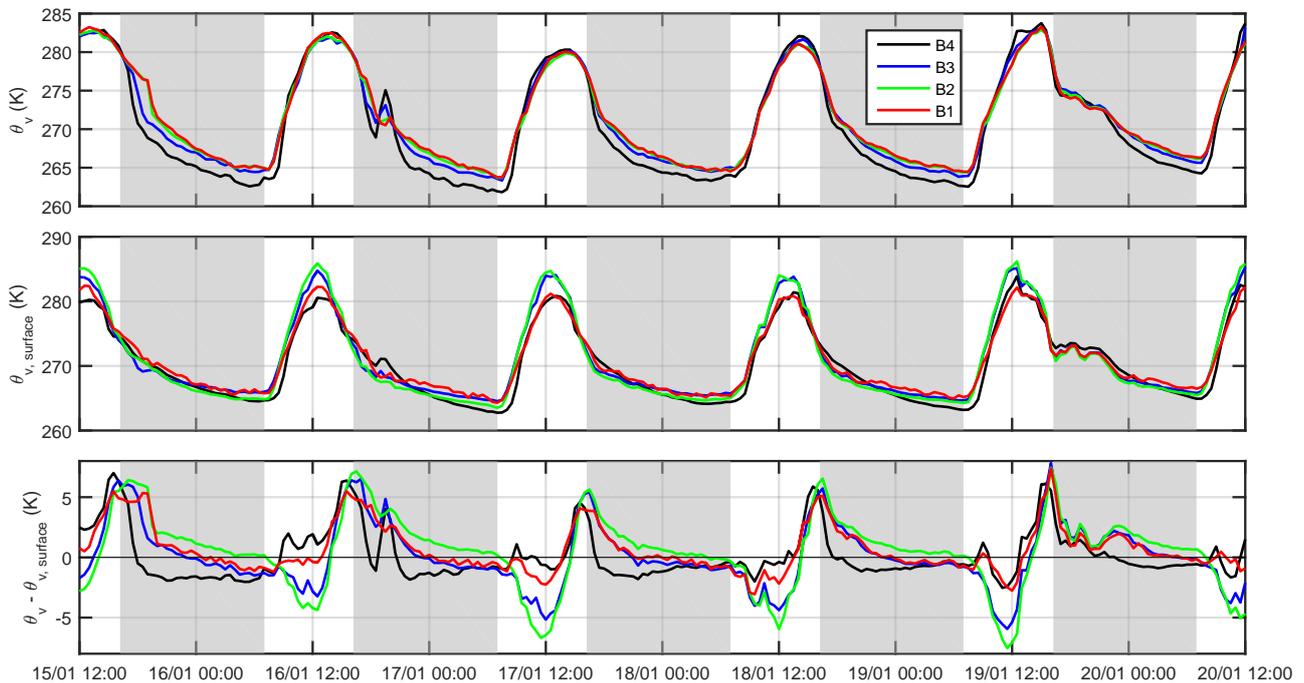


FIGURE 3.31 : Températures potentielles virtuelles à 2 m (en haut), à la surface (au milieu) et différence entre ces deux niveaux (en bas) pour les LEMS B1, B2, B3 et B4 (moyennes 30 minutes). Les parties grisées représentent les nuits.

Les différences de températures entre ces deux niveaux permettent en théorie de déterminer le signe du flux de chaleur entre la surface et l'atmosphère. Cependant, bien qu'il n'y ait pas de biais important (cf. annexe B) sur la mesure de l'émission infrarouge de la surface entre les 12 LEMS, l'estimation de la température de surface présente une incertitude parce que l'émissivité de la surface est définie a priori.

En journée, jusqu'en milieu d'après-midi, les valeurs sont en général négatives ($H > 0$). Pendant la nuit, les conditions stables se développent lorsque la surface subit un refroidissement radiatif ($H < 0$). Ce comportement est bien visible avec la LEMS B2 et dans une moindre mesure sur les LEMS B1 et B3 (les valeurs deviennent négatives en fin de nuit). La LEMS B4 est singulière puisque ses valeurs restent négatives pendant presque toute la nuit, ce qui signifie que la surface est plus chaude que l'air et donc que le refroidissement excessif de l'air (par rapport aux autres stations) n'est pas causé par le contact avec la surface. Cet écart de température observé pendant la nuit est alors probablement lié à une advection d'air plus froid, liée au drainage de l'air dans toute la partie haute de la vallée.

Situation particulière en partie amont de la vallée

Les températures significativement plus faibles observées sur la LEMS B4 ne sont donc pas liées à un refroidissement excessif par la surface mais à une advection d'air plus froid qui peut être associée au drainage dans tout le haut de la vallée. Or, les températures plus froides en haut de la vallée peuvent laisser penser que cet air stagne sans être redirigé vers la sortie de la vallée ou que son drainage est atténué, ce qui peut être causé par la géométrie de la vallée (Sheridan et al., 2014).

La comparaison des vitesses de vent observées sur les stations est représentée sur la figure 3.32. Pendant la nuit, les vitesses mesurées sur la LEMS B4 sont généralement supérieures à 1 m s^{-1} ce qui montre que le vent de vallée (venant de l'est à cause d'un changement d'orientation de l'axe de la vallée en cet endroit, cf. figure 3.22) est bien présent en haut de la vallée. De plus, les vitesses sont généralement plus fortes en haut de la vallée (LEMS B4) qu'en bas (LEMS B1 et B2) alors qu'il n'y a pas de tendance claire entre les vitesses des LEMS B3 et B4. L'air froid dans le fond de la vallée sur la partie haute peut donc s'écouler ce qui invalide l'hypothèse de la cold pool bien développée.

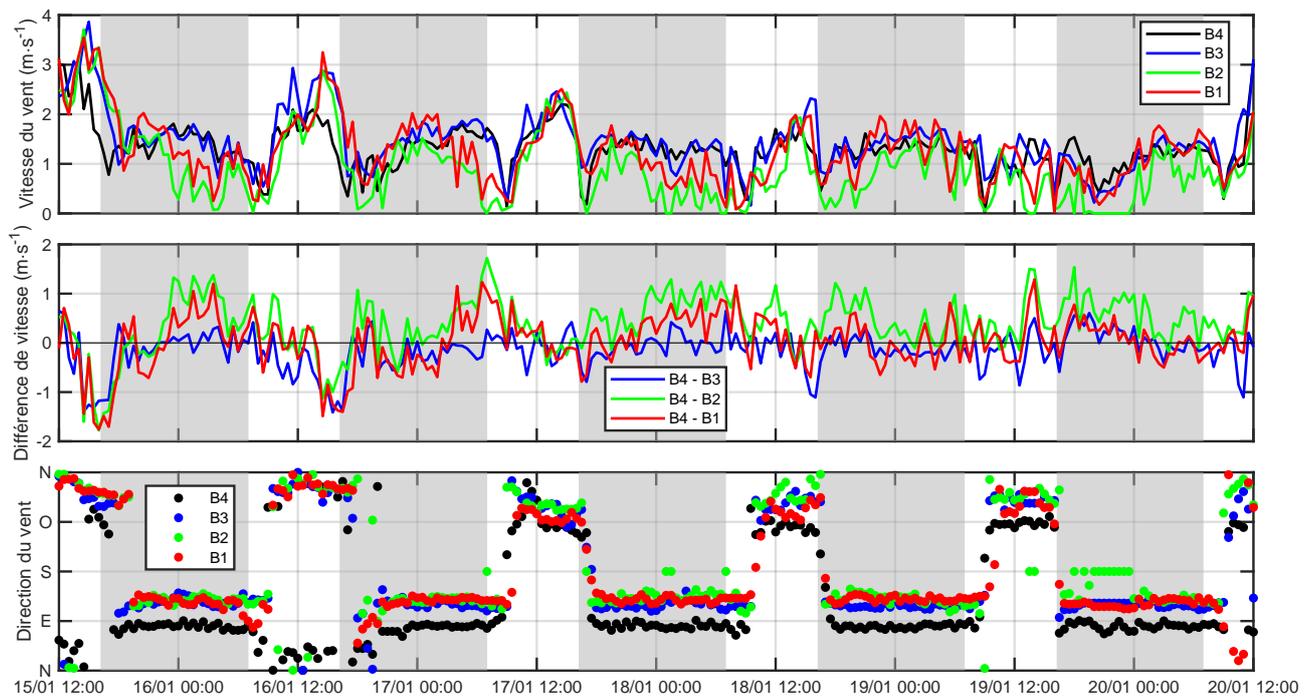


FIGURE 3.32 : Vitesse des vents mesurés à 2 m aux stations B1, B2, B3 et B4 (en haut), différence de vitesses entre celles mesurées à la station B4 et celles mesurées aux stations B1, B2 et B3 (au milieu) et directions du vent (en bas). Les données sont moyennées sur 30 minutes.

Différentes hypothèses permettant d'expliquer la structure de la stratification dans le fond de la vallée ont été étudiées. Les températures plus fortes observées dans la partie aval de la vallée (stations B1 et B2) sont liées à une augmentation de la turbulence dont l'origine n'est pas déterminée mais qui peut résulter d'une interaction entre les écoulements des vallées de Cadarache et de Durance. Les températures très froides observées dans la partie amont de la vallée (station B4) ne résultent pas de la présence d'une cold pool, puisqu'il existe bien un écoulement de vallée permettant le drainage de l'air, ni d'un refroidissement excessif au contact de la surface, puisque cette dernière est ici plus chaude que l'air. Une advection d'air plus froid résultant d'un drainage dans toute la partie haute de la vallée est alors l'hypothèse privilégiée.

3.4 Écoulements dans la vallée de Cadarache

Whiteman and Doran (1993) ont décrit quatre régimes qualifiant les interactions entre les vents dans et au-dessus d'une vallée (voir partie 1.2.1.c). Comme cela a déjà été montré dans la partie 2.2.3, les principaux mécanismes se produisant dans la vallée de Cadarache sont le « downward momentum transport » (mêmes directions pour les vents dans et au-dessus de la vallée), la canalisation (les vents présents au-dessus de la vallée sont déviés en entrant dans la vallée où ils s'alignent avec l'axe de la vallée) et les courants de densité produits par des effets thermiques. Au contraire des deux premiers mécanismes, les courants de densité supposent un découplage entre les vents dans et au-dessus de la vallée. Dans la partie suivante, nous tentons de déterminer les conditions conduisant aux situations de couplage/découplage des vents dans la vallée de Cadarache.

3.4.1 Couplage/découplage des écoulements dans et au-dessus de la vallée

Le couplage des vents est évalué avec une différence de direction entre les vents à 120 m (sodar SOD2) et à 10 m (MET01) :

$$\Delta\text{WD} = |\text{WD}_{120\text{ m}} - \text{WD}_{10\text{ m}}| \quad (3.1)$$

La figure 3.33 montre que ΔWD est dépendant de la vitesse du vent aussi bien au-dessus que dans la vallée. C'est un résultat attendu puisque les vents forts en altitude sont capables de forcer les directions des vents dans la vallée et de même la présence de vents forts dans la vallée est corrélée avec celle de vents forts en altitude. Il y a une vitesse seuil autour de 4 m s^{-1} pour les vents à 120 m au-delà de laquelle les vents sont couplés. De même, un seuil de 2 m s^{-1} apparaît pour les vents à 10 m. En dessous de ces vitesses seuil, les différences de direction peuvent prendre des valeurs très élevées.

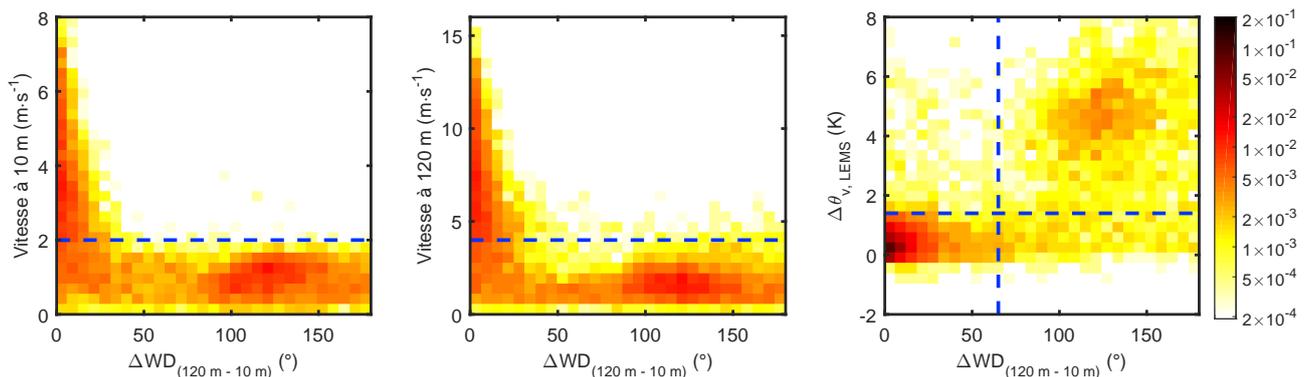


FIGURE 3.33 : Différence entre les directions de vents à 120 m (SOD2) et 10 m (MET01) versus la vitesse à 10 m et à 120 m, et la différence de température potentielle virtuelle entre les LEMS N4 et B3. Les droites pointillées bleues repèrent les vitesses de 2 m s^{-1} à 10 m, 4 m s^{-1} à 120 m, la stratification de $1,4\text{ K}$ et les différences de directions de 65° . Les données de toute la période KASCADE 2017 moyennées sur 30 minutes ont été utilisées.

La figure 3.33 montre également que les stratifications, mesurées sur le versant de la vallée, supérieures à $1,4\text{ K}$ sont associées à un découplage des vents au-dessus et dans la vallée. Les faibles ΔWD sont associées à de plus faibles stratifications. Ce seuil a été déterminé en appliquant la méthode de Duine et al. (2016) et en définissant les vents comme découplés lorsque

$\Delta\text{WD} > 65^\circ$. Ainsi, 82% des données vérifient l'une des deux conditions : soit $\Delta\text{WD} > 65^\circ$ et $\Delta\theta_{v,\text{LEMS}} > 1,4\text{K}$, soit $\Delta\text{WD} < 65^\circ$ et $\Delta\theta_{v,\text{LEMS}} < 1,4\text{K}$, c'est à dire que 82% des données sont situées dans les deux quadrants sur la figure 3.33 (à droite). Ce seuil est nettement inférieur à celui trouvé par Duine et al. (2016) (2,6 K), qui permet de distinguer les vents de vallée descendants à partir de la différence de température potentielle entre 110 et 2 m à la station station GBA, ce qui s'explique par le fait que les stratifications mesurées sur la station GBA sont supérieures à celles mesurées sur le versant de la vallée ($\Delta\theta_{v,\text{LEMS}}$) pour les conditions où la stratification est modérée (figure 3.27).

Dans les modèles théoriques de stratification décrits par Mahrt (2017), les vents couplés sont associés à des stratifications dans lesquelles les isentropes sont parallèles au terrain ce qui signifie que la stratification le long d'une pente est nulle et elle est stable au centre de la vallée. Au contraire, les vents découplés sont associés à des stratifications importantes le long des pentes. Cela correspond aux résultats présentés dans la figure 3.33 avec néanmoins un couplage des vents possible pour de faibles stratifications (jusqu'à 2 K entre les stations N4 et B3) le long des pentes.

Plusieurs études ont tenté de distinguer les conditions menant à un couplage/découplage des vents dans et au-dessus de la vallée sur la base d'une adaptation du nombre de Froude ou d'une profondeur adimensionnelle de la vallée $\hat{H} = 1/Fr$ (voir partie 1.2.1.d).

\hat{H} a été calculé pour la vallée de Cadarache à partir de moyennes sur 30 minutes de la vitesse U mesurée à 100 m avec SOD2, d'une hauteur caractéristique H de 100 m (profondeur moyenne de la vallée) et d'une pulsation de Brunt-Väisälä calculée à partir de la différence de température potentielle mesurée avec la station GBA (méthode du gradient adiabatique). Une valeur critique \hat{H} permet de distinguer les conditions pour lesquelles le vent présent au-dessus de la vallée est capable de forcer les vents dans la vallée. Les valeurs seuil trouvées dans de précédentes études, entre 0,3 et 1 (Bell and Thompson, 1980; Mason, 1987; Holden et al., 2000; LeMone et al., 2003; Vosper and Brown, 2008), voire plus pour les vents en altitude alignés avec l'axe de la vallée (Largeron, 2010), varient d'un site à l'autre ce qui montre que cette valeur seuil est une caractéristique intrinsèque de chaque site.

Pour la vallée de Cadarache, la valeur de seuil trouvée est $\hat{H} = 1$ (cohérent avec de précédentes études sur d'autres sites). Ce seuil (déterminé en appliquant la méthode de discrimination de Duine et al. (2016) avec une réussite de 82%) est visible pour la différence de directions entre les vents à 120 m et 10 m, mais pas pour la stratification sur le côté de la vallée (figure 3.34). De même que Largeron (2010), qui a trouvé des valeurs critiques très élevées dans le cas d'écoulements en altitude alignés avec l'axe de la vallée (\hat{H} compris entre 3,3 et 16,7), la valeur de ce seuil sur le site de Cadarache est légèrement plus grande pour ces vents ($\hat{H} = 1,5$ pour les vents en altitudes alignés avec la vallée à $\pm 20^\circ$).

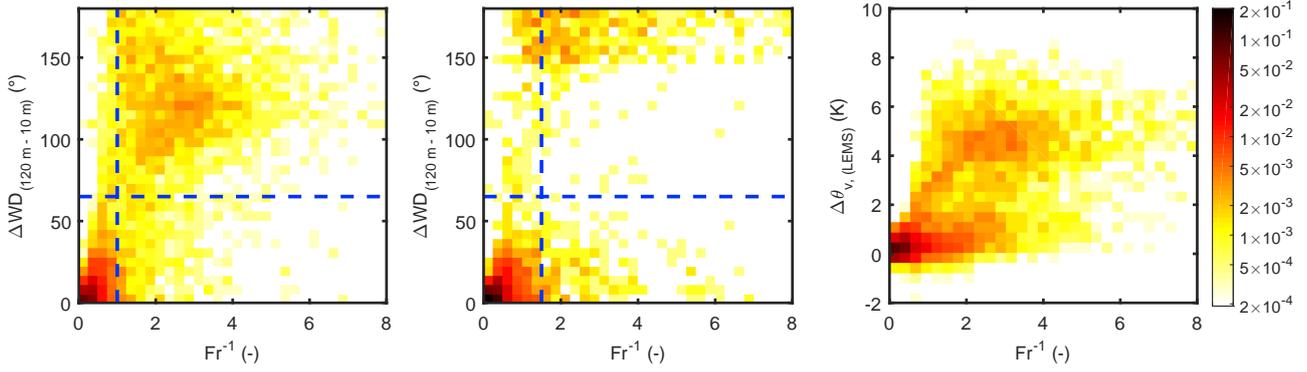


FIGURE 3.34 : Profondeur adimensionnelle de la vallée $\hat{H} = 1/Fr$ versus la différence de directions entre les vents à 120 m et 10 m pour tous les vents (à gauche) et uniquement pour les vents à 120 m alignés avec l'axe de la vallée à $\pm 20^\circ$ (au centre) et versus la différence de température potentielle virtuelle entre les LEMS N4 et B3 (à droite). Les droites pointillées bleues repèrent les valeurs $\hat{H} = 1$ (à gauche), $\hat{H} = 1,5$ (au centre) et les $\Delta WD_{(120\text{m}-10\text{m})} = 65^\circ$. Seules les données nocturnes ont été utilisées ici.

3.4.2 Hétérogénéité des directions des vents

L'étude de l'hétérogénéité des vents dans la vallée de Cadarache est réalisée à partir des observations des 12 LEMS. Ces stations mesurent la vitesse et la direction du vent avec un anémomètre à coupelles et une girouette qui présentent des limites pour l'observation des vents très faibles. L'inter-comparaison des mesures de vent des différentes stations n'ayant pas pu être effectuée à l'issue de KASCADE 2017, il est tout de même possible d'évaluer la qualité des mesures de quelques stations.

En effet, il y a deux sites sur lesquels se trouvaient à la fois une station LEMS et un anémomètre sonique ce qui a permis de comparer les mesures de vent, notamment pour les vents faibles. Les résultats de cette comparaison sont compilés dans l'annexe B. Comme attendu, il ressort que les directions mesurées avec les LEMS perdent en précision avec la baisse des vitesses. Un moyen de diminuer l'erreur de mesure est alors d'augmenter la période sur laquelle les données sont moyennées. Une moyenne sur 30 minutes permet d'obtenir une erreur médiane (par rapport aux mesures des stations AS) de 38° pour les vents inférieurs à $0,25 \text{ m s}^{-1}$ et 28° pour ceux compris entre $0,25$ et $0,50 \text{ m s}^{-1}$ sur le site AS1-B1. Ces valeurs sont respectivement de 54° et 31° sur le site AS3-B3. Les données relatives aux directions de vents mesurées sur les LEMS sont donc moyennées sur 30 minutes dans la suite du chapitre.

3.4.2.a Climatologie des vents

Comme pour l'hétérogénéité des températures, celle des directions de vent est estimée à partir de l'écart-type calculé sur les directions des 12 LEMS. La méthode de Yamartino (1984) est utilisée pour ce calcul parce qu'elle permet d'éviter les problèmes liés au fait que la direction est une valeur cyclique :

$$\sigma_{\text{WD}} = \arcsin(\epsilon) \left[1 + \left(\frac{2}{\sqrt{3}} - 1 \right) \epsilon^3 \right] \quad (3.2)$$

$$\text{avec : } \epsilon = \sqrt{1 - (s_a^2 + c_a^2)}; \quad s_a = \frac{1}{n} \sum_{i=1}^n \sin \text{WD}_i; \quad c_a = \frac{1}{n} \sum_{i=1}^n \cos \text{WD}_i$$

La figure 3.35 montre l'évolution de l'hétérogénéité des directions en fonction de plusieurs paramètres. L'hétérogénéité est faible quand les directions des vents au-dessus et dans la vallée sont proches (couplage des vents) ce qui montre que le couplage constaté entre les vents du sodar à 120 m et de la station MET01 est effectif dans le fond de la vallée mais également sur les flancs. À l'inverse, quand il y a découplage, l'hétérogénéité entre les stations augmente. Les stratifications supérieures à 1,4 K (seuil trouvé dans la partie 3.4.1) sont associées à de fortes hétérogénéités. Enfin, on retrouve le même seuil pour la profondeur adimensionnelle de la vallée ($\hat{H} = 1$, figure 3.35) que précédemment avec la différence de direction entre les vents au-dessus et dans la vallée. Finalement, une vision climatologique des observations de vent des LEMS, notamment à partir de roses des vents, permet d'améliorer la compréhension des hétérogénéités dans la vallée de Cadarache.

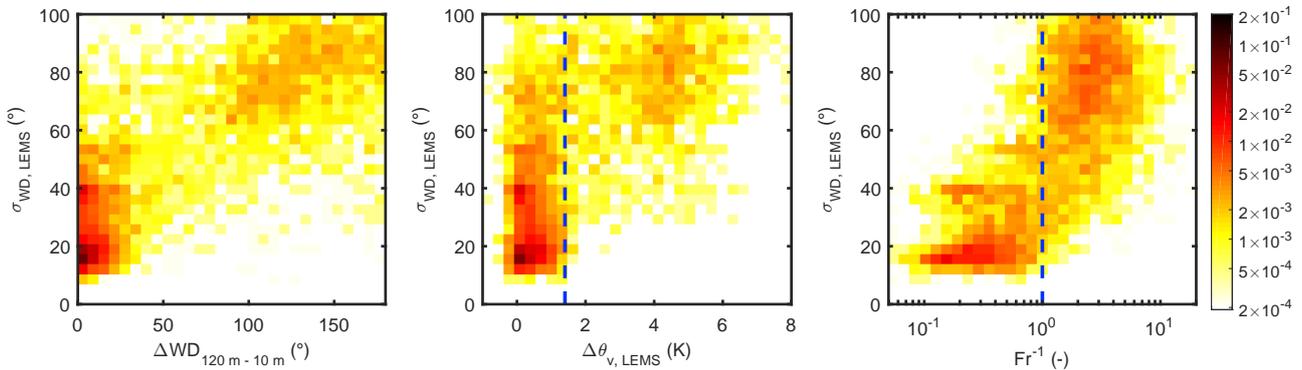


FIGURE 3.35 : Écart-type calculé sur les directions de vent des 12 LEMS en fonction de la différence de direction ΔWD entre la station MET01 et la mesure sodar à 120 m (à gauche), de la différence verticale de température potentielle virtuelle $\Delta\theta_v$ (au centre) et de la profondeur adimensionnelle de la vallée Fr^{-1} (à droite). Les lignes tiretées bleues repèrent les valeurs $\Delta\theta_v = 1,4\text{ K}$ sur le graphe du centre et $Fr^{-1} = 1$ sur le graphe de droite. Les données de toute la période KASCADE 2017 moyennées sur 30 minutes ont été utilisées.

Tout d'abord, notons que la LEMS N4 qui est placée sur la crête du côté nord de la vallée voit les mêmes vents que le sodar à 120 m (soit sensiblement la même altitude que celle de la LEMS N4) comme le montre la figure 3.36. La LEMS N4 mesure donc des vents qui sont représentatifs de ceux dans la vallée de la Durance (comme la station GBA et les sodars) et qui ne sont donc pas dépendants des effets locaux générés par la vallée de Cadarache. De même, la comparaison des directions des deux LEMS situées sur les crêtes (N4 sur le versant nord et S4 sur le versant sud) montre que les vents observés sur la station S4 sont également caractéristiques des vents observés au-dessus de la vallée. La seule exception concerne les vents de nord-est à la station LEMS N4 (vents de vallée de Durance) qui prennent une direction nord-nord-ouest sur la station S4. Cette rotation peut être la conséquence d'une redirection du vent de vallée de Durance par le relief local, puisque la station S4 se situe sur la crête d'un vallon orienté SSE-NNO, et/ou par le relief à plus grande échelle (montagne de la Vautubière, cf. la topographie locale représentée sur les figures 3.36 et 3.37).

Pour la suite de l'analyse, les données sont séparées suivant l'état de couplage ou découplage entre les vents dans et au-dessus de la vallée. Pour cela, la valeur critique de la profondeur adimensionnelle de la vallée ($\hat{H} = 1$) est utilisée.

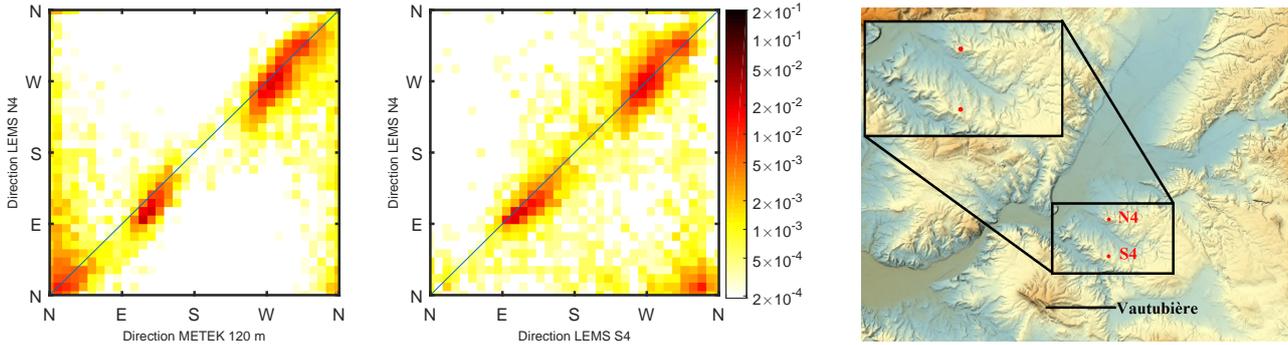


FIGURE 3.36 : Comparaison des directions observées sur les stations SOD2 (à 120 m a.g.l.) et N4 (à gauche) et S4 et N4 (au centre). Les positions de ces stations sont rappelées sur le graphe de droite. Les données de toute la période KASCADE 2017 moyennées sur 30 minutes ont été utilisées.

Couplage des vents : $\hat{H} < 1$

Les roses des vents des 12 LEMS pour les situations de couplage sont représentées sur la figure 3.38 (la position des stations est rappelée sur la figure 3.37). Dans ces conditions, les roses des vents des 12 stations sont similaires bien qu'il y ait quelques différences. Les stations N4 et S4 permettent de rendre compte des vents dominants au-dessus de la vallée de Cadarache (voir la partie 2.2 pour une description précise de ces vents). Elles sont donc importantes pour interpréter les vents observés sur les autres stations. Il est à noter que la LEMS S3, qui est placée sur la crête d'une vallée secondaire présente dans la vallée de Cadarache, voit les mêmes vents que les stations N4 et S4.

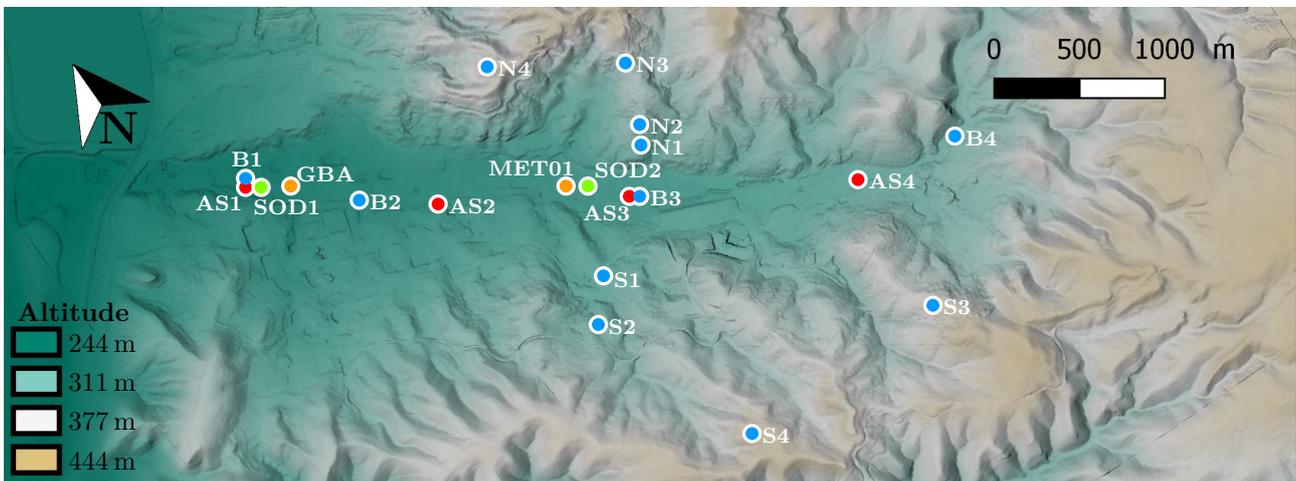


FIGURE 3.37 : Rappel de la position des stations dans la vallée de Cadarache (idem figure 3.1 avec une vue centrée sur la vallée de Cadarache).

Les vents observés sur les stations du fond de la vallée (LEMS B1 à B4) sont clairement alignés avec l'axe de la vallée. Les trois stations les plus basses sont situées à des endroits où la vallée est orientée SE–NO ce qui ne permet pas de distinguer le phénomène majeur (canalisation ou transport de quantité de mouvement vers le bas) conduisant à l'alignement des vents avec l'axe de la vallée puisque les vents en altitude sont déjà en grande partie orientés dans l'une ou l'autre des directions parallèles à l'axe de la vallée. En revanche, la vallée change de direction dans sa partie amont où elle s'oriente suivant un axe ENE–OSO au niveau de la LEMS B4. Or, les vents observés sur cette station restent alignés avec l'axe local de la vallée, ce qui signifie qu'ils sont canalisés par la géométrie de la vallée de façon à ce qu'ils soient alignés avec l'axe

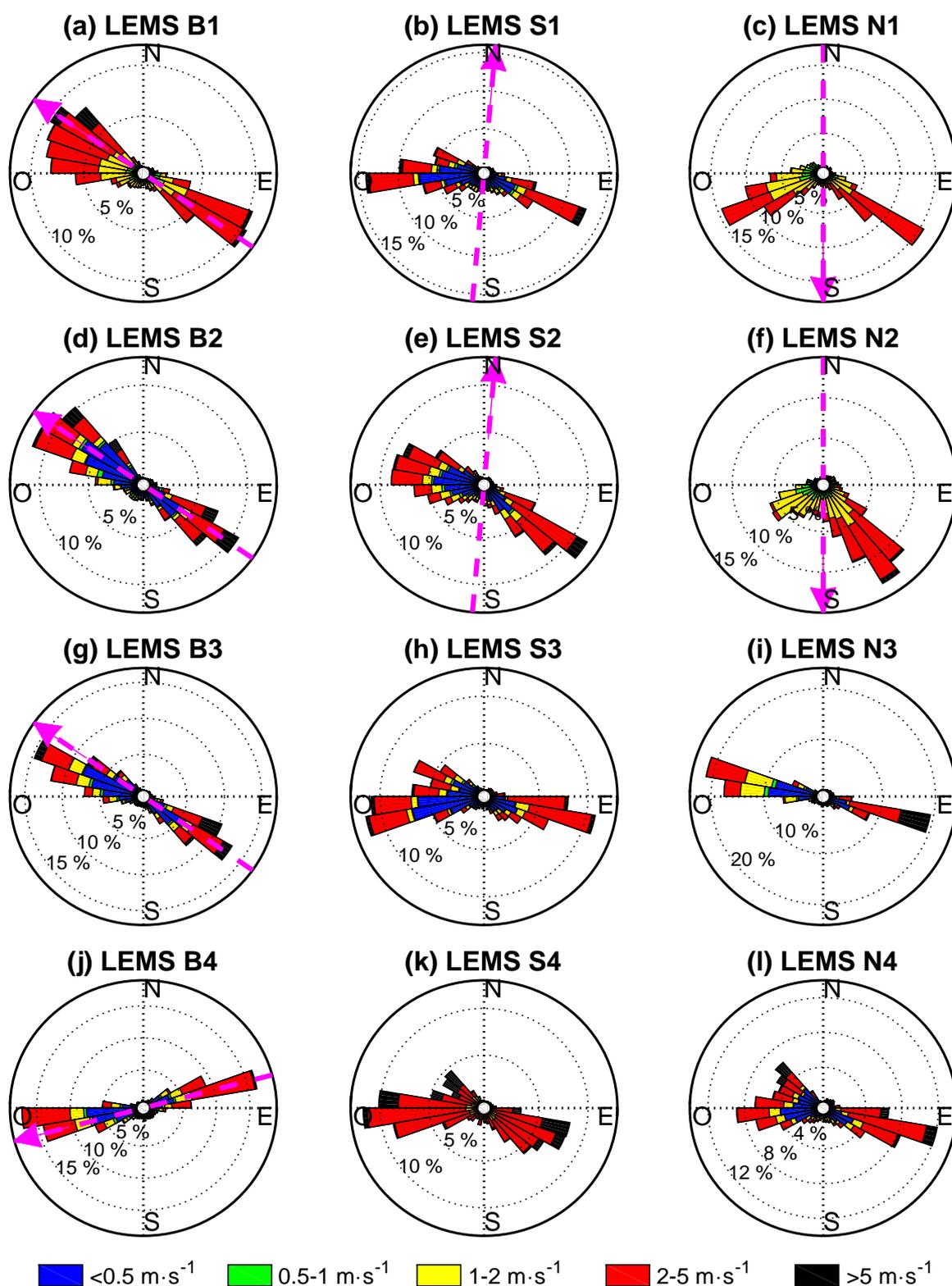


FIGURE 3.38 : Roses des vents pour les observations des 12 LEMS (moyennes 30 minutes) pour les conditions où $\hat{H} < 1$. Les flèches tiretées magenta indiquent soit l'axe descendant de la vallée pour les stations du fond de la vallée, soit la pente locale autour de la station (sauf lorsque la topographie locale ne permet pas de la définir de manière univoque).

de la vallée. La canalisation est donc le phénomène prépondérant sur cette station en situation de couplage entre les vents dans et au-dessus de la vallée.

Les cinq stations décrites dans ce paragraphe sont les stations les plus basses situées sur le

segment transverse à l'axe de la vallée. Les vents des stations S1 et S2 sont relativement bien alignés avec l'axe de la vallée. Pour les stations N1 et N2, les vents venant du sud-est sont bien alignés avec l'axe de la vallée de Cadarache mais les vents venant de l'ouest et du nord-ouest prennent une direction sud-ouest sous l'influence d'une canalisation par un petit vallon orienté NE-SO. Bien que le relief sur le site de la LEMS N3 soit particulièrement complexe avec deux pentes concurrentes, l'une orientée ONO-ESE et l'autre orientée N-S, et un positionnement près d'un col dont le franchissement se fait dans la direction ONO-ESE, les vents observés sur cette station semblent également avoir des directions forcées par le relief local. La canalisation des vents sur cette station, orientée de manière à franchir le col, est alors particulièrement bien visible avec très peu de dispersion des directions.

De manière générale, les situations de couplage produisent donc des vents qui ont des caractéristiques proches avec deux plages de directions préférentielles liées à la canalisation des vents par la vallée de Cadarache. Cependant, le relief particulièrement complexe impacte localement les directions des vents. Les vents de sud-est ne subissent pas de rotation majeure, à l'exception des stations B4 et N2 (environ 30° de différence par rapport aux vents observés sur la station N4), alors que les vents d'ouest et nord-ouest peuvent être beaucoup plus déviés (environ 80° de différence pour la station N2). En situation de couplage, l'hétérogénéité des directions de vent des 12 stations est faible, avec toutefois une variabilité un peu plus importante pour les vents venant du nord-ouest.

Découplage des vents : $\hat{H} > 1$

Dans les situations de découplage entre les vents dans et au-dessus de la vallée ($\hat{H} > 1$, figure 3.39) les roses des vents des 12 stations présentent d'importantes différences (figure 3.35).

Les deux vents typiquement rencontrés en altitude par conditions stables sont visibles sur les stations N4 et S4 : d'abord des vents d'ouest souvent présents en début de nuit, puis un courant de densité dans la vallée de la Durance en fin de nuit (Duine et al., 2017). Comme le laissait déjà entrevoir la figure 3.36, la rotation du vent de vallée de Durance entre les stations N4 et S4 est maintenant visible sur les roses des vents. Les quatre stations placées dans le fond de la vallée mesurent toutes un vent de vallée qui descend la vallée de Cadarache. Les vents observés sur les stations S1 et S2 ont également une origine thermique manifeste puisque ces vents descendent les pentes locales.

L'analyse des autres stations est plus complexe. Concernant la LEMS N3, son site est marqué par deux pentes locales concurrentes, orientées ESE-ONO et N-S. Le vent observé sur cette station comporte deux directions majoritaires qui sont orientées de façon à descendre l'une ou l'autre de ces pentes. La distinction entre les deux directions est corrélée à la présence ou non du vent de vallée de Durance. La présence de ce vent favorise les vents de nord sur la LEMS N3 alors que son absence est associée à la présence de vents d'ouest-nord-ouest. Une situation similaire se produit sur la station N2, mais avec une seule pente locale, puisque la présence d'un vent de vallée de Durance favorise les vents de nord sur cette LEMS. En revanche, son absence ne favorise pas une direction en particulier. Le vent de vallée de Durance est donc capable d'impacter les vents sur le versant nord de la vallée de Cadarache, lesquels sont donc moins découplés que sur les stations placées dans le fond ou sur le versant sud de la vallée.

Finalement, le découplage produit des vents aux directions majoritairement forcées par le relief local (les différentes orientations de pentes) alors que c'est le relief à plus grande échelle

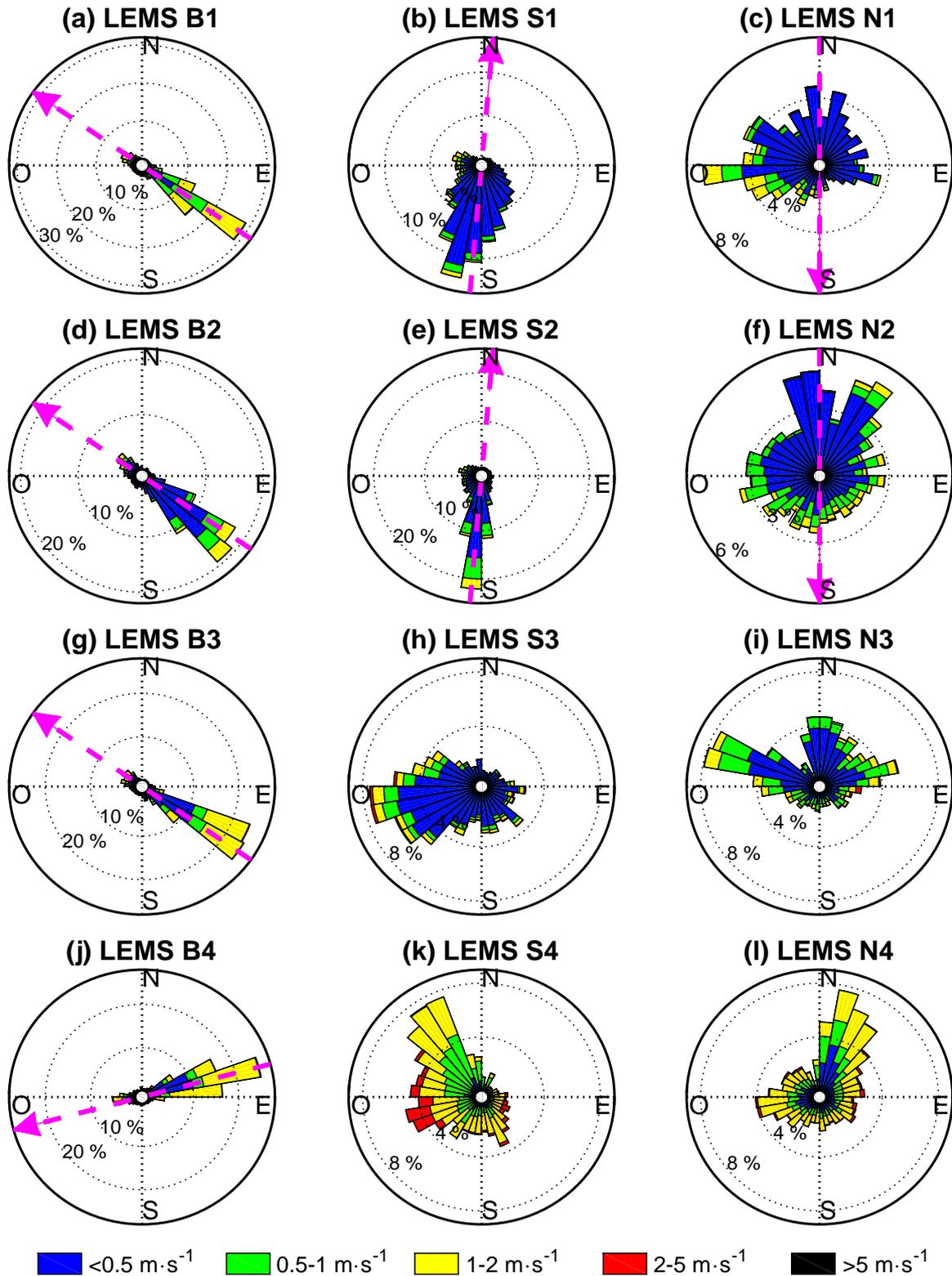


FIGURE 3.39 : Idem figure 3.38 pour les conditions où $\hat{H} > 1$.

(la vallée de Cadarache dans son ensemble) qui force les directions des stations placées dans la vallée en situation de couplage. Cela explique pourquoi l'hétérogénéité des directions de vent est plus importante quand il y a découplage des vents dans et au-dessus de la vallée.

3.4.3 Évolution temporelle du vent de vallée de Cadarache

La partie 3.4.2 a permis de montrer qu'il existe un système de vents bien développé dans la vallée de Cadarache par conditions stables avec des vents de pente (plus ou moins marqués suivant les situations et les localisations) sur les versants et un vent de vallée qui s'écoule dans le fond de la vallée. Le but maintenant est d'étudier la formation et la disparition de ces vents et notamment du vent de vallée. Pour cela, des moyennes sur 5 minutes des vents observés avec les anémomètres soniques sont utilisées, les directions mesurées étant plus précises que celles mesurées par les LEMS.

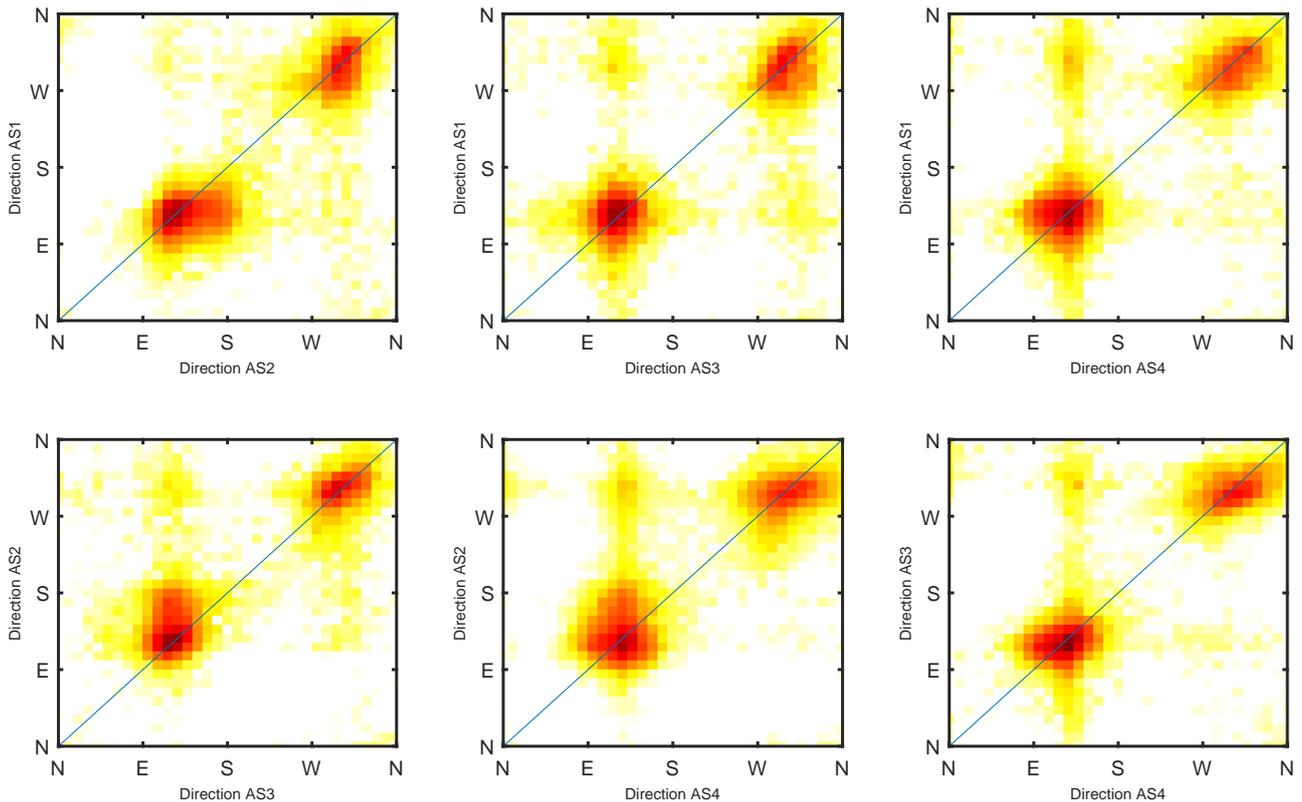


FIGURE 3.40 : Comparaison des directions des anémomètres soniques AS1, AS2, AS3 et AS4 (moyennes 5 minutes pour l'ensemble des données de la campagne KASCADE 2017).

Tout d'abord, les directions des quatre anémomètres soniques placés dans le fond de la vallée (AS1–4) sont comparées entre elles (figure 3.40). Il y a une bonne corrélation entre les directions des quatre stations.

Malgré ce bon accord global, il y a une dispersion expliquée par une variabilité temporelle importante. Même le vent de vallée qui est pourtant assez régulier en direction subit des changements brutaux de directions toujours visibles avec des moyennes 5 minutes (figure 3.41). Cette variabilité temporelle n'est pas surprenante parce que comme l'a montré [Mahrt \(2011\)](#), bien que la variabilité temporelle des vents diminue dans le cas d'un vent canalisé, elle augmente quand la vitesse de vent diminue, or le vent de vallée ici dépasse rarement des valeurs de 1 m s^{-1} .

La variabilité temporelle n'explique pas d'autres différences visibles sur la figure 3.40, notamment des situations de vents de sud-est (descendant la vallée) observés sur la station AS4 alors qu'ils sont opposés sur les autres stations. Cela peut-être le signe d'un décalage temporel dans l'apparition et/ou la disparition du vent de vallée par rapport aux autres stations (cette hypothèse sera analysée par la suite). De même, la station AS3 voit quelques vents avec des

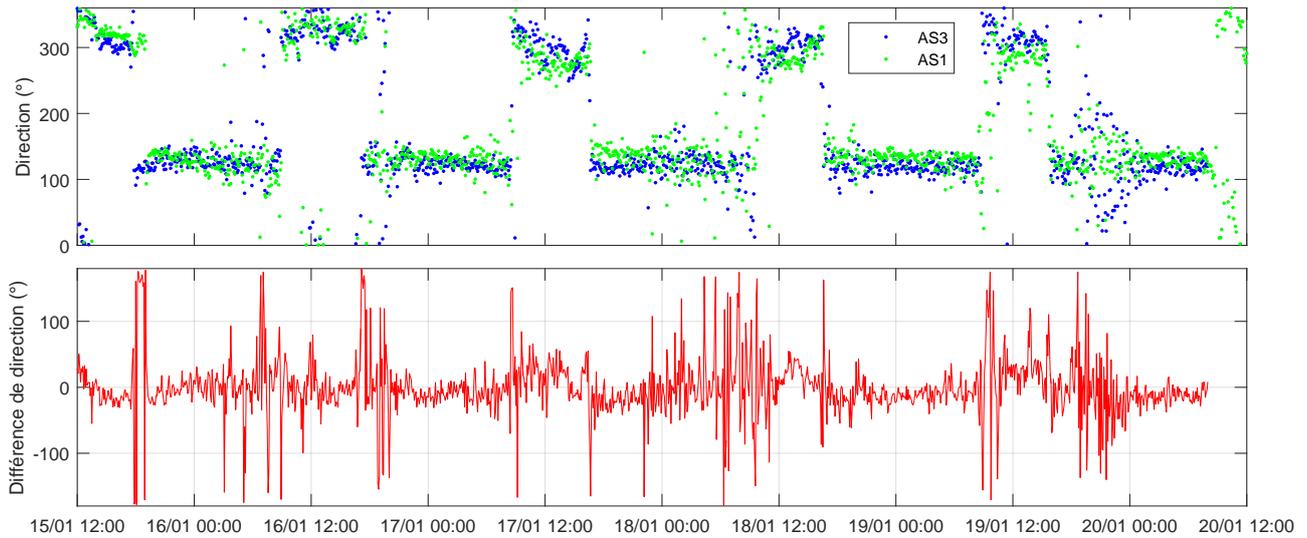


FIGURE 3.41 : En haut : directions observées sur les stations AS1 et AS3. En bas : différence de direction entre les stations AS1 et AS3. Ce sont des moyennes sur 5 minutes pour les deux graphes.

directions opposées à celles des stations AS2 et AS1. Les 2 stations les plus basses (AS1 et AS2) sont mieux corrélées et il n’y a pas de différences notables dans la fréquence d’apparition des vents de vallée ce qui montre que apparition et disparition de ce vent sont synchronisées sur ces deux stations.

La grande variabilité temporelle complique la détection précise des heures d’apparition et de disparition du vent de vallée. Cependant, sur la base de quelques critères simples, il est possible de les détecter sur chacune des quatre stations. Les critères sont les suivants :

- lorsque les conditions sont stables ($\Delta\theta_{GBA} > 1\text{ K}$), les vents ayant des directions comprises entre 90° et 170° , ce qui fait $\pm 40^\circ$ par rapport à l’orientation de l’axe de la vallée, sont considérés comme des vents de vallée (CDV) ;
- le vent CDV apparaît quand les directions sur 6 échéances consécutives (soit une durée de 30 minutes) ont des directions CDV ;
- il disparaît quand 6 directions consécutives ne sont pas comprises entre 90° et 170° .

Le résultat de cette méthode appliquée à la période du 15 au 20 janvier 2017 est représenté sur la figure 3.42 pour la station AS4. Les périodes de présence du vent de vallée détectées par cette méthode sont visuellement cohérentes avec les observations des directions du vent. La variabilité temporelle des directions est bien visible sur les cinq nuits présentées, et est particulièrement marquée en fin de nuit le matin du 16, en début de nuit le soir du 16 et au milieu de la nuit du 19 au 20 janvier où le vent de vallée est parfois intermittent (directions non comprises dans l’intervalle $90\text{--}170^\circ$). Malgré cela, la méthode diagnostique la présence du vent CDV sur ces périodes. Le choix d’analyser les directions sur des périodes de 30 minutes permet donc de ne pas avoir un signal pollué par de courtes intermittences durant lesquelles le vent CDV disparaît.

