

N° d'Ordre : D.U. 2359

# UNIVERSITÉ CLERMONT II – BLAISE PASCAL

U.F.R. Sciences et Technologies

ÉCOLE DOCTORALE DES SCIENCES FONDAMENTALES

N : 751

## THÈSE

présentée pour obtenir le grade de

DOCTEUR D'UNIVERSITÉ

*Spécialité : Physique de l'Atmosphère*

par Carole LEGORGEU

Diplômée des masters « Sciences de la Planète et de l'Environnement » spécialité Recherche *Océan, Atmosphère et Surfaces Continentales* et spécialité Professionnelle *Génie de l'Environnement – option AIR* (UPS, Toulouse III)

---

### AMÉLIORATION DES ESTIMATIONS QUANTITATIVES DES PRÉCIPITATIONS À HAUTES RÉOLUTIONS :

*comparaison de deux techniques combinant les observations et application à la vérification spatiale des modèles météorologiques.*

---

Soutenue publiquement le 18 juin 2013, devant la commission d'examen :

<b>Président :</b>	M. VAN BAELEN Joël	(Directeur de recherche, Laboratoire de Météorologie Physique, Aubière)
<b>Rapporteurs :</b>	M. ANDRIEU Hervé	(Directeur de recherche, Institut de Recherche Sciences et Techniques de la Ville, Nantes)
	M. BERNE Alexis	(Professeur assistant tenure track, École Polytechnique Fédérale de Lausanne, Lausanne)
<b>Examineur :</b>	M. BOUDEVILLAIN Brice	(Physicien-Adjoint, Laboratoire d'étude des Transferts en Hydrologie et Environnement, Grenoble)
<b>Invité :</b>	M. PERGAUD Julien	(Ingénieur de Recherche, NUMTECH, Aubière)
<b>Directeur de thèse :</b>	M. WOBROCK Wolfram	(Professeur, Laboratoire de Météorologie Physique, Aubière)



Thèse préparée au sein du  
Laboratoire de Météorologie Physique (OPGC/LaMP, UMR6016 CNRS/UBP, Aubière)



financée dans le cadre d'un projet collaboratif supporté par le Conseil Régional d'Auvergne et  
le Fonds Européen de Développement Régional en Auvergne.





*À mon papa,*



# REMERCIEMENTS

Je me revois à la sortie du lycée, mon bac en poche, me disant « fini la physique ! ». Je me revois assise derrière mon ordinateur à chercher une orientation pour ma troisième année de licence me disant « je ne ferais jamais une thèse »... Pourtant, après plus de neuf années d'étude et un passage dans le monde professionnel, me voilà à rédiger des remerciements pour finir mon manuscrit de thèse en physique de l'atmosphère.

Je commencerais par remercier par Wolfram WOBROCK pour avoir supervisé cette thèse un peu particulière. Avec mes idées nouvelles (souvent anticonformistes), il m'a, pour autant, accordé une entière liberté dans mes travaux et une grande confiance. Il a toujours trouvé un moment pour répondre à mes questions, même si parfois j'avais l'impression de parlé dans un langage inconnu. Je tiens également à témoigner toute ma reconnaissance à Julien PERGAUD, qui m'a également suivi durant ces trois années. Malgré son emploi du temps chargé, il m'a prêté une attention sans faille (d'autres n'en auraient pas fait autant). Même si ce n'était pas son domaine de prédilection, il a toujours réussi à trouver la lumière dans l'obscurité. Ses remarques (parfois déconcertantes), ses conseils et son soutien m'ont été d'une aide incroyable.

Mes remerciements s'adressent ensuite aux rapporteurs de cette thèse, MM. ANDRIEU et BERNE, pour l'intérêt qu'ils m'ont porté (tant sur l'évaluation de ces 190 pages que sur leurs remarques faites oralement) J'associe à ces remerciements, M. BOUDEVILLAIN, pour avoir examiné mon travail avec beaucoup d'intérêt, et M. VAN BAELEN, pour sa présidence lors de ma soutenance.

Cette thèse n'aurait pas abouti sans l'accès aux différentes bases de données. Concernant les observations en région cévenole, j'aimerais à nouveau remercier Brice BOUDEVILLAIN ainsi que Guy DELRIEU Guy pour avoir pris le temps de répondre à mes questions sur les mesures RADAR. De même, je remercie Yves POINTIN et Guillaume MIOCHE pour les données relatives à l'agglomération clermontoise. Enfin, je citerais Sandra BANSON pour toutes mes requêtes sur les données Météo-France et les analyses ECMWF.

Merci à l'ensemble des professeurs, chercheurs et ingénieurs du laboratoire qui m'ont accordé un peu de leur temps pour avoir répondu à mes questions diverses (scientifiques, administratives ou informatiques), qui ont participé à l'amélioration de mon travail et m'ont épaulé lors de la préparation de ma soutenance. Merci à tous ceux (anciens et présents) avec qui j'ai partagé des moments entre fou rire et doute. Ils sont nombreux et je ne pourrais tous les citer sans en oublier. J'aurais toutefois une pensée particulière pour Christelle B., Laurent D., Laurent L., Aura L., Yoann L., Céline P. et Frédéric T.

Mes plus grands remerciements iront à mes ami(e)s et ma famille. Pauline R., je ne trouverais jamais les mots pour te remercier de tout ce que tu m'as apporté (une chose est sûre, je bénis l'inventeur du forfait illimité !). Magalie B., Julien C., vous avez toujours répondu présents que ce soit dans la joie comme dans la peine. Je n'oublie pas Clémentine D., Julien N., Thomas B. et Vincent V. pour lesquels je garderais d'excellents souvenirs. Un immense merci à ma (farfelue extraordinaire) Véro B. et à mon (deuxième papa) Jean D. pour votre soutien infailible (qui ne pourra jamais être copié !). Je remercie également mon (futur Maître) Ludovic B. et ma (houbi-houba) Amandine B-M qui ont su m'offrir de très bons moments d'évasion (et bien plus). Enfin, j'adresse tout mon amour aux deux personnes qui sont ce que j'ai de plus cher au monde (et qui ont eu la bonté de corriger mon manuscrit) : mon frère (qui est tout ce que je voudrais être... enfin presque !) et ma maman (sans qui je ne serais rien).

« *Un bonheur que rien n'a entamé succombe à la moindre atteinte ; mais quand on doit se battre contre les difficultés incessantes, on s'aguerrit dans l'épreuve, on résiste à n'importe quels maux, et même si l'on trébuche, on lutte encore à genoux* ». Extrait de La vie heureuse de Sénèque.

Je finirais par cette citation qui traduit ce que m'a appris mon défunt père. Son courage et sa volonté « phénoménaux » ont été et resteront à jamais ma plus grande source d'inspiration. La vie n'est pas toujours facile, mais ce n'est pas pour autant qu'on a le droit de baisser les bras, il faut avancer coûte que coûte (quitte à y laisser quelques plumes). Il nous a toujours encouragé à poursuivre nos rêves... il a tout fait pour que je puisse réaliser les miens. Il ne le sait pas, mais c'est pour lui et grâce à lui que je clos ce chapitre de ma vie. Maintenant, c'est une autre page qui se tourne...



# TABLE DES MATIERES

Introduction.....	1
La pluie : Origines et Mesures .....	5
I-1 D'où vient la pluie ?.....	5
I-1.1 Origine de la pluie – partie I .....	5
I-1.2 Les types de pluie.....	6
I-2 Comment mesure-t-on la pluie ? .....	7
I-2.1 Mesures directes des précipitations.....	7
I-2.2 Mesures indirectes des précipitations .....	11
Reconstruction d'un champ de pluie : méthodologie.....	23
II-1 Contrôle des données pluviométriques .....	23
II-1.1 Version initiale .....	24
II-1.2 Modifications apportées.....	26
II-2 Reconstruction d'un champ spatial à partir de données ponctuelles .....	27
II-2.1 Le krigeage : théorie générale .....	28
II-2.2 La combinaison des données par krigeage.....	37
Reconstruction d'un champ de pluie : Application dans la région Cévennes- Vivarais.....	47
III-1 La région Cévennes-Vivarais .....	47
III-1.1 Description physique des événements cévenols .....	48
III-1.2 Le réseau d'observations.....	50
III-2 Spatialisation des données .....	52
III-3 CAS A : du 19 au 22 Octobre 2008.....	53
III-3.1 Situation météorologique .....	53
III-3.2 Le contrôle de qualité des données pluviométriques .....	55
III-3.3 Informations complémentaires.....	57
III-3.4 Reconstruction d'un champ de précipitation à l'échelle de l'événement.....	57
III-3.5 Reconstruction d'un champ de précipitation à différentes échelles temporelles .....	65
III-4 CAS B : du 31 Octobre au 2 Novembre 2008 .....	68
III-4.1 Situation météorologique .....	68
III-4.2 Le contrôle de qualité des données pluviométriques .....	68
III-4.3 Informations complémentaires.....	70
III-4.4 Reconstruction d'un champ de précipitation à l'échelle de l'événement.....	72
III-4.5 Reconstruction d'un champ de précipitation à différentes échelles temporelles .....	79
III-5 Test de sensibilité : rôle de la densité de pluviomètres.....	83
III-6 Conclusions.....	88
Reconstruction d'un champ de pluie : Application autour de l'agglomération clermontoise.....	91
IV-1 Clermont-Ferrand et son agglomération .....	91
IV-2 Le réseau d'observation .....	93

IV-2.1 Le RADAR .....	93
IV-2.2 Les réseaux de mesures au sol.....	94
IV-3 Spatialisation des données : définition du variogramme paramétrique ..	95
IV-3.1 Théorie générale.....	95
IV-3.2 Application aux observations pluviométriques de Clermont-Ferrand.....	97
IV-4 CAS HIVER : 16 Mars 2011.....	107
IV-4.1 Situation météorologique.....	107
IV-4.2 Les données.....	108
IV-4.3 Reconstruction d'un champ de précipitation à l'échelle de l'événement .....	111
IV-5 CAS ÉTÉ : 31 Août 2011 .....	120
IV-5.1 Situation météorologique.....	120
IV-5.2 Les données.....	121
IV-5.3 Reconstruction d'un champ de précipitation à l'échelle de l'événement .....	126
IV-6 Conclusions .....	132
Modélisation de la pluie avec WRF .....	135
V-1 Origine de la pluie – partie II .....	135
V-2 Les modèles météorologiques .....	137
V-2.1 Généralités.....	137
V-2.2 Modélisation des nuages et des précipitations .....	139
V-3 Le modèle à méso-échelle WRF .....	141
V-3.1 Généralités.....	141
V-4 Simulation idéalisée de la pluie avec le modèle WRF.....	144
V-4.1 Contexte de l'étude : les pluies orographiques .....	144
V-4.2 Initialisation du modèle .....	146
V-4.3 Les schémas de microphysique.....	149
Évaluation du modèle WRF : application sur les précipitations .....	161
VI-1 Bonne ou mauvaise prévision : telle est la question .....	161
VI-1.1 Petit historique des méthodes de vérification .....	161
VI-1.2 Les méthodes de vérification spatiale.....	163
VI-2 La technique SAL .....	167
VI-2.1 Définition mathématique .....	167
VI-2.2 Discussions .....	170
VI-3 Application de la technique SAL.....	172
VI-3.1 Paramétrisation du modèle .....	172
VI-3.2 CAS A : résultats de la technique SAL .....	173
VI-4 Autres méthodes de vérification .....	176
VI-4.1 Approche upscaling .....	176
VI-4.2 Approche traditionnelle.....	180
Conclusions générales .....	185
Perspectives .....	188
BIBLIOGRAPHIE .....	192
LISTE DES SYMBOLES .....	202



# Introduction

Il est estimé que près de 12 000 milliard de tonnes d'eau sont présents à chaque instant dans l'atmosphère, de quoi recouvrir l'ensemble de la planète d'une couche d'eau liquide de 2 cm. Cette eau atmosphérique est principalement sous forme de vapeur d'eau mais aussi sous forme de gouttelettes liquides et de cristaux de glace qui composent les nuages. Le terme « précipitations » désigne tous produits de la condensation de la vapeur d'eau (Météo-France 2011, 10–11). Sur le globe, les précipitations sont très inégalement réparties : certaines régions souffrent d'un déficit important rendant presque impossible la vie (famine, maladie, etc.) tandis que d'autres sont sujettes à des précipitations régulières parfois à l'origine de crues de cours d'eau et d'inondations ravageuses.

Dans le quatrième rapport de synthèse du Groupe Intergouvernemental d'étude sur l'Évolution du Climat (GIEC 2007, p. 8), les projections climatiques sur le XXI<sup>e</sup> siècle annoncent une hausse « *très probable* » de la fréquence des épisodes de fortes précipitations. Zhang *et al.* (2007) estiment que les forçages anthropiques ont contribué de façon significative à l'augmentation des précipitations dans les latitudes moyennes et à l'assèchement des tropiques. L'incertitude quant à l'évolution climatique a conduit les acteurs à réfléchir arduement sur ces problématiques, devançant ainsi les politiciens.

En juin 2010 et novembre 2011, le département du Var a connu deux inondations catastrophiques majeures. La première causa 23 morts, 2 disparus et 1.2 milliards d'euros de dégâts. La seconde, qui toucha tout le sud-est de la France, coûta la mort à 4 personnes et entre 500 et 800 millions d'euros de préjudices (Collombat 2012, p. 13). La Direction Générale de la Prévention des Risques, la Direction Générale de la Sécurité Civile et de la Gestion des Crises, Météo-France, le Service Central d'Hydrométéorologie et d'Appui à la Prévision des Inondations, les Services de Prévision des Crues, etc., nombreux sont les acteurs français visant à surveiller, informer, prévenir, alerter, protéger ou secourir les populations (Météo-France 2011, 4–9). À chaque événement dévastateur, des engagements politiques sont promus. Pourtant les règles de fonctionnement actuelles n'ont pas permis de préparer parfaitement les populations touchées. Cette constatation a récemment incité le Sénat à faire prendre conscience aux politiciens de la nécessité d'établir un plan d'action pour ces régions particulièrement sensibles.

En janvier 2012, la conférence des présidents a ainsi décidé la création d'une mission commune d'information. Après 9 mois d'investigation, les sénateurs ont présenté leur rapport intitulé « Se donner les moyens de ses ambitions : les leçons des inondations du Var et du sud-est de la France » dans lequel ils préconisent de repenser globalement la prévention et la gestion de crise (et d'après-crise) des risques liés aux inondations. Parmi 22 recommandations (Collombat 2012, 283–291) ils suggèrent notamment la nomination d'un établissement public d'aménagement et de gestion des eaux pour chaque bassin versant. Ces propositions devraient faire prochainement l'objet d'une proposition de loi (site Internet : [senat.fr](http://senat.fr)) car le cas varois n'est pas une exception : 19 000 communes françaises sont en zones inondables (Molga 2012).

Courant 2009, un consortium de 6 partenaires auvergnats (3 sociétés privées et 3 laboratoires de recherche publique) a été mis en place autour d'un projet collaboratif. À terme, l'objectif est la conception et le développement de plates-formes opérationnelles pour la gestion des risques environnementaux (Brocheton *et al.* 2012, p. 2). L'approche dans laquelle s'inscrivent les travaux de cette thèse s'axe sur la surveillance des phénomènes précipitants et de leur prévision afin d'anticiper et de réduire les risques associés. L'objectif de la thèse est double : d'une part améliorer les estimations des champs de pluie à partir des observations disponibles et d'autre part vérifier les prévisions issues des modèles météorologiques.

La première partie s'articule autour de l'étude des moyens d'observations et des méthodes permettant de déterminer la meilleure cartographie possible de la pluie au sol.

Le principe de la mesure des précipitations consiste à chiffrer la quantité d'eau atteignant la surface. La principale difficulté réside dans l'obtention d'un échantillon représentatif des quantités reçues par une région. Ainsi le choix des appareils, de leur emplacement et des méthodes d'exploitation font l'objet d'une attention particulière. Le Chapitre I présente les principaux instruments de mesures, leur principe de fonctionnement et les incertitudes associées. Le plus couramment utilisé est le pluviomètre qui fournit une quantité de pluie précise mais ponctuelle. Afin d'obtenir une information sur le caractère spatial des précipitations, il est nécessaire de disposer d'un large réseau de pluviomètres. Une manière de s'affranchir de ce problème est d'utiliser les RADAR météorologiques. Leur principe est basé sur la réflexion d'une onde radioélectrique par les gouttes d'eau présentes dans l'air. Le traitement du signal reçu permet d'obtenir une image détaillée des précipitations sur une large zone géographique (pouvant aller jusqu'à 200 km autour du RADAR) avec une

fine résolution spatiale (d'une cinquantaine de mètres à quelques kilomètres) et temporelle (d'une trentaine de secondes à plusieurs minutes). L'intensité des précipitations est déduite des mesures qui sont soumises à de nombreux problèmes, d'autant plus prononcés en régions montagneuses.

Les pluviomètres et RADAR météorologiques sont par conséquent complémentaires : l'un mesure la quantité des précipitations et l'autre donne leur répartition spatiale. Le défi est de trouver la meilleure façon de combiner ces deux informations pour obtenir une estimation du champ de pluie au sol la plus proche de la réalité. Nous nous sommes alors focalisés sur deux techniques géostatistiques qui seront détaillées dans le 0. Le cœur de ce manuscrit s'axe sur l'analyse des performances de ces deux techniques choisies. Des études ont alors été menées pour deux domaines localisés dans le Massif-Central. Le premier concerne la région Cévennes-Vivarais qui possède un dense réseau de pluviomètres et un RADAR performant (0). Le second couvre la ville de Clermont-Ferrand et ses alentours où sont disponibles des données d'une vingtaine de pluviomètres et un RADAR fonctionnant à très hautes résolutions spatiale et temporelle (Chapitre IV).

La seconde partie de la thèse est consacrée à la mise en place d'une méthodologie fiable permettant de comparer observations et simulations numériques. L'effort entrepris fut porté sur le modèle « Weather Research et Forecasting ». Ce nouveau standard de la simulation de l'atmosphère à méso-échelle est décrit dans le Chapitre V. Les capacités du modèle à reproduire les précipitations ont été testées au travers d'une situation idéalisée. Enfin, nous présenterons dans le 0 une revue des différentes méthodes de vérification. Parmi les nombreuses techniques existantes, celle retenue repose sur la quantification des erreurs de prévisions sur la structure, l'intensité et la localisation du champ de pluie (Wernli *et al.* 2008). Au travers d'un cas réel, nous exposerons les avantages des informations ainsi obtenues et les confronterons à celles de méthodes plus traditionnelles.

Les travaux présentés dans ce manuscrit s'orientent sur la quantification spatiale des précipitations à partir des observations et la quantification des erreurs de prévisions faites *via* un modèle météorologique. *In fine*, l'objectif est d'assurer le suivi en temps réel et la prévision des pluies à hautes résolutions sur de petites zones d'étude telles que le bassin versant de l'agglomération Clermontoise.



# Chapitre I

## La pluie : Origines et Mesures

Pendant longtemps, la pluie a constitué un grand mystère pour les hommes. Dans de nombreuses cultures, la formation des précipitations était associée à une volonté divine. Aujourd'hui encore perdurent des rituels destinés à invoquer la pluie dans l'espoir de prévenir de la sécheresse et d'assurer la protection des récoltes. Par exemple, dans les Balkans ou pour les indiens Hopis, une cérémonie de la danse de la pluie est effectuée ; pour les mayas Chorti, des faiseurs de pluie frappent des serpents considérés comme les messagers des Dieux de la pluie. Selon les régions du monde, l'attitude des populations par rapport à la pluie diffère. Dans les régions tempérées comme l'Europe, elle a tendance à avoir une connotation triste, négative voire destructive. Dans les régions sèches à désertiques, où les cours d'eau sont rares et la distribution en l'eau potable et l'irrigation sont conditionnées par les précipitations, elle est considérée comme une bénédiction, apportant purification, fertilité et source de vie. Comprendre l'origine des pluies fait partie de la volonté humaine depuis la nuit des temps.

Ce premier chapitre est scindé en deux parties. La première introduit brièvement le phénomène à l'origine des précipitations et les différents types de temps associés. La seconde partie, plus complète, décrit les principaux instruments utilisés pour mesurer la pluie, leur principe de fonctionnement et les erreurs associées à leur mesure.

### I-1 D'où vient la pluie ?

#### I-1.1 Origine de la pluie – partie I

L'eau de la pluie provient de la condensation de la vapeur d'eau présente dans l'air qui nous entoure. Or, l'atmosphère ne peut contenir qu'une quantité limitée de vapeur d'eau qui dépend de la pression (donc de l'altitude) et de la température. La variation de cette valeur limite, dite rapport de mélange à saturation, est décrite par deux principes (Yates 2006, p. 9) :

- l'augmentation de la température d'une particule d'atmosphère fait augmenter son rapport de mélange à saturation ;

- le soulèvement d'une particule d'atmosphère fait diminuer son rapport de mélange à saturation.

Si la quantité de vapeur d'eau dépasse ce rapport de mélange, l'atmosphère est dite « sursaturée ». Cette situation instable ne sera équilibrée que par la condensation de la vapeur d'eau excédentaire formant alors de très fines gouttes d'eau liquide ou des cristaux de glace. Un nuage est né.

Pour donner lieu à des précipitations, il faut que les gouttelettes nuageuses grossissent suffisamment (*i.e.* diamètre supérieur à 80 – 100  $\mu\text{m}$ ). Le processus de condensation est la première étape à la formation de la pluie mais à mesure que se poursuit l'ascendance, d'autres phénomènes entrent en jeu dans la croissance des gouttes (ex : coalescence, collection, collision, rupture, évaporation). En dessous de 0°C, des cristaux de glace apparaissent. Par des processus d'agrégation ou de givrage, ils formeront des flocons de neige, de la neige roulée (ou graupel) ou encore la grêle. Ainsi, au sein d'un même nuage, des particules liquides peuvent coexister avec des particules glacées. Les hydrométéores pris dans l'engrenage des courants verticaux peuvent alors subir jusqu'à plusieurs cycles au cours desquels la taille des particules augmente, jusqu'à ce que les particules soient contraintes à la gravité et tombent, ou bien diminue, jusqu'à possible disparition par évaporation ou sublimation sans même avoir touché le sol.

La formation des précipitations est un phénomène complexe combinant de nombreux processus physiques. Un nombre incommensurable de possibilités engageant chacun de ces mécanismes indépendamment induit qu'aucun événement ne se ressemble, tant sur la nature, la forme (= diamètre) ou le nombre des particules précipitantes que sur leur répartition spatiale et temporelle.

## I-1.2 Les types de pluie

En météorologie, nous distinguons deux types de précipitation : les pluies convectives et les pluies stratiformes. Nous les détaillons rapidement ici.

Les précipitations convectives résultent d'une ascension rapide des masses d'air chargées d'humidité. Ces événements sont généralement associés à des nuages à fort développement vertical où la phase glace est prédominante. Les précipitations résultantes sont : i) très localisées, isolées ou organisées, ii) de courte durée et surtout iii) de fortes intensités. Ce type de précipitation est très commun proche de l'équateur (sous forme d'orages) et des zones tropicales (à l'origine des cyclones). Aux latitudes moyennes, ces

précipitations sont synonymes de l'arrivée d'un front froid où de l'air froid avance rapidement vers de l'air plus chaud et plus lent qui est alors contraint de s'élever.

Les précipitations stratiformes trouvent leur origine dans le soulèvement lent et à grande échelle de l'humidité qui se condense de façon uniforme. De faibles intensités, ces événements couvrent une large zone et sont souvent persistants dans le temps. Ces pluies annoncent l'arrivée d'un front chaud où de l'air chaud avance vers de l'air plus froid et plus dense pour finalement le surmonter. Ce type de précipitation se retrouve également lors du soulèvement d'une masse d'air chaude et humide à l'abord d'un massif montagneux, c'est le phénomène de pluie orographique (décrit dans le paragraphe V-4.1).

Ces deux types de précipitations ne sont pas mutuellement exclusifs. En effet, il est possible de trouver des zones instables (*i.e.* convectives) au sein d'un nuage stratiforme précipitant ce qui donnera lieu à des averses plus fortes localement.

## **I-2 Comment mesure-t-on la pluie ?**

Cette seconde partie est consacrée à la description des principaux instruments de mesure. Nous distinguerons deux catégories : ceux qui mesurent directement la pluie et ceux qui déduisent des valeurs à partir des mesures. Nous décrirons ici brièvement le principe de fonctionnement de chaque instrument, les différents paramètres mesurés et les principales sources d'erreurs associées. Le lecteur intéressé trouvera d'avantage d'informations dans les références utilisées : Michaelides *et al.* (2009) ; Civate et Mandel (2008) ; Leroy (2000) ; Delobbe (2006) ; Parent du Châtelet (2003) ; Tridon (2011, chap. I).

### **I-2.1 Mesures directes des précipitations**

Ces types d'instruments sont classés selon qu'ils mesurent (Michaelides *et al.* 2009) :

- une quantité moyenne proportionnelle au volume intégré des précipitations, c'est le principe des pluviomètres ;
- un décompte et une caractérisation individuelle des particules précipitantes, c'est le principe des disdromètres.

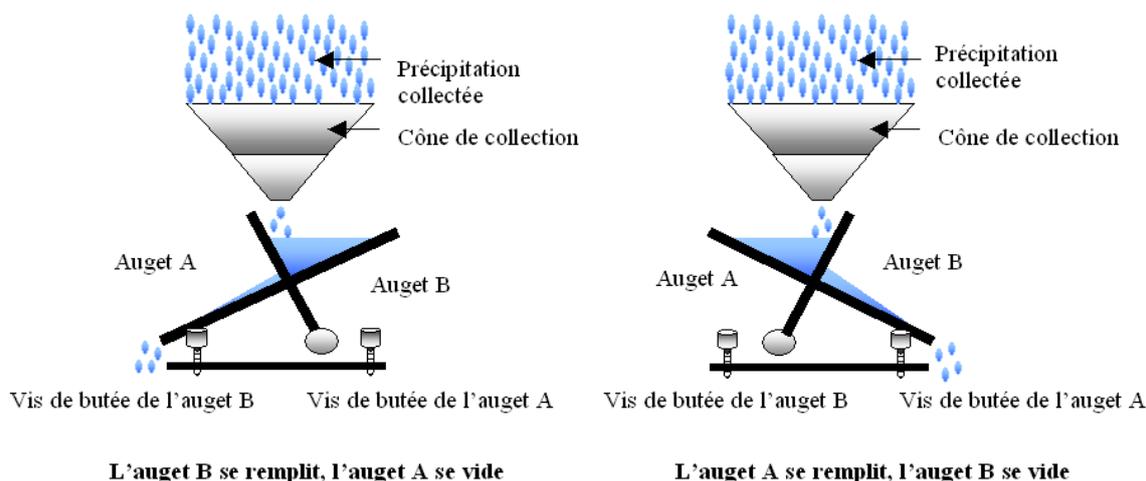


Figure I-1 : Schéma de principe des augets basculants.

Source : Civiata et Mandel (2008) © Météo-France.

### I-2.1.1 Mesure globale par pluviométrie

Le pluviomètre est l'instrument de mesure le plus utilisé notamment de part ses faibles coûts d'investissement. Aujourd'hui, le plus commun est le pluviomètre automatique à augets basculants.

#### a | Principe de fonctionnement des pluviomètres

Leur principe est simple : la pluie est recueillie par un collecteur terminé par un entonnoir qui dirige l'eau vers une pièce basculante composée de deux compartiments (ou augets). Un premier auget se remplit jusqu'à ce qu'une certaine quantité d'eau y soit accumulée (Figure I-1). Sous son poids, l'auget bascule pour se vider ce qui provoque la fermeture brève d'un circuit électronique. Dès lors, le deuxième auget se place en position de remplissage jusqu'au moment de se vider dès que la quantité limite est atteinte.

Les impulsions sont enregistrées par une station automatique et l'enregistrement continu permet de déterminer deux grandeurs :

- la hauteur de précipitation (exprimée en mm) définie comme l'épaisseur de la quantité d'eau liquide tombée sur une surface équivalente à 1 m<sup>2</sup> (1 litre m<sup>-2</sup> = 1 mm) et obtenue *via* le rapport du volume de l'auget sur la surface de collecte du pluviomètre ;
- le flux massique de pluie (ou intensité, exprimé en mm h<sup>-1</sup>) déduite de la fréquence des basculements sur une période de temps donnée et la contenance des augets.

Météo-France dispose d'un large réseau national de pluviomètres pourvus essentiellement d'un collecteur conique d'une surface réceptrice de 1 000 cm<sup>2</sup> et de deux

augets d'une contenance de 20 grammes d'eau (ou 20 cm<sup>3</sup>), soit une résolution de hauteur de précipitation de 0.2 mm. Le nombre de basculement est comptabilisé sur une période de temps de 6 minutes ou 1 heure. Il existe un second type de pluviomètres, moins conventionnel, avec un collecteur cylindrique d'une surface réduite à 400 cm<sup>2</sup>. Pour la même résolution de hauteur, les augets contiennent seulement 8 grammes d'eau. Notons que la taille de la surface collectrice n'a pas d'influence sur la mesure mais un pluviomètre avec une grande surface sera préféré pour la mesure des très forts événements pluvieux.

## b | Sources d'erreur associées aux mesures pluviométriques

Malgré tous leurs avantages, les pluviomètres ne sont pas exempts d'erreurs de mesure. L'Organisation Mondiale de la Météorologie (WMO 2012, p. II 22–24, II 35) dresse une liste non-exhaustive des erreurs possibles, accompagnée d'une estimation de l'erreur faite sur un basculement (liste complétée avec les incertitudes données dans Leroy (2000)) :

- déviation de la pluie par le vent, les turbulences créées par le collecteur ou les obstacles voisins (selon la force du vent, incertitude entre 0 et 30%) ;
- en cas de – très – forte pluie : perte d'eau lors du basculement des augets (selon le pluviomètre et l'intensité des pluies, incertitude entre 0.5 et 10%), décalage temporel dû à la limitation du débit à l'ajutage (incertitude nulle sur la quantité), rebondissement des augets (incertitude négligée) ;
- en cas de faible pluie ou intermittente : évaporation de l'eau sur les parois (incertitude inférieure à 0.5%) ou de l'eau contenue dans un auget en fin d'un événement et en attente du prochain ;
- adhérence de l'eau sur les parois du collecteur (incertitude inférieure à 0.5%), présence d'eau résiduelle dans les augets après leur basculement (incertitude négligée) ;
- attente de mouillage des augets (incertitude de 10% sur les premiers basculements) ;
- présence de précipitation solide (neige, grêle, etc.) si le pluviomètre ne dispose pas de systèmes chauffant ou présence de contaminants (feuilles d'arbres, poussières, excréments d'oiseaux, escargots, etc.) (incertitude pouvant aller jusqu'à 100%) ;
- localisation du pluviomètre dans son environnement : présence de pente, d'arbres ou bâtiments (incertitude arbitraire entre 0 et 30% et pouvant aller jusqu'à 100%) ;
- erreurs de nivellement des augets (incertitude inférieure à 3%) ou de non-horizontalité du pluviomètre (incertitude dépendante de l'angle faisant avec la verticale) ;

- erreurs propres au pluviomètre : asymétrie des augets (incertitude négligeable), détérioration et dérèglement du mécanisme (incertitude à long terme de  $\pm 5\%$ ).

La mesure d'un pluviomètre est connue à un basculement près suivant l'état de remplissage des augets au début mais aussi à la fin de la période de pluie. Une mesure est considérée significative s'il y a eu un décompte d'au moins trois basculements dans la période de temps considérée.

Hormis sa relative précision (estimée globalement à  $\pm 5\%$ ), l'inconvénient principal de cet instrument est la faible représentativité de sa mesure dans l'espace puisqu'il ne fournit qu'une mesure ponctuelle. Espacés de seulement quelques mètres, deux pluviomètres donnent des mesures équivalentes mais certainement pas identiques. Cette corrélation entre les pluviomètres diminue à mesure que la distance les séparant augmente. Pour observer le caractère hétérogène de la pluie, l'utilisation d'un réseau de pluviomètres est nécessaire et plus celui-ci est dense et bien agencé, meilleure est l'information résultante. La nature et la densité d'un réseau doit tenir compte des phénomènes locaux, de la précision désirée ou encore de la topographie. Le deuxième désavantage du pluviomètre est la nécessité d'une surveillance périodique et d'un entretien régulier. La comparaison des mesures avec des stations voisines permet de détecter toutes anomalies, ce constat justifie donc la mise en place d'un contrôle de qualité des données (comme celui décrit dans la partie II-1).

### *I-2.1.2 Mesure individuelle par disdrométrie*

Au cours de sa chute, une goutte de pluie accélère jusqu'à atteindre sa vitesse terminale de chute qui résulte de l'équilibre entre son poids (associé à son diamètre) et la résistance de l'air. Dans la littérature, plusieurs relations théoriques ont été proposées entre la vitesse terminale de chute et le diamètre des gouttes. De l'anglais « distribution drop meter », les disdromètres mesurent la distribution de chaque hydrométéore (classé en nombre selon leur taille et leur vitesse terminale) et en déduisent un taux de précipitation avec une meilleure résolution que les pluviomètres. Deux types d'instruments sont distingués :

- les disdromètres à impact qui mesurent individuellement le choc de chaque goutte sur une surface mobile exposée à la pluie. Lors de l'impact, l'énergie mécanique est convertie en impulsion électrique dont l'amplitude est proportionnelle à la vitesse des gouttes et leur diamètre. Le détail des conversions est propre à chaque instrument. Le dimensionnement de la surface détectrice est contraint par la limitation du nombre d'impacts par seconde (petite surface) et par l'échantillonnage des plus grosses gouttes (grande surface). En

outre, l'instrument doit également faire face à d'autres problèmes tels que des erreurs de calibration, des perturbations acoustiques dans l'environnement proche, un masquage des petites gouttes par les plus grosses, la présence de vent ou encore de flux turbulents, des vibrations produits lors de forts impacts ;

- les disdromètres optiques, plus récents, utilisent deux têtes, dont l'une transmet un faisceau laser plat, la seconde le recevant. Lorsque les hydrométéores (liquides ou solides) passent au travers de la zone d'échantillonnage, le degré d'atténuation du faisceau permet de déduire leur taille tandis que la durée d'extinction est reliée à leur vitesse de chute. Malgré la petite largeur du faisceau où finalement un petit nombre de gouttes est échantillonné, la résolution des mesures est supérieure à celle des pluviomètres (Tridon 2011, p. 25). La présence de vent ou de turbulences proches des détecteurs peut déformer la distribution des gouttes déduite des mesures. Il est également possible de voir apparaître un spectre secondaire mesuré à partir des gouttes ayant rebondi sur l'une des deux têtes.

Finalement, les pluviomètres et les disdromètres donnent une mesure ponctuelle mais précise de la pluie tombée au sol. La présence de vent ou de flux turbulents proches de la surface reste un problème majeur. Malgré les efforts menés, il n'existe pas encore de méthode permettant la correction de la sous-estimation occasionnée sur la quantité de précipitation pour les pluviomètres ou la déformation de la distribution de gouttes pour les disdromètres.

## I-2.2 Mesures indirectes des précipitations

Le terme RADAR est l'acronyme de l'expression anglaise « RAdio Detection et Ranging ». Basé sur le principe de télédétection active, ces appareils permettent de détecter et de localiser les pluies dans l'espace. Les sources d'erreurs associées aux mesures sont diverses, certaines affectent la mesure en elle-même, d'autres les estimations quantitatives des précipitations déduites de plusieurs relations et approximations.

En France, 80 % du territoire métropolitain est aujourd'hui couvert par un réseau de 26 RADAR opérationnels (site Internet : Météo-France et Ministère de l'Écologie et du Développement durable). Le traitement actuel des données collectées permet d'obtenir une carte des précipitations avec une résolution spatiale de l'ordre du kilomètre carré et une résolution temporelle de plusieurs minutes.

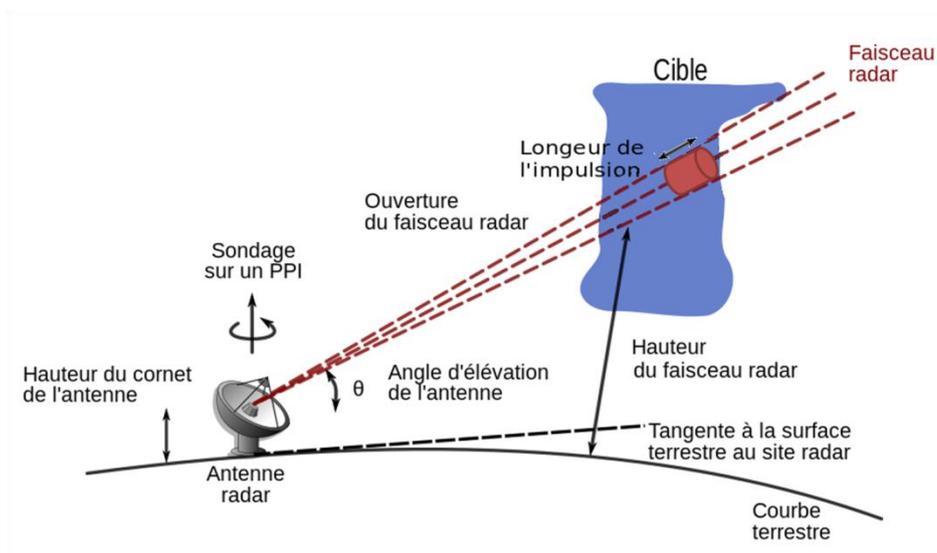


Figure I-2 : Fonctionnement général du RADAR météorologique en mode PPI.  
 Source : [http://fr.wikipedia.org/wiki/Radar\\_m%C3%A9t%C3%A9orologique](http://fr.wikipedia.org/wiki/Radar_m%C3%A9t%C3%A9orologique) (modifiée)

### I-2.2.1 Principe de fonctionnement du RADAR

Un RADAR est constitué de trois éléments principaux : un émetteur, un récepteur et une antenne. À intervalles réguliers, le système d'émission génère une impulsion électronique de forte puissance et haute fréquence pendant une très courte durée. L'antenne concentre les ondes électromagnétiques résultantes dans un faisceau (en forme de cône étroit) qui se propagent ensuite dans l'atmosphère à la vitesse de la lumière. Lorsqu'une onde rencontre une cible (ex : gouttes de pluie), une partie de l'énergie est diffusée dans toutes les directions dont celle de l'antenne qui capte alors un écho retour (ou écho RADAR). Le temps écoulé entre l'émission et la réception du signal informe sur la distance des précipitations par rapport au RADAR.

Il existe deux modes de balayage de l'antenne. Le plus commun est le mode PPI (de l'anglais « Plan Position Indicator ») où l'antenne pointe avec un certain angle d'élevation et tourne autour de l'axe azimutal en effectuant des rotations de 0 à 360° (Figure I-2). À l'inverse, en mode RHI (de l'anglais « Range Height Indicator »), l'angle azimutal est fixe mais l'antenne change d'angle d'élevation sondant l'atmosphère sur la verticale. Les RADAR dits volumétriques combinent ces deux modes : l'antenne effectue une rotation complète en PPI et ce pour différents angles d'élevation. Cette dernière stratégie permet d'obtenir un sondage en trois dimensions.

Concernant les récepteurs, deux principales catégories sont distinguées :

- les récepteurs dits incohérents qui mesurent uniquement l'amplitude du signal reçu donnant une estimation de l'intensité des précipitations ;
- les récepteurs cohérents (ou Doppler) qui déduisent en plus la vitesse radiale des cibles (et par déduction leur taille – pour les RADAR en visée verticale) à partir de la différence de phase du signal reçu.

Enfin, les nouveaux RADAR possédant un système de double polarisation apportent également l'information sur la forme relative des hydrométéores (analyse des signaux selon deux directions orthogonales : horizontale et verticale).

### I-2.2.2 L'équation RADAR

Soit une cible, considérée sphérique et située à une distance  $r_c$  par rapport au RADAR, intercepte sur une section  $S(r_c)$  le signal RADAR de puissance  $P_e$ . Si cette cible réémettait totalement cette énergie de manière isotrope (*i.e.* identique quelle que soit la direction), alors la puissance reçue par l'antenne RADAR (ayant un gain  $G$ ) pourrait s'écrire :

$$P_r(r_c) = \frac{P_e \cdot G^2 \cdot \lambda_{RADAR}^2}{(4\pi)^3 \cdot r_c^4} \cdot S(r_c)$$

En réalité, aucune cible météorologique ne diffuse les ondes électromagnétiques de manière isotrope. Pour s'affranchir de ce constat, on définit une section efficace de rétrodiffusion RADAR  $\sigma_e$  comme la surface d'une cible idéale qui renverrait au récepteur une puissance égale à celle effectivement reçue. En considérant que i) le volume d'échantillonnage  $\mathcal{V}$  correspond au produit de la section du faisceau (= ellipse de dimension horizontale  $\theta$  et verticale  $\phi$ ) par la longueur de l'impulsion (définie par la durée  $\tau$  d'une impulsion et la vitesse  $c$  de propagation de l'onde) et que ii) les cibles sont réparties uniformément dans ce volume  $\mathcal{V}$ , alors la section de rétrodiffusion de l'ensemble de ces cibles peut être décrite par :

$$\eta(r_c) = \frac{\sum_i \sigma_e(\text{cible}_i)}{\mathcal{V}}$$

Appelée réflectivité RADAR,  $\eta$  est reliée à la puissance du signal reçu  $P_r$  selon l'équation RADAR (Probert-Jones 1962) :

$$P_r(r_c) = \left[ \frac{P_e \cdot G^2 \cdot \lambda_{RADAR}^2 \cdot \theta \cdot \phi \cdot \tau \cdot c}{512 \cdot (2\ln 2) \cdot \pi^2} \right] \cdot \frac{\eta(r_c)}{r_c^2} \quad [I-1]$$

La partie entre crochets dépend des caractéristiques intrinsèques au RADAR, elle est appelée constante RADAR. Le terme  $(2 \ln 2)$  correspond à un facteur de forme assimilant le faisceau à une gaussienne.

Lorsqu'une onde électromagnétique interagit avec un corps dont les propriétés électriques diffèrent de celles du milieu traversé, une partie de l'énergie peut être absorbée par le corps et apparaît sous forme de chaleur. L'autre partie est diffusée dans toutes les directions. Dans le cas de la pluie (dont les particules sont considérées sphériques et de composition homogène), ce phénomène est décrit par l'application des équations de Maxwell à la diffusion d'une onde plane par une sphère homogène en milieu absorbant. La théorie de Mie donne une solution générale à ce problème (Tridon 2011, p. 31).

## a | De la puissance reçue au facteur de réflectivité RADAR

Ces dernières hypothèses permettent de simplifier la résolution de l'équation RADAR [I-1]. En considérant que la taille des cibles est très inférieure à la longueur d'onde  $\lambda_{RADAR}$  (approximation de Rayleigh), il est alors possible d'établir une relation entre la réflectivité  $\eta$  et une grandeur  $Z$ , appelée facteur de réflectivité, telle que :

$$\eta = \frac{\pi^5}{\lambda_{RADAR}^4} \cdot |K_w|^2 \cdot Z \quad [I-2]$$

La constante de normalisation de la réflectivité RADAR  $K_w$  dépend de la nature réfléchissante des cibles, de la longueur d'onde  $\lambda_{RADAR}$  et de la température du milieu. Sa valeur s'étend de 0.93 pour des gouttes d'eau liquide à 0.176 pour des particules de glace.

Le facteur de réflectivité RADAR  $Z$  est défini comme étant le moment d'ordre 6 de la distribution en taille des gouttes  $N$ , tel que :

$$Z = \int D^6 N(D) dD \quad [I-3]$$

où  $N(D)dD$  est le nombre de particules de diamètre entre  $D$  et  $D + dD$  par unité de volume d'air. Contrairement à la réflectivité  $\eta$  (équation [I-2]), le facteur de réflectivité  $Z$  est indépendant de la longueur d'onde  $\lambda_{RADAR}$ . Il est cependant tributaire des hydrométéores contenus dans le volume RADAR et plus leur nombre est élevé et leur diamètre grand, plus sa valeur sera importante. Les valeurs de  $Z$  (exprimées en  $\text{mm}^6 \text{m}^{-3}$ ) peuvent couvrir plusieurs ordres de grandeur. Ainsi pour faciliter sa lecture, une échelle logarithmique lui est associée, le décibel  $Z$  (ou dBZ), définie comme :

$$Z[\text{dBZ}] = 10 \log_{10} Z[\text{mm}^6 \text{m}^{-3}]$$

Les valeurs typiques de  $Z[\text{dBZ}]$  sont (Houze 1994, p. 112) : -30 à 0 pour des précipitations marginales à peine détectables ; 0 à 10 pour de la bruine, des très faibles pluies ou des faibles chutes de neige ; 10 à 30 pour des pluies modérées ou de fortes chutes de neige ; 30 à 45 pour

de la neige fondante ; 30 à 60 pour des pluies modérées à fortes et 60 à 70 ou plus pour de la grêle.

### b | Du facteur de réflectivité RADAR au taux de précipitation

La distribution en taille des gouttes  $N(D)$  permet également de déterminer le rapport de mélange des pluies  $q_r$  :

$$q_r = \frac{\pi \rho_l}{6 \rho} \int D^3 N(D) dD$$

où  $\rho_l$  est la densité de l'eau liquide et  $\rho$  la densité de l'air.

Le taux de précipitation  $\mathfrak{R}$  (exprimé en mm/h) est donné par :

$$\mathfrak{R} = \frac{\pi \rho_l}{6} \int v(D) D^3 N(D) dD \quad [\text{I-4}]$$

avec  $v(D)$  la vitesse de chute des gouttes de diamètre  $D$ . À partir des équations [I-3] et [I-4] nous remarquons qu'il existe un lien entre le facteur de réflectivité  $Z$  et le taux de précipitation  $\mathfrak{R}$ . Cette relation (dite «  $Z$ - $\mathfrak{R}$  ») peut prendre la forme :

$$Z = a \cdot \mathfrak{R}^b \quad [\text{I-5}]$$

où  $a$  et  $b$  sont des constantes positives.

### I-2.2.3 Sources d'erreurs associées aux mesures RADAR

Le RADAR est l'un des seuls systèmes d'observations permettant d'observer spatialement les précipitations en temps réel. Une estimation quantitative des mesures est affectée par différentes sources d'erreurs qui affectent à la fois les mesures directes et dérivées. Avant d'énumérer ces erreurs, il est important de rappeler l'influence de la longueur d'onde  $\lambda_{RADAR}$  sur les mesures RADAR.

En effet, l'interaction entre une onde électromagnétique et les hydrométéores dépend de la longueur d'onde (et donc de la fréquence d'opération) du RADAR. Or la longueur d'onde  $\lambda_{RADAR}$  influe sur trois paramètres (Tridon 2011, 33–34) :

- la réflectivité RADAR  $\eta$  (équation [I-2]) : plus  $\lambda_{RADAR}$  est courte, plus elle est sensible à des faibles facteurs de réflectivité  $Z$  et permet de détecter des fins hydrométéores ;
- le dimensionnement de l'antenne : pour garder un faible angle d'ouverture de l'antenne parabolique, plus la longueur d'onde est courte, plus le diamètre de l'antenne doit être petit. Par ailleurs, pour une même surface effective, plus  $\lambda_{RADAR}$  est courte, plus le gain de l'antenne est élevé et donc, pour une même quantité de puissance émise, la durée d'une impulsion peut être courte (impliquant une meilleure résolution) ;

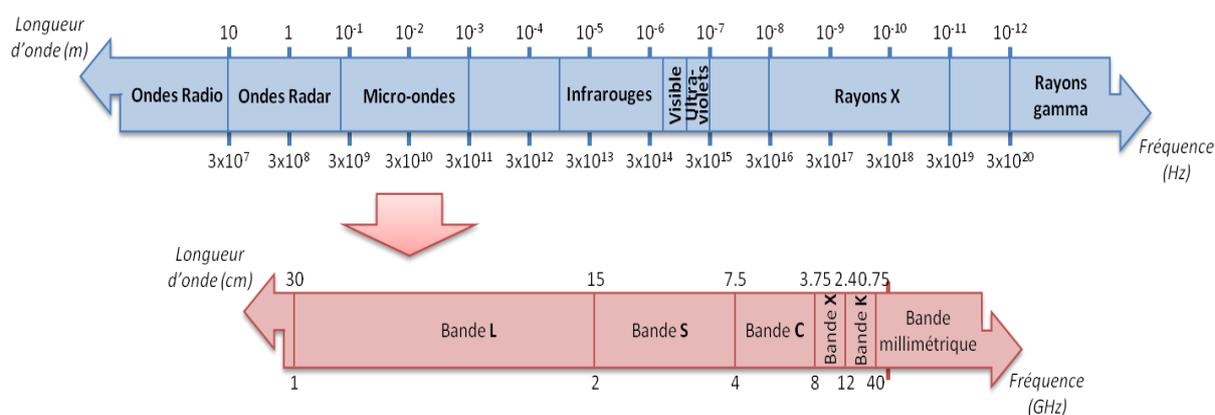


Figure I-3 : Gamme de longueurs d'onde et fréquences des RADAR météorologiques dans le spectre électromagnétique.