L’application de cette méthode aux quatre anémomètres soniques placés dans le fond de la vallée permet de détecter quelques décalages dans l’apparition et la disparition du vent de vallée entre les différentes stations (figure 3.43). Le vent de vallée apparaît d’abord dans la partie amont de la vallée (station AS4), puis atteint la station AS3 10 à 30 minutes plus tard. L’ordre d’apparition du vent sur les 2 stations du bas de la vallée varie suivant les nuits. Sur les nuits du 15 au 18 janvier, le vent apparaît d’abord sur la station AS2 puis sur la station AS1 alors que c’est l’inverse sur les 2 nuits suivantes. Cependant le décalage dans les apparitions du

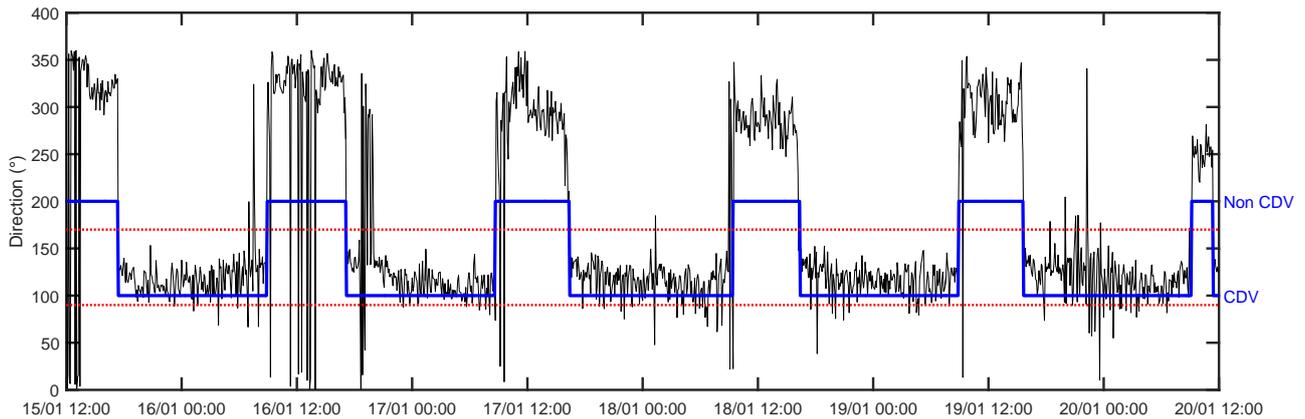


FIGURE 3.42 : Évolution temporelle de la direction du vent mesurée sur la station AS4 et moyennée sur 5 minutes (en noir) et estimation de la présence ou non du vent de vallée (en bleu). Les lignes pointillées rouges marquent les directions 90 et 170°.

vent CDV sur ces deux stations est faible, entre 10 et 20 minutes.

Le délai nécessaire pour que le vent de vallée s'établisse sur toute la longueur de la vallée est également variable suivant les nuits. Quarante-cinq minutes sont suffisantes dans la nuit du 19 au 20 janvier alors que 100 minutes (1h40) ont été nécessaires dans la nuit du 15 au 16 janvier. L'heure d'apparition par rapport au coucher du soleil est également variable.

Plus l'apparition du vent CDV au niveau de l'AS4 est tardive et plus le délai nécessaire à ce que ce vent s'établisse jusque dans le bas de la vallée est long. Par exemple, durant les deux premières journées présentées sur la figure 3.43, le vent de vallée apparaît sur l'AS4 après le coucher du soleil et il met entre 1h et 1h40 pour atteindre le bas de la vallée. Durant les trois jours suivants, le vent apparaît à la tombée de la nuit voire légèrement avant dans la nuit du 19 au 20 janvier, et le temps nécessaire à ce que ce vent atteigne le bas de la vallée varie entre 40 et 55 minutes. Les valeurs de \hat{H} permettent d'apprécier l'effet de la stratification et celui du forçage par les vents au-dessus de la vallée. Les débuts de nuit des deux premières journées sont marqués par de faibles valeurs de \hat{H} (développement tardif de la stratification et vent fort en altitude) ce qui explique l'apparition tardive du vent de vallée. Au contraire, les trois nuits suivantes sont marquées par de fortes valeurs de \hat{H} rapidement après le coucher du soleil ce qui est cohérent avec une apparition rapide du vent de vallée.

La disparition du vent de vallée se produit entre 1h30 et 2h30 après le lever du soleil en fonction des stations et des journées. Il n'apparaît pas clairement d'ordre particulier entre les 4 stations dans la disparition du vent, par contre le décalage entre les stations est très variable. Le matin du 16 janvier, la disparition du vent de vallée est quasi simultanée sur les 4 stations puisqu'il n'y a qu'un décalage de 5 minutes entre elles alors qu'il y a près de 1h30 le matin du 18 janvier. Pour les autres nuits, la disparition du vent de vallée est plus rapide (entre 20 et 25 minutes).

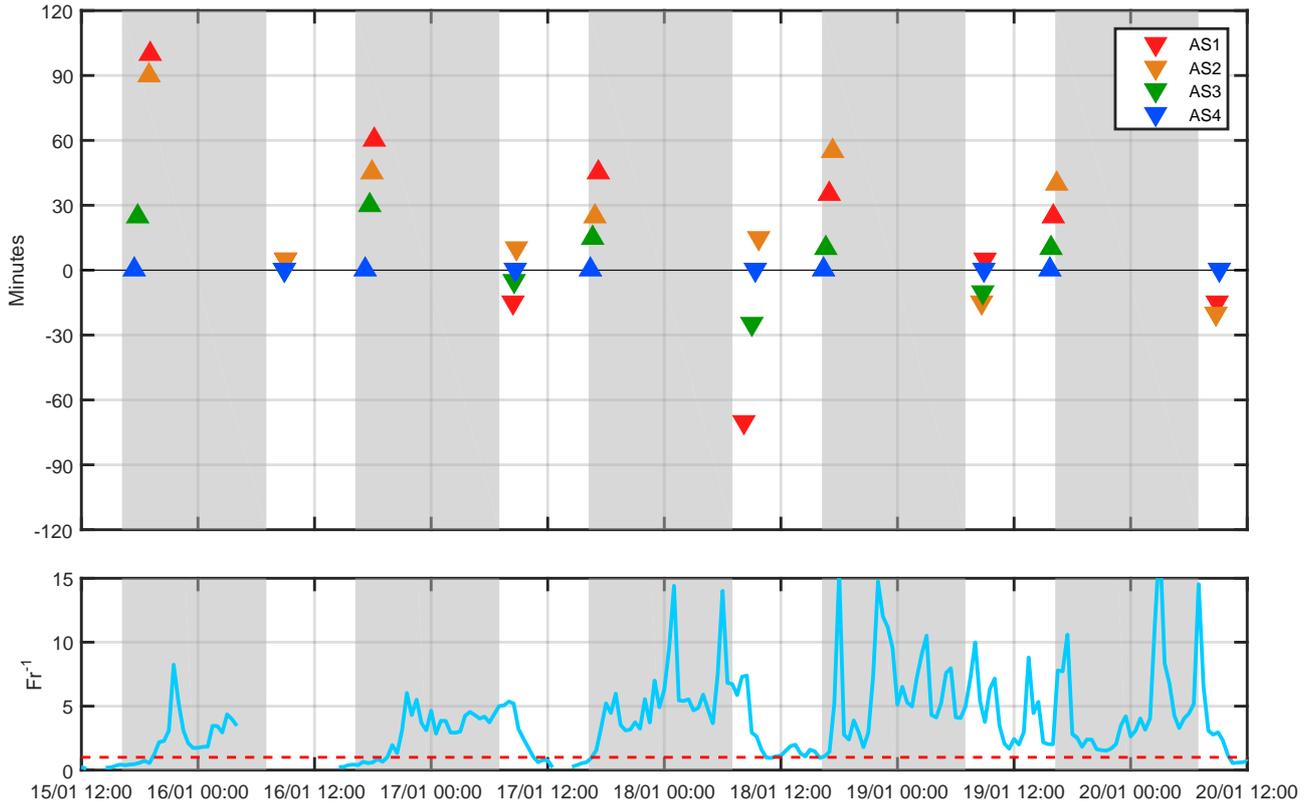


FIGURE 3.43 : En haut : heures d'apparition (triangle vers le haut) et de disparition (triangle vers le bas) du vent de vallée de Cadarache observé avec les anémomètres soniques. Les apparitions/disparitions du vent sont également représentées en ordonnée en minutes par rapport à l'apparition/disparition du vent observé sur la station AS4 afin de mieux visualiser les décalages entre les 4 stations qui sont peu lisibles sur l'axe des abscisses. En bas : profondeur adimensionnelle de la vallée $\hat{H} = Fr^{-1}$ en bleu (moyennes 30 minutes) et $\hat{H} = 1$ en ligne tiretée rouge. Les zones grisées représentent la nuit.

3.5 Conclusion

La campagne de mesures KASCADE 2017 a eu lieu sur le site de Cadarache pendant l'hiver 2016–2017. Elle fait suite à une précédente campagne, KASCADE 2013, dont le but était l'étude de la structure verticale de l'atmosphère par conditions stables dans la vallée de Cadarache. KASCADE 2017 avait donc pour but de compléter la connaissance apportée par KASCADE 2013 en spatialisant les observations dans la vallée de Cadarache afin d'en étudier l'hétérogénéité.

Les LEMS ont fourni une mesure de la stratification sur les versants de la vallée, ce qui a permis la comparaison des profils de température avec ceux effectués dans le centre de la vallée par radiosondage (et avec la station GBA). La stratification est de type cold pool marginale (proche de la cold pool mature) d'après les régimes théoriques de stratification décrits par Mahrt (2017), c'est à dire que les isentropes ne sont pas tout à fait horizontales. Ce régime est associé à des écoulements de pente qui permettent le drainage d'air froid. L'étude des écoulements dans la vallée de Cadarache a d'ailleurs montré que des vents de pente existent sur les versants de la vallée et alimentent un vent de vallée présent dans le fond, ce qui est cohérent avec le régime de cold pool marginale.

Whiteman and Doran (1993) ont décrit les différentes relations pouvant exister entre les

écoulements dans et au-dessus d'une vallée, avec notamment des cas de couplage ou découplage. Plusieurs études sur différentes vallées ont tenté de déterminer les conditions conduisant à l'une ou l'autre de ces situations. Pour cela, une valeur seuil de la profondeur adimensionnelle de la vallée ($\hat{H} = Fr^{-1}$), calculée à partir de la pulsation de Brunt-Väisälä, de la vitesse du vent au-dessus de la vallée et de la profondeur de la vallée, permettant cette distinction est recherchée. La valeur seuil associée à la vallée de Cadarache vaut $\hat{H} = 1$ ce qui est cohérent avec les valeurs trouvées dans les précédentes études.

Les stations avec anémomètres soniques situées dans le fond de la vallée ont permis d'apprécier les décalages temporels dans l'apparition et la disparition du vent de vallée. Le vent apparaît d'abord dans la partie la plus en amont de la vallée puis se développe progressivement le long de la vallée jusqu'à atteindre son débouché. Le délai nécessaire à ce que ce vent s'établisse sur toute la longueur de la vallée est variable suivant les nuits, entre 45 minutes et 1h40 pour les 5 nuits étudiées. Ces différences ont été attribuées aux vents forts au-dessus de la vallée qui contrarient la formation du vent dans le bas de la vallée. Pour les mêmes raisons, les heures d'apparition de ce vent par rapport au coucher du soleil sont variables d'une nuit à l'autre. Les vents plus forts en altitude retardent l'apparition du vent de vallée et rallongent le délai nécessaire à ce qu'il occupe toute la longueur de la vallée. En revanche, il n'y a pas de tendance claire sur la chronologie de disparition de ce vent.

Deux phénomènes particuliers ont été observés. Le premier s'est répété fréquemment durant la période d'observations continue. Lorsque l'atmosphère est stratifiée, les profils de température effectués sur une verticale dans le centre et sur les versants de la vallée montrent bien la stratification alors que les températures à 1,5 m augmentent en descendant le long de la vallée. Seule la station B3 (qui est située environ à mi-longueur de la vallée) a des températures cohérentes avec celles formant le profil sur les versants de la vallée. Les deux stations situées le plus bas dans la vallée connaissent des températures plus chaudes que celles de la station B3 ce qui a été associé à un mélange plus important dans la partie aval de la vallée. Les hypothèses pouvant expliquer cette augmentation du mélange dans le fond de la vallée sont une interaction avec les écoulements de la vallée de Durance ou une augmentation de la turbulence constatée dans le vent de vallée de Cadarache. Enfin, les températures plus froides observées sur la station située le plus en amont dans la vallée peuvent être associées à une stagnation plus importante d'air froid, bien que le drainage de cet air froid vers l'aval ait bien lieu. Ce phénomène reste encore à étudier pour en améliorer la compréhension.

Enfin, le deuxième phénomène particulier, les oscillations des directions du vent de vallée, a également été observé fréquemment. Ces oscillations ont principalement été observées sur les vents (composantes longitudinales et transverses à l'axe de la vallée, vitesse et direction) ainsi que de manière moins évidente sur les températures. Elles produisent des vents dont les directions oscillent généralement autour de la direction moyenne du vent CDV. La période d'oscillation est généralement de 10 minutes, mais de rares cas d'oscillations avec une période de 5 minutes ont été observés. Enfin, un cas d'intermittence dans la présence du vent CDV a été observé, avec une période de répétition d'environ une heure. Plusieurs études, sur d'autres sites, ont rapporté la présence d'oscillations dans le vent de vallée et l'intermittence dans la présence du vent CDV semble correspondre à la théorie énoncée par [Fleagle \(1950\)](#). Toutefois, l'étude de ce phénomène, qui ne faisait pas partie des objectifs de cette thèse, est restée descriptive et mérite un approfondissement afin de comprendre ses mécanismes de formations.

Chapitre 4

Prévision immédiate statistique du vent dans la vallée de Cadarache

L'étude des conditions stables pendant les campagnes de mesures KASCADE (cf. Duine et al. (2017) et le chapitre 3) a montré que la structure des écoulements dans et au-dessus de la vallée de Cadarache est particulièrement complexe. L'instrumentation continue disposée dans cette vallée n'est alors pas suffisante pour documenter cette hétérogénéité. KASCADE 2013 a ainsi permis d'observer la structure verticale de l'écoulement, et le découplage entre les vents dans et au-dessus de la vallée par conditions stables. Or, les besoins du CEA Cadarache pour le calcul de la dispersion atmosphérique portent en premier lieu sur les vents au niveau des sources d'émissions, qui sont majoritairement situées dans la vallée et relativement proches de la surface.

Duine et al. (2016) ont alors développé une technique de prévision immédiate visant à diagnostiquer la présence ou l'absence d'un vent de vallée à partir des observations continues réalisées avec la station GBA (voir la partie 2.4). La méthode a été validée, mais elle fournit une information limitée à la présence ou l'absence du vent dans le fond de la vallée puisque ni la vitesse ni la direction du vent ne sont calculées.

Dans le but d'améliorer cette première méthode de prévision immédiate, un réseau de neurones artificiels (RNA) a été utilisé pour calculer la direction et la vitesse du vent dans la vallée à partir des observations continues de la station GBA. Les RNA sont des outils statistiques qui permettent l'approximation de relations, notamment non-linéaires. Ils sont alors fréquemment utilisés dans le but de déterminer des fonctions reliant des variables, comme dans notre cas le vent dans la vallée et les observations de la station GBA. L'autre atout des RNA pour un tel calcul est leur disponibilité au travers de logiciels qui facilitent leur mise en œuvre, tel que MATLAB par exemple.

La méthode utilisée est décrite en détail ci-dessous, puis les résultats sont présentés sous la forme d'un article publié dans le Journal of Applied Meteorology and Climatology.

Plan du chapitre

4.1	Fonctionnement des RNA	110
4.2	Description de la structure des réseaux utilisés	111
4.3	Conclusion	112
4.4	Article	113

4.1 Fonctionnement des RNA

La propriété fondamentale des réseaux de neurones est l'approximation non linéaire parcimonieuse, c'est à dire qu'ils peuvent approcher une fonction non-linéaire en utilisant moins de paramètres ajustables (leur nombre est proportionnel au nombre de neurones) qu'une autre méthode statistique à précision égale. C'est pour cette raison qu'ils sont fréquemment utilisés comme outil de régression. Il y a toutefois quelques limites à l'utilisation d'un réseau de neurone. Ils sont très difficiles à interpréter physiquement. [McCann \(1992\)](#) dit des RNA qu'ils sont très utiles pour l'estimation de relations mais leur défaut est qu'ils sont presque impossibles à comprendre. Cependant, en construisant plusieurs réseaux avec différentes combinaisons de variables, il est possible de déterminer quelles variables sont les plus pertinentes pour représenter le phénomène étudié.

Les RNA sont des outils mathématiques composés de « neurones artificiels » et qui ont pour but de lier des variables par une relation non-linéaire (figure 4.1). Chaque neurone est une fonction qui effectue un calcul avec les différentes variables qui lui sont données en entrée et qui produit un résultat qui est appelé la sortie. Ce calcul est explicité dans l'équation 4.1 pour N variables d'entrée :

$$Y = f \left(W_0 \cdot b + \sum_{n=1}^N W_n \cdot X_n \right) \quad (4.1)$$

avec X_n la $n^{\text{ième}}$ variable d'entrée, W_n les paramètres ajustables appelés poids, b un biais (souvent égal à 1) et f la fonction d'activation (le plus souvent la fonction tangente hyperbolique ou une fonction linéaire). Un réseau de neurones est alors constitué de plusieurs de ces neurones qui sont interconnectés et distribués sous forme de couches, la dernière produisant la sortie du système (figure 4.1). Les sorties produites par ces neurones sont comparées à des valeurs cible, le but étant de se rapprocher le plus possible de la cible. Un score de performance est alors calculé pour évaluer la capacité du réseau à bien estimer les sorties souhaitées. Si ce score n'est pas suffisamment élevé, le réseau recommence le même calcul en changeant la valeur des poids : c'est la phase d'apprentissage.

Le nombre de couches et de neurones optimums sont dépendants du problème. Un trop grand nombre amène à un sur-apprentissage qui produit une fonction qui tend à restituer dans la sortie le bruit contenu dans les mesures, et un nombre trop faible conduit à un sous-apprentissage qui se traduit par une difficulté à estimer une relation entre les entrées et la sortie ciblée. Dans le cas du sur-apprentissage, le réseau donne de très bons résultats à partir du jeu de données qu'il utilise pour son apprentissage, mais il a du mal à généraliser sur un jeu de données indépendant de son apprentissage.

Le jeu de données utilisé pour construire la fonction du RNA doit être assez important et assez représentatif, de manière à inclure le plus possible de situations que le RNA est censé représenter. Il est divisé en trois parties :

- le jeu d'apprentissage qui sert à l'ajustement des poids ;
- le jeu de validation qui permet de calculer la performance du réseau (l'apprentissage se termine quand la performance calculée sur ces données atteint un niveau fixé ou n'augmente plus) ;
- le jeu de test, dont les données ne sont pas utilisées pour construire le RNA, mais qui est nécessaire pour la comparaison de différents RNA (sur des données indépendantes de l'apprentissage du RNA) dont les configurations (et donc l'apprentissage) sont différentes.

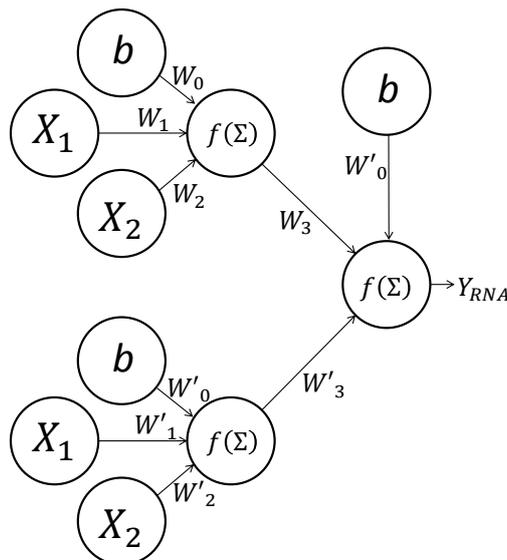


FIGURE 4.1 : Représentation schématique de la structure d'un réseau de neurones (de type « feedforward ») à deux couches possédant deux neurones cachés et un neurone de sortie produisant la sortie Y_{RNA} . Il y a deux variables en entrée du réseau (X_1 et X_2). Les W_i sont les poids, b est le biais et f est la fonction d'activation décrite dans l'équation 4.1.

4.2 Description de la structure des réseaux utilisés

Le RNA utilisé est intégré dans MATLAB (Beale et al., 1992) et il est appelé « perceptron multi-couche ». Les RNA de ce type sont très largement utilisés dans les études atmosphériques (Gardner and Dorling, 1998). Ils sont composés d'une ou plusieurs couches cachées et d'une couche de sortie (qui correspond à la sortie du réseau). Ce sont des réseaux de type « feedforward » (ou « non bouclés »), c'est-à-dire que le calcul se déroule uniquement dans le sens de la première couche vers la couche de sortie. De plus, chaque sortie de neurone est connectée à chaque neurone de la couche suivante pour laquelle il sert donc d'entrée.

Le RNA est appliqué dans le but de calculer la vitesse et la direction du vent en un point à mi-longueur de la vallée de Cadarache (position de la station MET01). Il a été choisi de calculer les composantes u et v du vent horizontal, puis d'en déduire la direction et la vitesse du vent, plutôt que directement ces derniers afin d'éviter les problèmes liés au caractère cyclique de la direction. Le réseau utilisé est donc composé d'une couche de sortie de 2 neurones (un pour chaque composante du vent). Les variables en entrée sont choisies parmi les grandeurs mesurées par la station GBA (à l'aide d'un processus qui sera décrit dans la partie 4.4 et l'annexe C).

Les choix de paramétrisation ont porté sur le nombre de neurones de la couche cachée, le type de fonction d'activation et l'algorithme d'apprentissage. Le nombre de neurones de la couche cachée a été déterminé à partir de différents essais dans lesquels ce nombre a varié de 5 à 50. Il s'est avéré qu'au-delà de 10 neurones les performances ne sont plus améliorées de façon significative. C'est donc cette configuration à 10 neurones qui a été retenue. Il est à noter que la phase d'apprentissage du RNA est d'autant plus rapide que le nombre de neurones est faible, cependant, le temps de calcul n'est que de quelques secondes ce qui explique que ce critère n'ait pas joué de rôle décisif dans le choix du nombre de neurones. Parmi les fonctions d'activation couramment utilisées (la fonction tangente hyperbolique, une fonction linéaire ou encore une

fonction seuil) c'est la fonction tangente hyperbolique qui a été choisie (faisant de ce RNA un perceptron multi-couche).

L'apprentissage, c'est-à-dire l'amélioration de la prévision d'une sortie par modification des poids, est effectué par un algorithme de descente de gradient appelé rétropropagation du gradient. Cet algorithme calcule le biais entre la sortie du réseau et la variable cible que l'on cherche à obtenir et propage l'erreur de la sortie vers l'entrée afin de calculer les modifications à apporter aux poids. Cette technique permet de calculer un gradient local de l'erreur par rapport aux poids, ceux-ci étant alors corrigés dans la direction de descente maximale du gradient. L'algorithme utilisé est celui de Levenberg-Marquardt parce qu'il est décrit comme étant l'algorithme permettant l'apprentissage le plus rapide (Beale et al., 1992) et que les autres algorithmes disponibles dans MATLAB qui ont été testés (méthode de Broyden-Fletcher-Goldfarb-Shanno et Scaled Conjugate Gradient) n'ont pas apporté d'amélioration significative des résultats.

La division du jeu de données est réalisée aléatoirement par le logiciel. Il est recommandé de garder une majorité de données dans le jeu d'apprentissage. Tout en respectant cette consigne, nous avons choisi de sélectionner un grand nombre de données pour le jeu de test afin d'élever la significativité de la performance calculée sur ces données. Ainsi, les proportions retenus sont 60 % pour le jeu d'apprentissage et 10 % pour le jeu de validation. Les 30 % restants composent le jeu de test, ce qui permet de garder un grand nombre de données indépendantes de l'apprentissage. Enfin, l'apprentissage du réseau est stoppé quand la performance (évaluée à partir de l'erreur quadratique moyenne) calculée sur le jeu de validation n'est plus améliorée pour six (valeur par défaut dans MATLAB) changements successifs des valeurs de poids (c'est à dire quand l'erreur quadratique moyenne ne diminue plus sur six itérations consécutives).

Les données sont préalablement normalisées par le logiciel dans l'intervalle -1 à 1 avant d'être entrées dans le réseau afin que toutes les variables aient le même ordre de grandeur. En effet, si deux variables étaient d'un ordre de grandeur différent, la variable la plus faible aurait moins d'impact sur la sortie.

La dernière étape de configuration du réseau consiste à diminuer la dimension des entrées en supprimant les variables redondantes ou inutiles. Ce choix est établi tout d'abord à partir de l'étude des corrélations entre les différentes entrées possibles, puis par une évaluation de l'apport de chaque variable au résultat global (voir la partie 4.4 et l'annexe C).

4.3 Conclusion

Dans l'article reproduit ci-après, le réseau de neurones artificiels, une fois configuré, est mis en œuvre sur des périodes pendant lesquelles on dispose de l'observation continue du vent dans le fond de la vallée nécessaire à l'apprentissage. Pour les 3 mois de la période de KASCADE 2013, les résultats montrent une nette amélioration par rapport à la précédente méthode de Duine et al. (2016) du point de vue du diagnostic de la présence ou non du vent de vallée, et de plus apportent une information supplémentaire, à savoir la vitesse et la direction du vent. Ces deux dernières variables sont représentées de manière satisfaisante puisque le biais pour la vitesse n'est que de $-0,28 \text{ m s}^{-1}$ et 84 % des directions sont prévues avec une erreur inférieure à 45° . La qualité des résultats obtenus sur un an de données est du même ordre (voir dans la partie 4.4 et plus en détail dans l'annexe C).

Bien que les résultats aient été jugés satisfaisants, quelques limitations ont été rencontrées. Premièrement, les directions des vents faibles sont les plus difficiles à reproduire du fait de la méthode de calcul basée sur les composantes. En effet, une erreur absolue (même faible) sur une composante d'un vent faible impacte la direction de manière significative. Deuxièmement, les vents avec une direction transverse à l'axe de la vallée sont les moins bien représentés pour plusieurs raisons : d'une part ces vents sont très faibles, donc leur direction est difficile à reproduire pour le RNA, d'autre part ils sont moins fréquents que les autres vents, et enfin ils ne résultent pas d'une situation bien identifiée de la météorologie locale.

Ces résultats ont été obtenus en utilisant seulement trois variables en entrée du RNA : la différence verticale de température potentielle entre 110 m et 2 m et les deux composantes du vent horizontal à 110 m mesurées sur la station GBA. Cela montre que le RNA est capable de calculer un vent local au fond de la vallée à partir de conditions représentatives de l'écoulement au-dessus de la vallée avec une précision satisfaisante. Pour cette raison, cette méthode a été appliquée aux simulations WRF à 3 km de résolution horizontale (décrites dans le chapitre 2) avec un objectif de descente d'échelle dans la prévision du vent. C'est ce qui sera présenté dans le chapitre 5.

4.4 Article

Local-Scale Valley Wind Retrieval Using an Artificial Neural Network Applied to Routine Weather Observations

FLORIAN DUPUY

Laboratoire d'Aérodynamique, Université de Toulouse, CNRS, UPS, Toulouse, and CEA, DEN, Cadarache, Laboratoire de Modélisation des Transferts dans l'Environnement, Saint-Paul-lès-Durance, France

GERT-JAN DUINE

Earth Research Institute, University of California, Santa Barbara, Santa Barbara, California

PIERRE DURAND

Laboratoire d'Aérodynamique, Université de Toulouse, CNRS, UPS, Toulouse, France

THIERRY HEDDE AND PIERRE ROUBIN

CEA, DEN, Cadarache, Laboratoire de Modélisation des Transferts dans l'Environnement, Saint-Paul-lès-Durance, France

ERIC PARDYJAK

Department of Mechanical Engineering, University of Utah, Salt Lake City, Utah

(Manuscript received 3 July 2018, in final form 14 February 2019)

ABSTRACT

We hereby present a new method with which to nowcast a thermally driven, downvalley wind using an artificial neural network (ANN) based on remote observations. The method allows the retrieval of wind speed and direction. The ANN was trained and evaluated using a 3-month winter-period dataset of routine weather observations made in and above the valley. The targeted valley winds feature two main directions (91% of the total dataset) that are aligned with the valley axis. They result from downward momentum transport, channeling mechanisms, and thermally driven flows. A selection procedure of the most pertinent ANN input variables, among the routine observations, highlighted three key variables: a potential temperature difference between the top and the bottom of the valley and the two wind components above the valley. These variables are directly related to the mechanisms that generate the valley winds. The performance of the ANN method improves on an earlier-proposed nowcasting method, based solely on a vertical temperature difference, as well as a multilinear regression model. The assessment of the wind speed and direction indicates good performance (i.e., wind speed bias of -0.28 m s^{-1} and 84% of calculated directions stray from observations by less than 45°). Major sources of error are due to the misrepresentation of cross-valley winds and very light winds. The validated method was then successfully applied to a 1-yr period with a similar performance. Potentially, this method could be used to downscale valley wind characteristics for unresolved valleys in mesoscale simulations.

1. Introduction

Low-level winds must be well described when studying atmospheric dispersion. Stable conditions are among those that most impact pollution because they tend to

trap emissions close to the surface, leading to high-pollutant concentrations. Local topography can generate a variety of terrain-induced phenomena near the surface (Chow et al. 2012), including channeling effects which dominate downward momentum transport (Whiteman and Doran 1993). Over complex topography, stable conditions can generate thermally driven flows along slopes and within valleys (Muñoz et al. 2013; Duine et al. 2017;

Corresponding author: Thierry Hedde, thierry.hedde@cea.fr

Simpson 1994) as well as produce cold pools (Burns and Chemel 2015; Lareau et al. 2013; Price et al. 2011; Clements et al. 2003), which also affect atmospheric dispersion. Numerous experiments have been conducted to study density-driven valley and slope flows and/or cold air pooling, including VTMX (Doran et al. 2002), COLPEX (Price et al. 2011), METCRAX (Whiteman et al. 2008), MATERHORN (Fernando et al. 2015), ASCOT (Clements et al. 1989), and more recently, the Passy project (Sabatier et al. 2018). Most of these and many other complex-terrain studies have been recently reviewed in Rotach et al. (2017).

The Katabatic Winds and Stability over Cadarache for Dispersion of Effluents (KASCADE) experiment (Duine et al. 2017) was conducted in the winter of 2013 in the Cadarache area (southeastern France), in order to investigate the influence of stability on local winds in a small valley. The site features a small valley, where the wind flow is insufficiently monitored by routine meteorological observations to properly calculate the dispersion conditions therein. The currently available meteorological observations can be used, however, to statistically assess winds in the narrow valley. Consequently, Duine et al. (2016, hereafter D16), developed a method for nowcasting the occurrence of downvalley winds that form under stable conditions in the Cadarache valley using observations from the KASCADE campaign. They found that among the permanent measurements available, a valley-scale vertical temperature gradient is a good indicator of the presence of downvalley winds. This method represents the first step toward downscaling local winds; it is, however, restricted to discriminating between the presence or absence of thermally driven winds and is unable to predict the wind speed in the valley. One practical and useful improvement would thus be to nowcast the actual wind speed and direction in the valley using the Cadarache routine observations.

In this paper, we build upon the work of D16 by applying an artificial neural network (ANN) to routine measurements in order to infer the valley winds in the Cadarache valley. Among the nowcasting tools for atmospheric studies, ANNs are increasingly used. Gardner and Dorling (1998) provide an overview of the scope of its utilization in atmospheric studies. An ANN is a statistical tool that reveals the correlation of variables, even if it is unaware of the physical link between the variables. Its principle is to find the best relationship between a set of variables (input variables, always available) so as to calculate an output that is the closest to target variables (observations in our case). For example, Delon et al. (2007) successfully determined nitrogen oxide soil emissions from environmental variables. ANNs have been used for the estimation of an interpolated wind speed field

from observations, (Philippopoulos and Deligiorgi 2012) and for predictions on short time scales using a feedforward technique (Cadenas and Rivera 2009; More and Deo 2003), with better performance than more traditional statistical methods. The study of Khosravi et al. (2018) relates the utilization of an ANN to derive the wind direction. In their study, these authors directly derived the wind direction from the ANN. However, we think that the continuous mathematical functions used in ANNs are unable to predict a cyclic variable such as the wind direction, which is prone to jumps at the $0/360^\circ$ (or $\pm 180^\circ$) transitions. For this reason, it seems more judicious to estimate the wind direction from the ANN-computed wind components.

The aim of this study is to build an ANN for nowcasting winds in the Cadarache valley, using continuously observed variables as inputs. The ANN enhances the nowcasting work of D16 by computing the actual wind speed and direction at one specific point in the Cadarache valley. The paper is organized as follows. In the second section, the study site and data used are described. The third section is dedicated to the methodology with a presentation of the ANN, its utilization, and how to assess its performance. The results of the ANN are then presented and discussed in the penultimate section of this paper, including a comparison with other techniques. There will be a conclusion in the last section.

2. Site and data

a. Site description

We focus our study on a region of southeastern France, approximately 50 km north of the Mediterranean Sea in the Alpine foothills that are composed of many small valleys and hills (Fig. 1). The Cadarache valley is oriented in a southeast to northwest direction. It is approximately 6 km long, 100 m deep, 1–2 km wide, and has a mean slope of 1.2° along its main axis. The winds in the valley have a clearly distinct diurnal pattern with up- and downvalley directions. The Cadarache valley is one of the many tributaries of the Durance valley, which is of a larger scale (in this area, it is oriented north-northeast to south-southwest, is approximately 67 km long, 200 m deep, and 5 km wide, and has a mean slope of 0.2° along its main axis).

Many meteorological conditions influence the wind patterns in the Cadarache valley. Under stable conditions and clear skies, a clear pattern of thermally driven up- and downvalley flows is found. This behavior was extensively documented during the KASCADE 2013 campaign. Other typical regional meteorological phenomena influencing the local wind patterns are the mistral (Bastin et al. 2005) and precipitation-related

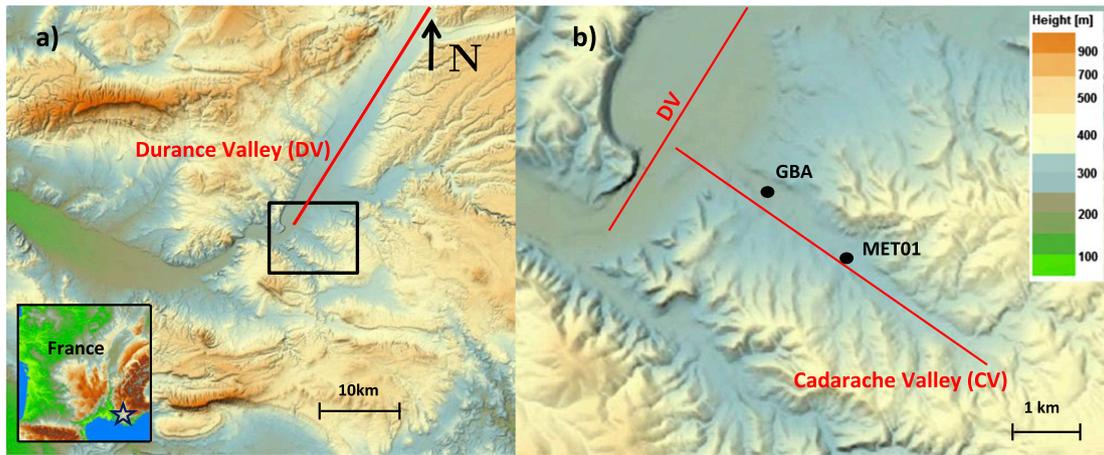


FIG. 1. (a) Topography of the region of Cadarache. (b) Zoomed-in view of the black box in (a). The red lines mark the Durance valley (DV) and the Cadarache valley (CV) axes, and the black dots represent the locations of the meteorological stations MET01 (286 m AGL) and GBA (265 m AGL). (Source: www.geoportail.gouv.fr.)

conditions. The mistral is a northerly regional wind that is often strong and is channeled along the Rhône River valley, spreading to the east and west when it approaches the Mediterranean Sea. A typical mistral situation is shown in Figs. 2c, 2f, and 2i for the period of

14–16 March 2013. Clear skies are generally associated with mistral conditions, and result in high values of shortwave radiation during the day and large radiative cooling at night. Despite this radiative cooling, the high winds prevent the formation of strong stable stratification.

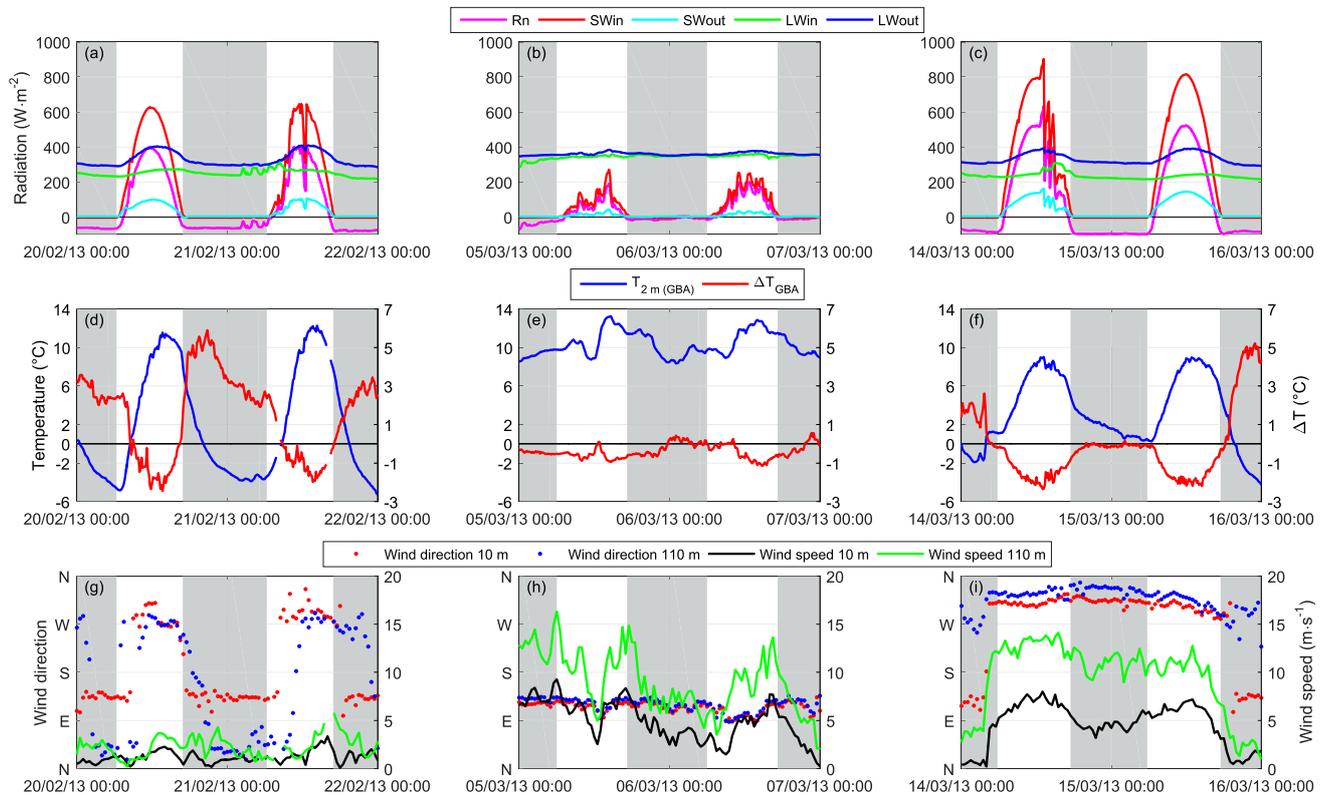


FIG. 2. Meteorological conditions observed in the Cadarache valley during three typical 2-day periods of the KASCADE campaign. (left) Fair-weather conditions, (center) cloudy/rainy conditions, and (right) mistral conditions. (a)–(c) Incoming (in) and outgoing (out) components for shortwave (SW) and longwave (LW) radiation, as well as net radiation Q_{net} measured at location MET01. (d)–(f) The 2-m temperature and the temperature difference between 2 and 110 m measured at GBA (Fig. 1). (g)–(i) The wind direction and wind speed measured at 110 (GBA) and 10 m (MET01). Gray stripes mark nighttime.

TABLE 1. Summary of observations made in the Cadarache valley: T is temperature, P is pressure, Rh is relative humidity, WS is wind speed, WD is wind direction, and u , v , and w correspond to the wind components. The height of measurement is above ground level. See Fig. 1 for stations location.

Station	Height (m)	Observed variables	Dates	Instrument
MET01	2	WS, WD	17 Feb 2015–17 Feb 2016	Campbell Sci. 05103 cup anemometer
	10	u , v , w	13 Dec 2012–16 Mar 2013	Young 81000 sonic anemometer
GBA	2	T	Continuously	Rotronic PT100 thermometer
		Rh	Continuously	Rotronic hygrometer
		P	Continuously	Vaisala PTB101C barometer
	110	WS, WD	Continuously	Metek sonic anemometer
		T	Continuously	Rotronic PT100 thermometer

The low-level winds are therefore forced by upper-level winds resulting in northerly winds at all heights. The region is also often subject to southeasterly winds that usually provoke precipitation events (Berthou et al. 2016), especially during the fall. In this case (Figs. 2b,e,h), the upper-level winds force the direction of the winds in the Cadarache valley to be oriented southeasterly. During the remainder of the time, the region is characterized by the frequent occurrence of weak-to-moderate winds and clear skies leading to stable conditions during calm nights and producing thermally driven downvalley winds. These conditions were investigated during the KASCADE experiment (winter of 2013; Duine et al. 2017). The main goal of KASCADE was to observe the evolution of characteristic winds in the Cadarache valley, particularly during stable conditions. During the daytime, winds are mostly westerly to northwesterly in orientation. This is a result of forcing at larger scales (Fig. 2i). After sunset, when the synoptic forcing is sufficiently weak (clear sky and light winds in altitude), stable conditions can occur, and a thermally driven downvalley wind develops in the Cadarache valley (Figs. 2a,d,g). Its thickness is about 50 m attaining a maximum speed at 25–30 m, as observed during the campaign on tethered-balloon profiles (Duine et al. 2017). The stable situation quickly fades away after sunrise, and is followed by the return of westerly winds. Above the Cadarache downvalley (hereafter CDV) wind, and under the same formation conditions, another wind develops between about 100 and 600 m AGL, the drain current descending the Durance valley with a north-northeast direction (Duine et al. 2017).

b. Observations

Two observation periods were used in this study. First, a 3-month period during the KASCADE 2013 experiment was used to compare the ANN nowcasting method to that of D16. This experiment was conducted during the winter season, with the highest occurrence of stable conditions. A 1-yr period was then used in order to include a larger range of meteorological conditions and to verify the performance of the ANN method.

1) DATA

The data used in this study, detailed in Table 1, comprise two main sets, all composed of hourly averaged observations (1899 exploitable samples over 2232 for the KASCADE 2013 experiment and 7991 over 8760 for the 1-yr period):

- The first set of observations are continuous measurements routinely made at the low end of the Cadarache valley (GBA site in Fig. 1). The GBA station is a 110-m tower that measures pressure, relative humidity, and temperature at 2 m, as well as wind speed and direction and temperature at 110 m, a level situated just above the adjacent ridgelines.
- The second set comprises data from temporary stations over limited periods. These were installed inside the Cadarache valley at the MET01 location near the middle of the valley, 1.6 km upvalley from the GBA station (see Fig. 1). Numerous observations were made here along a vertical profile during the KASCADE experiment between 13 December 2012 and 16 March 2013. Subsequently, MET01 was operated in the same place, instrumented at 2 m only and providing one entire year of data from 17 February 2015 to 17 February 2016.

2) WIND CHARACTERISTICS

We examine wind roses at 110 m during both observation periods (Figs. 3c,d). Both periods exhibit very similar features, with three major wind directions: westerly to northwesterly winds (mix of mistral and typical daytime winds; Figs. 2g,i), north-northeasterly winds aligned with the Durance valley and southeasterly winds generally associated with cloudy/rainy weather (Fig. 2h). The Cadarache valley clearly impacts the wind as nearly all winds observed at 2 m are aligned with the valley axis (Figs. 3a,b). The large-scale, westerly winds are channeled by the valley and tend to take a northwest (upvalley) direction. The southeasterly winds correspond to a mix of valley-channeled winds (Duine et al. 2017) and rainy weather conditions. No typical

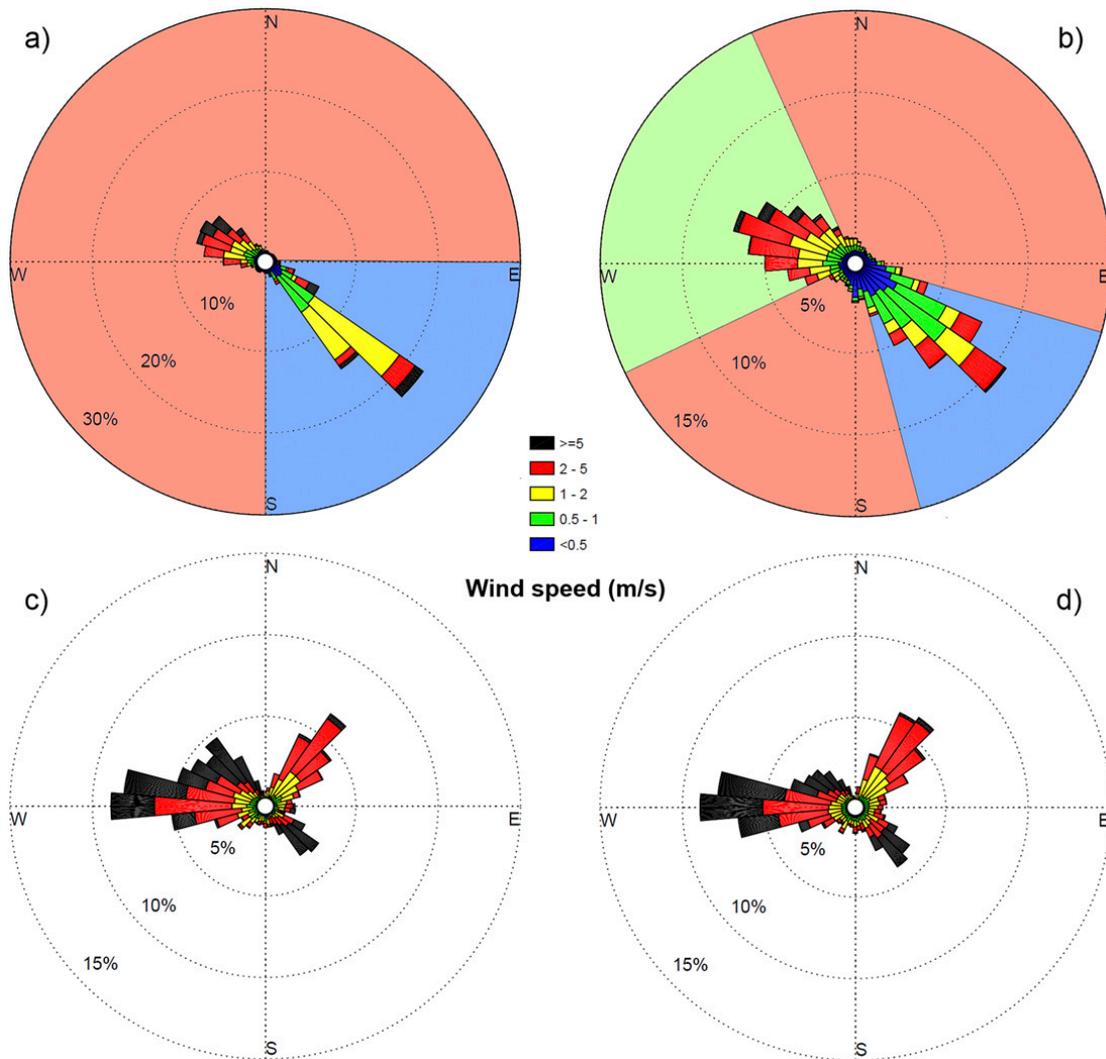


FIG. 3. Wind roses of hourly averaged observations at (a) 10 m from the KASCADE dataset, (b) 2 m from the 1-yr period, (c) 110 m from the KASCADE dataset, and (d) 110 m from the 1-yr period. The blue and red (salmon) sectors in (a) correspond to the classification of D16, whereas the three colors in (b) correspond to the classification used in this paper (see section 3b): the blue sector for the CDV winds extends from 105° to 165° , the green sector for the Cadarache upvalley winds extends from 245° to 335° , and red sectors include all other directions, corresponding to transverse winds.

meteorological conditions could be associated with cross-valley directions.

The conditions during the KASCADE period repeat themselves often with a high occurrence of CDV winds (50% of total dataset, defined as the blue sector on Fig. 3b) and upvalley winds (41% of dataset, green sector on Fig. 3b), and only 9% of cross-valley winds (red sectors on Fig. 3b). Over the 1-yr period, the directions at 2 m are more spread than during the KASCADE period. The channeling of winds is however still clearly visible. The CDV winds represent 38% of the full dataset; there are 40% of upvalley winds and 22% of cross-valley winds. The higher proportion of CDV winds during the KASCADE period (winter) can

be explained by the longer nights favorable for these winds.

We demonstrate the local wind characteristics by relating above-valley winds to within-valley winds (Whiteman and Doran 1993) for all winds (Fig. 4a), and we further discriminate between the low-wind conditions ($< 0.5 \text{ m s}^{-1}$; Fig. 4c) and other wind conditions ($> 0.5 \text{ m s}^{-1}$; Fig. 4b). Two characteristic valley winds appear (Fig. 4b): first, there are two clusters of winds with fairly identical directions at 2 and 100 m. They correspond to westerly-to-northwesterly and southeasterly directions, that is, similar to the orientation of the valley. This is the result of downward momentum transport combined with channeling by the valley walls. It is not easy to distinguish among these two

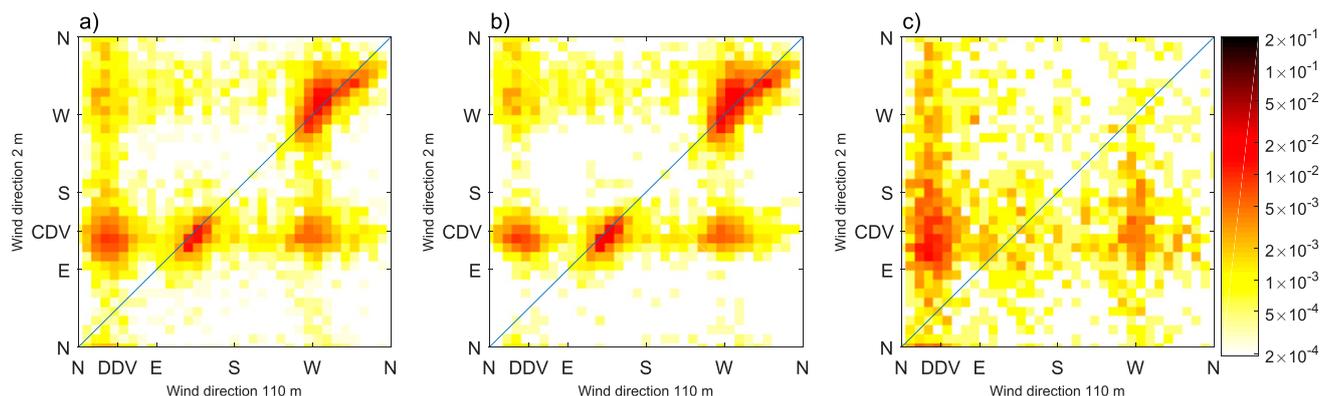


FIG. 4. Probability density showing a comparison of wind directions measured at 110 and 2 m (a) for the full 1-yr dataset, (b) after removing light winds ($<0.5 \text{ m s}^{-1}$ at 2 m), and (c) for light winds only. CDV represents the Cadarache downvalley direction and DDV the Durance downvalley direction.

mechanisms which one is responsible for the determination of directions in the valley because the above valley winds are almost in alignment with the valley axis. Second, there are 2-m CDV winds, with upper winds blowing either from the west or from the NNE. These CDV winds are thermally driven, during stable stratification periods. They are described in detail by Duine et al. (2017). The upper NNE winds are also thermally driven, but by the Durance valley (Duine et al. 2017). The cluster of 2-m CDV winds associated to upper westerly winds corresponds to the early stage of the CDV winds, in the first hours after sunset, when the Durance downvalley (DDV) wind is not yet established (Duine et al. 2017). Similarly, DDV winds persist several hours after sunrise, whereas CDV winds have already ceased, resulting in the dots of the upper-left-hand corners of Figs. 4a and 4b. For the weak wind conditions (Fig. 4c), downward momentum transport is not observed (as expected), but the other mechanisms described above are still present, even though the scatter is larger because of the higher variability of the wind direction in such conditions.

3. ANN methodology

An ANN is a statistical tool, the aim of which is to calculate a mathematical function linking two datasets available over the same period (Yegnanarayana 2009). It is a kind of machine learning algorithm, able to process complex data inputs. The ANN learns to perform tasks by considering examples, generally without being programmed with any task-specific rules. One dataset is composed of known variables [which is the dataset gathering continuous observations in this study described in section 2b(1)] while the other is considered a target to be reproduced [which are temporary observations of the wind in the valley described in section 2b(1)]. The ANN

starts with a randomly chosen function whose inputs are the known variables. This function is composed of weighted activation functions or neurons, interconnected between each other and arranged in layers.

Each of the two datasets is divided into three subsets: training, validation and test. The first step, applied to the training set, consists of fitting the abovementioned function to produce results as close as possible to the target dataset. This step is iterative and lasts until the performance calculated on the validation set reaches a given limit and no longer improves. The test set is useful for comparing different ANNs because it is the only set that is totally independent from the training of the ANN. The performance of the ANNs is calculated therefore only on the test set. Since ANNs are good interpolators but poor extrapolators (Gardner and Dorling 1998), the dataset used to train the ANN has to be as big and heterogeneous as possible, in order to encompass the widest range of cases that the ANN is expected to treat.

a. ANN configuration

The ANN used in this study is a multilayer perceptron (MLP) type (Beale et al. 2010). The software is included in the Deep Learning Toolbox of MATLAB. The MLP is frequently used in atmospheric sciences (Gardner and Dorling 1998). Its specific feature is that each neuron of a layer is connected with all neurons of the previous and next layers. It is composed of an input layer, at least one hidden layer and an output layer. Here, the input layer is composed of the training variables and there is a single hidden layer of 10 neurons. To support this choice, we verified that the performances were not improved by adding a second hidden layer, and we evaluated the evolution of the ANN performance when the number of the neurons in the hidden layer varied from 5 to 50. The output layer (target) is composed of two neurons that are expected to reproduce the horizontal wind components

at 2 m (u and v). These components are used to calculate wind speed and direction. The activation function is a hyperbolic tangent function because it offers more favorable results than other tested functions (e.g., log-sigmoid or linear functions), the Levenberg–Marquardt algorithm was chosen as training algorithm because it is described in the MATLAB toolbox user’s guide as the fastest training algorithm and performances are very similar for all tested algorithms (Levenberg–Marquardt, BFGS Quasi-Newton, and Scaled Conjugate Gradient). The pseudorandom generation of the initial weights is also tested by training several ANNs with different seeds. The seed producing the best results is selected.