- l'atténuation : ce phénomène est conditionné par la capacité des hydrométéores à réfléchir et/ou absorber l'onde électromagnétique qui dépend de la longueur d'onde et du diamètre des gouttes. Ainsi, plus la longueur d'onde est courte, plus l'atténuation est importante.

En partie pour ces raisons, les RADAR météorologiques opèrent le plus souvent dans la gamme de fréquence des micro-ondes (*i.e.* longueurs d'onde centimétriques – Figure I-3). Les RADAR en bande S sont les RADAR à précipitation par excellence. Opérant avec une grande longueur d'onde, ils sont peu concernés par l'atténuation du signal et peuvent en conséquence disposer d'une très longue portée ( $\approx 500$  km). Leur gros désavantage est qu'ils ont besoin d'une antenne imposante et surtout coûteuse. À l'opposé, les RADAR en bande X fonctionnent sur des plus courtes longueurs d'onde. Ils sont plus petits et donc facilement transportables (sur une remorque par exemple). L'utilisation de courtes impulsions leur permet d'obtenir une très bonne résolution radiale, de l'ordre de plusieurs dizaines de mètres. En revanche, ils sont fortement affectés par les phénomènes d'atténuation et ont donc une portée nettement limitée (en dessous de 60 km). Enfin, les RADAR en bande C sont un bon compromis entre les deux fonctionnements déjà cités. Il en résulte des RADAR de taille moyenne avec une portée convenable (de l'ordre de la centaine de kilomètre) mais pour lesquels l'atténuation ne peut pas être négligée. L'utilisation de la bande X est préférée pour les observations locales (*i.e.* à très fine résolution radiale) notamment en région montagneuse, la bande C est plus couramment utilisée pour l'observation en région tempérée et la bande S est surtout réservée pour les régions à fort risque de pluies intenses (Parent du Châtelet 2003).

## a | Hypothèses et approximations

Lors de l'établissement de l'équation RADAR (équation [I-1]), une première hypothèse a été nécessaire. En effet, le facteur de réflectivité  $\eta$  est défini comme étant la somme de la section de rétrodiffusion de chaque cible divisée par le volume RADAR. Cette relation n'est possible uniquement que s'il est admis que les cibles sont uniformément réparties dans le volume d'échantillonnage. Cette hypothèse n'est donc plus valable si le volume sondé présente différents types d'hydrométéores simultanément ou s'il est inégalement rempli. Une solution à ce problème est de disposer d'un RADAR avec une très fine résolution radiale (ex : bande X).

Par ailleurs, lorsque les conditions de l'approximation de Rayleigh ne sont plus respectées (c'est-à-dire quand la taille des gouttes est supérieure à la longueur d'onde du RADAR), les erreurs faites sur le facteur de réflectivité  $Z$  sont de l'ordre de 1.5 à 2 dBZ (Planche 2011, p. 155). En utilisant un RADAR en visée verticale avec un système Doppler, la mesure de la taille des gouttes permet de s'assurer de la validité de cette approximation.

## b | Étalonage du système RADAR

L'étalonnage précis des systèmes d'émission et de réception est nécessaire pour une mesure correcte de l'écho RADAR. À savoir qu'un étalonnage parfait est difficilement réalisable. L'utilisation d'une cible dont la réflectivité et la distance sont connues avec précision (ex : un pylône ou montagne) permet par exemple de détecter des variations éventuelles dans la mesure (Delobbe 2006, p. 10). Ces mesures sont nécessaires pour déterminer le plus exactement possible la valeur de la constante RADAR qui conditionne la mesure de la réflectivité (cf. équation [I-1]). Un contrôle régulier des mesures RADAR assure la stabilité du niveau d'étalonnage et garantit la bonne répétitivité des mesures dans le temps.

## c | Atténuation

L'onde électromagnétique émise par le RADAR se propage dans l'atmosphère et rencontre sur son passage des molécules de gaz atmosphériques, des gouttelettes nuageuses, des gouttes de pluie ou des cristaux de glace. L'absorption et la diffusion de l'onde par ces particules atténuent son intensité. Ces effets d'atténuation ne peuvent pas être négligés dans le cas de grosses gouttes ou grêlons. Les secteurs situés à l'arrière de ces cibles (qui jouent alors un rôle d'écran) peuvent être occultés (Figure I-4, erreur 3). De nombreuses méthodes ont été développées pour évaluer et corriger ces phénomènes d'atténuation, la plus connue étant celle de Hitschfeld et Bordan (1954).

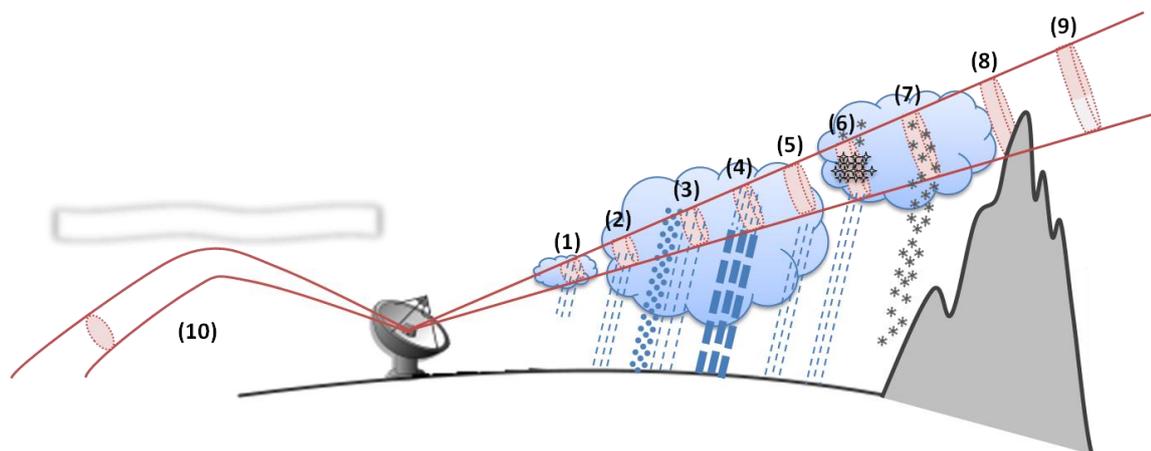


Figure I-4 : Principales sources d'erreur dans la mesure RADAR.

(1) évaporation ; (2) remplissage partiel ; (3) atténuation par grosses gouttes ; (4) grossissement par condensation ; (5) over-shooting ; (6) présence d'une bande brillante ; (7) sous-estimation dans la neige ; (8) échos de sol ; (9) blocage partiel ; (10) propagation anormale.

Le facteur de réflectivité RADAR  $Z$  dépend de la constante  $K_w$  (équation [I-2]) ainsi que du diamètre  $D$  des cibles et de leur distribution  $N(D)$  (équation [I-3]) et sa mesure. Entre un flocon de neige et une goutte de même masse, ces 3 variables diffèrent. En effet, le diamètre  $D$  d'un flocon est beaucoup plus grand que celui de la goutte, mais la constante  $K_w$  est beaucoup plus faible. En tombant plus lentement, la concentration des flocons est une plus grande concentration que les gouttes, mais celles-ci se peuvent se combiner par collision pour donner des plus grosses gouttes. Lorsque tous ces facteurs sont pris en compte, la différence de réflectivité RADAR est de 1 à 2 dB (Fabry et Zawadzki 1995) (Figure I-4, erreur 7).

La présence d'un radome installé pour protéger l'antenne contribue également à l'atténuation du signal. L'atténuation du radome sec est faible et connue, elle est prise en compte dans la constante RADAR. Cependant, lorsqu'il pleut, le ruissellement le long du radome peut engendrer une atténuation de plusieurs dBZ entraînant une sous-estimation du taux de précipitation sur tout le champ. Ce phénomène s'ajoute à l'atténuation produite par les pluies, il est donc difficile d'évaluer la contribution seule du radome humide.

#### d | Échos de sol, parasites et effet de masque

Pendant la phase d'émission, le rôle de l'antenne est de focaliser le signal en un faisceau conique. Ne pouvant pas être parfait, ce faisceau est constitué d'un faisceau principal, dirigé dans la direction souhaitée, et de lobes secondaires, de plus faible puissance, émis dans des directions transverses. Même en condition normales de propagation, toutes ces

ondes peuvent rencontrer des cibles au sol (qui va agir comme un miroir). Les résultants (appelés échos fixes ou échos de sol) contaminent alors l'image RADAR ponctuellement par de très fortes valeurs (Figure I-4, erreur 8). Une grande majorité de ces valeurs aberrantes peut être éliminée par l'application d'une carte d'échos de sol obtenue par beau temps. Il arrive cependant que des échos persistent. C'est souvent le cas lors de situations anticycloniques où les gradients verticaux de température et d'humidité bloquent toute propagation au-delà d'une certaine altitude contraignant le faisceau à se courber vers le bas (Figure I-4, erreur 10). Les effets de cette propagation anormale sont l'apparition de forts échos de sol inhabituels, plus difficiles à détecter et à corriger.

Les signaux renvoyés par les avions, oiseaux ou toute autre cible présente dans l'atmosphère apparaissent comme des échos parasites sur l'image RADAR. Ils sont facilement repérés et corrigés si le RADAR peut mesurer la vitesse des particules (*i.e.* récepteurs cohérents).

Enfin, la présence de bâtiments, d'arbres, d'éolienne (Agence Nationale des Fréquences 2005) ou de montagnes génère un effet d'occultation partielle ou totale de la région située à l'arrière de l'obstacle (Figure I-4, erreur 9). Cet effet de masque est d'autant plus significatif que l'obstacle est de grande dimension et est situé à proximité du RADAR. Les solutions sont alors variées : reconstitution de la mesure par une mesure à un angle d'élévation plus élevé, interpolation entre les cellules voisines ou application d'un facteur correcteur.

## e | Résolution spatiale ou radiale

Le signal reçu par le récepteur provient de l'ensemble des hydrométéores contenus dans un volume d'échantillonnage. Ce volume, assimilé à une portion de cône, est défini par i) la forme du faisceau qui s'élargit avec la distance au RADAR et ii) la résolution de l'impulsion qui est propre au système de réception. Avec la distance, ce volume croît. À grande distance, une particule fortement réfléchissante peut remplir que partiellement le faisceau RADAR, ce qui induit une sous-estimation de la mesure du facteur de réflectivité. Inversement, à courte distance, les mesures sont plus sensibles aux variations locales de la réflectivité associées aux structures plus fines du champ de précipitation.

## f | Altitude des mesures

Pour limiter les échos de sol, l'antenne est orientée selon un certain angle d'élévation. La rotondité de la Terre et la géométrie de la stratégie de sondage font que l'altitude des mesures augmente avec la distance au RADAR. À des altitudes élevées, les mesures sont de

moins représentatives des précipitations observées au sol. Plusieurs effets non désirés peuvent apparaître sur les mesures, tels que l'évaporation (partielle ou totale) des hydrométéores (Figure I-4, erreur 1), le grossissement des gouttes par condensation ou coalescence (Figure I-4, erreur 4), le remplissage partiel du faisceau où seule une partie du faisceau intercepte les précipitations (Figure I-4, erreur 2) ou l'over-shooting qui signifie que le faisceau RADAR pointant trop haut ne détecte pas les pluies qui se génèrent à plus basse altitude (Figure I-4, erreur 5). L'ensemble de ces effets obligent à réduire la portée du RADAR à des distances raisonnables. Par exemple, pour des applications hydrologiques où la portée est limitée à 80 ou 100 km alors que la portée pour la détection météorologique peut aller jusqu'à 300 km pour les phénomènes convectifs très étendus en altitude (Parent du Châtelet 2003).

### g | Profil vertical de la réflectivité et bande brillante

Au sein d'un nuage, la réflectivité RADAR peut être extrêmement forte. Cette zone spécifique, appelée bande brillante (Figure I-4, erreur 6), correspond à l'altitude de l'isotherme 0°C où la phase glace commence à fondre. Sur une certaine distance verticale, les précipitations se constituent de particules glacées dont la couche superficielle devient liquide. L'eau liquide étant plus réfléchive, le RADAR observe donc une couche d'altitude avec de fortes valeurs de réflectivité. En mode PPI, ce phénomène fait apparaître une structure annulaire autour du RADAR.

Une des méthodes de correction de ces effets d'altitude est l'estimation des profils verticaux de réflectivité obtenus soit par des modèles théoriques (Kirstetter 2008, chap. 3; Andrieu et Creutin 1995) soit à partir des mesures RADAR à visée verticale ou volumétrique. Ces profils peuvent être ainsi utilisés pour extrapoler les réflectivités mesurées en altitude vers la surface.

### h | Relation $Z$ - $\mathfrak{R}$

Dans le sous-paragraphe I-2.2.2-b nous avons vu qu'il est possible d'estimer le taux de précipitation  $\mathfrak{R}$  à partir du facteur de réflectivité  $Z$ . Or il n'existe pas de la relation universelle qui lie ces deux grandeurs. En effet, les caractères très dépendant de  $Z$  et de  $\mathfrak{R}$  à la distribution en taille des gouttes  $N(D)$  (inconnue et extrêmement variable spatialement et temporellement) oblige à faire davantage d'hypothèses augmentant l'erreur sur l'estimation finale. La caractérisation de  $N(D)$  a fait et fait encore l'objet de multiples recherches pour proposer la meilleure relation  $Z$ - $\mathfrak{R}$ . La conversion du facteur de réflectivité  $Z$  en taux de précipitation  $\mathfrak{R}$  peut être envisagée grâce à :

- l'utilisation d'une relation théorique qui satisfait la situation météorologique ; la plus connue étant celle de Marshall et Palmer (1948) où les constantes  $a$  et  $b$  sont respectivement fixées à 200 et 1.6 pour une pluie de type stratiforme ;
- le recalage des données du RADAR avec des données pluviométriques, ce qui nécessite alors un réseau suffisamment dense pour s'assurer d'une statistique significative.

Une autre approche serait basée sur l'utilisation d'un RADAR disposant d'un système polarisant permettant d'estimer le diamètre des gouttes en tout points. Toutes les erreurs sur l'hypothèse de la distribution des précipitations peuvent être ainsi limitées mais des incertitudes (liées à mesure des paramètres de polarisation) viennent s'ajouter sur l'estimation des pluies. L'utilisation d'une loi  $Z-R$  est donc source indirecte d'erreurs de part le choix porté sur la relation utilisée pour un événement pluvieux

La mesure de la pluie par télédétection RADAR est une technique très puissante mais souffre encore de problèmes parfois non négligeables. De nombreuses solutions ont été proposées pour diminuer ou corriger certains effets nuisibles à la mesure. Les avancées technologiques sont porteuses d'atouts quant à l'estimation quantitative des précipitations comme par exemple l'utilisation de la technique Doppler, le sondage volumétrique ou la diversité de polarisation. Au sol ou embarqués, les RADAR constituent aujourd'hui l'instrument de référence dans la mesure des pluies, ce malgré les incertitudes que cela suppose.

Au travers de ce chapitre, nous montrons que les deux types de mesures (pluviomètres/disdromètres et RADAR) sont complémentaires : l'un mesurant la quantité ponctuelle des pluies, l'autre donnant leur répartition spatiale. Nous nous proposons dans le chapitre suivant de trouver la meilleure façon de combiner ces deux informations pour obtenir une estimation du champ de pluie au sol la plus proche de la réalité.

---

## Les Estimateurs utilisés dans le Chapitre II

◆ *En statistique :*

Notons  $V_i$  la valeur observée et  $\tilde{V}_i$  la valeur estimée à la station  $i$ . Considérons  $N$  stations sur l'ensemble du domaine étudié, alors :

- l'erreur quadratique moyenne (MSE) mesure la moyenne des carrés des écarts entre les estimations et les observations. Elle est plus souvent exprimée sous la forme d'une racine carré (RMSE), telle que :

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (\tilde{V}_i - V_i)^2}$$

La RMSE est particulièrement appréciée pour détecter la présence de larges écarts, surtout si les écarts se compensent. Elle n'indique cependant pas la direction des écarts.

- le coefficient de corrélation (spatial) renseigne sur la relation linéaire (ou phase) entre le champ des valeurs estimées et celui des valeurs observées. Il est donnée par :

$$r = \frac{\sum_{i=1}^N (\tilde{V}_i - \bar{\tilde{V}}) \cdot (V_i - \bar{V})}{\sqrt{\sum_{i=1}^N (\tilde{V}_i - \bar{\tilde{V}})^2} \cdot \sqrt{\sum_{i=1}^N (V_i - \bar{V})^2}}$$

avec  $\bar{V}$  et  $\bar{\tilde{V}}$  les moyennes des valeurs observées et estimées. Un coefficient de corrélation nul indique que les deux champs n'ont aucun lien entre eux. Plus sa valeur s'éloigne de zéro (*i.e.* en s'approchant de 1 ou -1), meilleure est la corrélation (positive = en phase ou négative = en opposition de phase).

- l'intervalle de confiance (IC) permet d'apprécier le degré de précision de l'estimation d'un paramètre statistique sur l'échantillon. Il définit une marge d'erreurs aléatoires qui contient avec une forte probabilité la vraie valeur du paramètre estimé. L'IC autour d'une moyenne est défini par :

$$\text{IC} = \left[ \bar{V} - t_\varepsilon \frac{\sqrt{\frac{1}{N} \sum_{i=1}^N (V_i - \bar{V})^2}}{\sqrt{N}} ; \bar{V} + t_\varepsilon \frac{\sqrt{\frac{1}{N} \sum_{i=1}^N (V_i - \bar{V})^2}}{\sqrt{N}} \right]$$

avec  $\bar{V}$  la moyenne des valeurs observées et  $t_\varepsilon$  un quantile de la loi Student défini par le niveau de confiance  $\varepsilon$  souhaité.

◆ *En probabilité :*

L'espérance mathématique d'une variable aléatoire (notée  $E$ ) correspond à la moyenne des valeurs possibles pondérées par les probabilités associées à ces valeurs. C'est l'équivalent de la moyenne arithmétique d'une série statistique.

La variance (notée  $\text{Var}$ ) représente la moyenne des carrés des écarts à l'espérance de la variable aléatoire considérée. Elle permet de caractériser la dispersion des valeurs par rapport à la moyenne.

---

# Chapitre II

## Reconstruction d'un champ de pluie : méthodologie

Qu'il s'agisse de la densité d'un minéral dans des échantillons de sol, de la concentration d'un polluant dans l'atmosphère autour d'un site industriel à risque ou des précipitations mesurées au sol, ces observations ponctuelles sont de plus en plus souvent exploitées sous forme de champs spatiaux. De nombreuses méthodes statistiques ont été mises au point et développées dans l'espoir de reproduire au plus proche de la réalité le caractère spatial des variables étudiées. Certaines méthodes sont aujourd'hui largement utilisées en raison des excellents résultats obtenus.

Ce chapitre se focalise sur la reconstruction d'un champ de précipitation. Pour ce faire, nous avons choisi d'utiliser deux techniques qui permettent de combiner les observations. Leur description est précédée d'une partie nécessaire à la compréhension de leur fonctionnement général et est suivie d'une discussion expliquant nos motivations sur le choix de ces techniques. Avant cela, il est important de noter que toutes ces méthodes d'interpolation spatiale requièrent au préalable des données dont la fiabilité est avérée. Cette vérification est réalisée grâce à un contrôle de qualité de données. Nous introduirons ce deuxième chapitre par la description du critère que nous avons instauré dans nos traitements.

### II-1 Contrôle des données pluviométriques

Comme mentionné dans le premier chapitre, les instruments sont sujets à diverses erreurs de mesures. Toutefois, certaines d'entre elles peuvent être détectées *via* un ensemble de procédures visant à tester la validité des données, appelé contrôle de qualité (ou QC pour l'acronyme anglais de « Quality Control »). Sa mise en place peut s'avérer difficile, surtout pour les précipitations qui présentent un caractère très hétérogène dans l'espace et le temps. Nous nous limiterons ici au QC des observations pluviométriques.

Un exemple simple de QC est la mise en place d'un test vérifiant que les mesures de précipitation d'une station ne sont pas aberrantes. Ainsi, toutes valeurs de pluie négatives ou

valeurs dépassant un certain seuil (ex : 40mm/min comme suggéré par Zahumenský (2004)) seraient éliminées. L'étape supérieure est la comparaison des données avec une ou plusieurs stations voisines. Dans la littérature, nous pouvons trouver différentes procédures qui intègrent cette notion d'espace. Chaque QC est évidemment contestable, avec leurs parts d'atouts et d'inconvénients. Celui intégré dans nos traitements est basé sur les travaux de Hubbard (2000). Nous allons décrire ce QC dans sa version initiale avant de commenter les modifications que nous avons apportées.

## II-1.1 Version initiale

Le QC décrit par Hubbard (2000) se décompose en trois étapes. Les deux premières permettent de sélectionner, dans un voisinage de dix stations (notées  $j$ ), les cinq stations les plus représentatives des mesures de la station à tester (notée  $i$ ). La dernière étape vérifie alors la validité de chaque observation individuelle de la station  $i$  (notée  $V_i$ ) par rapport aux observations respectives des stations  $j$  sélectionnées (notée  $V_j$ ).

La première étape sélectionne parmi toutes les stations  $j: \{j \neq i\}$  disponibles, les dix plus proches stations (notées de 0 à 9) (Figure II-1). Pour les dix couples  $ij$  retenus, le coefficient de corrélation (noté  $r_{ij}^k$ ) entre la série temporelle de la station  $i$  (notée  $v_i$ ) et la série temporelle d'une station  $j$  (notée  $v_j$ ) est calculé pour trois configurations : un cas où  $v_i$  et  $v_j$  sont prises au même instant ( $k = 0$ ), un cas où  $v_j$  est retardée d'un certain laps de temps par rapport à  $v_i$  ( $k = -1$ ) et enfin un cas où  $v_j$  est avancée du même laps de temps par rapport à  $v_i$  ( $k = 1$ ). L'application de ce décalage temporel  $k$  permet de prendre en considération i) le déplacement des précipitations dans le temps tel que deux stations distantes ne débutent pas leurs mesures au même moment et ii) les erreurs éventuelles de synchronisation dans le datage des observations.

La seconde étape consiste à déterminer une régression linéaire entre les séries  $v_i$  et  $v_j$  pour chacun des trente cas (= 10 couples de stations  $ij$  pour 3 décalages  $k$ ) telle que :

$$\tilde{V}_{ij}^k = a_{ij}^k + b_{ij}^k \cdot V_j^k$$

où  $\tilde{V}_{ij}^k$  est une approximation linéaire de l'observation  $V_i$  par rapport à l'observation  $V_j$  en tenant compte du décalage temporel  $k$ . Pour chaque cas, la RMSE (notée  $RMSE_{ij}^k$ ) est calculée entre les observations  $V_i$  et les approximations  $\tilde{V}_{ij}^k$ . Une seconde sélection est alors faite pour ne garder que les cas  $ij^k$  pour lesquels le coefficient de corrélation  $r_{ij}^k$  est arbitrairement supérieur à 30%. Une troisième sélection ne conserve alors plus que cinq

couples  $ij^{k'}$  dont le décalage  $k':\{k' = -1, 0 \text{ ou } 1\}$  donne les plus faibles écarts (*i.e.* plus petites valeurs de  $RMSE_{ij}^{k'}$ ). Notons que dans le cas où une station  $i$  ne dispose pas exactement de cinq couples  $ij^{k'}$  alors aucune vérification n'est faite et la procédure continue pour une autre station.

La dernière étape consiste à contrôler individuellement toutes les observations  $V_i$  de la station  $i$  par comparaison avec celles des cinq stations alors retenues. Une observation est alors considérée comme « douteuse » si, pour les cinq couples  $ij^{k'}$ , sa valeur n'est pas comprise dans l'intervalle défini par :

$$\tilde{V}_{ij}^{k'} - F \cdot RMSE_{ij}^{k'} < V_i < \tilde{V}_{ij}^{k'} + F \cdot RMSE_{ij}^{k'}$$

La valeur du facteur  $F$  définit les limites de l'intervalle de confiance souhaité. Dans notre cas, l'utilisation d'une valeur de 5 est équivalente à un intervalle de confiance à 99.99994%.

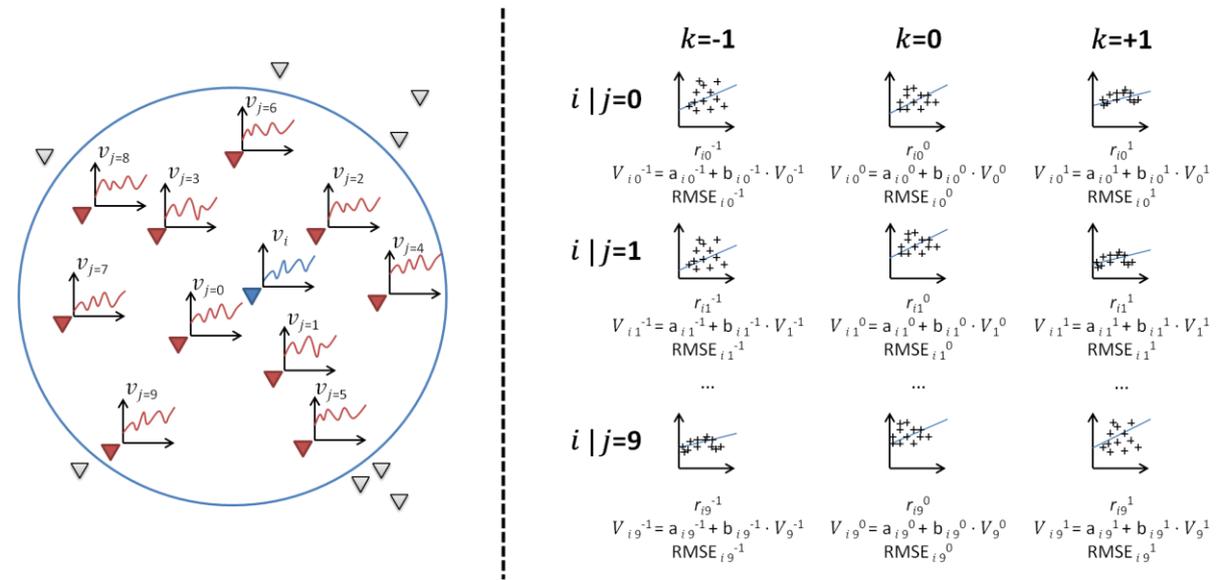


Figure II-1 : Principe du contrôle de qualité.

À droite : le principe repose sur la comparaison de la série temporelle d'une station à contrôler (en bleu, notée  $v_i$ ) avec celles des dix plus proches stations (en rouge, notées  $v_j:\{j = 0, \dots, 9, j \neq i\}$ ). Les stations les plus loin (en gris) ne sont pas prises en compte dans le QC. Un décalage temporel  $k:\{-1, 0, 1\}$  sera appliqué sur les séries temporelles  $v_j$ . À gauche : pour chaque couple  $ij^k$ , le coefficient de corrélation ( $r_{ij}^k$ ) est calculé ainsi que la régression linéaire  $\tilde{V}_{ij}^k = a_{ij}^k + b_{ij}^k \cdot V_j^k$  et la RMSE associées ( $RMSE_{ij}^k$ ).

Si la condition n'est pas respectée, nous définissons alors une valeur qui correspondrait à la valeur  $V_i$  que nous espérons avoir pour que celle-ci ne soit pas considérée comme douteuse. Cette valeur, dite « attendue », est donnée par l'expression :

$$V_i^{\text{attendue}} = \frac{\sum_j \left( (RMSE_{ij}^{k'})^2 \cdot \tilde{V}_{ij}^{k'} \right)}{\sum_j \left( (RMSE_{ij}^{k'})^2 \right)} \quad [\text{II-1}]$$

## II-1.2 Modifications apportées

La première modification apportée à ce QC se situe au niveau de la première sélection des trente cas  $ij^k$ . Basé sur les travaux de Feng *et al.* (2004), nous avons inclus un critère supplémentaire à savoir que les coefficients de corrélation  $r_{ij}^k$  doivent être statistiquement significatifs, c'est-à-dire que les différents couples  $ij^k$ , les  $r_{ij}^k$  doivent être compris dans l'intervalle de confiance à 95%. Un problème est alors survenu. En effet, pour les stations situées en bordure de domaine, la condition d'avoir exactement cinq couples  $ij^{k'}$  n'était pas toujours obtenue. Pour s'assurer que ces stations soient quand même contrôlées, nous avons décidé de réduire la condition à trois couples  $ij^{k'}$  minimum, en dessous de quoi, le contrôle de qualité n'est pas appliqué.

La seconde modification ajoutée concerne la détermination de la régression linéaire  $\tilde{V}_{ij}^k$ . Afin de s'assurer de la meilleure représentativité des observations, il est préférable, statistiquement parlant, d'utiliser de longues séries temporelles (*i.e.* plusieurs mois voire années consécutifs). En prenant en compte le caractère hétérogène et intermittent des précipitations, il est toutefois recommandé de différencier les mesures selon les phénomènes observés. Par exemple, Feng *et al.* (2004) ou Hubbard *et al.* (2005) ont appliqué de tels QC sur des cumuls journaliers, sur 49 et 30 ans respectivement et en comparant les données mois par mois. Ceci permet de ne pas exclure les rares événements extrêmes (peut être typiques à une certaine période de l'année) en moyennant l'ensemble des données. Ainsi par exemple, les quelques très forts orages d'été sont séparés des longues et faibles pluies stratiformes d'hiver plus récurrentes. En travaillant de la sorte, l'estimation de la régression linéaire est améliorée et le QC devient donc moins sélectif. Dans notre cas, nous avons opté pour une application sur un événement, c'est-à-dire que l'estimation de la droite de régression est propre à chaque épisode pluvieux étudié.

You *et al.* (2007) dénoncent l'utilisation de ce QC à cause d'un trop grand nombre de données alors considérées comme suspicieuses. Les auteurs estiment que cinq sixième de ces observations sont en réalité correctes. La mise en place de ce QC s'est néanmoins avérée bénéfique (comme nous le verrons dans les paragraphes III-3.2, III-4.2, IV-4.2 et IV-5.2). Ce QC modifié nous a permis d'écarter de façon automatique quelques données que nous aurions suspectées par un contrôle visuel de chaque série temporelle considérée.

La qualité des données étant validée, la mise en place d'une interpolation spatiale peut être réalisée dans les meilleures conditions.

## **II-2 Reconstruction d'un champ spatial à partir de données ponctuelles**

En théorie, tout phénomène naturel peut être mesuré en tout point de l'espace. Pour des raisons principalement économiques, les observations proviennent de seulement quelques sites de mesures ponctuels, souvent dispersés et inégalement répartis. Pour connaître la valeur de la variable étudiée dans une région dépourvue d'instruments, il est commun d'avoir recours à une technique d'interpolation spatiale. Il en existe plusieurs et sont souvent divisées en deux classes que nous décrivons brièvement ci-dessous.

La première catégorie regroupe les méthodes qui reposent sur des priorités purement mathématiques (ou géométriques) et ne prennent pas en compte les caractéristiques physiques du phénomène observé. La plus simple de ces méthodes (= méthode du plus proches voisin) consiste à partitionner l'espace en polygones de Thiessen ou triangles de Delaunay et à attribuer à chaque surface définie une valeur unique correspondant à la mesure du seul site qu'elle comprend. Pour les méthodes barycentriques (ou moyennes pondérées), la valeur en un point est estimée à partir d'une combinaison linéaire des valeurs des stations les plus proches et dont les poids attribués accordent plus d'influence aux stations les plus proches (ex : inverse de la distance). Ce premier groupe inclut également d'autres méthodes plus compliquées telles que l'utilisation de surface de tendance ou de fonctions splines. N'étant pas le sujet de nos travaux, ces méthodes ne seront pas détaillées dans ce manuscrit.

Dans le deuxième groupe, les méthodes stochastiques, ou plus usuellement géostatistiques, font appel à des modèles probabilistes qui découlent de l'analyse statistique des mesures. À l'origine, le géologue sud africain D.G. Krige (1919 – 2013) proposa une méthode statistique pour estimer la teneur en or d'un bloc de minerai à partir d'échantillons

extraits autour. Dans les années 1960, le français G. Matheron (1930 – 2000) reprit cette méthode et développa un outil qui permet d'analyser la continuité spatiale des teneurs. Il mit au point une méthode d'estimation qu'il appela « krigeage ». Aujourd'hui, cette méthode n'est plus seulement utilisée à des fins géologiques, mais également dans des domaines d'applications tels que la météorologie, l'océanographie, la santé, le génie civil ou encore la télécommunication.

En comparaison avec les méthodes barycentriques, le krigeage se différencie par la prise en compte de la structure spatiale des données. Son développement actuel permet d'intégrer différentes informations auxiliaires améliorant considérablement ses performances. D'un point de vue statistique, le krigeage est considéré comme étant la plus juste des méthodes d'interpolation spatiale.

Notre objectif est de reconstruire un champ de précipitation le plus proche possible de la réalité. Nous avons pour cela sélectionné deux techniques différentes de krigeage qui seront appliquées sur des observations pluviométriques et mesures RADAR. Dans ce qui suit, nous décrirons dans un premier paragraphe, la théorie générale du krigeage. Nous verrons que la mise en œuvre de cette méthode passe par une étape cruciale d'analyse des données destinée à décrire la structure spatiale du phénomène observé. Au travers d'une technique simple de krigeage, nous détaillerons la résolution du système d'équations associé. Le deuxième paragraphe porte sur les récentes techniques de krigeage combinant plusieurs variables et avec un focus plus particulier sur les deux techniques choisies. Enfin, nous discuterons des performances prometteuses rapportées dans plusieurs études inter-comparatives.

## II-2.1 Le krigeage : théorie générale

Le modèle de base du krigeage permet de décrire mathématiquement un phénomène naturel (*i.e.* aléatoire), homogène et continu spatialement dont une quantité  $V$  (ex : cumul de pluie) est mesurée en différents points de coordonnées  $\mathbf{s}$  dans un domaine spatial fini  $\mathcal{D}$  et qui peut être définie par :

$$V(\mathbf{s}) = \mu(\mathbf{s}) + \delta(\mathbf{s}) \quad \mathbf{s} \in D \quad [\text{II-2}]$$

où  $\mu(\mathbf{s})$  est une structure déterministe pour l'espérance de  $V(\mathbf{s})$  et  $\delta(\mathbf{s})$  est une fonction aléatoire stationnaire, d'espérance nulle et de structure de dépendance connue. Autrement dit,  $\mu(\mathbf{s})$  correspond au processus principal de la quantité  $V$  observable à grande échelle, tandis que  $\delta(\mathbf{s})$  est assimilée aux fluctuations de  $V$  à des échelles plus fines et qui comprend également les erreurs de mesures.

La formulation de  $\mu(\mathbf{s})$  définit le type de krigeage utilisé. Les trois formes les plus usuelles sont : le krigeage ordinaire ( $\mu =$  constante inconnue), le krigeage simple ( $\mu =$  constante connue) et le krigeage universel ( $\mu =$  combinaison linéaire fonction des localisations  $\mathbf{s}$ ).

La structure de la fonction  $\delta(\mathbf{s})$  n'est pas directement mesurable mais est déterminée à partir des données par l'intermédiaire de l'analyse variographique (ou détermination du variogramme expérimental). Cette étape importante permet de décrire la variabilité spatiale du phénomène observé au travers d'un modèle mathématique (ou détermination du variogramme paramétrique).

### II-2.1.1 L'analyse variographique

Le modèle sur lequel se base le krigeage suppose que la structure de dépendance spatiale de la fonction aléatoire  $\delta(\mathbf{s})$  soit connue, ce qui en pratique ne l'est pas. Une estimation peut être toutefois faite *via* une étude du comportement spatial des observations. Cette étape préalable est appelée analyse variographique.

Tout phénomène naturel est par nature unique, ce qui suppose une seule et unique estimation de  $\delta(\mathbf{s})$ . Ainsi pour rendre possible le krigeage, une hypothèse de stationnarité doit être émise, comme l'hypothèse de stationnarité intrinsèque impliquant que l'espérance (notée E) et la variance (notée Var) de la fonction aléatoire  $\delta(\mathbf{s})$  s'écrivent respectivement :

$$\begin{aligned} E[\delta(\mathbf{s}_1) - \delta(\mathbf{s}_2)] &= 0 \\ \text{Var}[\delta(\mathbf{s}_1) - \delta(\mathbf{s}_2)] &= 2\gamma(\mathbf{s}_1 - \mathbf{s}_2) = 2\gamma(\mathbf{h}) \end{aligned} \quad \mathbf{s}_1, \mathbf{s}_2 \in D \quad \text{[II-3]}$$

où  $\mathbf{h}$  représente le vecteur de distance qui sépare respectivement deux stations  $\mathbf{s}_1$  et  $\mathbf{s}_2$ . La quantité  $2\gamma(\mathbf{h})$  est nommée variogramme.

#### a | Attributs d'un variogramme

Un variogramme décrit la dépendance des valeurs de quantité  $2\gamma$  par rapport à la distance  $\mathbf{h}$ . Il est défini par trois grandeurs (Figure II-2) :

- l'effet de pépité ( $c_0$ ) qui correspond au comportement des données lorsque  $\mathbf{h}$  s'approche de 0. Ce caractère peut être continu ( $c_0 = 0$ ) ou discontinu ( $c_0 \neq 0$ ). Son existence peut être en partie expliquée par des variations non-mesurables à micro-échelle. L'inclusion d'un tel effet est difficile à déterminer, surtout si peu de stations sont proches les unes des autres ;

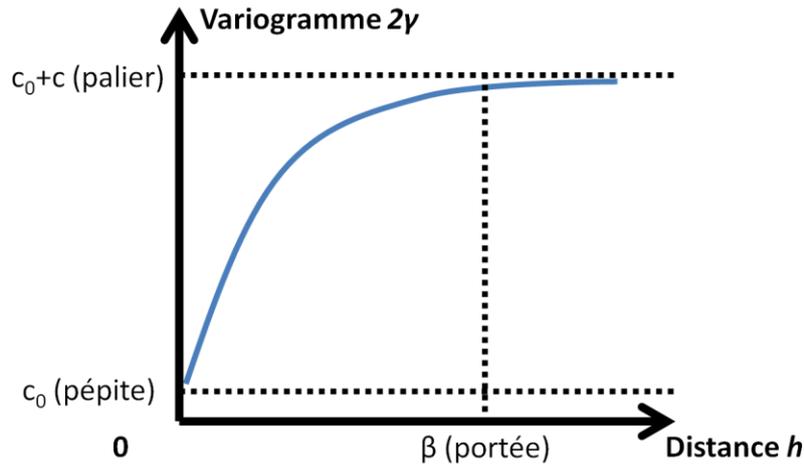


Figure II-2 : Attributs d'un variogramme.

- le palier ( $c_0 + c$ ), exact ou asymptotique, qui correspond à la valeur pour laquelle de  $2\gamma$  présente un plateau ;
- la portée ( $\beta$ ), exacte ou effective (*i.e.* 95% du palier atteint), qui correspond à la distance à partir de laquelle il n'y a plus de dépendance spatiale observée entre les données (*i.e.* pour laquelle le palier est atteint).

À noter qu'un variogramme qui ne dépend que de la norme de  $\mathbf{h}$  (ou distance euclidienne) est dit « isotrope ». À l'inverse, lorsqu'un phénomène s'organise selon une orientation privilégiée, il est alors préférable de considérer cet angle dans le calcul de  $\mathbf{h}$ . Le variogramme alors obtenu est dit « anisotrope ».

Par ailleurs, l'analyse du variogramme permet par exemple de faire une critique (ou QC) des données pluviométriques afin d'éliminer les mesures présentant des écarts anormaux avec des stations voisines (Kirstetter *et al.* 2010).

## b | Estimation du variogramme expérimental

Le variogramme décrit précédemment peut être réécrit comme :

$$\begin{aligned}
 2\gamma(\mathbf{s}_1 - \mathbf{s}_2) &= \text{Var}[\delta(\mathbf{s}_1) - \delta(\mathbf{s}_2)] \\
 &= \text{Var}[V(\mathbf{s}_1) - V(\mathbf{s}_2)] \text{ car } \mu \text{ n'est pas une fonction aléatoire} \\
 &= E\left[(V(\mathbf{s}_1) - V(\mathbf{s}_2))^2\right] - E[(V(\mathbf{s}_1) - V(\mathbf{s}_2))]^2 \\
 &= E\left[(V(\mathbf{s}_1) - V(\mathbf{s}_2))^2\right] - (\mu(\mathbf{s}_1) - \mu(\mathbf{s}_2))^2
 \end{aligned}$$

Dans le cas où  $\mu$  est une fonction constante (*i.e.* krigeage ordinaire ou krigeage simple), le deuxième terme s'annule et le variogramme peut directement être déduit des  $N$  observations. En revanche, lorsque  $\mu$  n'est pas une constante (*i.e.* krigeage universel), l'estimation du variogramme est rendue possible à partir des résidus  $e(\mathbf{s}_i) = V(\mathbf{s}_i) - \mu(\mathbf{s}_i)$ .

Le variogramme expérimental est calculé *via* un estimateur  $2\hat{\gamma}$ , dont le plus classique est défini par :

$$2\hat{\gamma}(\mathbf{h}) = \frac{1}{N} \sum_{i,j \neq i}^N [V(\mathbf{s}_i) - V(\mathbf{s}_j)]^2 \quad \text{ou} \quad 2\hat{\gamma}(\mathbf{h}) = \frac{1}{N} \sum_{i,j \neq i}^N [e(\mathbf{s}_i) - e(\mathbf{s}_j)]^2 \quad \mathbf{s}_i, \mathbf{s}_j \in \mathcal{D}$$

Cet opérateur est supposé non-biaisé mais la présence d'un carré peut fausser son estimation en présence de données erronées. Ce problème a incité certains auteurs à utiliser d'autres estimateurs, plus robustes tels que ceux proposés par Cressie (1993, p. 74) ou Li et Lake (1994). Tobin *et al.* (2011) précisent que ces derniers estimateurs conduisent à une meilleure estimation de  $\tilde{V}$  notamment pour les réseaux où les stations sont irrégulièrement réparties dans l'espace.

### c | Modélisation du variogramme expérimental

De par la nature aléatoire des observations, le variogramme expérimental  $\hat{\gamma}$  ne peut être utilisé directement. En pratique, il est utilisé une liste prédéfinie de modèles de semi-variogramme  $\tilde{\gamma}$ , dits « paramétriques » qui répondent aux exigences de la méthode du krigeage et dont les plus courants sont décrits mathématiquement ci-dessous. Comme illustré dans la Figure II-3, chacun se différencie par leur comportement proche de l'origine, leur structure générale et la façon dont est atteint le palier (s'il en possède un).

Le modèle **sphérique** (Figure II-3a), d'effet de pépité  $c_0$ , de palier  $c_0 + c$  et de portée  $\beta$  est décrit par la fonction :

$$\tilde{\gamma}(\mathbf{h}) = \begin{cases} c_0 & \mathbf{h} = 0 \\ c_0 + c \left[ \frac{3}{2} \left( \frac{\mathbf{h}}{\beta} \right) - \frac{1}{2} \left( \frac{\mathbf{h}}{\beta} \right)^2 \right] & 0 < \mathbf{h} \leq \beta \\ c_0 + c & \mathbf{h} > \beta \end{cases}$$

Le modèle **exponentiel** (Figure II-3b) est défini par un effet de pépité  $c_0$ , un palier  $c_0 + c$  et une portée effective de  $3\beta$  tel que :

$$\tilde{\gamma}(\mathbf{h}) = \begin{cases} c_0 & \mathbf{h} = 0 \\ c_0 + c \left[ 1 - \exp\left(-\frac{\mathbf{h}}{\beta}\right) \right] & \mathbf{h} > 0 \end{cases}$$

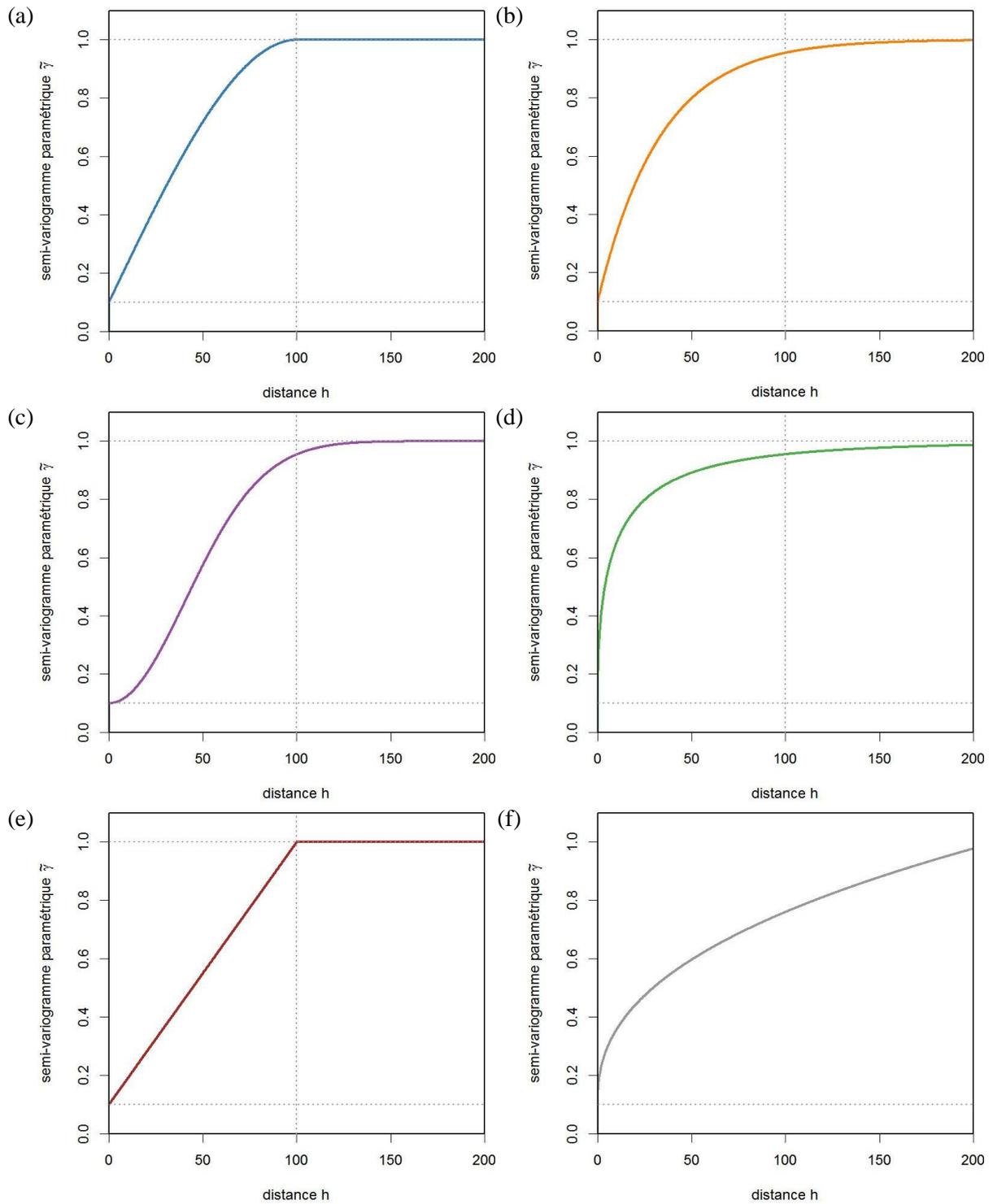


Figure II-3 : Exemples de modèle de semi-variogramme (effet de pépite = 0.1).  
 (a) sphérique, (b) exponentiel, (c) gaussien, (d) stable, (e) linéaire avec palier (portée = 100 ; palier = 1) et (f) puissance (exposant = 0.4 ; facteur d'échelle = 0.1).