The dataset was split into 60% training data, 10% validation data, and the remaining 30% was used as the test subset. With respect to the default values in MATLAB, we increased by 10% the amount of the validation and test sets (40% instead of 30%) in order to avoid over-training, which can occur when too much of the dataset is used for training. All input variables were normalized between -1 and $+1$ in order to avoid weight discrepancy problems, which occur when the order of magnitude of variables differs.

b. ANN performance evaluation

The ANN outputs are evaluated by comparing wind speed and direction with the observations (i.e., targets). The performance of the wind-direction nowcasts was measured by using (i) the DACC (Direction Accuracy; Santos-Alamillos et al. 2013) metric, which represents the proportion of horizontal winds which do not depart by more than 45° from observations, and (ii) the PC (Proportion Correct) metric which indicates the proportion of values correctly classified in different wind sectors defined from the observed wind rose to represent the main wind regimes. Two forms of PC are used. PC2, used by D16, considers two wind sectors: the southeast quarter for downvalley winds for the first class and all other directions for the second class. It is the only criterion used in D16’s nowcasting method. PC4 considers four classes: a first class for light winds (lower than 0.5 m s^{-1}) that have no well-defined direction, then a class of CDV winds, a class of upvalley winds, and a class of cross valley winds. The last three classes correspond to the blue, green, and red regions, respectively, in Fig. 3b. The three classes of PC4 representing a sector are chosen based on the wind rose observed at 2 m across the 1-yr period; these classes allow the discrimination of all valley winds whereas the PC2 metric is mostly adapted to assess the representation of the downvalley winds. The performance of the wind speed nowcast is assessed using the correlation coefficient, bias, and mean absolute

error (MAE) computed between observed values and ANN output.

All these metrics were calculated on the test dataset. To assess their sensitivity to the test dataset, a bootstrap method was applied using 10 000 subsamples taken from the test set. Each subsample was created using a random sampling with a replacement from the original test set. The function produced by the ANN was then applied to each subsample, producing 10 000 output datasets. The performance metrics were calculated for each output dataset to build a distribution for each metric. This distribution was used to calculate the mean value for each metric with a confidence interval (at 90%) defined as the interval between the 5th and 95th percentiles.

c. Choice of input variables

Dreyfus et al. (2002) have developed a method for selecting the most pertinent variables to use as input of an ANN. This is performed in two phases:

- 1) The correlation coefficients between each possible pair of variables are calculated, and when the values are high (absolute value higher than 0.9), the two corresponding parameters are considered redundant and one of them is removed under physical and statistical considerations.
- 2) In the second selection stage, the input dataset is constructed by iteration (stepwise forward regression; Dreyfus et al. 2002). A reference ANN is launched with all variables still available as inputs after the first stage (which we call ANN_{ref}). For the first iteration, “elementary” ANNs are defined, each one using a single input variable taken from among the input variables of ANN_{ref}. The performance metrics of these elementary ANNs are compared to each other so that the importance of each input variable can be determined. The input variable set is thus progressively constructed by adding one more input variable to the selected variables at each iteration. The performance of the ANNs is compared to the reference one at the end of each iteration. This process is continued until improvements in the performance metrics are no longer significant.

4. Results and discussion

We first assess the ability of the ANN method to nowcast winds in the Cadarache valley by using the 3-month KASCADE dataset. The performance is compared to the temperature threshold method of D16, and to a simple multilinear regression. The improved nowcasting associated with the ANN method compared to

TABLE 2. Summary of the D16, ANN, and MLR performance metrics. Mean absolute error (MAE): PC2 for the two D16 wind direction classes and PC4 for the four classes (see Fig. 2b). ANN_K and ANN_{K-LW} show the results for the 3-month KASCADE period with and without periods with wind speeds $<0.5 \text{ m s}^{-1}$, respectively, ANN_{K-DC} shows the results for the same period using only two variables relative to the diurnal cycle (hours relative to the closest sunset and sunrise), and MLR_K shows the results for the same 3-month KASCADE period for the multilinear regression. For each ANN, the mean value of the metrics is on the upper line and the 90% confidence interval is on the lower line.

Case	Direction indicators			Speed	Correlation		Speed (m s^{-1})	
	DACC	PC2	PC4		u	v	Bias	MAE
D16	—	0.78	—	—	—	—	—	—
ANN _K	0.84	0.89	0.84	0.93	0.95	0.93	-0.28	0.54
	[0.81, 0.87]	[0.86, 0.92]	[0.81, 0.87]	[0.92, 0.95]	[0.94, 0.97]	[0.92, 0.95]	[-0.33, -0.23]	[0.50, 0.57]
ANN _{K-LW}	0.92	0.95	0.86	0.94	0.97	0.95	-0.22	0.52
	[0.89, 0.94]	[0.93, 0.97]	[0.83, 0.89]	[0.92, 0.95]	[0.95, 0.98]	[0.93, 0.96]	[-0.27, -0.16]	[0.48, 0.56]
ANN _{K-DC}	0.55	0.64	0.57	0.49	0.64	0.53	-1.02	1.34
	[0.52, 0.59]	[0.60, 0.68]	[0.53, 0.61]	[0.42, 0.56]	[0.58, 0.69]	[0.47, 0.58]	[-1.14, -0.91]	[1.24, 1.44]
MLR _K	0.77	0.84	0.75	0.90	0.92	0.91	-0.26	0.66

the temperature threshold method is then assessed and discussed.

a. Selection of variables

Routine observations from the permanent 110-m-tower weather station (GBA in Table 1) are used as inputs for the ANN. They include observations at 110 m (the temperature $T_{110\text{m}}$ and wind), and at 2 m (the atmospheric pressure $P_{2\text{m}}$, temperature $T_{2\text{m}}$, and relative humidity $\text{Rh}_{2\text{m}}$). Moreover, the specific humidity calculated at 2 m $Q_{2\text{m}}$ and the potential temperature difference between 110 and 2 m $\Delta\theta$ are also used as additional input variables. The wind speed $U_{110\text{m}}$, as well as the northern and eastern components, $u_{110\text{m}}$ and $v_{110\text{m}}$, respectively, are retained. It is preferable to use the components as an input variable rather than the wind direction because the cyclic nature of the latter is difficult to handle with the ANN. Altogether, they constitute nine variables which are tested during the first step of selection ($\Delta\theta$, $u_{110\text{m}}$, $v_{110\text{m}}$, $U_{110\text{m}}$, $P_{2\text{m}}$, $Q_{2\text{m}}$, $T_{110\text{m}}$, $T_{2\text{m}}$, and $\text{Rh}_{2\text{m}}$). Three groups of three variables are potentially redundant ($\Delta\theta$, $T_{110\text{m}}$, and $T_{2\text{m}}$ for temperature observations; $u_{110\text{m}}$, $v_{110\text{m}}$, and $U_{110\text{m}}$ for wind observations; and $Q_{2\text{m}}$, $T_{2\text{m}}$, and $\text{Rh}_{2\text{m}}$ for 2-m observations). This redundancy is removed through the selection step which defines the most relevant variables as those that yield the best performance of the ANN. The most pertinent variables are defined as being those which allow the ANN to reach the best results.

Concerning temperatures, a high correlation coefficient (0.92) was found between $T_{2\text{m}}$ and $T_{110\text{m}}$. The correlation coefficients between $T_{2\text{m}}$ and $\Delta\theta$ (-0.66) and between $T_{110\text{m}}$ and $\Delta\theta$ (-0.32) clearly indicate the need to retain $T_{110\text{m}}$ and $\Delta\theta$ rather than $T_{2\text{m}}$ and $\Delta\theta$; thus $T_{2\text{m}}$ is removed. This reflects that the variation of $\Delta\theta$ in the diurnal cycle is controlled by that of $T_{2\text{m}}$ rather than that of $T_{110\text{m}}$. The

negative correlations are explained by the fact that $\Delta\theta$, which reflects the stratification close to the surface, is negative for daytime, convective conditions and positive for nighttime stable stratifications, whereas the air temperature (either at 2 m or at 100 m) is higher during the day than during the night. Concerning the other variables, they are all retained because their correlations are weaker (the maxima obtained are 0.70 between $Q_{2\text{m}}$ and $T_{110\text{m}}$, and 0.68 between $Q_{2\text{m}}$ and $T_{2\text{m}}$). Finally, eight variables remained after the first step of selection: $\Delta\theta$, $u_{110\text{m}}$, $v_{110\text{m}}$, $U_{110\text{m}}$, $P_{2\text{m}}$, $Q_{2\text{m}}$, $T_{110\text{m}}$, and $\text{Rh}_{2\text{m}}$.

The second phase of the input variable selection process starts with the construction of a reference ANN (ANN_{ref}), in which the remaining eight variables are used as input. The final input dataset is then constructed following the procedure described in section 3c. This process produced three input variables ($\Delta\theta$, $u_{110\text{m}}$, and $v_{110\text{m}}$) for the ANN, which resulted in 10-m wind nowcast without a significant loss of performance. This optimized ANN, which uses observations from the KASCADE dataset, is called ANN_K and its performance is presented in Table 2.

b. Results and comparison with other statistical methods

To assess the ability of the ANN method to nowcast winds in the Cadarache valley, we compared its performance to two other statistical methods: the D16 method and a multilinear regression (MLR).

Here, we will briefly review the D16 nowcasting method as a baseline for comparison to the proposed ANN method. Based on routine measurements from the 110-m GBA tower, D16 found a criterion for the diagnosis of the existence of a thermally driven downvalley wind in the Cadarache valley. They developed this method using data from the 3-month KASCADE

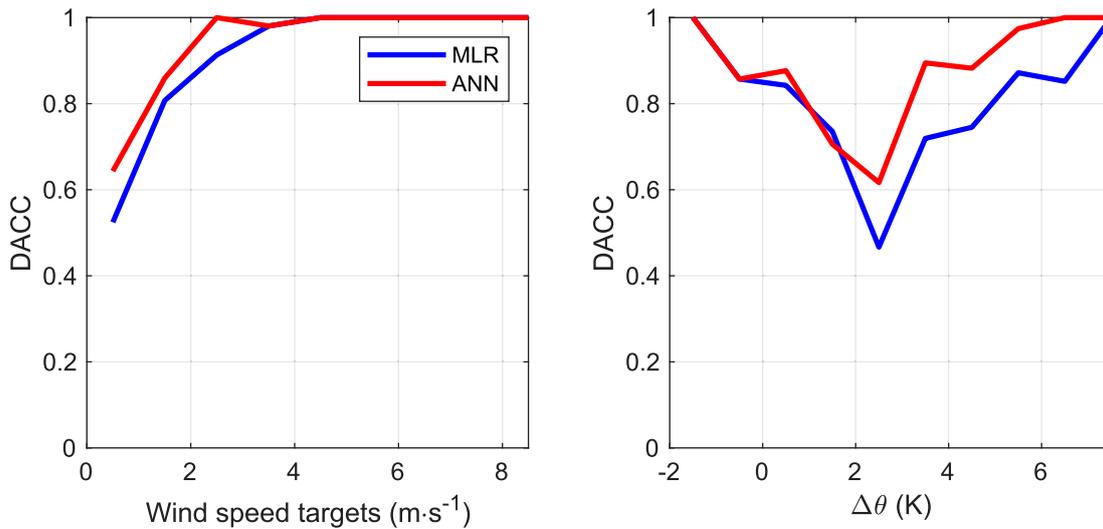


FIG. 5. Comparison of DACC values from the ANN_K (red) and MLR (blue) as a function of (left) the observed wind speed and (right) the observed potential temperature difference between 110 and 2 m. The plots are relative to the 3-month KASCADE period.

period. Under the **D16** criterion, if a potential temperature difference between 110 and 2 m exceeds 2.6 K, then a downvalley flow exists. The technique performs well, with 90% of accurate predictions (PC2 of 0.90, see details in [section 3b](#)). The method has been developed to highlight thermally driven, southeasterly winds in the Cadarache valley. These winds are not only thermally driven, however, as they can also result from a synoptic forcing [[section 2b\(2\)](#)]. To avoid the confusion between these two mechanisms, **D16** removed those winds which were synoptically forced. If all winds are included, the overall performance is somewhat reduced, with a PC2 of 0.78. The threshold for the temperature difference was also slightly modified, with a value of 2.5 K instead of 2.6 K.

The MLR is among the simplest methods to rely input and output parameters, hence it is natural to wonder what benefits arise due to ANNs with respect to MLR. To be comparable with the ANN, the MLR was operated with the same variables ($\Delta\theta$, u_{110m} , and v_{110m}) and the fitting equation was computed on the training and validation sets of the ANN. The MLR equations obtained for the u and v components are

$$u_{\text{MLR}} = 0.296 - 0.279\Delta\theta + 0.278u_{110m} - 0.275v_{110m}, \quad (1)$$

$$v_{\text{MLR}} = 0.057 + 0.166\Delta\theta - 0.095u_{110m} + 0.297v_{110m}. \quad (2)$$

The **D16** method only discriminates wind directions such as CDV or non-CDV according to the value of the vertical potential temperature difference on the 110-m GBA mast. As mentioned above, the ANN method goes beyond the **D16** method's ability to determine the

existence of downvalley winds and computes the two wind components. The performance of the three methods for the KASCADE dataset are summarized in [Table 2](#). The 90% confidence intervals on the ANN metrics correspond to the bootstrap results as explained in [section 3b](#).

The ANN_K achieves a PC2 of 0.89, with a confidence interval between 0.86 and 0.92. This is significantly better than that achieved by **D16** and by the MLR. The comparison of the ANN and MLR methods with the other metrics demonstrates the superiority of the ANN compared with the MLR. To determine for which conditions the differences between the two methods are the highest, we analyzed how their respective performances evolve when the wind or the stratification varies. This is presented in [Fig. 5](#) for the DACC metrics. The differences are higher for light winds ($<3 \text{ m s}^{-1}$) and very stable conditions ($\Delta\theta > 2 \text{ K}$). For high winds, and close-to-neutral or unstable stratifications, the two methods achieve comparable performances, which can be related to a higher coupling between the winds above and at the bottom of the valley, due to forced channeling (high winds) or turbulent mixing (unstable conditions).

The final comparison between observed and ANN nowcasted wind direction and speed is presented in [Fig. 6](#). The ANN reproduces the wind-direction pattern well, with two dominant winds (southeasterly and northwesterly winds), as well as the wind speed. These performances are quantified using several metrics indicated by ANN_K in [Table 2](#). The direction-related metrics all have success rates higher than 80% (PC2 = 0.89; PC4 = 0.84; DAAC = 0.84). The correlation coefficients for the two wind components as well as wind speed reach values

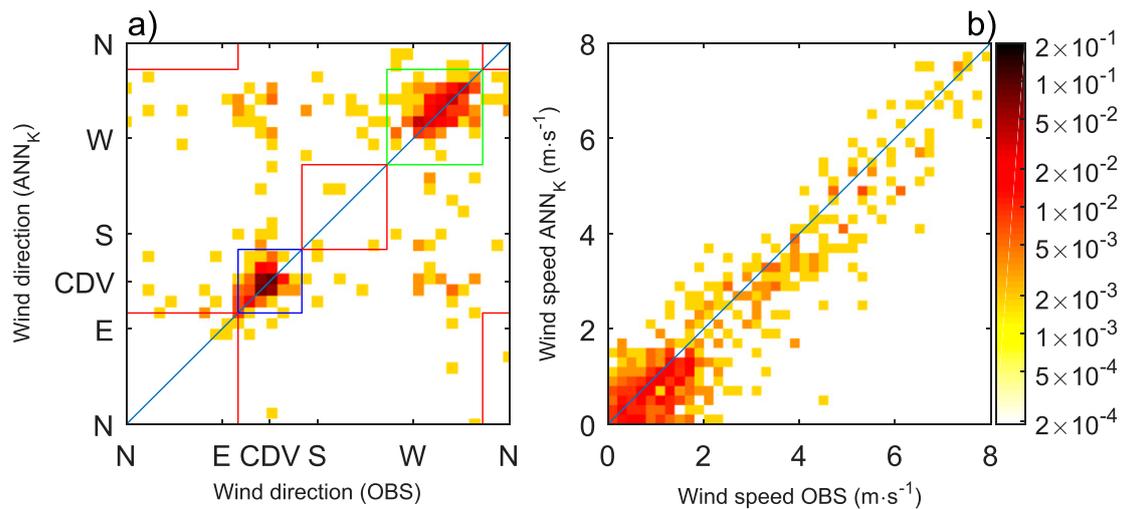


FIG. 6. Comparison of the 10-m (a) wind direction and (b) wind speed observed and nowcasted, represented by a probability density (color scale). Densities are calculated for wind direction bins of 10° and wind speed bins of 0.2 m s^{-1} . The red, green, and blue rectangles in (a) correspond to the three wind direction categories shown in Fig. 3.

between 0.93 and 0.95, and the bias on wind speed is as low as -0.28 m s^{-1} for a mean speed of 2.11 m s^{-1} . For these reasons, we conclude that a method using ANNs yields very good performance and allows a more reliable operational nowcasting than D16's $\Delta\theta$ criterion or a MLR technique.

c. Relative effect of ANN input parameters

Using only three variables ($u_{110\text{m}}$, $v_{110\text{m}}$, and $\Delta\theta$) as ANN inputs, the ANN is capable of calculating the 10-m wind speed and direction at one point in the Cadarache valley. One may, however, observe that the impact of each variable on the ANN result differs.

The ANN calculates the two wind components at 10 m, requiring that the wind components at 110 m be used. The wind at these levels has different patterns. At 10 m, it is strongly impacted by the local topography in the Cadarache valley, while the wind at 110 m reflects the influence of a larger area. Despite these differences, the correlation coefficients for the wind components at 10 and 110 m reach 0.80 and 0.81 for the u and v components, respectively, which reveals the importance of these variables for the calculation.

Duine et al. (2017) have revealed the high occurrence of CDV winds and they have demonstrated that they are induced by topography under stable conditions. The use of $\Delta\theta$, which is related to atmospheric stability, enables the ANN to improve the nowcasting of the CDV winds. As shown by D16, $\Delta\theta$ is strongly connected to the thermally driven downvalley wind. Increasing stable stratification is closely linked to the frequency of occurrence of downvalley winds as shown in Fig. 7b.

The selection of variables shows that the 10-m wind can be reproduced by taking into account the above-valley winds (by using the wind components at 110 m) and the generation of downvalley winds under stable conditions (using the $\Delta\theta$). This is consistent with the results shown in section 2b(2).

d. Effect of the diurnal cycle

In the boundary layer, most meteorological variables have a strong diurnal pattern. We should therefore ask whether a large part of the ANN output was a simple reproduction of this cycle. To highlight information carried from the diurnal cycle to the wind prediction, we made an additional ANN (called ANN_{K-DC}) with only two variables as input, namely, the times relative to the closest sunrise and sunset. The scores of this ANN are presented in Table 2 (row “ANN_{K-DC}”). The information contained in the diurnal cycle is considerable. It is far from being sufficient however to provide a satisfactory performance (i.e., at a level comparable to the performance obtained with the three meteorological parameters used as input for the ANN). The metrics are much lower than those of the ANN_K showing that the diurnal cycle does impact the ANN calculation but does not carry as much information as the three variables used by the ANN_K.

e. Limitation for light winds

Although the ANN is a good tool for nowcasting of low-level winds in the Cadarache valley, there are, however, a few observations for which the wind directions are poorly reproduced by the ANN_K (see Fig. 6a the dots outside the colored boxes). Given the very high correlation between

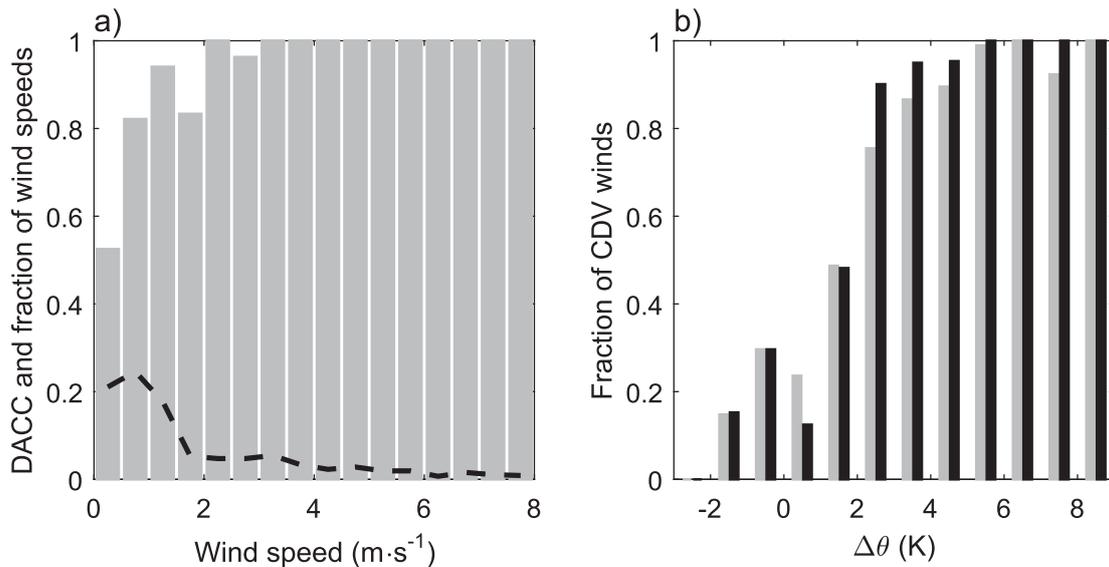


FIG. 7. (a) DACC values as a function of the ANN-calculated wind speed (gray bars) and the proportion of data in each wind speed bin (black dashed line). (b) Proportion of observed (gray) and ANN_K nowcasted (black) 10-m CDV winds as a function of the observed potential temperature difference between 110 and 2 m. The plots are relative to the 3-month KASCADE period.

observed and ANN_K-computed wind components (0.95 and 0.93 for u and v , respectively), these defects can be attributed to very weak winds for which the directions are highly sensitive to small variations of the components values. This is due to the indirect calculation of the wind direction (from wind components), which is impossible to compute as a cyclic function through the ANN. A small error made in one component for lower winds produces a larger error in the wind direction than for stronger winds. Wind directions for winds with a speed lower than 0.5 m s^{-1} were found to be badly reproduced by the ANN, with a DACC of 0.53 (Fig. 7a). The ANN_K performance indices were thus recalculated after wind speeds (from the ANN outputs) lower than 0.5 m s^{-1} were removed (note that the removal of light winds beforehand a learning step on the basis of target winds is impossible without introducing a bias). The corresponding row in Table 2 (ANN_{K-LW}) clearly shows the improvement of direction metrics DACC and PC2 above those of ANN_K. PC4 is only slightly improved because one of the four classes is dedicated to light winds, independent of their directions. Neither the correlation coefficients, nor the mean absolute error of the wind speed are significantly modified. The negative bias on the wind speed is reduced, showing that light winds were underestimated by ANN_K.

f. Performance of the ANN over a 1-yr period

To encompass a wider range of meteorological conditions, the ANN method is assessed over a 1-yr period. As indicated above, this period lasts from 17 February 2015

to 17 February 2016. To check if the synoptic meteorological conditions were not unusual during this period, and therefore if our results could be regarded as valid whatever the year considered, we checked the year-to-year variability of the wind observed at 110 m, which is representative of a scale larger than that of the valley. For the period between early 2010 and the end of 2017 (corresponding to the dataset at our disposal), the eight yearly wind roses appear to be very similar, and resemble that presented in Fig. 3d, with the same three main wind directions. The only year-to-year difference lies in the occurrence of the strongest SE winds, associated to bad weather episodes.

This new ANN, called ANN_{1yr}, has exactly the same structure as ANN_K. The selection step revealed that the same three variables are the most pertinent ones ($\Delta\theta$, $u_{110\text{m}}$, and $v_{110\text{m}}$). Its performances are presented in Table 3 (rows “ANN_{1yr}” and “ANN_{1yr-LW}”). These performances are slightly worse than those computed over the 3-month period (ANN_K and ANN_{K-LW}), but they are not noticeably different. The visual comparison of wind speed and direction (Fig. 8) highlights the same limitations as those with the KASCADE dataset. Specifically, the representation of the light winds and cross-valley winds remains a challenge.

g. Length of the dataset

In the previous section, we demonstrated the efficiency of the ANN method when applied to a dataset involving a larger diversity of input conditions. The increase of size from the 3-month to the 1-yr period

TABLE 3. As in Table 2, but for the 1-yr period. ANN_{1yr} and ANN_{1yr-LW} show the results with and without wind speeds lower than 0.5 m s⁻¹, respectively.

Case	Direction indicators			Speed	Correlation		Speed (m s ⁻¹)	
	DACC	PC2	PC4		<i>u</i>	<i>v</i>	Bias	MAE
ANN _{1yr}	0.81	0.88	0.83	0.92	0.94	0.86	-0.24	0.52
	[0.77, 0.84]	[0.85, 0.91]	[0.80, 0.86]	[0.90, 0.94]	[0.93, 0.96]	[0.83, 0.90]	[-0.29, -0.19]	[0.48, 0.55]
ANN _{1yr-LW}	0.90	0.96	0.83	0.91	0.95	0.88	-0.18	0.55
	[0.87, 0.93]	[0.93, 0.98]	[0.80, 0.87]	[0.88, 0.93]	[0.93, 0.96]	[0.85, 0.91]	[-0.25, -0.12]	[0.51, 0.60]

does not however improve the performance. Therefore, we assessed the impact of the length of the dataset on the ANN performance, with shorter datasets. These tests were conducted on the 3-month KASCADE dataset.

A total of 550 subsamples of the input dataset were created, ranging from 100% (1899 samples) to 5% (95 samples) of the entire dataset. For each reduced size, 50 random subsamples were created (the random

draw guarantees selected data from all seasons and hours, avoiding seasonal or diurnal cycle bias). This allowed us to investigate the effect of a sample size, particularly for the smallest datasets for which two different random samples can generate datasets without any common values. For each size, the overall performance of the ANN is characterized by the median value and the 5th and 95th percentiles computed over the 50 index values.

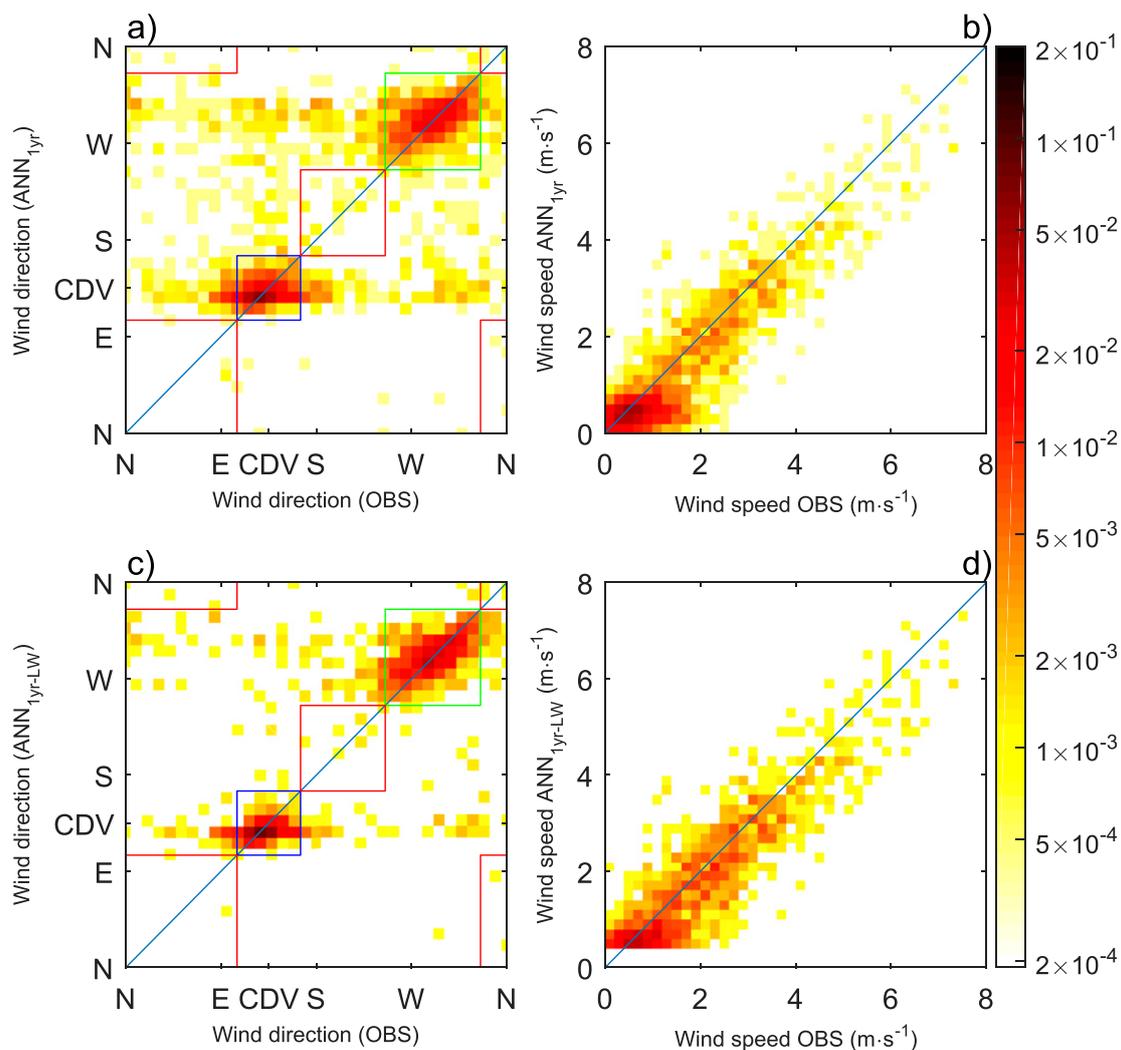


FIG. 8. As in Fig. 6, but for the (a),(b) ANN_{1yr} and (c),(d) ANN_{1yr-LW} outputs.

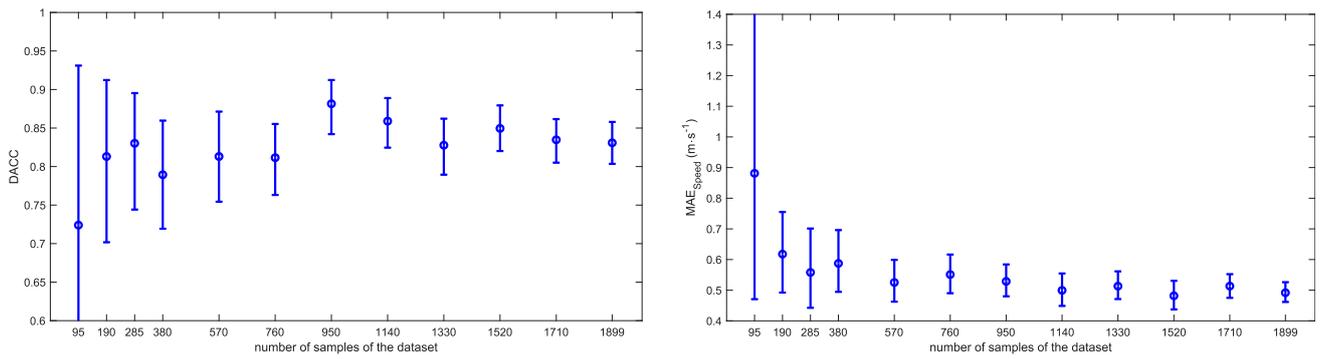


FIG. 9. Evolution of (a) DACC and (b) MAE calculated on the wind speed as a function of the size of the input dataset.

Figure 9 shows the evolution of DACC and mean absolute error for the wind speed as a function of the length of the dataset. For the longest datasets, there is no clear tendency in the evolution of performance with sample size. However, for the smallest sample sizes, the performance increases with the dataset size. Moreover, the dispersion (confidence interval), which is relatively constant beyond the 950 samples dataset, increases for the smaller sample sizes. Hence, the constant behavior of the performance metrics for datasets composed of 950–1899 samples indicates that the length of the KASCADE dataset can be considered as “long enough” to allow the ANN to sufficiently characterize the Cadarache wind patterns. The previous part demonstrated that an ANN trained with a longer dataset (e.g., the 1-yr period) neither improves the performance nor diminishes the confidence interval. From these results, we conclude that the KASCADE as well as the

1-yr-period datasets are long enough to ensure that the ANN performance is not impacted by the length of the dataset.

h. Excessive nowcast of channeling

Another limitation of the ANN is shown in Fig. 10, which exhibits the horizontal wind components’ occurrence frequency. ANN_K wind components (Fig. 10b) seem to be clustered in a small region of the graph, contained between the magenta lines (these have been simply hand drawn to show a visual appreciation of the larger spread of the observed data relative to the ANN data), indicating less spread wind directions than in the observations (Fig. 10a). For example, the observed east-southeasterly-to-northwesterly winds are rarely represented by the ANN (Table 4). This results in an excessive nowcast of channeled winds and a lack of cross-valley winds. We hypothesize two explanations for this behavior.

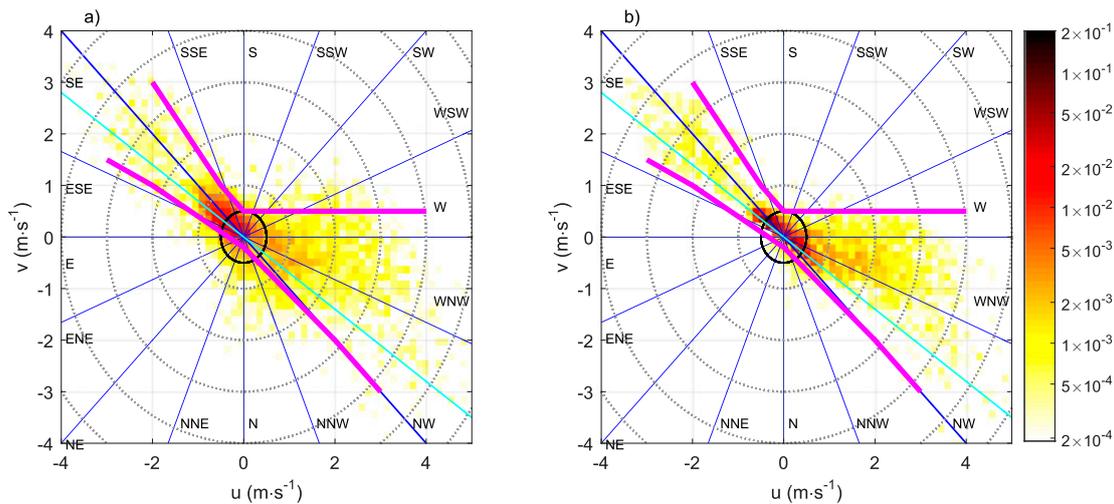


FIG. 10. Probability density of wind components at 2 m: (a) observed and (b) calculated by ANN_{1yr}. The blue lines and the corresponding labels indicate the wind directions. The cyan line shows the direction of the valley axis. The dotted circles represent the wind speeds at 1 m s⁻¹ intervals, and the solid black circle represents 0.5 m s⁻¹. The magenta lines indicate an “envelope” that encompasses most of the ANN winds.

TABLE 4. As in Table 2, but for the performance of ANN_K when considering only observed CDV winds [ANN_{K(CDV)}], only upvalley winds [ANN_{K(CUV)}], or only cross-valley winds [ANN_{K(XValley)}].

Case	Direction indicators			Correlation			Speed (m s ⁻¹)	
	DACC	PC2	PC4	Speed	<i>u</i>	<i>v</i>	Bias	MAE
ANN _{K(CDV)}	0.90 [0.87, 0.94]	0.91 [0.88, 0.94]	0.86 [0.82, 0.90]	0.95 [0.93, 0.97]	0.94 [0.92, 0.96]	0.92 [0.89, 0.94]	-0.17 [-0.23, -0.11]	0.43 [0.39, 0.47]
ANN _{K(CUV)}	0.86 [0.82, 0.90]	0.92 [0.88, 0.95]	0.88 [0.84, 0.91]	0.93 [0.90, 0.95]	0.90 [0.87, 0.92]	0.90 [0.87, 0.93]	-0.46 [-0.54, -0.38]	0.66 [0.59, 0.72]
ANN _{K(XValley)}	0.38 [0.27, 0.50]	0.65 [0.54, 0.76]	0.52 [0.40, 0.64]	0.77 [0.44, 0.92]	0.86 [0.60, 0.95]	0.31 [0.01, 0.57]	-0.01 [-0.16, 0.15]	0.48 [0.37, 0.61]

First, cross-valley winds are always light, and it was shown above that these are the most challenging types of winds for the ANN to predict. Second, contrary to up- and downvalley winds [section 2b(2)] there are no typical meteorological conditions associated with these winds, which therefore cannot be expressed as a mathematical relation between ANN input parameters.

This misrepresentation of cross-valley winds is visible when comparing the performances calculated on the three wind direction sectors (Table 4). The efficiency of representing cross-valley winds is much lower than for up- or downvalley winds. Indeed, performances of up- and downvalley winds are higher and very close, apart from for wind speed bias and mean absolute error, explained by stronger wind speeds for upvalley winds than for downvalley winds.

Apart from the narrowness of the two main wind-direction lobes in the ANN winds, the plots also reveal a nonperfect alignment between the two opposite lobes: downvalley winds are strictly parallel to the main axis of the valley, whereas upvalley directions cluster around 280°, that is, about a 25° counterclockwise from the valley axis. This behavior, mainly observed during the day, could result from the enlargement of the valley when approaching its mouth (see Fig. 1b), and/or from downward momentum transport: above-valley winds, when they originate from the northwest quadrant, span a rather wide sector (west to northwest).

5. Conclusions

The goal of this study was to develop a tool for nowcasting winds in a small valley embedded in a complex topography using operational meteorological observations. In these types of situations, small-scale complex topography can give rise to local circulations, which are neither observed by routine measurements nor forecasted by operational meteorological models. The targeted area is the Cadarache valley, which is 1 km wide, 100 m deep, and 6 km long. At a larger scale, the wind features two

main directions close to the direction of the valley axis, which is oriented 125°–305°. When it is strong enough, the large-scale wind is able to force the direction of low-level winds, while a decorrelation between low-level winds and above-valley winds has been observed during stable conditions, producing thermally driven winds with directions forced by the topography (explaining the alignment with the valley axis). On the other hand, few observations feature light cross-valley winds which cannot be associated with any typical meteorological conditions of the region.

The ANN method was used to improve the nowcasting method of D16, which aims to assess the occurrence of thermally driven winds in the Cadarache valley, by providing both the wind speed and direction predictions. A multilayer perceptron type ANN was used to calculate the wind within the Cadarache valley using continuous observations made above the valley depth. First, the ability of the ANN was assessed by comparing its performance to the nowcasting method of D16 and to a simple MLR. The dataset used gathers observations from the 3-month-long KASCADE experiment. The ANN method produces better results than D16 and MLR methods. Furthermore, the main improvement brought by the ANN is the calculation of the wind components which allow the calculation of actual wind speeds and directions, clearly indispensable for pollutant dispersion monitoring. The method performs well: the correlation coefficients calculated for the wind components (*u* and *v*) and wind speed reach 0.95, 0.93, and 0.93, respectively. The wind speed bias is -0.28 m s⁻¹ with a mean absolute error of 0.54 m s⁻¹. Furthermore, 84% of the wind directions are correctly nowcasted within ±45°.

This ANN uses only three variables ($\Delta\theta$, u_{110m} , and v_{110m}) from the routine observations made at a height greater than the valley depth, and therefore not representative of the local flow. Despite the differences between the wind patterns at 110 and 10 m, the wind components at 110 m are indispensable for the calculation of the winds at 10 m by the ANN. This is explained

by a high correlation of the components between the two heights (coefficients of correlation of 0.80 and 0.81 for the u and v components, respectively). The difference of temperature between 110 and 2 m is also fundamental for the calculation of the downvalley winds under thermally stratified conditions.

However, some challenges remain, especially in the wind direction. They are explained by the method used to calculate the wind direction, which is not a direct output of the ANN. Indeed, only the wind components are calculated by the ANN, so as to avoid problems with a cyclic variable such as the direction. This introduces a limitation for the calculation of the light winds direction because a small error in the components can lead to a large error in the direction. Therefore, the ANN ability to predict wind direction improves with increasing wind speed. The other limitation encountered with the ANN method is the excessive nowcast of channeled winds, which is explained by the fact that most wind directions away from the channeled corridor correspond to very light winds whose directions are challenging for the ANN to reproduce. There is no typical meteorological situation associated with these winds. The utilization of a longer dataset, over a 1-yr period, did not fix this problem. Conversely and surprisingly, studying the impact of shorter datasets on the ANN performance proved that the length of both the KASCADE period and 1-yr-period datasets were long enough to validate the ANN calculation. Furthermore, the ANN method applied to the 1-yr period produces results that are comparable to those of the KASCADE period.

From these results, we conclude that the ANN is a suitable tool for nowcasting and downscaling winds in the Cadarache valley using routine observations. Although ANNs have already been used to nowcast meteorological variables, to the knowledge of the authors, most of these studies focused on the possibility of representing satisfactorily the wind speed but missing the wind direction. This study shows that the utilization of ANNs is a promising technique for the nowcasting of a challenging variable such as the local winds in complex topography areas based on remote measurements. We plan to apply this method to numerical mesoscale modeling outputs in order to downscale local winds in an unresolved valley.

Acknowledgments. The KASCADE field experiment was made possible by the contributions of several institutions and sources of support: CEA, Sud-PACA region, and Fulbright Scholarship Program are thanked for the financial support; Laboratoire d'Aérodynamique (LA, University of Toulouse) is thanked for the provision of

instrumentation. LA and LMTE (CEA, Cadarache) are acknowledged for their participation in the field campaign, which was hosted by LMTE. The SEDOO, Observatoire Midi-Pyrénées, is acknowledged for its help in database management. We thank the reviewers for their constructive feedback on the manuscript.

REFERENCES

- Bastin, S., P. Drobinski, A. Dabas, P. Delville, O. Reitebuch, and C. Werner, 2005: Impact of the Rhône and Durance valleys on sea-breeze circulation in the Marseille area. *Atmos. Res.*, **74**, 303–328, <https://doi.org/10.1016/j.atmosres.2004.04.014>.
- Beale, M., M. T. Hagan, and H. B. Demuth, 2010: Neural network toolbox 7: User's guide. MathWorks, 951 pp.
- Berthou, S., and Coauthors, 2016: Influence of submonthly air–sea coupling on heavy precipitation events in the western Mediterranean basin. *Quart. J. Roy. Meteor. Soc.*, **142**, 453–471, <https://doi.org/10.1002/qj.2717>.
- Burns, P., and C. Chemel, 2015: Interactions between downslope flows and a developing cold-air pool. *Bound.-Layer Meteor.*, **154**, 57–80, <https://doi.org/10.1007/s10546-014-9958-7>.
- Cadenas, E., and W. Rivera, 2009: Short term wind speed forecasting in la Venta, Oaxaca, México, using artificial neural networks. *Renewable Energy*, **34**, 274–278, <https://doi.org/10.1016/j.renene.2008.03.014>.
- Chow, F. K., S. F. De Wekker, and B. J. Snyder, 2012: *Mountain Weather Research and Forecasting: Recent Progress and Current Challenges*. Springer Science & Business Media, 750 pp.
- Clements, C. B., C. D. Whiteman, and J. D. Horel, 2003: Cold-air-pool structure and evolution in a mountain basin: Peter Sinks, Utah. *J. Appl. Meteor.*, **42**, 752–768, [https://doi.org/10.1175/1520-0450\(2003\)042<0752:CSAEIA>2.0.CO;2](https://doi.org/10.1175/1520-0450(2003)042<0752:CSAEIA>2.0.CO;2).
- Clements, W. E., J. A. Archuleta, and P. H. Gudiksen, 1989: Experimental design of the 1984 ascot field study. *J. Appl. Meteor.*, **28**, 405–413, [https://doi.org/10.1175/1520-0450\(1989\)028<0405:EDOTAF>2.0.CO;2](https://doi.org/10.1175/1520-0450(1989)028<0405:EDOTAF>2.0.CO;2).
- Delon, C., D. Serça, C. Boissard, R. Dupont, A. Dutot, P. Laville, P. De Rosnay, and R. Delmas, 2007: Soil no emissions modelling using artificial neural network. *Tellus*, **59B**, 502–513, <https://doi.org/10.1111/j.1600-0889.2007.00254.x>.
- Doran, J., J. D. Fast, and J. Horel, 2002: The VTMX 2000 campaign. *Bull. Amer. Meteor. Soc.*, **83**, 537–551, [https://doi.org/10.1175/1520-0477\(2002\)083<0537:TVC>2.3.CO;2](https://doi.org/10.1175/1520-0477(2002)083<0537:TVC>2.3.CO;2).
- Dreyfus, G., J. Martinez, M. Samuelides, M. B. Gordon, F. Badran, S. Thiria, and L. Herault, 2002: *Réseaux de Neurones: Méthodologie et Applications*. Eyrolles, 408 pp.
- Duine, G.-J., T. Hedde, P. Roubin, and P. Durand, 2016: A simple method based on routine observations to nowcast down-valley flows in shallow, narrow valleys. *J. Appl. Meteor. Climatol.*, **55**, 1497–1511, <https://doi.org/10.1175/JAMC-D-15-0274.1>.
- , —, —, —, M. Lothon, F. Lohou, P. Augustin, and M. Fourmentin, 2017: Characterization of valley flows within two confluent valleys under stable conditions: Observations from the KASCADE field experiment. *Quart. J. Roy. Meteor. Soc.*, **143**, 1886–1902, <https://doi.org/10.1002/qj.3049>.
- Fernando, H., and Coauthors, 2015: The Materhorn: Unraveling the intricacies of mountain weather. *Bull. Amer. Meteor. Soc.*, **96**, 1945–1967, <https://doi.org/10.1175/BAMS-D-13-00131.1>.
- Gardner, M. W., and S. Dorling, 1998: Artificial neural networks (the multilayer perceptron)—A review of applications in the

- atmospheric sciences. *Atmos. Environ.*, **32**, 2627–2636, [https://doi.org/10.1016/S1352-2310\(97\)00447-0](https://doi.org/10.1016/S1352-2310(97)00447-0).
- Khosravi, A., R. Koury, L. Machado, and J. Pabon, 2018: Prediction of wind speed and wind direction using artificial neural network, support vector regression and adaptive neuro-fuzzy inference system. *Sustainable Energy Technol. Assess.*, **25**, 146–160, <https://doi.org/10.1016/j.seta.2018.01.001>.
- Lareau, N. P., E. Crosman, C. D. Whiteman, J. D. Horel, S. W. Hoch, W. O. Brown, and T. W. Horst, 2013: The persistent cold-air pool study. *Bull. Amer. Meteor. Soc.*, **94**, 51–63, <https://doi.org/10.1175/BAMS-D-11-00255.1>.
- More, A., and M. Deo, 2003: Forecasting wind with neural networks. *Mar. Structures*, **16**, 35–49, [https://doi.org/10.1016/S0951-8339\(02\)00053-9](https://doi.org/10.1016/S0951-8339(02)00053-9).
- Muñoz, R. C., M. J. Falvey, M. Araya, and M. Jacques-Coper, 2013: Strong down-valley low-level jets over the Atacama Desert: Observational characterization. *J. Appl. Meteor. Climatol.*, **52**, 2735–2752, <https://doi.org/10.1175/JAMC-D-13-063.1>.
- Philippopoulos, K., and D. Deligiorgi, 2012: Application of artificial neural networks for the spatial estimation of wind speed in a coastal region with complex topography. *Renewable Energy*, **38**, 75–82, <https://doi.org/10.1016/j.renene.2011.07.007>.
- Price, J., and Coauthors, 2011: COLPEX: Field and numerical studies over a region of small hills. *Bull. Amer. Meteor. Soc.*, **92**, 1636–1650, <https://doi.org/10.1175/2011BAMS3032.1>.
- Rotach, M. W., and Coauthors, 2017: Investigating exchange processes over complex topography: The Innsbruck Box (i-Box). *Bull. Amer. Meteor. Soc.*, **98**, 787–805, <https://doi.org/10.1175/BAMS-D-15-00246.1>.
- Sabatier, T., A. Paci, G. Canut, Y. LARGERON, A. Dabas, J.-M. Donier, and T. Douffet, 2018: Wintertime local wind dynamics from scanning Doppler lidar and air quality in the Arve River valley. *Atmosphere*, **9**, 118, <https://doi.org/10.3390/atmos9040118>.
- Santos-Alamillos, F., D. Pozo-Vázquez, J. Ruiz-Arias, V. Lara-Fanego, and J. Tovar-Pescador, 2013: Analysis of WRF Model wind estimate sensitivity to physics parameterization choice and terrain representation in Andalusia (southern Spain). *J. Appl. Meteor. Climatol.*, **52**, 1592–1609, <https://doi.org/10.1175/JAMC-D-12-0204.1>.
- Simpson, J. E., 1994: *Sea Breeze and Local Winds*. Cambridge University Press, 252 pp.
- Whiteman, C. D., and J. C. Doran, 1993: The relationship between overlying synoptic-scale flows and winds within a valley. *J. Appl. Meteor.*, **32**, 1669–1682, [https://doi.org/10.1175/1520-0450\(1993\)032<1669:TRBOSS>2.0.CO;2](https://doi.org/10.1175/1520-0450(1993)032<1669:TRBOSS>2.0.CO;2).
- , and Coauthors, 2008: METCRAX 2006: Meteorological experiments in Arizona's meteor crater. *Bull. Amer. Meteor. Soc.*, **89**, 1665–1680, <https://doi.org/10.1175/2008BAMS2574.1>.
- Yegnanarayana, B., 2009: *Artificial Neural Networks*. PHI Learning Pvt. Ltd., 476 pp.

Chapitre 5

Descente d'échelle statistique de simulations méso-échelle pour la prévision d'un vent local

Les simulations WRF (de l'outil « Mistrau ») présentées dans le chapitre 2 ont une résolution horizontale de 3 km sur la région de Cadarache. Cette résolution permet une représentation grossière de la vallée de la Durance mais pas de la vallée de Cadarache. La prévision des vents y est alors crédible seulement en cas de forçage élevé par les vents en altitude et en journée lorsque les vents en altitude sont alignés avec la vallée, mais pas en conditions stables lorsque les vents sont très dépendants d'effets locaux concernant des échelles nettement inférieures à 3 km.

De précédentes études ont montré qu'il est possible d'appliquer des post-traitements aux sorties de simulations dans le but de représenter des phénomènes sous-maille (Sheridan et al., 2010; Focken et al., 2002; Lei et al., 2009; Ahn et al., 2012; Liu and Coulibaly, 2011; Kumar et al., 2012; Coulibaly et al., 2005). De plus, dans le chapitre 4, il a été démontré qu'un réseau de neurones artificiels peut être utilisé pour le calcul du vent dans la vallée de Cadarache à partir d'observations représentatives d'une échelle supérieure à celle de la vallée (en l'occurrence représentatives de la vallée de la Durance). Ce chapitre présente donc une tentative de descente d'échelle de prévisions numériques à l'aide d'un réseau de neurones artificiels dans le but d'améliorer la prévision des vents dans la vallée de Cadarache. Il est composé d'une partie visant à présenter la méthodologie appliquée pour la construction du RNA, d'une comparaison des résultats à ceux du chapitre 4 (prévision immédiate), d'une analyse de l'apport de ce post-traitement aux simulations WRF et d'une analyse des erreurs produites par le RNA.

Plan du chapitre

5.1	Construction du réseau de neurones artificiels	132
5.1.1	La stratégie	132
5.1.2	Sélection des variables d'entrée	132
5.2	Performance de prévision du vent avec le RNA	135
5.2.1	Comparaison aux résultats du RNA–Observations	135
5.2.2	Apport du RNA par rapport aux simulations WRF	135
5.2.3	Analyse des erreurs	140
5.3	Conclusion	145

5.1 Construction du réseau de neurones artificiels

5.1.1 La stratégie

Le but du chapitre 4 était d'avoir une représentation instantanée du vent dans la vallée de Cadarache ; l'objectif est maintenant d'obtenir une prévision jusqu'à 48 h de ce vent. Pour cela, ce ne sont plus des observations qui sont utilisées comme variables en entrée du RNA mais les sorties des prévisions WRF. En revanche, le reste de la structure générale du réseau est identique à celle présentée dans le chapitre 4 pour le réseau utilisant des observations comme variables d'entrée :

- les observations utilisées comme cible pour l'apprentissage du RNA sont les observations du vent à 2 m sur la station MET01 sur la même période du 17 février 2015 au 17 février 2016 ;
- les deux sorties du RNA ont pour objectif de représenter les composantes u et v du vent à 2 m ;
- le nombre de neurones de la couche cachée est de 10 (chiffre obtenu après avoir testé des RNA avec un nombre de neurones cachés allant de 5 à 50) ;
- la fonction d'activation est la fonction tangente hyperbolique ;
- le jeu de données est divisé en trois sous-ensembles : 60 % des données pour le jeu d'apprentissage, 10 % pour le jeu de validation et les 30 % restants pour le jeu de test.

Les variables utilisées sont extraites des simulations WRF suivant la méthode décrite dans la partie 2.3.1. Seules les échéances horaires de +24 h à +47 h des prévisions quotidiennes sont utilisées afin de ne prendre en compte ni la période de spin-up (temps de stabilisation de la simulation), ni les prévisions plus lointaines qui perdent en précision. L'ensemble constitue donc un jeu de données d'entrée horaire sur une période d'un an. Les données sont ensuite calculées par interpolation bilinéaire au point de coordonnées (43,685°N ; 5,760°E) qui correspond à la position de la station MET01. Enfin, les variables 3D ont été calculées à différents niveaux altimétriques. La méthode appliquée pour le choix des variables d'entrée du RNA est détaillée dans la partie suivante.

5.1.2 Sélection des variables d'entrée

Une première sélection de variables est faite a priori, parmi les nombreuses variables produites par les simulations WRF, sur la base de la connaissance que l'on a de l'aérodynamique locale de la vallée de Cadarache. Notamment, il a été montré dans les chapitres 2 et 3 que ces vents peuvent être couplés ou découplés de l'écoulement au-dessus de la vallée. Les vents couplés ont alors des caractéristiques proches des vents à l'échelle régionale alors que les vents découplés sont générés localement par des effets thermiques. La première sélection de variables consiste donc à retenir des variables qui peuvent être corrélées à la présence d'un type de vent en particulier (tableau 5.1).

Ainsi, la vitesse du vent dans et au-dessus de la vallée de Cadarache peut être une bonne indication du couplage/découplage des vents (voir la partie 3.4.1). Les composantes et la vitesse du vent à trois niveaux ont donc été sélectionnées comme variables d'entrée : à 10 m parce que c'est le niveau le plus bas des sorties WRF, et donc le niveau qui en théorie doit représenter le mieux le vent dans le fond de la vallée ; à 110 m parce que le RNA appliqué aux observations a montré que ce paramètre était déterminant ; à 300 m parce que c'est à cette hauteur que se

trouve le maximum du jet de vent de Durance (Duine et al., 2017) et que ce vent est, comme le vent de vallée de Cadarache, d'origine thermique.

Comme la différence de température potentielle sur la verticale a été une variable importante pour la prévision immédiate du vent (chapitre 4), celle-ci ainsi que d'autres variables caractérisant la stabilité de l'atmosphère ont aussi été sélectionnées dans le but de donner au RNA une information pouvant l'aider à identifier les périodes de découplage pendant lesquelles les vents ont une origine thermique. Ainsi, un nombre de Richardson global (équation 1.21), qui est un indicateur de la stabilité, calculé entre 110 m et 2 m, la vitesse de frottement (équation 1.29) qui est un indicateur des contraintes de surface (et donc de la stabilité de l'écoulement) ainsi que la hauteur de couche limite atmosphérique (définie comme étant le niveau auquel l'énergie cinétique turbulente devient inférieure à $0,01 \text{ m}^2 \text{ s}^{-2}$ (Banks et al., 2016)) sont également testées comme variables d'entrée.

De plus, il a été montré dans la partie 3.4.3 que le vent de vallée peut apparaître et disparaître à des heures variables. Cependant, Duine et al. (2017) ont montré qu'il est fréquent qu'il apparaisse autour du coucher du soleil et qu'il disparaisse environ 2 heures après le lever du soleil. C'est pourquoi les écarts de temps entre l'échéance et le coucher et lever de soleil les plus proches ont aussi été utilisées comme variables d'entrée. Enfin, des variables météorologiques fondamentales telles que la température potentielle, la pression, l'humidité relative, la température de surface et la pluie ont été rajoutées à la liste de variables.

Vent et forçage de grande échelle	<ul style="list-style-type: none"> – Rayonnement net (Rn) – Vitesse (U) et composantes (u et v) du vent à 10 m, 110 m et 300 m
Stabilité de l'atmosphère	<ul style="list-style-type: none"> – Différence verticale de température potentielle entre 110 m et 2 m ($\Delta\theta$) – Hauteur de couche limite ($PBLH$) – Vitesse de frottement (u_*) – Nombre de Richardson global entre 110 m et 2 m (Ri) (comme dans Duine et al. (2016))
Cycle diurne	<ul style="list-style-type: none"> – Écart de temps relatif au coucher/lever de soleil le plus proche ($Sunset$ et $Sunrise$)
Variables météorologiques fondamentales	<ul style="list-style-type: none"> – Pression à 2 m au-dessus du sol ($P_{2\text{m}}$) – Humidité spécifique à 2 m au-dessus du sol ($Q_{2\text{m}}$) – Températures potentielles à 2 m et 110 m au-dessus du sol ($\theta_{2\text{m}}$ et $\theta_{110\text{m}}$) – Température de la surface (T_{surf}) – Pluie

Tableau 5.1 : Revue des variables sélectionnées a priori parmi celles disponibles en sortie de simulation WRF.

La réduction du nombre de variables d'entrée est ensuite effectuée en deux étapes, de la même manière qu'au chapitre précédent, en suivant la méthodologie de Dreyfus et al. (2002). Le détail de ces deux étapes de sélection est présenté dans l'annexe C.

La première étape est basée sur les corrélations entre les différentes variables testées comme

potentielles entrées du réseau. De fortes corrélations ont été trouvées entre plusieurs couples de variables (tableau 5.2). Pour les coefficients de corrélation supérieurs à 0,9, les variables correspondantes sont considérées comme redondantes et l'une d'entre elles peut être retirée de la sélection.

(a)

	$Q_{2\text{m}}$	$\theta_{2\text{m}}$	$\theta_{110\text{m}}$	T_{surf}
$Q_{2\text{m}}$	–	0,75	0,80	0,69
$\theta_{2\text{m}}$	0,75	–	0,97	0,99
$\theta_{110\text{m}}$	0,80	0,97	–	0,94
T_{surf}	0,69	0,99	0,94	–

(b)

	u_*	U_{10}	u_{10}	v_{10}	U_{110}	u_{110}	v_{110}	U_{300}	u_{300}	v_{300}
u_*	–	0,94	0,37	–0,01	0,73	0,34	0,12	0,53	0,30	–0,02
U_{10}	0,94	–	0,39	–0,01	0,87	0,37	0,09	0,67	0,32	–0,08
u_{10}	0,37	0,39	–	–0,44	0,30	0,93	–0,45	0,18	0,83	–0,55
v_{10}	–0,01	–0,01	–0,44	–	–0,01	–0,33	0,83	–0,07	–0,26	0,59
U_{110}	0,73	0,87	0,30	–0,01	–	0,29	0,08	0,83	0,23	–0,10
u_{110}	0,34	0,37	0,93	–0,33	0,29	–	–0,40	0,15	0,93	–0,56
v_{110}	0,12	0,09	–0,45	0,83	0,08	–0,40	–	0,04	–0,35	0,85
U_{300}	0,53	0,67	0,18	–0,07	0,83	0,15	0,04	–	0,04	–0,12
u_{300}	0,30	0,32	0,83	–0,26	0,23	0,93	–0,35	0,04	–	–0,50
v_{300}	–0,02	–0,08	–0,55	0,59	–0,10	–0,56	0,85	–0,12	–0,50	–

Tableau 5.2 : Revue des coefficients de corrélation pour certains couples de variables sélectionnées a priori parmi les différentes variables disponibles en sortie de simulation WRF. Les valeurs supérieures à 0,9 sont indiquées par des cellules rouges et celles supérieures à 0,7 par des cellules jaunes.

Les variables $\theta_{2\text{m}}$, $\theta_{110\text{m}}$ et T_{surf} sont fortement corrélées. $\theta_{2\text{m}}$ étant la variable qui affiche les coefficients de corrélation les plus élevés avec les deux autres variables, elle est la première à être supprimée. Il existe également une corrélation importante avec l'humidité spécifique à 2 m. De plus, la température de surface affiche des corrélations (positives ou négatives) élevées avec d'autres variables (supérieures à 0,55 avec Rn , $PBLH$, $Sunset$, $Sunrise$, $\Delta\theta$ et u_*). Pour cette raison, la température de surface est retirée pour garder la température potentielle à 110 m qui apporte plus de variabilité par rapport aux autres variables testées comme entrées du RNA.

Il existe une forte corrélation entre les vents aux différentes hauteurs, ce qui s'explique par le fait que les vents dans les simulations sont souvent forcés par les vents sus-jacents (figure 2.16). Les corrélations entre les vents à 110 m et ceux aux autres altitudes sont plus fortes que celle entre les vents à 10 m et 300 m, ce qui justifie de retirer les vents à 110 m.

Il y a aussi une forte corrélation entre la vitesse à 10 m et la vitesse de frottement ce qui vient du lien existant entre les deux variables. Ces deux variables ont aussi une corrélation importante avec la hauteur de couche limite, avec un coefficient de corrélation plus important pour la vitesse de frottement, c'est pourquoi cette variable est retirée et la vitesse à 10 m est conservée.

Finalement, il reste 16 variables relativement décorréliées utilisables comme entrées du RNA

après la première étape de sélection : $Q_{2\text{m}}$, $P_{2\text{m}}$, $\theta_{110\text{m}}$, $\Delta\theta$, Rn , Ri , la pluie, la hauteur de couche limite, les heures relatives aux lever et coucher de soleil les plus proches et les composantes u et v ainsi que la vitesse des vents à 10 m et 300 m.

La deuxième étape de sélection consiste à construire le jeu de données d'entrée le plus pertinent : la présentation du processus étant quelque peu fastidieuse, les détails sont fournis dans l'annexe C, ainsi que les résultats obtenus. Parmi les 16 variables sélectionnées lors de la première étape, 5 ont été jugées pertinentes lors de cette deuxième étape : la différence de température potentielle entre 110 m et 2 m ($\Delta\theta$), les composantes du vent à 10 m ($u_{10\text{m}}$ et $v_{10\text{m}}$), le rayonnement net (Rn) et l'humidité spécifique ($Q_{2\text{m}}$).

5.2 Performance de prévision du vent avec le RNA

5.2.1 Comparaison aux résultats du RNA–Observations

Dans cette partie, les performances du RNA utilisant les cinq variables issues des simulations WRF sont comparées à celles du RNA utilisant les observations en entrée (voir le chapitre 4). Les deux jeux de données d'entrée (observations et variables WRF) couvrent la même période d'un an décrite dans la partie 2.3.

Les performances des deux RNA sont synthétisées dans le tableau 5.3. Suivant une méthodologie identique à ce qui a été mis en place au chapitre 4, les calculs de performance ont été réalisés à la fois sur le jeu de test complet et après élimination des vents inférieurs à $0,5\text{ m s}^{-1}$ produits par le RNA. Le retrait des vents faibles entraîne alors une augmentation des performances du RNA (à l'exception du PC4 qui classe déjà à part les vents faibles) ce qui montre une nouvelle fois que les vents faibles sont les plus difficiles à représenter pour le RNA. En revanche, les performances du RNA qui utilise les variables WRF en entrée sont significativement plus faibles que celles de celui qui utilise les observations. Cela peut être causé par les erreurs de prévision portées par les variables WRF utilisées comme entrées du RNA. Cette hypothèse est étudiée plus en détail dans la partie 5.2.3. Le problème de sur-représentation des vents alignés avec l'axe de la vallée, déjà rencontré dans le cas du RNA basé sur les observations, est de nouveau visible.

5.2.2 Apport du RNA par rapport aux simulations WRF

La figure 5.1 permet de visualiser l'évolution de la représentation des prévisions des directions et des vitesses de vent entre les simulations WRF brutes (figures c et d) et après application du RNA (figures a et b). Comme cela a été vu dans la partie 2.3.4.b, les vents d'origine thermique qui descendent la vallée de Cadarache sont mal représentés dans les simulations brutes et l'alignement fréquent des vents avec l'axe de la vallée dans les observations n'est alors pas bien représenté. Le RNA corrige en partie cela, mais à l'inverse il produit ces vents trop fréquemment. Les vitesses sont largement surestimées en sortie de WRF avec un biais de $+1,07\text{ m s}^{-1}$ (tableau 5.4). Le RNA permet de nettement réduire ce biais sur la vitesse ($-0,34\text{ m s}^{-1}$) ainsi que d'augmenter la précision sur la prévision des directions (par exemple le DACC passe de 0,49 à 0,68). Tous les indicateurs statistiques utilisés s'améliorent ainsi avec l'utilisation du RNA (tableau 5.4). En revanche, le RNA produit trop de vents alignés avec l'axe de la vallée (figure 5.1.a).

Case	DACC	PC2	PC4	Corrélation			Vitesse (m s^{-1})	
				Vitesse	u	v	Biais	MAE
RNA_{WRF}	0,68 [0,66 ; 0,70]	0,76 [0,75 ; 0,78]	0,71 [0,69 ; 0,73]	0,78 [0,76 ; 0,80]	0,78 [0,76 ; 0,80]	0,70 [0,66 ; 0,73]	-0,34 [-0,37 ; -0,31]	0,62 [0,60 ; 0,65]
$\text{RNA}_{\text{WRF-LW}}$	0,75 [0,72 ; 0,77]	0,84 [0,82 ; 0,86]	0,66 [0,63 ; 0,68]	0,76 [0,73 ; 0,79]	0,80 [0,77 ; 0,82]	0,73 [0,69 ; 0,76]	-0,30 [-0,34 ; -0,25]	0,75 [0,71 ; 0,78]
$\text{RNA}_{1\text{yr}}$	0,81 [0,77 ; 0,84]	0,88 [0,85 ; 0,91]	0,83 [0,80 ; 0,86]	0,92 [0,90 ; 0,94]	0,94 [0,93 ; 0,96]	0,86 [0,83 ; 0,90]	-0,24 [-0,29 ; -0,19]	0,52 [0,48 ; 0,55]
$\text{RNA}_{1\text{yr-LW}}$	0,90 [0,87 ; 0,93]	0,96 [0,93 ; 0,98]	0,83 [0,80 ; 0,87]	0,91 [0,88 ; 0,93]	0,95 [0,93 ; 0,96]	0,88 [0,85 ; 0,91]	-0,18 [-0,25 ; -0,12]	0,55 [0,51 ; 0,60]

Tableau 5.3 : Récapitulatif des résultats pour les différents RNA construits en utilisant un an de données : RNA_{WRF} pour le réseau utilisant les variables WRF comme variables d'entrée ; $\text{RNA}_{\text{WRF-LW}}$ pour le même réseau après retrait des vents inférieurs à $0,5 \text{ m s}^{-1}$ produits par le RNA ; $\text{RNA}_{1\text{yr}}$ pour le réseau utilisant le jeu d'observations d'un an (cf. chapitre 4) ; $\text{RNA}_{1\text{yr-LW}}$ pour le même réseau après retrait des vents produits par le RNA inférieurs à $0,5 \text{ m s}^{-1}$. Pour chaque mesure de performance, la valeur moyenne est indiquée ainsi que son intervalle de confiance à 90 % (entre crochets).

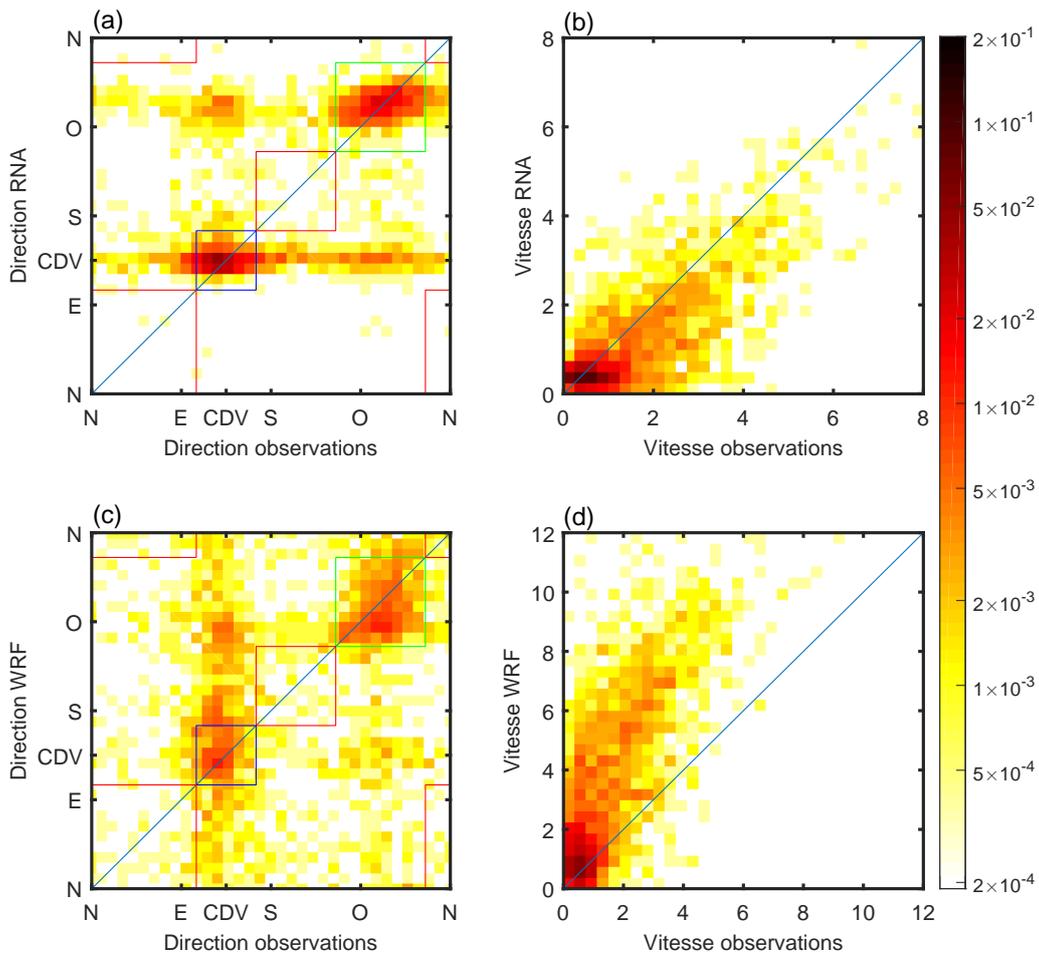


FIGURE 5.1 : Comparaison des prévisions de vent aux observations, représentée en densité de probabilité. En haut : en appliquant le RNA aux prévisions WRF ; en bas : prévisions WRF brutes. (a) et (c) pour les directions, (b) et (d) pour les vitesses. Les figures (c) et (d) sont l'équivalent de la figure 2.14, mais uniquement pour les données appartenant au jeu de test du RNA.

Cas	DACC	PC2	PC4	Corrélation			Vitesse (m s^{-1})	
				Vitesse	u	v	Biais	MAE
WRF	0,50	0,68	0,47	0,74	0,69	0,59	+1,07	1,32
RNA_{WRF}	0,68	0,76	0,71	0,78	0,78	0,70	-0,34	0,62
	[0,66 ; 0,70]	[0,75 ; 0,78]	[0,69 ; 0,73]	[0,76 ; 0,80]	[0,76 ; 0,80]	[0,66 ; 0,73]	[-0,37 ; -0,31]	[0,60 ; 0,65]

Tableau 5.4 : Récapitulatif de l'apport du RNA par comparaison des résultats de performance des simulations WRF brutes (ligne WRF) et du RNA utilisant les variables WRF comme entrée avec (ligne RNA_{WRF}). Tous ces indicateurs sont calculés seulement sur le jeu de test, ce qui explique les légères différences avec le tableau 2.5 en ce qui concerne les performances des simulations WRF. Pour chaque mesure de performance, la valeur moyenne est indiquée ainsi que (entre crochets) son intervalle de confiance à 90 %.

L'apport du RNA est également visible sur la figure 5.2 qui présente les distributions cumulées des erreurs de direction et de vitesse, pour les prévisions WRF brutes et les deux RNA appliqués respectivement aux observations et aux prévisions WRF. La zone grise permet de visualiser le gain significatif apporté par le RNA par rapport aux simulations brutes. Ainsi, les erreurs de direction inférieures à 20° représentent 27 % des cas en sortie de WRF contre 46 % après traitement par le RNA. La valeur médiane d'erreur de direction est de 46° en sortie de WRF, elle est quasiment divisée par deux (24°) après traitement par le RNA. La baisse de l'erreur commise sur les vitesses est également nettement visible sur la figure 5.2. Par exemple, la valeur du 75^{ième} centile de l'erreur est de $2,05 \text{ m s}^{-1}$ pour les prévisions WRF brutes contre $0,83 \text{ m s}^{-1}$ après traitement par le RNA.

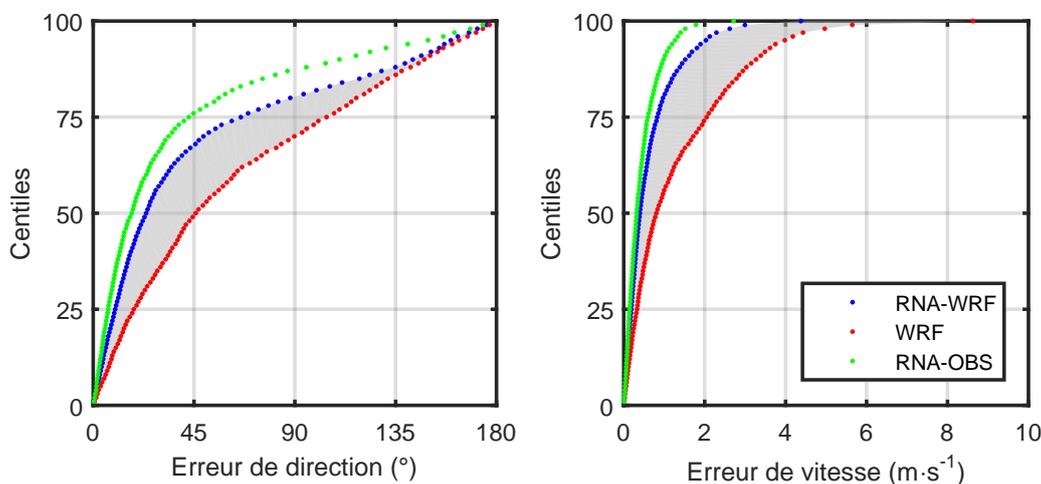


FIGURE 5.2 : Erreurs de direction (à gauche) et de vitesse (à droite) dans le calcul des vents par les prévisions WRF (en rouge), le RNA utilisant les prévisions WRF en entrée (en bleu) et le RNA utilisant les observations en entrée (en vert). La distribution cumulée des erreurs est représentée avec les centiles. La zone grise indique le gain produit par l'utilisation du RNA dans la prévision du vent par rapport aux simulations WRF brutes.

La figure 5.2 permet également de comparer les erreurs produites par les RNA utilisant soit les observations soit les variables WRF en entrée. Les performances sur la vitesse sont proches pour les deux réseaux, les erreurs étant légèrement inférieures en utilisant les observations comme entrées. Il en est de même pour la direction.

La figure 5.3 permet de visualiser l'évolution des performances avec la vitesse du vent. Les

simulations WRF ont leurs meilleures performances sur les directions pour les vents les plus forts. Comme cela a été montré dans les chapitres précédents, les vents forts en altitude ont plus de capacité à forcer les directions des vents les plus bas dans la vallée, qui ont alors des directions proches de celles des vents en altitude. Cela explique que les directions soient mieux représentées par vent fort dans les simulations malgré le manque de représentation du relief local.

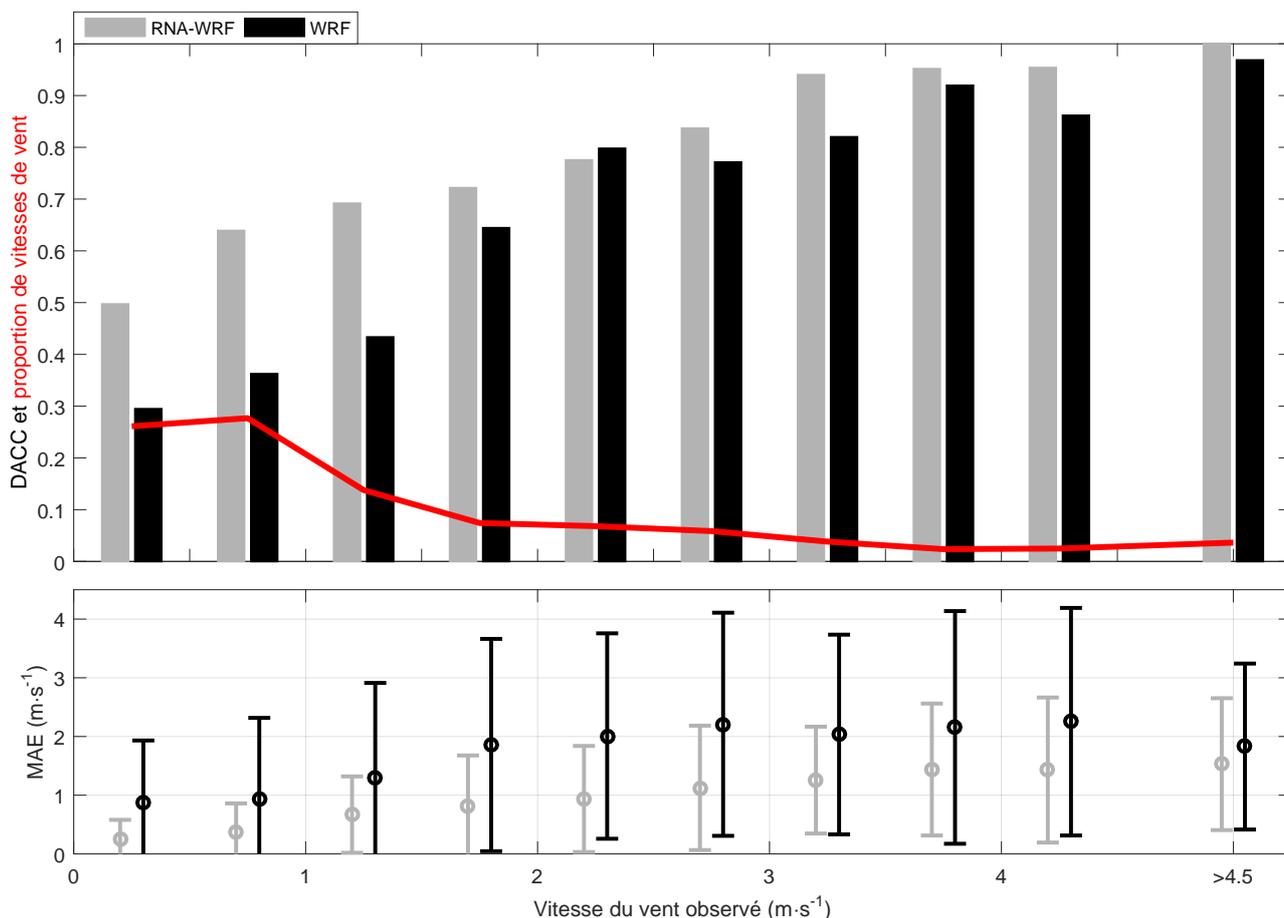


FIGURE 5.3 : En haut : évolution du DACC par classe de vitesse du vent observé à 2 m (largeur de classe de $0,5 \text{ m}\cdot\text{s}^{-1}$ et la dernière classe pour les vents supérieurs à $4,5 \text{ m}\cdot\text{s}^{-1}$) observé à 2 m et proportion de données composant chaque classe de vent (ligne rouge). En bas : Erreur absolue moyenne et écart-type de l'erreur sur la prévision de la vitesse. Sur les deux graphes, les tracés sont en noir pour les vents bruts prévus par WRF et en gris pour les vents calculés par le RNA à partir des prévisions WRF.

Pour les simulations WRF brutes, le DACC est constamment supérieur à 0,8 pour les vents supérieurs à $3,0 \text{ m}\cdot\text{s}^{-1}$ et vaut 1 pour les vents supérieurs à $5,5 \text{ m}\cdot\text{s}^{-1}$ alors qu'il est inférieur à 0,4 pour les vents inférieurs à $1,0 \text{ m}\cdot\text{s}^{-1}$. La marge d'amélioration est faible pour les vents forts ; malgré cela l'application du RNA permet de gagner encore en précision par rapport aux résultats des simulations WRF brutes. Bien que les valeurs de DACC restent plus faibles pour les vents les plus faibles (entre 0,5 et 0,7 pour les vents inférieurs à $1,5 \text{ m}\cdot\text{s}^{-1}$), c'est sur ces derniers que l'amélioration est la plus importante, ce qui est particulièrement intéressant car ce sont les vents majoritairement observés dans la vallée en conditions stables.

Concernant l'erreur faite sur la vitesse dans les simulations WRF (figure 5.3 bas), elle est importante (MAE toujours proche de ou supérieure à $1 \text{ m}\cdot\text{s}^{-1}$ et supérieure à $2 \text{ m}\cdot\text{s}^{-1}$ pour les vents forts de $2 \text{ m}\cdot\text{s}^{-1}$ et plus) et augmente avec la vitesse du vent. Le RNA permet de corriger

en partie cette erreur pour toutes les vitesses de vent.

Afin d'identifier les erreurs de prévision de direction que le RNA est capable de corriger (par rapport à l'erreur des simulations WRF brutes), les directions WRF sont comparées aux directions observées en distinguant les conditions menant à une bonne (à $\pm 45^\circ$) ou mauvaise prévision par le RNA (figure 5.4). Les données constituant la figure (a) (qui contient toutes les données du jeu de test du RNA) sont tracées séparément : pour les bonnes prévisions par le RNA sur la figure (b), et sur la figure (c) pour les mauvaises.

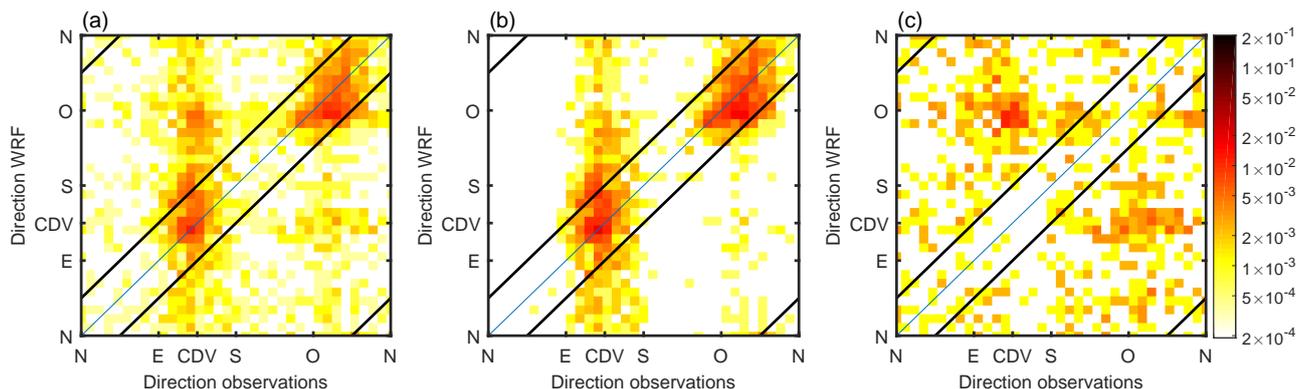


FIGURE 5.4 : (a) idem figure 5.1.c pour l'ensemble des données du jeu de test, (b) uniquement pour les données du jeu de test dont les directions calculées par le RNA sont correctes selon le critère DACC (déviations inférieures à 45° par rapport à la direction observée), et (c) pour les mauvaises prévisions par le RNA (directions calculées par le RNA éloignées de plus de 45° des observations). Les traits noirs délimitent les zones dans lesquelles les directions comparées ne diffèrent pas de plus de 45° , c'est à dire entre les deux lignes centrales et dans les coins supérieur gauche et inférieur droit.

Le DACC de WRF brut vaut 0,5, ce qui signifie que 50 % de directions sont bien prévues à $\pm 45^\circ$. Le RNA conserve cette bonne prévision pour 89 % des cas, ce qui est visible sur la figure puisque la majorité des points contenus entre les deux lignes noires centrales et dans les coins supérieur gauche et inférieur droit sont présents sur la figure (b). Les 11 % de données initialement bien prévues par WRF et que le RNA ne reproduit pas bien sont majoritairement les vents transverses à l'axe de la vallée et quelques vents de nord-ouest. Les bonnes prévisions initiales sur les vents CDV sont bien conservées avec le RNA.

Ensuite, parmi les 50 % de directions mal prévues par WRF, 47 % sont corrigées par le RNA (à $\pm 45^\circ$). Ces données correspondent majoritairement aux vents de sud-est (ce sont ces vents qui sont les moins bien représentés par les simulations WRF, 36 % de réussite contre 64 % pour les vents d'ouest-nord-ouest, voir la partie 2.3.4.b) comme cela est visible sur la figure 5.4.b. Les vents qui ne sont pas bien corrigés par le RNA (figure 5.4.c) sont majoritairement des vents pour lesquels WRF intervertit les directions sud-est et nord-ouest ainsi que les directions transverses à la vallée (même problème rencontré avec les observations en entrée, cf chapitre 4).

Enfin, un point important dans l'évaluation des capacités du RNA concerne le degré de réalisme de la relation entre le vent au-dessus de la vallée et au fond de celle-ci. Ceci est illustré sur la figure 5.5, qui présente les relations entre le vent à 2 m et le vent à 110 m, successivement pour les observations, pour les simulations brutes et pour le RNA appliqué aux simulations. Dans les simulations brutes, les vents de sud-est à 2 m ne sont présents que lorsqu'il y a concomitamment à 110 m des vents de sud-est (forçage) ou des vents de nord-est (vents d'origine

thermique descendant les reliefs situés au sud et à l'est de la vallée suffisamment grands pour être représentés dans la simulation, voir la partie 2.3.4.d) mais, contrairement aux observations, pas lorsque les vents en altitude viennent d'ouest-nord-ouest. Le RNA permet de produire ces vents et de manière générale il reproduit les différents liens qui existent entre les vents à 110 m et 2 m : canalisation des vents par la vallée de Cadarache lorsque les vents sont couplés avec les écoulements au-dessus de la vallée et vent d'origine thermique descendant la vallée lorsque les vents sont découplés.

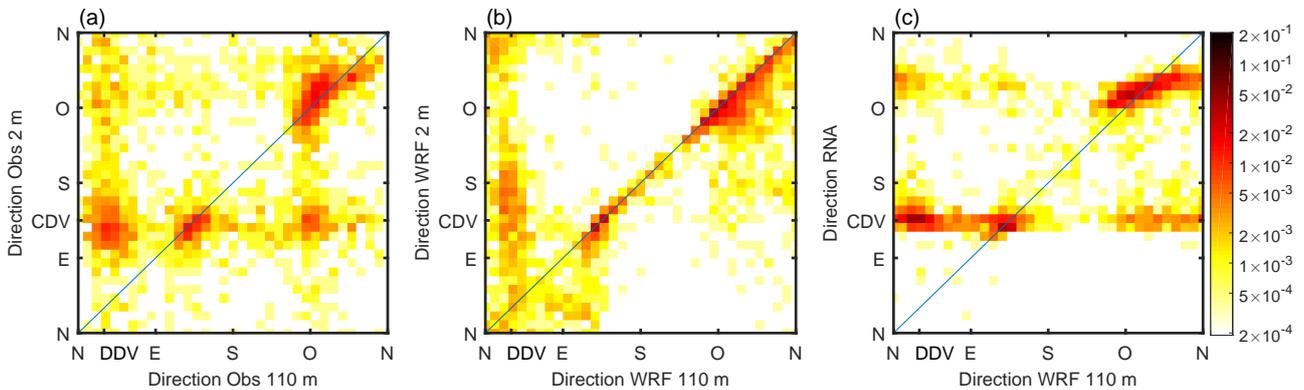


FIGURE 5.5 : Relation entre les directions des vents dans et au-dessus de la vallée de Cadarache : observations (a), prévisions WRF brutes (b) et prévisions WRF post-traitées par le RNA (c). Les figures (a) et (b) sont équivalentes aux figures 2.7.(a) et 2.16.(a) mais uniquement pour les données du jeu de test.

5.2.3 Analyse des erreurs

Il y a deux sources d'erreurs dans le calcul du RNA qui utilise les variables WRF comme entrées : l'erreur liée au calcul du RNA et celle liée à l'incertitude sur les variables d'entrée. La première est causée entre autres par la pertinence des variables utilisées (c'est l'erreur que l'on a aussi en utilisant les observations en entrée, cf. chapitre 4). La deuxième est causée par les erreurs de prévision des variables utilisées en entrée du RNA (fidélité par rapport à la réalité). Ces erreurs ne sont pas systématiques comme un biais constant qu'un RNA pourrait corriger facilement mais dues aux aléas classiques de la prévision (cf. <http://www.meteofrance.fr/>), et l'apprentissage du RNA est alors rendu plus difficile du fait de l'apport d'informations erronées et décorréelées de la réalité observée.

Parmi les erreurs existant dans les simulations WRF, un problème récurrent est une sous-estimation de l'amplitude thermique au cours du cycle diurne dont les causes peuvent être liées à la résolution dans les simulations, aux erreurs de flux turbulents près de la surface ou encore à une mauvaise représentation des caractéristiques de la surface (Cheng and Steenburgh, 2005; Hanna and Yang, 2001; Mass et al., 2002; Massey et al., 2014). Or, une erreur de ce type a un impact sur la représentation de la différence verticale de température potentielle (voir aussi la partie 2.3.4.f). La mauvaise représentation du relief dans les simulations impacte également le calcul des vents près de la surface. Une erreur sur le rayonnement net peut être causée par une mauvaise prévision du rayonnement solaire descendant, qui est régulièrement sur-estimé à cause d'un problème de représentation des nuages dans les simulations (Jiménez et al., 2016; Ruiz-Arias et al., 2016).

L'erreur de prévision peut donc avoir plusieurs sources, certaines sont plutôt des biais telle une mauvaise représentation de la surface (hétérogénéité du sol, relief), c'est dans le but de

corriger ces erreurs sur les vents près de la surface que le RNA est appliqué, certaines sont plus aléatoires comme certains mauvais forçages méso-échelle. Les effets d'un mauvais forçage par les vents présents au-dessus de la vallée et l'erreur de prévision de la différence verticale de température potentielle sont étudiés dans la suite du chapitre.

5.2.3.a Erreur de forçage au-dessus de la vallée

Afin d'étudier l'effet de l'erreur de forçage sur le calcul du RNA, le jeu de données d'entrée a été séparé en deux sous-ensembles en fonction de la qualité du forçage par le vent présent au-dessus de la vallée (à 110 m au-dessus du sol puisqu'une mesure de vent est disponible en continu à cette hauteur). Le forçage est jugé correct lorsque la direction du vent diffère de l'observation de moins de 45° . Les prévisions WRF du vent à 2 m sont bien meilleures dans le cas d'un bon forçage, comme on peut s'en rendre compte en comparant les lignes WRF du tableau 5.5. Seuls le biais et l'erreur absolue moyenne sont plus élevés avec un bon forçage ce qui s'explique par le fait que les bons forçages sont associés à des vents plus forts ($2,96 \text{ m s}^{-1}$ en moyenne contre $1,70 \text{ m s}^{-1}$ pour les vents associés à un mauvais forçage), or l'erreur absolue commise sur ces vents est plus élevée.

Un post-traitement RNA est effectué sur chacun des deux sous-ensembles de données d'entrée (4990 échantillons pour le sous-ensemble regroupant les données correspondant aux bons forçages contre 3032 pour les mauvais forçages). Le RNA avec le bon forçage atteint des performances qui sont très proches de celles du RNA qui utilise les observations en entrées (lignes RNA_{WRF} et $\text{RNA}_{1\text{YR}}$ dans le tableau 5.5). Les deux RNA ont des performances similaires pour les indicateurs de directions ainsi que pour le biais sur la vitesse. Cela montre qu'avec une bonne prévision des vents à 110 m, le RNA est capable de corriger les erreurs de simulation au fond de la vallée jusqu'à atteindre une performance équivalente à celle de l'utilisation d'observations. Cependant, les coefficients de corrélation sur les composantes et la vitesse ainsi que l'erreur absolue moyenne sur la vitesse sont légèrement inférieurs pour le RNA utilisant les prévisions WRF.

En revanche, le RNA appliqué aux mauvais forçages produit des performances significativement inférieures. Cela montre bien que l'erreur de prévision dans WRF impacte les performances du RNA. Ainsi, la réussite d'une descente d'échelle de simulations WRF avec un RNA pour le calcul du vent dans la vallée de Cadarache est conditionnée par l'utilisation d'une information à l'échelle supérieure qui soit de bonne qualité.

	Cas	DACC	PC2	PC4	Corrélation			Vitesse (m s^{-1})	
					Vitesse	u	v	Biais	MAE
Forçage bon	WRF	0,61	0,75	0,57	0,76	0,80	0,70	+1,27	1,48
	RNA _{WRF}	0,79 [0,77 ; 0,81]	0,86 [0,84 ; 0,87]	0,80 [0,78 ; 0,81]	0,81 [0,79 ; 0,83]	0,88 [0,87 ; 0,89]	0,79 [0,76 ; 0,81]	-0,23 [-0,27 ; -0,20]	0,63 [0,60 ; 0,66]
Forçage faux	WRF	0,33	0,55	0,31	0,52	0,14	0,18	+0,71	1,03
	RNA _{WRF}	0,66 [0,63 ; 0,68]	0,76 [0,73 ; 0,78]	0,73 [0,70 ; 0,75]	0,34 [0,25 ; 0,42]	0,44 [0,38 ; 0,50]	0,41 [0,34 ; 0,47]	-0,42 [-0,46 ; -0,37]	0,61 [0,57 ; 0,65]
-	RNA _{1yr}	0,81 [0,77 ; 0,84]	0,88 [0,85 ; 0,91]	0,83 [0,80 ; 0,86]	0,92 [0,90 ; 0,94]	0,94 [0,93 ; 0,96]	0,86 [0,83 ; 0,90]	-0,24 [-0,29 ; -0,19]	0,52 [0,48 ; 0,55]

Tableau 5.5 : Récapitulatif des résultats de performance des simulations WRF (lignes WRF) et du RNA utilisant les variables WRF comme entrée (lignes RNA_{WRF}) pour un bon ou un mauvais forçage au-dessus de la vallée (prévision de la direction du vent à 110 m correcte à $\pm 45^\circ$). La ligne RNA_{1yr} rappelle les résultats du RNA utilisant les observations en entrée. Pour chaque mesure de performance, la valeur moyenne est indiquée ainsi que (entre crochets) son intervalle de confiance à 90 %.

5.2.3.b Erreur de prévision de la différence de température potentielle

La différence de température potentielle entre les niveaux 110 m et 2 m, représentative de la stratification locale, est une variable déterminante pour la discrimination des vents d'origine thermique qui descendent la vallée de Cadarache (Duine et al., 2016) ainsi que pour la prévision immédiate du vent dans la vallée (voir le chapitre 4). L'étude de cette variable dans les prévisions WRF (partie 2.3.4.f) a montré que la stratification, globalement correcte, était cependant entachée de certaines erreurs, en particulier dans le phasage des transitions entre conditions stables et instables, et inversement, ainsi que sur la représentation de l'intensité pour les conditions les plus stratifiées. Il est donc intéressant d'évaluer l'impact que l'erreur de prévision sur cette variable a sur la précision du calcul du RNA.

Pour cela, le jeu de données d'entrées composé de variables WRF est modifié : la différence verticale de température potentielle issue de WRF est remplacée par la même variable issue des observations. Le RNA utilisant ce jeu de données est appelé RNA_{WRF/Obs} puisque ses entrées sont composées de quatre variables issues des simulations WRF et d'une issue des observations. Le tableau 5.6 montre l'impact sur les résultats. Les vitesses et les composantes sont similaires entre les deux RNA. En revanche, le calcul de la direction est plus précis en utilisant la différence de température potentielle observée. C'est un effet des vents faibles, pour lesquels l'erreur sur les vitesses reste limitée mais l'erreur sur la direction peut être importante. Or les vents de vallée en stratification stable (CDV) sont souvent faibles. Ceci montre qu'une partie de l'erreur du RNA sur ce paramètre est liée à l'erreur de prévision sur la différence de température potentielle.

Cette performance globale est analysée plus en détail à partir de la figure 5.6 qui permet de visualiser la qualité des prévisions de direction du vent en fonction de la valeur de la différence de température potentielle. Pour commencer, il est intéressant d'étudier la performance du RNA_{WRF} qui utilise uniquement des variables WRF comme entrées. Pour ce réseau, les performances sont les meilleures pour des atmosphères très stratifiées (DACC moyen de 0,79 pour $\Delta\theta > 4\text{K}$) ou au contraire instables (DACC de 0,82 pour $\Delta\theta < 0\text{K}$).

D'après les observations, plus la stratification est élevée, plus la probabilité que le vent soit d'origine thermique et descende la vallée est élevée (cf. figure 5.6). Notons que la proportion de

Cas	DACC	PC2	PC4	Corrélation			Vitesse (m s^{-1})	
				Vitesse	u	v	Biais	MAE
RNA_{WRF}	0,68 [0,66 ; 0,70]	0,76 [0,75 ; 0,78]	0,71 [0,69 ; 0,73]	0,78 [0,76 ; 0,80]	0,78 [0,76 ; 0,80]	0,70 [0,66 ; 0,73]	-0,34 [-0,37 ; -0,31]	0,62 [0,60 ; 0,65]
$\text{RNA}_{\text{WRF}/\text{Obs}}$	0,73 [0,71 ; 0,75]	0,83 [0,81 ; 0,85]	0,73 [0,71 ; 0,75]	0,79 [0,76 ; 0,81]	0,81 [0,78 ; 0,83]	0,69 [0,66 ; 0,72]	-0,35 [-0,38 ; -0,32]	0,60 [0,58 ; 0,63]

Tableau 5.6 : Récapitulatif des résultats de performance du RNA utilisant les variables WRF comme entrée (ligne RNA_{WRF}) et avec les mêmes variables WRF à l'exception de la différence de température potentielle observée (ligne $\text{RNA}_{\text{WRF}/\text{Obs}}$). Pour chaque mesure de performance, la valeur moyenne est indiquée ainsi que (entre crochets) son intervalle de confiance à 90 %.

vents CDV commence à augmenter significativement pour les $\Delta\theta > 2,5 \text{ K}$ ce qui est cohérent avec le seuil de $2,6 \text{ K}$ trouvé par [Duine et al. \(2016\)](#) et qui permet de diagnostiquer ces vents. Cela est généralement reproduit par le RNA. De même, plus la stratification est faible et plus la proportion de vents d'ouest à nord-ouest (CUV) augmente ce qui facilite la tâche du RNA (forçage direct par le vent sus-jacent) et explique que les performances sont également bonnes. Cependant, pour les conditions très stables les bonnes prévisions (à $\pm 45^\circ$) sont principalement dues aux vents CDV qui sont majoritaires. Mais le RNA est également capable de bien calculer le vent CDV qui est minoritaire en conditions instables. Pour ce dernier, cela s'explique par le fait que les vents CDV peuvent également résulter d'un forçage par les vents forts au-dessus de la vallée (notamment par conditions nuageuses ou pluvieuses), or un forçage synoptique fort facilite le calcul de la direction (cf. figure 5.3).

Le RNA produit ses moins bonnes performances (pour les directions) pour les atmosphères modérément stratifiées (DACC moyen de 0,62 pour $\Delta\theta$ compris entre $0,5 \text{ K}$ et 4 K), ce qui s'explique par le fait qu'aucune des trois classes de direction de vent n'est alors majoritaire, ce qui complique l'apprentissage du RNA. De plus, la proportion de vents transverses à l'axe de la vallée est plus importante, or nous avons vu précédemment que ces vents sont très mal reproduits par le RNA (car faibles et peu nombreux).

La figure 5.6 permet de visualiser l'impact de l'erreur de prévision de la différence de température potentielle sur le calcul de la direction du vent par le réseau. Les performances des deux RNA restent très proches pour les atmosphères neutres et instables ($\Delta\theta_{\text{OBS}} < 0,5 \text{ K}$, cf. tableau 5.7 et la quasi-identité des barres noires et grises sur la figure 5.6), ce qui s'explique par une bonne prévision du $\Delta\theta$ dans les simulations (cf. figure 5.7 et partie 2.3.4.f). Pour ces conditions, la stratification prévue par WRF donne d'aussi bons résultats que la stratification observée.

	RNA_{WRF}	$\text{RNA}_{\text{WRF}/\text{Obs}}$
$\Delta\theta_{\text{OBS}} < 0,5 \text{ K}$	0,78	0,77
$0,5 \text{ K} < \Delta\theta_{\text{OBS}} < 1,5 \text{ K}$	0,60	0,65
$1,5 \text{ K} < \Delta\theta_{\text{OBS}} < 4 \text{ K}$	0,63	0,62
$\Delta\theta_{\text{OBS}} > 4 \text{ K}$	0,79	0,86

Tableau 5.7 : Récapitulatif des valeurs de DACC calculées sur différents intervalles de $\Delta\theta_{\text{OBS}}$ pour les RNA_{WRF} et $\text{RNA}_{\text{WRF}/\text{Obs}}$.

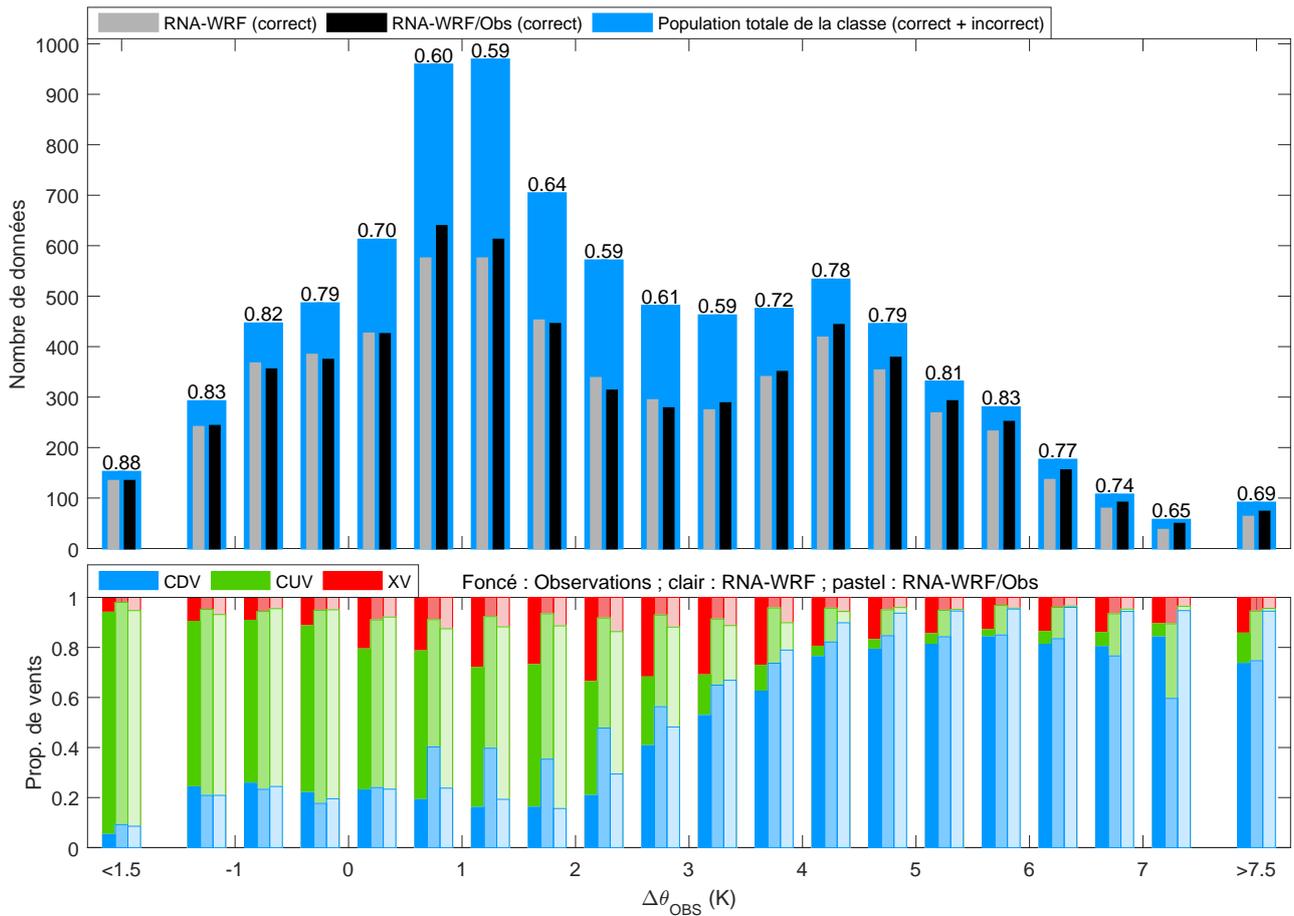


FIGURE 5.6 : En haut : distribution des données en fonction de la valeur de $\Delta\theta_{OBS}$ (en bleu, pour des classes de 0,5 K), nombre de données pour lesquelles le RNA_{WRF} (en gris) et le $RNA_{WRF/Obs}$ (en noir) produisent de bonnes prévisions de la direction du vent (à $\pm 45^\circ$) et valeur de DACC calculée sur le RNA_{WRF} (chiffres en haut des barres bleues), qui correspondent donc au rapport de la barre grise sur la bleue. En bas : proportion de directions de vents (observés) dans chacune des trois classes (CDV en bleu, CUV pour Cadarache upvalley en vert et vents transverses en rouge, définies à partir de la rose des vents observés à 2 m) par intervalle de $\Delta\theta_{OBS}$. Pour chaque valeur de $\Delta\theta_{OBS}$, le triplet de barres représente les observations (à gauche en couleurs foncées), le RNA_{WRF} (au centre) et le $RNA_{WRF/Obs}$ (à droite, couleur pastel).

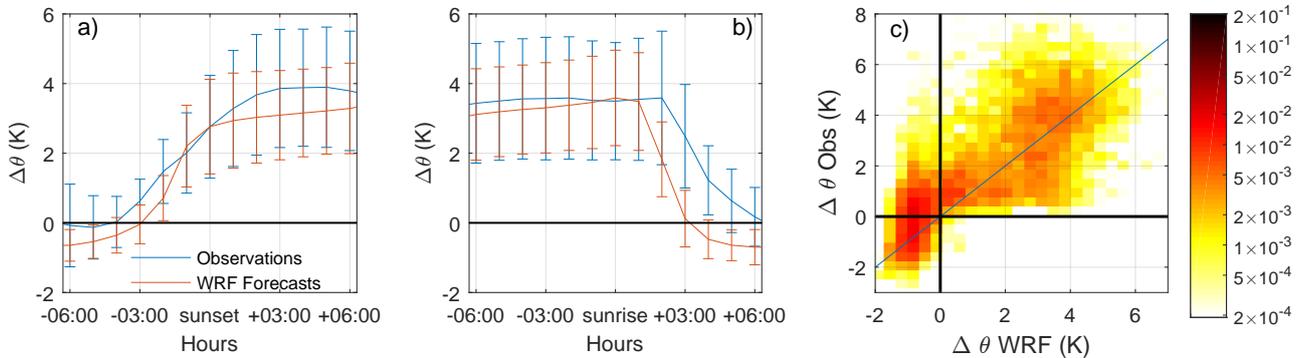


FIGURE 5.7 : Rappel de la figure 2.19. (a) : moyenne de la différence de température potentielle entre 110 m et 2 m présentée relativement à l'heure du coucher de soleil. En bleu pour les observations et en rouge pour la prévision. Les barres verticales représentent l'écart-type ; (b) : idem mais représenté par rapport au lever de soleil ; (c) : densité de valeurs montrant la comparaison entre les observations et les prévisions. Les données horaires de la période de 1 an ont été utilisées pour tracer cette figure.