Sur la même base, le modèle **gaussien** (Figure II-3c) se différencie par sa portée effective de  $\sqrt{3}\beta$  :

$$\tilde{\gamma}(\mathbf{h}) = \begin{cases} c_0 & \mathbf{h} = 0 \\ c_0 + c \left[ 1 - \exp\left(-\frac{\mathbf{h}^2}{\beta^2}\right) \right] & \mathbf{h} > 0 \end{cases}$$

tandis que le modèle **stable** (Figure II-3d) possède une portée effective de  $9\beta$  :

$$\tilde{\gamma}(\mathbf{h}) = \begin{cases} c_0 & \mathbf{h} = 0 \\ c_0 + c \left[ 1 - \exp\left(-\sqrt{\frac{\mathbf{h}}{\beta}}\right) \right] & \mathbf{h} > 0 \end{cases}$$

Toute combinaison de modèles est également possible mais demande une analyse plus approfondie. Un exemple simple est donné par la Figure II-3f avec un modèle **linéaire avec palier** :

$$\tilde{\gamma}(\mathbf{h}) = \begin{cases} c_0 + \frac{c}{\beta} \mathbf{h} & \mathbf{h} \geq 0 \text{ et } \mathbf{h} < \beta : \text{modèle linéaire} \\ c_0 + c & \mathbf{h} \geq \beta : \text{modèle palier} \end{cases}$$

Contrairement aux modèles cités précédemment, le modèle **puissance** ne possède pas de palier. Il se définit par un effet de pépite  $c_0$ , un exposant  $\nu$  et un facteur d'échelle  $m$  (Figure II-3e) selon la formule :

$$\tilde{\gamma}(\mathbf{h}) = \begin{cases} c_0 & \mathbf{h} = 0 \\ c_0 + m\mathbf{h}^\nu & \mathbf{h} \geq 0 \text{ et } 0 \leq \nu < 2 \end{cases}$$

#### d | Détermination du variogramme paramétrique

La détermination du variogramme paramétrique  $\tilde{\gamma}$  ne se limite pas seulement à sélectionner un modèle mathématique qui décrit le mieux l'allure du variogramme expérimental  $\hat{\gamma}$  puisqu'il faut également estimer les trois paramètres associés (*i.e.* effet de pépite, palier et portée). Cette identification peut se faire intuitivement avec un œil d'expert ou alors être réalisé à l'aide d'une méthode d'estimation. La plus courante est celle des moindres carrés qui cherche à minimiser l'erreur quadratique entre la quantité  $V$  et son estimation  $\tilde{V}$  en certains points de validation  $\mathbf{s}_k$  issus soit d'observations collectées à partir d'un second réseau de mesures indépendant, soit par validation croisée. Cette dernière technique consiste à créer deux échantillons à partir du jeu de données initial dont l'un est utilisé pour l'interpolation et le second pour la validation. Cette procédure est employée lorsque le nombre de données disponibles est suffisamment important (*i.e.* un grand nombre de stations). Dans le cas contraire, la technique de validation croisée dite « leave-one-out » (ou LOO) consiste à soustraire une à une les observations. L'interpolation est faite à partir des

stations restantes et l'estimation est ensuite comparée à la valeur soutirée. L'opération est ensuite répétée autant de fois que nécessaire pour que la totalité des observations soit testée.

Les procédures de détermination du variogramme paramétrique sont nombreuses. Pour les utilisateurs expérimentés, une automatisation est souvent proscrite. En effet, l'emploi de telles « boîtes noires » peut conduire à des estimations négatives de  $V$  (peu recommandé pour des cumuls de pluie par exemple). Elles permettent cependant de tester plusieurs modèles pour choisir alors le plus optimal.

En résumé, il n'existe donc pas de méthode universelle pour déterminer le variogramme paramétrique, mais celle-ci doit être choisie de façon à être la plus optimale selon l'application souhaitée.

## e | Discussion

Comme décrit précédemment, la détermination du variogramme paramétrique (dits « 1D ») se fait communément *via* la distance  $h$  séparant deux stations. Mais ce n'est pas une méthode unique. Par exemple, Lebel *et al.* (1987) et Berne *et al.* (2004) emploient un deuxième paramètre pour la détermination du variogramme paramétrique (2D) : la fenêtre temporelle  $T$  (*i.e.* durée de cumul des données). Kebaili Bargaoui et Chebbi (2009) utilisent l'intensité des précipitations comme troisième paramètre, donnant alors un caractère unique au variogramme paramétrique (3D).

Haberlandt (2007) ou Verworn et Haberlandt (2010) démontrent que la détermination du variogramme paramétrique 1D n'a que peu d'influence sur les performances du krigeage tant que, comme le souligne Erdin (2009, p. 55), les modèles mathématiques utilisés représentent suffisamment correctement la structure du phénomène observé. Ils montrent également que l'utilisation d'un variogramme spécifique à un événement ne montre pas d'amélioration justifiant sa mise en place. Par ailleurs, Tobin *et al.* (2011) montrent que, quelle que soit la méthode géostatistique employée, les estimations obtenues à l'aide d'un variogramme anisotrope (quand cela est légitimé) reproduisent correctement de la structure spatiale des pluies. En effet certains détails ayant une orientation préférentielle peuvent être manqués avec un variogramme isotrope classique.

### II-2.1.2 Krigeage des données

Quelle que soit la technique du krigeage employée, l'estimation de  $V$  (notée  $\tilde{V}$ ) au point souhaité de coordonnées  $\mathbf{s}_0$  prendra toujours la forme d'une combinaison linéaire des  $N$  observations ponctuelles de coordonnées  $\mathbf{s}_i$  telle que :

$$\tilde{V}(\mathbf{s}_0) = a + \sum_{i=1}^N \lambda_i V(\mathbf{s}_i) \quad \mathbf{s}, \mathbf{s}_0 \in D \quad [\text{II-4}]$$

Seule variante entre les différentes techniques, les poids  $\lambda_i$  et la constante  $a$  qui sont déterminés de façon à ce que i) l'espérance et la variance de krigeage<sup>⊕</sup> existent, ii) l'estimation est exempte de biais (soit une espérance nulle) et iii) la variance de krigeage est minimale. En général, le krigeage est par définition une prédiction linéaire, non biaisée et de variance minimale.

Isaaks et Srivastava (1989, 301–313) informent que le choix des valeurs des paramètres du variogramme paramétrique influe sur les résultats du krigeage (estimation et/ou variance). Les auteurs montrent que :

- la valeur du palier n'affecte pas l'estimation du krigeage mais la variance de krigeage. Celle-ci varie alors proportionnellement au changement de la valeur du palier (ex : une augmentation d'un facteur 3 du palier triple la variance de krigeage) ;
- le choix du modèle et plus particulièrement sa forme au voisinage de l'origine est un facteur d'erreur non négligeable. Par exemple, pour un phénomène extrêmement continu (ex : couches géologiques, champ gravimétrique), il sera préféré un modèle avec une forme parabolique (ex : modèle gaussien). Cependant en accordant plus d'importance aux stations les plus proches, il est possible de former un effet d'écran qui peut entraîner l'apparition de poids  $\lambda_i$  négatifs (et donc des estimations négatives) pour les stations les plus lointaines ;
- l'effet de pépite joue sur la distribution des poids  $\lambda_i$  et par conséquent sur l'estimation et la variance de krigeage ;
- la portée a un effet mineur sur l'attribution des poids  $\lambda_i$  mais entraîne un changement notable sur l'estimation et la variance de krigeage.

### II-2.1.3 Exemple du krigeage ordinaire

Dans le cadre d'un krigeage ordinaire (ou OK pour l'acronyme anglais de « Ordinary Kriging »), la quantité  $\mu$  de l'équation [II-2] est considérée comme une constante inconnue. Sous l'hypothèse de stationnarité intrinsèque de la fonction  $\delta$  (cf. équation [II-3]), la contrainte de non-biais implique pour l'équation [II-4] que :

---

<sup>⊕</sup> C'est un indicateur de la précision de l'estimation réalisée, quantifiant la dispersion possible de la valeur vraie, mais inconnue, autour de la valeur estimée. Il est dépendant du modèle géostatistique utilisé, par exemple du variogramme, et de la répartition des données mais pas de leurs valeurs. *Source : Wikipédia.*

$$a = 0 \quad \text{et} \quad \sum_{i=1}^N \lambda_i^{OK} = 1$$

et que la variance de krigeage est alors définie par :

$$\begin{aligned} \text{Var}_{OK}(\mathbf{s}_0) &= \text{E} \left[ \left( \tilde{V}_{OK}(\mathbf{s}_0) - V(\mathbf{s}_0) \right)^2 \right] & \mathbf{s}_0 \in \mathcal{D} \\ &= 2 \sum_{i=1}^N \lambda_i^{OK} \tilde{\gamma}(\mathbf{s}_0 - \mathbf{s}_i) - \sum_{i=1}^N \sum_{j=1}^N \lambda_i^{OK} \lambda_j^{OK} \tilde{\gamma}(\mathbf{s}_i - \mathbf{s}_j) & \mathbf{s}_0, \mathbf{s}_i, \mathbf{s}_j \in \mathcal{D} \end{aligned}$$

où  $\tilde{\gamma}(\mathbf{s}_i - \mathbf{s}_j)$  est la valeur déduite du semi-variogramme paramétrique  $\tilde{\gamma}$  pour la distance entre les points  $\mathbf{s}_i$  et  $\mathbf{s}_j$  et où  $\tilde{\gamma}(\mathbf{s}_0 - \mathbf{s}_i)$  est la valeur calculée pour la distance entre le point souhaité  $\mathbf{s}_0$  et le point  $\mathbf{s}_i$ .

L'étape suivante est de trouver la valeur des poids  $\lambda_i^{OK}$  de façon à minimiser la variance de krigeage  $\text{Var}_{OK}$  tout en sachant que leur somme vaut 1. Pour faciliter la résolution d'un tel système, une fonction auxiliaire  $f(\lambda_i^{OK}, \ell)$ , contenant la variance de krigeage et un multiplicateur de Lagrange  $\ell$ , est définie :

$$f(\lambda_i^{OK}, \ell) = \text{Var}_{OK}(\mathbf{s}_0) - 2\ell \left( \sum_{i=1}^N \lambda_i^{OK} - 1 \right) \quad \mathbf{s}_0 \in \mathcal{D} \quad [\text{II-5}]$$

La minimisation de [II-5] revient à trouver les solutions pour lesquelles les dérivés partielles par rapport à  $\lambda_i^{OK}$  et  $\ell$  s'annulent :

$$\begin{cases} \frac{\partial([\text{II5}])}{\partial \lambda_i^{OK}} = 0 \\ \frac{\partial([\text{II5}])}{\partial \ell} = 0 \end{cases} = \begin{cases} \sum_{i=1}^N \lambda_i^{OK} \tilde{\gamma}(\mathbf{s}_i - \mathbf{s}_j) + \ell = \tilde{\gamma}(\mathbf{s}_0 - \mathbf{s}_i) \quad \forall j = 1, \dots, N \\ \sum_{i=1}^N \lambda_i^{OK} = 1 \end{cases} \quad [\text{II-6}]$$

La solution du système [II-6] de  $(N + 1)$  équations à  $(N + 1)$  inconnues permet de déterminer les  $N$  poids  $\lambda_i^{OK}$ . L'estimation de  $V$  au point en  $\mathbf{s}_0$  est alors donnée par :

$$\tilde{V}_{OK}(\mathbf{s}_0) = \sum_{i=1}^N \lambda_i^{OK} V(\mathbf{s}_i) \quad \mathbf{s}_0, \mathbf{s}_i \in \mathcal{D}$$

pour laquelle la variance de krigeage est :

$$\text{Var}_{OK}(\mathbf{s}_0) = \sum_{i=1}^N \lambda_i^{OK} \tilde{\gamma}(\mathbf{s}_0 - \mathbf{s}_i) + \ell \quad \mathbf{s}_0, \mathbf{s}_i \in \mathcal{D}$$

Afin de faciliter la réalisation du krigeage (et son insertion dans nos programmes), il est préférable de passer sous forme matriciel. Le système d'équations [II-6] est alors réduit à :

$$\begin{aligned}
 \mathbf{A}\boldsymbol{\lambda}^{OK} &= \mathbf{B}, \quad \text{avec} \\
 \mathbf{A}_{N+1,N+1} &= \begin{bmatrix} \tilde{\gamma}(\mathbf{s}_1 - \mathbf{s}_1) & \cdots & \tilde{\gamma}(\mathbf{s}_1 - \mathbf{s}_N) & 1 \\ \vdots & \vdots & \vdots & \vdots \\ \tilde{\gamma}(\mathbf{s}_N - \mathbf{s}_1) & \cdots & \tilde{\gamma}(\mathbf{s}_N - \mathbf{s}_N) & 1 \\ 1 & \cdots & 1 & 0 \end{bmatrix}, \\
 \boldsymbol{\lambda}_{N+1,1}^{OK} &= \begin{bmatrix} \lambda_1^{OK} \\ \vdots \\ \lambda_N^{OK} \\ \ell \end{bmatrix} \quad \text{et} \quad \mathbf{B}_{N+1,1} = \begin{bmatrix} \tilde{\gamma}(\mathbf{s}_0 - \mathbf{s}_1) \\ \vdots \\ \tilde{\gamma}(\mathbf{s}_0 - \mathbf{s}_N) \\ 1 \end{bmatrix}
 \end{aligned}$$

La matrice  $\mathbf{A}$  doit être inversible telle que les poids  $\lambda_i^{OK}$  et  $\ell$  sont obtenus par :

$$\boldsymbol{\lambda}^{OK} = \mathbf{A}^{-1}\mathbf{B} \quad [\text{II-7}]$$

La variance de krigeage est donnée par :

$$\text{Var}_{OK}(\mathbf{s}_0) = \mathbf{B}^T \boldsymbol{\lambda}^{OK} \quad \mathbf{s}_0 \in \mathcal{D} \quad [\text{II-8}]$$

En résumé, le krigeage consiste en une estimation linéaire sans biais et à variance minimale. Parmi toutes les techniques associées, aucune ne ressort cependant comme étant universellement meilleure. Les performances de chacune d'elles dépendent de nombreux facteurs (nature des données, détermination des paramètres du variogramme, etc.). Mais, l'avantage majeur de ce type d'interpolation est le calcul simultané de l'estimation et des erreurs associées. L'incertitude calculée n'inclut cependant pas ni l'erreur faite lors de la détermination du variogramme paramétrique ni les erreurs de mesures.

## II-2.2 La combinaison des données par krigeage

Dans le cadre de nos travaux, nous nous sommes focalisés sur l'estimation de la pluie. En interpolant seulement les observations ponctuelles des pluviomètres, le champ spatial résultant est rarement proche de la réalité. Les mesures RADAR permettent de mieux appréhender le caractère hétérogène des précipitations malgré d'importantes erreurs d'estimation. De nombreuses méthodes ont alors été développées pour combiner ces deux types de mesures. Ehret (2003, p. 82) les ordonnent en quatre catégories :

- les méthodes qui ajustent les données RADAR par un ou plusieurs facteurs spécifiques pour obtenir une meilleure correspondance avec les données pluviométriques (ex : Brandes (1975)) ;
- les méthodes qui utilisent les données pluviométriques pour l'améliorer les estimations des mesures RADAR (ex : optimisation de la loi Z- $\mathfrak{R}$ ) ;

- les méthodes utilisant conjointement les données par des méthodes géostatistiques ;
- toutes les autres méthodes (ex : réseaux de neurones, filtres de Kalman, classification par type de temps comme le propose Gottardi *et al.* (2012)).

Parmi l'ensemble des ces méthodes, nous nous sommes focalisées sur les méthodes géostatistiques et plus particulièrement sur deux techniques : le krigeage avec dérive externe et la fusion conditionnée. Toutes deux sont décrites ici.

### II-2.2.1 Le krigeage avec dérive externe

#### a | Théorie générale

Le krigeage avec modèle de tendance ou krigeage avec dérive externe (ou KED pour l'acronyme anglais de « Kriging with External Drift ») est employé lorsque la variable  $\mu$  de l'équation [II-2] est une combinaison de  $K$  variables auxiliaires, notées  $Y_k: \{k = 1, \dots, K\}$ . Ces données dites « externes » sont reliées linéairement à  $V$  et leur valeurs sont connues à la fois aux points  $\mathbf{s}_i$  et au point souhaité  $\mathbf{s}_0$ . La structure  $\mu$  est alors définie par :

$$\mu(\mathbf{s}_i) = \sum_{k=0}^K \omega_k Y_k(\mathbf{s}_i) \quad \mathbf{s}_i \in \mathcal{D}$$

$$V(\mathbf{s}) = [\omega_0 Y_0(\mathbf{s}) + \omega_1 Y_1(\mathbf{s}) + \dots + \omega_K Y_K(\mathbf{s})] + \delta(\mathbf{s}) \quad \mathbf{s} \in \mathcal{D}$$

où les  $\omega_k$  sont des coefficients inconnus qui devront être déterminés. Notons que si  $K = 0$  et  $Y_0(\mathbf{s}) \equiv 1$ , alors  $\mu(\mathbf{s}) = \omega_0$  est une constante inconnue et le système obtenu est celui du krigeage ordinaire présenté précédemment. Nos travaux ont été menés avec  $K = 1$  variable externe pour laquelle les  $Y_1(\mathbf{s}_i)$  sont les mesures RADAR extraites aux points  $\mathbf{s}_i$  considérés.

Tout comme le OK, la résolution du système du KED n'est possible qu'avec l'hypothèse de stationnarité intrinsèque de la fonction  $\delta$  et sous la contrainte de non-biais (cf. équation [I-3]). Le système final est alors formé de  $(N + 1 + K = 1)$  équations à  $(N + 1 + K = 1)$  inconnues. Pour simplifier sa résolution, il est souvent supposé que  $Y_0(\cdot) = 1$ , ce qui permet de faire un lien avec le système du krigeage ordinaire. Sous sa forme matricielle, le système de krigeage du KED est donné comme suit :

$$\mathbf{A}\boldsymbol{\lambda}^{KED} = \mathbf{B}, \quad \text{avec}$$

$$\mathbf{A}_{N+1+K=1, N+1+K=1} = \begin{bmatrix} \tilde{\gamma}(s_1 - s_1) & \cdots & \tilde{\gamma}(s_1 - s_n) & 1 \\ \vdots & \vdots & \vdots & \vdots \\ \tilde{\gamma}(s_N - s_1) & \cdots & \tilde{\gamma}(s_N - s_N) & 1 \\ 1 & \cdots & 1 & 0 \\ [Y_1(s_1) & \cdots & Y_1(s_N) & 0] \end{bmatrix} \begin{bmatrix} Y_1(s_1) \\ \vdots \\ Y_1(s_N) \\ 0 \\ [0] \end{bmatrix},$$

$$\boldsymbol{\lambda}_{N+1+K=1,1}^{KED} = \begin{bmatrix} \lambda_1^{KED} \\ \vdots \\ \lambda_N^{KED} \\ \ell \\ \ell_1 \end{bmatrix} \quad \text{et} \quad \mathbf{B}_{N+1+K=1,1} = \begin{bmatrix} \tilde{\gamma}(s_0 - s_1) \\ \vdots \\ \tilde{\gamma}(s_0 - s_N) \\ 1 \\ Y_1(s_0) \end{bmatrix}$$

Les parties grisées s'apparentent au système matriciel du krigeage ordinaire présenté précédemment. La matrice  $\mathbf{A}$  doit être une matrice inversible de sorte que les poids  $\lambda_i^{KED}$  soient obtenus par :

$$\boldsymbol{\lambda}^{KED} = \mathbf{A}^{-1} \mathbf{B} \quad [\text{II-9}]$$

La variance de krigeage est alors donnée par :

$$\text{Var}_{KED}(\mathbf{s}_0) = \mathbf{B}^T \boldsymbol{\lambda}^{KED} \quad \mathbf{s}_0 \in \mathcal{D} \quad [\text{II-10}]$$

## b | Discussion

Notons que pour certaines variables  $Y_k$ , la relation avec  $V$  est parfois plus complexe qu'une simple relation linéaire. Dans de pareils cas, il faudrait effectuer une transformation des données externes pour rendre cette relation linéaire. Verworn et Haberlandt (2010) préconisent alors l'utilisation d'une transformation logarithme pour les données de réflectivité RADAR. Seulement, après avoir mis en œuvre le krigeage sur les données transformées, une transformation inverse doit être réalisée et un facteur correctif doit être appliqué pour assurer le non-biais des prévisions. Ce facteur dépend du choix de la transformation, des paramètres employés et du type de krigeage utilisé (Baillargeon 2005, p. 44).

Le KED est considéré comme une excellente technique d'interpolation spatiale mais l'estimation du variogramme paramétrique est un point déterminant dans son utilisation. En effet,  $\mu$  n'étant plus une constante, le variogramme expérimental  $\hat{\gamma}$  doit être estimé à partir des résidus  $e(\mathbf{s}_i)$ . Cette tâche peut s'avérer très difficile car  $\mu$  n'est en pratique pas connue et doit être estimée au préalable (Webster et Oliver 2008, p. 205). De nombreux auteurs ont toutefois tenté de trouver une solution. Par exemple, Cressie (1993, p. 166) propose une estimation des  $\omega_k$  à l'aide d'une succession de techniques des moindres carrés. Pour Velasco-Forero *et al.* (2009), la tendance est estimée en tous points  $\mathbf{s}_0$  en interpolant (par krigeage ordinaire) les valeurs en  $\mathbf{s}_i$  d'une variable secondaire et les résidus sont déduits de la soustraction entre le champ de l'estimation de la tendance et le champ initial de la variable secondaire.

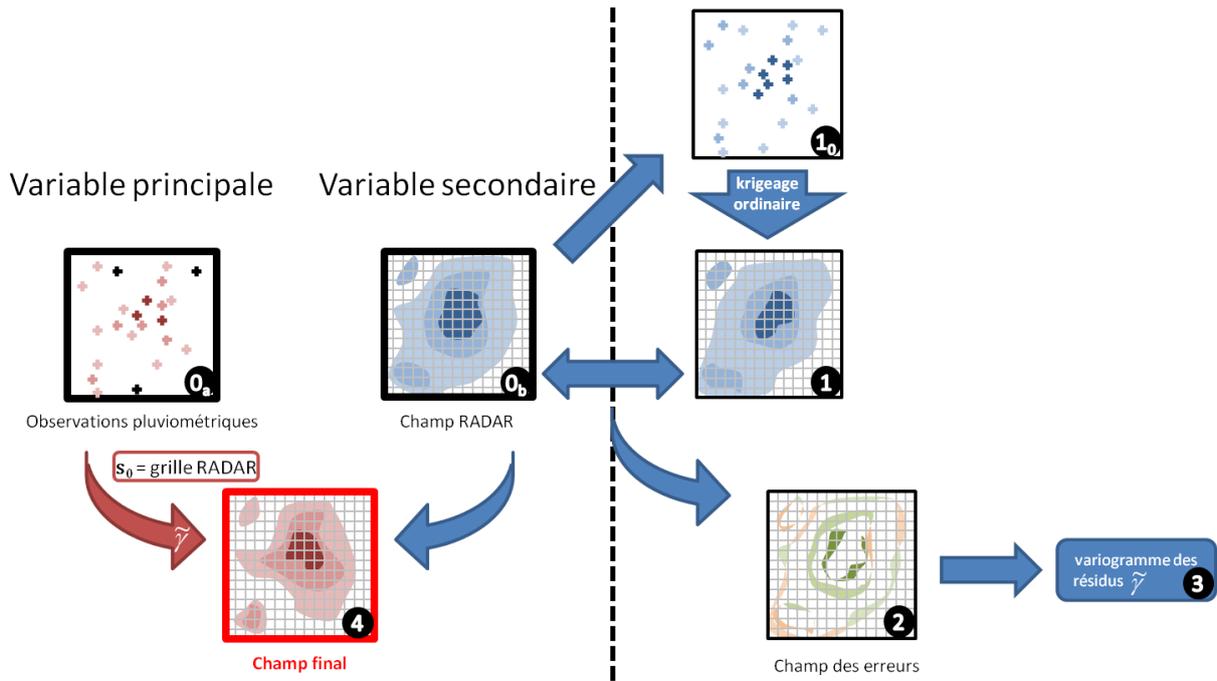


Figure II-4 : Illustration des étapes du KED selon Velasco-Forero *et al.* (2009).  
 $(0_a)$  = observations pluviométriques ;  $(0_b)$  = mesures RADAR ;  $(1_0)$  = pixels RADAR  $(0_b)$  aux plus proches des localisations des pluviomètres  $(0_a)$  ; (1) = interpolation par krigeage ordinaire des pixels RADAR  $(1_0)$  ; (2) = champs des erreurs entre l'interpolation des données RADAR (1) et les données RADAR  $(0_b)$  ; (3) recherche du variogramme paramétrique des résidus ; (4) interpolation des données pluviométriques  $(0_a)$ .

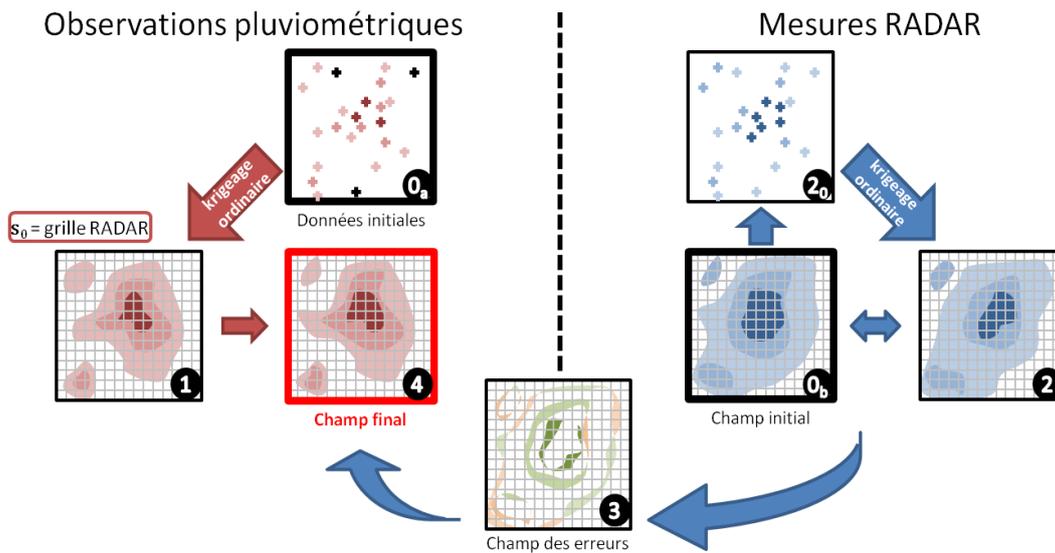


Figure II-5 : Illustration des étapes du MERG.  
 $(0_a)$  = observations pluviométriques ;  $(0_b)$  = mesures RADAR ; (1) = interpolation par OK des observations pluviométriques  $(0_a)$  ;  $(2_0)$  = pixels RADAR  $(0_b)$  aux plus proches des localisations des pluviomètres  $(0_a)$  ; (2) = interpolation par krigeage ordinaire des pixels RADAR  $(2_0)$  ; (3) = champs des erreurs entre l'interpolation des données RADAR (2) et les données RADAR  $(0_b)$  ; (4) = correction du champ interpolé issu des données pluviométriques (1) par le champ des erreurs (3).

Schématiquement, la technique du KED selon Velasco-Forero *et al.* (2009) peut se décomposer en 4 étapes (Figure II-4). À partir de l'image RADAR ( $\mathbf{0}_b$ ), les valeurs des pixels aux points des pluviomètres sont récupérées ( $\mathbf{1}_0$ ) puis interpolées sur la grille RADAR par krigeage ordinaire (1). La différence entre le champ RADAR observé ( $\mathbf{0}_b$ ) et celui interpolé (1) donne le champ des erreurs (2) qui sera utilisé pour déterminer le variogramme paramétrique des résidus (3). Le champ final (4) est obtenu par la résolution du système d'équations en considérant les observations pluviométriques ( $\mathbf{0}_a$ ) comme variable principale, les mesures RADAR ( $\mathbf{0}_b$ ) comme variable secondaire et le variogramme des résidus (3) comme fonction de structure.

### II-2.2.2 La fusion conditionnée

#### a | Théorie générale

Comme le KED, la technique de fusion conditionnée (ou MERG pour l'acronyme anglais de « conditional MERGing ») est classée dans la catégorie des techniques géostatistiques multi-variables. L'idée principale de cette récente technique est que l'information spatiale des données RADAR peut être utilisée indépendamment pour conditionner l'interpolation des pluviomètres (Ehret 2003, p. 99). En effet, en considérant qu'un pluviomètre mesure ponctuellement la quantité réelle de la pluie au sol et qu'un RADAR fournit une image de la répartition spatiale réelle des précipitations, l'estimation doit contenir la structure spatiale donnée par le RADAR et dont la valeur quantitative est donnée par les pluviomètres (Sinclair et Pegram 2005).

Cette technique se décompose en 4 grandes étapes présentées dans la Figure II-5. En premier lieu, les observations ponctuelles des pluviomètres ( $\mathbf{0}_a$ ) sont interpolées sur une grille similaire à celle du RADAR par krigeage ordinaire (OK) (1). Parallèlement, les valeurs des pixels RADAR ( $\mathbf{0}_b$ ) aux points des pluviomètres sont récupérées ( $\mathbf{2}_0$ ) pour être interpolées spatialement par OK (2). À chaque point de grille, le champ des erreurs (3) est calculé à partir du champ RADAR observé ( $\mathbf{0}_b$ ) et celui interpolé (2). Ce champ est ensuite appliqué au champ obtenu par l'interpolation des données pluviométriques (1). Le champ de pluie final (4) est supposé suivre le champ moyen des observations pluviométriques ( $\mathbf{0}_a$ ) tout en conservant la structure des mesures RADAR ( $\mathbf{0}_b$ ).

## b | Discussion

Dans ce qui suit, nous commenterons quelques considérations qui ont été faites dans l'algorithme initial et que nous avons délibérément choisi de ne pas prendre en compte dans nos travaux de recherches.

Premièrement, dans l'étude initiale (Ehret 2003, p. 100), le champ résultant de l'interpolation par OK des données pluviométriques (1) et des pixels de l'image RADAR aux points des pluviomètres (2) sont obtenus à partir de variogrammes paramétriques différents. Cette considération amène donc à construire deux variogrammes paramétriques : celui des données pluviométriques et celui des mesures RADAR. Dans nos cas d'étude nous avons opté pour l'utilisation d'un unique variogramme pour les deux interpolations. C'est une façon de s'assurer que champ des erreurs (3) faites lors de l'interpolation des données RADAR correspond aux mêmes erreurs faites avec la même technique d'interpolation (et donc du même variogramme) des données pluviométriques.

Le second point soulevé par Ehret (2003, 100–101) porte sur la différence du volume d'échantillonnage entre les observations faites par les pluviomètres ou les RADAR. En effet, l'un observe la pluie en un point, le second la mesure sur un volume intégré le long d'un faisceau que nous avons assimilé à un point de grille. C'est pourquoi l'auteur préconise un krigeage « par blocs » (*i.e.* moyenne sur une surface ou volume) au lieu d'un krigeage par points comme utilisé dans nos réalisations. Le krigeage par blocs est particulièrement apprécié lorsqu'une estimation est désirée pour une surface telle qu'un bassin versant hydrologique. L'estimation par bloc est généralement plus exacte (Isaaks et Srivastava 1989, p. 337) mais comme le note Isaaks et Srivastava (1989, p. 326), le désavantage d'un tel système est qu'il nécessite davantage de calculs (et donc de temps pour sa résolution) qu'un système par points. Pour nos cas d'étude, la résolution spatiale des grilles RADAR étant suffisamment fine, l'utilisation d'un krigeage par blocs n'est donc pas justifiée.

La troisième remarque pointe le problème d'extrapolation des données en tout point en dehors du réseau de pluviomètres (*i.e.* loin de toutes stations). Il est alors possible de définir une zone d'influence de ces observations au-delà de laquelle l'estimation finale doit davantage se fier aux données RADAR plutôt qu'aux données pluviométriques. Pour se faire, Ehret (2003, p. 101) propose d'appliquer un facteur correctif au champ des données pluviométriques interpolées (1) qui correspond à l'inverse de la variance de krigeage (faible à proximité des stations et d'autant plus grand que l'on s'en éloigne). Par ailleurs, pour une meilleure correspondance du champ final avec les données RADAR, l'auteur applique un

masque tel que pour tous les pixels de l'image RADAR ( $\mathbf{0}_b$ ) sans pluie (*i.e.* valeur nulle), la valeur finale du pixel associée du champ final ( $\mathbf{4}$ ) est ramenée automatiquement à 0. Frezghi et Smithers (2008) démontrent qu'avec cet algorithme, la valeur des pluviomètres ( $\mathbf{0}_a$ ) n'est pas conservée dans le champ final ( $\mathbf{4}$ ). C'est pourquoi nous avons décidé de ne pas implémenter cette modification dans notre analyse.

Enfin, Ehret (2003, 102–103) montre que le champ final ( $\mathbf{4}$ ) dépend fortement du calcul du champ des erreurs ( $\mathbf{3}$ ). La plus intuitive des méthodes est de calculer la différence entre le champ RADAR interpolé et observé (*i.e.*  $(\mathbf{2}) - (\mathbf{0}_b)$ ). Cette simple soustraction amène toutefois à la présence de valeurs négatives dans le champ final ( $\mathbf{4}$ ). L'auteur propose alors deux calculs alternatifs :

- l'utilisation du rapport entre les deux champs (*i.e.*  $(\mathbf{3}) = \frac{(\mathbf{2})}{(\mathbf{0}_b)}$ ) ;
- l'utilisation du rapport logarithmique entre les deux champs auxquels est ajouté un facteur de 1.5 pour éviter toutes discontinuités de la fonction logarithme au voisinage de 0 (*i.e.*  $(\mathbf{3}) = \frac{\log((\mathbf{2})+1.5)}{\log((\mathbf{0}_b)+1.5)}$ ).

En testant ces deux propositions, nous nous sommes rendu compte d'un problème majeur : pour les pixels de l'image RADAR ( $\mathbf{0}_b$ ) dont les mesures de la pluie sont très faibles (*i.e.* le dénominateur tend vers 0), les différences avec le champ interpolé ( $\mathbf{2}$ ) entraînent des valeurs finales complètement irréalistes. Ainsi, pour éviter tous désagréments, nos champs des déviations ( $\mathbf{3}$ ) sont simplement obtenus à partir du calcul de la différence entre le champ RADAR interpolé et observé. Nous avons également implémenté un contrôle des valeurs de chaque pixel du champ final ( $\mathbf{4}$ ) qui ramène automatiquement toutes valeurs négatives à 0.

### II-2.2.3 Discussion sur la comparaison des deux techniques

Dans ce dernier paragraphe, nous résumerons les principales conclusions obtenues dans différentes études trouvées dans la littérature. Ceci nous permet de positionner les performances du KED et du MERG parmi d'autres méthodes d'interpolation spatiale (simples ou complexes), en se focalisant plus particulièrement sur l'étude des précipitations.

#### a | Le choix des variables secondaires

De nombreuses études de comparaisons ont été menées dans l'espoir de trouver la méthode donnant les meilleures estimations avec le minimum d'erreur. Goovaerts (2000) a été l'un des premiers à montrer l'importance de l'utilisation de variables externes dans les

méthodes d'interpolation. En s'appuyant sur la relation entre précipitation et topographie, il conclue dans sa comparaison climatique de diverses méthodes que i) les erreurs d'estimations sont les plus grandes pour les méthodes n'incluant pas l'altitude et ii) les méthodes géostatistiques à plusieurs variables doivent être préférées aux simples régressions linéaires. L'auteur indique par ailleurs que les résultats obtenus avec le cokrigeage (technique plus complexe et plus coûteuse en calcul) présentent moins de détails que le KED. Des résultats analogues sont retrouvés dans l'étude de Moral (2010). Dans une étude similaire, Haberlandt (2007) rapporte que les performances du KED sont incontestables mais celles-ci dépendent de la variable externe utilisée. Verworn et Haberlandt (2010) ont ainsi testé différentes variables et avancent que les meilleures estimations sont obtenues grâce à l'implémentation des données RADAR, devançant ainsi les mesures issues d'un éventuel second réseau de pluviomètres (= champ interpolé) et la topographie. Erdin (2009, p. 55) conforte ces remarques. Tobin *et al.* (2011) ont également expérimenté l'utilisation de sorties de modèles numériques mais les résultats dépendent fortement de la qualité des ré-analyses employées. Dans leur étude, Velasco-Forero *et al.* (2009) comparent trois techniques intégrant les données RADAR comme variable auxiliaire (OK des pluviomètres dont le variogramme est déterminé à partir des données RADAR, KED et cokrigeage). Tous les résultats obtenus sont satisfaisants. Toutefois, la technique du KED reproduit avec plus d'exactitude le caractère spatial des pluies observées par le RADAR, notamment pour les stations éloignées (*i.e.* grandes distances sans pluviomètres) (Velasco-Forero *et al.* 2008).

## b | KED vs MERG

La technique du KED ressort souvent comme étant la technique d'interpolation, combinant observations pluviométriques et données RADAR et donnant les meilleurs résultats par rapport à des méthodes d'interpolation plus simples telles que le plus proche voisin, l'inverse de la distance ou le OK. Delobbe *et al.* (2008) et Goudenhoofdt et Delobbe (2009) ont évalué les performances du KED par rapport entre-autre à la technique du MERG, plus récente et donc moins répandue. Encore une fois, les estimations par KED (avec comme seule variable externe les données RADAR) devancent toutes les autres, à l'exception faite du MERG qui donnent des résultats très similaires avec un moindre coût de calcul.

## c | Tests de sensibilités

En poussant un peu plus loin leurs recherches, Delobbe *et al.* (2008) et Goudenhoofdt et Delobbe (2009) montrent que les résultats obtenus par KED et MERG semblent être les techniques les moins sensibles à i) la saison (et donc du type de précipitation), ii) la distance

par rapport au RADAR ou iii) la densité du réseau d'observations pluviométriques. Les résultats révèlent également que i) les estimations par OK sont toutes aussi bonnes que celles des meilleures méthodes de combinaison dans le cas de précipitations stratiformes étendues (*i.e.* en hiver) et ii) plus le réseau d'observations est dense, plus les performances des méthodes de combinaison sont accrues. Ce dernier point est également observé par (Velasco-Forero *et al.* 2008) qui estiment que lorsque le nombre de stations pluviométriques est extrêmement faible, les images RADAR sont préférables aux champs obtenus par KED ou MERG.

Tous les résultats obtenus au travers des diverses études montrent la suprématie des méthodes de combinaison de données (et plus particulièrement du KED) ce quels que soient :

- la région géographique concernée : en plaine ou en montagne, pour une surface centralisée sur quelques centaines de kilomètres autour d'un RADAR ou une surface telle un petit pays comme la Suisse ou la Belgique ;
- les événements considérés : d'une durée de quelques heures à quelques décennies, différenciés ou non selon la saison ou le type de pluie ;
- le type de réseaux d'observations : petit ou grand nombre de pluviomètres, principaux ou secondaires, un ou plusieurs RADAR.

L'observation d'un phénomène naturel, le plus souvent ponctuellement, peut être estimée en tout point de l'espace à l'aide de méthode d'interpolation. Le krigeage est la méthode la plus courante. Ce choix provient d'un souci de réalisme puisque c'est l'une des seules à prendre en compte la dépendance spatiale du phénomène observé. Depuis sa découverte, nombreux sont ceux qui ont contribué à son développement. Aujourd'hui il est possible d'apporter diverses informations complémentaires dans l'espoir de rapprocher l'estimation de la réalité. Par d'exemple, les champs de précipitation ne sont plus seulement obtenus à partir des observations pluviométriques mais en les combinant avec les mesures des RADAR météorologiques, des données de topographie voire des données issues de modèles numériques. Dans le cadre de nos travaux, nous nous sommes focalisés sur deux techniques de combinaison (KED et MERG) qui ont fait leur preuve à travers différentes études. Dans les deux chapitres qui suivent, nous présenterons les résultats respectifs au travers de deux domaines d'étude distincts. Nous verrons que chacune de ces techniques a ses avantages comme ses inconvénients

---

## Les Estimateurs statistiques utilisés dans le Chapitre III

Notons  $V_i$  la valeur observée et  $\tilde{V}_i$  la valeur estimée à la station  $i$ . Considérons  $N$  stations sur l'ensemble du domaine étudié, alors :

→ l'erreur moyenne (ME) correspond à la moyenne des écarts entre les valeurs estimées et observées. Également appelée biais, elle est définie par :

$$ME = \frac{1}{N} \sum_{i=1}^N (\tilde{V}_i - V_i)$$

Pour le cas d'une estimation « parfaite », la ME arbore une valeur de 0. Dans le cas contraire, il est possible d'obtenir une ME nulle si les écarts se compensent entre eux.

→ l'erreur quadratique moyenne (MSE) mesure la moyenne des carrés des écarts entre les estimations et les observations. Elle est plus souvent exprimée sous la forme d'une racine carrée (RMSE), telle que :

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (\tilde{V}_i - V_i)^2}$$

Par rapport à la ME, la RMSE accorde plus de poids aux larges écarts. Elle est particulièrement appréciée pour détecter la présence de larges écarts, surtout si les écarts se compensent. Elle n'indique cependant pas la direction des écarts.

→ l'écart-type mesure la dispersion des valeurs de la variable considérée. Pour les observations, il est exprimé par :

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (V_i - \bar{V})^2}$$

avec  $\bar{V}$  la moyenne des valeurs observées.  $\sigma^2$  définit la variance de l'échantillon considéré.

→ le coefficient de corrélation (spatial) renseigne sur la relation linéaire (ou phase) entre le champ des valeurs estimées et celui des valeurs observées. Il est donné par :

$$r = \frac{\sum_{i=1}^N (\tilde{V}_i - \bar{\tilde{V}}) \cdot (V_i - \bar{V})}{\sqrt{\sum_{i=1}^N (\tilde{V}_i - \bar{\tilde{V}})^2} \cdot \sqrt{\sum_{i=1}^N (V_i - \bar{V})^2}}$$

avec  $\bar{V}$  et  $\bar{\tilde{V}}$  les moyennes des valeurs observées et estimées. Un coefficient de corrélation nul indique que les deux champs n'ont aucun lien entre eux. Plus sa valeur s'éloigne de zéro (*i.e.* en s'approchant de 1 ou -1), meilleure est la corrélation (positive = en phase ou négative = en opposition de phase).

---

# Chapitre III

## Reconstruction d'un champ de pluie : Application dans la région Cévennes- Vivarais

Dans ce chapitre, nous nous proposons d'appliquer les deux techniques d'interpolation spatiale combinant les observations pluviométriques et mesures RADAR décrites dans le 0. Nos premiers cas d'étude portent sur un phénomène météorologique atypique : les épisodes cévenols. Nous verrons dans une première partie que la singularité des événements précipitants et les enjeux (parfois humains) qui en découlent ont incité les autorités à mettre en place un important réseau d'observation. Nous détaillerons ensuite les analyses quant à la reconstruction des champs de pluie de deux cas extrêmes. Les deux techniques d'interpolation considérées seront alors confrontées entre elles afin de mieux comprendre leur comportement et leur performance. Pour les deux dates étudiées, nous présenterons une comparaison qualitative que nous avons accompagnée par une validation quantitative. Enfin, nous testerons la sensibilité de chacune des techniques à la période temporelle et la densité du réseau pluviomètre au sol.

### III-1 La région Cévennes-Vivarais

A l'extrême sud-est du Massif central (Figure III-1), les Cévennes forment une chaîne de basses et moyennes montagnes à cheval sur les départements de la Lozère et du Gard. Elles sont prolongées au nord-est par les monts du Vivarais qui englobe une partie de l'Ardèche et de la Haute-Loire. La région Cévennes-Vivarais s'étend par delà les plaines, sur la rive droite du Rhône, du nord de Valence jusqu'à la mer Méditerranée.

Au cours de l'automne, cette zone est le théâtre d'intenses événements pluvieux où la topographie joue un rôle très important. Ces précipitations se produisent lors de situations météorologiques typiques de méso-échelle (étendue spatiale = 2 à 2 000 km, durée de vie = plusieurs heures à quelques jours). Même si aujourd'hui ces épisodes sont bien connus, leurs

prévisibilités restent encore très limitées de par la complexité des processus impliqués à plus fines échelles. De surcroît, les caractéristiques géologiques de la région font que lorsque de fortes pluies surviennent sur les versants cévenols, l'eau s'infiltré mais les sols sont rapidement saturés car peu profonds (Audard-Vincendon 2010, p. 6). Par ailleurs, la configuration des bassins versants provoque des réactions soudaines et puissantes des cours d'eau (dites « crues éclairs ») dont les conséquences sont aggravées par la présence d'une plaine fortement peuplée en aval du relief (Yates 2006, p. 3)

### III-1.1 Description physique des événements cévenols

Le terme d'« événement cévenol » correspond à un mécanisme météorologique précis et particulier à la région. Il désigne des pluies convectives d'intensités moyennes à fortes, d'une longue durée de vie (jusqu'à 36 - 48 heures). La région Cévennes-Vivarais présente deux facteurs favorables au développement de la convection (Yates 2006, p. 10) : i) la présence de la mer Méditerranée, source d'humidité et source de chaleur en automne de part son inertie thermique et ii) la présence de relief qui, en forçant le soulèvement des masses d'air, favorise la formation des nuages et de précipitations. Ces événements sont reliées à trois types de dynamique de grande échelle (Kirstetter 2008, p. 9; Yates 2006, p. 11) :

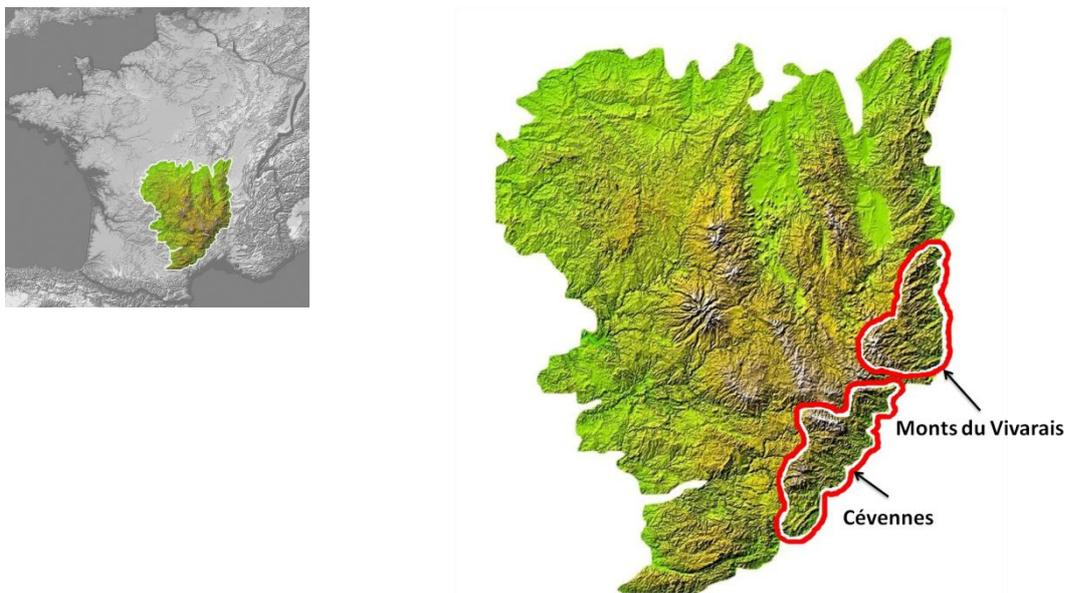


Figure III-1 : Carte du Massif central avec localisation des Cévennes et des monts du Vivarais.