En revanche, l'utilisation de la valeur de stratification observée permet de mieux représenter les vents pour les stratifications comprises entre 0,5 K et 1,5 K. Elles sont principalement associées aux transitions entre les conditions instables de la journée et stables nocturnes et vice versa (figure 5.7 et voir la partie 2.3.4.f). Ces stratifications sont mal prévues par WRF puisqu'elles s'étalent entre -2 K et 4 K à cause d'un déphasage des transitions dans la simulation. La figure 5.8 permet d'en apprécier l'impact sur les résultats du RNA. Notons que les bonnes performances des deux RNA (figure 5.8.a) sont caractérisées par une bonne prévision du $\Delta\theta$ (erreur absolue moyenne de 1,07 K). Pour les stratifications observées comprises entre 0,5 K et 1,5 K, l'utilisation de la variable observée comme entrée du RNA permet de nettement améliorer la prévision lorsque la stratification simulée est surestimée (voir les pixels rouges sur les figures 5.8.c et 5.8.d entre les deux lignes cyan du bas). La plus grande proportion de vents CDV avec le RNA_{WRF} (visible sur la figure 5.6) s'explique donc par ces cas de surestimation de la stratification dans les simulations WRF.

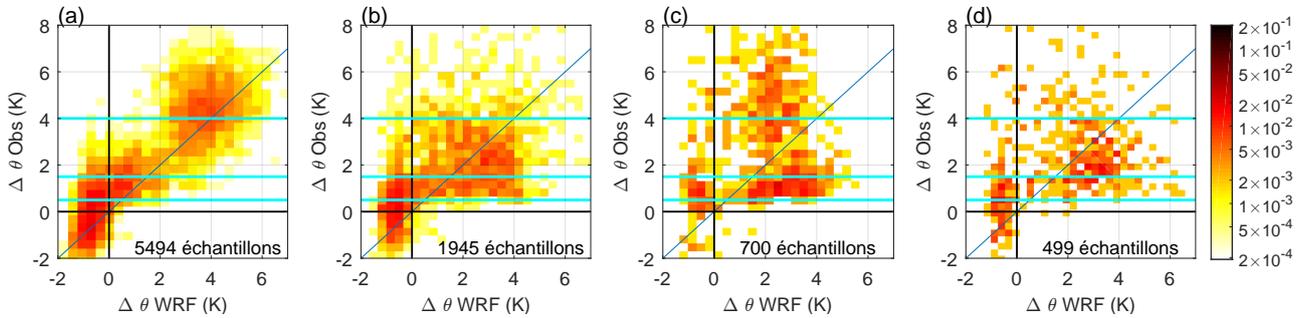


FIGURE 5.8 : Idem figure 5.7.c, (a) pour les échantillons produisant une bonne prévision de la direction du vent (à $\pm 45^\circ$) pour les deux RNA, (b) pour une mauvaise prévision sur les deux RNA, (c) pour une bonne prévision en utilisant le RNA_{WRF/Obs} et une mauvaise avec le RNA_{WRF} et (d) pour une bonne prévision en utilisant le RNA_{WRF} et une mauvaise avec le RNA_{WRF/Obs}. Les lignes cyan repèrent les $\Delta\theta_{OBS}$ de 0,5, 1,5 et 4 K.

Ensuite, les deux RNA sont très proches pour les $\Delta\theta_{OBS}$ compris entre 1,5 K et 4 K (tableau 5.7). Il y a pourtant, comme pour les stratifications comprises entre 0,5 K et 1,5 K, trop de vents CDV produits par le RNA_{WRF}, ce qui est en partie corrigé par le RNA_{WRF/Obs}. Cependant, ce gain est contrebalancé par une perte équivalente liée à des stratifications surestimées par WRF (voir les pixels rouges sur la figure 5.8.d entre les deux lignes cyan du haut).

Enfin, l'utilisation de la variable observée apporte un gain important pour les $\Delta\theta_{OBS}$ supérieurs à 4 K (figure 5.6 et tableau 5.7). Dans les observations, les vents CDV sont majoritaires pour les $\Delta\theta_{OBS} > 4$ K, or le $\Delta\theta_{WRF}$ ne représente pas toujours bien l'intensité de la stratification et ces vents sont sous-représentés avec le RNA_{WRF}. L'apport de l'observation se fait alors quand la stratification est sous-estimée dans WRF (voir les points rouges au-dessus de la ligne cyan supérieure sur la figure 5.8.c). La fréquence d'apparition des vents CDV est alors plus proche des observations avec le RNA_{WRF/Obs}.

5.3 Conclusion

Après avoir calculé avec succès le vent dans la vallée de Cadarache à partir d'observations réalisées en continu au-dessus de la vallée et à l'aide d'un réseau de neurones artificiels dans un but de prévision immédiate (chapitre 4), l'objectif dans ce chapitre a été d'appliquer cette

méthode pour produire une prévision de ce vent. Pour cela, les sorties de WRF à 3 km de résolution horizontale ont été utilisées. Le RNA permet alors de faire une descente d'échelle permettant la prévision d'un vent local non résolu par les simulations.

Le RNA appliqué aux simulations WRF connaît les mêmes limitations que celui utilisant des observations en entrée, à savoir une médiocre précision des directions de vents faibles et une sur-production de vents alignés avec l'axe de la vallée. Cependant, cette méthode de descente d'échelle statistique permet de nettement améliorer la qualité de prévision du vent dans la vallée par rapport aux prévisions WRF brutes, aussi bien pour la direction que pour la vitesse. Par exemple, la proportion de directions bien prévues à $\pm 45^\circ$ (DACC) passe de 49 % à 68 % et le biais sur la vitesse passe de $+1,07 \text{ m s}^{-1}$ à $-0,34 \text{ m s}^{-1}$ respectivement pour les prévisions WRF brutes et la prévision avec le RNA.

Toutefois, ces performances restent inférieures à celles obtenues avec les observations en entrées du RNA. La raison avancée pour expliquer cette différence est l'utilisation en entrée du RNA de variables comprenant des erreurs de prévision non systématiques. En effet, il a été montré qu'un forçage correct par les vents présents au-dessus de la vallée permet au RNA de produire des résultats bien meilleurs (équivalents à ceux obtenus avec le RNA utilisant des observations en entrée) que dans le cas où ce forçage n'est pas bon. De plus, il a également été montré l'importance de fournir des variables d'entrée réalistes au RNA, puisque dans le cas de la stratification stable, l'utilisation d'une variable réaliste permet de significativement améliorer la fiabilité du calcul. Il sera donc important à l'avenir de travailler à améliorer la représentation du vent au-dessus de la vallée et de la stratification dans les simulations WRF à 3 km de résolution.

Conclusion et perspectives

Conclusions générales

Le vent et les conditions de stratification de l'atmosphère gouvernent le comportement de la dispersion atmosphérique de substances émises près du sol. La stratification influe sur le niveau de turbulence qui lui-même pilote le mélange des substances dans l'atmosphère. Sur les terrains au relief complexe, la stratification pilote également l'apparition de phénomènes aérologiques particuliers et localisés tels que les vents de pente, impactant ainsi doublement la dispersion. Ces écoulements aérologiques présentent alors une grande variabilité spatiale qu'il est important de comprendre afin d'améliorer leur calcul par les modèles numériques et ainsi de mieux les prévoir.

Cette thèse a porté sur l'étude des vents sur le site de Cadarache et notamment dans la vallée éponyme. Cette vallée d'une longueur de 6 km, de largeur 1 km et dont les versants s'élèvent en moyenne à une centaine de mètres, débouche dans la vallée de la Durance, dont les dimensions sont d'un ordre de grandeur supérieur. La campagne de mesure KASCADE 2013 a été menée pendant l'hiver 2013 ; elle a permis d'étudier la structure verticale de la stratification et des écoulements. Une superposition de deux vents de vallée a alors été observée par nuit claire et vent calme, le premier s'écoulant dans les 50 premiers mètres dans la vallée de Cadarache et le second descendant la vallée de la Durance entre environ 100 et 400 m au-dessus du sol. Des simulations numériques à 3 km de résolution horizontales, qui constituent les prévisions météorologiques effectuées quotidiennement par le CEA, permettent de représenter le vent dans la vallée de la Durance mais pas celui associé à la vallée de Cadarache, trop petite pour être résolue.

Les objectifs de cette thèse étant l'amélioration de la prévision du vent sur la région de Cadarache en vue d'utiliser ces données pour des calculs de dispersion atmosphérique, le travail a été effectué suivant deux approches : expérimentale avec une seconde campagne KASCADE dont le but était d'améliorer la connaissance des écoulements, et numérique visant à la descente d'échelle des observations et des prévisions opérationnelles par modélisation statistique.

Les résultats expérimentaux : KASCADE 2017

La campagne KASCADE 2017 a permis de compléter les connaissances acquises à partir de la campagne effectuée en 2013 en élargissant les observations à l'ensemble de la vallée de Cadarache afin d'en étudier l'hétérogénéité spatiale, permettant ainsi d'observer le développement spatial et temporel de la stratification et des vents de vallée par conditions stables. Les observations continues, qui se sont étalées de mi-décembre à fin mai ont été réalisées à l'aide de 21 stations météorologiques. Ces stations ont été placées majoritairement suivant deux axes, une partie en fond de vallée le long de son axe de manière à suivre l'écoulement longitudinal, l'autre partie suivant un axe transverse de manière à observer les vents et la stratification sur les versants. De plus, des profils verticaux réalisés par radiosondage et ballon captif ont été effectués lors de 14 POI de 24 h afin d'observer la structure verticale de l'atmosphère lors de

nuits favorisant la formation de conditions stables.

À l’instar de KASCADE 2013, des conditions stables ont fréquemment été observées et ont conduit à la formation du vent de vallée de Cadarache. Ce vent apparaît généralement au coucher du soleil et disparaît environ 2 heures après le lever du soleil. De même, un vent de vallée de Durance a été observé au-dessus. La couche située entre ces deux vents de vallée est le siège d’un cisaillement qui génère une turbulence mécanique responsable localement d’une baisse de l’intensité de la stratification.

L’hétérogénéité des températures dans la vallée a été étudiée avec les 12 LEMS à partir de la méthode développée par [LeMone et al. \(2003\)](#). En journée, elle est fortement liée à l’orientation des pentes locales, les stations placées en adret sur le versant nord présentant donc les températures les plus élevées. L’hétérogénéité atteint ensuite un minimum au coucher du soleil, signe d’un passage par la neutralité, puis elle évolue proportionnellement à l’intensité de la stratification pendant la nuit et ce jusqu’à environ 2 heures après le lever du soleil.

Le type de stratification (suivant la classification de [Mahrt \(2017\)](#)) en place dans la vallée par conditions stables a été déterminé par comparaison des températures observées sur les versants de la vallée par les LEMS et en son milieu par radiosondage et par la station GBA. La stratification est de type cold pool marginale (proche de la cold pool mature), c’est à dire que les isentropes ne sont pas horizontales. Ce régime est associé à des écoulements de pente qui permettent le drainage d’air froid. L’étude des écoulements dans la vallée a d’ailleurs montré que des vents de pente existent sur les versants et alimentent un vent de vallée dans le fond, ce qui est cohérent avec le régime de cold pool marginale.

L’apparition et la disparition de ce vent de vallée ont été étudiées sur 5 nuits consécutives grâce aux stations avec anémomètres soniques situées dans le fond de la vallée. Elles ont permis d’apprécier des décalages temporels le long de la vallée. Le vent de vallée apparaît d’abord dans la partie la plus en amont puis se développe progressivement le long de la vallée jusqu’à atteindre son débouché. Le délai nécessaire à ce que ce vent s’établisse sur toute la longueur de la vallée varie suivant les nuits entre 45 minutes et 1h40. Ces différences ont été attribuées au forçage par le vent au-dessus de la vallée qui contrarie la formation du vent dans le bas de la vallée. Pour les mêmes raisons, les heures d’apparition de ce vent par rapport au coucher du soleil sont variables d’une nuit à l’autre. Les vents forts en altitude retardent alors l’apparition du vent de vallée et rallongent le délai nécessaire à ce qu’il occupe toute la longueur de la vallée. En revanche, il n’y a pas de tendance claire sur la chronologie de disparition de ce vent.

Des oscillations du vent de vallée (composantes longitudinale et transverse à l’axe de la vallée, vitesse et direction) ont fréquemment été observées. Elles produisent des directions oscillant autour de la direction moyenne du vent CDV avec une amplitude de 90 à 100° et avec une période qui est autour de 10 minutes, sauf en de rares cas où une période de 5 minutes a été observée. Des oscillations ont aussi été observées, mais de manière moins systématique sur les températures. Enfin, un cas d’intermittence dans la présence du vent CDV a été observé, avec une période de répétition d’environ une heure. Plusieurs études, sur d’autres sites, ont rapporté la présence d’oscillations dans le vent de vallée et l’intermittence dans la présence du vent CDV semble correspondre à la théorie énoncée par [Fleagle \(1950\)](#). Toutefois, l’étude de ce phénomène d’oscillation, qui ne faisait pas partie des objectifs initiaux de cette thèse, est restée descriptive et mérite un approfondissement afin d’en comprendre les mécanismes de formation.

Alors que dans la vallée la stabilité produit généralement un écoulement constitué de vents

descendant les pentes, certaines situations conduisent à un autre type de vent, notamment lorsque l'écoulement dans la vallée est fortement lié à celui au-dessus. [Whiteman and Doran \(1993\)](#) ont décrit les différentes relations pouvant exister entre ces deux écoulements, avec notamment des cas de couplage ou découplage. Plusieurs études sur différentes vallées ont tenté de déterminer les conditions conduisant à l'une ou l'autre de ces situations à partir d'une valeur seuil de la profondeur adimensionnelle de la vallée ($\hat{H} = Fr^{-1}$). Dans la vallée de Cadarache, le calcul de ce paramètre a conduit à la valeur seuil $\hat{H} = 1$ ce qui est cohérent avec les précédentes études. Elle a été déterminée en appliquant la méthode développée par [Duine et al. \(2016\)](#) avec un taux de réussite de 82 %.

Enfin, il a été observé que la distribution spatiale de la stratification présente des singularités. Les profils de température, qu'ils soient observés sur la verticale au centre de la vallée, ou bien près de la surface en remontant les versants de chaque côté, sont cohérents avec le schéma général d'une cold-pool mature ou marginale. En revanche, si l'on suit le fond de la vallée, on observe une stratification inversée, avec des températures près de la surface qui augmentent en descendant vers l'embouchure. En situation stable, les stations situées en aval dans la vallée connaissent ainsi des températures plus chaudes que celle du milieu, ce qui a été associé à un mélange plus important dans la partie aval de la vallée. Les hypothèses soulevées pour expliquer ce phénomène sont une interaction avec les écoulements de la vallée de Durance ou une augmentation de la turbulence générée par le vent de vallée de Cadarache. Une autre singularité a été observée à l'extrémité amont de la vallée, où les températures les plus froides ont été relevées. Ce phénomène reste encore à étudier pour en améliorer la compréhension, car le drainage de cet air froid vers l'aval semble bien se produire, ce qui devrait en limiter le refroidissement à cet endroit.

La modélisation statistique

L'instrumentation permanente disposée dans la vallée de Cadarache n'est pas adaptée à l'observation des hétérogénéités qui ont été documentées durant les deux campagnes de mesures KASCADE. KASCADE 2013 a notamment permis d'observer la structure verticale de l'écoulement, et ainsi de constater les différences entre les vents observés à 110 m sur la station permanente et ceux mesurés près de la surface dans la vallée par conditions stables. Or, les besoins du CEA à Cadarache pour le calcul de la dispersion atmosphérique portent en premier lieu sur les vents au niveau des sources d'émissions, c'est à dire près de la surface.

[Duine et al. \(2016\)](#) avaient développé une technique de prévision immédiate visant à diagnostiquer la présence ou l'absence d'un vent catabatique d'origine thermique dans la vallée de Cadarache à partir des observations permanentes qui y sont réalisées. La méthode avait été validée, mais elle fournit une information limitée à la présence ou absence de ce vent puisque ni la vitesse ni la direction ne sont calculées.

Dans le but de franchir une étape supplémentaire, un réseau de neurones artificiels a été utilisé pour calculer la direction et la vitesse du vent. Pour cela, le RNA a été construit de façon à reproduire les composantes du vent observées ponctuellement dans la vallée, permettant ensuite le calcul de la vitesse et de la direction, à partir des observations permanentes. Sa mise en œuvre comprend une étape de sélection des variables utilisées par le RNA visant à retenir les plus pertinentes vis à vis du calcul du vent. Ainsi, seulement trois variables sont nécessaires : la différence verticale de température potentielle mesurée entre 110 m et 2 m et les deux composantes du vent horizontal mesurées à 110 m. Les composantes du vent à 110 m présentent une corrélation non négligeable avec celles observées près de la surface (coefficients

de corrélations de 0,80 et 0,81 respectivement pour u et v), et servent donc de base au calcul alors que $\Delta\theta$ a été identifié comme la variable permettant d'améliorer le calcul des vents CDV.

Dans un premier temps, les résultats du RNA ont été comparés à ceux obtenus par la méthode de [Duine et al. \(2016\)](#). Pour les 3 mois de la période de KASCADE 2013, il y a une nette amélioration du point de vue du diagnostic de la présence ou non du vent de vallée, le taux de réussite passant de 78 % à 89 %. De plus, la vitesse et la direction du vent sont représentées de manière satisfaisante puisque le biais pour la vitesse n'est que de $-0,28 \text{ m s}^{-1}$ et 84 % des directions sont prévues avec une erreur inférieure à 45° . L'extension de la méthode à un jeu de données d'un an produit des résultats de même qualité. Cependant, bien que ces résultats aient été jugés satisfaisants, quelques limitations ont été rencontrées :

- les directions des vents faibles sont les plus difficiles à reproduire, ce qui est associé à la méthode de calcul basée sur les composantes, une erreur absolue (même faible) sur une composante d'un vent faible impactant la direction de manière significative ;
- les vents ayant des directions transverses à l'axe de la vallée sont les moins bien représentés, et ce pour plusieurs raisons : d'une part ces vents sont très faibles et donc plus difficiles à reproduire pour le RNA, d'autre part ils sont moins fréquents que les autres vents, et ils ne résultent pas d'une situation bien identifiée de la météorologie locale.

Ainsi, le RNA est capable de calculer avec une précision satisfaisante un vent local au fond de la vallée de Cadarache à partir de conditions représentatives de l'écoulement au-dessus de la vallée et de la stratification globale de la couche correspondant à la profondeur de la vallée. Bien que les RNA aient déjà été utilisés pour le calcul de nombreuses variables météorologiques dont la vitesse du vent, ceci représente, à notre connaissance, la première utilisation visant au calcul de la direction du vent. Cette méthode a ensuite été appliquée aux simulations WRF à 3 km de résolution horizontale avec un objectif de descente d'échelle dans la prévision du vent.

Le RNA appliqué aux simulations WRF connaît les mêmes limitations que celui utilisant des observations en entrée. Cependant, la qualité de prévision du vent dans la vallée est nettement améliorée par rapport aux simulations brutes, aussi bien pour la direction que pour la vitesse. Ainsi, la proportion de directions bien prévues à $\pm 45^\circ$ passe de 49 % à 68 % et le biais sur la vitesse passe de $+1,07 \text{ m s}^{-1}$ à $-0,34 \text{ m s}^{-1}$ respectivement pour les prévisions WRF brutes et la prévision par le RNA. Le potentiel des RNA utilisés comme outil de post-traitement de simulations méso-échelle a donc été démontré pour le cas de la descente d'échelle visant à prévoir un vent local. Il s'agit là d'une méthode relativement originale d'exploitation de prévisions opérationnelles.

Toutefois, ces performances restent inférieures à celles obtenues avec les observations en entrée du RNA. La raison avancée pour expliquer cette différence est l'utilisation de variables WRF comprenant des erreurs de prévision non systématiques. En effet, il a été montré qu'un forçage par les vents corrects au-dessus de la vallée permet au RNA de produire des résultats bien meilleurs (équivalents à ceux obtenus avec le RNA utilisant des observations en entrée) que dans le cas où ce forçage n'est pas bon. De plus, il a également été montré l'importance de fournir au réseau des stratifications réalistes pour significativement améliorer la précision du calcul. Il est donc important d'améliorer la représentation du vent au-dessus de la vallée et de la stratification locale dans les simulations WRF à 3 km de résolution.

Perspectives

Quelques interrogations subsistent concernant la compréhension de certains phénomènes observés pendant la campagne KASCADE 2017 et présentés dans cette thèse. Tout d'abord, les différentes hypothèses avancées pour expliquer le comportement particulier de la stratification dans le fond de la vallée par conditions stables sont à vérifier. Les données disponibles suggèrent un réchauffement par l'augmentation du mélange dans la partie aval de la vallée et une advection d'air plus froid dans la partie amont. Ce comportement s'est répété très régulièrement, mais il n'a été observé qu'avec les 4 stations LEMS placées dans le fond de la vallée, c'est pourquoi de nouvelles observations avec une multiplication des capteurs dans le fond de la vallée permettraient une meilleure description de la stratification le long de la vallée.

L'interaction entre les écoulements des vallées de Cadarache et de Durance est une hypothèse avancée pour expliquer l'augmentation de la turbulence dans la partie aval de la vallée de Cadarache. Le mât de 110 m, qui se situe à cet endroit, va prochainement être instrumenté sur plusieurs niveaux pour la mesure de la température, de l'humidité et des vents par anémomètres soniques 3D. Les profils obtenus ainsi de façon continue étendront largement la documentation qui a pu être réalisée par ballon captif lors des POI de KASCADE 2013. On disposera en effet d'un profil de la turbulence pour toutes les conditions de stratification.

Il reste également à améliorer la compréhension du phénomène responsable des oscillations observées dans et en dehors de la vallée de Cadarache. Les théories existantes concernant les mécanismes de génération de telles oscillations n'ont pas pu être validées. Vu la complexité apparente de ces mécanismes sur le site de Cadarache, il serait donc intéressant de se rapprocher de la communauté des spécialistes en mécanique des fluides afin d'explorer de nouvelles techniques d'analyse. En outre, la future instrumentation du mât de 110 m permettra d'observer l'étendue verticale de ces oscillations.

L'analyse des observations de KASCADE 2017 se poursuit également à l'université de l'Utah dans une étude basée sur l'utilisation de RNA actuellement en cours ([Gunawardena et al., np](#)), avec un objectif de prévision immédiate de variables mesurées par les LEMS. Dans cette étude, les données de certaines stations sont utilisées comme entrées du RNA qui est ensuite chargé de reproduire les conditions observées sur les autres stations. La recherche du groupe de stations permettant la meilleure reproduction des conditions sur les autres stations pourrait servir de base à une stratégie de déploiement optimal de l'instrumentation dans la vallée.

Par ailleurs, les RNA ont montré leur potentiel comme outil de post-traitement de prévisions à méso-échelle ce qui a permis d'effectuer une descente d'échelle statistique visant à prévoir le vent dans la vallée de Cadarache. Cependant, une limite de cette méthode est que le calcul ne produit qu'une prévision au point d'apprentissage. L'application de la méthode de descente d'échelle aux données des nombreuses stations utilisées pendant KASCADE 2017 pourrait alors permettre de produire un champ de vent de surface réparti dans la vallée.

Il a également été montré que la qualité de la prévision du RNA dépend de celle des données utilisées en entrées. Notamment, une amélioration de la prévision des vents au-dessus de la vallée et de l'intensité de la stratification permettra d'obtenir de meilleurs résultats de prévision avec le RNA.

Enfin, l'alternative aux méthodes statistiques pour la descente d'échelle dans la prévision des écoulements reste naturellement la modélisation déterministe qui doit être réalisée à très

haute résolution (environ 100 m, voire moins) pour prendre correctement en compte les hétérogénéités du terrain, et notamment le relief et l'occupation des sols, avec le réalisme nécessaire à l'obtention de prévisions fiables. Dans cette perspective, les observations de KASCADE 2017 constituent un jeu de données adéquat pour la validation de telles simulations dans la vallée de Cadarache. En particulier, les trois premières POI, réalisées du 17 au 20 janvier 2017, présentent de bonnes données pour l'étude des écoulements en conditions stables. De telles simulations sur ces trois POI permettraient alors d'améliorer la compréhension des écoulements dans la vallée, leurs phases de mise en place et de disparition, leurs interactions avec l'écoulement d'échelle supérieure, et permettraient d'établir les bases d'une plateforme mobilisable à terme au niveau opérationnel.

Bibliographie

- Ahn, J.-B., Lee, J., and Im, E.-S. (2012). The reproducibility of surface air temperature over South Korea using dynamical downscaling and statistical correction. *Journal of the Meteorological Society of Japan. Ser. II*, 90(4) :493–507. <https://doi.org/10.2151/jmsj.2012-404>.
- Arnfield, A. J. (2003). Two decades of urban climate research : a review of turbulence, exchanges of energy and water, and the urban heat island. *International Journal of Climatology*, 23(1) :1–26. <https://doi.org/10.1002/joc.859>.
- Banks, R. F., Tiana-Alsina, J., Baldasano, J. M., Rocadenbosch, F., Papayannis, A., Solomos, S., and Tzanis, C. G. (2016). Sensitivity of boundary-layer variables to PBL schemes in the WRF model based on surface meteorological observations, lidar, and radiosondes during the HygrA-CD campaign. *Atmospheric Research*, 176 :185–201. <https://doi.org/10.1016/j.atmosres.2016.02.024>.
- Barry, R. G. (1992). *Mountain weather and climate*. Routledge Physical Environment Series. <https://doi.org/10.1017/CBO9780511754753>.
- Bastin, S., Drobinski, P., Dabas, A., Delville, P., Reitebuch, O., and Werner, C. (2005). Impact of the Rhône and Durance valleys on sea-breeze circulation in the Marseille area. *Atmospheric Research*, 74(1-4) :303–328. <https://doi.org/10.1016/j.atmosres.2004.04.014>.
- Beale, M. H., Hagan, M. T., and Demuth, H. B. (1992). *Neural Network Toolbox User's Guide*. The MathWorks Inc, r2015b edition.
- Bell, R. C. and Thompson, R. O. (1980). Valley ventilation by cross winds. *Journal of Fluid Mechanics*, 96(4) :757–767. <https://doi.org/10.1017/S0022112080002340>.
- Berthou, S., Mailler, S., Drobinski, P., Arsouze, T., Bastin, S., Béranger, K., Flaounas, E., Lebeaupin Brossier, C., Somot, S., and Stéfanon, M. (2016). Influence of submonthly air–sea coupling on heavy precipitation events in the Western Mediterranean basin. *Quarterly Journal of the Royal Meteorological Society*, 142(S1) :453–471. <https://doi.org/10.1002/qj.2717>.
- Beyrich, F., Richter, S., Weisensee, U., Kohsiek, W., Lohse, H., De Bruin, H., Foken, T., Göckede, M., Berger, F., Vogt, R., et al. (2002). Experimental determination of turbulent fluxes over the heterogeneous LITFASS area : Selected results from the LITFASS–98 experiment. *Theoretical and Applied Climatology*, 73(1-2) :19–34. <https://doi.org/10.1007/s00704-002-0691-7>.
- Brioude, J., Arnold, D., Stohl, A., Cassiani, M., Morton, D., Seibert, P., Angevine, W., Evan, S., Dingwell, A., Fast, J. D., Easter, R. C., Pisso, I., Burkhardt, J., and Wotawa, G. (2013). The Lagrangian particle dispersion model FLEXPART-WRF version 3.1. *Geoscientific Model Development*, 6(6) :1889–1904. <https://doi.org/10.5194/gmd-6-1889-2013>.

- Brost, R. and Wyngaard, J. (1978). A model study of the stably stratified planetary boundary layer. *Journal of the Atmospheric Sciences*, 35(8) :1427–1440. [https://doi.org/10.1175/1520-0469\(1978\)035<1427:AMSOTS>2.0.CO;2](https://doi.org/10.1175/1520-0469(1978)035<1427:AMSOTS>2.0.CO;2).
- Businger, J. A., Wyngaard, J. C., Izumi, Y., and Bradley, E. F. (1971). Flux-profile relationships in the atmospheric surface layer. *Journal of the Atmospheric Sciences*, 28(2) :181–189. [https://doi.org/10.1175/1520-0469\(1971\)028<0181:FPRITA>2.0.CO;2](https://doi.org/10.1175/1520-0469(1971)028<0181:FPRITA>2.0.CO;2).
- Buzzi, M. (2008). *Challenges in operational numerical weather prediction at high resolution in complex terrain*. Thèse de doctorat, ETH Zurich. <https://doi.org/10.3929/ethz-a-005698833>.
- Cadenas, E. and Rivera, W. (2009). Short term wind speed forecasting in La Venta, Oaxaca, México, using artificial neural networks. *Renewable Energy*, 34(1) :274–278. <https://doi.org/10.1016/j.renene.2008.03.014>.
- Center, E. M. (2003). The GFS atmospheric model. *National Centers for Environmental Prediction Office Note*, 442 :14.
- Cheng, W. Y. and Steenburgh, W. J. (2005). Evaluation of surface sensible weather forecasts by the WRF and the Eta models over the Western United States. *Weather and Forecasting*, 20(5) :812–821. <https://doi.org/10.1175/WAF885.1>.
- Chou, M.-D. and Suarez, M. J. (1994). An efficient thermal infrared radiation parameterization for use in general circulation models. Tech. Memo. Vol. 3 104606, NASA.
- Clements, C. B., Whiteman, C. D., and Horel, J. D. (2003). Cold-air-pool structure and evolution in a mountain basin : Peter Sinks, Utah. *Journal of Applied Meteorology*, 42(6) :752–768. [https://doi.org/10.1175/1520-0450\(2003\)042<0752:CSAEIA>2.0.CO;2](https://doi.org/10.1175/1520-0450(2003)042<0752:CSAEIA>2.0.CO;2).
- Clements, W. E., Archuleta, J. A., and Hoard, D. E. (1989). Mean structure of the nocturnal drainage flow in a deep valley. *Journal of Applied Meteorology*, 28(6) :457–462. [https://doi.org/10.1175/1520-0450\(1989\)028<0457:MSOTND>2.0.CO;2](https://doi.org/10.1175/1520-0450(1989)028<0457:MSOTND>2.0.CO;2).
- Conangla, L., Cuxart, J., Jiménez, M. A., Martínez-Villagrasa, D., Miró, J. R., Tabarelli, D., and Zardi, D. (2018). Cold-air pool evolution in a wide Pyrenean valley. *International Journal of Climatology*, 38(6) :2852–2865. <https://doi.org/10.1002/joc.5467>.
- Coulibaly, P., Dibike, Y. B., and Anctil, F. (2005). Downscaling precipitation and temperature with temporal neural networks. *Journal of Hydrometeorology*, 6(4) :483–496. <https://doi.org/10.1175/JHM409.1>.
- Cros, B., Durand, P., Cachier, H., Drobinski, P., Frejafon, E., Kottmeier, C., Perros, P., Peuch, V.-H., Ponche, J.-L., Robin, D., et al. (2004). The ESCOMPTE program : an overview. *Atmospheric Research*, 69(3-4) :241–279. <https://doi.org/10.1016/j.atmosres.2003.05.001>.
- Cuxart, J. (2015). When can a high-resolution simulation over complex terrain be called LES ? *Frontiers in Earth Science*, 3 :87. <https://doi.org/10.3389/feart.2015.00087>.
- Deacon, E. (1968). The levelling error in Reynolds stress measurement. *Bulletin of the American Meteorological Society*, 49(8) :836.
- Delon, C., Serça, D., Boissard, C., Dupont, R., Dutot, A., Laville, P., De Rosnay, P., and Delmas, R. (2007). Soil NO emissions modelling using artificial neural network. *Tellus B*, 59(3) :502–513. <https://doi.org/10.1111/j.1600-0889.2007.00254.x>.

- Doran, J., Fast, J. D., and Horel, J. (2002). The VTMX 2000 campaign. *Bulletin of the American Meteorological Society*, 83(4) :537–551. [https://doi.org/10.1175/1520-0477\(2002\)083<0537:TVC>2.3.CO;2](https://doi.org/10.1175/1520-0477(2002)083<0537:TVC>2.3.CO;2).
- Dreyfus, G., Martinez, J.-M. Samuelides, M., Gordon, M., Badran, F., Thiria, S., and Hérault, L. (2002). *Réseaux de neurones – Méthodologie et applications*. Eyrolles.
- Dubayah, R. and Loechel, S. (1997). Modeling topographic solar radiation using GOES data. *Journal of Applied Meteorology*, 36(2) :141–154. [https://doi.org/10.1175/1520-0450\(1997\)036<0141:MTSRUG>2.0.CO;2](https://doi.org/10.1175/1520-0450(1997)036<0141:MTSRUG>2.0.CO;2).
- Duine, G.-J. (2015). *Characterization of down-valley winds in stable stratification from the KASCADE field campaign and WRF mesoscale simulations*. Thèse de doctorat, Université Toulouse III Paul Sabatier. <https://tel.archives-ouvertes.fr/tel-01263986>.
- Duine, G.-J., Hedde, T., Roubin, P., and Durand, P. (2016). A simple method based on routine observations to nowcast down-valley flows in shallow, narrow valleys. *Journal of Applied Meteorology and Climatology*, 55(7) :1497–1511. <https://doi.org/10.1175/JAMC-D-15-0274.1>.
- Duine, G.-J., Hedde, T., Roubin, P., Durand, P., Lothon, M., Lohou, F., Augustin, P., and Fourmentin, M. (2017). Characterization of valley flows within two confluent valleys under stable conditions : observations from the KASCADE field experiment. *Quarterly Journal of the Royal Meteorological Society*, 143(705) :1886–1902. <https://doi.org/10.1002/qj.3049>.
- Dyer, A. (1981). Flow distortion by supporting structures. *Boundary-Layer Meteorology*, 20(2) :243–251. <https://doi.org/10.1007/BF00119905>.
- Dyer, A. and Hicks, B. (1972). The spatial variability of eddy fluxes in the constant flux layer. *Quarterly Journal of the Royal Meteorological Society*, 98(415) :206–212. <https://doi.org/10.1002/qj.49709841518>.
- Entekhabi, D., Rodriguez-Iturbe, I., and Castelli, F. (1996). Mutual interaction of soil moisture state and atmospheric processes. *Journal of Hydrology*, 184(1-2) :3–17. [https://doi.org/10.1016/0022-1694\(95\)02965-6](https://doi.org/10.1016/0022-1694(95)02965-6).
- Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., et al. (2007). The Shuttle Radar Topography Mission. *Reviews of Geophysics*, 45(2). <https://doi.org/10.1029/2005RG000183>.
- Fernando, H., Pardyjak, E., Di Sabatino, S., Chow, F., De Wekker, S., Hoch, S., Hacker, J., Pace, J., Pratt, T., Pu, Z., et al. (2015). The MATERHORN : Unraveling the intricacies of mountain weather. *Bulletin of the American Meteorological Society*, 96(11) :1945–1967. <https://doi.org/10.1175/BAMS-D-13-00131.1>.
- Fleagle, R. G. (1950). A theory of air drainage. *Journal of Meteorology*, 7(3) :227–232. [https://doi.org/10.1175/1520-0469\(1950\)007<0227:ATOAD>2.0.CO;2](https://doi.org/10.1175/1520-0469(1950)007<0227:ATOAD>2.0.CO;2).
- Focken, U., Lange, M., Mönnich, K., Waldl, H.-P., Beyer, H. G., and Luig, A. (2002). Short-term prediction of the aggregated power output of wind farms – a statistical analysis of the reduction of the prediction error by spatial smoothing effects. *Journal of Wind Engineering and Industrial Aerodynamics*, 90(3) :231–246. [https://doi.org/10.1016/S0167-6105\(01\)00222-7](https://doi.org/10.1016/S0167-6105(01)00222-7).

- Gardner, M. and Dorling, S. (1998). Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences. *Atmospheric Environment*, 32(14) :2627–2636. [https://doi.org/10.1016/S1352-2310\(97\)00447-0](https://doi.org/10.1016/S1352-2310(97)00447-0).
- Georgelin, M., Bougeault, P., Black, T., Brzovic, N., Buzzi, A., Calvo, J., Casse, V., Desgagné, M., El-Khatib, R., Geleyn, J.-F., et al. (2000). The second COMPARE exercise : A model intercomparison using a case of a typical mesoscale orographic flow, the PYREX IOP3. *Quarterly Journal of the Royal Meteorological Society*, 126(564) :991–1029. <https://doi.org/10.1002/qj.49712656410>.
- Goger, B., Rotach, M. W., Gohm, A., Fuhrer, O., Stiperski, I., and Holtslag, A. A. (2018). The impact of three-dimensional effects on the simulation of turbulence kinetic energy in a major Alpine valley. *Boundary-Layer Meteorology*, 168(1) :1–27. <https://doi.org/10.1007/s10546-018-0341-y>.
- Grell, G. A., Peckham, S. E., Schmitz, R., McKeen, S. A., Frost, G., Skamarock, W. C., and Eder, B. (2005). Fully coupled « online » chemistry within the WRF model. *Atmospheric Environment*, 39(37) :6957 – 6975. <https://doi.org/10.1016/j.atmosenv.2005.04.027>.
- Gryning, S.-E., Mahrt, L., and Larsen, S. (1985). Oscillating nocturnal slope flow in a coastal valley. *Tellus A*, 37(2) :196–203. <https://doi.org/10.1111/j.1600-0870.1985.tb00281.x>.
- Gunawardena, N., Pardyjak, E., Durand, P., and Hedde, T. (n.p.). Data filling of wind fields in complex terrain for high-resolution nowcasting. En préparation.
- Gunawardena, N., Pardyjak, E., Stoll, R., and Khadka, A. (2018). Development and evaluation of an open-source, low-cost distributed sensor network for environmental monitoring applications. *Measurement Science and Technology*, 29(2) :024008. <https://doi.org/10.1088/1361-6501/aa97fb>.
- Haltiner, G. (1971). *Numerical weather prediction*. Wiley.
- Hanna, S. R. and Yang, R. (2001). Evaluations of mesoscale models’ simulations of near-surface winds, temperature gradients, and mixing depths. *Journal of Applied Meteorology*, 40(6) :1095–1104. [https://doi.org/10.1175/1520-0450\(2001\)040<1095:EOMMSO>2.0.CO;2](https://doi.org/10.1175/1520-0450(2001)040<1095:EOMMSO>2.0.CO;2).
- Hoch, S. W. and Whiteman, C. D. (2010). Topographic effects on the surface radiation balance in and around Arizona’s Meteor Crater. *Journal of Applied Meteorology and Climatology*, 49(6) :1114–1128. <https://doi.org/10.1175/2010JAMC2353.1>.
- Holden, J., Derbyshire, S., and Belcher, S. (2000). Tethered balloon observations of the nocturnal stable boundary layer in a valley. *Boundary-Layer Meteorology*, 97(1) :1–24. <https://doi.org/10.1023/A:1002628924673>.
- Hong, S.-Y. and Lim, J.-O. J. (2006). The WRF single-moment 6-class microphysics scheme (WSM6). *J. Korean Meteor. Soc*, 42(2) :129–151.
- Hong, S.-Y., Noh, Y., and Dudhia, J. (2006). A new vertical diffusion package with an explicit treatment of entrainment processes. *Monthly Weather Review*, 134(9) :2318–2341. <https://doi.org/10.1175/MWR3199.1>.
- Hsieh, C.-I., Katul, G., and Chi, T.-w. (2000). An approximate analytical model for footprint estimation of scalar fluxes in thermally stratified atmospheric flows. *Advances in Water Resources*, 23(7) :765–772. [https://doi.org/10.1016/S0309-1708\(99\)00042-1](https://doi.org/10.1016/S0309-1708(99)00042-1).

- Janjić, Z. I. (1994). The step-mountain eta coordinate model : Further developments of the convection, viscous sublayer, and turbulence closure schemes. *Monthly Weather Review*, 122(5) :927–945. [https://doi.org/10.1175/1520-0493\(1994\)122<0927:TSMECM>2.0.CO;2](https://doi.org/10.1175/1520-0493(1994)122<0927:TSMECM>2.0.CO;2).
- Jarraud, M. (2008). Guide to meteorological instruments and methods of observation (WMO-No. 8). *World Meteorological Organisation : Geneva, Switzerland*.
- Jiménez, M. A. and Cuxart, J. (2014). A study of the nocturnal flows generated in the north side of the Pyrenees. *Atmospheric Research*, 145 :244–254. <https://doi.org/10.1016/j.atmosres.2014.04.010>.
- Jiménez, P. A., Alessandrini, S., Haupt, S. E., Deng, A., Kosovic, B., Lee, J. A., and Delle Monache, L. (2016). The role of unresolved clouds on short-range global horizontal irradiance predictability. *Monthly Weather Review*, 144(9) :3099–3107. <https://doi.org/10.1175/MWR-D-16-0104.1>.
- Jiménez, P. A. and Dudhia, J. (2012). Improving the representation of resolved and unresolved topographic effects on surface wind in the WRF model. *Journal of Applied Meteorology and Climatology*, 51(2) :300–316. <https://doi.org/10.1175/JAMC-D-11-084.1>.
- Jiménez, P. A. and Dudhia, J. (2013). On the ability of the WRF model to reproduce the surface wind direction over complex terrain. *Journal of Applied Meteorology and Climatology*, 52(7) :1610–1617. <https://doi.org/10.1175/JAMC-D-12-0266.1>.
- Jiménez-Esteve, B., Udina, M., Soler, M. R., Pepin, N., and Miró, J. R. (2018). Land use and topography influence in a complex terrain area : A high resolution mesoscale modelling study over the Eastern Pyrenees using the WRF model. *Atmospheric Research*, 202 :49–62. <https://doi.org/10.1016/j.atmosres.2017.11.012>.
- Kaimal, J. and Haugen, D. (1969). Some errors in the measurement of Reynolds stress. *Journal of Applied Meteorology*, 8(3) :460–462. [https://doi.org/10.1175/1520-0450\(1969\)008<0460:SEITMO>2.0.CO;2](https://doi.org/10.1175/1520-0450(1969)008<0460:SEITMO>2.0.CO;2).
- Kalverla, P. C., Duine, G.-J., Steeneveld, G.-J., and Hedde, T. (2016). Evaluation of the Weather Research and Forecasting model in the Durance valley complex terrain during the KASCADE field campaign. *Journal of Applied Meteorology and Climatology*, 55(4) :861–882. <https://doi.org/10.1175/JAMC-D-15-0258.1>.
- Kani, S. P. and Ardehali, M. (2011). Very short-term wind speed prediction : a new artificial neural network–Markov chain model. *Energy Conversion and Management*, 52(1) :738–745. <https://doi.org/10.1016/j.enconman.2010.07.053>.
- Kolmogorov, A. N. (1941). Dissipation of energy in locally isotropic turbulence. In *Dokl. Akad. Nauk SSSR*, volume 32, pages 16–18. <https://doi.org/10.1098/rspa.1991.0076>.
- Kumar, J., Brooks, B.-G. J., Thornton, P. E., and Dietze, M. C. (2012). Sub-daily statistical downscaling of meteorological variables using neural networks. *Procedia Computer Science*, 9 :887–896. <https://doi.org/10.1016/j.procs.2012.04.095>.
- Largerion, Y. (2010). *Dynamique de la couche limite atmosphérique stable en relief complexe. Application aux épisodes de pollution particulaire des vallées alpines*. Thèse de doctorat, Université de Grenoble. <https://tel.archives-ouvertes.fr/tel-00606115/>.

- Largeron, Y. and Staquet, C. (2016). Persistent inversion dynamics and wintertime PM10 air pollution in Alpine valleys. *Atmospheric Environment*, 135 :92 – 108. <https://doi.org/10.1016/j.atmosenv.2016.03.045>.
- Largeron, Y., Staquet, C., and Chemel, C. (2013). Characterization of oscillatory motions in the stable atmosphere of a deep valley. *Boundary-Layer Meteorology*, 148(3) :439–454. <https://doi.org/10.1007/s10546-013-9825-y>.
- Legain, D., Bousquet, O., Douffet, T., Tzanos, D., Moulin, E., Barrié, J., and Renard, J.-B. (2013). High-frequency boundary layer profiling with reusable radiosondes. *Atmospheric Measurement Techniques*, 6 :2195–2205. <https://doi.org/10.5194/amt-6-2195-2013>.
- Lei, M., Shiyan, L., Chuanwen, J., Hongling, L., and Yan, Z. (2009). A review on the forecasting of wind speed and generated power. *Renewable and Sustainable Energy Reviews*, 13(4) :915–920. <https://doi.org/10.1016/j.rser.2008.02.002>.
- LeMone, M. A., Ikeda, K., Grossman, R. L., and Rotach, M. W. (2003). Horizontal variability of 2-m temperature at night during CASES-97. *Journal of the Atmospheric Sciences*, 60(20) :2431–2449. [https://doi.org/10.1175/1520-0469\(2003\)060<2431:HVOMTA>2.0.CO;2](https://doi.org/10.1175/1520-0469(2003)060<2431:HVOMTA>2.0.CO;2).
- Leukauf, D., Gohm, A., and Rotach, M. W. (2016). Quantifying horizontal and vertical tracer mass fluxes in an idealized valley during daytime. *Atmospheric Chemistry and Physics*, 16(20) :13049–13066. <https://doi.org/10.5194/acp-16-13049-2016>.
- Lilly, D. K. (1967). The representation of small-scale turbulence in numerical simulation experiments. In *Proceedings of IBM scientific computing symposium on environmental science*, pages 195–210, Yorktown Heights, N.Y. <https://doi.org/10.5065/D62R3PMM>.
- Liu, X. and Coulibaly, P. (2011). Downscaling ensemble weather predictions for improved week-2 hydrologic forecasting. *Journal of Hydrometeorology*, 12(6) :1564–1580. <https://doi.org/10.1175/2011JHM1366.1>.
- Lorente-Plazas, R., Jiménez, P. A., Dudhia, J., and Montávez, J. P. (2016). Evaluating and improving the impact of the atmospheric stability and orography on surface winds in the WRF model. *Monthly Weather Review*, 144(7) :2685–2693. <https://doi.org/10.1175/MWR-D-15-0449.1>.
- Mahrt, L. (1981). The early evening boundary layer transition. *Quarterly Journal of the Royal Meteorological Society*, 107(452) :329–343. <https://doi.org/10.1002/qj.49710745205>.
- Mahrt, L. (2011). Surface wind direction variability. *Journal of Applied Meteorology and Climatology*, 50(1) :144–152. <https://doi.org/10.1175/2010JAMC2560.1>.
- Mahrt, L. (2017). Stably stratified flow in a shallow valley. *Boundary-Layer Meteorology*, 162(1) :1–20. <https://doi.org/10.1007/s10546-016-0191-4>.
- Malardel, S. (2005). *Fondamentaux de météorologie*. Cépadués éditions.
- Mason, P. (1987). Diurnal variations in flow over a succession of ridges and valleys. *Quarterly Journal of the Royal Meteorological Society*, 113(478) :1117–1140. <https://doi.org/10.1002/qj.49711347804>.
- Mass, C. F., Ovens, D., Westrick, K., and Colle, B. A. (2002). Does increasing horizontal resolution produce more skillful forecasts? *Bulletin of the American Meteorological Society*, 83(3) :407–430. [https://doi.org/10.1175/1520-0477\(2002\)083<0407:DIHRPM>2.3.CO;2](https://doi.org/10.1175/1520-0477(2002)083<0407:DIHRPM>2.3.CO;2).

- Massey, J. D., Steenburgh, W. J., Hoch, S. W., and Knievel, J. C. (2014). Sensitivity of near-surface temperature forecasts to soil properties over a sparsely vegetated dryland region. *Journal of Applied Meteorology and Climatology*, 53(8) :1976–1995. <https://doi.org/10.1175/JAMC-D-13-0362.1>.
- Matzinger, N., Andretta, M., Gorsel, E. V., Vogt, R., Ohmura, A., and Rotach, M. (2003). Surface radiation budget in an Alpine valley. *Quarterly Journal of the Royal Meteorological Society*, 129(588) :877–895. <https://doi.org/10.1256/qj.02.44>.
- McCann, D. W. (1992). A neural network short-term forecast of significant thunderstorms. *Weather and Forecasting*, 7(3) :525–534. [https://doi.org/10.1175/1520-0434\(1992\)007<0525:ANNSTF>2.0.CO;2](https://doi.org/10.1175/1520-0434(1992)007<0525:ANNSTF>2.0.CO;2).
- McNider, R. T. (1982). A note on velocity fluctuations in drainage flows. *Journal of the Atmospheric Sciences*, 39(7) :1658–1660. [https://doi.org/10.1175/1520-0469\(1982\)039<1658:ANOVFI>2.0.CO;2](https://doi.org/10.1175/1520-0469(1982)039<1658:ANOVFI>2.0.CO;2).
- Milton, S. and Wilson, C. (1996). The impact of parameterized subgrid-scale orographic forcing on systematic errors in a global NWP model. *Monthly Weather Review*, 124(9) :2023–2045. [https://doi.org/10.1175/1520-0493\(1996\)124<2023:TIO PSS>2.0.CO;2](https://doi.org/10.1175/1520-0493(1996)124<2023:TIO PSS>2.0.CO;2).
- Mirocha, J., Kosović, B., and Kirkil, G. (2014). Resolved turbulence characteristics in large-eddy simulations nested within mesoscale simulations using the Weather Research and Forecasting Model. *Monthly Weather Review*, 142(2) :806–831. <https://doi.org/10.1175/MWR-D-13-00064.1>.
- Mlawer, E. J., Taubman, S. J., Brown, P. D., Iacono, M. J., and 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>.
- More, A. and Deo, M. (2003). Forecasting wind with neural networks. *Marine Structures*, 16(1) :35–49. [https://doi.org/10.1016/S0951-8339\(02\)00053-9](https://doi.org/10.1016/S0951-8339(02)00053-9).
- Muñoz, R. C., Falvey, M. J., Araya, M., and Jacques-Coper, M. (2013). Strong down-valley low-level jets over the Atacama desert : Observational characterization. *Journal of Applied Meteorology and Climatology*, 52(12) :2735–2752. <https://doi.org/10.1175/JAMC-D-13-063.1>.
- Muñoz-Esparza, D., Kosović, B., Mirocha, J., and van Beeck, J. (2014). Bridging the transition from mesoscale to microscale turbulence in numerical weather prediction models. *Boundary-Layer Meteorology*, 153(3) :409–440. <https://doi.org/10.1007/s10546-014-9956-9>.
- Nakanishi, M. and Niino, H. (2006). An improved Mellor–Yamada level-3 model : Its numerical stability and application to a regional prediction of advection fog. *Boundary-Layer Meteorology*, 119(2) :397–407. <https://doi.org/10.1007/s10546-005-9030-8>.
- Neff, W. and King, C. (1987). Observations of complex-terrain flows using acoustic sounders : experiments, topography, and winds. *Boundary-Layer Meteorology*, 40(4) :363–392. <https://doi.org/10.1007/BF00116103>.
- Niu, G.-Y., Yang, Z.-L., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., Kumar, A., Manning, K., Niyogi, D., Rosero, E., et al. (2011). The community Noah land surface model with multiparameterization options (Noah-MP) : 1. Model description and evaluation with local-scale measurements. *Journal of Geophysical Research : Atmospheres*, 116(D12109). <https://doi.org/10.1029/2010JD015139>.

- Paci, A., Staquet, C., Allard, J., Barral, H., Canut, G., Cohard, J.-M., Jaffrezo, J.-L., Martinet, P., Sabatier, T., Troude, F., et al. (2017). La campagne Passy-2015 : dynamique atmosphérique et qualité de l'air dans la vallée de l'Arve. *Pollution Atmosphérique*, 231–232. <https://doi.org/10.4267/pollution-atmospherique.5903>.
- Philippopoulos, K. and Deligiorgi, D. (2012). Application of artificial neural networks for the spatial estimation of wind speed in a coastal region with complex topography. *Renewable Energy*, 38(1) :75–82. <https://doi.org/10.1016/j.renene.2011.07.007>.
- Pinto, J., Parsons, D., Brown, W., Cohn, S., Chamberlain, N., and Morley, B. (2006). Coevolution of down-valley flow and the nocturnal boundary layer in complex terrain. *Journal of Applied Meteorology and Climatology*, 45(10) :1429–1449. <https://doi.org/10.1175/JAM2412.1>.
- Pleim, J. E. (2007). A combined local and nonlocal closure model for the atmospheric boundary layer. Part I : Model description and testing. *Journal of Applied Meteorology and Climatology*, 46(9) :1383–1395. <https://doi.org/10.1175/JAM2539.1>.
- Pond, S. (1968). Some effects of buoy motion on measurements of wind speed and stress. *Journal of Geophysical Research*, 73(2) :507–512. <https://doi.org/10.1029/JB073i002p00507>.
- Pope, S. B. (2001). Turbulent flows. *Measurement Science and Technology*, 12(11) :2020. <https://doi.org/10.1088/0957-0233/12/11/705>.
- Porch, W. M., Clements, W. E., and Coulter, R. L. (1991). Nighttime valley waves. *Journal of Applied Meteorology*, 30(2) :145–156. [https://doi.org/10.1175/1520-0450\(1991\)030<0145:NVW>2.0.CO;2](https://doi.org/10.1175/1520-0450(1991)030<0145:NVW>2.0.CO;2).
- Price, J., Vosper, S., Brown, A., Ross, A., Clark, P., Davies, F., Horlacher, V., Claxton, B., McGregor, J., Hoare, J., et al. (2011). COLPEX : field and numerical studies over a region of small hills. *Bulletin of the American Meteorological Society*, 92(12) :1636. <https://doi.org/10.1175/2011BAMS3032.1>.
- Princevac, M., Hunt, J., and Fernando, H. (2008). Quasi-steady katabatic winds on slopes in wide valleys : Hydraulic theory and observations. *Journal of the Atmospheric Sciences*, 65(2) :627–643. <https://doi.org/10.1175/2007JAS2110.1>.
- Ronda, R., Van den Hurk, B., and Holtslag, A. (2002). Spatial heterogeneity of the soil moisture content and its impact on surface flux densities and near-surface meteorology. *Journal of Hydrometeorology*, 3(5) :556–570. [https://doi.org/10.1175/1525-7541\(2002\)003<0556:SHOTSM>2.0.CO;2](https://doi.org/10.1175/1525-7541(2002)003<0556:SHOTSM>2.0.CO;2).
- Rontu, L. (2006). A study on parametrization of orography-related momentum fluxes in a synoptic-scale NWP model. *Tellus A*, 58(1) :69–81. <https://doi.org/10.1111/j.1600-0870.2006.00162.x>.
- Rotach, M. W., Calanca, P., Graziani, G., Gurtz, J., Steyn, D. G., Vogt, R., Andretta, M., Christen, A., Cieslik, S., Connolly, R., et al. (2004). Turbulence structure and exchange processes in an Alpine Valley : The Riviera project. *Bulletin of the American Meteorological Society*, 85(9) :1367–1386. <https://doi.org/10.1175/BAMS-85-9-1367>.
- Ruiz-Arias, J. A., Arbizu-Barrena, C., Santos-Alamillos, F. J., Tovar-Pescador, J., and Pozo-Vázquez, D. (2016). Assessing the surface solar radiation budget in the WRF model : A spatiotemporal analysis of the bias and its causes. *Monthly Weather Review*, 144(2) :703–711. <https://doi.org/10.1175/MWR-D-15-0262.1>.

- Salcedo-Sanz, S., Perez-Bellido, A. M., Ortiz-García, E. G., Portilla-Figueras, A., Prieto, L., and Correoso, F. (2009). Accurate short-term wind speed prediction by exploiting diversity in input data using banks of artificial neural networks. *Neurocomputing*, 72(4-6) :1336–1341. <https://doi.org/10.1016/j.neucom.2008.09.010>.
- Santos-Alamillos, F., Pozo-Vázquez, D., Ruiz-Arias, J., Lara-Fanego, V., and Tovar-Pescador, J. (2013). Analysis of WRF model wind estimate sensitivity to physics parameterization choice and terrain representation in Andalusia (Southern Spain). *Journal of Applied Meteorology and Climatology*, 52(7) :1592–1609. <https://doi.org/10.1175/JAMC-D-12-0204.1>.
- Santos-Alamillos, F., Pozo-Vázquez, D., Ruiz-Arias, J., and Tovar-Pescador, J. (2015). Influence of land-use misrepresentation on the accuracy of WRF wind estimates : evaluation of GLCC and CORINE land-use maps in southern Spain. *Atmospheric Research*, 157 :17–28. <https://doi.org/10.1016/j.atmosres.2015.01.006>.
- Schmidli, J., Böing, S., and Fuhrer, O. (2018). Accuracy of simulated diurnal valley winds in the Swiss Alps : Influence of grid resolution, topography filtering, and land surface datasets. *Atmosphere*, 9(5) :196. <https://doi.org/10.3390/atmos9050196>.
- Sheridan, P., Smith, S., Brown, A., and Vosper, S. (2010). A simple height-based correction for temperature downscaling in complex terrain. *Meteorological Applications*, 17(3) :329–339. <https://doi.org/10.1002/met.177>.
- Sheridan, P., Vosper, S., and Brown, A. (2014). Characteristics of cold pools observed in narrow valleys and dependence on external conditions. *Quarterly Journal of the Royal Meteorological Society*, 140(679) :715–728. <https://doi.org/10.1002/qj.2159>.
- Skamarock, W., Klemp, J., Dudhia, J., Gill, D., Barker, D., Duda, M., Huang, X., Wang, W., and Powers, J. (2008). A description of the advanced research WRF version 3. Note technique NCAR/TN-475+STR, 113pp, NCAR. <https://doi.org/10.5065/D68S4MVH>.
- Stull, R. B. (2012). *An introduction to boundary layer meteorology*, volume 13. Springer Science & Business Media. <https://doi.org/10.1007/978-94-009-3027-8>.
- Sukoriansky, S., Galperin, B., and Perov, V. (2005). Application of a new spectral theory of stably stratified turbulence to the atmospheric boundary layer over sea ice. *Boundary-Layer Meteorology*, 117(2) :231–257. <https://doi.org/10.1007/s10546-004-6848-4>.
- Taha, H. (1997). Urban climates and heat islands : albedo, evapotranspiration, and anthropogenic heat. *Energy and Buildings*, 25(2) :99–103. [https://doi.org/10.1016/S0378-7788\(96\)00999-1](https://doi.org/10.1016/S0378-7788(96)00999-1).
- Udina, M., Soler, M. R., and Sol, O. (2017). A modeling study of a trapped lee-wave event over the Pyrénées. *Monthly Weather Review*, 145(1) :75–96. <https://doi.org/10.1175/MWR-D-16-0031.1>.
- Van Gorsel, E., Vogt, R., Christen, A., and Rotach, M. W. (2003). Low frequency temperature and velocity oscillations in katabatic winds. In *Ext. Abstr., Vol A, Int. Conf. Alpine Meteor. and MAP Meeting*, pages 251–254.
- Vionnet, V., Bélair, S., Girard, C., and Plante, A. (2015). Wintertime subkilometer numerical forecasts of near-surface variables in the Canadian Rocky Mountains. *Monthly Weather Review*, 143(2) :666–686. <https://doi.org/10.1175/MWR-D-14-00128.1>.

- Vosper, S. and Brown, A. (2008). Numerical simulations of sheltering in valleys : the formation of nighttime cold-air pools. *Boundary-Layer Meteorology*, 127(3) :429–448. <https://doi.org/10.1007/s10546-008-9272-3>.
- Wagner, J., Gohm, A., and Rotach, M. (2015). Influence of along-valley terrain heterogeneity on exchange processes over idealized valleys. *Atmospheric Chemistry and Physics*, 15(12) :6589–6603. <https://doi.org/10.5194/acp-15-6589-2015>.
- Wagner, J. S., Gohm, A., and Rotach, M. W. (2014). The impact of horizontal model grid resolution on the boundary layer structure over an idealized valley. *Monthly Weather Review*, 142(9) :3446–3465. <https://doi.org/10.1175/MWR-D-14-00002.1>.
- Weber, R. O. and Kaufmann, P. (1998). Relationship of synoptic winds and complex terrain flows during the MISTRAL field experiment. *Journal of Applied Meteorology*, 37(11) :1486–1496. [https://doi.org/10.1175/1520-0450\(1998\)037<1486:ROSWAC>2.0.CO;2](https://doi.org/10.1175/1520-0450(1998)037<1486:ROSWAC>2.0.CO;2).
- Whiteman, C. D. (1990). Observations of thermally developed wind systems in mountainous terrain. In *Atmospheric processes over complex terrain*, pages 5–42. Springer. https://doi.org/10.1007/978-1-935704-25-6_2.
- Whiteman, C. D. (2000). *Mountain meteorology : fundamentals and applications*. Oxford University Press. [https://doi.org/10.1659/0276-4741\(2001\)021\[0200:MMFAA\]2.0.CO;2](https://doi.org/10.1659/0276-4741(2001)021[0200:MMFAA]2.0.CO;2).
- Whiteman, C. D., Allwine, K. J., Fritschen, L. J., Orgill, M. M., and Simpson, J. R. (1989). Deep valley radiation and surface energy budget microclimates. Part I : Radiation. *Journal of Applied Meteorology*, 28(6) :414–426. [https://doi.org/10.1175/1520-0450\(1989\)028<0427:DVRASE>2.0.CO;2](https://doi.org/10.1175/1520-0450(1989)028<0427:DVRASE>2.0.CO;2).
- Whiteman, C. D. and Doran, J. C. (1993). The relationship between overlying synoptic-scale flows and winds within a valley. *Journal of Applied Meteorology*, 32(11) :1669–1682. [https://doi.org/10.1175/1520-0450\(1993\)032<1669:TRBOSS>2.0.CO;2](https://doi.org/10.1175/1520-0450(1993)032<1669:TRBOSS>2.0.CO;2).
- Whiteman, C. D., Eisenbach, S., Pospichal, B., and Steinacker, R. (2004a). Comparison of vertical soundings and sidewall air temperature measurements in a small Alpine basin. *Journal of Applied Meteorology*, 43(11) :1635–1647. <https://doi.org/10.1175/JAM2168.1>.
- Whiteman, C. D., Haiden, T., Pospichal, B., Eisenbach, S., and Steinacker, R. (2004b). Minimum temperatures, diurnal temperature ranges, and temperature inversions in limestone sinkholes of different sizes and shapes. *Journal of Applied Meteorology*, 43(8) :1224–1236. [https://doi.org/10.1175/1520-0450\(2004\)043<1224:MTDTRA>2.0.CO;2](https://doi.org/10.1175/1520-0450(2004)043<1224:MTDTRA>2.0.CO;2).
- Wilczak, J. M., Oncley, S. P., and Stage, S. A. (2001). Sonic anemometer tilt correction algorithms. *Boundary-Layer Meteorology*, 99(1) :127–150. <https://doi.org/10.1023/A:1018966204465>.
- Wilks, D. S. (2011). *Statistical methods in the atmospheric sciences*, volume 100. Academic Press. <https://doi.org/10.1016/C2017-0-03921-6>.
- Wippermann, F. (1981). On the construction of orographically influenced wind roses for given distributions of the large-scale wind. *Beitr. Phys. Atmos.*, 54 :492–501.
- Wippermann, F. (1984). Air flow over and in broad valleys : Channeling and counter-current. *Beiträge zur Physik der Atmosphäre*, 57(1) :92–105.

-
- Wyngaard, J. C. (2004). Toward numerical modeling in the "Terra Incognita". *Journal of the Atmospheric Sciences*, 61(14) :1816–1826. [https://doi.org/10.1175/1520-0469\(2004\)061<1816:TNMITT>2.0.CO;2](https://doi.org/10.1175/1520-0469(2004)061<1816:TNMITT>2.0.CO;2).
- Yamartino, R. J. (1984). A comparison of several "single-pass" estimators of the standard deviation of wind direction. *Journal of Climate and Applied Meteorology*, 23(9) :1362–1366. [https://doi.org/10.1175/1520-0450\(1984\)023<1362:ACOSPE>2.0.CO;2](https://doi.org/10.1175/1520-0450(1984)023<1362:ACOSPE>2.0.CO;2).
- Zängl, G. (2005). Formation of extreme cold-air pools in elevated sinkholes : An idealized numerical process study. *Monthly Weather Review*, 133(4) :925–941. <https://doi.org/10.1175/MWR2895.1>.
- Zhou, B., Simon, J. S., and Chow, F. K. (2014). The convective boundary layer in the terra incognita. *Journal of the Atmospheric Sciences*, 71(7) :2545–2563. <https://doi.org/10.1175/JAS-D-13-0356.1>.

Annexes



Annexe A

Étude de sensibilité sur le calcul de la stratification

Cette annexe présente une étude de sensibilité pour évaluer l'impact des différentes approximations faites dans le calcul de l'intensité de la stratification à partir des mesures de la station GBA. En effet, cette station ne dispose pas de mesure d'humidité ni de pression à 110 m (le niveau haut), ce qui ne permet pas de calculer en toute rigueur une différence de température potentielle virtuelle entre 110 m et 2 m. Différentes méthodes visant à approximer la valeur de cette stratification sont alors étudiées.

A.1 Théorie

Les notations suivantes sont utilisées :

- ρ la densité de l'air (en kg m^{-3});
- ρ_d la densité de l'air sec (en kg m^{-3});
- ρ_v la densité de la vapeur d'eau (en kg m^{-3});
- θ la température potentielle (en K);
- θ_v la température potentielle virtuelle (en K);
- C_p la capacité calorifique de l'air à pression constante (en $\text{J K}^{-1} \text{kg}^{-1}$);
- C_{pd} la capacité calorifique de l'air sec à pression constante : $C_{pd} \approx 1004,67 \text{ J K}^{-1} \text{kg}^{-1}$;
- C_{pv} la capacité calorifique de la vapeur d'eau à pression constante : $C_{pv} \approx 1875 \text{ J K}^{-1} \text{kg}^{-1}$;
- e la pression partielle de vapeur (en Pa);
- P la pression atmosphérique (en Pa);
- m_h la masse molaire de l'air humide (en kg mol^{-1});
- R^* la constante universelle des gaz parfaits : $R^* = 8,314\,462\,1 \text{ J mol}^{-1} \text{K}^{-1}$;
- R la constante de l'air humide : $R = R^*/m_h$ (en $\text{J K}^{-1} \text{kg}^{-1}$);
- R_d la constante universelle pour l'air sec : $R_d \approx 287,04 \text{ J K}^{-1} \text{kg}^{-1}$;
- R_h la constante de la vapeur d'eau : $R_h = 461,5 \text{ J K}^{-1} \text{kg}^{-1}$;
- r_v le rapport de mélange de l'humidité $r_v = \rho_v/\rho_d$;
- T la température (en K);
- T_v la température virtuelle (en K);
- 1 et 2 les indices correspondants aux hauteurs z_1 et z_2 (en m).

A.1.1 Calcul théorique de la stratification

La stratification en densité est calculée par le gradient de température potentielle virtuelle. Son calcul ainsi que les approximations considérées sont détaillés dans les calculs qui suivent. Par définition, le rapport de mélange et la densité de l'air s'écrivent :

$$r_v = \rho_v / \rho_d \quad (\text{A.1})$$

$$\rho = \rho_d + \rho_v \quad (\text{A.2})$$

L'équation des gaz parfaits pour l'air humide, la vapeur d'eau et l'air sec s'écrit alors :

$$P = \rho RT = \rho_d (1 + r_v) RT \quad (\text{A.3})$$

$$e = \rho_v R_h T = \rho_d r_v R_h T \quad (\text{A.4})$$

$$P - e = \rho_d R_d T \quad (\text{A.5})$$

Par définition de la température potentielle virtuelle :

$$\rho RT = \rho R_d T_v \quad \Leftrightarrow \quad T_v = T \frac{R}{R_d} \quad (\text{A.6})$$

En combinant les équations A.3 A.4 et A.5, on obtient :

$$(1 + r_v) R - r_v R_h = R_d \quad \Leftrightarrow \quad R = R_d \frac{1 + (r_v R_h / R_d)}{1 + r_v} \quad (\text{A.7})$$

En remplaçant l'équation A.7 dans l'équation A.6, on obtient :

$$T_v = T \frac{1 + r_v / \epsilon}{1 + r_v} \quad \text{avec} \quad \epsilon = R_d / R_h \approx 0,622 \quad (\text{A.8})$$

Donc, en considérant que $r_v \ll 1$, on obtient au premier ordre :

$$T_v \approx (1 + 0,61r_v) T \quad (\text{A.9})$$

et de même on obtient :

$$\theta_v \approx (1 + 0,61r_v) \theta \quad (\text{A.10})$$

avec la température potentielle θ qui est définie par :

$$\theta = T \left(\frac{P_{\text{réf}}}{P} \right)^{R/C_p} \quad (\text{A.11})$$

et avec le rapport R/C_p qui s'écrit :

$$R/C_p = \frac{R_d}{C_{pd}} \frac{1 + r_v / \epsilon}{1 + (r_v C_{pv} / C_{pd})} \approx 0,2854 (1 - 0,24r_v) = k(r_v) \quad (\text{A.12})$$

La différence verticale de température potentielle virtuelle vaut alors :

$$\theta_{v2} - \theta_{v1} \approx T_2 (1 + 0,61r_{v2}) \left(\frac{P_{\text{réf}}}{P_2} \right)^{k2(r_{v2})} - T_1 (1 + 0,61r_{v1}) \left(\frac{P_{\text{réf}}}{P_1} \right)^{k1(r_{v1})} \quad (\text{A.13})$$

A.1.2 Exemples

L'objectif étant d'évaluer la sensibilité du calcul de l'intensité de la stratification en considérant différentes approximations, le calcul de l'équation A.13 est comparé à une méthode qui consiste à déterminer la pression au niveau z_2 en considérant l'approximation hydrostatique sur la couche considérée. Pour cela, la densité ρ_d est moyennée sur la couche et est calculée pour une atmosphère sèche :

$$P_{2, \text{hydro}} = P_1 + \rho_d g (z_1 - z_2) \quad \text{avec} \quad \rho_d = \frac{P_1 + P_2}{R_d (T_1 + T_2)} \quad (\text{A.14})$$

Une autre approximation consiste à retirer la dépendance à l'humidité dans le paramètre $k(r_v)$. Ainsi, le calcul de la différence de température potentielle virtuelle s'écrit :

$$\theta_{v2} - \theta_{v1} \approx T_2 (1 + 0,61r_{v2}) \left(\frac{P_{\text{réf}}}{P_2} \right)^{R_d/C_{pd}} - T_1 (1 + 0,61r_{v1}) \left(\frac{P_{\text{réf}}}{P_1} \right)^{R_d/C_{pd}} \quad (\text{A.15})$$

De plus, différents cas de figure sont considérés (voir le tableau A.1) afin d'apprécier l'impact de ces approximations sur le calcul de la différence de température potentielle virtuelle. Dans les trois premiers cas du tableau, il n'y a pas d'humidité, ce qui signifie que $\Delta\theta_v = \Delta\theta$. Les grandes différences de température entre les trois cas ne sont alors liées qu'aux différences de pression. Les écarts de $\Delta\theta_v$ sont quant à eux très faibles.

Dans les trois cas suivants, l'humidité est supposée homogène entre les niveaux z_1 et z_2 ce qui est une approximation que l'on peut être amené à faire lorsque l'humidité n'est mesurée qu'à un seul des deux niveaux considérés. Le calcul de $\Delta\theta_v$ est alors donné dans l'équation A.16. Les résultats sont alors très proches de ceux obtenus sans humidité.

$$\theta_{v2} - \theta_{v1} \approx (1 + 0,61r_v) \left(T_2 \left(\frac{P_{\text{réf}}}{P_2} \right)^{k(r_v)} - T_1 \left(\frac{P_{\text{réf}}}{P_1} \right)^{k(r_v)} \right) \quad \text{avec} \quad r_v = r_{v1} = r_{v2} \quad (\text{A.16})$$

Dans les deux derniers cas, il y a un gradient vertical d'humidité, dont l'intensité est plus ou moins marqué, ce qui produit des différences plus importantes par rapport aux six cas précédents. Cela montre l'importance de la connaissance de l'humidité aux deux niveaux considérés pour le calcul de la stratification.

L'approximation consistant en la suppression de l'effet de l'humidité dans le calcul de la température potentielle (indices $_d$ dans le tableau A.1) n'a pas d'effet important sur le calcul de la stratification.

Enfin, l'utilisation de l'approximation hydrostatique pour le calcul de la pression au niveau z_2 produit des différences de température potentielle virtuelle très proches des valeurs obtenues avec la valeur de pression P_2 ce qui montre que, dans le cas où une mesure de pression est manquante au niveau z_2 , cette approximation peut être utilisée sans que l'erreur produite soit importante.

Le calcul de la stratification a donc des sensibilités variables par rapport aux différentes approximations appliquées. L'approximation hydrostatique et la suppression de la dépendance à l'humidité dans le paramètre $k(r_v)$ n'ont que peu d'impact sur le calcul et elles peuvent donc être appliquées sans craindre d'introduire d'erreurs importantes qui rendraient difficile une interprétation. Dans l'hypothèse où une mesure d'humidité est manquante, le fait de considérer

une humidité constante sur la couche ou de ne pas prendre en compte l'humidité dans le calcul donne des approximations proches l'une de l'autre, mais qui surestiment la stratification d'autant plus que le gradient d'humidité réel est fort.

	1. Pas d'humidité, P standard	2. Pas d'humidité, P forte	3. Pas d'humidité, P faible	4. r_v faible et homogène	5. r_v moyen et homogène	6. r_v fort et homogène	7. Fort gradient de r_v	8. Faible gradient de r_v
T_1	270 K	270 K	270 K	270 K	270 K	270 K	270 K	270 K
T_2	272 K	272 K	272 K	272 K	272 K	272 K	272 K	272 K
r_{v_1}	0 g kg ⁻¹	0 g kg ⁻¹	0 g kg ⁻¹	2 g kg ⁻¹	5 g kg ⁻¹	10 g kg ⁻¹	5 g kg ⁻¹	4 g kg ⁻¹
r_{v_2}	0 g kg ⁻¹	0 g kg ⁻¹	0 g kg ⁻¹	2 g kg ⁻¹	5 g kg ⁻¹	10 g kg ⁻¹	2 g kg ⁻¹	3 g kg ⁻¹
P_1	1013 hPa	1030 hPa	980 hPa	1013 hPa	1013 hPa	1013 hPa	1013 hPa	1013 hPa
P_2	1001 hPa	1018 hPa	968 hPa	1001 hPa	1001 hPa	1001 hPa	1001 hPa	1001 hPa
δP_1	-13 hPa	-30 hPa	20 hPa	-13 hPa	-13 hPa	-13 hPa	-13 hPa	-13 hPa
δP_2	-1 hPa	-18 hPa	32 hPa	-1 hPa	-1 hPa	-1 hPa	-1 hPa	-1 hPa
$\delta P_{2, \text{hydro}}$	-0,3 hPa	-17,1 hPa	32,3 hPa	-0,3 hPa	-0,3 hPa	-0,3 hPa	-0,3 hPa	-0,3 hPa
θ_{v_1}	269,0 K	267,7 K	271,5 K	269,3 K	269,8 K	270,6 K	269,8 K	269,7 K
$\theta_{v_1, d}$	269,0 K	267,7 K	271,6 K	269,3 K	269,8 K	270,6 K	269,8 K	269,7 K
θ_{v_2}	271,9 K	270,6 K	274,5 K	272,3 K	272,8 K	273,6 K	272,3 K	272,4 K
$\theta_{v_2, d}$	271,9 K	270,6 K	274,5 K	272,3 K	272,8 K	273,6 K	272,3 K	272,4 K
$\theta_{v_2, \text{hydro}}$	272,0 K	270,7 K	274,5 K	272,3 K	272,8 K	273,6 K	272,3 K	272,5 K
$\Delta\theta_v$	2,9 K	2,9 K	3,0 K	3,0 K	3,0 K	3,0 K	2,5 K	2,7 K
$\Delta\theta_{v, \text{hydro}}$	3,0 K	3,0 K	3,0 K	3,0 K	3,0 K	3,0 K	2,5 K	2,8 K
$\Delta\theta_{v, d}$	2,9 K	2,9 K	3,0 K	2,9 K	2,9 K	2,9 K	2,4 K	2,8 K

Tableau A.1 : Récapitulatif des résultats obtenus lors du calcul de la différence verticale de température potentielle virtuelle entre deux niveaux (z_1 et z_2) séparés de 100 m à l'aide de l'équation A.13. Les calculs ont été réalisés en considérant différents états de l'atmosphère, définis par les températures T , rapports de mélange r_v et pressions P aux deux niveaux. La température potentielle virtuelle au niveau 2 a successivement été calculée en considérant la pression P_2 (θ_{v_2}) puis en calculant la pression à ce niveau à partir de l'approximation hydrostatique ($\theta_{v_2, \text{hydro}}$, voir l'équation A.14 pour le détail du calcul). Enfin, les température potentielles virtuelles portant l'indice d sont calculées à partir de l'équation A.15 qui consiste à négliger la dépendance à l'humidité du paramètre $k(r_v)$.

A.2 Calculs à partir du ballon captif (KASCADE 2013)

Dans cette partie, l'étude de sensibilité est réalisée à partir des profils verticaux (de température, pression et humidité) obtenus à l'aide du ballon captif déployé pendant la campagne KASCADE 2013. Ces profils ont été effectués uniquement durant la nuit et par conditions stables. Ce sont donc les conditions pour lesquelles les gradients verticaux sont les plus importants par comparaison à ceux que l'on observerait en journée dans une CLA mélangée.

Il y a un total de 789 profils disponibles pour lesquels les données ont été lissées sur 5 m à partir du niveau 2,5 m (voir Duine (2015) pour le détail des traitements effectués). Les différences de température sont calculées entre les niveaux 2,5 m et 112,5 m parce que ce sont les niveaux les plus proches des altitudes de mesure de la station GBA (2 et 110 m).

– Différence de température potentielle virtuelle : $\Delta\theta_v$

Tout d'abord, la température potentielle virtuelle est calculée en faisant quelques approximations puisqu'il est supposé que la dépendance à l'humidité du paramètre $k(r_v)$ est nulle (équation A.15) et que le rapport de mélange est approximé par la valeur de l'humidité spécifique ($r_v \approx q$).

– Différence de température potentielle : $\Delta\theta$

Le calcul de l'intensité de la stratification à partir d'une différence de température potentielle revient à ne pas prendre en compte l'effet de l'humidité de l'air sur sa densité (c'est équivalent au calcul dans une atmosphère sèche réalisée dans la partie A.1). Le biais commis sur le calcul de l'intensité de la stratification en comparaison avec la différence de température potentielle virtuelle est alors de $-0,05\text{ °C}$ et l'erreur absolue moyenne vaut $0,07\text{ °C}$ (figure A.1). Ces différences sont très faibles ce qui montre que la différence verticale de température potentielle est un bon indicateur de l'intensité de la stratification quand il n'est pas possible de calculer des températures potentielles virtuelles.

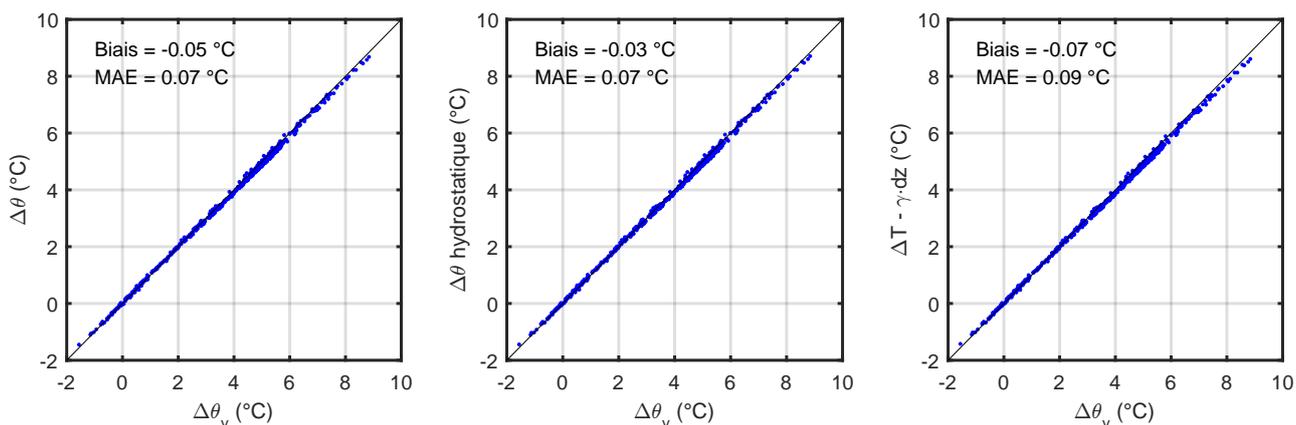


FIGURE A.1 : Comparaison entre la différence de température potentielle virtuelle et des différences de température potentielle entre 112,5 m et 2,5 m calculées de différentes manières. À gauche, les températures potentielle sont directement calculées à partir de la température et de la pression avec la formule de l'équation A.11. Au centre, la température potentielle à 112,5 m est calculée à partir d'une pression déterminée en considérant l'équilibre hydrostatique. À droite, un gradient adiabatique standard ($\gamma = -9,8\text{ °C km}^{-1}$) est ajouté à la différence de température entre les deux niveaux.

L'humidité a un rôle important dans le calcul de la stratification, notamment lorsque son gradient vertical est important (voir le tableau A.1). Or, les humidités spécifiques mesurées

dans les 110 premiers mètres avec le ballon captif varient peu avec l'altitude, les différences dépassant rarement $\pm 1 \text{ g kg}^{-1}$ (figure A.2), ce qui explique les faibles écarts entre $\Delta\theta_v$ et $\Delta\theta$.

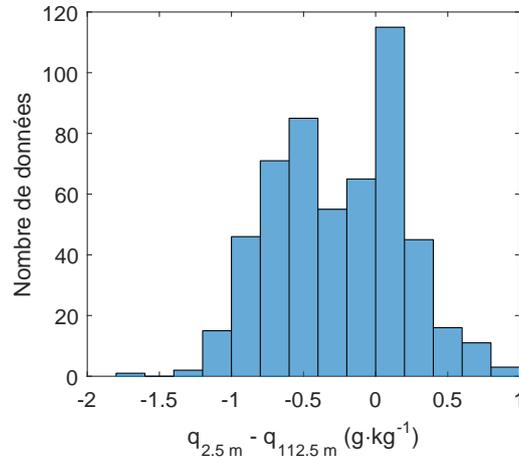


FIGURE A.2 : Distribution des différences d'humidité spécifique entre les niveaux 2,5 m et 112,5 m.

– **Approximation hydrostatique :** $\Delta\theta_{\text{hydrostatique}}$

Dans le cas où il n'y a pas de mesure de pression sur le point haut servant à calculer la différence de température, comme c'est le cas sur la station GBA, cette pression peut être estimée en considérant une approximation hydrostatique (équation A.14) où ρ est la densité moyenne de l'air dans la couche d'atmosphère considérée. Cette valeur estimée de la pression permet de calculer une température potentielle au point haut puis une différence de température potentielle. Le biais commis sur le calcul de l'intensité de la stratification en comparaison avec la différence de température potentielle virtuelle est alors de $-0,03^\circ\text{C}$ et l'erreur absolue moyenne vaut $0,07^\circ\text{C}$ (figure A.1), ce qui est équivalent au calcul avec la pression réellement mesurée et ce qui montre donc l'efficacité de cette méthode pour l'estimation de l'intensité de la stratification sur cette couche d'atmosphère en l'absence de mesure de pression à 110 m.

– **Gradient adiabatique :** $\Delta T - \gamma dz$

Cette dernière méthode consiste à utiliser une approximation de la transformation isentropique avec une valeur constante du paramètre $\gamma = \frac{dT}{dz} = -9,8^\circ\text{C km}^{-1}$. Dans la vallée de Cadarache, cette hypothèse produit des stratifications aux intensités très proches de celles obtenues avec une différence de températures potentielles virtuelles : le biais vaut $-0,07^\circ\text{C}$ et l'erreur absolue moyenne vaut $0,09^\circ\text{C}$ (figure A.1).

Les trois méthodes produisent des résultats similaires et très proches de ceux produits par le calcul de la différence de température potentielle virtuelle pour l'estimation de l'intensité de la stratification. Elles sont donc parfaitement adaptées aux mesures de la station GBA sur laquelle il n'y a ni mesure de pression ni mesure d'humidité à 110 m.

Annexe B

Inter-comparaison

La période d'inter-comparaison a débuté le 3 juin 2017 et s'est terminée le 3 juillet 2017. Les mesures des stations LEMS, MET01 et AS3 (les données des stations AS1, AS2, AS4 et AS5 sont manquantes pour cette période) sont comparées sur la base de moyennes 10 min.

	<i>LEMS</i>	<i>MET01</i>	<i>AS3</i>	référence
T_{2m}	$\pm 0,3$ à $0,5$ °C à 25 °C	$\pm 0,1$ °C	$\pm 0,1$ °C	AS3
RH_{2m}	$\pm 2,0$ à $4,5$ %	$\pm 0,8$ %	$\pm 0,8$ %	AS3
P_{2m}	± 1 hPa	$\pm 0,5$ à $1,5$ mbar		MET01
WS_{2m}	± 1 m s ⁻¹	-	$\pm 0,05$ m s ⁻¹	AS3
<i>Rayonnement</i>			-	MET01
$T_{surface}$	$\pm 0,6$ °C	-	-	LEMS

Tableau B.1 : Liste des observations disponibles pour la période d'inter-comparaison avec la précision de chaque capteur. La colonne référence indique la station utilisée comme référence pour la comparaison avec les autres stations. Pour la température de surface, la référence est calculée en faisant la moyenne des valeurs des 12 LEMS. La référence pour la vitesse du vent est calculée en faisant la moyenne entre les 3 anémomètres soniques de la station AS3.

B.1 Température à 2 m

Les stations MET01 et AS3 ont des capteurs de température qui sont en théorie plus précis que ceux des LEMS (tableau B.1). La station AS3 a été calibrée plus récemment que la station MET01, c'est donc cette station qui est prise comme référence pour les mesures de température à 2 m.

Les températures pendant la période d'inter-comparaison vont de $5,6$ °C à $36,6$ °C, alors qu'elles vont de $-11,5$ °C à $27,6$ °C pendant la période KASCADE 2017.

Les T_{2m} sont relativement similaires entre toutes les stations. Les valeurs de RMSE sont pour toutes les LEMS inférieures à la précision des capteurs et les biais sont très faibles (inférieurs à la précision du capteur de la station AS3, tableau B.2). Les RMSE calculées après correction avec les coefficients a et b (équation B.1) ne sont que très légèrement inférieures aux valeurs calculées sur les données non-corrigées. C'est aussi visible sur les graphes de la figure B.2 où les points bleus (non-corrigés) et verts (corrigés) sont très proches. En conséquence, il n'est pas utile d'appliquer une correction sur les données de température.

$$y = ax + b \quad \text{en } ^\circ\text{C} \quad (\text{B.1})$$

	a	b	R^2	Biais	RMSE	RMSE corrigé
MET01	0,99	0,28	1,00	0,09	0,24	0,21
LEMS-B1	0,99	0,39	1,00	0,05	0,25	0,23
LEMS-B2	0,98	0,37	1,00	0,02	0,28	0,26
LEMS-B3	0,99	0,27	1,00	0,00	0,26	0,25
LEMS-B4	0,98	0,44	1,00	-0,01	0,27	0,24
LEMS-N1	0,98	0,43	1,00	-0,03	0,30	0,27
LEMS-N2	0,98	0,43	1,00	-0,03	0,28	0,25
LEMS-N3	0,99	0,21	1,00	0,00	0,22	0,21
LEMS-N4	0,99	0,28	1,00	-0,04	0,26	0,24
LEMS-S1	0,99	0,24	1,00	-0,02	0,24	0,23
LEMS-S2	0,99	0,41	1,00	0,07	0,25	0,22
LEMS-S3	0,99	0,21	1,00	-0,03	0,19	0,17
LEMS-S4	0,99	0,35	1,00	0,03	0,22	0,20

Tableau B.2 : Résultats des comparaisons des T_{2m} entre la station AS3 et les stations MET01 et LEMS. Les valeurs a et b correspondent aux coefficients de la régression linéaire (équation B.1 avec y , x et b en $^\circ\text{C}$). R^2 est le coefficient de détermination correspondant à la régression linéaire. RMSE est la racine de l'erreur quadratique moyenne calculée pour les données non corrigées et RMSE corrigé est la valeur calculée pour les données corrigées à avec les coefficients produits par la régression linéaire.

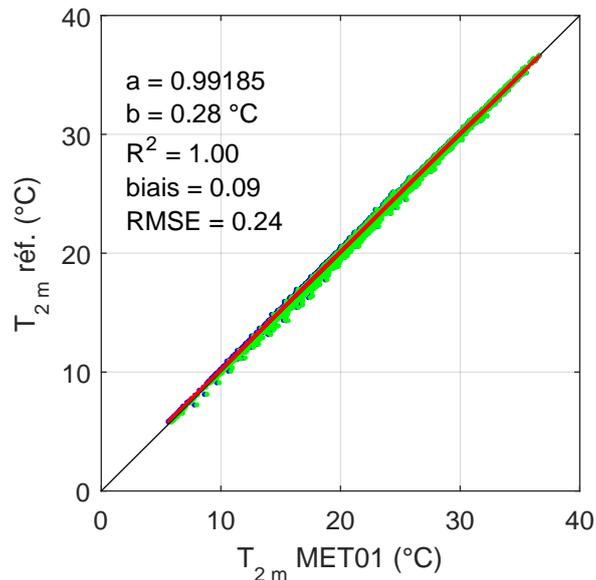


FIGURE B.1 : Comparaison de T_{2m} mesurée à AS3 et MET01 (points bleus) et valeurs corrigées (points verts). Droite de la régression linéaire (en rouge) et première bissectrice (en noir).

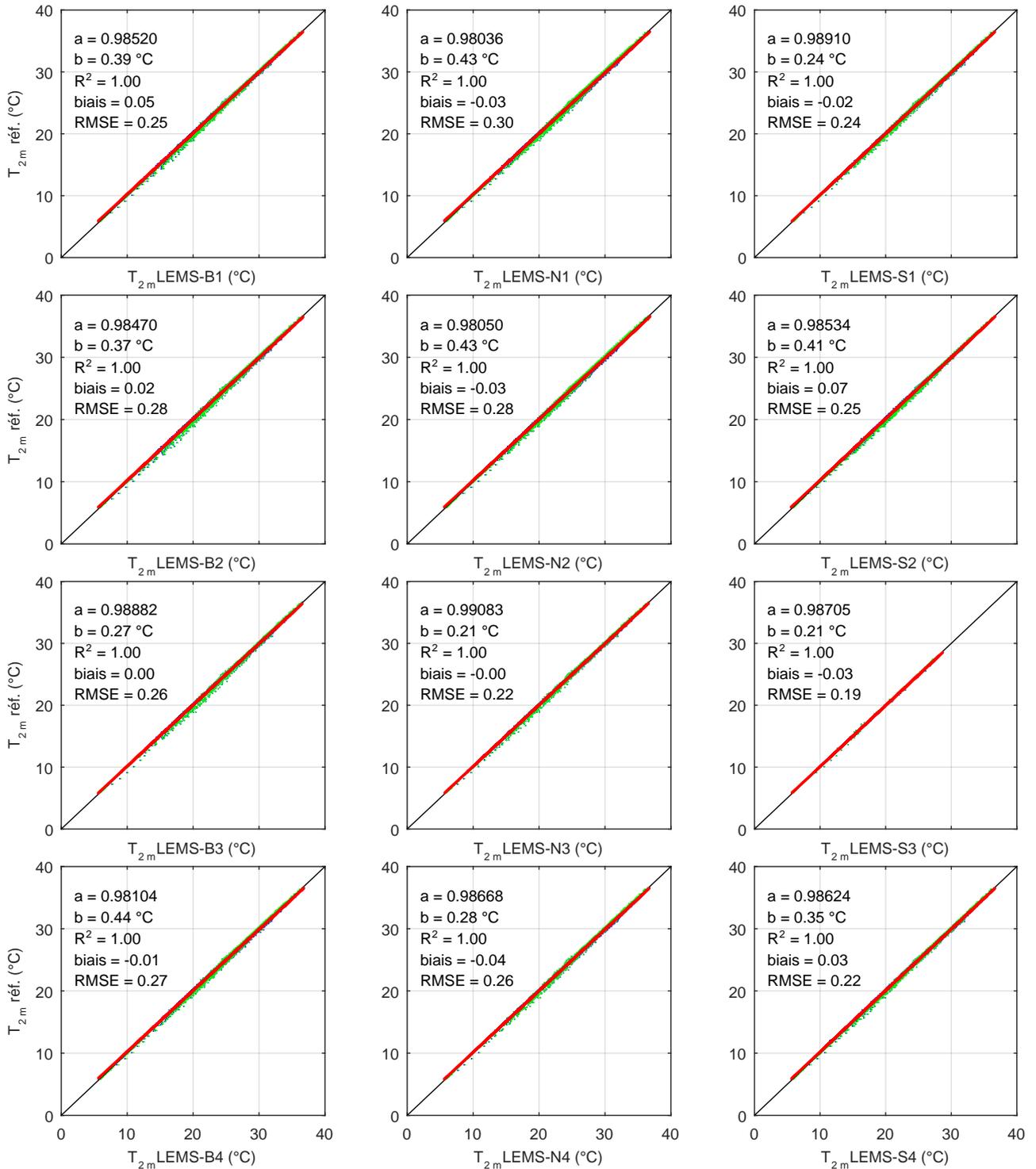


FIGURE B.2 : Comparaison de T_{2m} mesuré à AS3 et aux LEMS (points bleus) et valeurs corrigées (points verts). Droite de la régression linéaire (en rouge) et droite 1 : 1 (en noir).

B.2 Humidité relative à 2 m

La station AS3 est prise comme référence pour les mesures d'humidité relative à 2 m pour les mêmes raisons qui ont fait que cette station a été choisie comme référence pour les températures à 2 m.

Les humidités pendant la période d'inter-comparaison vont de 20,9 % à 98,8 % pour la station de référence, alors que pendant la période KASCADE 2017 elles s'évaluaient de 14,1 % à

99,1 %.

Les RH_{2m} sont proches entre toutes les stations avec des biais souvent inférieurs à la précision du capteur de la station AS3 ($\pm 0,8\%$). Seules les LEMS B1 et N4 et la station MET01 ont des biais supérieurs (jusque $2,76\%$). Les humidités ne sont utilisées que pour calculer des températures potentielles virtuelles. Les erreurs d'humidité relative impactent donc la valeur de l'humidité spécifique qui est utilisée pour le calcul de la température virtuelle :

$$T_v = T (1 + 0.61q) \quad (\text{B.2})$$

$$q = 0.622 \left(\frac{RH \cdot e_s}{100 \cdot P} \right) \quad (\text{B.3})$$

$$e_s = e_{s0} \exp \left(\frac{l_v}{R_v} \left(\frac{1}{T_0} - \frac{1}{T} \right) \right) \quad \text{formule de Clapeyron} \quad (\text{B.4})$$

où $e_{s0} = 6,11$ hPa est la pression de vapeur saturante pour un état de référence pris à la température de référence $T_0 = 273,15$ K, $l_v = 2,5 \times 10^6$ J kg⁻¹ la chaleur latente de vaporisation de l'eau, et $R_v = 461,5$ J K kg⁻¹ la constante des gaz pour la vapeur d'eau. La gamme de pressions enregistrées au cours de la campagne KASCADE 2017 s'étale de 963,0 à 1003,5 hPa. D'après les équations B.2 B.3 et B.4, les pressions faibles et les humidités relatives fortes sont les conditions qui impactent le plus fortement le calcul de la température virtuelle. À 26 °C (la température maximale enregistrée pendant KASCADE 2017) et à 963 hPa, des humidités relatives de 95 et 100 % (une différence de 5 points a été choisie parce que le RMSE maximal vaut 4,35) produisent des températures virtuelles qui diffèrent de 0,2 °C. Ce même écart calculé à une température de 0 °C vaut 0,03 °C ce qui est négligeable par rapport à la précision des capteurs.

Les humidités enregistrées pendant KASCADE ont des valeurs maximales de 100 % pour toutes les stations LEMS. En appliquant les corrections avec les coefficients des régressions linéaires (tableau B.3) les valeurs maximales dépassent les 100 %. Le faible impact des erreurs de mesure de l'humidité relative et le problème de saturation rencontré avec la correction des données à partir de la régression linéaire incitent donc à ne pas corriger ces données.

	a	b	R^2	Biais	RMSE	RMSE corrigé
MET01	1,03	0,16	1,00	2,01	2,30	0,88
LEMS-B1	1,16	-6,15	1,00	2,76	4,35	1,49
LEMS-B2	1,06	-4,38	1,00	-0,68	1,70	0,88
LEMS-B3	1,03	-2,44	1,00	-0,83	1,23	0,70
LEMS-B4	1,07	-4,59	1,00	-0,73	1,77	0,90
LEMS-N1	1,08	-5,08	1,00	-0,30	1,87	0,84
LEMS-N2	1,05	-4,08	1,00	-0,93	1,66	0,83
LEMS-N3	1,12	-6,65	1,00	0,14	2,45	0,88
LEMS-N4	1,04	-4,12	1,00	-1,49	1,86	0,65
LEMS-S1	1,13	-7,05	1,00	0,58	2,73	0,83
LEMS-S2	1,09	-5,67	1,00	-0,18	2,08	0,91
LEMS-S3	1,08	-4,73	1,00	-0,46	1,91	0,77
LEMS-S4	1,10	-5,82	1,00	-0,08	2,23	1,08

Tableau B.3 : Pareil que le tableau B.2 pour RH_{2m} entre la station AS3 et les stations MET01 et LEMS. Les coefficients a et b sont issus de l'équation B.1 dans laquelle y , x et b sont en %.

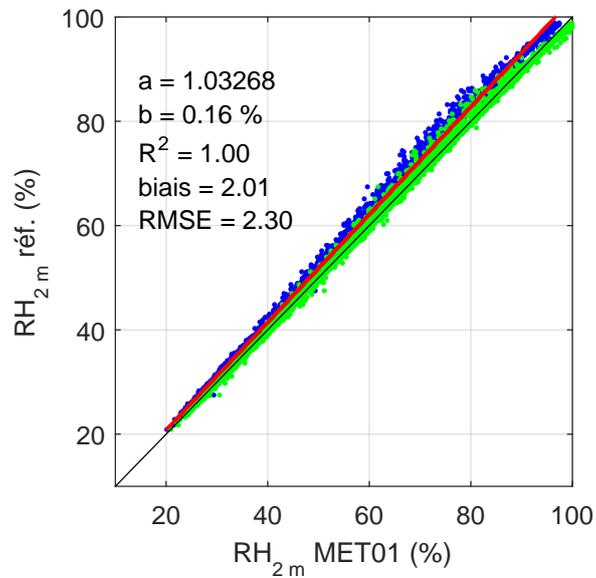


FIGURE B.3 : Comparaison de RH_{2m} mesuré à AS3 et MET01 (points bleus) et valeurs corrigées (points verts). Droite de la régression linéaire (en rouge) et droite 1 : 1 (en noir).

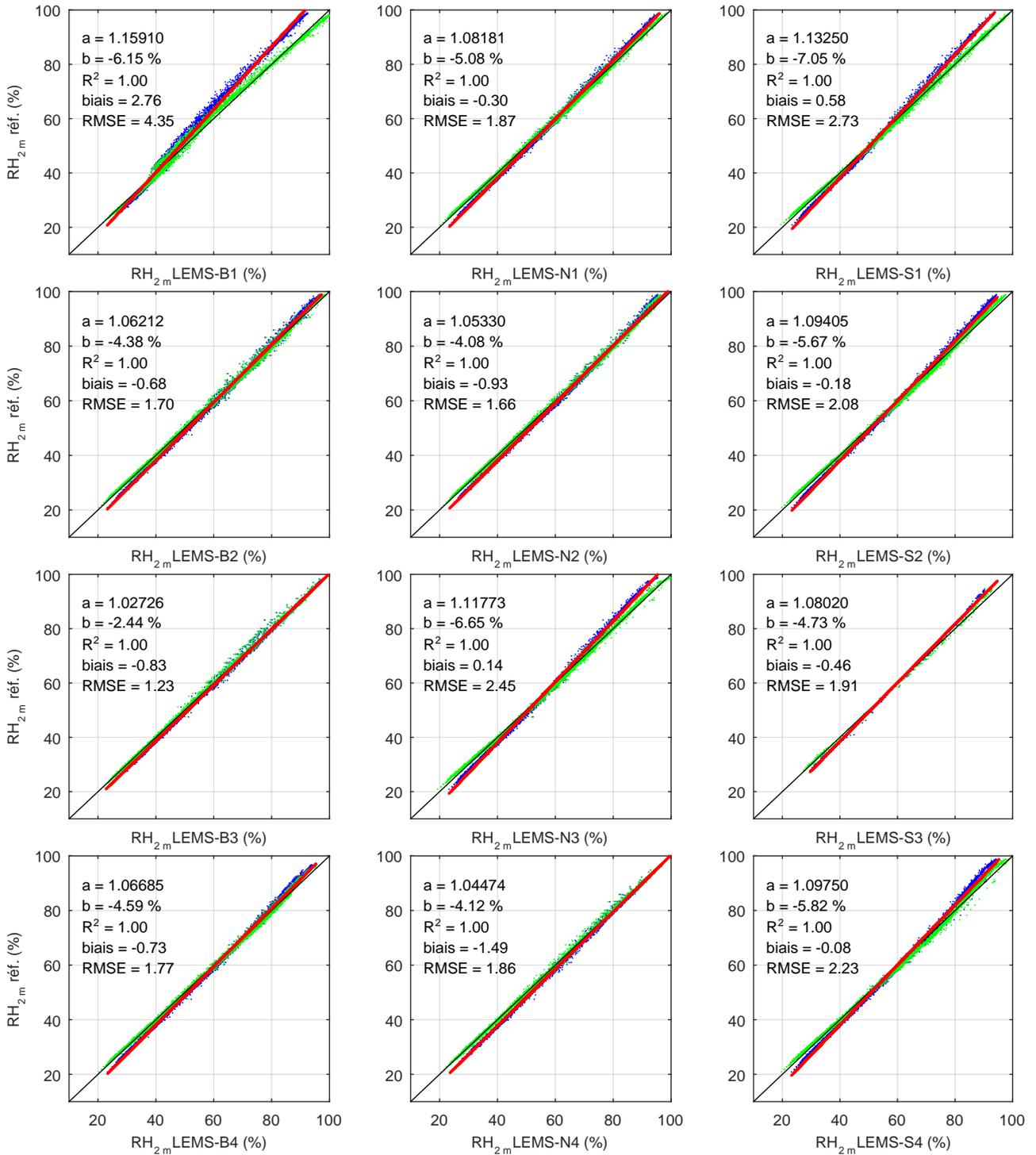


FIGURE B.4 : Comparaison de RH_{2m} mesuré à AS3 et aux LEMS (points bleus) et valeurs corrigées (points verts). Droite de la régression linéaire (en rouge) et première bissectrice (en noir).

B.3 Pression à 2 m

Les valeurs de pression de la station AS3 affichent des paliers (utilisation d'un entier en hPa par l'acquisition). Elles ne sont donc pas utilisées dans cette inter-comparaison. Les pressions pendant la période d'inter-comparaison vont de 968,5 hPa à 989,3 hPa pour la station de référence (MET01), alors que pendant la période KASCADE 2017 elles s'évaluaient de 963 à 1003,5 hPa.

La figure B.5 permet de voir un décalage net entre les pressions mesurées par les LEMS et celles mesurées à la station MET01. Les LEMS produisent des valeurs avec un biais négatif compris entre 0,77 hPa et 1,32 hPa ce qui est inférieur à la précision du capteur de la station MET01. Les pentes calculées avec la régression linéaire ainsi que les coefficients de détermination sont très proche de 1.

La pression n'est utilisée que pour le calcul des températures potentielles. Or, à une température de 0 °C une différence de 2 hPa autour de la pression 1000 hPa ne produit qu'une différence de température de 0,15 °C ce qui est faible par rapport à la précision des capteur de température des LEMS ($\pm 0,3$ à $0,5$ °C). De plus, dans l'étude, les températures des LEMS sont comparées entre elles. Or, les différences de pression entre les 12 LEMS sont beaucoup plus faibles. Enfin, l'objectif de l'expérience est l'étude des conditions stables pendant lesquelles les différences de température liées à la stratification sont beaucoup plus fortes. Pour toutes ces raisons, il n'est pas utile de corriger les données de pression des LEMS.

	a	b	R^2	Biais	RMSE	RMSE corrigé
LEMS-B1	0,994 60	4,54	1,00	-0,77	0,80	0,23
LEMS-B2	0,995 74	3,15	0,99	-1,03	1,07	0,28
LEMS-B3	0,998 23	0,42	1,00	-1,32	1,33	0,17
LEMS-B4	1,002 76	-3,83	0,99	-1,11	1,14	0,25
LEMS-N1	0,995 85	3,10	1,00	-0,98	1,00	0,19
LEMS-N2	0,996 31	2,31	1,00	-1,32	1,32	0,14
LEMS-N3	1,002 71	-3,65	0,99	-0,98	1,01	0,23
LEMS-N4	1,000 69	-1,85	1,00	-1,17	1,19	0,19
LEMS-S1	0,998 60	0,31	1,00	-1,06	1,08	0,19
LEMS-S2	0,996 35	2,49	1,00	-1,10	1,12	0,19
LEMS-S3	1,002 84	-3,63	0,99	-0,85	0,88	0,24
LEMS-S4	0,997 46	1,32	1,00	-1,18	1,21	0,24

Tableau B.4 : Résultats des comparaisons entre la pression mesurée à MET01 (référence) et aux LEMS. Les coefficients a et b sont issus de l'équation B.1 dans laquelle y , x et b sont en hPa.

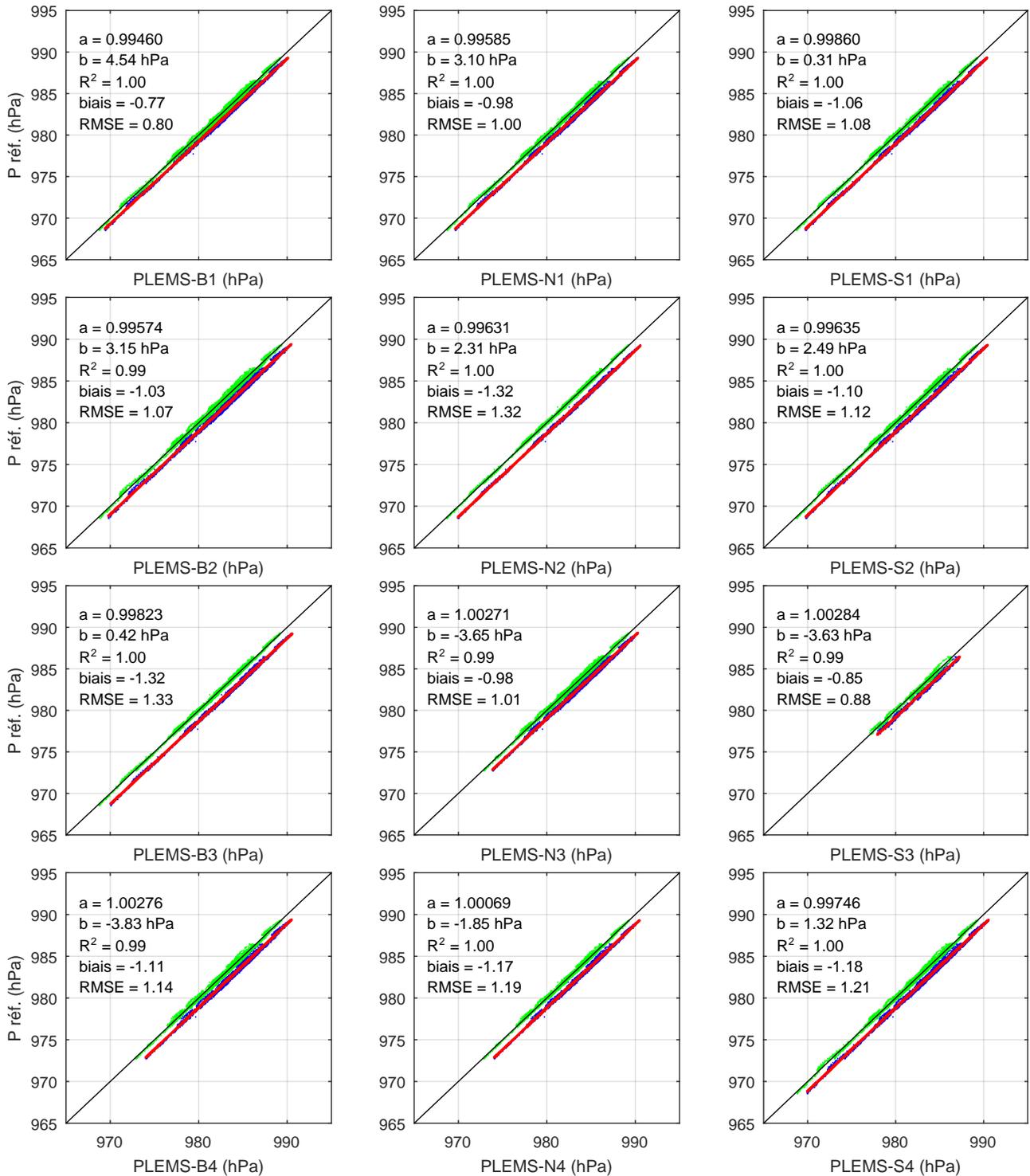


FIGURE B.5 : Comparaison entre la pression mesurée à MET01 et aux LEMS (points bleus) et valeurs corrigées (points verts). Droite de la régression linéaire (en rouge) et droite 1 : 1 (en noir).