Source : Wikipedia (modifiée)

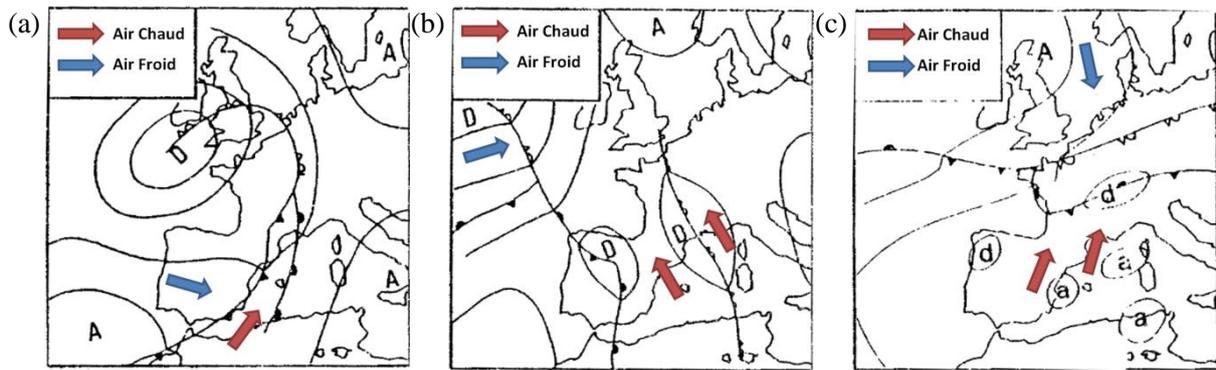


Figure III-2 : Les trois types de situations météorologiques associées à des précipitations dans les Cévennes : (a) les perturbations cévenoles, (b) les retours d'air méditerranéen et (c) les marais barométriques.

Source : Yates (2006, p. 11) © Tourasse (1981) (modifiées)

- le passage d'une perturbation d'ouest lié à un système dépressionnaire au large de l'Islande et où le déplacement du front froid vers l'est est gêné par deux anticyclones, positionnés respectivement généralement au dessus de la Russie et au nord de l'Afrique (Figure III-2a) ;
- la présence d'une zone de basse pression centrée sur le Golfe de Gascogne ou le nord de l'Espagne, entraînant de puissants flux de secteur sud à sud-sud-est qui, en se heurtant à l'orographie, génèrent par soulèvement la formation de pluies au-dessus des reliefs (Figure III-2b) ;
- les cas de situations de marais barométriques où la présence de maxima et/ou minima locaux de pression peuvent conduire à des convergences des basses couches (Figure III-2c).

Depuis dix ans, une douzaine de phénomènes extrêmes ont été enregistrés et quasiment tous ont été reconnus comme catastrophes naturelles. Pour prévenir ces risques, la surveillance de ces événements est un enjeu essentiel et dont la priorité est axée sur la localisation et la mesure des intensités des zones pluvieuses. Mis en place en 2000 par un consortium de laboratoires de recherche français et étrangers, l'OHMCV (acronyme pour Observatoire Hydrométéorologique Méditerranéen Cévennes-Vivarais, [www.ohmcv.fr](http://www.ohmcv.fr)) se consacre à l'étude de ces pluies intenses et des crues éclair engendrées. Il s'oriente également sur le développement de techniques d'observations (ex : Delrieu *et al.* 2009) et la validation par des modèles météorologiques (ex : Nuissier *et al.* 2008; Ducrocq *et al.* 2008) parfois couplés à des modèles hydrologiques (ex : Vincendon *et al.* 2009).

### III-1.2 Le réseau d'observations

Les organismes opérationnels partenaires de l'OHMCV mettent à disposition un réseau dense d'instruments regroupant notamment des pluviomètres et des RADAR météorologiques (Figure III-3).

Les données pluviométriques (triangles pointant vers le haut) proviennent des réseaux de Météo-France, du Service de Prévision des Crues Méditerranée-Ouest, du Service de Prévision des Crues du Grand Delta et de l'unité Électricité De France/Direction Technique Générale. Au total, ce sont plus de 300 stations disponibles, soit une densité moyenne d'un pluviomètre par  $12 \times 12 \text{ km}^2$  (site Internet : OHMCV). Les capteurs sont principalement des pluviomètres automatiques à augets basculeurs. La base de temps commune à l'ensemble de ces appareils est le pas de temps horaire, même si certains d'entre eux fournissent des mesures avec une meilleure résolution (5 minutes). Ce réseau est complété par environ 220 pluviomètres de Météo-France fournissant des données journalières (triangles pointant vers le bas).

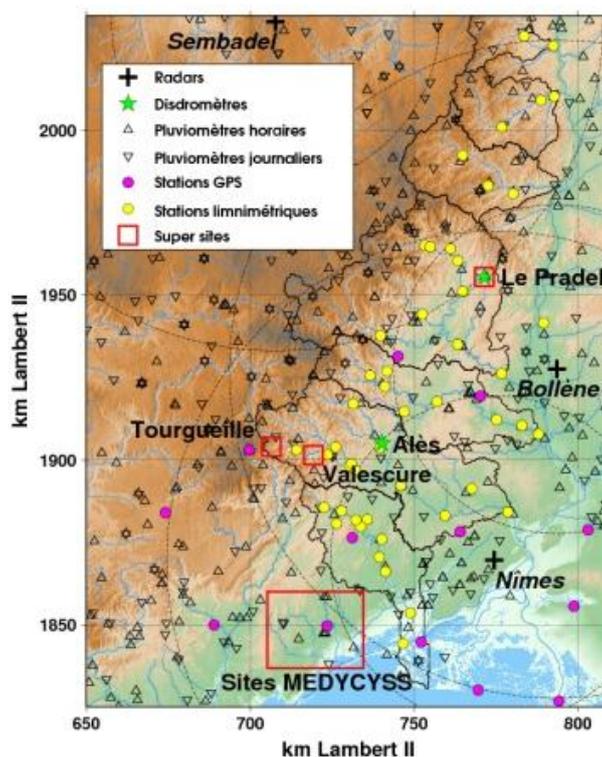


Figure III-3 : Fenêtre d'observation et réseaux de mesures de l'OHMCV.  
Source : [http://www.ohmcv.fr/P234\\_cv.php?lang](http://www.ohmcv.fr/P234_cv.php?lang)

La région Cévennes-Vivarais est couverte par trois RADAR météorologiques (croix noires) localisés à Bollène (Vaucluse), Nîmes (Gard) et Sembadel (Haute-Loire). Tous trois sont membres du réseau national de RADAR opérationnels de Météo-France. Pour les cas d'étude que nous présentons dans ce chapitre, seules les données du RADAR de Bollène étaient disponibles. Ce RADAR fonctionne en bande S, ce qui permet de s'affranchir de l'atténuation du signal. Son protocole de sondage volumétrique consiste en une série de trois balayages panoramiques PPI de six angles d'élévation (de 0.4 à 6°), chacun d'une durée de 5 minutes. Deux grandeurs sont mesurées : le facteur de réflectivité moyen et la variance du signal RADAR tir à tir.

Le post-traitement des données est assuré par le logiciel « Traitements Régionalisés et Adaptatifs de Données RADAR pour l'Hydrologie » développé par le Laboratoire d'étude des Transferts en Hydrologie et Environnement (Grenoble). La chaîne d'algorithmes se compose de quatre étapes :

- la détermination du domaine de détection du RADAR : identification des zones de masque, détermination des échos fixes sur des images par temps sec, vérification de l'étalonnage du RADAR (via l'étude des échos de relief) ;
- l'identification dynamique des échos parasites (échos fixes, avions, interférences, ...) et les types de pluie rencontrés le long des profils verticaux ;
- le calcul des facteurs correcteurs à appliquer sur la réflectivité pour la corriger des effets de masques et des hétérogénéités verticales ;
- l'estimation quantitative de la pluie : projection au sol (via une moyenne pondérée sur la verticale), application d'une relation  $Z-R$  optimisée, moyenne horaire des sondages à 15 minutes, application d'un schéma d'advection (pour limiter les effets de crénelages), élimination des échos fixes résiduels.

En sortie, une image composite d'intensités pluvieuses moyennes est élaborée toutes les heures sur une grille de maille rectangulaire de  $1 \times 1 \text{ km}^2$  (projection conique conforme de Lambert II étendu – cf. par exemple Snyder (1987, 104–107)). Pour plus d'informations sur la chaîne de traitement, le lecteur pourra se référer à Kirstetter (2008, chap. 2), Delrieu *et al.* (2009) ou Bouilloud *et al.* (2010).

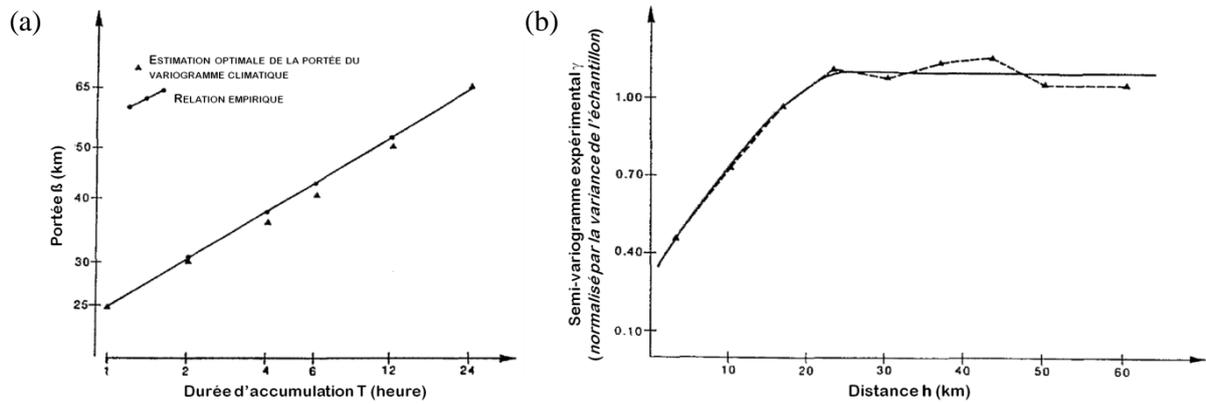


Figure III-4 : (a) Estimation de la relation entre la portée ( $\beta$ ) et la durée d'accumulation ( $T$ ) pour le cas des Cévennes et (b) détermination du variogramme paramétrique (trait gras) à partir du variogramme expérimental (trait pointillé) pour  $T = 1$  h.  
 Source : Lebel *et al.* (1987) (modifiées)

## III-2 Spatialisation des données

Les événements affectant la région ont depuis longtemps été observés. Dès lors, de nombreuses études ont été menées dans l'espoir d'obtenir une représentation spatiale la plus proche possible de la réalité. Pour ce faire, l'étape de la détermination du variogramme paramétrique  $\tilde{\gamma}$  est essentielle. Dans le cadre de nos recherches, nous nous sommes reposés sur les études de Lebel et Bastin (1985) et Lebel *et al.* (1987). Le contexte et la procédure menant à l'estimation des paramètres du variogramme paramétrique seront abordés en détail dans le paragraphe IV-3.1. Nous admettons que le variogramme  $\tilde{\gamma}$  est décrit par un modèle **sphérique** isotrope dont :

- la portée  $\tilde{\beta}$  (en km) est donnée en fonction de la durée d'accumulation des précipitations  $T$  (en heure) telle que  $\tilde{\beta} = 25 \times T^{0.3}$  (Figure III-4a) ;
- l'effet de pépite  $\tilde{c}_0$  n'est pas considéré (*i.e.* valeur nulle) ;
- le facteur  $\tilde{\alpha}$  (assimilé au palier) est égal à 1.12 fois la variance de l'échantillon (Figure III-4b).

L'automne 2008 a été marqué par deux épisodes de crues significatives sur les Cévennes : les 21 et 22 octobre puis le 1er et 2 novembre. Ces deux événements ont été retenus de par leurs dissemblances, notamment sur la localisation des pluies mais également sur la qualité des estimations des mesures RADAR. Les champs de précipitation reconstruits

ont alors été analysés de façon à mieux appréhender les atouts et faiblesses du krigeage avec dérive externe (KED) et de la fusion conditionnée (MERG). L'analyse des deux cas sera la même. Nous commencerons par une description de la situation météorologique, suivie d'un bref rappel sur les données utilisées et d'une analyse des données pluviométriques. Nous exposerons ensuite les résultats et conclusions obtenus lors de la reconstruction du champ de pluie au pas de temps<sup>⊕</sup> de l'événement ainsi que sur des échelles temporelles plus courtes. Enfin, nous présenterons à la fin de ce chapitre quelques expérimentations effectuées pour tester la sensibilité des différentes techniques d'estimation notamment face à la densité du réseau de pluviomètres.

### III-3 CAS A : du 19 au 22 Octobre 2008

#### III-3.1 Situation météorologique

Après un début d'automne 2008 déficitaire en pluies sur les Cévennes, un premier épisode intense survient du 20 au 22 Octobre. La Figure III-5 donne les cartes d'analyses de la situation météorologique pour ces deux dates. Durant cette période, deux anticyclones (croix rouges), positionnés au nord des Açores et en Russie occidentale, bloquent une perturbation liée à une dépression située au large de la Scandinavie (croix bleue). Une zone frontale s'étend depuis l'Atlantique à l'est des Açores jusqu'au nord-est de l'Europe septentrionale. Pendant ces trois jours, un front froid (ligne bleue) se propage lentement vers l'est.

Dans la région cévenole, un vent de sud/sud-ouest s'installe dans les basses couches (flèche orange). En provenance de la mer Méditerranée, ce flux chaud et chargé en humidité vient buter sur le relief déclenchant localement des processus de convection (ligne verte = ligne de grain). Ces cellules se développent dans l'après-midi du 20 Octobre (Figure III-6a) et perdureront une quinzaine d'heures (Bousquet 2009). L'arrivée du front froid survient dans la nuit du 21 au 22 et s'accompagne par une bande de pluie qui s'ajoute à l'activité convective déjà présente sur les contreforts des Cévennes, avant de poursuivre sa course vers l'est, étalant les précipitations plus au nord sous forme de pluies stratiformes (Figure III-6b).

Notre zone d'intérêt s'étend sur 110 km d'ouest en est et sur 130 km du nord au sud (Figure III-7). Au total, 269 stations au pas de temps horaires sont disponibles.

<sup>⊕</sup> Par les termes « pas de temps », « durée de cumul ou d'accumulation », « période, fenêtre, échelle temporelle », nous considérons que les données initiales (pluviométriques ou RADAR) sont sommées de façon à obtenir des données cumulées sur le laps de temps considérés.

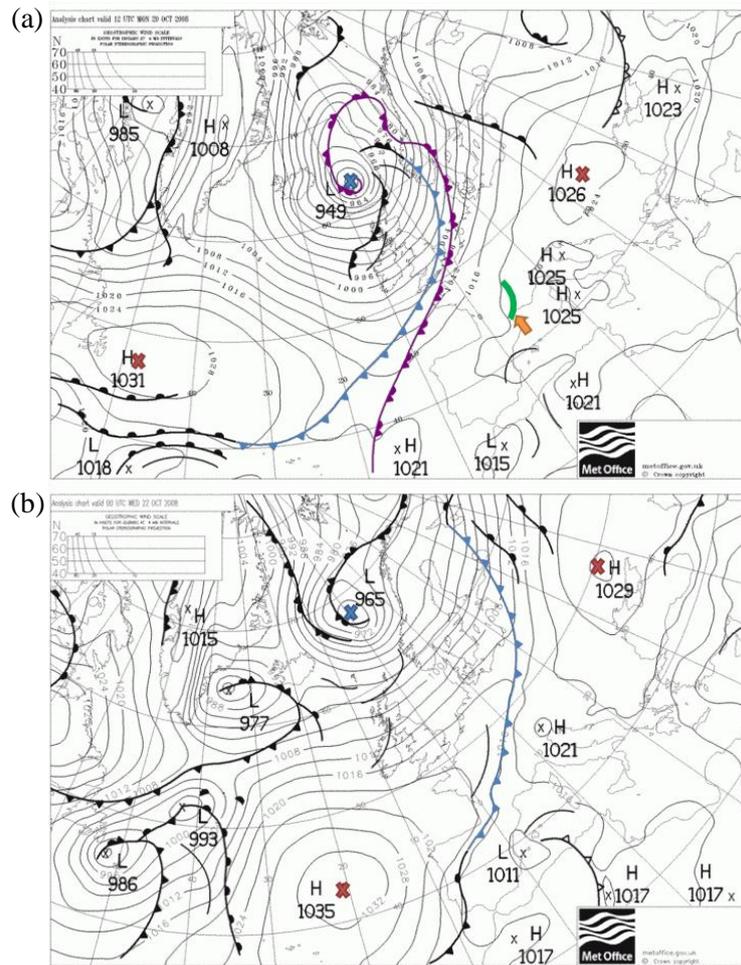


Figure III-5 : Carte d'analyse (a) du 20 Octobre 2008 à 12 h UTC et (b) du 22 Octobre 2008 à 00 h UTC.

Source : Met Office (archive : <http://www.meteo60.fr/archives.php> | coloriées)

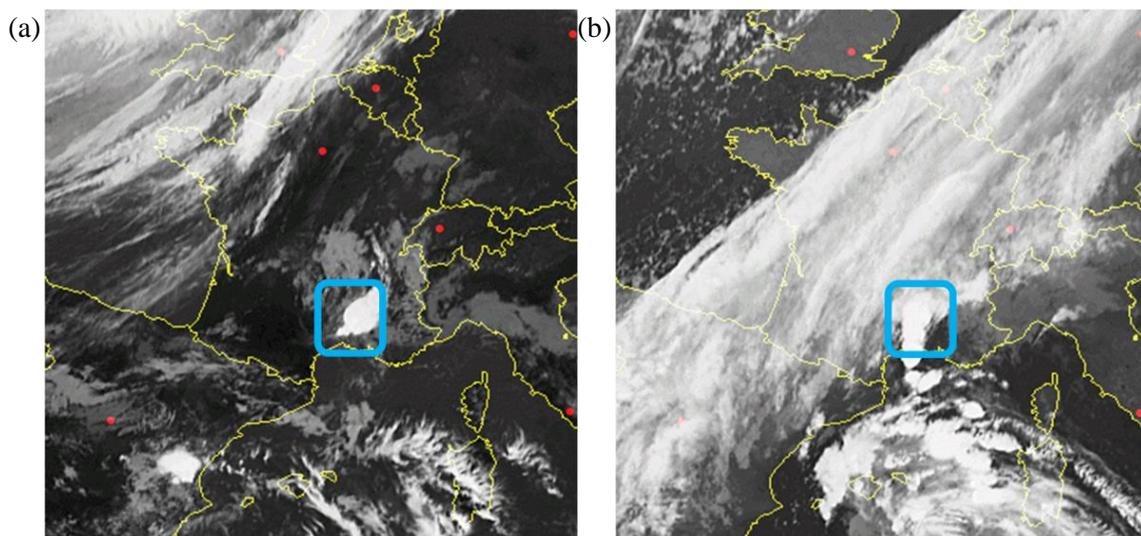


Figure III-6 : Image satellite infrarouge (a) du 20 Octobre 2008 à 13 h 15 UTC et (b) du 21 Octobre 2008 à 13 h 15 UTC.

Source : EUMETSAT (archive : <http://www.meteo60.fr/archives.php> | zoomées)

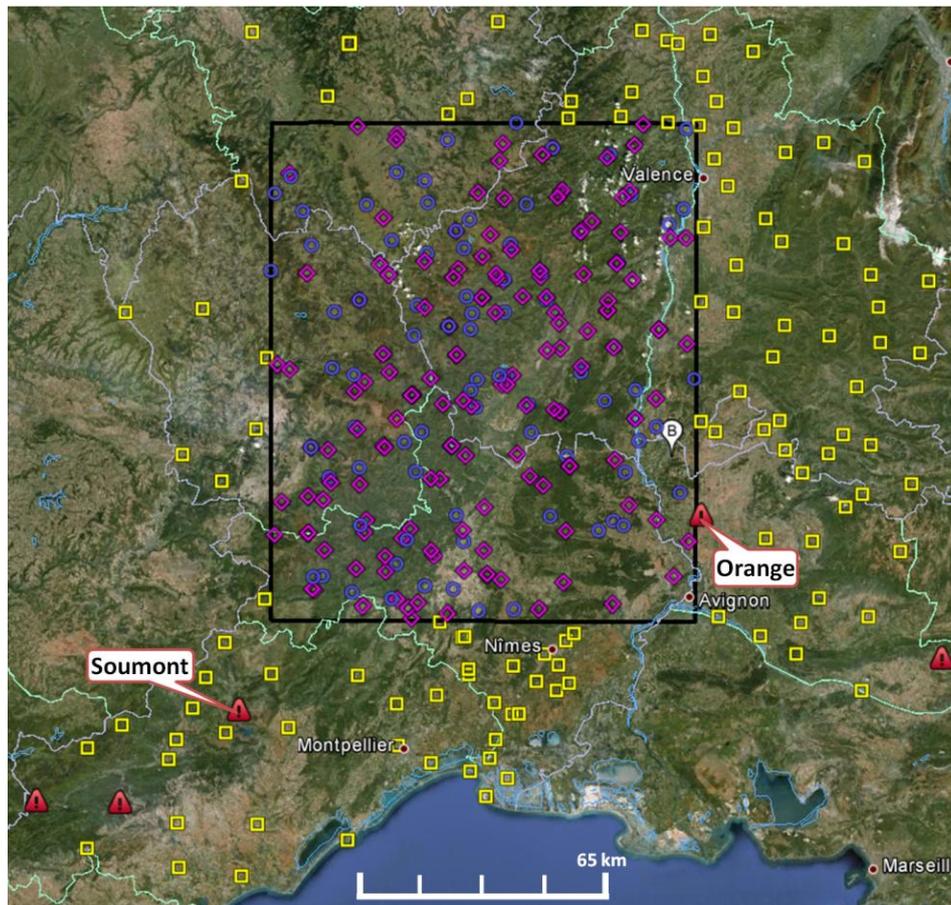


Figure III-7 : Localisation des instruments pour le cas A = du 19 au 21 Octobre 2008. Le rectangle noir représente la zone d'intérêt couverte par 149 pluviomètres horaires (marqueurs roses) et 84 pluviomètres journaliers (marqueurs bleus). En dehors de cette zone, 120 pluviomètres horaires sont également disponibles (marqueurs jaunes). La lettre B dans le marqueur blanc indique l'emplacement de l'unique RADAR de Bollène disponible pour ce cas. Les marqueurs rouges localisent les stations horaires où au moins une observation est considérée comme suspecte d'après le contrôle de qualité appliqué. Source : Google Earth.

### III-3.2 Le contrôle de qualité des données pluviométriques

Le contrôle de qualité (QC) des données pluviométriques appliqué a mis en exergue cinq observations douteuses provenant de cinq stations différentes (nous n'en montrerons ici que deux). La Figure III-8 montre l'observation douteuse enregistrée à la station #245 (Soumont – Hérault) accompagnée des observations mesurées à plus ou moins 2 heures ainsi celles des cinq stations voisines les plus représentatives. La valeur observée dépasse alors de plus de 11 fois la valeur attendue (équation [II-1]). Sur le même principe, la Figure III-9 montre l'observation douteuse mesurée à la station #172 (Orange – Vaucluse). Pour cette station, la mesure est 20 fois supérieure à celle attendue.

Détectées par le QC, ces observations sont supprimées de la base de données.

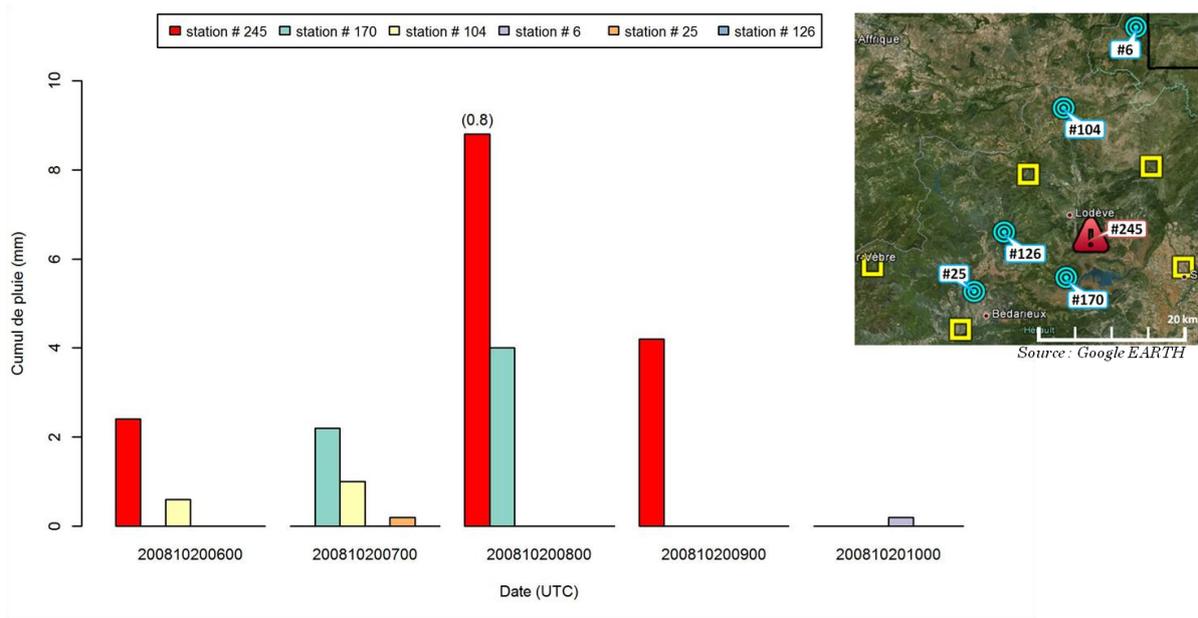


Figure III-8 : Évolution temporelle des cumuls horaires (en mm) de la station #245 (marqueur rouge sur la carte de droite) et ses 5 plus représentatives stations voisines (marqueurs cyans sur la carte de droite) du 20 Octobre 2008 de 06 h UTC à 10 h UTC. L'observation à 08 h UTC est celle suspectée. L'indication entre parenthèses donne une approximation de la valeur attendue.

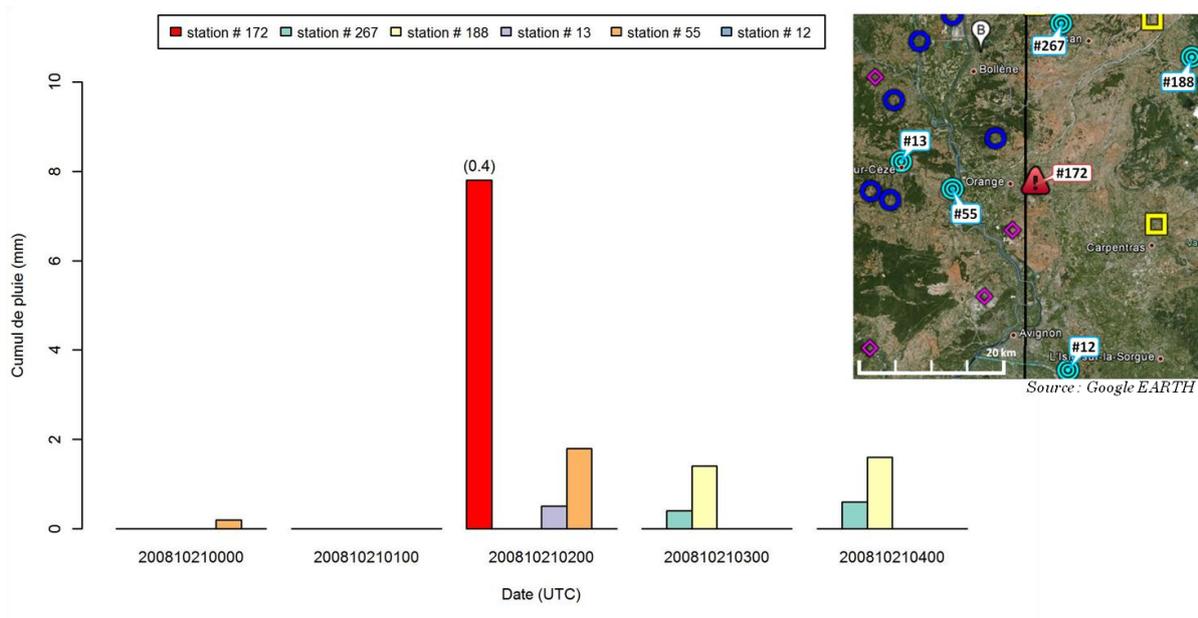


Figure III-9 : Évolution temporelle des cumuls horaires (en mm) de la station #172 (marqueur rouge sur la carte de droite) et ses 5 plus représentatives stations voisines (marqueurs cyans sur la carte de droite) du 21 Octobre 2008 de 00 h UTC à 04 h UTC. L'observation à 02 h UTC est celle suspectée. L'indication entre parenthèses donne une approximation de la valeur attendue.

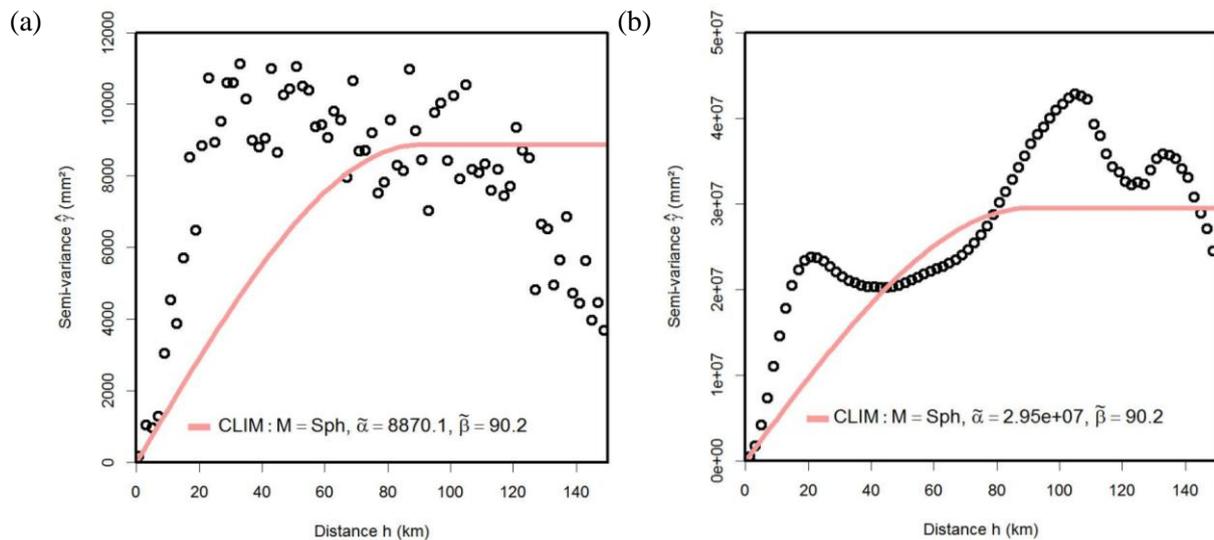


Figure III-10 : CAS A = Semi-variogramme expérimental (a) des observations pluviométriques et (b) des résidus obtenus à partir des mesures RADAR. ( $T = 72$  h). Les variogrammes paramétriques considérés sont représentés par les courbes rouges.

### III-3.3 Informations complémentaires

Pour ce cas, la relation  $Z$ - $\mathfrak{R}$  (équation [I-5]) utilisée est donnée par l'équation (Bouilloud *et al.* 2010; Boudevillain 2011) :

$$Z = 325 \cdot \mathfrak{R}^{1.4}$$

Au pas de temps de l'événement ( $T = 72$  h), les variogrammes paramétriques considérés, identiques pour les observations (technique du krigeage ordinaire) et les résidus (technique du krigeage avec dérive externe – selon l'approche de Velasco-Forero *et al.* (2009)), sont représentés par la Figure III-10.

### III-3.4 Reconstruction d'un champ de précipitation à l'échelle de l'événement

#### III-3.4.1 Validation qualitative

Dans un premier temps, nous nous sommes focalisés sur les cumuls des précipitations sur les 3 jours de l'événement observés par les pluviomètres (Figure III-11a) et le RADAR (Figure III-11b). Ces cartes montrent que la convection s'est organisée sur un axe sud-ouest/nord-est, principalement en aval du massif. Le maximum est enregistré à La Grand-Combe (Gard) avec un cumul de 502 mm en 72 heures.

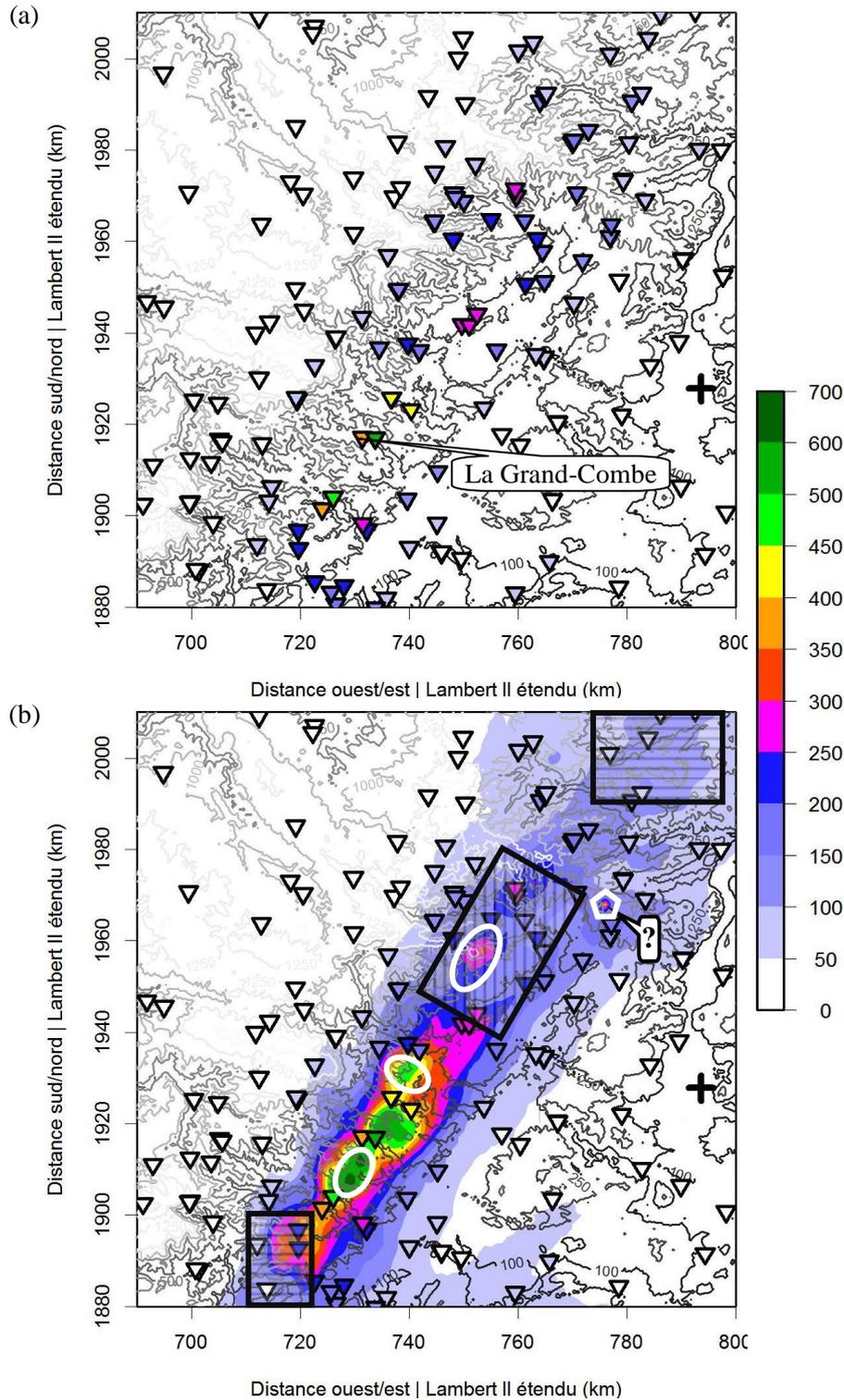


Figure III-11 : CAS A = Champ de pluie (en mm) observé par (a) les pluviomètres et (b) le RADAR. (T = 72 h).

La croix noire indique la position du RADAR de Bollène. La topographie de la région est rajoutée sous forme d'iso-contours en nuances de gris, du plus foncé pour les altitudes proches de la surface de la mer, au plus clair pour les altitudes atteignant les 1 250 m. Sur la figure (b), les observations pluviométriques sont reportées par les valeurs contenues dans les triangles pointant vers le bas. Les rectangles hachurés horizontalement (verticalement) montrent une surestimation (sous-estimation) des mesures RADAR par rapport aux pluviomètres. Les ellipses blanches localisent de petites structures de pluie non détectables par le réseau de pluviomètres.

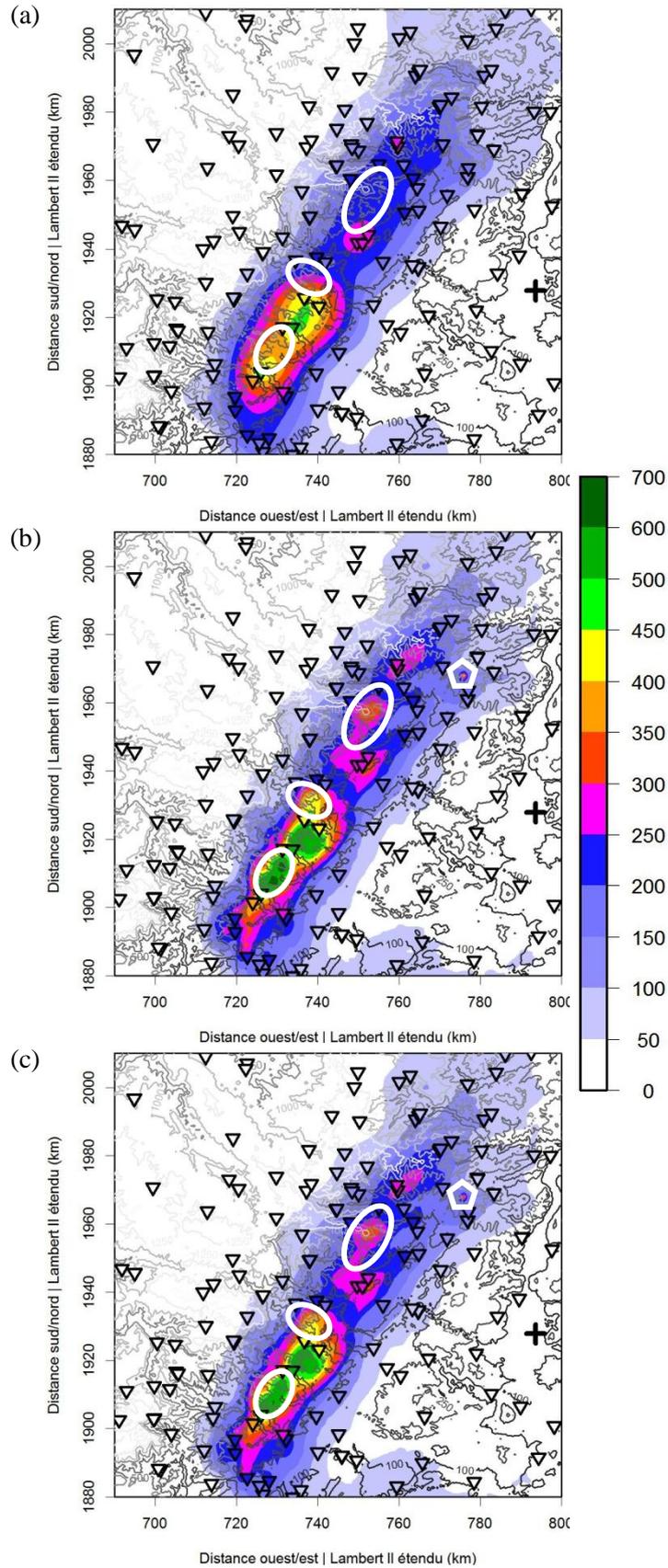


Figure III-12 : CAS A = Champ de pluie (en mm) interpolé par (a) OK\_RG, (b) KED et (c) MERG. (T = 72 h).

À première vue, les deux systèmes de mesures présentent une forte similitude, notamment sur la localisation des pluies. En termes de cumul, certaines régions montrent toutefois des différences. Les plus marquantes ont été mises en exergue sur la Figure III-11b au travers des rectangles hachurés. Nous pouvons également constater que le RADAR détecte la présence de très fines structures de pluie qui, localement, possèdent d'importants cumuls (*cf.* ellipses). Notons que ces zones sont dépourvues de pluviomètres au sol. Malgré les corrections apportées lors du post-traitement, certaines mesures RADAR douteuses semblent persister (*cf.* pentagone marqué d'un point d'interrogation).

La Figure III-12 montre le champ spatial de pluie reconstruit respectivement par les techniques du krigeage ordinaire des données pluviométriques seules (OK\_RG), du krigeage avec dérive externe (KED) et la fusion conditionnée (MERG).

L'interpolation par OK\_RG (Figure III-12a) dessine des structures de pluie très régulières. La combinaison des données par KED (Figure III-12b) ou MERG (Figure III-12c) intègre l'hétérogénéité spatiale des précipitations mesurée avec le RADAR (Figure III-11b). De plus, les petites cellules pluvieuses observées uniquement par le RADAR (*cf.* ellipses blanches) se redessinent dans le champ final (tout comme les pixels RADAR que nous considérons comme douteux – *cf.* pentagone). Enfin, les cumuls estimés semblent avoir été ajustés afin de s'accorder davantage avec les observations pluviométriques.

L'influence des pluviomètres ou du RADAR est observable sur tout le champ final. Pour mieux se rendre compte des effets associés à cette combinaison, la Figure III-13 montre un zoom du domaine étudié au niveau du coin inférieur gauche. Nous remarquons alors que :

- pour les mesures RADAR (Figure III-13a) présentant une surestimation (sous-estimation), les valeurs finales par KED ou MERG (Figure III-13 c et d) ont été abaissées (augmentées) pour tendre vers les observations pluviométriques (rectangles hachurés) ;
- dans les zones dépourvues de stations pluviométriques au sol, la structure du champ final s'apparente davantage aux mesures RADAR (ellipses).

Notons que, pour les deux techniques (KED et MERG), l'influence de la variable auxiliaire semble différente. Ici, le MERG paraît plus sensible aux mesures RADAR que le KED. Ce propos est illustré à l'intérieur de l'ellipse où le maximum de pluie enregistré par le RADAR est mieux reproduit au travers de la technique du MERG. Le champ de pluie mesuré par le RADAR et celui résultant du MERG sont proches, alors que le champ obtenu par KED semble être davantage contraint par les observations pluviométriques (*cf.* octogone)

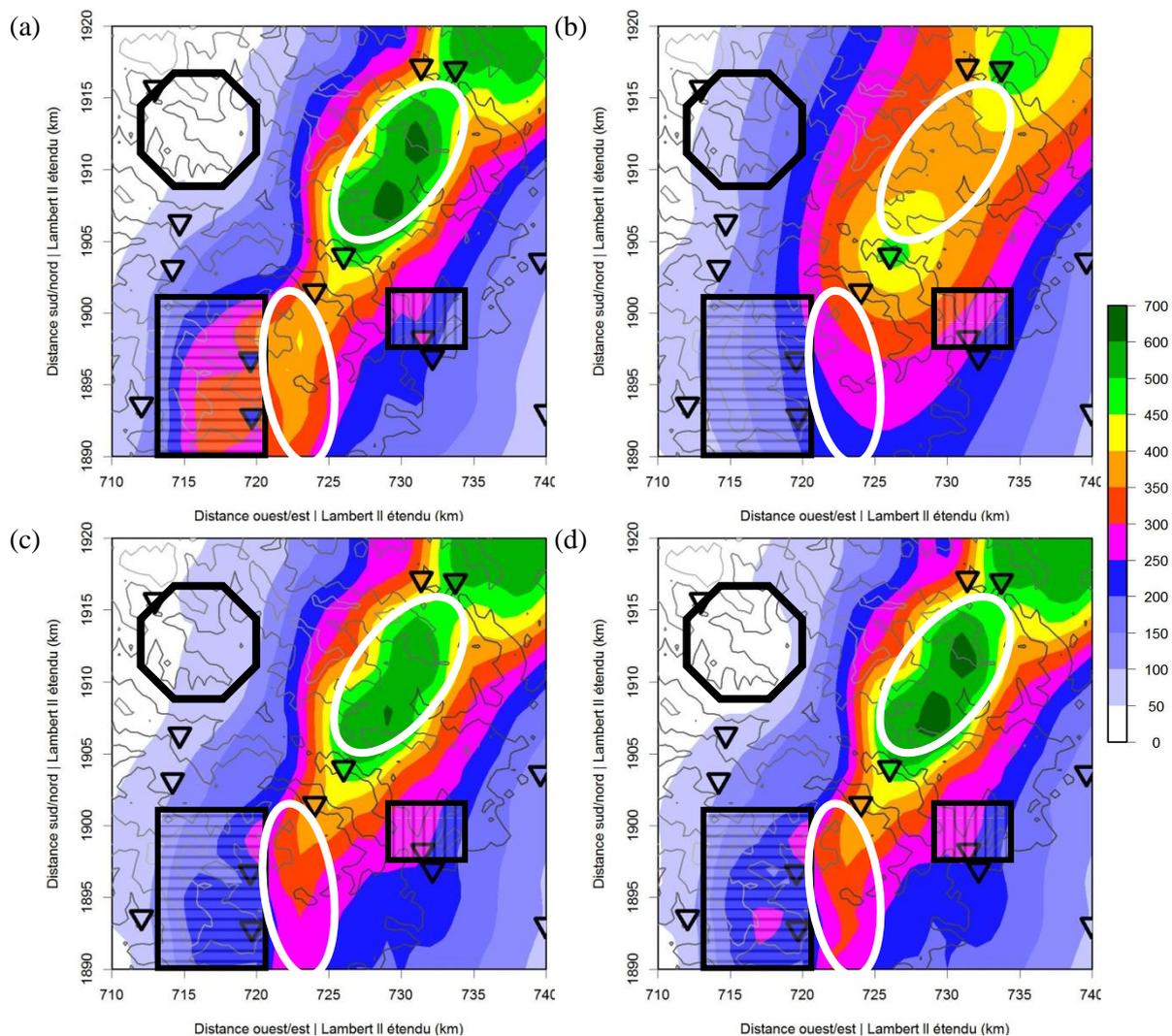


Figure III-13 : CAS A = Zoom du champ de pluie (en mm) (a) mesuré par le RADAR et interpolé par (b) OK\_RG, (c) KED et (d) MERG. (T = 72 h).

Les observations pluviométriques sont reportées par les valeurs contenues dans les triangles pointant vers le bas. Les rectangles hachurés horizontalement (verticalement) montrent les conséquences de l'abaissement (l'augmentation) des mesures RADAR pour tendre vers les observations pluviométriques. Les ellipses localisent les inclusions les plus marquantes des structures uniquement observées par le RADAR. L'octogone localise des effets aux origines antinomiques sur les champs issus de la combinaison de données.

Les deux techniques montrent des résultats prometteurs. La grande ressemblance entre le champ résultant du KED et du MERG ne permet pas cependant de les dissocier. Toutefois, il semblerait qu'une distinction peut être faite quant à la sensibilité du MERG aux mesures RADAR. Nous retiendrons que la technique du KED et du MERG prennent en considération i) la valeur des observations pluviométriques et ii) les structures hétérogènes des mesures RADAR notamment dans les zones situées hors de la portée des pluviomètres (contrairement à la technique du OK\_RG).

### III-3.4.2 Validation quantitative

Le Tableau III-1 donne une première comparaison entre les différents champs de précipitation sur la valeur maximale, la proportion de pixels ayant un cumul supérieur à 50 mm (soit la première classe de l'échelle de couleur considérée) et la quantité totale de pluie tombée au sol.