B.4 Rayonnement SW descendant à 2 m

La comparaison des rayonnements n'est pas évidente parce que l'horizontalité des capteurs n'est pas garantie sur toutes les stations. De plus, l'orientation des capteurs a quelquefois changé au cours de l'expérience à cause de coups de vent. La comparaison des données des différents capteurs est donc simplement utile pour vérifier qu'il n'y a pas de biais important, mais elle ne peut en aucun cas permettre de corriger les données récoltées pendant l'expérience. Les régressions linéaires affichent des coefficients de détermination proche de 1 avec des pentes proches de 1 également. De plus, les biais sont relativement faibles ce qui montre que les 12 stations LEMS produisent des valeurs cohérentes avec la station MET01.

	a	b	R^2	Biais	RMSE	RMSE corrigé
LEMS-B1	1,018 25	4,80	0,99	10,28	27,33	24,56
LEMS-B2	1,008 51	4,91	0,99	7,49	34,53	33,58
LEMS-B3	1,022 52	2,80	0,99	9,61	31,94	29,50
LEMS-B4	1,051 08	1,01	1,00	16,19	30,72	20,01
LEMS-N1	1,027 85	1,62	1,00	10,03	22,15	17,41
LEMS-N2	1,066 80	-3,14	1,00	16,59	33,40	19,39
LEMS-N3	1,002 38	4,39	1,00	5,12	22,43	21,83
LEMS-N4	1,000 59	19,35	0,94	19,52	82,77	80,43
LEMS-S1	0,996 54	7,22	1,00	6,16	24,82	24,01
LEMS-S2	1,009 89	4,71	0,99	7,71	37,66	36,71
LEMS-S3	1,047 72	-1,25	1,00	12,71	26,12	16,86
LEMS-S4	1,041 63	0,13	1,00	12,55	25,61	17,55

Tableau B.5 : Résultats des comparaisons entre le rayonnement mesuré à MET01 (référence) et aux LEMS. Les coefficients a et b sont issus de l'équation B.1 dans laquelle y , x et b sont en $W m^{-2}$.

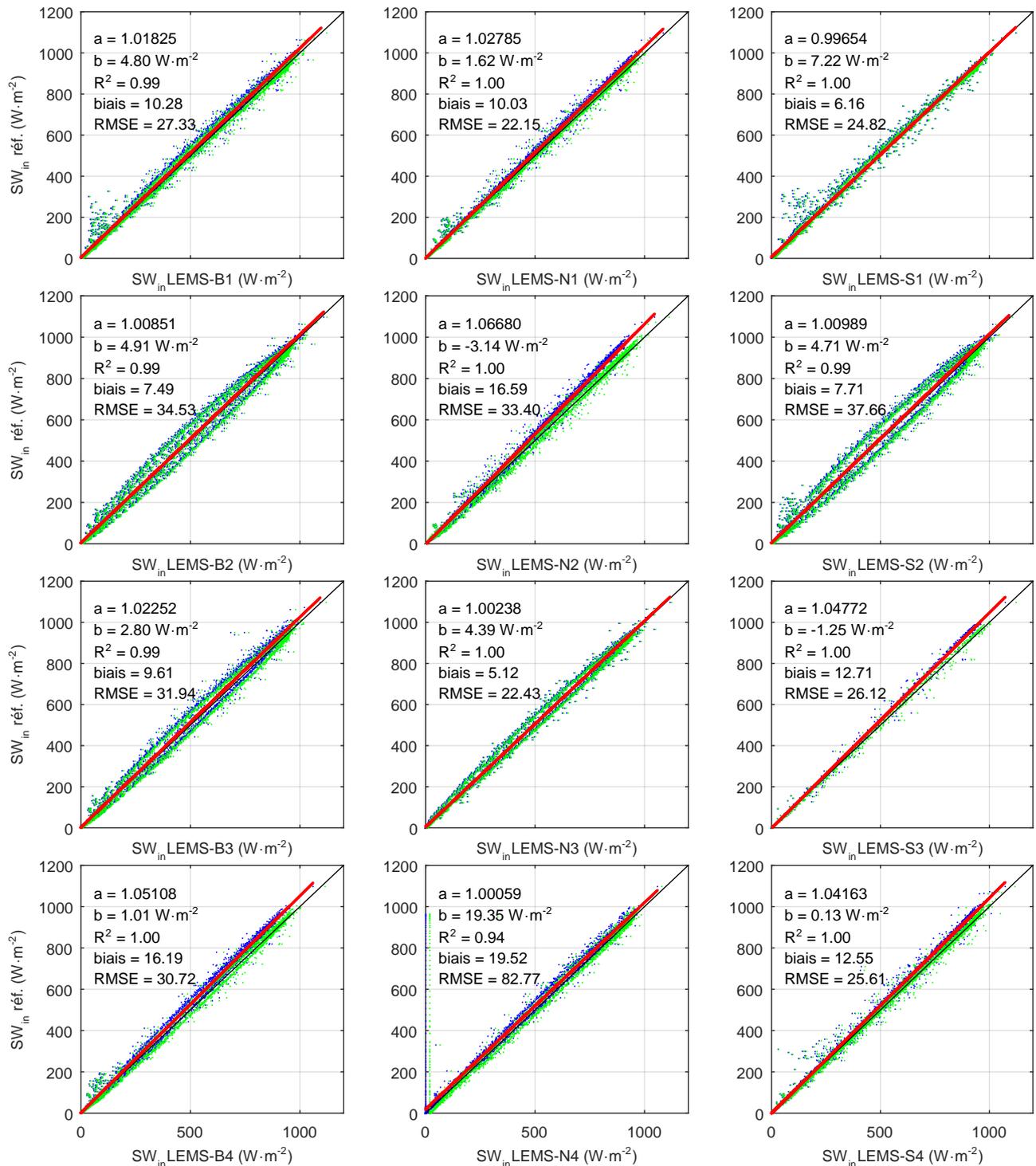


FIGURE B.6 : Comparaison entre le rayonnement mesuré à MET01 et aux LEMS (points bleus) et valeurs corrigées (points verts). Droite de la régression linéaire (en rouge) et première bissectrice (en noir).

B.5 Module du vent à 2 m

Les trois anémomètres soniques de la station AS3 étaient espacés de moins d'un mètre chacun, à environ 2 m du sol. Les anémomètres des LEMS étaient placés à environ 2 m du sol et les stations étaient espacées d'environ 3 – 4 m chacune. Il y a eu un problème d'acquisition avec les mesures de vent pour 8 des 12 LEMS pendant la période d'inter-comparaison. Seules les LEMS N1, N2, S4 et B1 ont enregistré des valeurs de vent.

La référence pour cette comparaison est calculée à partir des mesures des anémomètres soniques. Les trois anémomètres sont identiques, donc aucun n'a de raison d'être pris comme référence plus qu'un autre. La moyenne des trois anémomètres est donc prise comme référence. Chaque anémomètre sonique est ensuite comparé à la référence. La gamme de vitesses pendant la période d'inter-comparaison est de $0,02 \text{ m s}^{-1}$ à $4,95 \text{ m s}^{-1}$ pour les valeurs de référence.

Les vitesses de vent pour les anémomètres soniques, pris indépendamment, sont très cohérentes avec la moyenne des trois anémomètres (qui sert de référence). Les droites de régression sont très proches de la première bissectrice avec des coefficients de déterminations très proches de 1, les biais et erreurs quadratiques moyennes sont très faibles (tableau B.6). Les vitesses de vent de ces anémomètres n'ont donc pas besoin d'être corrigées.

	a	b	R^2	Biais	RMSE	RMSE corrigé
AS3#1	1,00	-0,01	1,00	0,00	0,03	0,03
AS3#2	1,00	0,00	1,00	0,01	0,02	0,01
AS3#3	0,99	0,00	1,00	-0,01	0,03	0,03

Tableau B.6 : Résultats des comparaisons des modules du vent à 2 m entre les anémomètres soniques de la station AS3 et la moyenne de ces 3 anémomètres. Dans l'équation B.1, y , x et b sont en m s^{-1} .

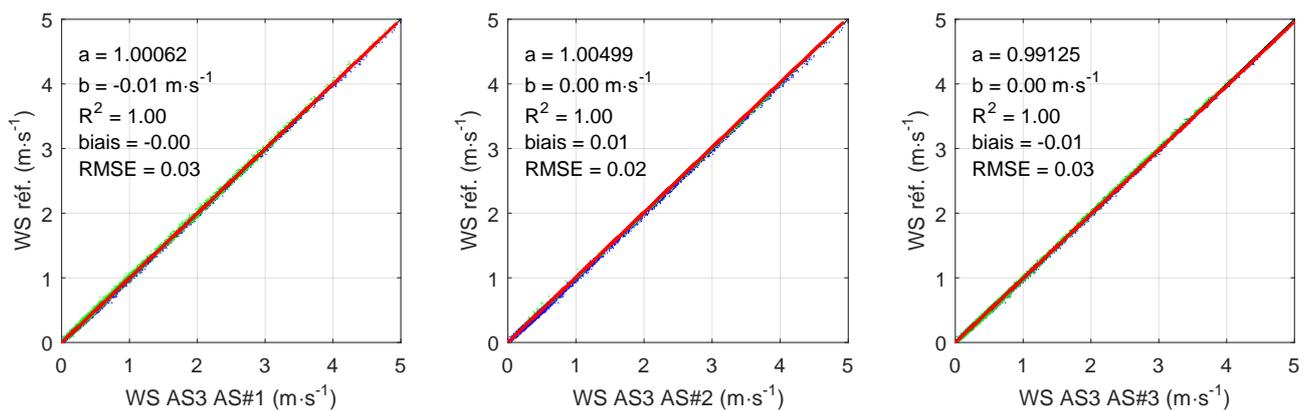


FIGURE B.7 : Comparaison du module du vent à 2 m mesuré avec les anémomètres soniques de la station AS3 (points bleus) et valeurs corrigées (points verts). De gauche à droite : 10 m, 5 m et 2 m. Droite de la régression linéaire (en rouge) et droite 1 : 1 (en noir).

Pour les LEMS, la gamme de mesure est de $0,5 \text{ m s}^{-1}$ à 89 m s^{-1} (d'après la fiche constructeur). Les vents inférieurs à $0,5 \text{ m s}^{-1}$ mesurés par les LEMS ont donc été supprimés pour cette inter-comparaison pour que ces vents faibles n'impactent pas le calcul de la régression et des coefficients de correction.

Le biais pour ces stations est compris entre $0,2 \text{ m s}^{-1}$ et $0,4 \text{ m s}^{-1}$ ce qui est meilleur que la précision annoncée par le constructeur ($\pm 1,0 \text{ m s}^{-1}$).

Pour les LEMS, les régressions polynomiales d'ordre 4 (équation B.5) donnent de meilleurs résultats que les régressions linéaires. C'est parce que la précision des mesures de vent des anémomètres à coupelle pour les vents faibles se détériore avec la diminution du module. Les corrections apportées aux vitesses de vent permettent de nettement diminuer l'erreur de mesure. Cependant, puisque seulement les données de 4 LEMS sont disponibles, il est préférable de ne pas corriger ces valeurs pour garder une cohérence entre toutes les LEMS.

$$y = a_1x^4 + a_2x^3 + a_3x^2 + a_4x + b \quad (\text{B.5})$$

	a_1	a_2	a_3	a_4	b	R^2	Biais	RMSE	RMSE corrigé
LEMS-N1	0,01	-0,09	0,47	-0,12	0,48	0,98	-0,43	0,48	0,17
LEMS-N2	0,01	-0,12	0,56	-0,08	0,48	0,99	-0,19	0,24	0,15
LEMS-S4	0,01	-0,09	0,43	0,19	0,26	0,98	-0,23	0,28	0,15
LEMS-B1	0,01	-0,08	0,39	0,24	0,23	0,98	-0,23	0,28	0,17

Tableau B.7 : Résultats des comparaisons des modules du vent à 2 m entre les LEMS et la moyenne des 3 anémomètres soniques de la station AS3. Les coefficients a_1 , a_2 , a_3 , a_4 et b sont issus de l'équation B.5 dans laquelle y , x et b sont en m s^{-1} .

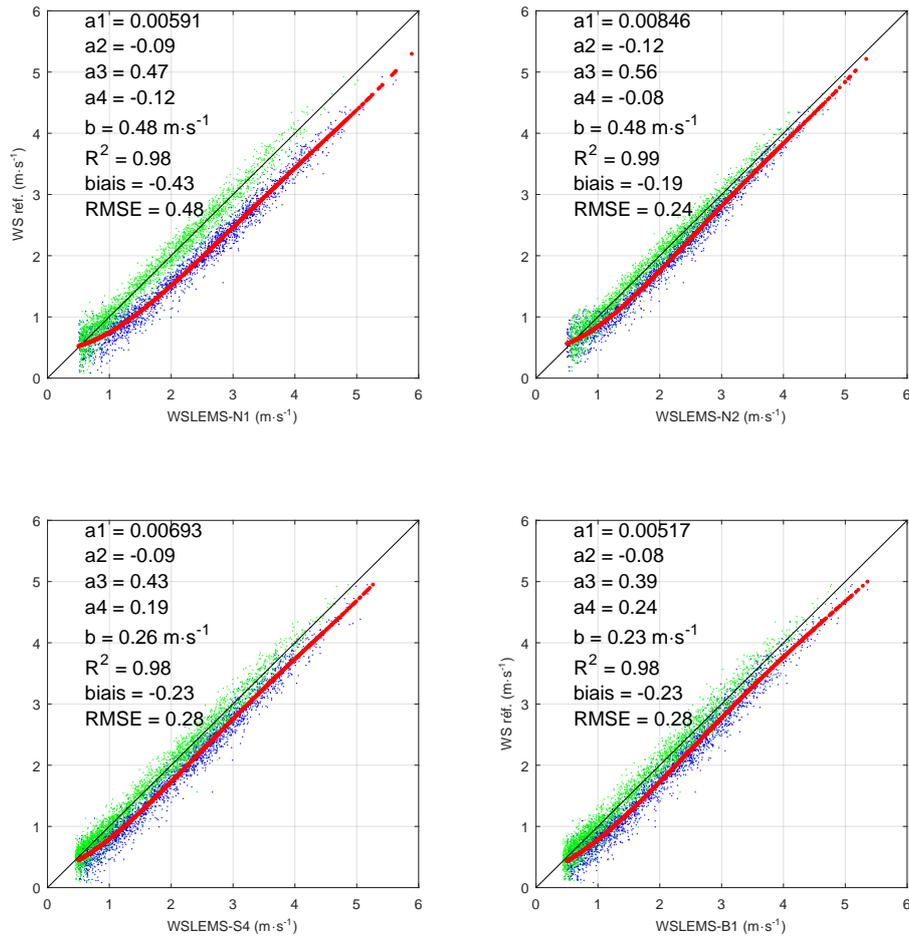


FIGURE B.8 : Comparaison du module du vent à 2m mesuré aux LEMS (points bleus) et valeurs corrigées (points verts). Droite de la régression polynomiale d'ordre 4 (en rouge) et droite 1 : 1 (en noir). Les valeurs a_n et b sont les coefficients de l'équation équation B.5.

B.6 Direction du vent à 2 m

La comparaison des directions n'a pas été effectuée sur la base du jeu de données de l'intercomparaison mais sur l'ensemble des données de la période KASCADE 2017. Il y a deux sites sur lesquels se trouvaient à la fois une station LEMS et un anémomètre sonique ce qui permet de comparer les mesures de vent, notamment pour les vents faibles. La figure B.9 permet de comparer les directions obtenues par les LEMS et les anémomètres soniques sur ces deux sites.

Tout d'abord, le biais entre les stations (qui peut être associé aux différences d'alignement des capteurs par rapport à une direction de référence) est très faible puisqu'il n'est que de $0,5^\circ$ entre les stations AS1 et B1 et de $-2,6^\circ$ entre les stations AS3 et B3 ce qui montre que l'orientation des stations est très proche sur les deux sites. En prenant pour référence les mesures des anémomètres soniques, on peut associer la différence de direction entre les stations à l'erreur de mesure commise par la LEMS. On voit alors qu'il y a une nette augmentation de l'erreur commise par les LEMS avec la diminution de la vitesse (figure B.9). La médiane de l'erreur est de 44° pour le site AS1–B1 (respectivement et 48° pour le site AS3–B3) pour les vents inférieurs à $0,25 \text{ m}\cdot\text{s}^{-1}$, mais elle diminue rapidement avec l'augmentation de la vitesse : 35° (41°) pour les vents compris entre $0,25\text{--}0,5 \text{ m}\cdot\text{s}^{-1}$, puis 23° (28°) pour les vents compris entre $0,5\text{--}0,75 \text{ m}\cdot\text{s}^{-1}$ pour finalement atteindre un palier en-dessous de 5° au-delà de $2 \text{ m}\cdot\text{s}^{-1}$.

Les données tracées sur la figure B.9 sont des moyennes sur 5 minutes. Les données moyen-

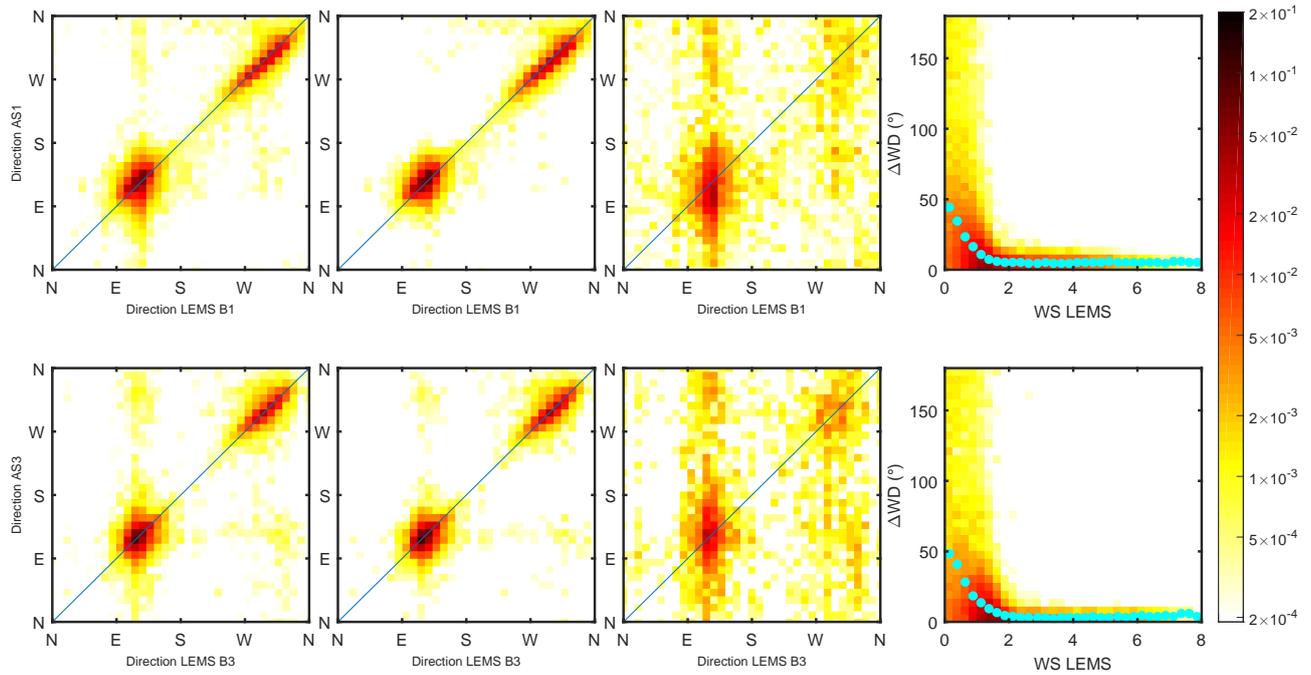


FIGURE B.9 : Comparaison des directions mesurées par la LEMS B1 (resp. B3) et l’anémomètre sonique AS1 (resp. AS3), en haut (resp. en bas). À gauche pour toutes les données, au centre gauche pour les vents (mesurés par la LEMS) supérieurs à $0,5 \text{ m s}^{-1}$ et au centre droit pour les vents inférieurs à $0,5 \text{ m s}^{-1}$. Les figures de droite montrent l’évolution de la différence de direction entre les deux types d’instrument en fonction de la vitesse mesurée sur la LEMS. Les points bleus montrent les valeurs médianes calculées pour des intervalles de vitesse de vent de $0,25 \text{ m s}^{-1}$. Toutes les données sont des moyennes sur 5 minutes et l’ensemble des données de KASCADE 2017 a été utilisé.

nées sur 30 minutes apportent une amélioration pour les vents les plus faibles : erreurs médianes de 38° et 28° pour les vents inférieurs à $0,25 \text{ m s}^{-1}$ et compris entre $0,25$ – $0,50 \text{ m s}^{-1}$ respectivement sur le site AS1–B1 et 54° et 31° sur le site AS3–B3. Cela montre donc que pour les vents faibles les directions observées avec les LEMS sont à interpréter avec précaution puisqu’il peut y avoir des erreurs importantes. Une solution pour réduire l’incertitude sur les mesures de direction des LEMS est d’utiliser des moyennes sur 30 minutes plutôt que des moyennes sur 5 minutes.

B.7 Température de surface

Les LEMS sont les seules stations à fournir une mesure de température de surface. La précision des capteurs est de $\pm 0,6^\circ\text{C}$ à une température ambiante de 25°C et pour un objet de température comprise entre 15°C et 35°C . La précision chute à $\pm 2^\circ\text{C}$ pour toute la gamme de température d’utilisation (-10 – 50°C). La gamme de températures enregistrées pendant la période d’inter-comparaison va de $6,8$ à $51,0^\circ\text{C}$ (de $-12,5$ à $54,9^\circ\text{C}$ pendant KASCADE).

Tous les capteurs ont été orientés vers le sud, pour éviter que l’ombre du mât ne gêne la mesure, montés à 50 cm au-dessus du sol. Pour l’inter-comparaison, toutes les stations étaient placées dans une « prairie » qui était constituée en juin d’herbe sèche méditerranéenne. La visualisation des données des 12 stations sur le mois de comparaison montre qu’aucune des stations n’a de comportement singulier. La valeur de référence est donc calculée en faisant la moyenne des 12 stations.

Les biais restent inférieurs à la précision du capteur ($\pm 2^\circ\text{C}$) et les coefficients de détermination, très proches de 1, montrent que les valeurs des différentes stations sont cohérentes entre elles. Les coefficients b issus des régressions linéaires affichent d'importants biais à 0°C ce qui incite à corriger ces données. Cependant, les graphes montrent que les différents capteurs ont un très bon accord pour les températures les plus faibles. Les fortes valeurs des coefficients b sont provoquées par les erreurs commises sur les températures les plus fortes. De plus, l'application de ces coefficients pour la correction de ces données, bien qu'elle supprime le biais, augmente les erreurs pour les températures faibles. Or, le but de l'expérience étant l'étude des conditions stables rencontrées pendant la nuit, il est préférable de ne pas corriger ces données pour garder un bon accord sur les températures les plus basses.

	a	b	R^2	Biais	RMSE	RMSE corrigé
LEMS-B1	0,97	0,54	1,00	-0,16	0,48	0,34
LEMS-B2	0,98	0,04	1,00	-0,37	0,48	0,24
LEMS-B3	1,04	-0,83	1,00	0,27	0,65	0,40
LEMS-B4	0,93	1,47	1,00	-0,38	1,05	0,58
LEMS-N1	1,05	-0,71	1,00	0,66	0,99	0,49
LEMS-N2	1,03	-0,10	1,00	0,76	0,89	0,29
LEMS-N3	1,05	-1,04	1,00	0,16	0,59	0,30
LEMS-N4	0,95	0,61	0,99	-0,79	1,38	0,96
LEMS-S1	1,00	-0,30	1,00	-0,34	0,45	0,30
LEMS-S2	1,02	0,29	1,00	0,79	0,89	0,34
LEMS-S3	0,95	0,80	1,00	-0,35	0,76	0,44
LEMS-S4	0,97	0,27	1,00	-0,60	0,73	0,21

Tableau B.8 : Résultats des comparaisons des températures de surface entre les LEMS et la moyenne des 12 LEMS qui sert de référence. Les coefficients a et b sont issus de l'équation B.1 dans laquelle y , x et b sont en $^\circ\text{C}$.

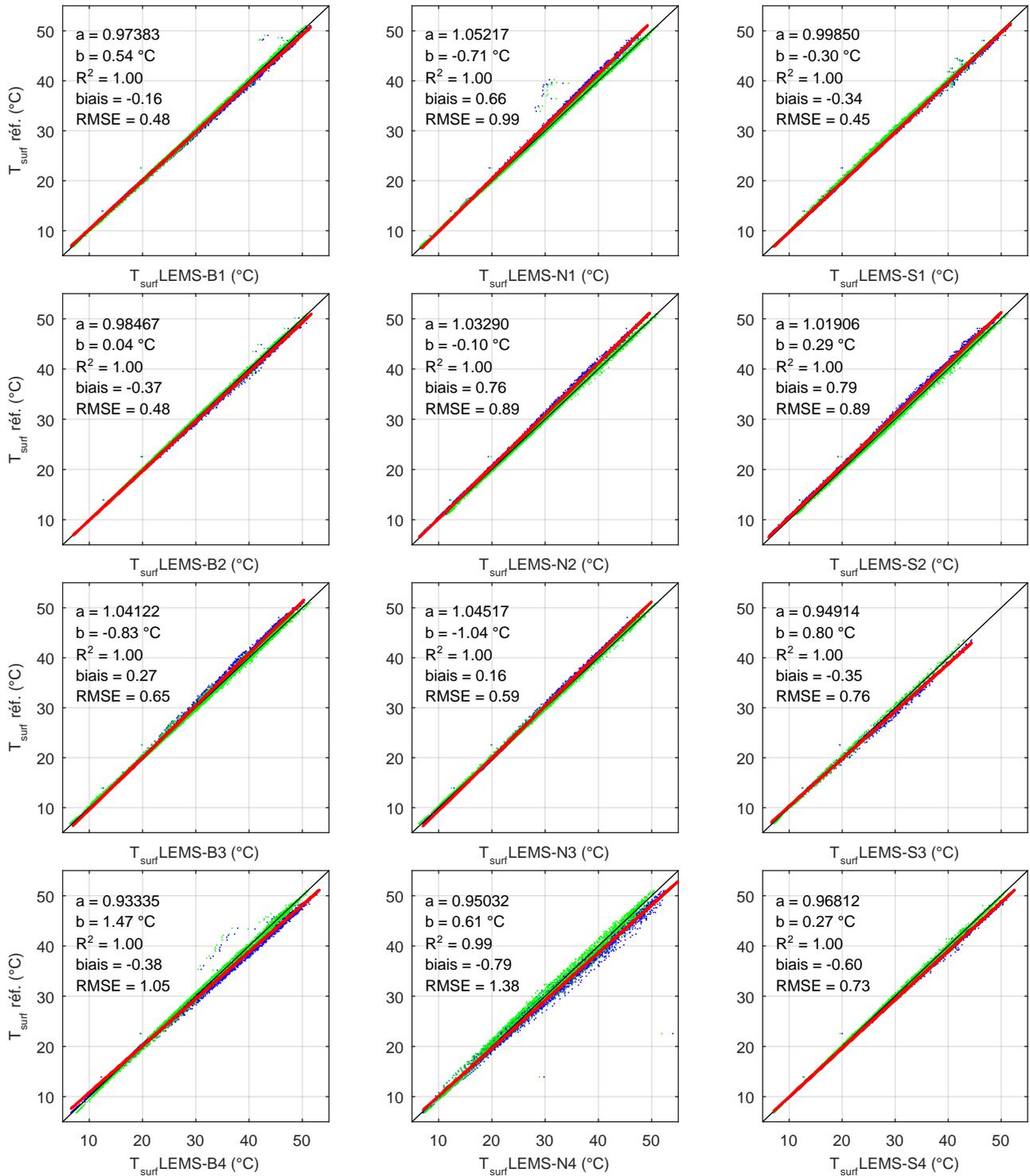


FIGURE B.10 : Comparaison de $T_{surface}$ mesuré à AS3 et aux LEMS (points bleus) et valeurs corrigées (points verts). Droite de la régression linéaire (en rouge) et première bissectrice (en noir).

B.8 Ballon captif et radiosondages

Le ballon captif et les radiosondages sont les deux seules sources d'information sur le profil vertical de température. Une comparaison entre les deux profils est donc possible sur les 50 premiers mètres au-dessus du sol. En revanche, il n'y a pas eu d'inter-comparaison entre les différents capteurs placés à la même altitude. La comparaison est donc directement effectuée par comparaison des profils obtenus lors des différentes POIs. Cela signifie donc que les comparaisons ne contiennent que des données de nuit durant lesquelles les conditions étaient stables.

Le ballon captif était placé sur le même site que la station MET01 qui possède des mesures de température à 2 et 10 m. Il est donc possible de comparer les mesures de température des deux premiers niveaux du ballon captif (environ 1 et 10 m) avec les observations de la station MET01. La précision des capteurs de la station MET01 est meilleure que celle des capteurs TinyTags utilisés sur le ballon captif ($\pm 0,1\text{ °C}$ contre $\pm 0,35\text{ °C}$) et il y a un bon accord entre les mesures de température de la station MET01 et des LEMS (partie B.1). Les mesures prises comme références sont donc celles de la station MET01.

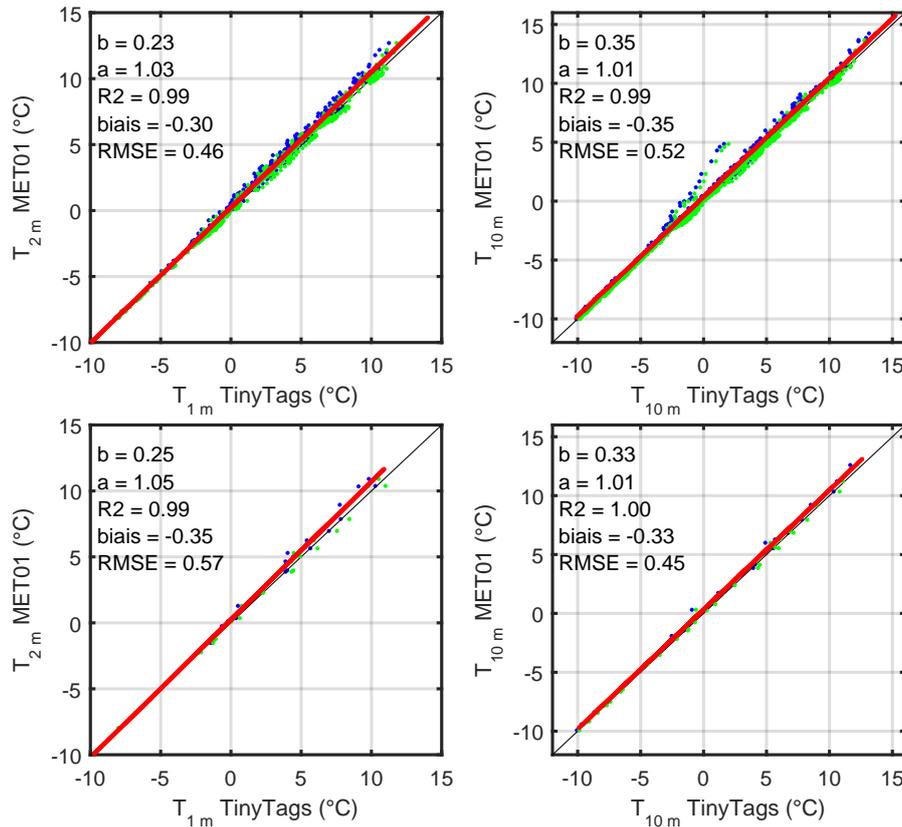


FIGURE B.11 : Comparaison entre les températures à 2 et 10 m mesurées avec le ballon captif et la station MET01 (points bleus) et valeurs corrigées (points verts). Droite de la régression linéaire (en rouge) et droite 1 : 1 (en noir). En haut pour l'ensemble des données disponibles et en bas uniquement pour les données correspondant aux heures de lâcher des radiosondages.

Il y a un bon accord entre les mesures de température des deux stations (figure B.11). Le biais est négatif à 2 m comme à 10 m, mais il reste compris dans la marge de précision des capteurs TinyTags ($\pm 0,35\text{ °C}$). De plus, l'erreur est fréquemment importante en début de nuit (environ 1 à 2 heures après le coucher du soleil) puis elle diminue rapidement ensuite, ce qui explique les écarts importants pour certaines données.

Les profils effectués avec le ballon captif sont donc validés sur les deux premiers niveaux de mesure par comparaison avec les mesures de la station MET01. Cependant, les capteurs ayant une précision assez faible, il subsiste une erreur non négligeable notamment pour le calcul d'une différence verticale de température. Bien que les coefficients de la régression linéaire calculée entre les données des deux stations permettent de corriger le biais et de réduire l'écart quadratique moyen, il est préférable de ne pas appliquer cette correction afin de garder une cohérence entre les mesures des 6 niveaux du ballon captif. Le ballon captif peut néanmoins être utilisé pour la comparaison avec les radiosondages en gardant à l'esprit que les différences peuvent être importantes par rapport aux températures enregistrées sur la station MET01. On admet d'étendre aux niveaux supérieurs la confiance obtenue sur les deux premiers niveaux, validant ainsi les données observées avec le ballon captif.

La comparaison des profils de température potentielle du ballon captif avec ceux des radiosondages permet donc d'évaluer ces derniers sur les 50 premiers mètres au-dessus du sol (figures B.13 et B.14). Les radiosondages affichent fréquemment des profils neutres ou instables près du sol au contraire du ballon captif et de la station MET01. Il y a donc un biais positif important près du sol, et le biais minimal entre le ballon captif et les radiosondages intervient entre 20 et 30 m de hauteur (biais négatif de $-0,16\text{ }^\circ\text{C}$ à 30 m, cf. figure B.12). De plus, bien que les capteurs du ballon captif aient aussi des erreurs de mesure non négligeables, elles sont bien moindres comparées à celles commises par le radiosondage. Il est donc préférable de ne pas prendre en compte les mesures de température des radiosondages sur les 30 premiers mètres au-dessus du sol.

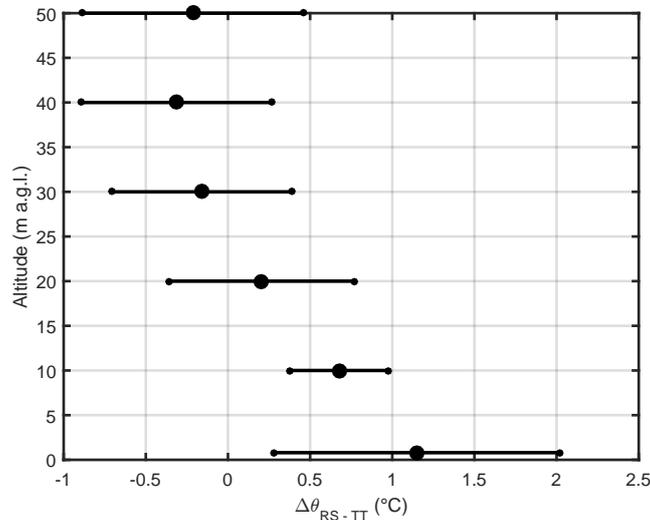


FIGURE B.12 : Biais et écart-type autour de la moyenne des différences de température potentielle entre les mesures effectuées par les radiosondages et par les TinyTags du ballon captif.

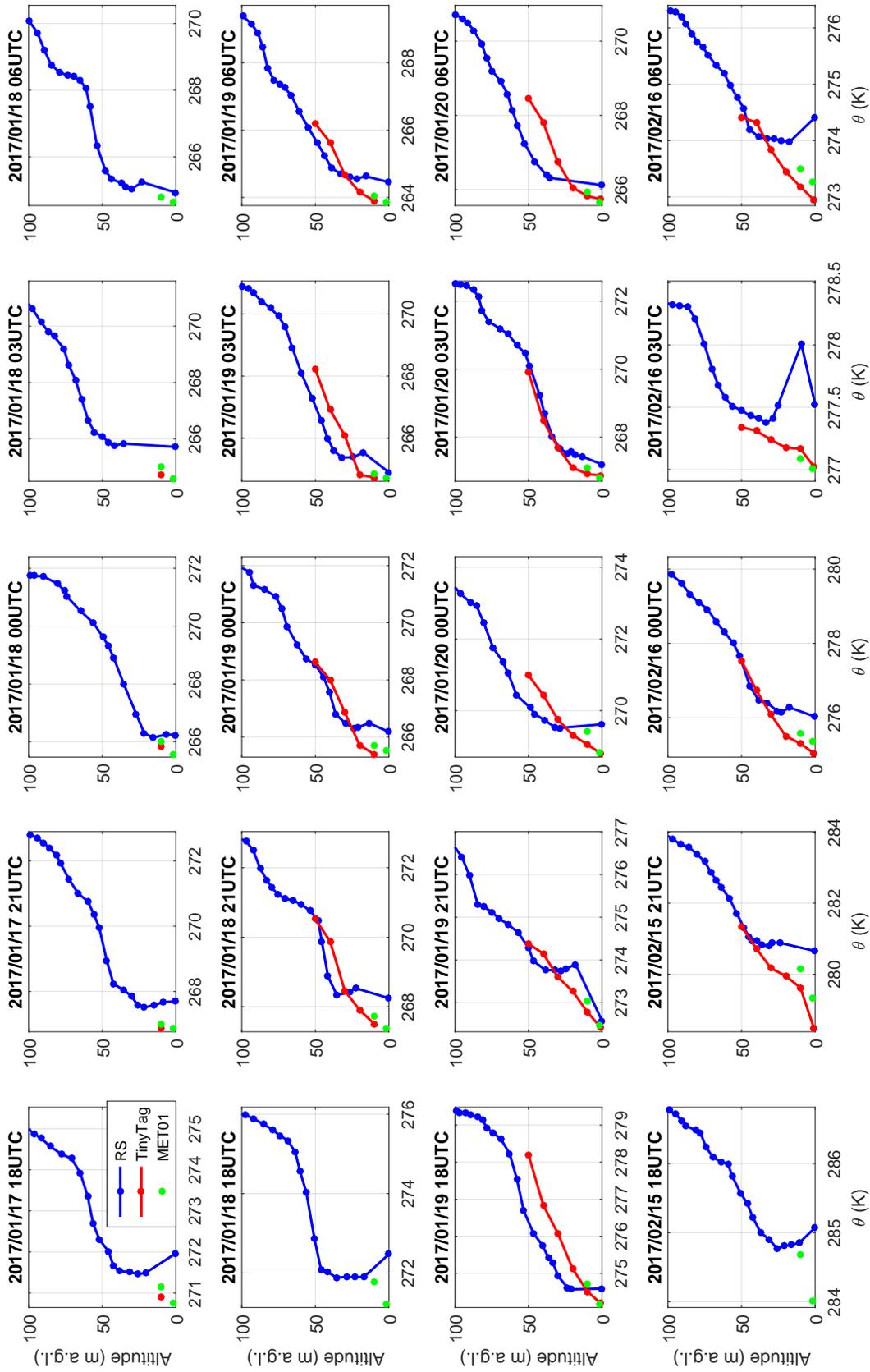


FIGURE B.13 : Comparaison des profils de température potentielle mesurés avec le ballon captif (en rouge),t avec les radiosondages (en bleu) et avec la station MET01 (en vert).

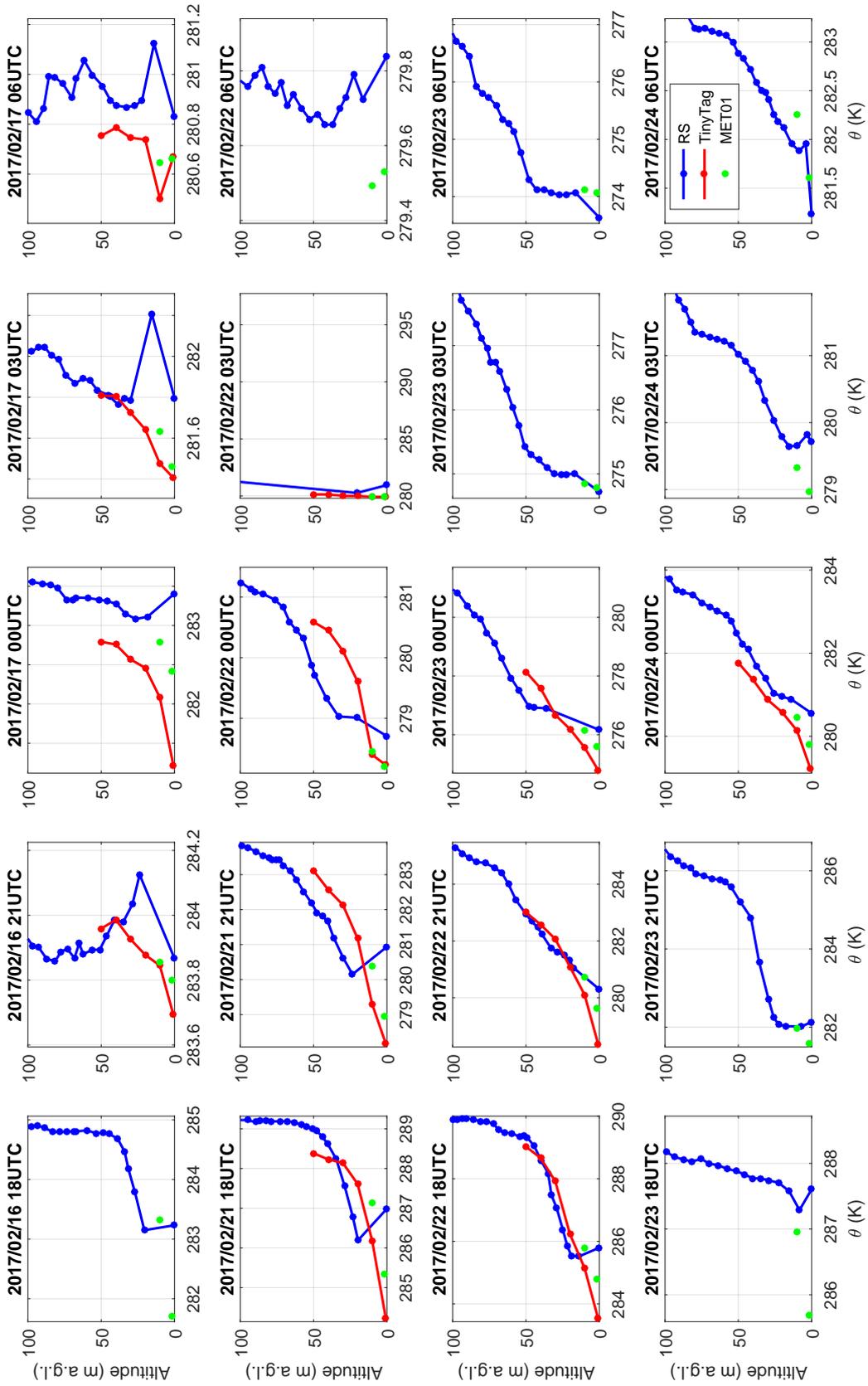


FIGURE B.14 : Comparaison des profils de température potentielle mesurés avec le ballon captif (en rouge),^t avec les radiosondages (en bleu) et avec la station MET01 (en vert).

Annexe C

Composition des jeux de données d'entrée pour les RNA

La qualité des performances d'un RNA est dépendante de la composition de son jeu de données d'entrée. Ainsi, les variables en entrée du RNA doivent être choisies méthodiquement et le jeu de données doit être suffisamment grand afin d'obtenir des résultats satisfaisants. Cette annexe présente le processus de sélection des variables d'entrée des RNA utilisés au cours de la thèse et dont les résultats sont détaillés dans les chapitres 4 et 5. Ensuite, la longueur des jeux de données utilisés ne pouvant pas être étendue, des tests sont effectués afin d'évaluer l'impact de la taille du jeu de données sur la performance du RNA.

C.1 Sélection des variables

C.1.1 Observations sur la période de un an

Les résultats de la sélection des variables pour le RNA utilisant des observations en entrées (cf. chapitre 4) sont présentés dans cette partie. La période considérée s'étend sur un an, du 17 février 2015 au 17 février 2016 et les variables testées proviennent des observations réalisées à la station GBA (voir le tableau 2.1). Il y a ainsi 9 variables testées comme potentielles entrées : la vitesse du vent à 110 m U_{110m} , la température à 2 m T_{2m} , la température à 110 m T_{110m} , la différence de température potentielle entre 110 m et 2 m $\Delta\theta$, l'humidité relative à 2 m RH , la pression calculée au niveau de la mer SLP , l'humidité spécifique à 2 m Q_{2m} , et les composantes du vent à 110 m u_{110m} et v_{110m} .

La sélection des variables d'entrée a été effectuée en deux étapes en suivant la méthodologie décrite par [Dreyfus et al. \(2002\)](#).

C.1.1.a Première étape – relations inter-variables

La première étape consiste à supprimer les variables qui sont fortement corrélées (le seuil de 0,9 pour le coefficient de corrélation a été retenu). Le tableau C.1 contient les valeurs des coefficients de corrélation calculés entre les neuf variables testées sur la période de un an.

Une corrélation importante existe entre les températures mesurées à 2 m et 110 m ($R = 0,97$) ce qui signifie qu'une de ces deux variables peut être supprimée. De plus, ces deux variables servent aussi au calcul de la différence de température potentielle entre 110 m et 2 m ($\Delta\theta$) qui est une autre des variables testées. L'évolution au cours du cycle diurne de la stabilité est pilotée par la variation de la température de surface (sous l'effet des échanges radiatifs), ce qui explique que le lien entre T_{2m} et $\Delta\theta$, bien que modéré ($R = 0,57$), soit plus important que le

	U_{110m}	T_{2m}	T_{110m}	$\Delta\theta$	RH	SLP	Q_{2m}	u_{110m}	v_{110m}
U_{110m}	–	0,24	0,15	–0,38	–0,44	–0,29	–0,05	0,47	–0,20
T_{2m}	–	–	0,97	–0,57	–0,67	–0,33	0,74	0,19	0,20
T_{110m}	–	–	–	–0,34	–0,61	–0,31	0,75	0,16	0,20
$\Delta\theta$	–	–	–	–	0,50	0,24	–0,30	–0,19	–0,10
RH	–	–	–	–	–	0,22	–0,03	–0,47	0,01
SLP	–	–	–	–	–	–	–0,29	–0,20	–0,01
Q_{2m}	–	–	–	–	–	–	–	–0,15	0,27
u_{110m}	–	–	–	–	–	–	–	–	–0,44
v_{110m}	–	–	–	–	–	–	–	–	–

Tableau C.1 : Coefficients de corrélation entre les différentes variables choisies comme entrées sur la période de 1 an. Les cellules rouges montrent les valeurs supérieures à 0,9 et les jaunes montrent les valeurs supérieures à 0,7.

lien entre T_{110m} et $\Delta\theta$ ($R = 0,34$). Pour cette raison, T_{2m} est supprimée lors de cette première étape de sélection.

C.1.1.b Seconde étape – choix des variables les plus pertinentes

Méthodologie

La deuxième étape de sélection consiste à sélectionner les variables les plus pertinentes pour le calcul du RNA. Tout d'abord, un RNA « étalon » est construit en utilisant comme entrées les N variables encore disponibles après la première étape de sélection. À partir de là, deux manières de procéder sont possibles (Dreyfus et al., 2002).

La première est une méthode d'élimination. Elle consiste à créer de nouveaux jeux de données à $N - 1$ entrées en retirant chaque variable à tour de rôle, ce qui représente un total de N jeux de données. Des RNA sont lancés en utilisant ces jeux de données et leurs performances sont comparées à celles du RNA « étalon ». Si les performances en utilisant un jeu de données réduit ne sont pas dégradées, cela signifie que la variable retirée n'est pas significative et peut donc être retirée. L'opération est ensuite répétée avec des jeux de données à $N - 2$ variables jusqu'à arriver à une combinaison de variables d'entrées qu'il n'est plus possible de réduire sans impacter significativement les performances du calcul du RNA.

La seconde est une méthode additive. Elle consiste à construire le jeu de données d'entrée optimal en ajoutant progressivement les variables les plus pertinentes. Il faut commencer avec les jeux de données les plus simples, c'est à dire avec une seule variable. Cela représente N jeux de données à 1 variable. Les performances des RNA utilisant ces jeux de données sont comparées entre elles, et la variable produisant les meilleurs résultats est sélectionnée. Un nouveau groupe de $N - 1$ jeux de données à 2 entrées est créé, chacun étant composé de la première variable sélectionnée et d'une autre variable à tester. Le jeu de données est ainsi construit pas à pas jusqu'à ce que les performances du RNA deviennent comparables à celles du RNA « étalon ».

Les deux méthodes ont été testées. Au final, la seconde méthode (construction du jeu de données optimal) a été utilisée parce qu'on a constaté que les différences entre les N RNA utilisant les jeux de données à $N - 1$ variables (méthode d'élimination) sont faibles, ce qui rend la sélection trop délicate.

La significativité des résultats est évaluée en utilisant la méthode du bootstrap. Cette méthode permet d'évaluer la robustesse de la fonction f_{obs} produite par le RNA. Elle consiste à créer, à partir du jeu de test, des sous-échantillons de même taille que celui-ci par tirage aléatoire avec remise. Un grand nombre de sous-échantillons sont créés (10000 dans notre cas) et la fonction f_{obs} produite par le RNA leur est appliquée. Les performances ($DACC$, PC , ...) produites par la fonction f_{obs} sur chacun de ces sous-échantillons sont calculées ce qui permet de déterminer un intervalle de confiance pour chaque indicateur (Wilks, 2011).

Résultats

Jeux de données à 1 variable

Les résultats des RNA utilisant une seule variable en entrée sont présentés sur la figure C.1, sur laquelle sont également représentés les résultats du RNA « étalon », c'est-à-dire le RNA qui utilise comme entrées les 8 variables restantes après la première étape de sélection. Comme on pouvait s'y attendre, les performances des RNA avec une seule variable sont nettement inférieures à celles du RNA étalon. D'autre part, il n'y a pas une variable qui produise les meilleurs résultats pour tous les indicateurs statistiques.

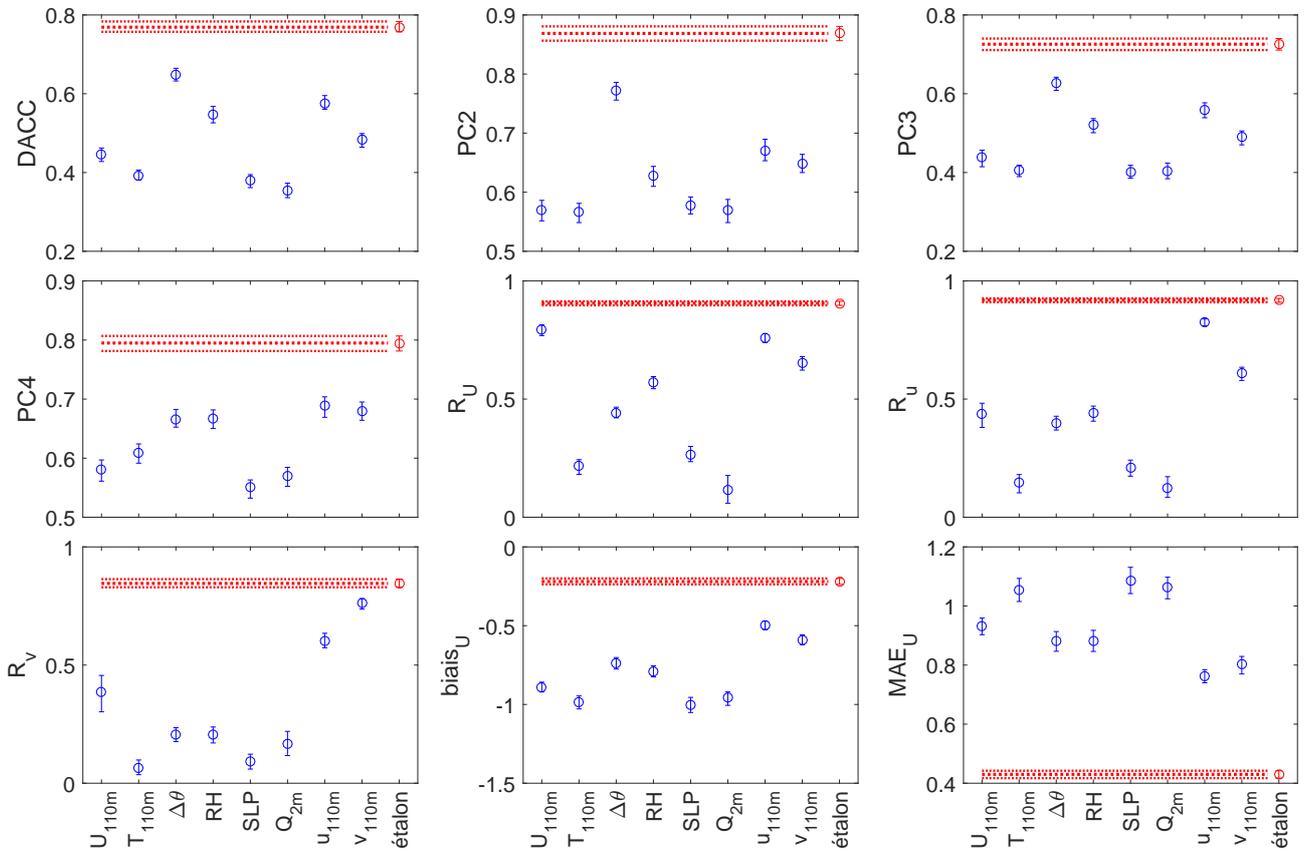


FIGURE C.1 : Performances pour les RNA qui utilisent en entrée des jeux de données constitués d'une seule variable. Le nom de la variable sur l'abscisse indique la variable utilisée. Les performances du RNA « étalon » sont en rouge. Les barres représentent l'intervalle de confiance à 90% calculé avec la méthode du bootstrap. Les lignes pointillées rouges indiquent la valeur moyenne et l'intervalle de confiance du RNA étalon.