Comme constaté précédemment, le maximum de pluie enregistré par le RADAR est mieux reproduit par la technique du MERG (+6.7 mm) que par le KED (-46.1 mm). En comparaison, nous notons une différence de -162.4 mm entre le maximum rapporté par la technique OK\_RG et les mesures RADAR.

Le champ de précipitation obtenu par le KED, plus lisse que celui déduit par la technique du MERG, semble être (très légèrement) plus large. Cette remarque est faite en comparant la surface du champ de pluie. En effet, nous obtenons une différence de -0.5 % entre le champ du KED et du MERG. Ramenée à l'échelle du domaine, cette différence équivaut à une surface de -64 km<sup>2</sup>. À titre de comparaison, le champ de pluie mesuré par le RADAR est plus grand de 465 et 529 km<sup>2</sup> par rapport au champ du KED et du MERG respectivement. Par analogie l'interpolation par OK\_RG donne un champ plus grand de 260 et 324 km<sup>2</sup>.

Nous noterons enfin qu'en termes de quantité totale de pluie tombée sur le domaine, le champ obtenu par KED et MERG sont plus faibles d'à peine plus de 1% avec le champ mesuré par le RADAR. En comparaison avec le champ OK\_RG, nous notons une différence plus faible de près de 4 à 5 mégatonnes d'eau respectivement.

Tableau III-1 : CAS A = Comparatif des différents champs spatiaux obtenus. (T = 72 h). La surface du champ de pluie correspond au pourcentage du nombre de pixels ayant une valeur supérieure ou égale à 50 mm (première classe de l'échelle de couleur considérée).

	RADAR	OK_RG	KED	MERG
Valeur maximale	660.4 mm	498.0 mm <sup>Ⓢ</sup>	614.3 mm	667.1 mm
Surface du champ de pluie	42.7 %	41.3 %	39.5 %	39.0 %
Quantité de pluie tombée	1.061 Gt	1.052 Gt	1.048 Gt	1.047 Gt

<sup>Ⓢ</sup> Notons que la valeur maximale du OK\_RG est inférieure à celle observée à La-Grand-Combe. Cette différence s'explique par le fait que ce pluviomètre n'est pas un point de la grille d'interpolation considérée, la valeur maximale est donc lissée lors de l'interpolation.

Nous avons voulu pousser plus loin notre analyse. Le Tableau III-2 présente trois scores statistiques résultants de la comparaison des observations au sol avec, d'une part la valeur du pixel du champ RADAR le plus proche des pluviomètres et d'autre part les estimations aux points des pluviomètres par OK\_RG, KED et MERG selon la technique du « leave-one-out » (LOO, cf. sous-paragraphe II-2.1.1-d). Nous considérons ici les cumuls observés par les pluviomètres comme valeurs de référence.

Première constatation : seules les mesures RADAR sous-estiment le champ de pluie par rapport aux pluviomètres (erreur moyenne négative). La forte valeur de la RMSE signifie la présence de très larges erreurs en moyenne. Pourtant ces mesures concordent globalement avec celles des pluviomètres (*i.e.* très fort coefficient de corrélation). Les estimations par OK\_RG présentent une légère surestimation (ME positive et plus faible que pour les mesures RADAR) mais ces erreurs se compensent puisqu'elles impliquent une très forte RMSE (plus importante que les mesures RADAR). Enfin, le caractère spatial des observations pluviométriques semble être conservé lors de l'interpolation (*i.e.* très fort coefficient de corrélation mais légèrement moins bon que pour les mesures RADAR).

Les résultats obtenus par KED et MERG montrent une légère amélioration des scores par rapport à la technique du OK\_RG. La comparaison des deux techniques montre que : i) le MERG présente la plus faible erreur moyenne (-0.2 mm) et ii) ce sont les estimations par KED qui présentent une plus faible compensation des écarts (RMSE = -0.7 mm). Enfin, c'est également la technique du KED qui représente le mieux les variations spatiales observées par les pluviomètres avec une plus forte corrélation (à  $10^{-3}$  près).

Tableau III-2 : CAS A = Comparaison entre les observations du réseau principal de pluviomètres et les estimations des différents champs spatiaux. (T = 72 h).  
Les indications (-) et (+) indiquent respectivement le moins bon et le meilleur score.

	RADAR	OK_RG	KED	MERG
Erreur moyenne – ME	-1.5 mm (-)	0.8 mm	0.3 mm	0.1 mm (+)
RMSE	31.7 mm	32.1 mm (-)	27.6 mm (+)	28.3 mm
Coefficient de corrélation	0.953	0.951 (-)	0.965 (+)	0.963

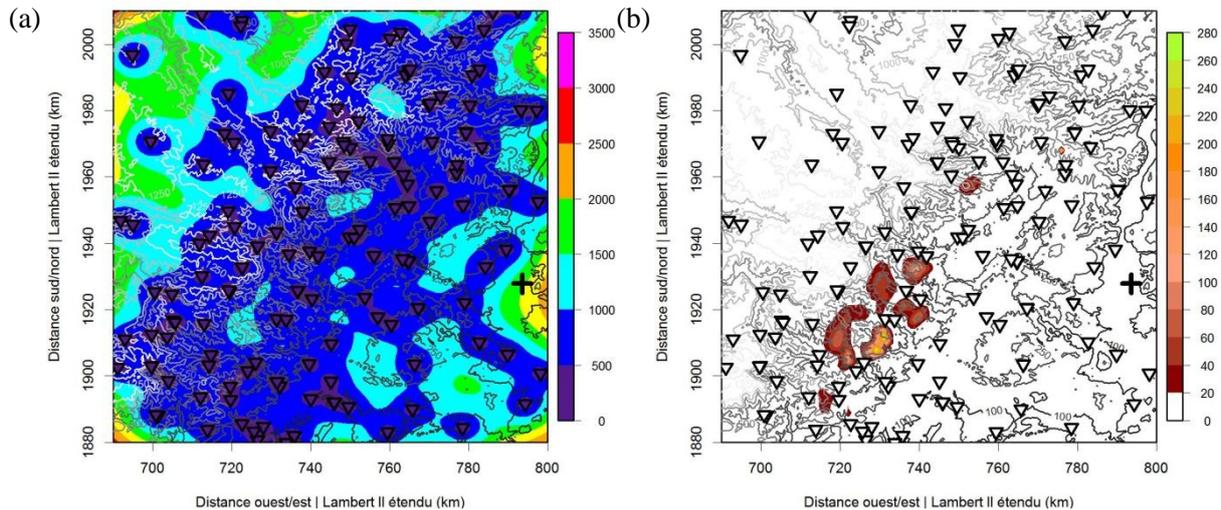


Figure III-14 : CAS A = (a) Champ de la variance de krigeage (en  $\text{mm}^2$ ) associée au krigeage ordinaire et (b) différence (en  $\text{mm}^2$ ) entre la variance de krigeage associée au KED et celle associée au krigeage ordinaire. ( $T = 72 \text{ h}$ ).

*La variance de krigeage à la position des pluviomètres est reportée dans les triangles pointant vers le bas.*

Comme nous l'avons mentionné dans le paragraphe II-2.1, le krigeage fournit une valeur estimée en tout point de l'espace souhaité. L'erreur commise est inconnue mais la variance de cette erreur, appelée variance de krigeage (ou VK), est calculée systématiquement.

La Figure III-14a présente le champ de la VK associée à la technique du krigeage ordinaire ( $\text{Var}_{OK}$  – équation [II-8] – associée à OK\_RG et au MERG). Nous observons bien qu'à proximité des pluviomètres,  $\text{Var}_{OK}$  est la plus faible et devient importante à mesure que nous nous en éloignons. En limite de domaine, elle est maximale à cause d'un manque d'information (*i.e.* stations voisines). Sur tout le champ, la VK moyenne est estimée à  $939 \text{ mm}^2$ , soit près de 12 fois inférieure à la variance de l'échantillon ( $= 10\,925 \text{ mm}^2$ ).

La VK associée à la technique du KED ( $\text{Var}_{KED}$  – équation [II-10]) étant très proche de  $\text{Var}_{OK}$ , nous avons décidé de la représenter sous forme de différence (Figure III-14b). Nous constatons alors que les dissemblances se localisent principalement dans les zones où sévissent de petites structures de pluie observées par le RADAR (*cf.* Figure III-11b et Figure III-13a).

Pour résumer, les techniques combinant les observations sont donc à privilégier. Qualitativement, les estimations et la variance de krigeage du KED et du MERG sont très proches. Néanmoins, nous avons souligné une plus forte sensibilité des mesures RADAR dans les estimations du MERG, notamment dans des zones dépourvues de pluviomètres.

Quantitativement, les quelques différences notifiées entre les deux techniques (KED et MERG) restent insignifiantes pour que nous puissions réellement les départager au pas de temps de l'événement.

### III-3.5 Reconstruction d'un champ de précipitation à différentes échelles temporelles

Après avoir testé les performances des différentes techniques au pas de temps de l'événement, nous nous sommes intéressés à des échelles de temps plus petites. La Figure III-15 présente trois scores statistiques obtenus par comparaison des observations pluviométriques avec les mesures RADAR (courbes bleues) et les estimations LOO par OK\_RG (courbes jaunes), KED (courbes vertes) et MERG (courbes roses) pour des durées  $T$  de cumul allant de 1 à 24 heures. Ici encore, nous considérons les observations pluviométriques comme valeurs de référence.

Quelle que soit la technique considérée, l'erreur moyenne (ME – Figure III-15a) montre une augmentation avec la durée de cumul  $T$ . Comme remarqué précédemment, seules les mesures RADAR présentent une sous-estimation (*i.e.* valeurs négative) qui s'amplifie presque linéairement avec  $T$ . Concernant les techniques d'interpolation spatiale, la progression de la ME en fonction de  $T$  est moins marquée. À titre indicatif, entre  $T = 1$  et 24 heures, nous relevons une augmentation d'un facteur 8.5 pour l'OK\_RG, 5.5 pour le KED et de seulement 3.3 pour le MERG. Face à la technique du KED, le MERG obtient les plus faibles valeurs de ME, avec des valeurs ne dépassant pas 0.2 mm quelle que soit  $T$ .

Comme la ME, tous les champs affichent également une augmentation de la RMSE avec la période de temps (Figure III-15b). Nous remarquons que la progression est d'autant plus marquée pour les petits pas de temps (*i.e.*  $T$  inférieure à 6 heures). La technique du OK\_RG est affectée par de larges erreurs, plus importantes en moyenne que les erreurs associées aux mesures RADAR. L'utilisation du KED ou du MERG abaisse sensiblement les valeurs de la RMSE mais aucune distinction n'est pas encore observable entre ces deux techniques.

Enfin, le coefficient de corrélation (Figure III-15c) tend progressivement vers de meilleures valeurs (*i.e.* plus proches de 1) en même temps que la durée du pas de temps croît pour toutes les techniques. De plus, tous les scores affichent initialement ( $T = 1$  h) une forte corrélation (valeurs supérieures à 0.8). Nous constatons ici également que les plus fortes augmentations se manifestent pour les petits pas de temps. Quelle que soit la durée de cumul

considérée, les mesures RADAR sont plus cohérentes avec les observations pluviométriques que ne le sont les estimations faites par OK\_RG (*i.e.* courbe jaune sous la courbe bleue). Enfin, la technique du KED ou du MERG obtiennent toujours les meilleurs résultats. Les différences entre les estimations par KED et MERG sont ici encore trop petites (entre 0.9 et  $3.3 \cdot 10^{-3}$ ) pour pouvoir vraiment les dissocier.

Finalement, la combinaison des données (pluviomètres et RADAR) présente des estimations de meilleure qualité (notamment en termes de corrélation) par rapport aux mesures RADAR seules ou au krigeage des données de pluviomètres (OK\_RG). Les estimations par la technique du MERG présentent un biais (ME) plus faible en comparaison à celles du KED alors que les erreurs quadratiques moyennes (RMSE) sont du même ordre.

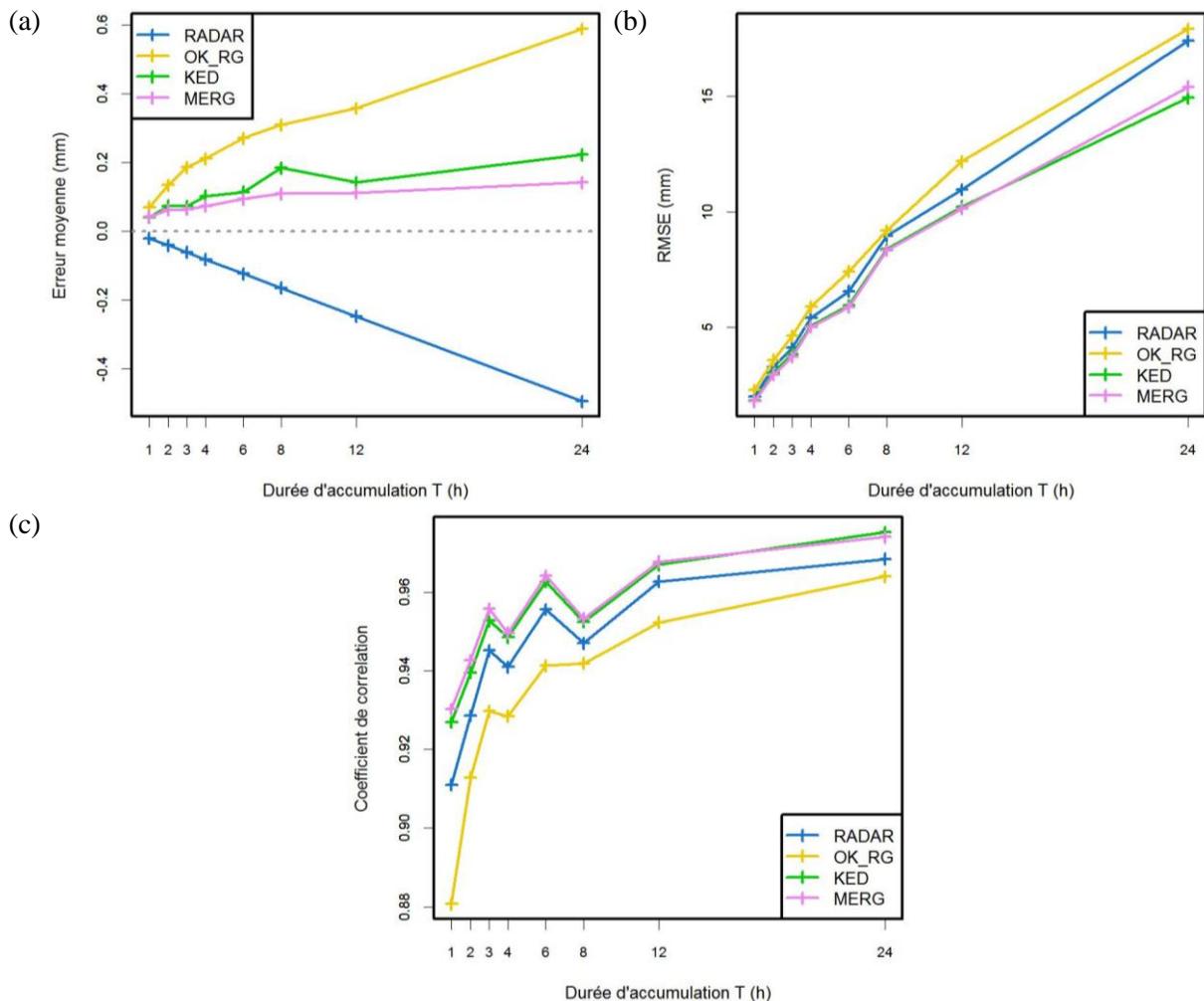


Figure III-15 : CAS A = Évolution (a) de l'erreur moyenne, (b) de la RMSE et (c) du coefficient de corrélation en fonction de la fenêtre temporelle T (en heure).  
 Remarque : La ligne en pointillé gris affiche la valeur du score indiquant une « parfaite » correspondance entre les estimations et les observations.

Tableau III-3 : CAS A = Comparaisons entre les observations du réseau secondaire de pluviomètres et les estimations des différents champs spatiaux. (T = 24 h).  
 Les indications (-) et (+) indiquent respectivement le moins bon et le meilleur score.

	RADAR	OK_RG	KED	MERG
Erreur moyenne – ME	-1.5 mm (-)	-0.5 mm (+)	-1.3 mm	-1.5 mm
RMSE	20.6 mm	21.7 mm (-)	17.2 mm	16.6 mm (+)
Coefficient de corrélation	0.946	0.935 (-)	0.960	0.962 (+)

En considérant un pas de temps T = 24 h, il nous est alors possible d'utiliser le second réseau de pluviomètres couvrant notre domaine d'étude (Figure III-7, marqueurs bleus). Le Tableau III-3 présente trois scores statistiques obtenus en comparant les observations des pluviomètres quotidiens avec les estimations obtenues par le cumul des mesures horaires.

La première remarque qui peut être faite est que, quelle que soit la technique employée, toutes les estimations sont sous-estimées (*i.e.* erreurs moyennes négatives). Les résultats obtenus par la technique du OK\_RG sont les moins biaisées (ME la plus faible) pourtant ces erreurs semblent se compenser (plus forte valeur de la RMSE). À l'inverse, même si les mesures RADAR présentent la plus grande ME, les grands écarts sont légèrement moins prédominants (RMSE plus faible). Ici encore, les estimations par KED ou MERG obtiennent les meilleurs résultats. La technique du MERG possède certes une plus grande ME (+0.2 mm) par rapport au KED mais la RMSE (-0.6 mm) et le coefficient de corrélation ( $+2 \cdot 10^{-3}$ ) sont optimisés. Notons que ces résultats s'accordent (à l'exception de la ME) avec ceux obtenus par la validation LOO (Figure III-15).

Pour conclure sur ce premier cas d'étude, la combinaison des observations pluviométriques et mesures RADAR présente un avantage certain, tant visuellement que quantitativement. Nous avons pu constater que les estimations par KED ou MERG donnent des résultats satisfaisants, de surcroît s'améliorant avec la durée de cumul T. Par validation LOO, les écarts entre les deux techniques sont insignifiants pour pouvoir les départager. La validation avec un réseau de pluviomètres secondaires conduit à une légère préférence de la technique du MERG qui enregistre la plus faible RMSE et un coefficient de corrélation légèrement meilleur. La principale différence entre le KED et le MERG semble se jouer sur l'intégration des mesures RADAR et notamment sur la reproduction des petites structures de pluie dans des zones dépourvues de pluviomètres. Il semble s'avérer que l'influence de ces

mesures paraissait plus importante dans la technique du MERG que dans le KED. Ceci étant, nous notons pour ce cas de meilleures estimations quantitatives par le RADAR en comparaison aux pluviomètres (OK\_RG).

## III-4 CAS B : du 31 Octobre au 2 Novembre 2008

### III-4.1 Situation météorologique

Dix jours plus tard, un nouvel épisode de pluie touche à nouveau les Cévennes. Un profond thalweg d'altitude aborde le pays à partir du 31 octobre et isole une goutte froide sur la France et l'Espagne dans la journée du 2 novembre (Audard-Vincendon 2010, p. 98) Sur les cartes d'analyses (Figure III-16), une dépression (croix bleue) se forme sur la péninsule Ibérique et stationnera un temps sur l'ouest de la France accompagnée d'un fort flux de sud/sud-est sur le pourtour méditerranéen (flèches orange). La présence de zones de haute pression (croix rouges) au dessus de l'Europe occidentale et l'est de l'Atlantique Nord ralentit la propagation de la perturbation. Une zone frontale s'étend du nord de l'Espagne vers le nord de l'Europe centrale et se décale progressivement en direction du nord-nord-est.

La barrière montagneuse des Cévennes-Vivarais et les forts vents du sud favorisent de fortes ascendances. L'advection d'air chaud et humide en provenance de la Méditerranée génère d'intenses cellules orageuses successives sur les Cévennes (Figure III-17). Après d'importants cumuls de pluie dans la journée du 1er novembre, l'activité pluvieuse se décale vers la région Rhône-Alpes (au nord-est) et Provence-Alpes-Côte d'Azur (au sud-est) dans la soirée du 2 novembre.

Notre zone d'intérêt s'étend sur 100km d'ouest en est et sur 130km du nord au sud (Figure III-18). L'évènement est observé par 311 stations au pas de temps horaires.

### III-4.2 Le contrôle de qualité des données pluviométriques

L'application du contrôle de qualité permet d'écarter deux observations issues de deux stations différentes : la station #55 (Cavaillon) et la station #63 (Chateauneuf-du-Pape). La Figure III-19 et la Figure III-20 montrent ces observations douteuses accompagnées des observations mesurées à plus ou moins 2 heures ainsi que celles des cinq stations voisines les plus représentatives. Pour ces deux cas, les valeurs observées dépassent respectivement plus de 3 et 7 fois les valeurs attendues (équation [II-1]).

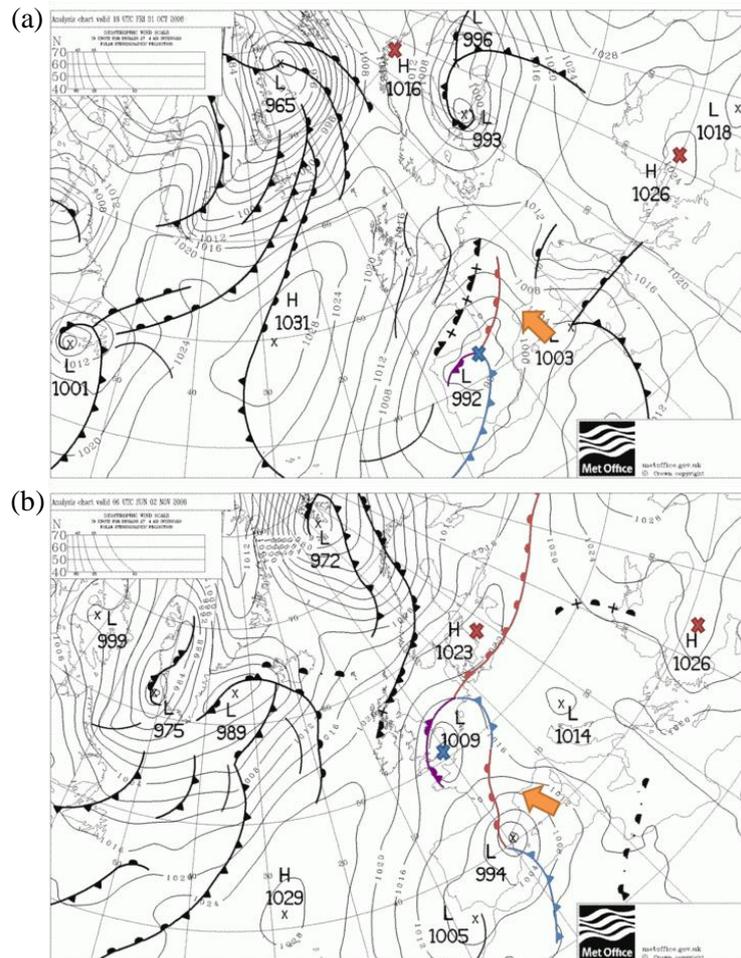


Figure III-16 : Carte d'analyse (a) du 31 Octobre 2008 à 18 h UTC et (b) du 2 Novembre 2008 à 06 h UTC.

Source : Met Office (archive : <http://www.meteo60.fr/archives.php> | coloriées)

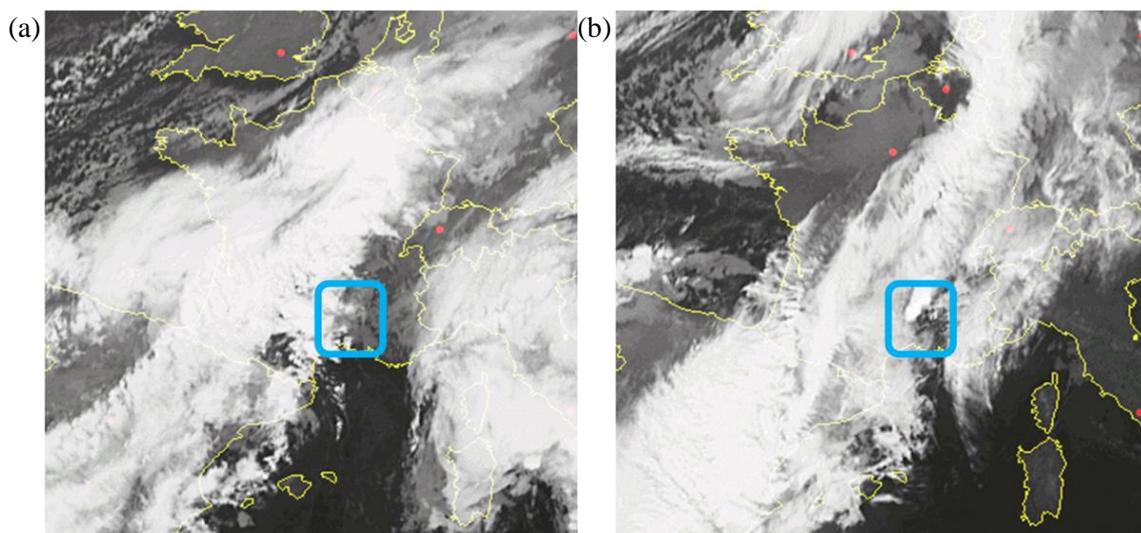


Figure III-17 : Images satellites infrarouge (a) du 31 Octobre 2008 à 20 h 15 UTC et (b) du 1er Novembre 2008 à 22 h 15 UTC.

Source : EUMETSAT (archive : <http://www.meteo60.fr/archives.php> | zoomées)

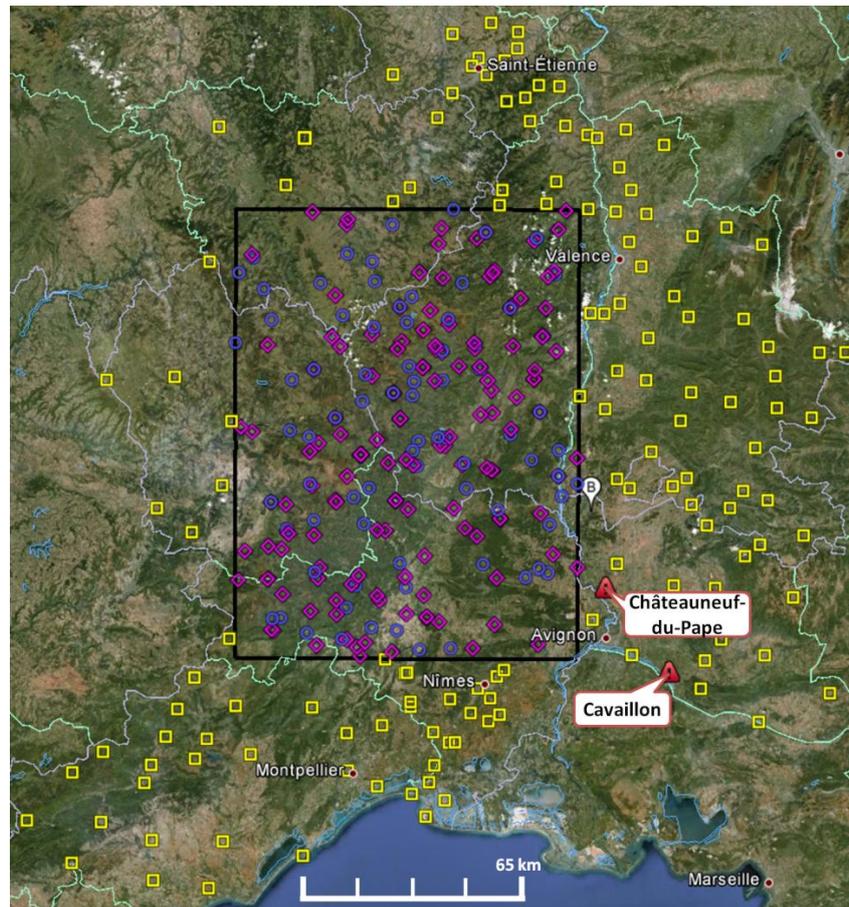


Figure III-18 : Localisation des instruments pour le cas B = du 31 Octobre (06 h UTC) au 2 Novembre 2008 (06 h UTC).

Le rectangle noir représente la zone d'intérêt couverte par 142 pluviomètres horaires (marqueurs roses) et 79 pluviomètres journaliers (marqueurs bleus). En dehors de cette zone, 142 pluviomètres horaires sont également disponibles (marqueurs jaunes). La lettre B dans le marqueur blanc indique l'emplacement de l'unique RADAR de Bollène disponible pour ce cas. Les marqueurs rouges localisent les stations horaires où au moins une observation est considérée comme suspecte d'après le contrôle de qualité appliqué. Source : Google Earth.

### III-4.3 Informations complémentaires

Pour ce cas, la relation  $Z$ - $\mathfrak{R}$  (équation [I-5]) utilisée est (Bouilloud *et al.* 2010; Boudevillain 2011) :

$$Z = 75 \cdot \mathfrak{R}^{1.85}$$

Au pas de temps de l'événement ( $T = 48$  h), les variogrammes paramétriques considérés, identiques pour les observations (technique du krigeage ordinaire) et les résidus (technique du KED), sont représentés par la Figure III-21.

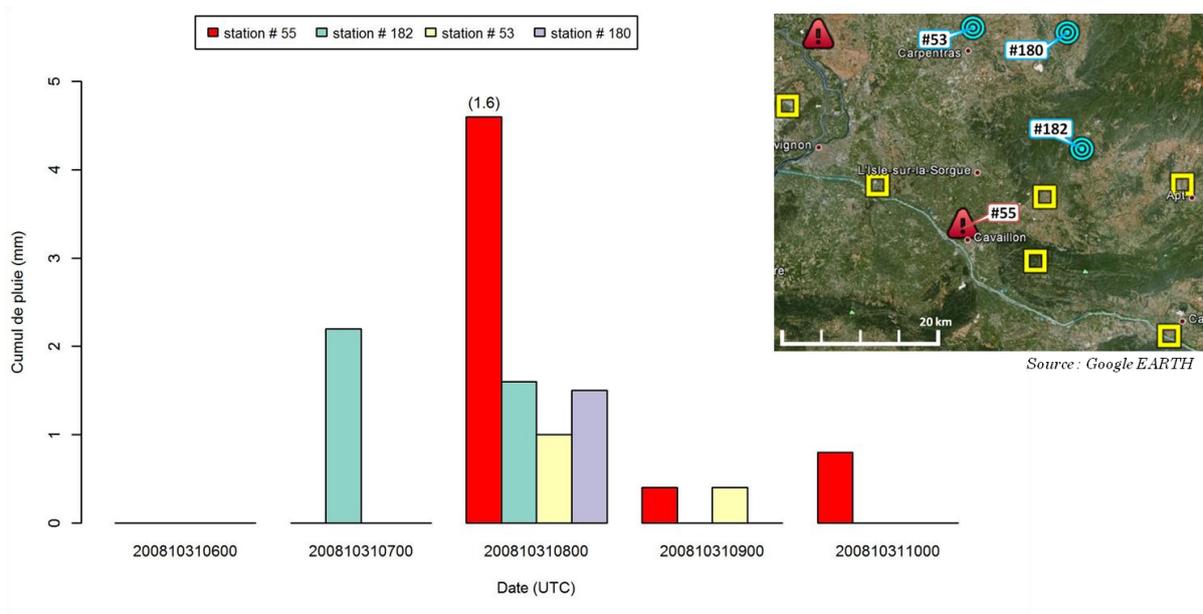


Figure III-19 : Évolutions temporelles des cumuls horaires de la station #55 (marqueur rouge sur la carte de droite) et ses 5 plus représentatives stations voisines (marqueurs cyans sur la carte de droite) du 31 Octobre 2008 de 06 h UTC à 10 h UTC. L'observation à 08 h UTC est celle suspectée. L'indication entre parenthèses donne une approximation de la valeur attendue.

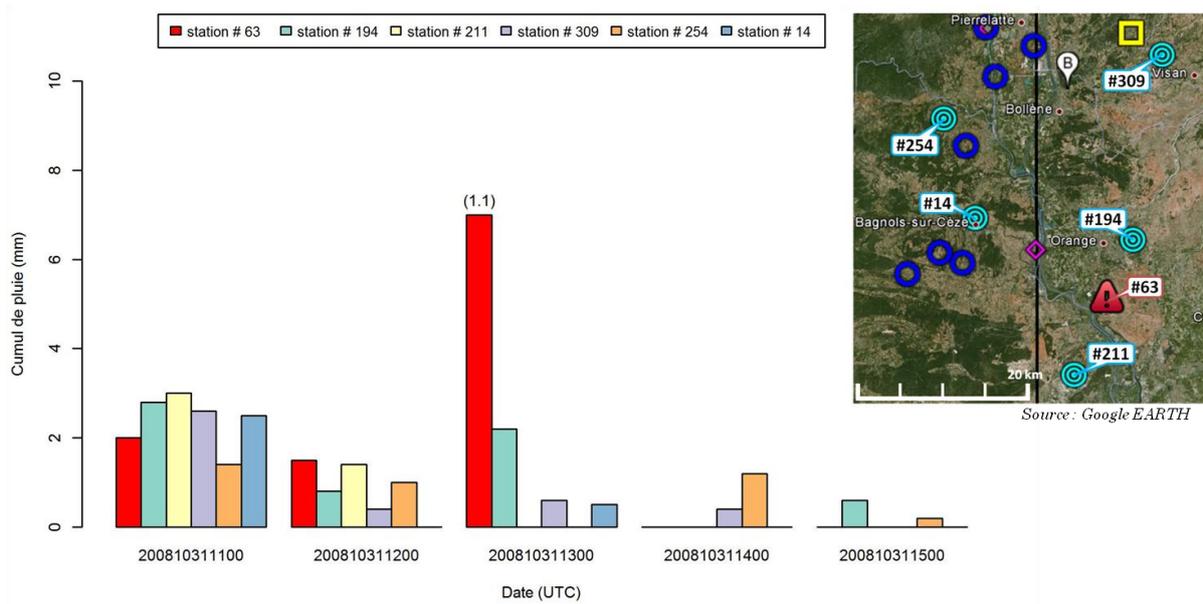


Figure III-20 : Évolutions temporelles des cumuls horaires de la station #63 (marqueur rouge sur la carte de droite) et ses 5 plus représentatives stations voisines (marqueurs cyans sur la carte de droite) du 31 Octobre 2008 de 11 h UTC à 15 h UTC. L'observation à 13 h UTC est celle suspectée. L'indication entre parenthèses donne une approximation de la valeur attendue.

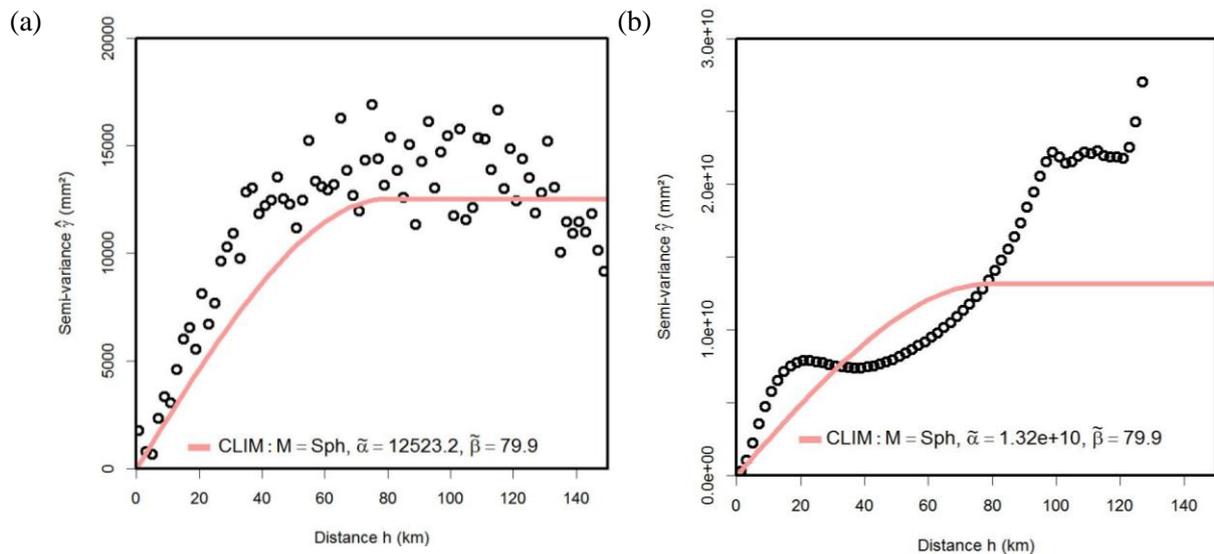


Figure III-21 : CAS B = Semi-variogramme expérimental (a) des observations pluviométriques et (b) des résidus obtenus à partir des mesures RADAR. ( $T = 48$  h). Les variogrammes paramétriques considérés sont représentés par les courbes rouges.

## III-4.4 Reconstruction d'un champ de précipitation à l'échelle de l'événement

### III-4.4.1 Validation qualitative

La Figure III-22 présente les valeurs des observations pluviométriques et les mesures RADAR pour ce second cas d'étude. Sur les 2 jours consécutifs, les précipitations se sont accumulées sur les contreforts du massif selon un axe sud-ouest/nord-ouest. Le maximum de précipitation est enregistré à Villefort (Lozère) avec un cumul dépassant 550 mm en 48 h.

Les pluies se localisent le long du relief et les maximums s'observent sur les pentes au vent. Les observations pluviométriques et mesures RADAR présentent une certaine similitude. Comme évoqué dans le cas A, l'avantage de l'utilisation RADAR est la détection de petites cellules précipitantes dans des zones dépourvues d'instrument au sol (*cf.* ellipses). Seulement, nous notons pour ce cas la présence d'importantes discordances en termes de quantité. Pour les mesures RADAR, les précipitations s'étendent sur une large zone en aval du relief avec des cumuls supérieurs à 50 – 100 mm (sur 48 heures) or celles-ci n'apparaissent pas dans les observations pluviométriques (*cf.* zone hachuré horizontalement). À l'inverse, seuls les pluviomètres mesurent de la pluie en amont du domaine (*cf.* zone hachurée verticalement).

La Figure III-23 montre le champ de pluie au sol reconstruit par la technique du OK\_RG, du KED et du MERG.

Comme pour le cas A, la combinaison des données par KED (Figure III-23b) ou MERG (Figure III-23c) permet d'intégrer les hétérogénéités du champ de pluie mesurées qui sont observés uniquement par le RADAR (Figure III-22b). Ces détails ne sont pas retrouvés dans l'interpolation des observations pluviométriques par OK\_RG (Figure III-23a). Nous constatons par ailleurs que les valeurs des cumuls ont été ajustées de sorte que les deux zones sans pluie en amont (*cf.* mesures RADAR) et aval (*cf.* pluviomètres) du relief soient redessinées dans le champ final (zones hachurées).

Les résultats obtenus par la technique du KED et du MERG sont encore très similaires. En zoomant au centre du domaine (Figure III-24), nous pouvons mieux appréhender les différences entre les quatre estimations faites. Comme pour le cas A, nous remarquons que :

- pour les régions où les mesures RADAR (Figure III-24a) présentent une surestimation (sous-estimation) des cumuls par rapport aux observations pluviométriques (OK\_RG), les estimations par KED ou MERG (Figure III-24 c et d) ont été abaissées (augmentées) pour s'aligner aux observations pluviométriques (*cf.* rectangles hachurés) ;
- loin des pluviomètres, la structure du champ final s'apparente davantage aux mesures RADAR (*cf.* ellipses).

#### III-4.4.2 Validation quantitative

Le Tableau III-4 présente une première comparaison entre les différents champs de pluie en termes de valeur maximale, proportion de pixels ayant un cumul supérieur à 50 mm (soit la première classe de l'échelle de couleur considérée) ou quantité totale de pluie tombée au sol. Nous montrons alors que :

- le cumul de pluie maximum enregistré par OK\_RG présente, en comparaison avec les mesures RADAR, une différence +34.5 mm (*contre -162.4 mm pour le cas A*). C'est le MERG qui affiche les plus faibles différences (soit +17.6 mm par rapport à OK\_RG et +52.1 mm par rapport aux mesures RADAR *contre +6.7 mm pour le cas A*) par rapport au KED (soit +33.6 mm par rapport à OK\_RG et +68.1 mm par rapport aux mesures RADAR *contre -46.1 mm pour le cas A*) ;

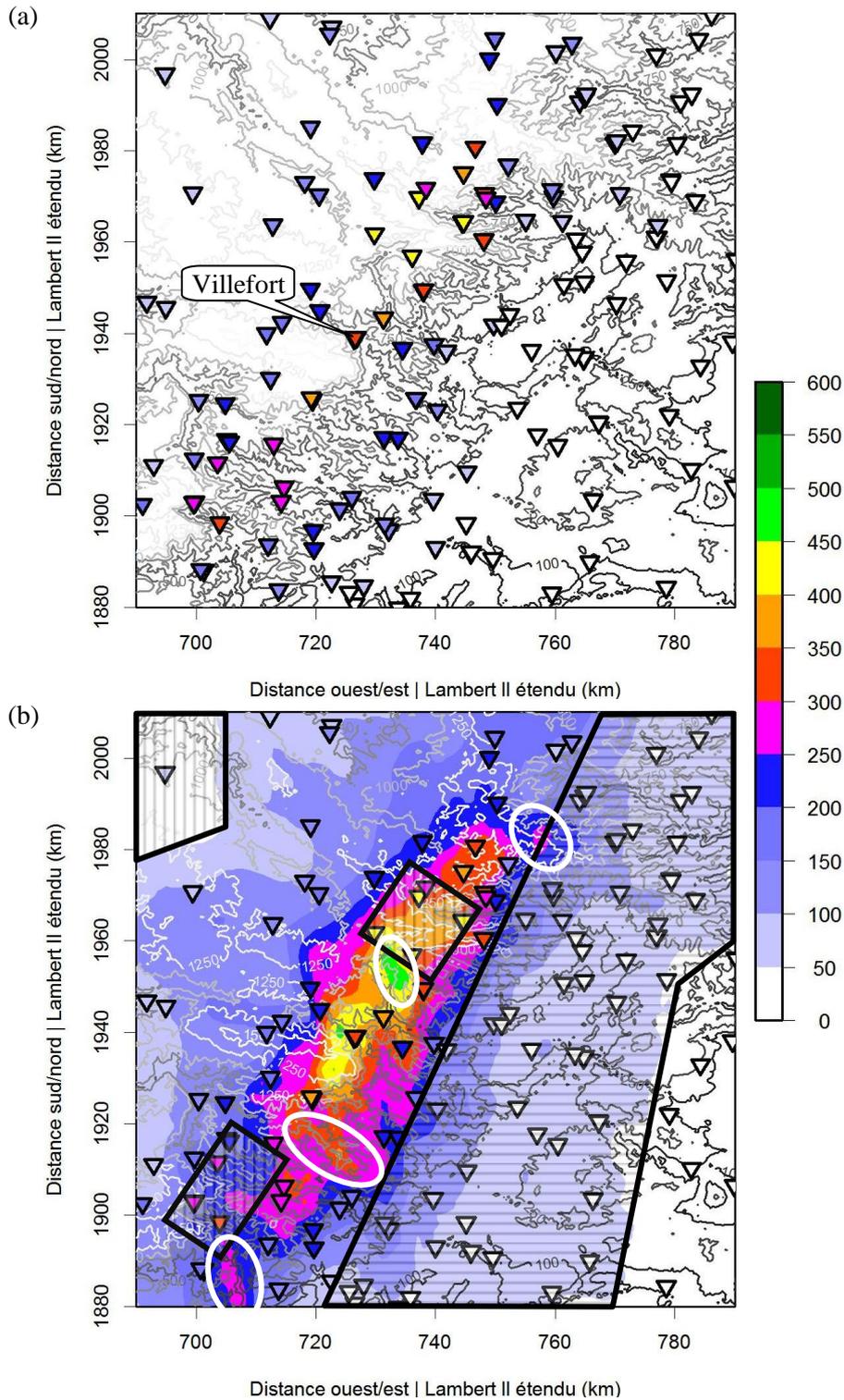


Figure III-22 : CAS B = Champ de pluie (en mm) observé par (a) les pluviomètres et (b) le RADAR. (T = 48 h).

La topographie de la région est rajoutée sous forme d'iso-contours en nuances de gris, du plus foncé pour les altitudes proches de la surface de la mer, au plus clair pour les altitudes atteignant les 1 250 m. Sur la figure (b), les observations pluviométriques sont reportées par les valeurs contenues dans les triangles pointant vers le bas. Les rectangles hachurés horizontalement (verticalement) montrent une surestimation (sous-estimation) des mesures RADAR par rapport aux pluviomètres. Les ellipses blanches localisent de petites structures de pluie non détectables par le réseau de pluviomètres.

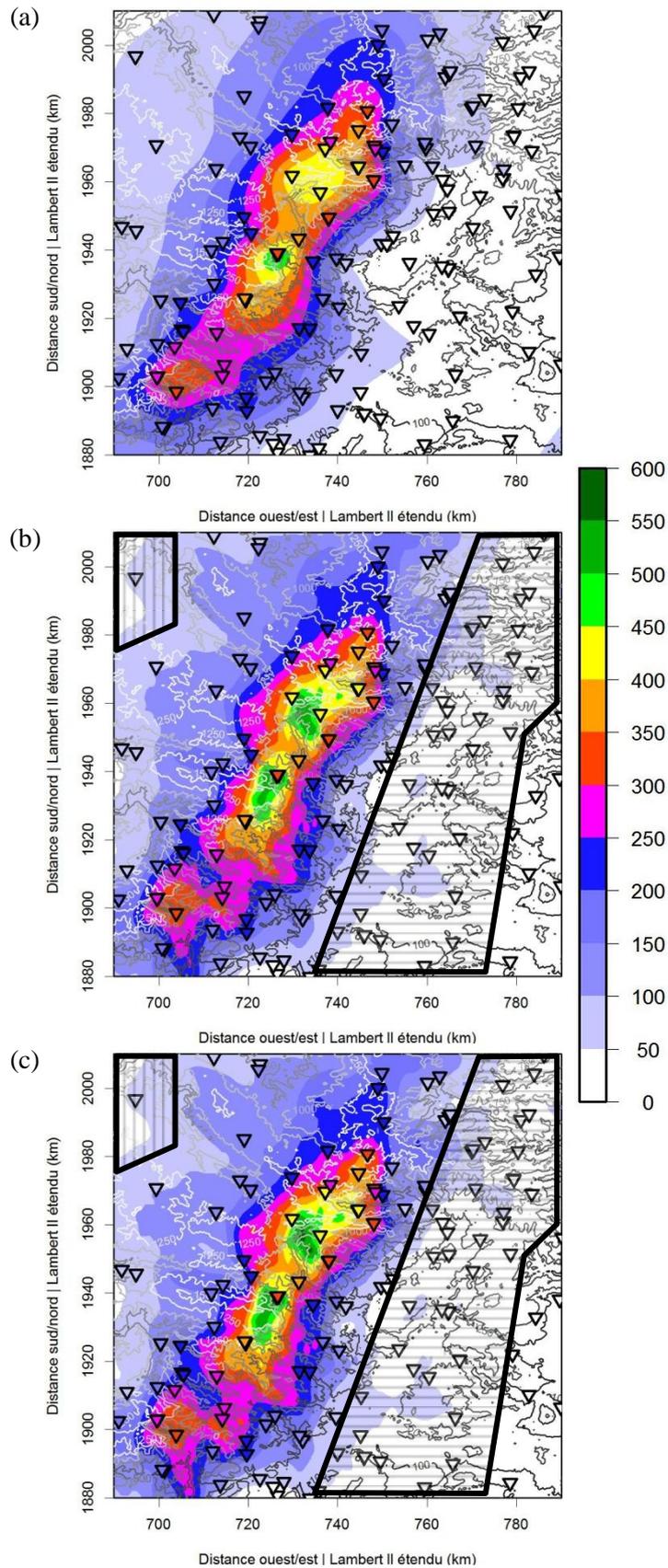


Figure III-23 : CAS B = Champ de pluie (en mm) interpolé par (a) OK\_RG, (b) KED et (c) MERG. (T = 48 h).

*Les hachures horizontales (verticales) montrent l'ajustement des cumuls vers les pluviomètres (RADAR).*

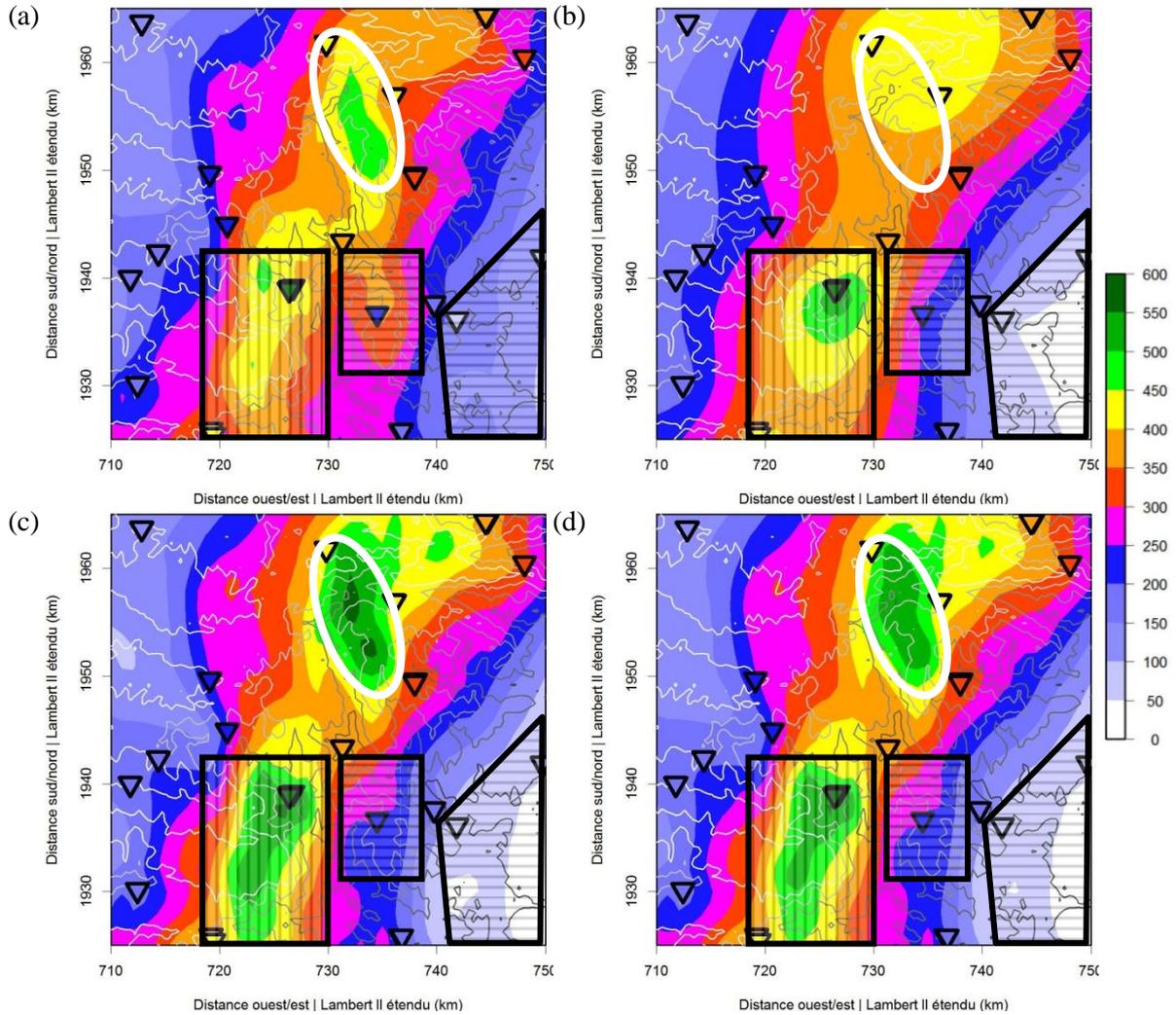


Figure III-24 : CAS B = Zoom du champ de pluie (en mm) (a) mesuré par le RADAR et interpolé par (b) OK\_RG, (c) KED et (d) MERG. (T = 48 h).

Les observations pluviométriques sont reportées par les valeurs contenues dans les triangles pointant vers le bas. Les rectangles hachurés horizontalement (verticalement) montrent les conséquences de l'abaissement (l'augmentation) des mesures RADAR pour tendre vers les observations pluviométriques. L'ellipse localise l'inclusion la plus marquante d'une structure observée par le RADAR.

Tableau III-4 : CAS B = Comparatif des différents champs spatiaux obtenus. (T = 48 h). La surface du champ de pluie correspond au pourcentage du nombre de pixels ayant une valeur supérieure ou égale à 50 mm (première classe de l'échelle de couleur considérée).

	RADAR	OK_RG	KED	MERG
Valeur maximale	501.8 mm	536.3 mm	569.9 mm	553.9 mm
Surface du champ de pluie	86.5 %	67.3 %	66.5 %	66.6 %
Quantité de pluie tombée	1.777 Gt	1.639 Gt	1.647 Gt	1.646 Gt

- les surfaces des champs de pluie obtenus par KED ou MERG sont très proches de celle estimée par OK\_RG (surface équivalente plus petite de 110 et 97 km<sup>2</sup> respectivement *contre 260 et 324 km<sup>2</sup> pour le cas A*). Le problème souligné sur les mesures RADAR estimant une large zone de pluie en amont du relief est retrouvé ici. Couvrant la grande majorité du domaine, le champ de pluie du RADAR n'offre donc aucun point de comparaison avec les autres champs estimés (surface équivalente plus grande de 2 642 km<sup>2</sup> par rapport au KED, 2 630 km<sup>2</sup> par rapport au MERG et de 2 542 km<sup>2</sup> par rapport au OK\_RG) ;
- la quantité totale de pluie du champ du KED et du MERG, très similaires (différence de seulement 1 mégatonne), s'apparentent davantage avec celle issue de l'interpolation par OK\_RG (différences de 8 et 7 mégatonnes respectivement *contre près de 5 et 4 mégatonnes pour le cas A*). Le champ mesuré par le RADAR présente certes un cumul maximum plus faible, la plus large surface de la structure de pluie implique cependant une différence de plus de 7 % avec le champ du KED et du MERG (*contre 1% pour le cas A*).

*Dans le cas A, nous avons constaté que la technique du MERG accordait plus d'importance aux mesures RADAR que le KED. Dans le cas présent, nous ne pouvons pas l'avancer. Entre les deux techniques (KED et MERG), aucune différence n'est réellement observée ni sur la valeur du cumul maximum (différence de seulement 16 mm) ni sur la taille du champ de pluie (différence d'à peine 0.1%) ni sur la quantité totale de pluie tombée au sol (différence de 1 mégatonne d'eau sur le domaine).*

En considérant les observations pluviométriques comme référence, le Tableau III-5 présente trois scores statistiques obtenus par la comparaison des observations au sol avec d'une part la valeur du pixel du champ mesuré par le RADAR le plus proche des pluviomètres et d'autre part les estimations LOO par OK\_RG, KED et MERG aux points des pluviomètres.

Tableau III-5 : CAS B = Comparaison entre les observations du réseau principal de pluviomètres et les estimations des différents champs spatiaux. (T = 48 h).  
*Les indications (-) et (+) indiquent respectivement le moins bon et le meilleur score.*

	RADAR	OK_RG	KED	MERG
Erreur moyenne- ME	11.0 mm (-)	1.3 mm	1.2 mm (+)	1.2 mm (+)
RMSE	51.2 mm (-)	45.8 mm	46.3 mm	44.5 mm (+)
Coefficient de corrélation	0.929 (-)	0.932	0.932	0.938 (+)

Ici, toutes les techniques surestiment le champ de pluie par rapport aux pluviomètres (erreur moyenne positive). Les très fortes valeurs de la ME et de la RMSE sur les mesures RADAR sont expliquées par l'imposante présence de pluie en aval du relief. Hormis ce souci de quantification, le champ spatial mesuré concorde toutefois avec celui des mesures au sol (très fort coefficient de corrélation). Exception faite de la ME, nous observons le même comportement pour les estimations par OK\_RG (*i.e.* grande RMSE et forte corrélation). Comme pour le cas A, le champ reconstruit par KED et MERG montrent une légère amélioration dans les scores. Les ME associées sont identiques et les coefficients de corrélation sont très proches ( $+10^{-2}$  pour le MERG). Avec une plus forte valeur de la RMSE (+1.8 mm), les erreurs faites par la technique du KED sont cependant plus importantes en moyenne que celles du MERG. Notons par ailleurs que la RMSE obtenue par OK\_RG est même plus faible (-0.5 mm par rapport au KED).

La Figure III-25a présente le champ de la variance de krigeage (VK) associée à la technique du krigeage ordinaire ( $Var_{OK}$  – pour le OK\_RG et le MERG). Comme pour le cas A, nous observons que  $Var_{OK}$  augmente à mesure que l'on s'éloigne des pluviomètres (en particulier en limite de domaine). Par rapport au cas A, nous notons que la VK affiche de plus fortes valeurs. Pourtant en moyennant  $Var_{OK}$  sur tout le champ, nous obtenons une valeur près de 11 fois inférieure par rapport à la variance de l'échantillon (1 465 vs 15 936  $mm^2$ ) (*contre 12 fois inférieure pour le cas A*).

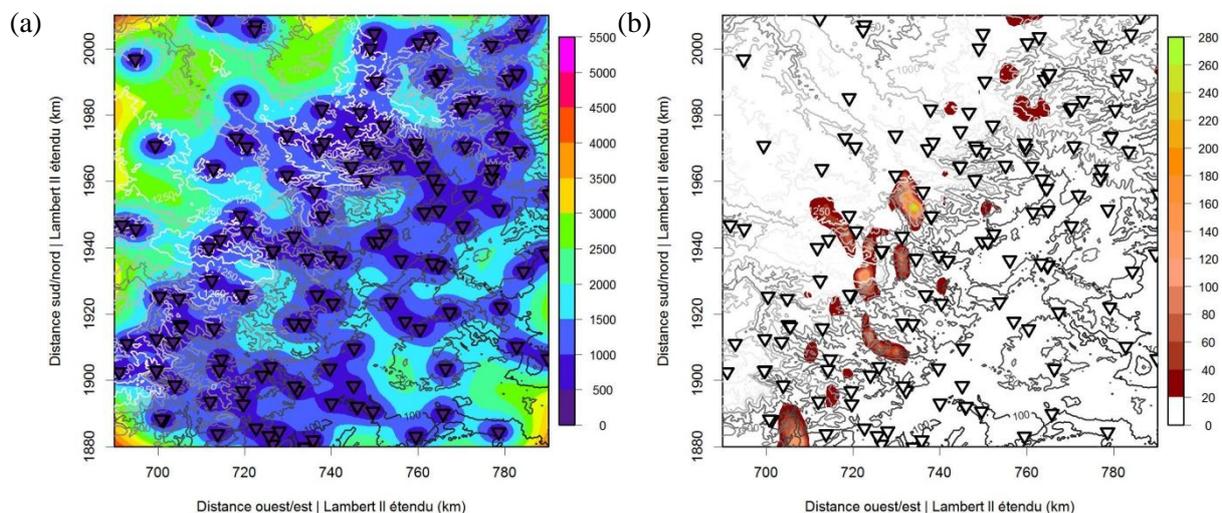


Figure III-25 : CAS B = (a) Champ de la variance de krigeage (en  $mm^2$ ) associée au krigeage ordinaire et (b) différence (en  $mm^2$ ) entre la variance de krigeage associée au KED et celle associée au krigeage ordinaire. (T = 48 h).  
 La variance de krigeage à la position des pluviomètres est reportée dans les triangles pointant vers le bas.

La VK associée à la technique du KED ( $Var_{KED}$ ) est très proche, nous la présentons sous forme de différences avec  $Var_{OK}$ . Les discordances sont principalement observées dans des zones dépourvues de pluviomètres.

En résumé, les techniques combinant les observations estiment avec le moins d'erreurs le champ de précipitation. L'intégration des mesures RADAR permet de redessiner le caractère hétérogène des précipitations observé par le RADAR, détails qui échappent au champ obtenu par OK\_RG. À l'échelle de l'événement, les deux techniques étudiées (KED et MERG) présentent des comportements très semblables tant visuellement que quantitativement.

*Dans le cas A, nous avons noté que le MERG accordait davantage d'importance aux mesures RADAR.* Dans le cas présenté, les estimations par le RADAR dessinent une large zone de précipitation en amont du relief qui n'est pourtant pas observée par les pluviomètres. Pourtant ce comportement n'est pas identifié sur les estimations par MERG. Les estimations par KED montrent toutefois une faiblesse. En effet, les résultats de la validation par LOO affichent ici des écarts en moyenne plus larges que ceux obtenus par l'interpolation OK\_RG.

### III-4.5 Reconstruction d'un champ de précipitation à différentes échelles temporelles

La Figure III-26 présente trois scores statistiques obtenus par comparaison des observations pluviométriques avec les mesures RADAR (courbes bleues) et les estimations LOO par OK\_RG (courbes jaunes), KED (courbes vertes) et MERG (courbes roses) pour différentes durées de cumul T.

Quelle que soit la technique utilisée, l'erreur moyenne (ME – Figure III-26a) augmente avec T. Les mesures RADAR présentent une très forte surestimation (*i.e.* valeurs positives) qui s'accroît linéairement avec la période de cumul. Les ME de l'ensemble des techniques d'interpolation (OK\_RG, KED et MERG) présentent de faibles valeurs (toutes inférieures à 0.8 mm) et dont les évolutions en fonction de T sont très similaires. À T = 1 h, les mesures RADAR affichent la plus forte valeur de ME (= 3, 3.5 et 3.8 fois supérieure à celle du OK\_RG, MERG et KED respectivement). Les estimations par KED donnent les plus faibles valeurs de ME.

*Dans le cas A, nous avons constaté que i) les mesures RADAR présentaient une ME négative mais plus faible que celle des estimations par OK\_RG, ii) la combinaison des données*

par KED ou MERG abaissait de façon distinctive les valeurs de la ME (qui ne dépassaient pas 0.2 mm) et iii) les plus faibles valeurs de ME s'affichaient pour le MERG.

Comme pour la ME, tous les champs estimés présentent une augmentation avec la période T de la RMSE (Figure III-26b). Les mesures RADAR montrent certes une très grande ME, pourtant les écarts sont en moyenne du même ordre de grandeur que ceux obtenus par les techniques d'interpolation (OK\_RG, KED ou MERG). Notons que ces dernières présentent des évolutions de RMSE très similaires (différence maximale observée : +1.5 mm à T = 12 h entre OK\_RG et MERG). La combinaison des données par KED ou MERG n'a pas un grand impact sur la RMSE. Nous remarquons par ailleurs que les valeurs de la RMSE sont presque doublées par rapport aux valeurs obtenues pour le cas A.

*Dans le cas A, nous avons noté pour le cas A que i) la technique par OK\_RG était affectée en moyenne par de plus larges écarts que les mesures RADAR et ii) la combinaison des données par KED ou MERG affichait des valeurs de RMSE identique et en moyenne plus faible d'un facteur moyen de 1.2 par rapport aux observations seules.*

Enfin, pour l'ensemble des estimations, le coefficient de corrélation (Figure III-26c) tend vers de meilleures valeurs (*i.e.* plus proches de 1) à mesure que T croît. Notons que tous les scores affichent initialement de très fortes corrélations (supérieures à 0.87). Au-delà de T = 6 h, l'amélioration du coefficient de corrélation est moins prononcée que pour les plus petits pas de temps. Nous observons que la corrélation des mesures RADAR est ici inférieure à celle observée avec les estimations par OK\_RG (*i.e.* la courbe bleue sous la courbe jaune). Sur la ME ou la RMSE, les effets (quantitatifs) liés à la combinaison des données étaient minces en comparaison des résultats obtenus par OK\_RG. Or, le coefficient de corrélation prouve la supériorité des estimations par KED ou MERG. La technique par MERG semble modestement devancer celle du KED (différences entre 0.8 et  $3.2 \cdot 10^{-3}$ ).

*Dans le cas A, nous avons remarqué que i) les mesures RADAR étaient plus cohérentes spatialement que ne le sont les estimations par OK\_RG, ii) les estimations par KED ou MERG obtenaient les meilleurs résultats et iii) les différences (entre 0.9 et  $3.3 \cdot 10^{-3}$ ) étaient trop petites pour les dissocier.*

Pour ce second cas d'étude, les mesures RADAR présentent un petit problème d'estimation par rapport aux observations pluviométriques. De plus, ces erreurs s'amplifient avec la période de temps considéré (*cf.* ME, RMSE). Ces écarts sont cependant en moyenne tout aussi importants que ceux calculés avec les différentes techniques d'interpolation

(OK\_RG, KED ou MERG). L'importante surestimation des mesures RADAR n'optimisent donc pas la minimisation des erreurs d'estimations associées (ME et RMSE pour le KED ou MERG proches de celles des estimations par OK\_RG). Ici, seul le coefficient de corrélation met en exergue l'atout de la combinaison des données par KED ou MERG. Comme pour le cas A, les différences entre les deux techniques sont trop subtiles pour les dissocier.

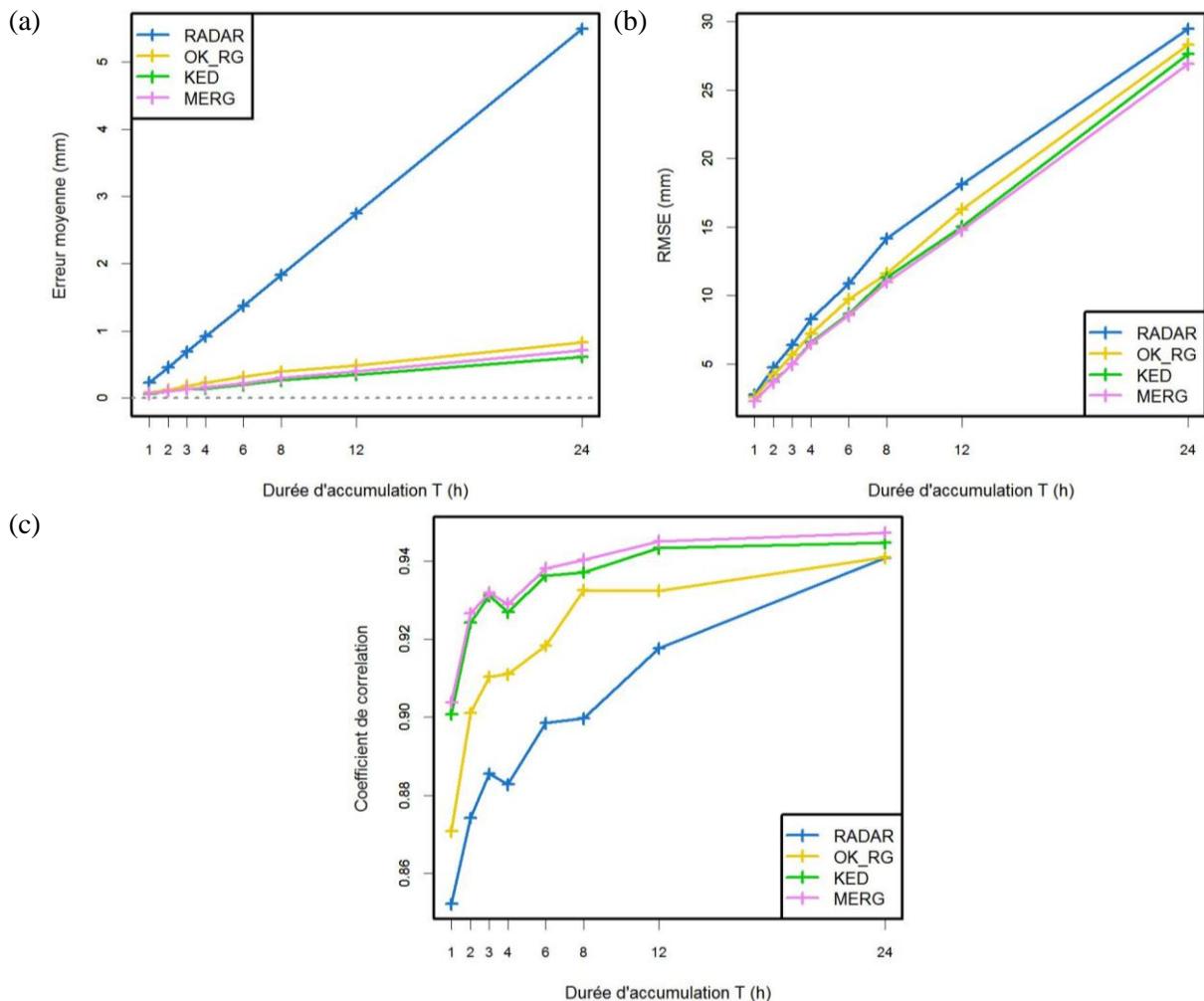


Figure III-26 : CAS B = Évolution (a) de l'erreur moyenne, (b) de la RMSE, et (c) du coefficient de corrélation en fonction de la fenêtre temporelle T (en heure).  
 Remarque : La ligne en pointillé gris affiche la valeur du score indiquant une « parfaite » correspondance entre les estimations et les observations.

Tableau III-6 : CAS B = Comparaison entre les observations du réseau secondaire de pluviomètres et les estimations des différents champs spatiaux. (T = 24 h).  
 Les indications (-) et (+) indiquent respectivement le moins bon et le meilleur score.

	RADAR	OK_RG	KED	MERG
Erreur moyenne – ME	1.0 mm (+)	-2.6 mm	-3.6 mm (-)	-3.1 mm
RMSE	29.5 mm (-)	25.9 mm	25.4 mm	24.5 mm (+)
Coefficient de corrélation	0.934 (-)	0.946	0.948	0.952 (+)

La validation des estimations avec les observations du réseau secondaire de pluviomètres (Figure III-18, marqueurs bleus) est présentée par le Tableau III-6. Les données RADAR sont les seules à présenter une surestimation (*i.e.* erreur moyenne positive), pourtant ce sont les estimations les moins biaisées en particulier par rapport avec celles par OK\_RG. Il s'avère cependant que les erreurs associées au RADAR se compensent puisque les larges écarts sont les plus prédominants (*cf.* RMSE). Les fortes valeurs des coefficients de corrélation montrent que les mesures RADAR ou les estimations par OK\_RG s'accordent spatialement avec les observations au sol. Les meilleurs résultats (à l'exception du biais) sont obtenus par la technique du KED et du MERG. Par rapport au KED, le MERG manifeste une ME (-0.5 mm) et une RMSE (-0.9 mm) plus faible ainsi qu'un coefficient de corrélation (légèrement) supérieur ( $+4 \cdot 10^{-3}$ ).

*Dans le cas A, nous avons avancé que i) les mesures RADAR avaient un biais (négatif) plus grand que les estimations par OK\_RG mais présentaient des meilleurs valeurs pour la RMSE et le coefficient de corrélation, ii) les estimations par KED ou MERG obtenaient les meilleurs résultats et iii) par rapport au KED, le MERG présentait un biais plus grand (+0.2 mm) mais une plus faible RMSE (-0.6 mm) et un meilleur coefficient de corrélation ( $+2 \cdot 10^{-3}$ ).*

Pour conclure sur le cas B, le champ mesuré par le RADAR présente un léger problème d'estimation (= large zone de précipitation en amont du relief) qui vient dégrader l'ensemble des scores statistiques employés dans notre analyse. La qualité des estimations par KED ou MERG ne semblent pourtant pas en être affectées. En termes de ME ou RMSE, nous n'observons pas d'améliorations notables par rapport aux estimations par simple OK\_RG. Néanmoins, la combinaison des données s'avère bénéfique tant visuellement que spatialement

(cf. coefficient de corrélation). Les différences par validation LOO entre les deux techniques sont insignifiantes pour pouvoir réellement les départager. La comparaison des estimations avec les observations du second réseau conduit cependant à une légère préférence pour la technique du MERG face au KED.

Les performances des techniques combinant les observations ont été prouvées au travers de deux événements intenses distincts. Les estimations par la technique du KED ou du MERG sont très proches mais celles du MERG semblent plus fiables (*i.e.* les erreurs sont plus faibles). Pour les départager, nous avons voulu pousser un peu plus loin l'expérience en étudiant la sensibilité des deux techniques en fonction de la densité de pluviomètres.

### **III-5 Test de sensibilité : rôle de la densité de pluviomètres**

L'objectif de ce test est d'évaluer les performances des estimations par OK\_RG, KED et MERG en fonction du nombre de stations disponibles pour l'interpolation. L'analyse présentée ici est basée sur des données cumulées à  $T = 12$  h. Les études faites précédemment ont été réalisées avec la totalité des stations disponibles (*i.e.* 149 stations pour le cas A et 142 pour le cas B). Ici, nous avons réduit progressivement ce nombre en supprimant à chaque fois un nombre défini de stations<sup>⊕</sup> de la base de données initiale. Les stations éliminées sont ensuite utilisées comme valeurs de référence pour une validation quantitative. Pour rendre le test statistiquement stable, nous avons répété l'opération 10 fois. La Figure III-27 et la Figure III-28 présentent les résultats de cette validation pour le cas A et le cas B respectivement.

Remarque : Les courbes associées aux mesures RADAR sont données à titre indicatif. La variation du nombre de stations n'intervenant pas dans l'estimation de la pluie par le RADAR, ce test n'a donc pas d'influence sur les valeurs du score (qui doivent rester constantes). Les quelques fluctuations observées sont expliquées par la localisation (aléatoire) des stations servant à la validation. En effet, les corrections apportées lors du post-traitement des données RADAR sont localement différentes. En d'autres termes, d'un pixel de l'image RADAR à un autre, les erreurs d'estimation associées ne sont pas identiques.

---

<sup>⊕</sup> Ces nombres correspondent à : 16.7, 20, 25, 33.3, 40, 50, 60, 66.7, 75, 80 et 83.3% du nombre total de stations.

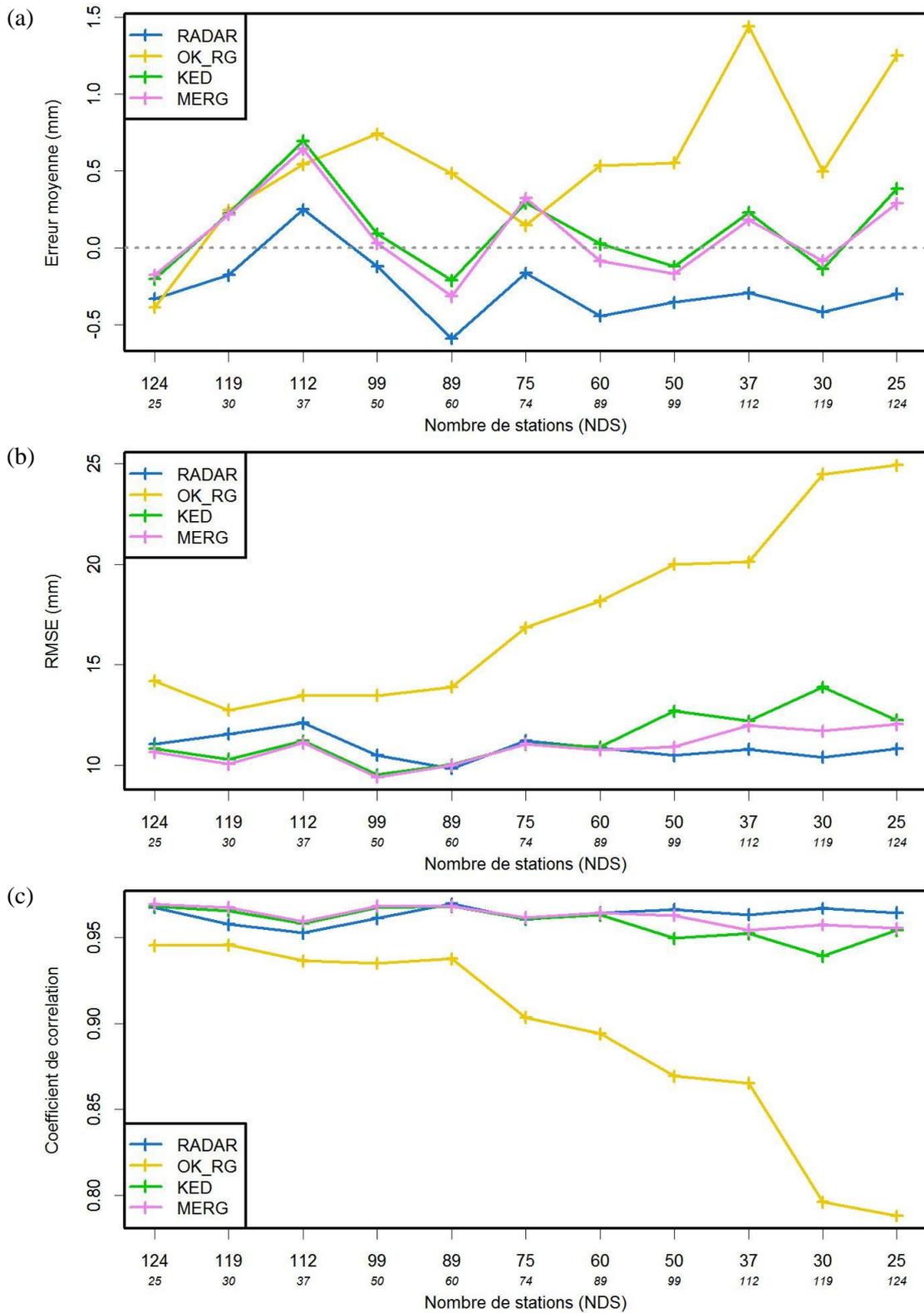


Figure III-27 : Cas A = Évolution de (a) l'erreur moyenne, (b) la RMSE et (c) le coefficient de corrélation en fonction du nombre de stations utilisées pour l'interpolation. (T = 12 h).

Remarque : la ligne inférieure sur l'axe des abscisses indique le nombre de stations alors utilisées pour la validation. La somme des stations utilisées pour l'interpolation et la validation correspond aux 149 stations horaires disponibles. À titre d'information : la surface du domaine d'étude est de 14 300 km<sup>2</sup>, soit une densité initiale de 1 pluviomètre pour 96 km<sup>2</sup>.

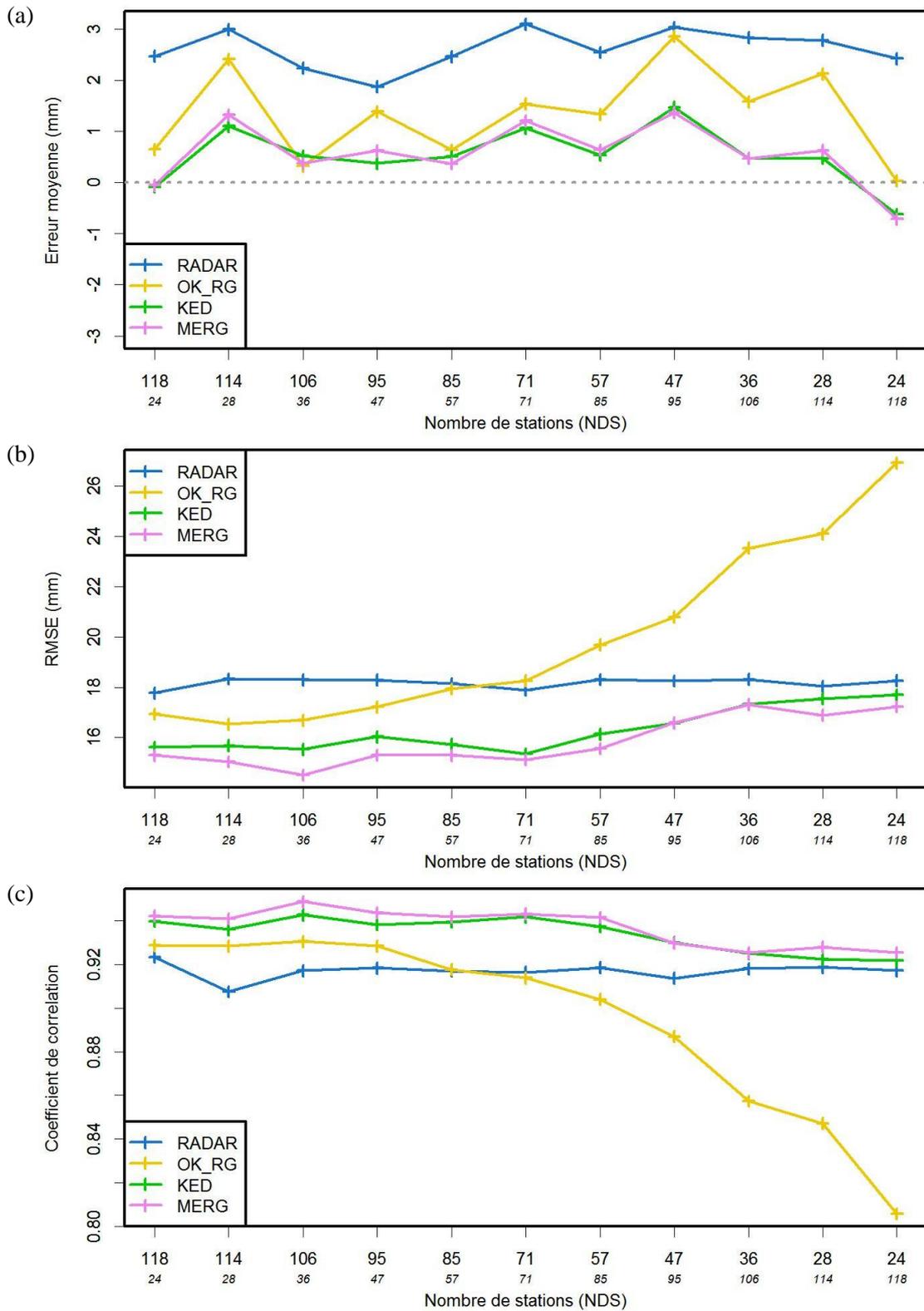


Figure III-28 : Cas B = Évolution de (a) l'erreur moyenne, (b) la RMSE et (c) le coefficient de corrélation en fonction du nombre de stations utilisées pour l'interpolation. (T = 12 h).

La somme des stations utilisées pour l'interpolation et la validation correspond aux 142 stations horaires disponibles. À titre d'information : la surface du domaine d'étude est de 13 000 km<sup>2</sup>, soit une densité initiale de 1 pluviomètre pour 91 km<sup>2</sup>.

Le nombre de stations (ou NDS) ne semble pas avoir d'influence sur l'erreur moyenne (ME, Figure III-27a). En cohérence avec ce que nous avons conclu précédemment, nous retrouvons bien un biais négatif uniquement pour les mesures RADAR<sup>⊕</sup>. Les estimations par OK\_RG ne semblent pas montrer de dépendance avec le NDS. Par rapport aux mesures RADAR, nous notons que les ME associées présentent de plus fortes valeurs et des variations légèrement plus prononcées. Ce comportement est principalement mis en cause par quatre stations pluviométriques, toutes situées au cœur de l'événement pluvieux et qui enregistrent les plus forts cumuls<sup>⊕⊕</sup>. Enfin, nous constatons que les ME calculées à partir des estimations par KED ou MERG présentent des évolutions similaires. De plus, la combinaison des données permet de minimiser le biais (valeurs oscillant entre -0.3 et 0.7 mm).

Concernant la RMSE (Figure III-27b), les estimations par OK\_RG, présentent initialement une valeur supérieure à celle des mesures RADAR (+3.1 mm pour un NDS de 124). Au cours du test, celles-ci affichent des écarts quadratiques inversement proportionnelle aux NDS (+14.1 mm pour un NDS de 25). Cette dégradation est également observée pour les estimations par KED ou MERG mais de façon beaucoup moins prononcée. Tant que le NDS reste supérieure à 60 stations, les estimations par KED ou MERG présentent une RMSE légèrement plus faible que celle des mesures RADAR. En dessous de cette limite, l'utilisation du RADAR est à préférer à toute technique d'interpolation.

Enfin, le coefficient de corrélation (Figure III-27c) montre une forte corrélation spatiale entre les différentes estimations et les observations pluviométriques. Toutefois, plus le NDS diminue, plus les estimations par OK\_RG se dégradent. Il en va de même pour les estimations par KED ou MERG mais toujours avec moins d'ampleur et uniquement pour un NDS inférieur à 60. Avec une densité moindre, le coefficient de corrélation des mesures RADAR devanent ceux des estimations par KED ou MERG.

Pour ce premier cas d'étude (cas A), la RMSE et le coefficient de corrélation se dégradent à mesure que le NDS diminue. La dépendance la plus marquée est celle observée avec les estimations par OK\_RG. La technique du KED ou du MERG présentent de meilleurs résultats pour une densité supérieure à 1 pluviomètre pour 238 km<sup>2</sup> (⇔ NDS de 60). Avec un

<sup>⊕</sup> En approfondissant notre analyse, nous avons trouvé que les deux plus importantes variations observées (NDS = 112 et 89) sont exclusivement associées à une même et unique station (de coordonnées  $x = 742$  km,  $y = 1936$  km) pour laquelle est enregistrée une différence avec les observations pluviométriques de presque -150 mm. Sur les 10 séries aléatoires réalisées, cette valeur est malheureusement prise en compte plusieurs fois (respectivement 7 et 6 fois), augmentant dès lors la ME en conséquence.

<sup>⊕⊕</sup> En s'affranchissant de ces stations, le champ résultant de l'interpolation par OK\_RG est fortement lissé affichant même une sous-estimation importante (non montré ici).

NDS plus faible, il est préférable d'opter pour l'utilisation du RADAR seul afin de minimiser les erreurs. Ces conclusions s'accordent avec celles de Velasco-Forero *et al.* (2008) discutées dans le sous-paragraphe II-2.2.3

Une étude similaire a été réalisée pour notre second cas d'étude (*i.e.* cas B, remarqué par des mesures RADAR erronées sur une large partie du domaine). Nous résumons ici les principaux résultats obtenus de l'analyse :

- pour la ME (Figure III-28a) : les plus fortes valeurs sont données par les mesures RADAR. Les estimations par OK\_RG affichent de plus faibles valeurs<sup>⊕</sup>. Les meilleurs résultats sont obtenus avec les estimations par KED et MERG. Les valeurs des ME associées oscillent toutefois entre -0.7 et 1.5 mm ;
- pour la RMSE (Figure III-28b) et le coefficient de corrélation (Figure III-28c) : les estimations par OK\_RG se dégradent à mesure que le NDS diminue. Les scores associés égalent les scores des mesures RADAR pour un NDS de 85. La technique du KED et du MERG présentent, pour toutes les valeurs de NDS, les meilleurs résultats. Une légère dégradation des scores est observée pour des NDS inférieurs à 57 ( $\Leftrightarrow$  densité de 1 pluviomètre pour 228 km<sup>2</sup>). Néanmoins la combinaison des données reste bénéfique par rapport aux estimations RADAR (qui *contrairement au cas A* n'égalent pas les scores alors obtenus).

Ce test montre les performances avérées des deux techniques combinant les observations. Il nous a surtout permis de mettre en avant une dépendance de la densité de pluviomètres sur les estimations par OK\_RG et de façon moins prononcée sur les estimations par KED et MERG. Ce comportement semble apparaître uniquement en dessous d'une certaine densité de pluviomètre (1 pour 238 km<sup>2</sup> pour le cas A, 1 pour 228 km<sup>2</sup> pour le cas B). En dessous de cette limite et si la qualité des estimations le permet (*cf.* cas A), l'utilisation des mesures RADAR seules est recommandée. Dans le cas contraire (*cf.* cas B), il est préférable d'opter pour la combinaison des données quelle que soit la densité du réseau de pluviomètres afin de minimiser les erreurs d'estimations.

---

<sup>⊕</sup> Ces estimations présentent néanmoins d'importantes variations que nous pouvons associer ici encore à la suppression répétée des stations au cœur de l'événement.

## III-6 Conclusions

Nous avons montré au travers de deux cas d'étude distincts, l'avantage de prendre en compte une seconde variable dans les procédures d'interpolation. En effet, les estimations par KED ou MERG donnent des résultats très satisfaisants par rapport aux estimations issues de l'interpolation des observations pluviométriques seules (OK\_RG) ou aux mesures RADAR.

Le champ reconstruit par la combinaison des données montre d'une part que les structures hétérogènes du champ de pluie observées par le RADAR sont conservées et d'autre part que les estimations finales sont ajustées de façon à s'apparenter à celles des observations pluviométriques. Les champs de variance de krigeage associés certifient la bonne représentativité des estimations (*i.e.* valeurs plus faible que celle de la variance de l'échantillon). Cependant la grande similarité entre les deux techniques étudiées ne nous permet pas de les dissocier.

Les analyses quantitatives ont mis en exergue la qualité des estimations par la technique du KED et du MERG puisque les erreurs associées sont minimisées. De plus, en considérant de grandes fenêtres temporelles, ces estimations sont davantage optimisées. La validation avec un second réseau de pluviomètre (différent de celui utilisé pour l'interpolation) ont timidement montré la supériorité du MERG face au KED.

Dans le premier cas d'étude, nous avons constaté que l'intégration des mesures RADAR semblait davantage marquée dans la technique du MERG par rapport au KED. Ce comportement est observé avec des mesures RADAR dont les erreurs d'estimation sont inférieures à celles de l'interpolation des pluviomètres (OK\_RG). Nous avons testé cette dépendance au travers d'un second cas d'étude qui présente un problème d'estimation du champ de pluie par le RADAR. Il s'est avéré que la combinaison des données n'est que peu sensible aux les estimations quantitatives des précipitations du RADAR puisque le problème en question est corrigé dans les champs finaux.

Enfin, nous avons montré que les estimations par OK\_RG sont dépendantes de la densité de pluviomètres disponibles pour l'interpolation. Un comportement similaire (mais nettement moins prononcé) est observé sur les estimations par KED et MERG. Cependant celui-ci semble se manifester uniquement en dessous d'une certaine densité de pluviomètre.

Pour finir, la technique de reconstruction spatiale par MERG présente une particularité qui n'a pas été abordée dans ce chapitre. En effet, la technique du KED est par définition (*cf.*

sous-paragraphe II-2.2.1) strictement limitée au domaine observé par la variable auxiliaire (ici assimilée au champ RADAR). Sa mise en place nécessite d'une part de connaître en des points connus les valeurs de la variable principale (ici les observations pluviométriques) et secondaire (ici les pixels RADAR les plus proches) et d'autre part de connaître les valeurs de la variable secondaire en tout point de la grille d'interpolation. Les observations en dehors du domaine ne sont donc pas prises en compte. Ce constat est notamment observable sur les champs de la variance de krigeage dont les valeurs sont maximales en bordure de domaine. Or, la technique du krigeage ordinaire (*i.e.* OK\_RG et MERG) permet de considérer ces stations externes.

Cette intégration peut s'avérer être un atout incontestable, notamment pour des domaines d'étude possédant un réseau de pluviomètres plus restreint spatialement. L'étude de tels cas fait l'objet du prochain chapitre où nous proposons une analyse des différentes techniques (KED et MERG) avec les données du bassin clermontois.

---

## Les Estimateurs statistiques utilisés dans le Chapitre IV

Notons  $V_i$  la valeur observée et  $\tilde{V}_i$  la valeur estimée à la station  $i$ . Considérons  $N$  stations sur l'ensemble du domaine étudié, alors :

→ l'erreur moyenne (ME) correspond à la moyenne des écarts entre les valeurs estimées et observées. Également appelée biais, elle est définie par :

$$ME = \frac{1}{N} \sum_{i=1}^N (\tilde{V}_i - V_i)$$

Pour le cas d'une estimation « parfaite », la ME arbore une valeur de 0. Dans le cas contraire, il est possible d'obtenir une ME nulle si les écarts se compensent entre eux.

→ l'erreur quadratique moyenne (MSE) mesure la moyenne des carrés des écarts entre les estimations et les observations. Elle est plus souvent exprimée sous la forme d'une racine carré (RMSE), telle que :

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (\tilde{V}_i - V_i)^2}$$

Par rapport à la ME, la RMSE accorde plus de poids aux larges écarts. Elle est particulièrement appréciée pour détecter la présence de larges écarts, surtout si les écarts se compensent. Elle n'indique cependant pas la direction des écarts.

→ l'écart-type mesure la dispersion des valeurs de la variable considérée. Pour les observations, il est exprimé par :

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (V_i - \bar{V})^2}$$

avec  $\bar{V}$  la moyenne des valeurs observées.  $\sigma^2$  définit la variance de l'échantillon considéré.

→ le coefficient de corrélation (spatial) renseigne sur la relation linéaire (ou phase) entre le champ des valeurs estimées et celui des valeurs observées. Il est donnée par :

$$r = \frac{\sum_{i=1}^N (\tilde{V}_i - \bar{\tilde{V}}) \cdot (V_i - \bar{V})}{\sqrt{\sum_{i=1}^N (\tilde{V}_i - \bar{\tilde{V}})^2} \cdot \sqrt{\sum_{i=1}^N (V_i - \bar{V})^2}}$$

avec  $\bar{V}$  et  $\bar{\tilde{V}}$  les moyennes des valeurs observées et estimées. Un coefficient de corrélation nul indique que les deux champs n'ont aucun lien entre eux. Plus sa valeur s'éloigne de zéro (*i.e.* en s'approchant de 1 ou -1), meilleure est la corrélation (positive = en phase ou négative = en opposition de phase).

L'échantillon des  $N$  observations (ou estimations) ordonnées par valeurs croissantes peut être découpé en  $q$  intervalles réguliers. On nomme quantiles, les valeurs arrondies marquant les limites entre deux intervalles consécutifs. Le terme centiles est donné pour 100 quantiles. Ainsi, le dernier centile sépare les 99 % valeurs inférieures des observations (ou estimations).

---

# Chapitre IV

## Reconstruction d'un champ de pluie : Application autour de l'agglomération clermontoise

Dans le chapitre précédent, nous avons présenté une première application de deux techniques de spatialisation combinant les observations pluviométriques et mesures RADAR. Les performances de chacune d'elles ont été testées au travers de deux événements extrêmes survenus dans une large région bien instrumentée. Nous nous proposons dans ce chapitre de réaliser une étude similaire mais pour une plus petite région où les pluies sont moins intenses et surtout où le réseau d'observation est fortement réduit. Dans les deux premières parties qui composent ce second chapitre d'application, nous décrirons la situation géographique et hydrologique de la région et présenterons brièvement le réseau d'observation. Nous nous consacrerons ensuite à la reconstruction des champs de pluie au travers de deux événements. Le cœur de ce chapitre se divise en trois parties. La première se focalise sur la détermination du variogramme paramétrique nécessaire pour l'application des différentes techniques géostatistiques considérées, le krigeage avec dérive externe (KED) et la fusion conditionnée (MERG). Les deux parties suivantes présentent les deux cas d'étude et les résultats obtenus.

### IV-1 Clermont-Ferrand et son agglomération

En bordure nord du Massif central (Figure IV-1), le paysage atypique du département du Puy-de-Dôme peut être découpé en trois régions. Au nord-ouest, le plateau des Combrailles forme une vaste zone de collines et de gorges. Elle est délimitée par la chaîne des Puys créée par l'alignement de plus de 80 volcans selon un axe nord-sud. Au centre, le puy de Dôme domine ses voisins du haut de ses 1 464 m. Ce massif volcanique se poursuit au sud par les monts Dore et le plateau du Cézallier. Au cœur du département, un imposant lac formé par l'affaissement du socle granitique a laissé place aujourd'hui à une large plaine très fertile, la Limagne. En bordure ouest, à seulement 12 km à vol d'oiseau du puy de Dôme, se dresse la

ville de Clermont-Ferrand et son agglomération de communes (*cf.* étoile). De l'autre côté, l'Allier et son affluent la Dore, creusent le terrain façonnant méandres et îlots. Enfin, à l'est du département, parallèle à la chaîne des Puys, une seconde ligne de crête s'élève, formée par les monts du Livradois et du Forez.