La variable $\Delta\theta$ produit les meilleurs résultats pour le *DACC*, le *PC2* et le *PC3*. La seule variable u permet de bien représenter la composante u du vent dans la vallée de Cadarache ($R = 0,82$). De même pour la variable v qui permet de bien reproduire la composante v du vent observé dans la vallée de Cadarache ($R = 0,77$). C'est assez logique puisque les composantes à 2 m et 110 m sont corrélées ($R = 0,79$ pour u et $R = 0,72$ pour v). Les composantes à 110 m produisent également les meilleurs résultats du point de vue du module (biais, MAE).

Les critères de direction (*DACC* et *PC*) sont les critères prépondérants pour le choix des variables. Ainsi, pour cette première itération, c'est la différence de température entre 110 m et 2 m qui a été sélectionnée.

Jeux de données à 2 variables

Sur la deuxième itération, dont les résultats sont présentés sur la figure C.2, la combinaison de la variable $\Delta\theta$ avec les variables $u_{110\text{m}}$ ou $v_{110\text{m}}$ produit des résultats bien meilleurs que n'importe quelle autre combinaison de variables, et ce pour tous les indicateurs statistiques. Les performances de la composante u étant généralement plus élevées que celles de la composante v , c'est donc cette variable qui est sélectionnée.

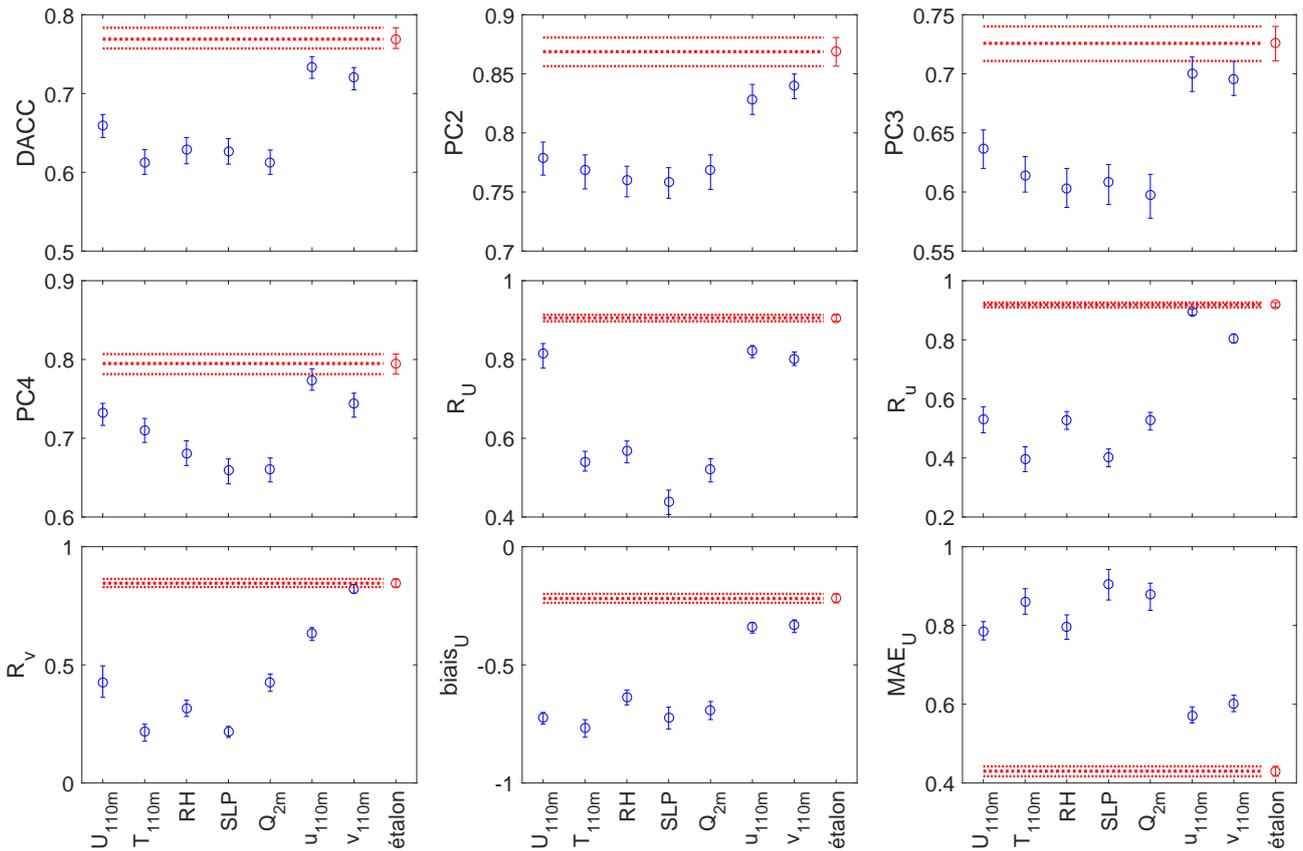


FIGURE C.2 : Idem figure C.1 pour des jeux de données d'entrée constitués de $\Delta\theta$ et d'une seconde variable dont le nom est indiqué sur l'abscisse.

Jeux de données à 3 variables

Sur la troisième itération, dont les résultats sont présentés sur la figure C.3, l'ajout de la variable $v_{110\text{m}}$ produit de meilleurs résultats que n'importe quelle autre combinaison de variables, et ce pour tous les indicateurs statistiques (sauf le *PC4*). C'est donc cette variable qui

est sélectionnée. On notera que la vitesse à 110 m produit aussi de bons résultats.

De plus, le RNA à 3 variables d'entrée ($\Delta\theta$, u_{110m} et v_{110m}) produit des résultats qui sont proches de ceux produits par le RNA à 8 variables, comme on peut le constater sur la figure C.3. Il n'est donc pas nécessaire d'ajouter de nouvelles variables ce qui signifie que la seconde étape de sélection est terminée.

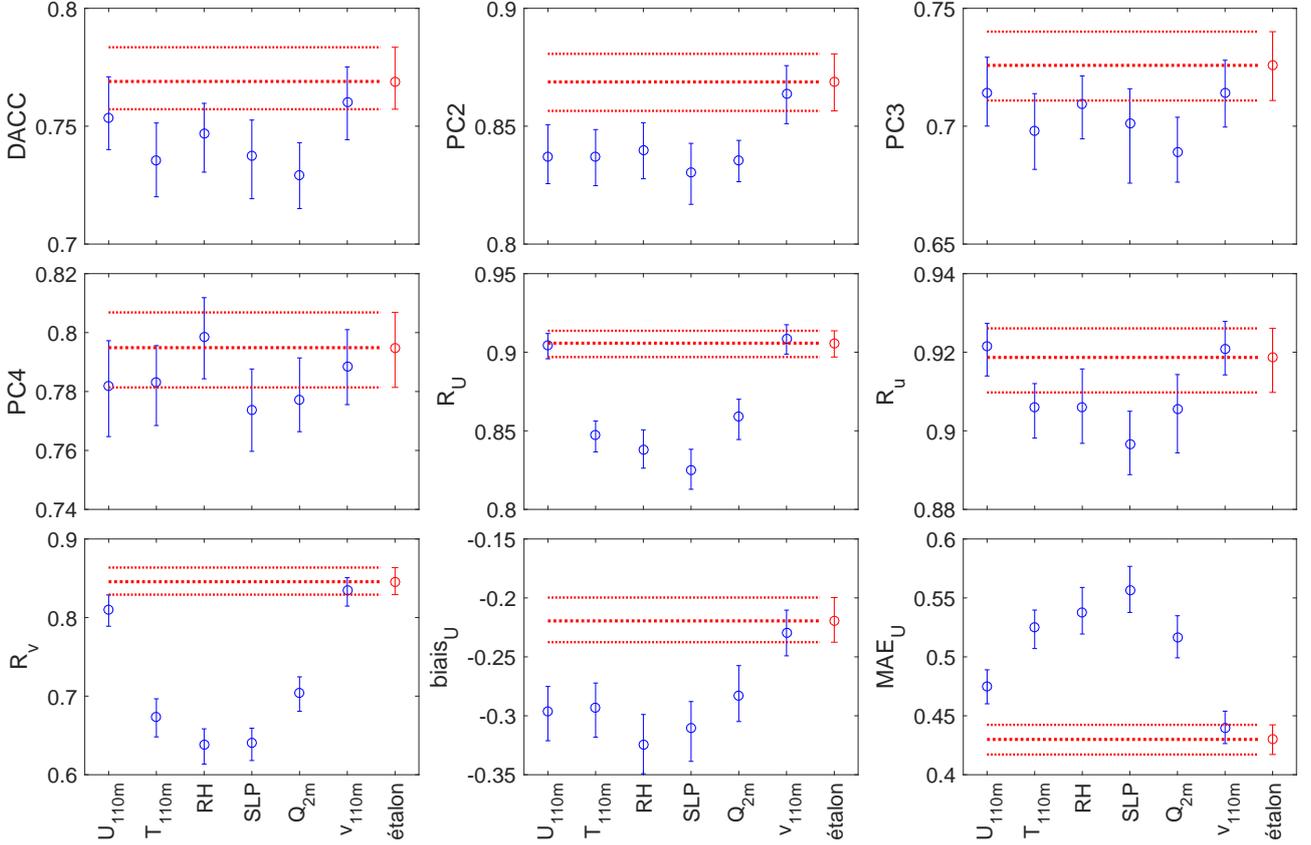


FIGURE C.3 : Idem figure C.1 pour des jeux de données d'entrée constitués de $\Delta\theta$ et u_{110m} et d'une troisième variable dont le nom est indiqué sur l'abscisse.

C.1.2 Données WRF

Les résultats de la sélection des variables pour le RNA utilisant des variables WRF en entrée (cf. chapitre 5) sont présentés dans cette partie. La période considérée s'étend sur un an, du 17 février 2015 au 17 février 2016 et les variables testées proviennent des simulations WRF (voir le tableau 5.1). La sélection des variables d'entrée a été effectuée en suivant la même méthodologie en deux étapes que dans la partie C.1.1.

C.1.2.a Première étape – relations inter-variables

L'argumentaire concernant le choix des variables sur la base des relations inter-variables ayant déjà été présenté dans le chapitre 5, seuls les tableaux contenant les coefficients de corrélation sont présentés dans cette annexe (tableaux C.2, C.3 et C.4). Les données sont présentées sur trois tableaux pour des raisons pratiques vu le grand nombre de variables testées.

Le tableau C.4 regroupe les coefficients de corrélation entre les variables qui sont testées dans la deuxième étape de sélection.

	$Q_{2\text{ m}}$	$P_{2\text{ m}}$	$\theta_{2\text{ m}}$	$\theta_{110\text{ m}}$	$\Delta\theta$	T_{surf}	Rn	Ri	$Pluie$	u_*	$PBLH$	$Sunrise$	$Sunset$
$Q_{2\text{ m}}$	—	-0,186	0,745	0,796	-0,087	0,691	0,215	-0,006	0,019	0,099	0,137	0,222	-0,233
$P_{2\text{ m}}$	-0,186	—	-0,226	-0,183	0,239	-0,169	-0,097	0,038	-0,175	-0,265	-0,268	-0,136	0,091
$\theta_{2\text{ m}}$	0,745	-0,226	—	0,970	-0,478	0,989	0,683	-0,033	-0,089	0,506	0,587	0,553	-0,547
$\theta_{110\text{ m}}$	0,796	-0,183	0,970	—	-0,250	0,935	0,546	-0,016	-0,109	0,345	0,441	0,445	-0,404
$\Delta\theta$	-0,087	0,239	-0,478	-0,250	—	-0,559	-0,749	0,075	-0,039	-0,771	-0,742	-0,595	0,717
T_{surf}	0,691	-0,169	0,989	0,935	-0,559	—	0,758	-0,034	-0,092	0,556	0,630	0,581	-0,620
Rn	0,215	-0,097	0,683	0,546	-0,749	0,758	—	-0,040	-0,039	0,692	0,748	0,581	-0,693
Ri	-0,006	0,038	-0,033	-0,016	0,075	-0,034	-0,040	—	0,001	-0,065	-0,067	-0,050	0,035
$Pluie$	0,019	-0,175	-0,089	-0,109	-0,039	-0,092	-0,039	0,001	—	-0,036	-0,035	-0,021	0,017
u_*	0,099	-0,265	0,506	0,345	-0,771	0,556	0,692	-0,065	-0,036	—	0,887	0,527	-0,507
$PBLH$	0,137	-0,268	0,587	0,441	-0,742	0,630	0,748	-0,067	-0,035	0,887	—	0,559	-0,497
$Sunrise$	0,222	-0,136	0,553	0,445	-0,595	0,581	0,581	-0,050	-0,021	0,527	0,559	—	-0,501
$Sunset$	-0,233	0,091	-0,547	-0,404	0,717	-0,620	-0,693	0,035	0,017	-0,507	-0,497	-0,501	—

Tableau C.2 : Coefficients de corrélation entre les variables choisies comme potentielles entrées du RNA (les variables relatives au vent sont rajoutées par la suite). Les cellules rouges montrent les valeurs supérieures à 0,9 et les jaunes montrent les valeurs supérieures à 0,7.

	Q_2 m	P_2 m	θ_{110} m	$\Delta\theta$	R_n	R_i	P_{luie}	u_*	$PBLH$	$Sunrise$	$Sunset$	U_{10} m	u_{10} m	v_{10} m	U_{110} m	u_{110} m	v_{110} m	U_{300} m	u_{300} m	v_{300} m
Q_2 m	—	-0,186	0,796	-0,087	0,215	-0,006	0,019	0,099	0,137	0,222	-0,233	0,062	-0,126	0,282	0,039	-0,092	0,321	-0,058	-0,060	0,305
P_2 m	-0,186	—	-0,183	0,239	-0,097	0,038	-0,175	-0,265	-0,268	-0,136	0,091	-0,287	-0,163	0,010	-0,278	-0,170	0,002	-0,267	-0,188	0,102
θ_{110} m	0,796	-0,183	—	-0,250	0,546	-0,016	-0,109	0,345	0,441	0,445	-0,404	0,239	0,071	0,221	0,094	0,082	0,260	-0,079	0,090	0,226
$\Delta\theta$	-0,087	0,239	-0,250	—	-0,749	0,075	-0,039	-0,771	-0,742	-0,595	0,717	-0,591	-0,266	0,049	-0,357	-0,232	-0,166	-0,227	-0,226	-0,090
R_n	0,215	-0,097	0,546	-0,749	—	-0,040	-0,039	0,692	0,748	0,581	-0,693	0,449	0,178	0,043	0,117	0,126	0,190	-0,060	0,146	0,200
R_i	-0,006	0,038	-0,016	0,075	-0,040	—	0,001	-0,065	-0,067	-0,050	0,035	-0,062	-0,035	0,010	-0,097	-0,022	0,001	-0,074	-0,014	0,009
P_{luie}	0,019	-0,175	-0,109	-0,039	-0,039	0,001	—	-0,036	-0,035	-0,021	0,017	-0,026	-0,056	-0,016	-0,004	-0,057	0,004	0,033	-0,055	0,033
u_*	0,099	-0,265	0,345	-0,771	0,692	-0,065	-0,036	—	0,887	0,527	-0,507	0,939	0,367	-0,006	0,734	0,338	0,123	0,529	0,303	-0,017
$PBLH$	0,137	-0,268	0,441	-0,742	0,748	-0,067	-0,035	0,887	—	0,559	-0,497	0,783	0,297	0,060	0,561	0,258	0,186	0,392	0,226	0,074
$Sunrise$	0,222	-0,136	0,445	-0,595	0,581	-0,050	-0,050	—	0,559	—	-0,501	0,419	0,193	0,043	0,229	0,178	0,181	0,047	0,174	0,163
$Sunset$	-0,233	0,091	-0,404	0,717	-0,693	0,035	0,017	-0,507	-0,497	-0,501	—	-0,299	-0,092	0,017	-0,070	-0,083	-0,171	0,039	-0,085	-0,170
U_{10} m	0,062	-0,287	0,239	-0,591	0,449	-0,062	-0,026	0,939	0,783	0,419	-0,299	—	0,386	-0,005	0,873	0,371	0,086	0,673	0,322	-0,082
u_{10} m	-0,126	-0,163	0,071	-0,266	0,178	-0,035	-0,056	0,367	0,297	0,193	-0,092	0,386	—	-0,435	0,296	0,931	-0,453	0,177	0,830	-0,548
v_{10} m	0,282	0,010	0,221	0,049	0,443	0,010	-0,016	-0,006	0,060	0,043	0,017	-0,005	-0,435	—	-0,009	-0,333	0,830	-0,066	-0,258	0,594
U_{110} m	0,039	-0,278	0,094	-0,357	0,117	-0,097	-0,004	0,734	0,561	0,229	-0,070	0,873	0,296	-0,009	—	0,291	0,075	0,832	0,228	-0,095
u_{110} m	-0,092	-0,170	0,082	-0,232	0,126	-0,022	-0,057	0,338	0,258	0,178	-0,083	0,371	0,931	-0,333	0,291	—	-0,402	0,154	0,926	-0,556
v_{110} m	0,321	0,002	0,260	-0,166	0,190	0,001	0,004	0,123	0,186	0,181	-0,171	0,086	-0,453	0,830	0,075	-0,402	—	0,038	-0,345	0,846
U_{300} m	-0,058	-0,267	-0,079	-0,227	-0,060	-0,074	0,033	0,529	0,392	0,047	0,039	0,673	0,177	-0,066	0,832	0,154	0,038	—	0,037	-0,123
u_{300} m	-0,060	-0,188	0,090	-0,226	0,146	-0,014	-0,055	0,303	0,226	0,174	-0,085	0,322	0,830	-0,258	0,228	0,926	-0,345	0,037	—	-0,504
v_{300} m	0,305	0,102	0,226	-0,090	0,200	0,009	0,033	-0,017	0,074	0,163	-0,170	-0,082	-0,548	0,594	-0,095	-0,556	0,846	-0,123	-0,504	—

Tableau C.3 : Idem figure C.2 après ajout des variables relatives au vent et retrait des variables redondantes.

	Q_2 m	P_2 m	θ_{110} m	$\Delta\theta$	R_n	R_i	$Pluie$	$PBLH$	$Sunrise$	$Sunset$	U_{10} m	u_{10} m	v_{10} m	U_{300} m	u_{300} m	v_{300} m
Q_2 m	—	-0,186	0,796	-0,087	0,215	-0,006	0,019	0,137	0,222	-0,233	0,062	-0,126	0,282	-0,058	-0,060	0,305
P_2 m	-0,186	—	-0,183	0,239	-0,097	0,038	-0,175	-0,268	-0,136	0,091	-0,287	-0,163	0,010	-0,267	-0,188	0,102
θ_{110} m	0,796	-0,183	—	-0,250	0,546	-0,016	-0,109	0,441	0,445	-0,404	0,239	0,071	0,221	-0,079	0,090	0,226
$\Delta\theta$	-0,087	0,239	-0,250	—	-0,749	0,075	-0,039	-0,742	-0,595	0,717	-0,591	-0,266	0,049	-0,227	-0,226	-0,090
R_n	0,215	-0,097	0,546	-0,749	—	-0,040	-0,039	0,748	0,581	-0,693	0,449	0,178	0,043	-0,060	0,146	0,200
R_i	-0,006	0,038	-0,016	0,075	-0,040	—	0,001	-0,067	-0,050	0,035	-0,062	-0,035	0,010	-0,074	-0,014	0,009
$Pluie$	0,019	-0,175	-0,109	-0,039	-0,039	0,001	—	-0,035	-0,021	0,017	-0,026	-0,056	-0,016	0,033	-0,055	0,033
$PBLH$	0,137	-0,268	0,441	-0,742	0,748	-0,067	-0,035	—	0,559	-0,497	0,783	0,297	0,060	0,392	0,226	0,074
$Sunrise$	0,222	-0,136	0,445	-0,595	0,581	-0,050	-0,021	0,559	—	-0,501	0,419	0,193	0,043	0,047	0,174	0,163
$Sunset$	-0,233	0,091	-0,404	0,717	-0,693	0,035	0,017	-0,497	-0,501	—	-0,299	-0,092	0,017	0,039	-0,085	-0,170
U_{10} m	0,062	-0,287	0,239	-0,591	-0,693	0,035	0,017	-0,497	0,419	-0,299	—	0,386	-0,005	0,673	0,322	-0,082
u_{10} m	-0,126	-0,163	0,071	-0,266	0,178	-0,035	-0,056	0,297	0,193	-0,092	0,386	—	-0,435	0,177	0,830	-0,548
v_{10} m	0,282	0,010	0,221	0,049	0,043	0,010	-0,016	0,060	0,043	0,017	-0,005	-0,435	—	-0,066	-0,258	0,594
U_{300} m	-0,058	-0,267	-0,079	-0,227	-0,060	-0,074	0,033	0,392	0,047	0,039	0,673	0,177	-0,066	—	0,037	-0,123
u_{300} m	-0,060	-0,188	0,090	-0,226	0,146	-0,014	-0,055	0,226	0,174	-0,085	0,322	0,830	-0,258	0,037	—	-0,504
v_{300} m	0,305	0,102	0,226	-0,090	0,200	0,009	0,033	0,074	0,163	-0,170	-0,082	-0,548	0,594	-0,123	-0,504	—

Tableau C.4 : Idem figure C.2 pour les 16 variables gardées après cette première étape de sélection.

C.1.2.b Seconde étape – choix des variables les plus pertinentes

Jeux de données à 1 variable

Les performances des RNA utilisant une seule variable sont significativement plus faibles que ceux du RNA « étalon ». Il ne se dégage pas une variable qui produise de meilleurs résultats pour tous les indicateurs statistiques (figure C.4).

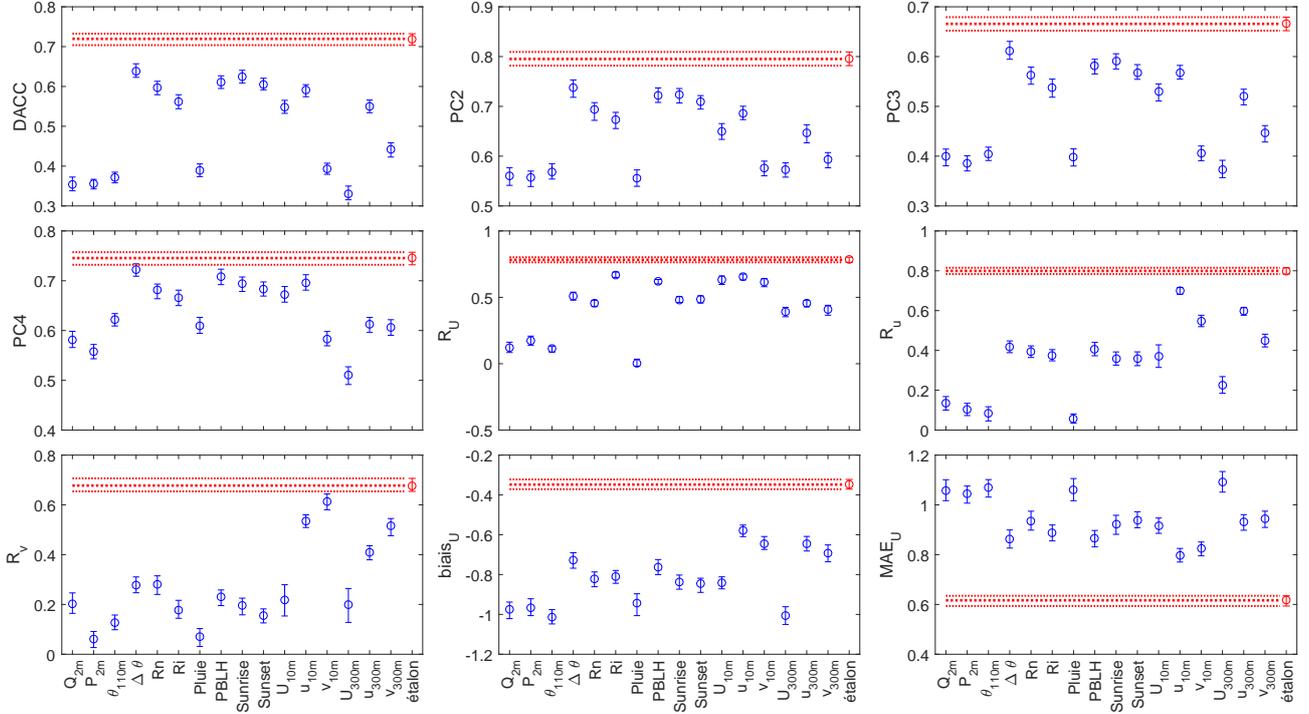


FIGURE C.4 : Performances pour les RNA qui utilisent des jeux de données constitués d'une seule variable. Le nom de la variable sur l'axe x indique la variable utilisée. Les performances du RNA « étalon » sont en rouge. Les barres représentent l'intervalle de confiance à 90 % calculé avec la méthode de bootstrap. Les lignes pointillées rouges indiquent la valeur moyenne et l'intervalle de confiance du RNA étalon.

La variable $\Delta\theta$ produit les meilleurs résultats pour le $DACC$, le $PC2$ et le $PC3$ et $PC4$. La variable u_{10} obtient les meilleures performances pour l'erreur absolue moyenne sur la vitesse, le biais sur la vitesse et la corrélation sur la composante u . De même pour la variable v_{10} qui permet de bien reproduire la composante v du vent observé dans la vallée de Cadarache. C'est assez logique puisque les composantes des vents observés et simulés à 10 m sont corrélées ($R = 0,69$ pour u et $R = 0,57$ pour v). Les composantes à 10 m produisent également les meilleurs résultats du point de vue du module (biais, MAE).

Pour cette première itération, la différence de température potentielle entre 110 m et 2 m a été sélectionnée principalement sur la base des indicateurs statistiques relatifs à la direction ($DACC$ et PC).

Jeux de données à 2 variables

Pour la deuxième itération (figure C.5), la combinaison de la variable $\Delta\theta$ avec la variable u_{10m} produit de meilleurs résultats que n'importe quelle autre combinaison de variables, et ce pour 6 des 9 indicateurs statistiques ($DACC$, $PC2$, $PC3$, $PC4$, MAE et biais). L'ajout du module à 10 m produit les meilleurs résultats pour le coefficient de corrélation sur le module,

mais les différences ne sont pas significatives par rapport aux RNA qui utilisent les composantes à 10 m ou 300 m. L'ajout de la composante u à 10 m ou à 300 m permet de mieux simuler la composante u (R_u) ; il en est de même pour la composante v avec les composantes v . Néanmoins, l'ajout de la composante u_{10} produisant de meilleures performances, c'est donc cette variable qui est sélectionnée.

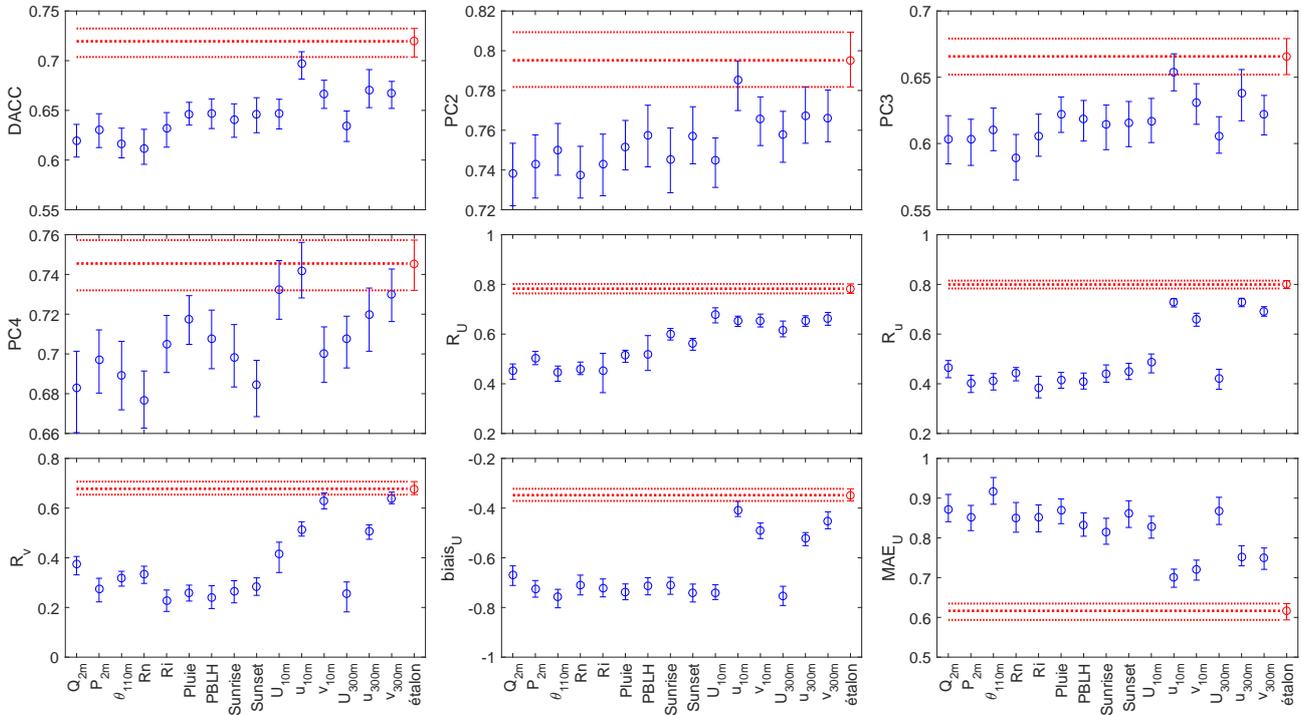


FIGURE C.5 : Idem figure C.4 pour des jeux de données constitués de $\Delta\theta$ et d'une seconde variable dont le nom est indiqué sur l'abscisse.

Jeux de données à 3 variables

Pour la troisième itération (figure C.6), les différentes combinaisons de 3 variables sont très proches les unes des autres, notamment sur la direction. Pour les PC , plusieurs RNA atteignent des performances très proches du RNA étalon, avec des différences qui ne sont pas statistiquement significatives. Les composantes v à 10 m et 300 m permettent d'atteindre des valeurs de R_v très proches de celle du RNA étalon. Pour les autres critères, il reste des différences significatives avec le RNA étalon ce qui montre qu'aucun jeu de données à 3 variables n'est suffisant. Au final, la variable la plus pertinente sur cette itération est la composante du vent v à 10 m.

Jeux de données à 4 variables

Pour la quatrième itération (figure C.7), toutes les combinaisons de variables produisent des résultats proches du RNA « étalon » et les différences ne sont pas toujours significatives. Le rayonnement net est sélectionné comme nouvelle variable parce que c'est la variables qui produit les résultats les plus proches du RNA étalon. Cependant, il reste des différences statistiquement significatives avec le RNA étalon. Il faut donc continuer la construction du jeu de données en ajoutant une nouvelle variable.

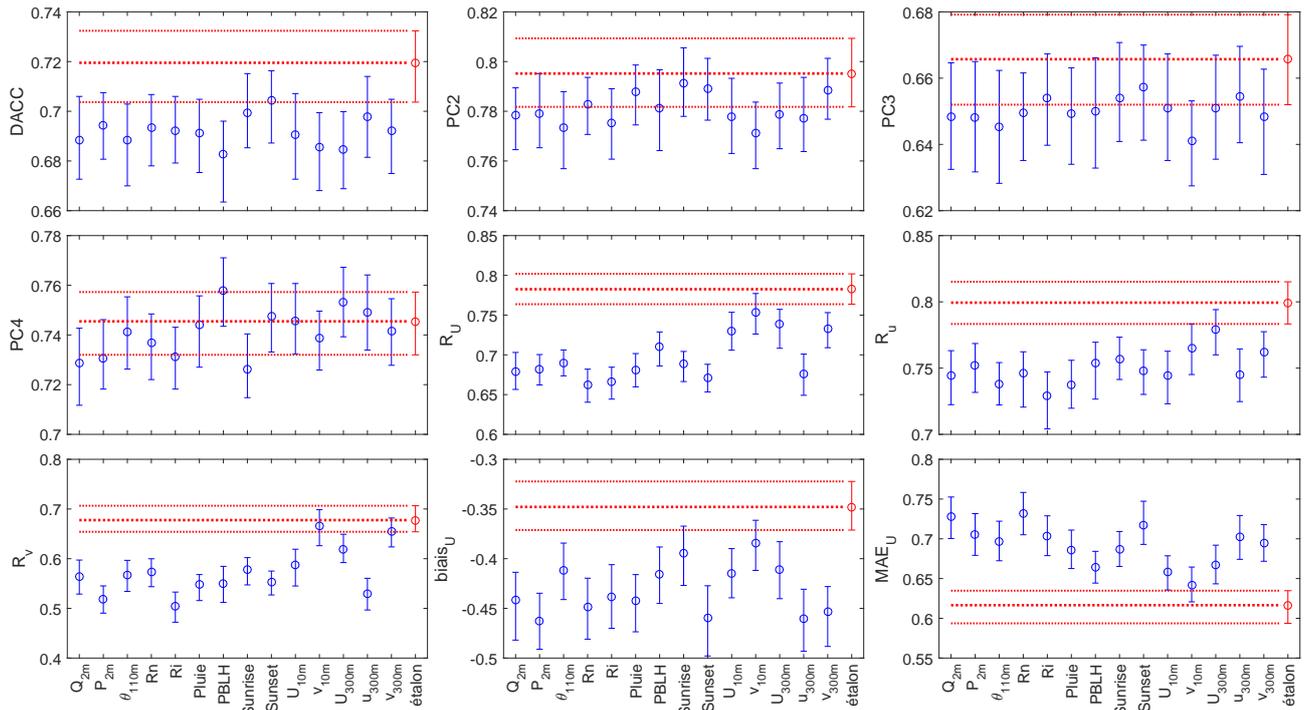


FIGURE C.6 : Idem figure C.4 pour des jeux de données d'entrée constitués de $\Delta\theta$, $u_{110\text{m}}$ et d'une troisième variable dont le nom est indiqué sur l'abscisse.

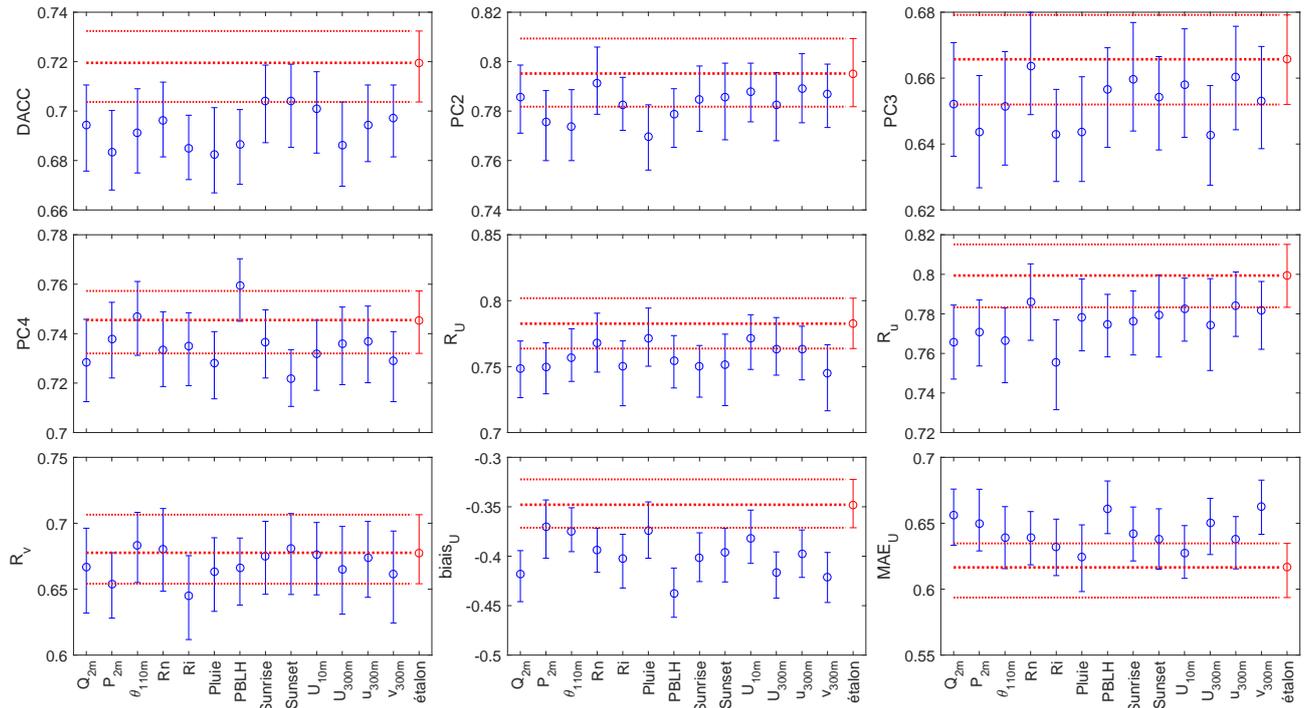


FIGURE C.7 : Idem figure C.4 pour des jeux de données d'entrée constitués de $\Delta\theta$, $u_{110\text{m}}$, $v_{110\text{m}}$ et d'une quatrième variable dont le nom est indiqué sur l'abscisse.

Jeux de données à 5 variables

Pour la cinquième itération (figure C.8), toutes les combinaisons de variables produisent des résultats proches du RNA « étalon », les différences étant inférieures à l'intervalle de confiance. Cependant, l'humidité spécifique à 2 m est la seule variable pour laquelle tous les critères statistiques sont compris dans l'intervalle de confiance du RNA étalon. C'est donc cette variable qui est choisie et qui clôt la construction du jeu de données d'entrée. En conclusion, les 5 variables

retenues pour l'alimentation du RNA à partir des simulations WRF sont : $\Delta\theta$, u_{110m} , v_{110m} , Rn et Q_{2m} .

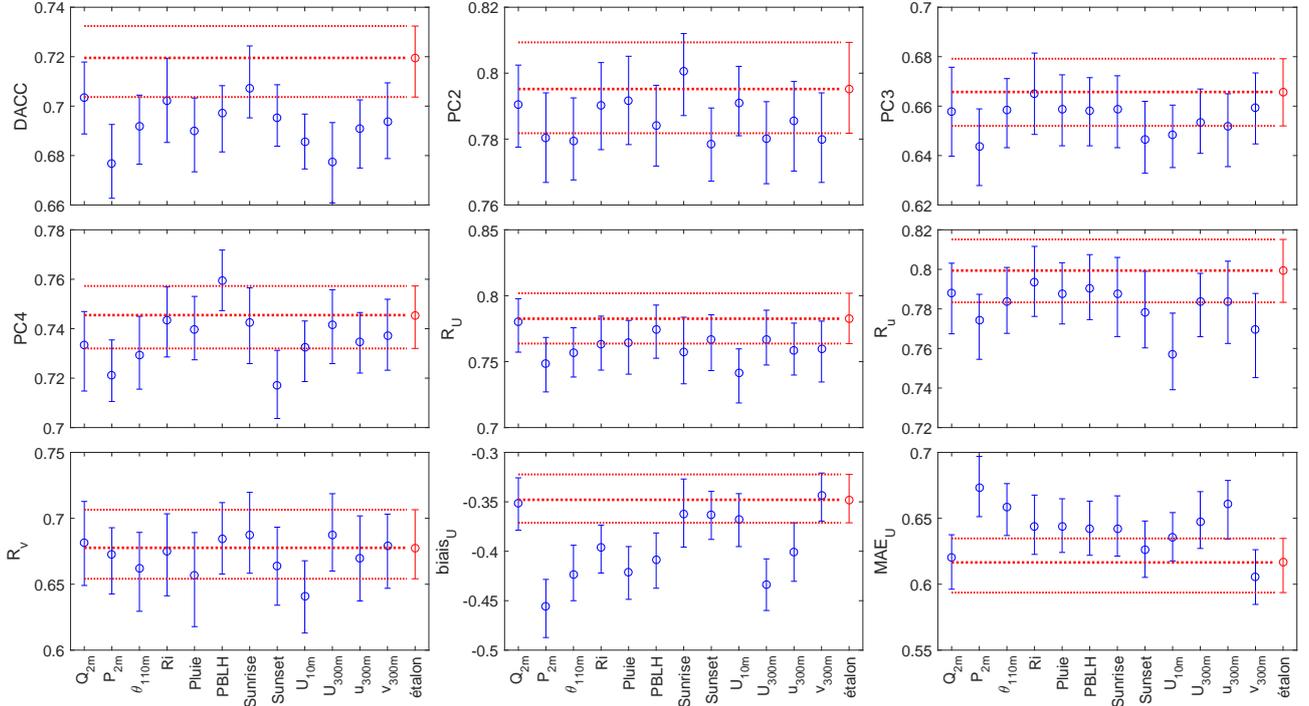


FIGURE C.8 : Pareil que la figure C.4 pour des jeux de données d'entrée constitués de $\Delta\theta$, u_{110m} , v_{110m} , Rn et d'une cinquième variable dont le nom est indiqué sur l'abscisse.

C.2 Longueur du jeu de données

L'évolution des performances du RNA en fonction de la taille du jeu de données utilisé en entrée a été étudiée sur la période de 3 mois de la campagne KASCADE 2013. En théorie, l'apprentissage d'un RNA s'améliore en augmentant la taille du jeu de données. Il n'est pas possible pour nous d'augmenter la taille du jeu de données. La méthode mise en œuvre a donc été de réduire la taille du jeu de données d'entrée et d'observer une éventuelle baisse de performance.

Des jeux de données d'entrée de tailles réduites ont été créés par tirage aléatoire (sans remise) dans le jeu complet. Le tirage aléatoire permet de sélectionner des données sur toute la période en ne favorisant aucune saison en particulier. Les jeux réduits sont composés de 90 %, 80 %, 70 %, 60 %, 50 %, 40 %, 30 %, 25 %, 20 %, 15 %, 10 % et 5 % du jeu complet qui contient 1899 échantillons pour la période KASCADE 2013 (environ 3 mois d'observations à cadence horaire). Ces jeux de données seront par la suite appelés *datNN*, où NN représente le pourcentage de données (donc *dat90*, *dat80*, *dat70*, ...).

Pour les jeux de données les plus gros, l'impact du tirage aléatoire sur la composition du jeu est faible parce que quel que soit le tirage aléatoire il y aura toujours beaucoup de données en commun. Par exemple, pour deux *dat90* construits à partir de 2 tirages aléatoires différents, il y aura au moins 80 % des données du jeu complet qui seront présentes dans les 2 jeux réduits. Pour les plus petits jeux de données il y aura beaucoup plus de différences liées au tirage puisque pour les jeux de taille inférieure à 50 % du jeu complet, il peut n'y avoir aucune donnée identique entre 2 tirages différents. Il est donc important de réaliser un nombre de tirages suffisant

pour tester l'effet de la composition du jeu d'entrée sur la performance du RNA. Cinquante jeux d'entrée différents ont été créés pour chaque taille réduite. Ces jeux de données sont appelés $datNN_1, datNN_2, datNN_3, \dots, datNN_{50}$. Des RNA sont ensuite lancés avec ces différents jeux de données en entrée (avec la même graine¹ pour l'initialisation du tirage aléatoire et la même division des données et set de test, validation et apprentissage).

Pour les jeux de données les plus grands, la dispersion des performances des RNA qui sont lancés avec ces différents jeux d'entrée est faible (exemple pour les $dat90_n$: tableau C.5). C'est assez logique puisque les différents $dat90_n$ ont beaucoup de données en commun. En revanche, les différences sont plus importantes pour les plus petits (exemple pour les $dat5_n$: tableau C.5) ce qui montre que la constitution du jeu de données par tirage aléatoire peut créer des jeux de données facilitant ou compliquant l'apprentissage du RNA. Il est donc préférable de créer plusieurs jeux de données d'entrée pour prendre cet effet en compte.

	<i>DACC</i>	<i>PC2</i>	<i>PC3</i>	<i>PC4</i>	R_U	R_u	R_v	<i>Biais</i>	<i>EQM</i>
Moy _{90%}	0,8274	0,8806	0,8119	0,8206	0,9291	0,9510	0,9073	-0,1706	0,4953
$\sigma_{90\%}$	0,0171	0,0117	0,0167	0,0196	0,0051	0,0044	0,0326	0,0382	0,0197
Moy _{5%}	0,6966	0,8117	0,6793	0,6641	0,8202	0,8517	0,7488	-0,2681	0,9453
$\sigma_{5\%}$	0,1673	0,1130	0,1567	0,1732	0,1667	0,1792	0,2843	1,1239	0,9005

Tableau C.5 : Moyennes (Moy) et dispersion (σ) des performances calculées pour les RNA lancés avec les jeux $dat90_1, \dots, dat90_{50}$ et $dat5_1, \dots, dat5_{50}$ en entrée.

Au final, 50 RNAs sont lancés pour chaque taille de sous-échantillon puis un calcul bootstrap est appliqué pour chaque RNA (100 pour chaque RNA), ce qui fait un total de 5000 calculs pour chaque taille réduite de jeu de données. Une moyenne et un intervalle de confiance à 90 % sont calculés à partir de ces 5000 calculs. La figure C.9 montre l'évolution des performances en fonction du nombre d'échantillons composants le jeu de données d'entrée.

La taille du jeu de données utilisé en entrée semble avoir peu d'importance sur les performances moyennes sauf pour les deux échantillons les plus petits. En revanche, l'intervalle de confiance est petit pour les jeux les plus grands mais il grandit progressivement avec la diminution de la taille du jeu de données.

Cette augmentation de la dispersion des performances n'est pas uniquement liée à la diminution de la taille du jeu de données, il y a aussi une part de cette dispersion qui provient des différences de performances liées au tirage aléatoire. L'écart-type calculé sur les 5000 valeurs (équation C.1), résulte en partie du bootstrap et en partie du tirage aléatoire. En calculant un écart-type indépendant pour chacun des jeux de données et en faisant la moyenne de ces écarts-types, on peut considérer que la valeur ne contient plus que la partie liée au bootstrap (équation C.2).

$$\sigma_1 = \sqrt{\frac{1}{n \times m} \sum_{i=1}^n \sum_{j=1}^m (x_{i,j} - \bar{x})^2} \quad \text{avec} \quad \bar{x} = \frac{1}{n \times m} \sum_{i=1}^n \sum_{j=1}^m x_{i,j} \quad (\text{C.1})$$

$$\sigma_2 = \frac{1}{n} \sum_{i=1}^n \sqrt{\frac{1}{m} \sum_{j=1}^m (x_{i,j} - \bar{x}_i)^2} \quad \text{avec} \quad \bar{x}_i = \frac{1}{m} \sum_{j=1}^m x_{i,j} \quad (\text{C.2})$$

1. l'indication utilisée pour la génération pseudo-aléatoire de nombres

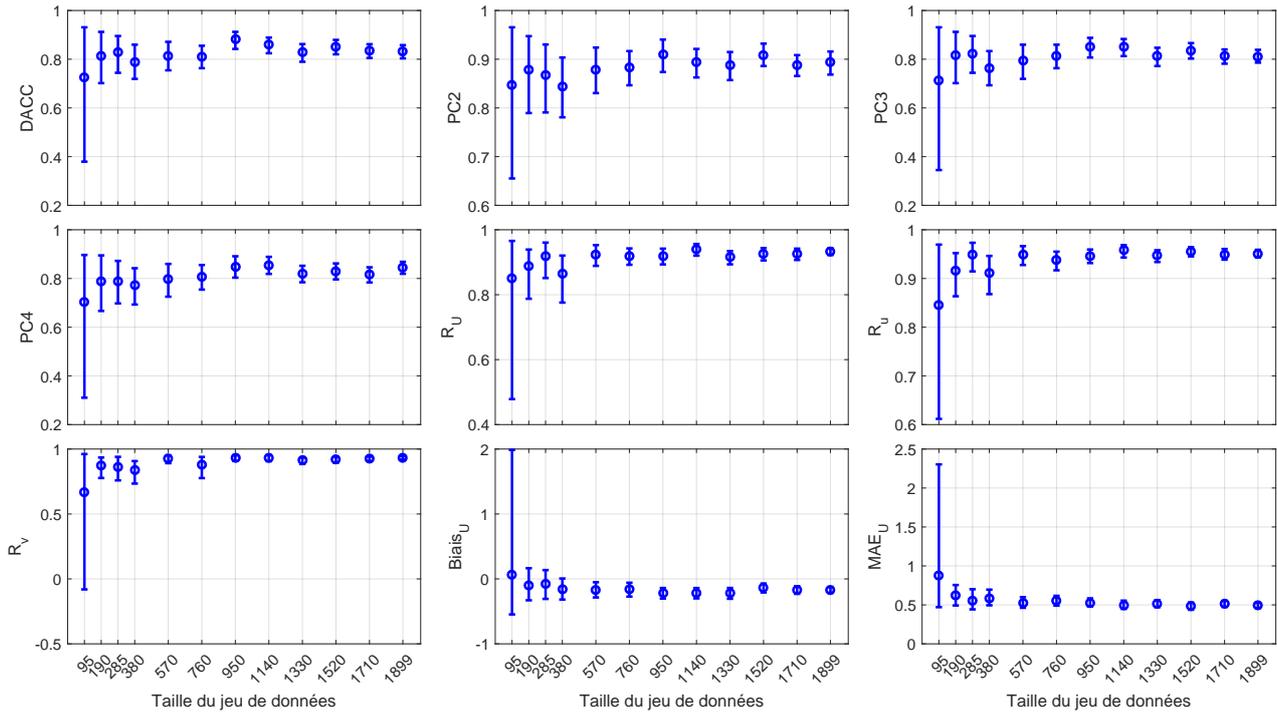


FIGURE C.9 : PC à 2, 3 et 4 classes, DACC, MAE et biais sur la vitesse et coefficients de corrélation pour la vitesse et les composantes u et v pour des RNA qui utilisent des jeux de données de tailles différentes (95 valeurs correspondent à 5% du jeu complet, 190 valeurs à 10%, ...).

avec $n = 50$, le nombre de tirages aléatoires effectués pour chaque taille de jeu réduit, $m = 100$ le nombre de calculs bootstrap effectués pour chaque RNA et x les indicateurs statistiques ($DACC$, PC , ...).

Ces 2 valeurs d'écart-type calculées pour les plus grands jeux de données sont très proches (tableau C.6), ce qui est attendu puisque les différents jeux contiennent beaucoup de données en commun et donc la variabilité liée au tirage aléatoire est faible. Au contraire, pour les plus petits jeux de données, avec peu de données en commun, les différences entre σ_1 et σ_2 sont plus importantes ce qui montre l'importance du tirage aléatoire dans la dispersion des valeurs des indicateurs statistique.

Il est donc difficile de conclure sur la taille idéale du jeu de données en diminuant le nombre d'échantillons. Un jeu de données réduit peut donner de très bonnes performances s'il contient beaucoup de données faciles à analyser pour le RNA (vents alignés avec la vallée, vents forts) ou de très mauvaises performances si le tirage produit plus de données difficiles à analyser. En revanche, le fait que les performances moyennes ne diminuent pas et que les écart-types n'augmentent pas pour les jeux de données de 100% à 50% montre qu'un jeu de données de 3 mois est suffisamment grand pour ne pas impacter la performance du RNA.

	<i>DACC</i>	<i>PC2</i>	<i>PC3</i>	<i>PC4</i>	R_U	R_u	R_v	<i>Biais</i>	<i>EQM</i>
σ_1 100%	0,0166	0,0134	0,0170	0,0151	0,0070	0,0052	0,0085	0,0279	0,0187
σ_2 100%	0,0166	0,0134	0,0171	0,0152	0,0071	0,0052	0,0085	0,0280	0,0187
σ_1 90%	0,0171	0,0129	0,0179	0,0195	0,0102	0,0064	0,0103	0,0360	0,0237
σ_2 90%	0,0156	0,0122	0,0162	0,0176	0,0101	0,0062	0,0097	0,0294	0,0218
σ_1 80%	0,0174	0,0144	0,0188	0,0204	0,0114	0,0058	0,0134	0,0422	0,0282
σ_2 80%	0,0162	0,0134	0,0173	0,0178	0,0106	0,0056	0,0123	0,0325	0,0241
σ_1 70%	0,0220	0,0174	0,0232	0,0209	0,0124	0,0072	0,0153	0,0500	0,0274
σ_2 70%	0,0185	0,0152	0,0186	0,0175	0,0116	0,0069	0,0130	0,0351	0,0207
σ_1 60%	0,0198	0,0172	0,0205	0,0209	0,0112	0,0076	0,0219	0,0482	0,0324
σ_2 60%	0,0179	0,0158	0,0176	0,0172	0,0105	0,0074	0,0136	0,0356	0,0267
σ_1 50%	0,0213	0,0202	0,0242	0,0259	0,0149	0,0085	0,0136	0,0500	0,0314
σ_2 50%	0,0194	0,0171	0,0207	0,0222	0,0139	0,0081	0,0108	0,0413	0,0269
σ_1 40%	0,0283	0,0223	0,0288	0,0299	0,0150	0,0118	0,0527	0,0645	0,0386
σ_2 40%	0,0251	0,0206	0,0257	0,0249	0,0141	0,0110	0,0462	0,0551	0,0351
σ_1 30%	0,0352	0,0284	0,0414	0,0407	0,0198	0,0116	0,0190	0,0690	0,0414
σ_2 30%	0,0298	0,0243	0,0309	0,0297	0,0191	0,0110	0,0179	0,0533	0,0375
σ_1 20%	0,0420	0,0371	0,0433	0,0451	0,0435	0,0478	0,0551	0,0993	0,0617
σ_2 20%	0,0361	0,0343	0,0377	0,0398	0,0410	0,0255	0,0476	0,0748	0,0509
σ_1 15%	0,0484	0,0420	0,0482	0,0538	0,0361	0,0230	0,0580	0,1402	0,0794
σ_2 15%	0,0403	0,0363	0,0408	0,0438	0,0278	0,0166	0,0468	0,0842	0,0610
σ_1 10%	0,0642	0,0510	0,0669	0,0712	0,0457	0,0283	0,0484	0,1513	0,0799
σ_2 10%	0,0540	0,0443	0,0520	0,0554	0,0388	0,0249	0,0448	0,0975	0,0629
σ_1 5%	0,1559	0,0943	0,1684	0,1682	0,1801	0,1176	0,3419	0,6858	0,5188
σ_2 5%	0,0830	0,0688	0,0831	0,0824	0,0571	0,0962	0,1766	0,1987	0,1426

Tableau C.6 : Performances calculées pour les RNA lancés avec les jeux $dat90_1, \dots, dat90_{50}$ en entrée.