L'eau dans cette région apparaît comme un élément important de part la présence de nombreux cours d'eau, lacs, étangs, marais et tourbières qui agrémentent le paysage. Ces zones humides restent cependant synonymes d'inondations en cas de forts événements pluvieux. Le socle granitique est la formation géologique prédominante. Les sols n'étant pas ou peu perméables, les eaux pluviales sont alors rapidement évacuées par ruissellement.

Si globalement l'eau tombe abondamment en Auvergne (site Internet : DREAL - Auvergne), elle est inégalement répartie du fait du relief conjugué à différentes influences climatiques. Sur les Combrailles et les versants Ouest de la chaîne de volcans, l'influence océanique génère une forte pluviométrie accentuée par les effets du relief. Au centre et à l'est, l'influence continentale se caractérise par de plus faibles précipitations. À l'abri des reliefs, Clermont-Ferrand est une ville des plus sèches de France (Choisnel et Payen 1988, 32–41). Néanmoins, de fortes pluies d'orage sont fréquentes en fin de printemps et en été. Le dernier exemple marquant en date du samedi 26 mai 2012, jour d'ouverture au public du Panoramique des Dômes, un orage éclata en fin d'après-midi causant des coulées de boue qui ont tout emporté sur leur passage jusqu'au ballast du nouveau train (Dubois 2012).

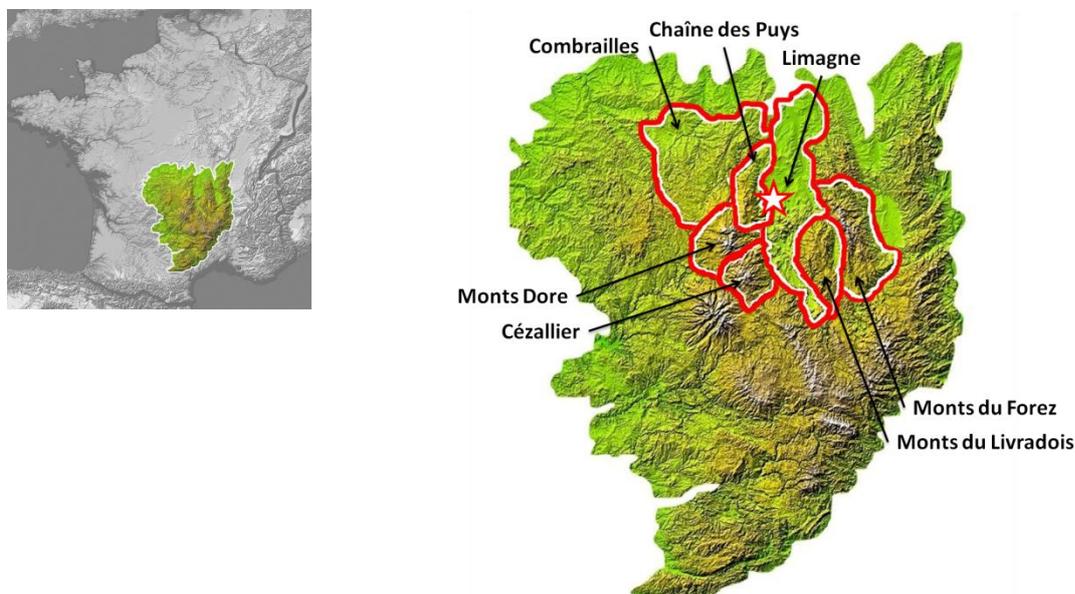


Figure IV-1 : Carte du Massif central avec localisation des principales régions montagneuses autour de la ville de Clermont-Ferrand (étoile).

Source : Wikipedia (modifiée)

## IV-2 Le réseau d'observation

Hors du réseau national opéré par Météo-France, l'étude des précipitations au LaMP (Laboratoire de Météorologie Physique, Aubière) est assurée par un radar météorologique et différents réseaux de mesures au sol (*cf.* Figure IV-2). Dans ce qui suit, nous les décrirons brièvement.

### IV-2.1 Le RADAR

Le territoire clermontois est couvert par un RADAR météorologique situé sur le toit de l'Observatoire de Physique du Globe de Clermont-Ferrand (Aubière). Il fonctionne en bande X et à très hautes résolutions spatiale ( $\approx 58$  mètres) et temporelle (= 30 secondes). Son protocole de sondage balaye l'atmosphère selon un seul PPI fixé à  $5^\circ$ . Non Doppler et non polarisant, il mesure uniquement le facteur de réflectivité moyen sur un rayon d'environ 20 km. Le post-traitement des données est assuré par le LaMP. La chaîne d'algorithme comprend (Mioche et Van Baelen 2013) :

- l'agrégation des données à 30 secondes pour simuler une résolution temporelle de 5 min ;
- l'élimination des échos de sol à partir d'une image établie par temps sec et sans pluie ;
- la correction des effets d'atténuation (ou EA) le long du faisceau du RADAR selon l'approche de Hitschfeld et Bordan (1954) avec une limite de 10 dBZ pour assurer la stabilité de l'algorithme (Delrieu *et al.* 1999) ;
- l'application de la relation  $Z-R$  (équation [I-5]) de Marshall et Palmer (1948).

En sortie, une image d'intensités pluvieuses moyennes est élaborée sur une grille en coordonnées polaires (projection conique conforme de Lambert II étendu – *cf.* par exemple Snyder (1987, 104–107) ou le site Internet de l'IGN pour les algorithmes de changement de systèmes et transformations de coordonnées).

Depuis son installation en 2006, le RADAR a subi plusieurs modifications techniques (ex : changement de l'antenne parabolique, de la puissance de l'émetteur, etc.), des améliorations du système d'acquisition des données (ex : diminution de la résolution angulaire de  $2$  à  $1^\circ$ , etc.) ainsi que des mises à jours du logiciel de post-traitement. Tous ces changements ont affectés d'une certaine manière la qualité des données (tout en cherchant à l'améliorer). Sa localisation et son protocole de sondage sont restés les mêmes.

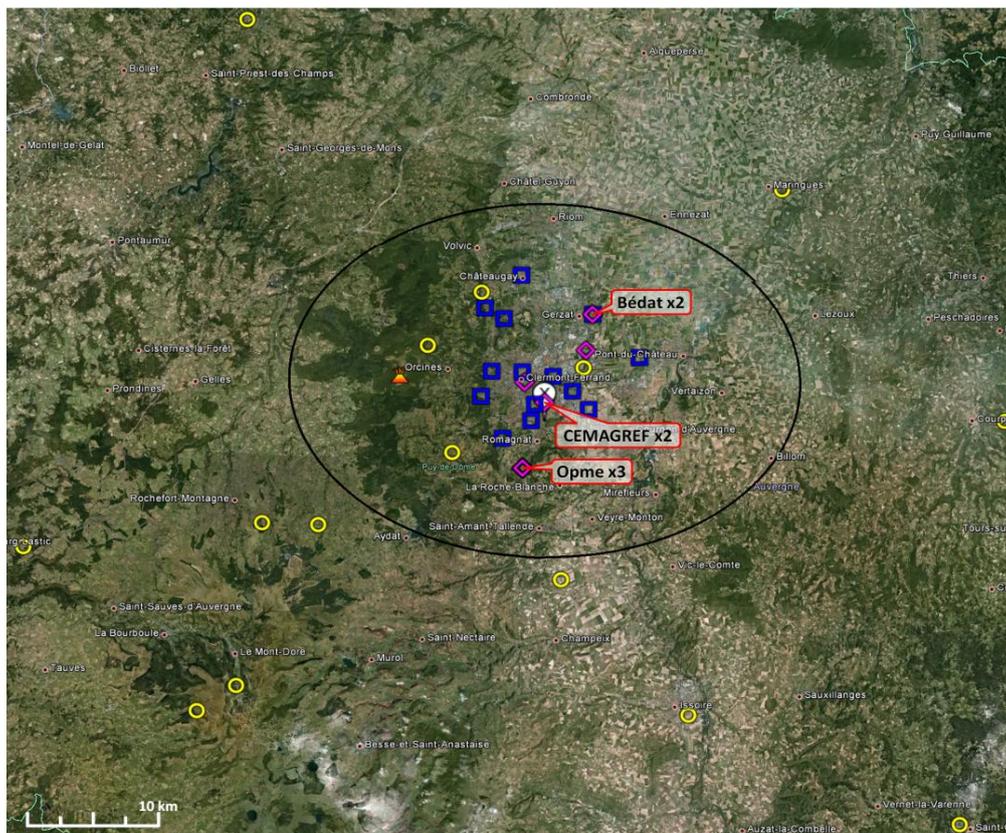


Figure IV-2 : Fenêtre d'observation et réseaux de mesures.

*L'ellipse noire représente la zone couverte par le RADAR du LaMP (localisé au centre = marqueur blanc). Sont également indiquées les positions des instruments du réseau principal (résolution à 5 min) gérés par le LaMP (marqueurs roses ; 4 Parsivels® et 5 pluviomètres) et Clermont-Communauté (marqueurs bleus ; 14 pluviomètres). Météo-France fournit les données du réseau de pluviomètres secondaires (résolution à 1 h) localisés par les marqueurs jaunes (une station située plus au sud n'est pas reportée). Le Puy de Dôme est représenté par le marqueur orange. Source : Google Earth.*

## IV-2.2 Les réseaux de mesures au sol

Les données pluviométriques proviennent de deux réseaux distincts. Le réseau principal est composé de 23 instruments dont la résolution temporelle est de 5 minutes. Ces observations sont gérées par deux entités : le LaMP et le Service Assainissement de Clermont-Communauté. La mise à disposition de ces données varie entre 1 et 3 mois. Pour trois positions (CEMAGREF, Opme et Bédât), 2 ou 3 instruments sont localisés à moins de 30 mètres. Cette distance étant inférieure à la résolution du RADAR, nous avons décidé de moyenniser ces données pour n'obtenir qu'une seule mesure par localisation. Ce réseau est complété par 16 pluviomètres automatiques horaires (sur les 20 qui couvrent le département) administrés par Météo-France (qui regroupe les instruments de Météo-France, de la Direction

Générale de la Prévention des Risques, de la Direction Départementale de l'Équipement et de l'Institut National de la Recherche Agronomique).

Pour s'assurer de l'exactitude des données de nos cas d'étude, nous avons appliqué un contrôle de qualité (décrit dans la partie II-1) uniquement sur les observations du réseau de pluviomètres principal.

## IV-3 Spatialisation des données : définition du variogramme paramétrique

L'application des techniques de krigeage requiert au préalable l'identification du variogramme paramétrique  $\tilde{\gamma}$  (cf. paragraphe II-2.1). Il est parfois compliqué de choisir parmi les plusieurs approches existantes pour définir le meilleur variogramme paramétrique. Dans le cadre des événements cévenols étudiés dans le chapitre précédent, nous avons admis les résultats de l'approche décrite par Lebel et Bastin (1985) (cf. section III-2). Pour les événements touchant le bassin clermontois, aucune étude de reconstruction spatiale n'a été réalisée jusqu'à présent. Nous devons donc dans un premier temps déterminer le variogramme paramétrique qui soit en accord avec les données clermontoises. Nous avons choisi d'effectuer ce travail avec la même approche que pour les cas cévenols.

Dans un premier paragraphe, nous reviendrons sur le contexte et la procédure menant à la détermination du variogramme paramétrique. Nous exposerons ensuite les résultats obtenus sur les données du bassin clermontois avant de révéler les deux variogrammes que nous avons retenus.

### IV-3.1 Théorie générale

La détermination des paramètres du variogramme paramétrique  $\tilde{\gamma}$  est fonction de l'espace et du temps. En suivant l'approche dite « climatologique » proposée par Bastin *et al.* (1984),  $\tilde{\gamma}$  peut alors être décomposé tel que :

$$\tilde{\gamma}(\mathbf{h}, t) = \alpha(t) \cdot \Gamma(\mathbf{h}, \beta)$$

Ainsi, tous les termes indépendants du temps  $t$  sont inclus dans le facteur  $\Gamma(\mathbf{h}, \beta)$  (aussi appelé variogramme « climatologique ») ; les termes non-stationnaires (mais invariants dans l'espace) sont eux concentrés dans le facteur  $\alpha(t)$ . L'avantage de cette approche est que les poids  $\lambda_i$  liés à l'estimation par krigeage (équations [II-7] et [II-9]) sont calculés une seule et

unique fois (en relation avec le variogramme  $\Gamma$ ) tandis que la variance de krigeage s'adapte en fonction de l'événement (en lien avec le facteur  $\alpha$ ).

Sur ce principe, Lebel et Bastin (1985) ont développé deux critères : la moyenne de l'erreur quadratique d'interpolation (ou MSIE pour l'acronyme anglais de « Mean Squared Interpolation Error ») et la moyenne de l'erreur quadratique standardisée (ou SMSE pour l'acronyme anglais de « Standardized Mean Square Error »). Le premier critère permet de déterminer les paramètres du variogramme climatologique  $\Gamma$  (*i.e.* modèle mathématique  $M$ , portée  $\beta$  et effet de pépite  $c_0$ ) minimisant les erreurs liées à l'interpolation. Les hypothèses associées au second critère facilitent alors le calcul du facteur  $\alpha$ . Les auteurs ont montré les performances de ces critères notamment dans le cadre de la spatialisation des données de pluie. Leur étude s'est focalisée sur un bassin versant cévenol couvert par  $N = 35$  pluviomètres (localisés par leurs coordonnées  $\mathbf{s}$ ) et sur les  $L = 103$  plus fortes mesures horaires (notées  $V_l$ ) enregistrées entre 1975 et 1980.

Pour la détermination du variogramme climatologique, les auteurs ont considéré deux modèles  $M$  (décrits dans la section II-2.1.1-c) : le modèle **sphérique** et **puissance**. La MSIE dépend du modèle  $M$  et la portée  $\beta$  et est exprimée par :

$$\text{MSIE}^M(\beta) = \frac{1}{L} \sum_{l=1}^L \text{SIE}_l^M(\beta) \quad [\text{IV-1}]$$

$$\text{avec } \text{SIE}_l^M(\beta) = \frac{1}{N} \sum_{i=1}^N \frac{\left( V_l(\mathbf{s}_i) - \tilde{V}_l^M(\mathbf{s}_i, \beta) \right)^2}{\sigma_l^2} \quad [\text{IV-2}]$$

où  $\tilde{V}_l^M(\mathbf{s}_i, \beta)$  est une estimation de la quantité de pluie (pour l'évènement  $l$ )  $V_l$  en considérant le modèle  $M$  et la portée  $\beta$  au pluviomètre localisé par ses coordonnées  $\mathbf{s}_i$  et où  $\sigma_l^2$  est la variance de l'échantillon pour l'évènement  $l$ .

Dans l'équation [IV-2], la standardisation des données par la variance de chaque évènement permet de faire tendre la valeur du palier du variogramme expérimental vers 1. Ainsi, le variogramme climatologique  $\Gamma(\mathbf{h}, \beta)$  est déterminé en s'affranchissant de ce paramètre (*i.e.*  $\Gamma(\mathbf{h}, \beta)$  est défini seulement par sa portée et son effet de pépite).

Pour chaque modèle  $M$ , la portée  $\tilde{\beta}$  associée est sélectionnée parmi une liste de valeurs prédéfinies de façon à ce que la MSIE (équation [IV-1]) soit minimale. Le modèle finalement retenu devra avoir la plus petite valeur de MSIE.

Initialement considéré comme nul, l'effet de pépité  $\tilde{c}_0$  est ensuite sélectionné parmi une liste de valeurs prédéfinies tel que, pour le modèle  $M$  et la portée  $\tilde{\beta}$ , la MSIE reste minimale.

Le facteur  $\alpha(t)$  doit être évalué pour chaque événement, il est donc fortement lié à la variance de l'échantillon de chaque événement  $\sigma_t^2$  (connue). Il est possible de l'exprimer par :

$$\alpha(t) = \alpha_0 \cdot \sigma_t^2$$

La SMSE permet alors d'estimer le facteur inconnu  $\alpha_0$  par :

$$\tilde{\alpha}_0 = \frac{1}{L} \frac{1}{N} \sum_{l=1}^L \sum_{i=1}^N \frac{(V_l(\mathbf{s}_i) - \tilde{V}_l^M(\mathbf{s}_i, \tilde{\beta}))^2}{\sigma_t^2} \cdot \frac{1}{s_l^M(\mathbf{s}_i, \tilde{\beta})} \quad [\text{IV-3}]$$

$$\text{avec } s_l^M(\mathbf{s}_i, \tilde{\beta}) = - \sum_{j_1=1}^N \sum_{j_2=1}^N \lambda_{j_1}(\tilde{\beta}) \cdot \lambda_{j_2}(\tilde{\beta}) \cdot \Gamma(\mathbf{h}_{j_1 j_2}, \tilde{\beta}) \quad \text{et} \quad \lambda_{j_1=j_2=i} = -1$$

où les  $\lambda$  sont les poids attribués aux  $N$  stations par la méthode du krigeage en tenant compte des paramètres du variogramme climatologique  $\Gamma(\mathbf{h}, \tilde{\beta})$  considéré.

Enfin, après avoir identifié le variogramme paramétrique  $\tilde{\gamma}(\mathbf{h}, t)$  pour des observations horaires, Lebel *et al.* (1987) ont cherché à établir une relation entre les paramètres alors déterminés (portée exclusivement) avec la durée d'accumulation des données (notée T). Nous essayerons pareillement de voir alors si de telles relations peuvent être observées sur nos données.

### IV-3.2 Application aux observations pluviométriques de Clermont-Ferrand

Afin d'identifier le variogramme paramétrique  $\tilde{\gamma}$  propre aux observations du bassin clermontois, nous avons appliqué la technique décrite précédemment sur les données brutes (*i.e.* sans QC). De plus, pour s'apparenter aux travaux de Lebel et Bastin (1985), nous avons travaillé avec les observations du réseau de pluviomètre principal cumulées sur  $T = 1$  h. Nous avons sélectionné les  $L = 100$  plus importants cumuls de pluie survenus entre 2006 et 2011.

Pour le calcul du variogramme expérimental  $\hat{\gamma}$ , nous avons considéré 16 classes de distances entre de 0 à 15 km, de largeurs inégales comprises entre 500 m à 1.5 km. La Figure IV-3a représente l'histogramme de l'ensemble des couples de données en fonction des classes de distance considérées. Il nous assure alors une distribution semblable à une loi normale. Le variogramme expérimental est donné par la Figure IV-3b.

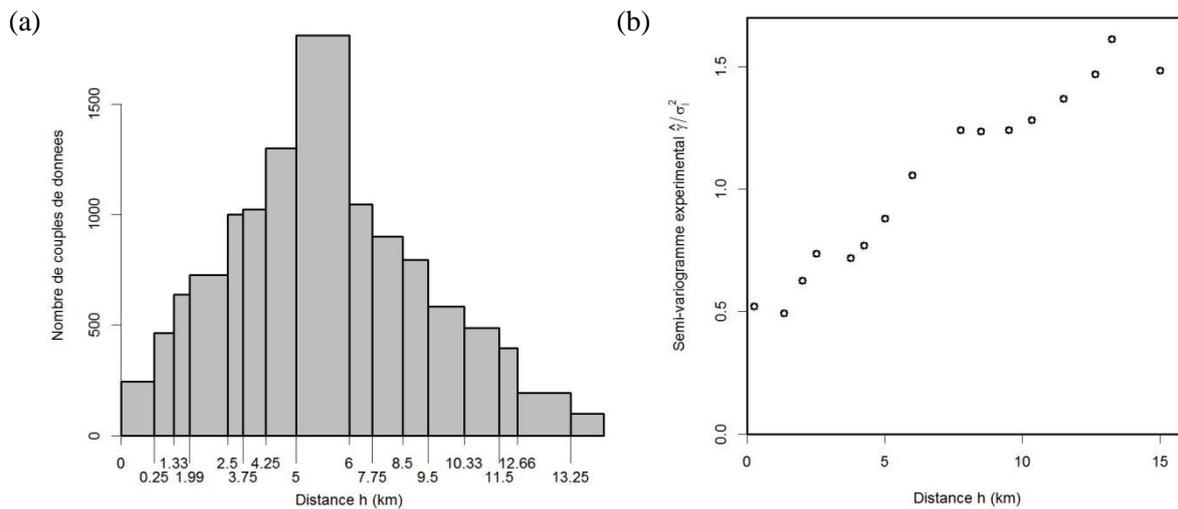


Figure IV-3 : (a) Nombre de couples de données utilisées en fonction des classes de distance considérées et (b) semi-variogramme expérimental  $\hat{\gamma}$  (normalisé par la variance de l'échantillon  $\sigma_l^2$ ) pour les  $L = 100$  plus importants événements. ( $T = 1$  h).

### IV-3.2.1 Identification du variogramme climatologique (portée & pépite)

Pour déterminer le variogramme climatologique  $\Gamma$ , nous avons considéré 6 modèles mathématiques  $M$  (définis dans la section II-2.1.1-c) : les modèles **sphérique** (sph) et **puissance** (pow) employés dans Lebel et Bastin (1985) ainsi que les modèles **exponentiel** (exp), **gaussien** (gau), **stable** (sta) et **linéaire avec palier** (lws). Parmi une liste de valeurs prédéfinies (assimilées aux classes de distances déterminées pour le variogramme expérimental aux quelles nous avons ajouté 5 distances supplémentaires entre 17 et 30 km), l'estimation de la portée  $\tilde{\beta}$  est donnée par la minimisation de la MSIE (équation [IV-1]). Les résultats obtenus sont présentés dans la Figure IV-4.

Les modèles **sphérique** et **exponentiel** présentent des comportements similaires : une forte diminution de la MSIE pour des portées croissantes de 0.25 à 6 km, au-delà une très légère ré-augmentation de la MSIE passée la portée optimale  $\tilde{\beta}$  (*i.e.* MSIE minimale). Nous noterons que la portée  $\tilde{\beta}$  est plus grande pour le modèle **exponentiel** (25 km) que pour le modèle **sphérique** (12.66 km). Avec un comportement très proche et une portée  $\tilde{\beta}$  intermédiaire (17 km), le modèle **linéaire avec palier** montre des petites fluctuations singulières entre 4.25 et 15 km. L'évolution de la MSIE avec le modèle **stable** paraît plus lente. Nous noterons que la portée  $\tilde{\beta}$  (30km) est plus grande que celles estimées par les autres modèles. Le modèle **gaussien** présente des valeurs de MSIE extrêmement fortes (atteignant presque 104 pour une portée de 15 km) et possède la plus petite portée  $\tilde{\beta}$  (2.5 km).

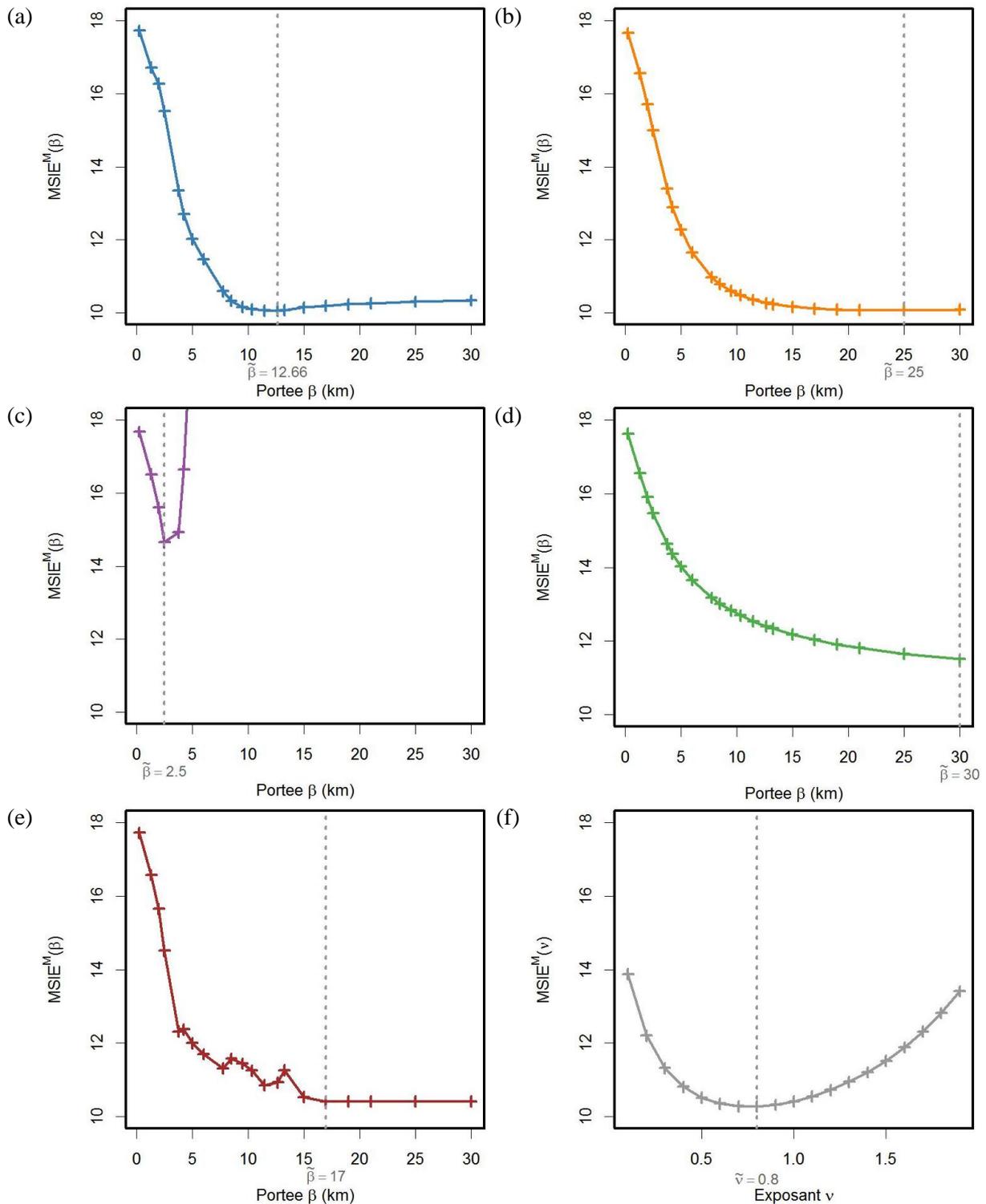


Figure IV-4 : Valeur de la MSIE en fonction de la portée  $\beta$  (ou exposant  $\nu$ ) pour les modèles : (a) sphérique, (b) exponentiel, (c) gaussien, (d) stable, (e) linéaire avec palier et (f) puissance. ( $T = 1h$ ).

La ligne en tirets gris indique l'estimation de la portée  $\hat{\beta}$  (ou exposant  $\hat{\nu}$ ) pour laquelle la MSIE est minimale.

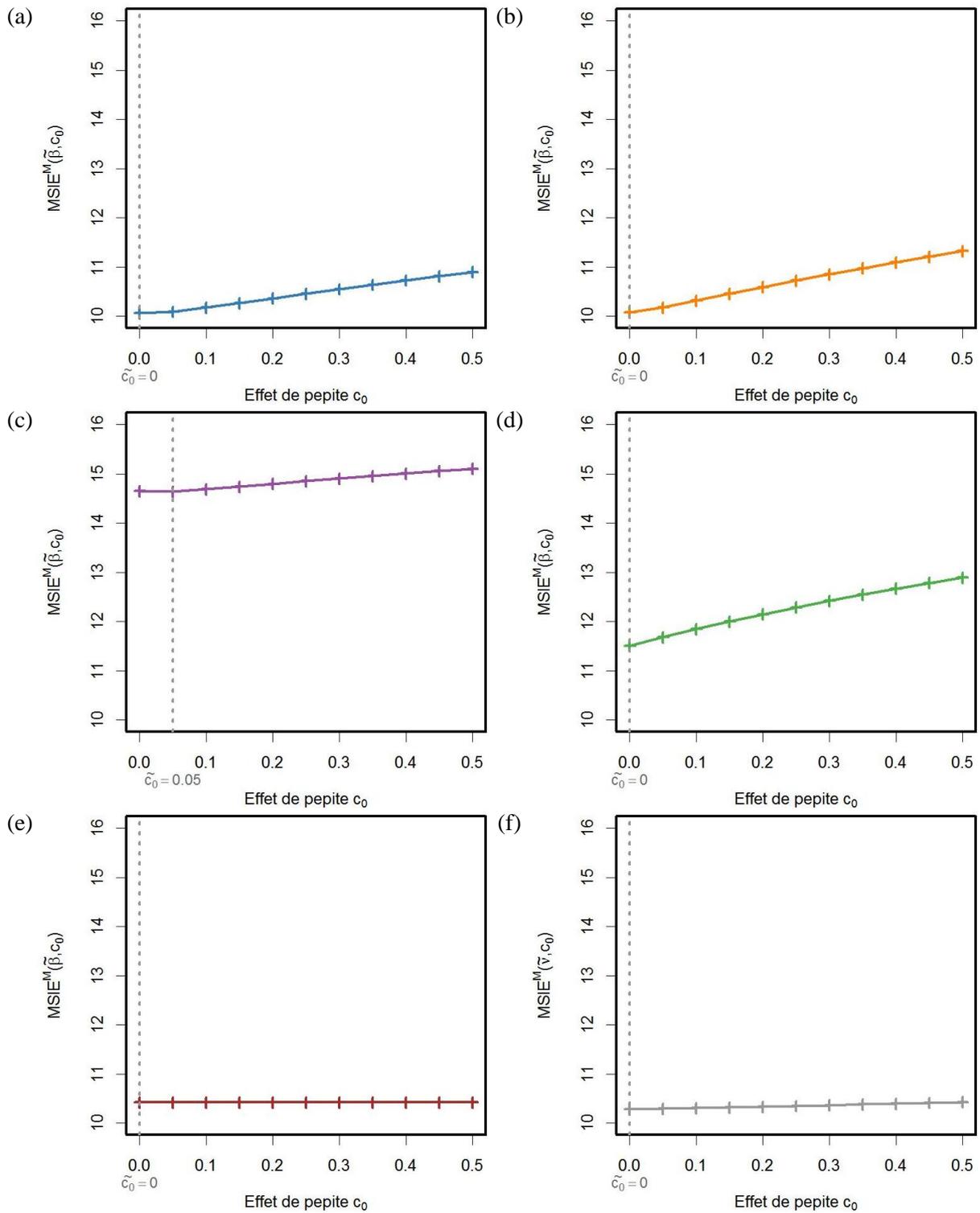


Figure IV-5 : Valeur de la MSIE( $\tilde{\beta}$  ou  $\tilde{\nu}$ ) en fonction de l'effet de pépite  $c_0$  pour les modèles : (a) sphérique, (b) exponentiel, (c) gaussien, (d) stable, (e) linéaire avec palier et (f) puissance. ( $T = 1h$ ).  
 La ligne en tirets gris indique l'estimation de l'effet de pépite  $\tilde{c}_0$  pour laquelle la MSIE est minimale.

Enfin, avec le modèle **puissance**, l'évolution de la MSIE décrit une parabole, avec un minimum local pour un exposant  $\tilde{\nu}$  de 0.8.

Afin de déterminer la valeur de l'effet de pépité  $\tilde{c}_0$ , l'étude est répétée pour les différents modèles en considérant la portée  $\tilde{\beta}$  estimée. Les résultats obtenus sont présentés dans la Figure IV-5. Seul le modèle **gaussien** montre une (très) légère amélioration de la MSIE pour un effet de pépité de 0.05. Pour tous les autres modèles, l'inclusion d'un effet de pépité détériore la valeur de la MSIE de façon plus ou moins prononcée. Pour ces modèles, nous ne considérerons donc pas d'effet de pépité pour les données horaires.

Le Tableau IV-1 reporte les paramètres estimés (portée et effet de pépité) pour les 6 modèles considérés ainsi que la valeur de la MSIE associée

### IV-3.2.2 Détermination du facteur $\alpha_0$

Pour chacun des 6 modèles de variogramme identifiés, le critère de la SMSE nous a permis d'estimer le facteur  $\tilde{\alpha}_0$  (équation [IV-3]) que nous résumons dans le Tableau IV-2.

Tableau IV-1 : Estimation de la portée et de l'effet de pépité pour les 6 modèles considérés et la valeur de la MSIE associée. (T = 1h).

Modèle $M$	Portée $\tilde{\beta}$	Effet de pépité $\tilde{c}_0$	MSIE	Rang
<b>Sphérique</b>	12.66 km	0	10.05	1
<b>Exponentiel</b>	25.00 km	0	10.07	2
<b>Gaussien</b>	2.50 km	0.05	14.64	6
<b>Stable</b>	30.00 km	0	11.51	5
<b>Linéaire + Palier</b>	17.00 km	0	10.42	4
<b>Puissance</b>	exposant $\tilde{\nu}$ : 0.8	0	10.28	3

Tableau IV-2 : Estimation du facteur  $\tilde{\alpha}_0$  pour les 6 modèles considérés. (T = 1h).

Modèle $M$	Palier $\tilde{\alpha}_0$
<b>Sphérique</b>	1.47
<b>Exponentiel</b>	1.52
<b>Puissance</b>	facteur d'échelle $\tilde{m}$ : 0.18
<b>Linéaire + Palier</b>	2.98
<b>Stable</b>	0.56
<b>Gaussien</b>	0.53

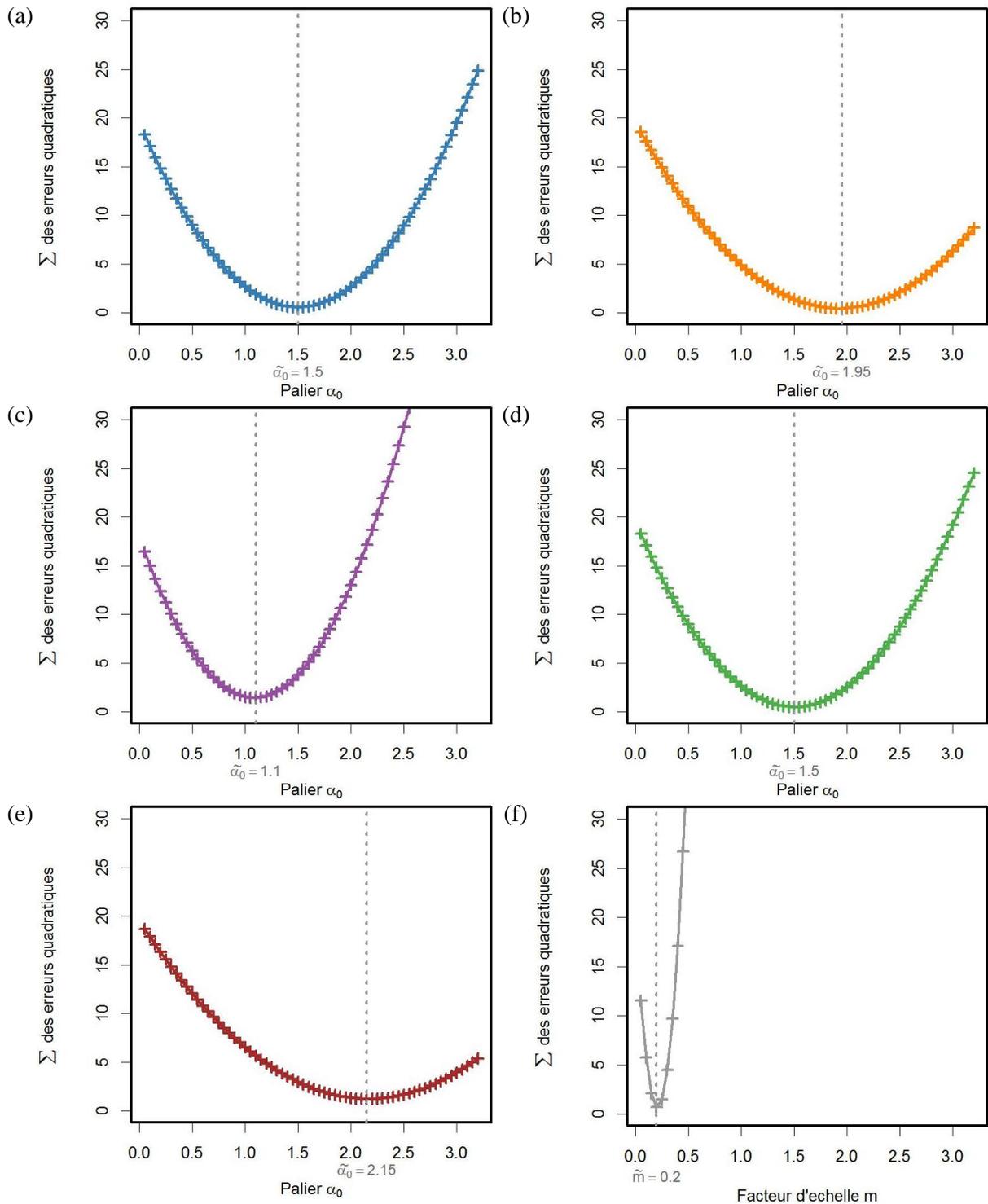


Figure IV-6 : Valeur de la SEQ en fonction du palier  $\alpha_0$  (ou facteur d'échelle  $m$ ) pour les modèles : (a) sphérique, (b) exponentiel, (c) gaussien, (d) stable, (e) linéaire avec palier et (f) puissance. ( $T = 1h$ ).

La ligne en tirets gris indique l'estimation du palier  $\tilde{\alpha}_0$  (ou facteur d'échelle  $\tilde{m}$ ) pour laquelle la SEQ est minimale.

Abstraction faite du modèle **puissance**, nous remarquons que tous les modèles affichent une valeur du palier loin de la valeur attendue de la standardisation des données (équation [IV-2] – *i.e.* 1). Par ailleurs, pour les modèles **stable** et **gaussien**, la valeur du palier  $\widetilde{\alpha}_0$  atteint tout juste la valeur du variogramme expérimental pour la plus petite classe de distance considérée (*cf.* Figure IV-3b).

Nous avons alors mis en place une seconde estimation du facteur  $\widetilde{\alpha}_0$  basée sur une approche des moindres carrés. Ainsi, pour un modèle de variogramme, le palier  $\widetilde{\alpha}_0$  est choisi parmi une liste de valeur prédéfinie tel que la somme des erreurs quadratiques (ou SEQ) entre les valeurs du variogramme expérimental  $\hat{\gamma}$  et celles du variogramme climatologique  $\Gamma$  soit minimale. Les résultats obtenus sont présentés dans la Figure IV-6.

Pour tous les modèles, l'évolution de la SEQ affiche une forme parabolique. Le minimum local donne alors la meilleure estimation du facteur  $\widetilde{\alpha}_0$  (assimilé à la valeur du palier ou du facteur d'échelle pour le modèle puissance). Nous constatons que pour le modèle **sphérique** et **puissance**, les estimations sont quasiment identiques à celles déduites par la SMSE (Tableau IV-2). Pour les modèles **exponentiel**, **gaussien** et **stable**, nous observons une augmentation de la valeur de  $\widetilde{\alpha}_0$  (respectivement +0.43, +0.57 et +0.94). À l'inverse, pour le modèle **linéaire avec palier**, nous notons une diminution (-0.83).

### IV-3.2.3 Choix du variogramme paramétrique

Accompagnés d'une estimation du variogramme expérimental  $\hat{\gamma}$ , les 6 modèles de variogrammes paramétriques  $\tilde{\gamma}$  alors identifiés sont représentés sur la Figure IV-7.

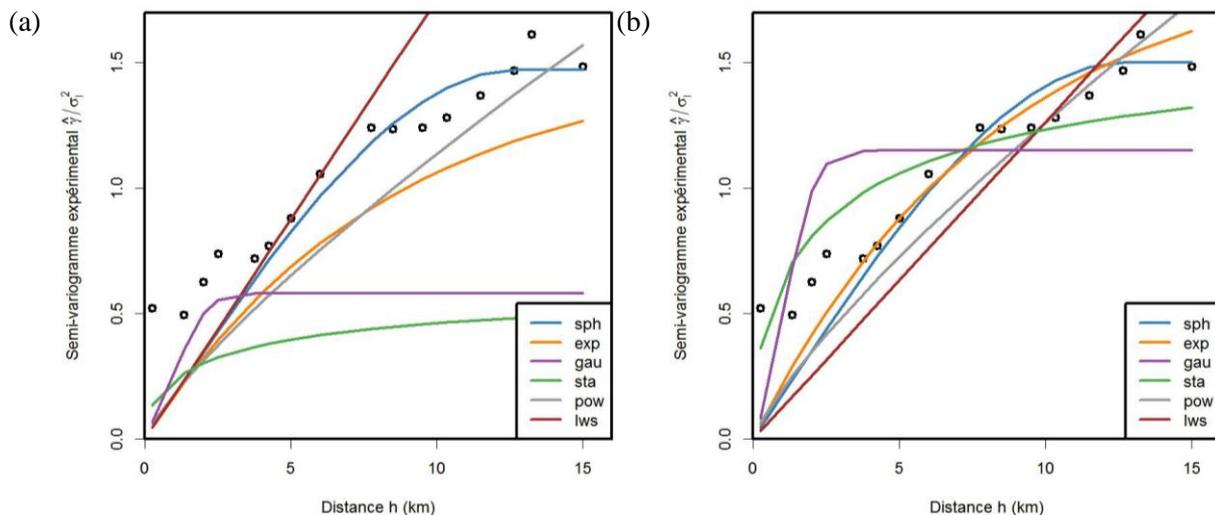


Figure IV-7 : Représentation du semi-variogramme expérimental  $\hat{\gamma}$  (normalisé par la variance de l'échantillon  $\sigma_l^2$ ) et les 6 modèles  $\Gamma(\mathbf{h}, \tilde{\beta})$  identifiés avec une estimation du palier  $\widetilde{\alpha}_0$  (ou facteur d'échelle  $\tilde{m}$ ) par (a) SMSE ou (b) SEQ. (T = 1h).

Nous noterons alors l'influence des modifications apportées par l'estimation par la technique SEQ par rapport à la technique SMSE. La Figure IV-7a montre qu'avec une estimation par SMSE du facteur  $\alpha_0$ , seul le modèle **sphérique** apparaît comme le plus représentatif. La valeur du palier semble être sous-estimée pour les modèles **exponentiel**, **gaussien** et **stable** et surestimée pour le modèle **linéaire avec palier**.

L'ajustement de la valeur du facteur  $\tilde{\alpha}_0$  par SEQ (Figure IV-7b) améliore visuellement les résultats. Nous distinguons alors trois groupes : i) les modèles **sphérique** et **exponentiel**, ii) les modèles **puissance** et **linéaire avec palier** et iii) les modèles **gaussien** et **stable**. Le premier groupe étant celui qui semble le mieux s'accorder au variogramme expérimental (sans distinction particulière entre les deux modèles). Le dernier groupe est celui dont les comportements correspondent le moins à celui des observations.

À partir des résultats obtenus, nous avons fait un choix sur le modèle de variogramme paramétrique que nous utiliserons pour nos données. Sur les 6 candidats présentés, nous avons écarté les modèles **gaussien** et **stable**, car i) leur comportement est très peu représentatif de nos données (*cf.* Figure IV-7b), ii) leur valeur de la portée  $\tilde{\beta}$  est respectivement trop petite et trop grande et iii) leur valeur de MSIE associée est des plus importantes (*cf.* Tableau IV-1). Le modèle **linéaire avec palier** présente certes une grande valeur du facteur  $\tilde{\alpha}_0$  (par SMSE ou SEQ) pourtant nous l'avons éliminé car son comportement n'est pas très réaliste. Malgré des résultats encourageant, le modèle **puissance** a l'inconvénient de ne pas posséder de palier. Ce comportement peut porter à défaut dans la reconstruction des champs de pluie notamment lorsque le domaine n'est pas entièrement couvert par le réseau d'observation (ce qui est le cas du bassin clermontois – *cf.* Figure IV-2 où tout le pourtour du domaine d'étude est dépourvu de pluviomètres). Au final, il nous reste les modèles **sphérique** et **exponentiel** au comportement très proche. Le Tableau IV-3 résume les paramètres estimés pour les deux modèles que nous avons retenus.

Tableau IV-3 : Variogrammes paramétriques  $\tilde{\gamma}$  retenus pour représenter les données du bassin clermontois. (T = 1h).

Modèle $M$	Portée $\tilde{\beta}$	Effet de pépité $\tilde{c}_0$	Palier $\tilde{\alpha}_0$ (SEQ)
<b>Sphérique</b>	12.66 km	0	1.50
<b>Exponentiel</b>	25.00 km	0	1.95

#### IV-3.2.4 Relation avec la durée d'accumulation des données

Ayant identifié le variogramme paramétrique  $\tilde{\gamma}$  pour les observations cumulées sur 1 heure, nous avons répété le processus pour différentes durées d'accumulation  $T$ . La Figure IV-8 présente les résultats obtenus pour les deux modèles retenus.

Nous observons une augmentation de la portée  $\tilde{\beta}$  avec  $T$  pour les deux modèles (Figure IV-8 a et b). Les droites de régression considérées semblent satisfaisantes (*i.e.* coefficient de corrélation  $r$  supérieur à 0.5). Nous notons que la portée est plus longue pour le modèle **exponentiel**. En effet, d'après les relations empiriques obtenues, la portée  $\tilde{\beta}$  du modèle **exponentiel** s'échelonne entre 14.56 et 25.65 km tandis que celle du modèle **sphérique** s'étend entre 8.98 et 15.83 km.

Pour l'effet de pépité  $\tilde{c}_0$  (Figure IV-8 c et d), nous constatons trois régimes. Ici, le comportement observé entre les deux modèles diffère. Le modèle **sphérique** présente une diminution de la valeur de la pépité de 0.1 à 0 entre 5 et 30 minutes. De 30 minutes jusqu'à 1 heure, il n'est nul besoin d'inclure un effet de pépité pour améliorer les interpolations. Au-delà, la valeur de l'effet de pépité augmente progressivement pour atteindre 0.19 à  $T = 24$  heures. Les relations qui décrivent ces trois régimes sont accompagnées d'un excellent coefficient de corrélation (*i.e.* supérieur à 0.9). Le modèle **exponentiel** affiche 3 paliers à valeur constante : 0 pour  $T$  entre 30 minutes et 3 heures, 0.05 entre 4 heures et 8 heures et enfin 0.1 après 12 heures. Nous noterons que globalement les estimations à partir des deux modèles sont améliorées si la valeur de l'effet de pépité augmente en même temps que  $T$ .

Enfin, pour les deux modèles, nous identifions deux régimes antinomiques pour le palier  $\tilde{\alpha}_0$  (Figure IV-8 e et f) : une augmentation entre 5 minutes et 3 heures suivit d'une diminution plus lente au-delà. Les deux régressions identifiées sont associées à d'excellents coefficients de corrélation (*i.e.* supérieurs à 0.9). Nous remarquons que la valeur du palier estimée par la régression linéaire est plus petite pour le modèle **sphérique** que pour le modèle **exponentiel** (1.73 contre 2.17 pour  $T = 3$  heures).

Finalement, quelle que soit le modèle, nous observons une influence certaine de la durée de cumul des données sur les estimations de la portée  $\tilde{\beta}$ , de l'effet de pépité  $\tilde{c}_0$  et du palier  $\tilde{\alpha}_0$ . Avec précaution, nous avançons une augmentation de la valeur de la portée avec la période d'observation. Pour l'effet de pépité, nous définissons 3 régimes distincts qui diffèrent selon le modèle employé. Enfin, nous notons une augmentation de la valeur du palier jusqu'à  $T = 3$  heures puis une diminution au-delà.

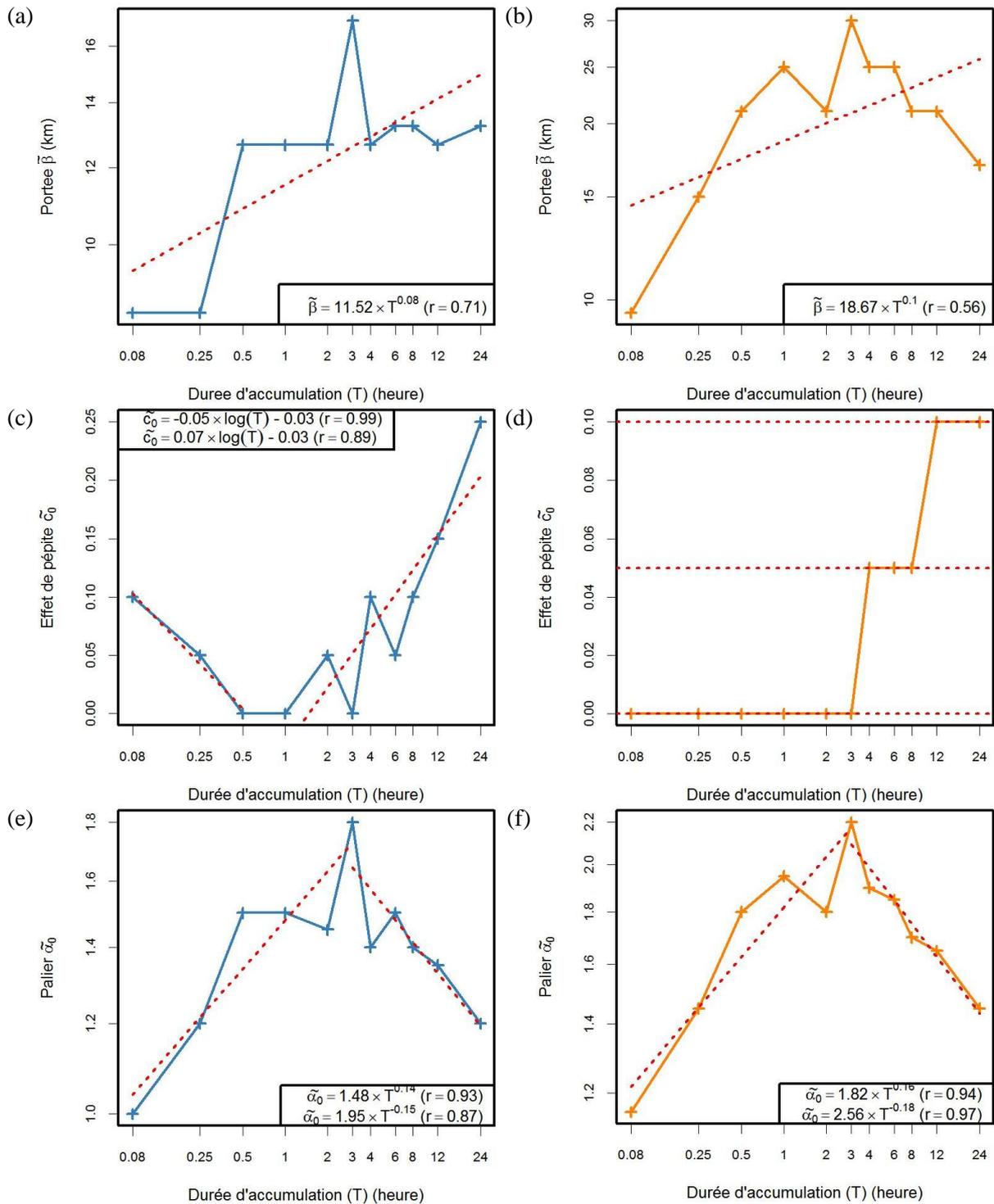


Figure IV-8 : Relation entre la portée  $\tilde{\beta}$ , l'effet de pépité  $\tilde{c}_0$  et le palier  $\tilde{\alpha}_0$  (estimé par SEQ) avec la durée d'accumulation (T) pour le modèle sphérique (a, c et e) et exponentiel (b, d et f).

La courbe en pointillé rouge représente la droite de régression entre les grandeurs. Les expressions mathématiques et les coefficients de corrélation associés  $r$  sont donnés dans les encadrés noirs.

Tableau IV-4 : Récapitulatif des variogrammes paramétriques  $\tilde{\gamma}$  utilisés pour les données du bassin clermontois. (T en heure).

Modèle	Sphérique	Exponentiel
Portée $\tilde{\beta}$ (km)	$18.67 \times T^{0.1}$	$11.52 \times T^{0.08}$
Effet de pépite $\tilde{c}_0$	$\begin{cases} 0 & T \leq 3 \text{ h} \\ 0.05 & 3 \text{ h} < T \leq 8 \text{ h} \\ 0.1 & T > 8 \text{ h} \end{cases}$	$\begin{cases} -0.05 \times \log(T) - 0.03 & T < 30 \text{ min} \\ 0 & 30 \text{ min} \leq T \leq 1 \text{ h} \\ 0.07 \times \log(T) - 0.03 & T > 1 \text{ h} \end{cases}$
Palier $\tilde{\alpha}_0$	$\begin{cases} 1.82 \times T^{0.16} & T < 3 \text{ h} \\ 2.56 \times T^{-0.18} & T \geq 3 \text{ h} \end{cases}$	$\begin{cases} 1.48 \times T^{0.14} & T \leq 3 \text{ h} \\ 1.95 \times T^{-0.15} & T > 3 \text{ h} \end{cases}$

Les deux modèles présentent un comportement et des résultats similaires. Nous noterons juste une plus grande portée  $\tilde{\beta}$  pour le modèle exponentiel. Pour les cas étudiés, nous avons considéré l'ensemble des régressions linéaires déduites, bien que certaines soient contestables (*cf.* coefficients de corrélation moyens pour les estimations de la portée). Les différents paramètres des deux variogrammes paramétriques  $\tilde{\gamma}$  considérés sont rappelés dans le Tableau IV-4.

Ayant identifié nos variogrammes paramétriques, nous pouvons dès lors appliquer la méthode du krigeage pour l'estimation du champ de pluie. La disponibilité des données nous a forcé à nous focaliser sur les événements survenus durant l'année 2011. Un peu plus d'une cinquantaine d'événements ont été enregistrés. Nous en présentons ici deux : un cas HIVER (le 16 mars 2011) et un cas ÉTÉ (le 31 août 2011). Au travers de ces deux épisodes pluvieux, nous verrons si nous pourrions différencier les deux techniques étudiées.

## IV-4 CAS HIVER : 16 Mars 2011

### IV-4.1 Situation météorologique

Au cours de la journée du 16 mars 2011, un marais barométrique s'est installé sur une grande partie de la France (Figure IV-9a). Au centre de l'Italie, un système dépressionnaire est associé à une zone frontale qui s'étend du nord de l'Italie à la Lybie. En amont, un front occlus (ligne violette) se développe. Dans un flux continental (*i.e.* nord-est = flèche orange), ce dernier passera au dessus de la région de Clermont-Ferrand en fin d'après midi. Il s'accompagnera par le développement de cellules convectives de faibles intensités (Figure IV-9b).

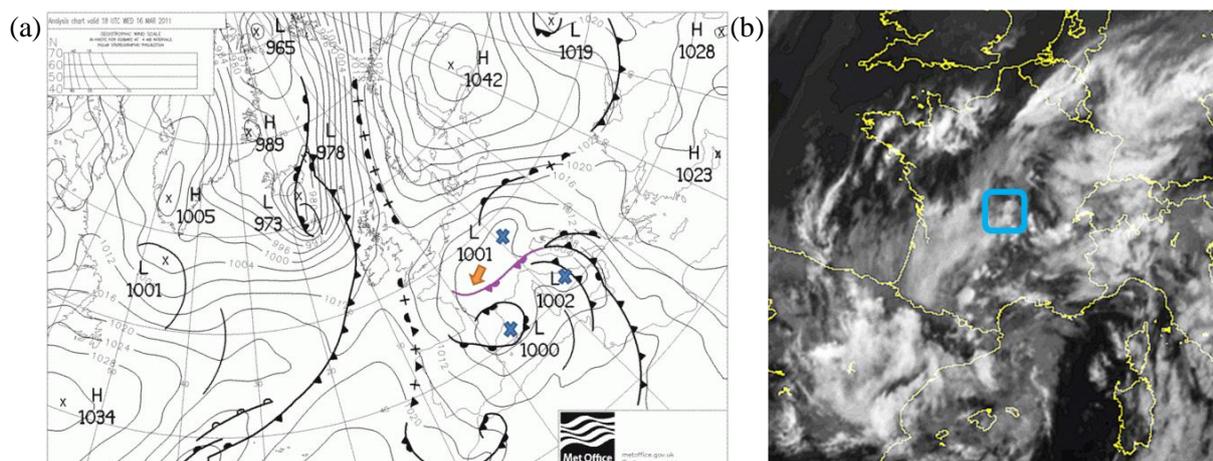


Figure IV-9 : CAS HIVER = Situation météorologique du 16 mars 2011 : (a) carte d'analyse à 17 h UTC ; (b) image satellite infrarouge à 17 h 30 UTC.  
Sources : Met Office et EUMETSAT (archive : <http://www.meteo60.fr/archives.php> | a - colorisée | b - zoomée)

## IV-4.2 Les données

L'événement considéré fut enregistré par les différents instruments sur 4 heures consécutives : de 17 h 05 à 21 h 05 UTC. La Figure IV-10 montre l'évolution des cumuls de pluie (bruts ou contrôlés) observés par les pluviomètres pour trois stations respectivement localisées au nord, au centre et au sud de notre domaine (cf. Figure IV-12a). Sont également reportées la valeur du plus proche pixel RADAR associés pour lesquelles nous distinguons les mesures sans ou avec la correction des effets d'atténuation (EA).

Première constatation : pour les trois stations, comme pour l'ensemble des stations du réseau de pluviomètre principal, le contrôle de qualité des observations appliqué ne suspecte aucune donnée (*i.e.* les courbes rouge foncé et rouge clair se superposent).

Située au nord de notre réseau de pluviomètres où sévissent principalement les précipitations, la station Blanzat (Figure IV-10a) affiche un cumul de près de 30 mm en 4 heures. Nous noterons ici que les cumuls finaux diffèrent de plus de 10 mm entre les deux mesures RADAR. La prise en compte des EA permet d'être davantage en accord avec les observations pluviométriques.

La station Crouël (Figure IV-10b), localisée au centre de notre domaine (et proche du RADAR) enregistre de plus faibles cumuls avec seulement 5 mm en 4 heures. Les deux mesures RADAR s'accordent aux observations pluviométriques. Les EA sont ici beaucoup moins prononcés (*i.e.* les deux courbes bleu foncé et bleu clair sont très similaires).

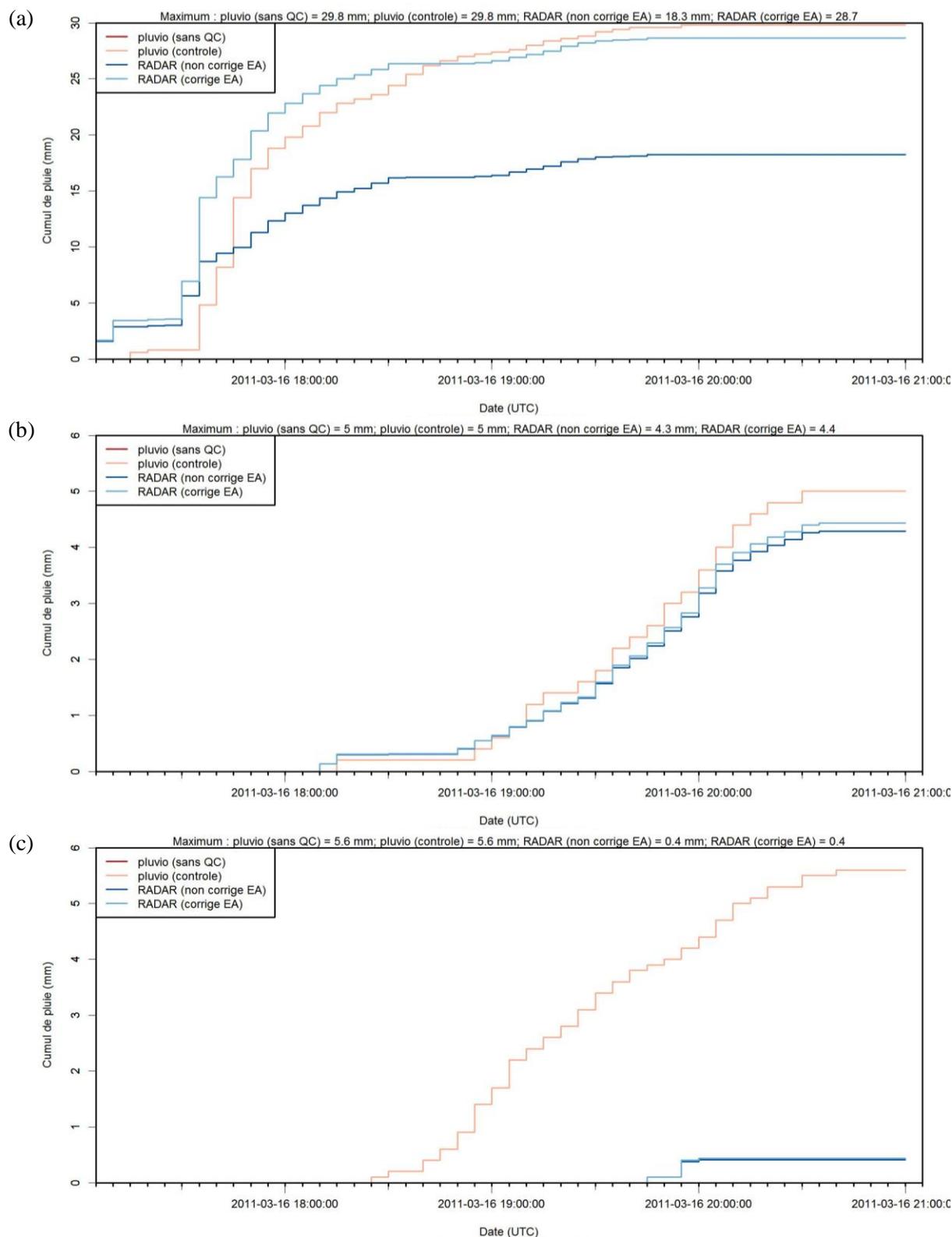


Figure IV-10 : CAS HIVER = Évolution temporelle des cumuls à 5 minutes (en mm) mesurés à : (a) Blanzat (secteur NO), (b) Crouël (secteur E) et (c) Opme (secteur SE). La courbe en rouge foncé (clair) correspond aux observations pluviométriques brutes (soumises au QC). La courbe en bleu foncé (clair) donne la valeur du pixel RADAR le plus proche de la localisation du pluviomètre non corrigée (corrigée) des EA. Les cumuls totaux enregistrés sont reportées en haut de chaque image.

Enfin, à l'extrême sud de notre réseau de pluviomètres, la station Opme (Figure IV-10c) enregistrent un cumul final de 5.6 mm en 4 heures, (équivalent à ce qui est observé à Crouël). Les mesures RADAR présentent ici un problème sur le cumul total puisqu'il est négligeable et les quelques mesures ne correspondent guère avec les mesures au sol.

Concernant les autres stations du réseau (non montré), les correspondances entre les mesures RADAR et les observations pluviométriques sont variées. Nous n'observons pas de lien évident entre la localisation des stations et la (bonne ou mauvaise) concordance des données. En général, nous avons constaté que la correction des EA est à privilégier. Par ailleurs, pour certaines stations, nous suspectons des décalages temporels entre 5 et 10 minutes. Ainsi, pour limiter les effets indésirables liés à ces problèmes de synchronisation, nous avons choisi de ne travailler qu'exclusivement sur le pas de temps de l'événement.

Pour ce premier cas d'étude, les résultats présentés ci-après ont été obtenus avec des variogrammes paramétriques pour les observations (technique du krigeage ordinaire) et les résidus (technique du krigeage avec dérive externe – selon l'approche de Velasco-Forero *et al.* (2009)) définis par un modèle **exponentiel**. Ils sont représentés dans la Figure IV-11.

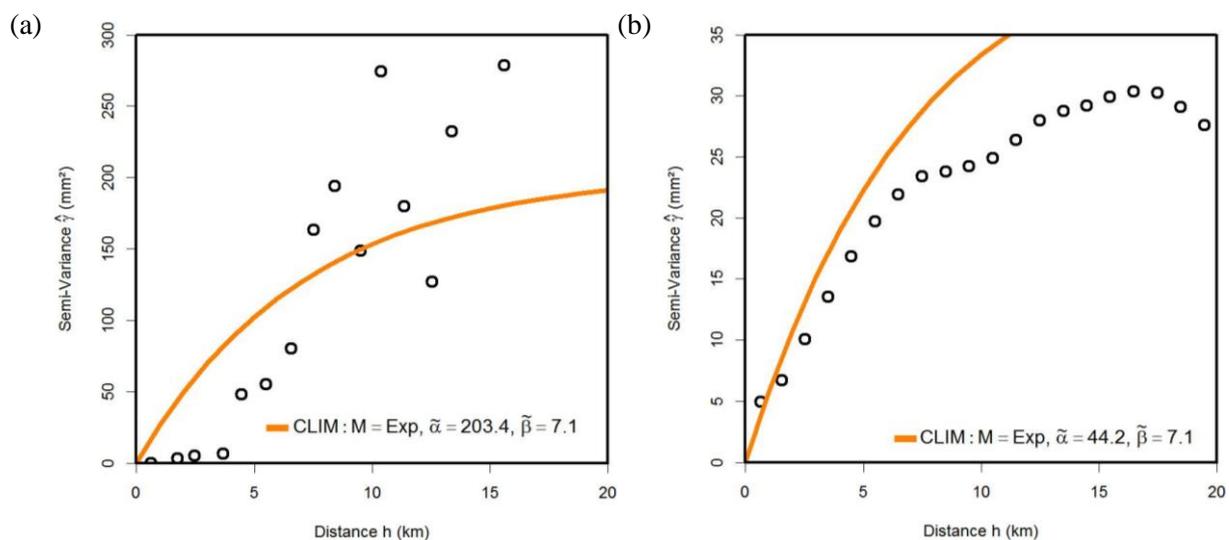


Figure IV-11 : CAS HIVER = Semi-variogramme expérimental (a) des observations pluviométriques et (b) des résidus obtenus à partir des mesures RADAR. (T = 4 h). Les variogrammes paramétriques considérés sont représentés par les courbes en orange.

La Figure IV-12 présente le champ de pluie cumulé sur  $T = 4$  h observé par les pluviomètres et par le RADAR (situé au centre de l'image). Les pluies s'organisent selon un axe nord-nord-est sud-sud-ouest, avec une phase d'intensification sur les versants au vent (*i.e.* côté est) de la chaîne des puys. Le maximum est enregistré au nord du réseau de pluviomètres.

Nous constatons alors que les mesures RADAR présentent d'importantes sous-estimations que nous incluons ou non les effets d'atténuation (*cf.* ellipses). Nous ne pouvons pas assurément prétendre que ce problème est à l'origine d'une mauvaise estimation de la loi  $Z-R$  puisque pour une station au cœur de l'événement (Figure IV-10a), la correspondance entre les observations pluviométriques et mesures RADAR semble correcte ; le problème est donc d'un autre ordre (non identifié). Par ailleurs, nous notons la présence de pixels épars qui viennent entacher l'image RADAR. Ces pixels s'apparentent à des échos de sol non éliminés correctement.

## IV-4.3 Reconstruction d'un champ de précipitation à l'échelle de l'événement

### IV-4.3.1 Validation qualitative

La Figure IV-13 montre le champ de pluie au sol reconstruit par la technique du krigeage ordinaire des données pluviométriques seules (OK\_RG), du krigeage avec dérive externe (KED) et de la fusion conditionnée (MERG).

L'interpolation par OK\_RG (Figure IV-13a) localise correctement la structure principale au nord du domaine. Or les excroissances du champ de pluie (ellipses blanches) ne sont pas reproduites. Ceci s'explique par un manque d'information au sol. En effet, ces petites structures se situent dans des régions dépourvues de pluviomètres. En dehors du réseau d'observation (forme en pointillé noir), les extrapolations étalent les précipitations jusqu'en bordure de domaine en lien avec celles observées par les pluviomètres en limite du réseau. L'interpolation par OK\_RG ne peut donc pas reproduire les structures précipitantes situées hors du réseau de pluviomètres.

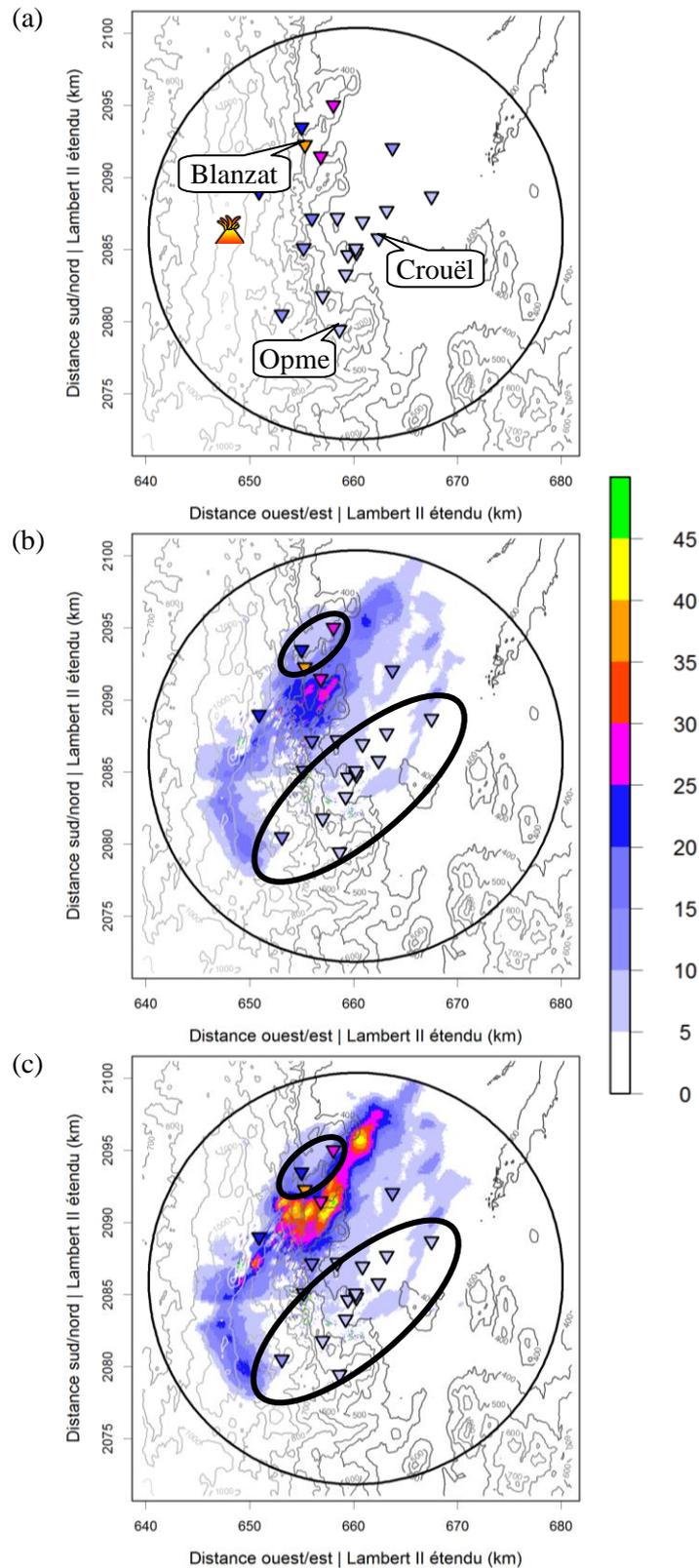


Figure IV-12 : CAS HIVER = Champ de pluie (en mm) observé par (a) les pluviomètres et le RADAR (b) en excluant et (c) en corrigeant les EA. (T = 4 h).

*La topographie de la région est rajoutée sous forme d'iso-contours en nuances de gris, du plus foncé pour les altitudes proches de 100 m, au plus clair pour les altitudes atteignant plus de 1 400 m (marqueur volcan = puy de Dôme). Sur les figures (b-c), les observations pluviométriques sont reportées par les valeurs contenues dans les triangles pointant vers le bas.*

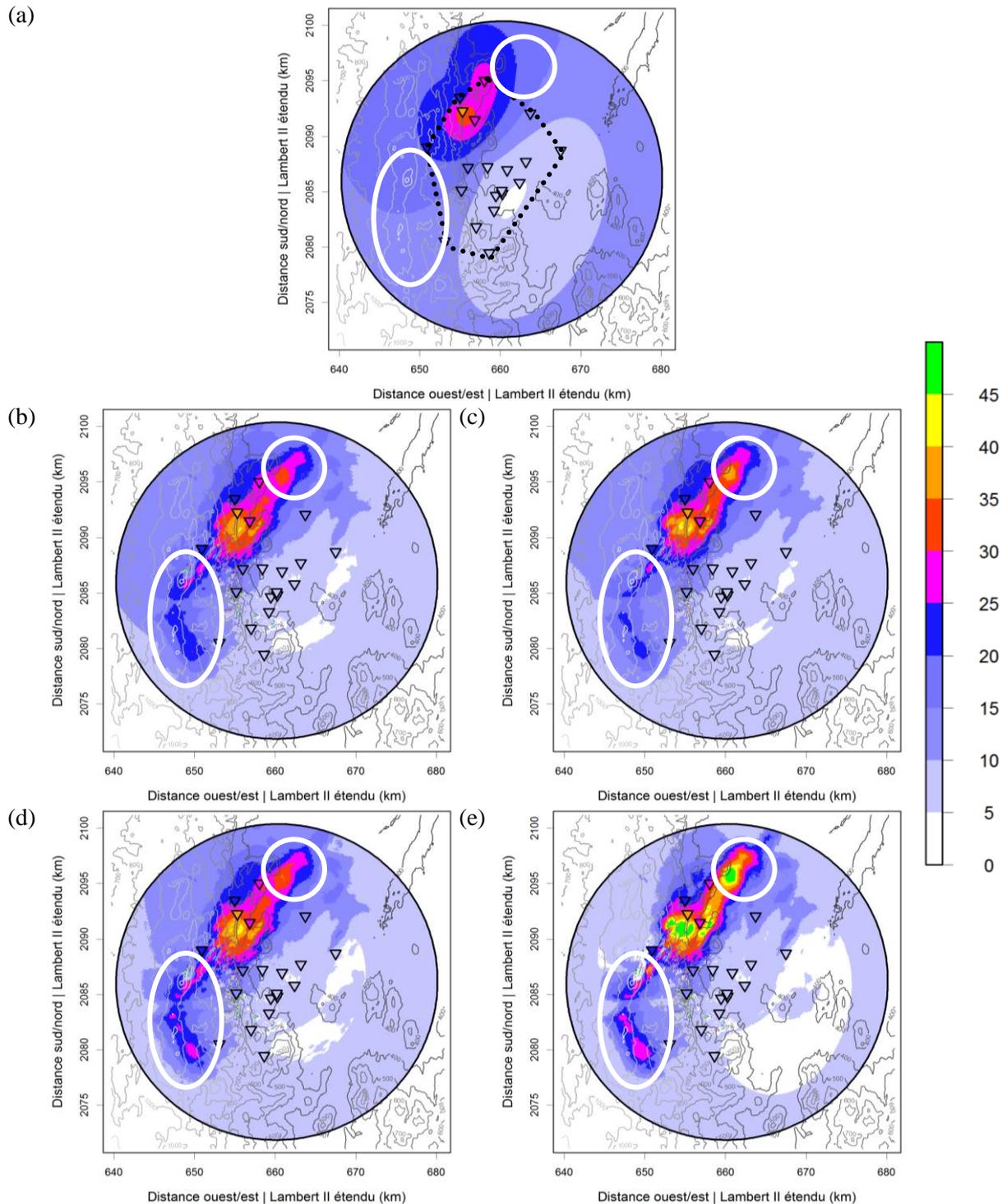


Figure IV-13 : CAS HIVER = Champ de pluie (en mm) interpolé par (a) OK\_RG, (b et c) KED et (d et e) MERG. (T = 4 h).

Les figures (c et e) sont issues des mesures RADAR corrigées des EA, contrairement aux figures (b et d). L'ellipse et le cercle localisent des structures de pluie observées par le RADAR, non détectables par le réseau de pluviomètres.

L'intégration des mesures RADAR dans la technique du KED (Figure IV-13 b et c) et du MERG (Figure IV-13 d et e) permet de retrouver le caractère spatial des pluies sur le champ final, avec la distinction des plus petites structures. En utilisant les mesures RADAR sans (Figure IV-13 b et d) ou avec (Figure IV-13 c et e) corrections des EA, la technique du MERG surestime les cumuls de pluie en comparaison avec la technique du KED. Pour les deux techniques, l'ajustement des valeurs de cumul pour s'accorder avec les observations pluviométriques est faussé par le manque d'information en dehors du réseau de pluviomètres : en bordure de domaine, nous observons un cumul non nul semblable à celui estimé par OK\_RG. Notons toutefois que la technique du MERG laisse alors apparaître des taux plus faibles en bordure nord-ouest et dessine une zone sans pluie au sud-est légèrement plus large.

La prise en compte des EA présente des effets singuliers. Tout d'abord, pour le KED (Figure IV-13 b vs c), nous constatons des effets opposés sur une même image. Par exemple, sur la structure située au nord-est (cercle), la correction des EA augmente la valeur du cumul alors que, pour la structure située au sud-ouest, nous observons une diminution des cumuls totaux. De plus, nous constatons que la zone sans pluie au sud-est du domaine semble plus restreinte lorsque nous considérons les EA. Pour la technique du MERG (Figure IV-13 d vs e), la correction des EA augmente les valeurs de cumul sur toutes les structures de pluie observées (cumuls supérieurs à 45 mm en 4 heures). Sur les bordures du domaine, la prise en compte des EA diminue visiblement les cumuls. De plus, nous observons l'apparition d'une zone sans pluie au sud-est est nettement plus large.

En conclusion du 0, nous avons notifié que la technique du krigeage ordinaire (*i.e.* OK\_RG et MERG) permettait de considérer les observations situées en dehors du domaine d'étude, ce qui n'est pas possible avec la technique du KED. Cette remarque est mise en exergue au travers de la Figure IV-14. Située à proximité ou à près de 50 km du RADAR, l'intégration de 12 stations supplémentaires est visuellement un atout. Pour la technique du OK\_RG, tout comme pour le MERG, nous supprimons de ce fait les effets incommodes des extrapolations : toute la partie sud-est est désormais sans pluie et les cumuls en bordures nord-ouest du domaine sont fortement diminués. Pour la technique du MERG (Figure IV-14 b et c), la proximité des stations en bordure sud provoque une réduction des cumuls. L'exemple le plus marquant est illustré sur la structure en arrière du puy de Dôme (ellipse noire – à comparer avec la Figure IV-13 d et e). En corrigeant des EA, nous arrivons ainsi à restreindre d'avantage l'étendue spatiale du champ de pluie et nous nous rapprochons davantage de la structure mesurée par le RADAR.

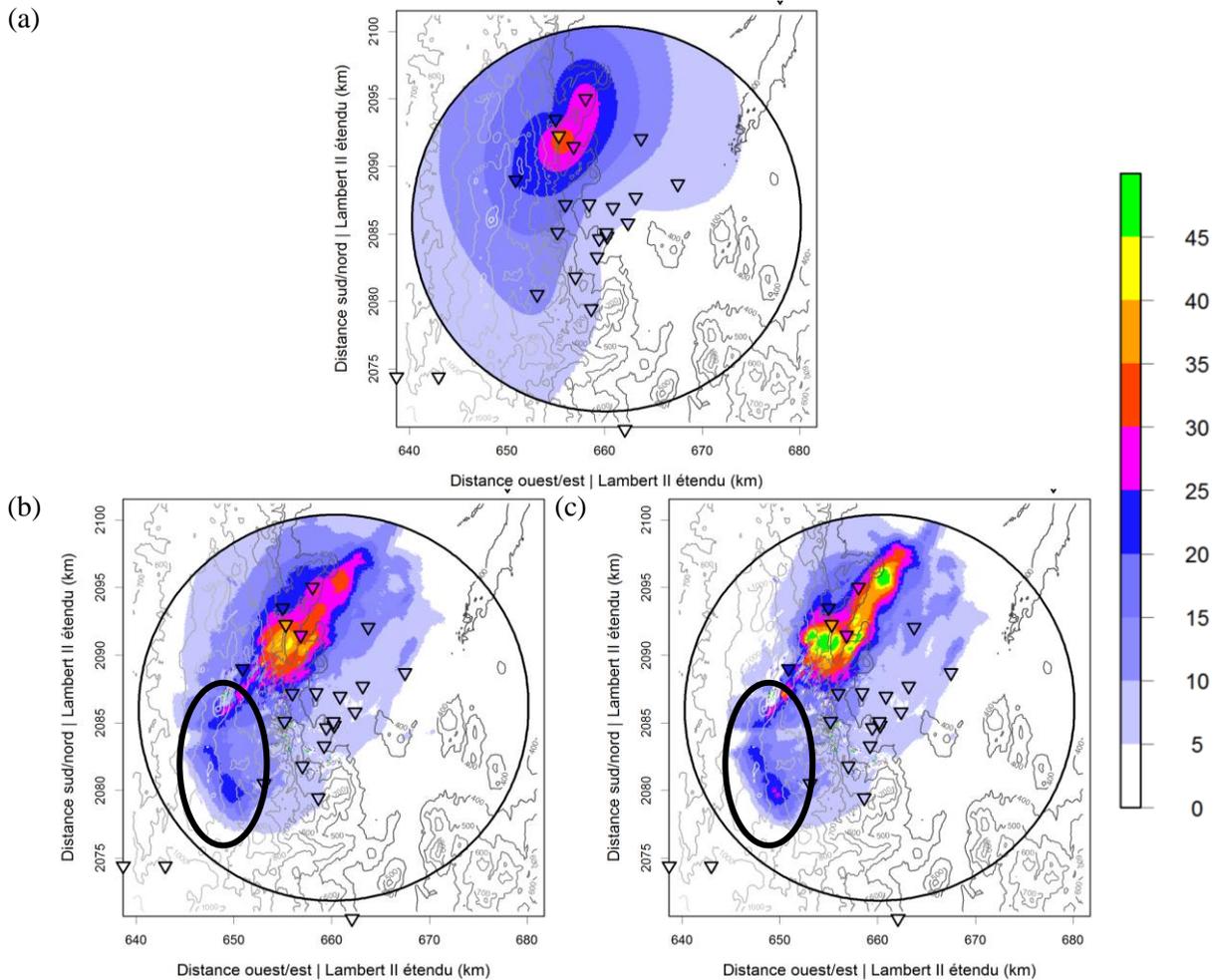


Figure IV-14 : CAS HIVER = Champ de pluie (en mm) interpolé par (a) OK\_RG et (b et c) MERG en incluant les stations situées hors du domaine d'étude. (T = 4 h). Les figures (b et c) sont issues des mesures RADAR sans et avec correction des EA. Seules 4 stations circonvoisines (sur 12) apparaissent sur la figure.

L'intérêt de la prise en compte de ces stations est également observable sur la variance de krigeage (VK). La Figure IV-15 (a et b) présente le champ de la VK associée à la technique du krigeage ordinaire ( $Var_{OK}$  – cf. équation [II-8] – pour OK\_RG et MERG). À proximité des pluviomètres,  $Var_{OK}$  semble identique selon que nous considérons ou non les observations hors du domaine d'étude. À mesure que nous nous éloignons du réseau de pluviomètres, la VK se dégrade et devient maximale en limite de domaine. L'intégration des stations circonvoisines permet de limiter sensiblement cet effet. En effet, la VK moyenne est diminuée de plus de 10 % ( $\overline{Var_{OK}} = 89.2$  et  $79.8 \text{ mm}^2$  sans et avec les stations circonvoisines).

Pour la technique du KED, nous avons décidé de représenter le champ de la VK associé ( $Var_{KED}$  – cf. équation [II-10] – Figure IV-15 c et d) sous forme de différence avec

$Var_{OK}$  (sans prise en compte des données externes). Les zones de dissemblances les plus importantes sont associées aux structures précipitantes observées par le RADAR. En bordure nord du domaine d'étude de plus fortes valeurs de  $Var_{KED}$  sont observées tandis qu'au sud, les différences sont plus faibles. Nous constatons également que  $Var_{KED}$  est localement diminuée lorsque les mesures RADAR sont corrigées des EA (ellipse en pointillé). Par ailleurs, nous retrouvons la présence des pixels résiduels mentionnée plus haut (ellipse en trait plein).

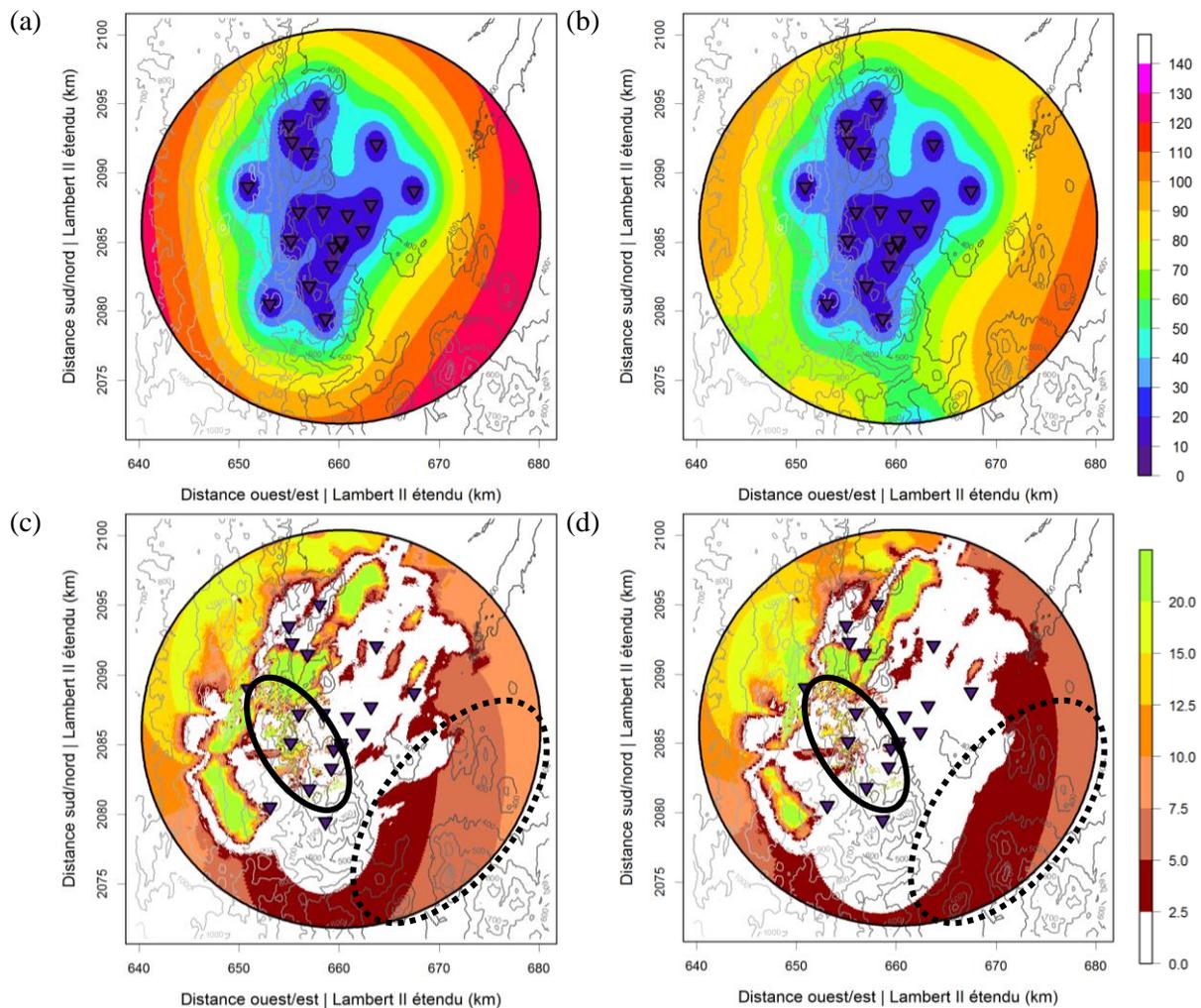


Figure IV-15 : CAS HIVER = (a et b) Champ de la variance de krigeage (en  $mm^2$ ) associée au krigeage ordinaire sans et avec inclusion des stations extérieures au domaine d'étude et (c et d) différence entre la variance de krigeage associée au KED et celle associée au krigeage ordinaire, sans ou avec correction des EA (en  $mm^2$ ). ( $T = 4 h$ ). La variance de krigeage à la position des pluviomètres est reportée dans les triangles pointant vers le bas. L'ellipse en trait plein localise quelques pixels résiduels issus des mesures RADAR et qui viennent entacher l'estimation du champ final.

Qualitativement, la technique du MERG montre des résultats prometteurs concernant la reconstruction du champ de précipitation. Elle permet de retrouver i) une ébauche de la structure hétérogène mesurée par le RADAR et ii) des estimations quantitatives proches des observations pluviométriques. Bien que la densité du réseau d'observations au sol soit importante (= 1 pluviomètre pour 66 km<sup>2</sup> – contre 1 pour plus de 90 km<sup>2</sup> pour les cas des cévenols étudiés), le problème ici est la petitesse du réseau de pluviomètres qui ne couvre qu'une partie du champ RADAR. En conséquence, les extrapolations sont importantes et les estimations en périphérie de domaine sont alors biaisées. Contrairement à la technique du KED, le MERG permet de tenir compte des observations extérieures au domaine. Pour le cas présenté ici, le champ reconstruit par MERG affiche certes des cumuls plus importants dans la zone de pluie principale mais les valeurs sur les bords du domaine sont plus faibles. Cette dernière remarque est également visible sur la variance de krigeage associée.

#### IV-4.3.2 Validation quantitative

Le Tableau IV-5 donne une première comparaison succincte entre les différents champs de précipitation où sont reportées à la fois la valeur maximale, la proportion de pixels ayant un cumul supérieur à 5 mm (soit la première classe de l'échelle de couleur considérée) et la quantité totale de pluie tombée au sol. Le problème lié à la présence de pixels résiduels à très fortes valeurs (constatés sur les mesures RADAR et le champ reconstruit par KED et MERG) a été contournée en utilisant uniquement les pixels dont les valeurs sont comprises entre le premier et dernier centile. Les échantillons considérés sont composés de l'ensemble des champs (*i.e.* RADAR, OK\_RG, KED et MERG) pour lesquels nous avons distingué les cas sans ou avec correction des EA.

Tableau IV-5 : CAS HIVER = Comparatif des différents champs spatiaux. (T = 4 h). Les valeurs des premier et dernier centiles sont indiquées dans la première ligne. La surface du champ de pluie correspond au pourcentage du nombre de pixels ayant une valeur supérieur ou égale à 5 mm. L'astérisque (\*) indique les cas où les stations extérieures au domaine d'étude ont été prises en compte.

	sans correction EA [0 - 31.7] mm			avec correction EA [0 - 35.5] mm		
	Valeur maximale	Surface du champ	Quantité tombée	Valeur maximale	Surface du champ	Quantité tombée
RADAR	30.0 mm	31.3 %	2.718 Mt	33.5 mm	34.0 %	3.330 Mt
OK_RG	31.7 mm	95.3 %	12.160 Mt	34.8 mm	95.3 %	12.081 Mt
OK_RG*	31.7 mm	65.0 %	8.411 Mt	34.8 mm	64.9 %	8.338 Mt
KED	31.7 mm	94.3 %	10.368 Mt	35.1 mm	96.3 %	10.123 Mt
MERG	31.7 mm	92.3 %	10.033 Mt	35.5 mm	82.4 %	8.425 Mt
MERG*	31.7 mm	59.5 %	6.339 Mt	35.4 mm	54.9 %	5.163 Mt

Comme constaté précédemment, la correction des EA sur les mesures RADAR induit une surface de pluie plus grande (= +9%) amenant plus de 600 kilotonnes de pluie supplémentaires et un maximum relevé de presque 12%. Concernant les estimations par OK\_RG, les valeurs maximales sont proches de celles mesurées par le RADAR mais dont les superficies sont triplées et les quantités totales sont plus que triplées (avec correction des EA) voire quadruplées (sans correction des EA). L'inclusion des stations extérieures au domaine (OK\_RG\*) permet de limiter à la fois la superficie et la quantité de pluie totale.

Sans la correction des mesures RADAR, nous notons une forte similitude des valeurs maximales pour le champ du KED et du MERG en comparaison du champ par OK\_RG. Pour la technique du KED, nous notons une diminution de la superficie et de la quantité des précipitations. La différence est davantage marquée avec la technique du MERG et plus encore lorsque les stations circonvoisines sont incluses (MERG\*).

Avec la correction des EA, nous observons une légère augmentation des valeurs des maximums par rapport au champ OK\_RG avec les deux techniques combinant les observations. Le champ du KED affiche une surface de pluie légèrement plus grande pour une intensité équivalente. C'est encore avec la technique du MERG que nous obtenons la plus petite surface de pluie avec le minimum de quantité de pluie.

Malgré des différences infimes, c'est le champ reconstruit par MERG qui présente une structure de pluie la plus en accord avec le champ mesuré par le RADAR. Nous notons que la superficie et la quantité totale des pluies sont fortement diminuées lorsque nous corrigeons les EA mais surtout lorsque nous incluons les stations situées hors du domaine d'étude (MERG\* vs OK\_RG : surface = -42 %, intensité = - 57 %).

Le Tableau IV-6 présente trois scores statistiques résultants de la comparaison des observations au sol avec d'une part la valeur du pixel du champ RADAR le plus proche des pluviomètres et d'autre part les estimations aux points des pluviomètres par OK\_RG, KED et MERG selon la technique du LOO (*cf.* paragraphe II-2.1.1-d). Pour ce faire, nous avons utilisé les pluviomètres comme référence. Nous avons également distingué les cas où la correction des EA a été, ou non, prise en compte. Enfin, nous avons rajouté les scores obtenus lorsque les observations situées hors du domaine d'étude ont été intégrées. Les deux techniques concernées ont été marquées d'un astérisque.

Tableau IV-6 : CAS HIVER = Comparaison entre les observations du réseau principal de pluviométriques et les différents champs spatiaux. (T = 4 h).  
L'astérisque (\*) indique les cas où les stations extérieures au domaine d'étude ont été prises en compte. Les indications (-) et (+) indiquent respectivement le moins bon et le meilleur score.

		Erreur moyenne		RMSE		Coef. corrélation		
		ME (mm)		(mm <sup>2</sup> )				
RADAR	sans	avec correction EA	6.5 (-)	4.7(-)	9.2 (-)	7.0 (-)	0.789 (-)	0.834 (-)
OK_RG			-0.1(+)		4.4 (+)		0.875 (+)	
OK_RG*			0.7		4.6		0.872	
KED	sans	avec correction EA	-0.4	-0.7	5.1	5.7	0.852	0.839
MERG	sans	avec correction EA	-0.1(+)	-0.3	4.6	5.5	0.870	0.846
MERG*	sans	avec correction EA	0.7	0.5	4.8	5.5	0.861	0.843

Seules des mesures RADAR surestiment grandement les observations pluviométriques (grandes ME positives), parmi lesquelles nous notons la présence de larges erreurs (grandes RMSE). Nous remarquons que la correction des EA permet de diminuer de presque 28% la ME et de 24% la RMSE. Le coefficient de corrélation spatial est également amélioré (mais reste inférieur à 0.9). En comparaison avec les différentes techniques de reconstruction spatiale du champ de pluie, les mesures RADAR corrigées ou non des EA, affichent les moins bons scores. Les estimations par OK\_RG donnent quantitativement les meilleurs résultats<sup>⊕</sup> : plus faibles valeurs de ME et RMSE et meilleur coefficient de corrélation. Nous constatons que l'intégration des stations extérieures au domaine d'étude détériore légèrement les scores obtenus (ME = +0.8 mm ; RMSE = +0.2 mm<sup>2</sup> ; corrélation = -3 10<sup>-3</sup>).

Concernant les techniques combinant les observations, nous remarquons que le MERG affiche des scores supérieurs à ceux du KED, avec de plus faibles écarts d'estimation en moyenne (cf. ME et RMSE) et une meilleure concordance avec les mesures au sol (cf. coefficient de corrélation). *A contrario* des mesures RADAR, la prise en compte des EA pour le KED et le MERG dégrade les scores obtenus. Les scores résultants de la prise compte des observations situées hors du domaine d'étude (MERG\*) sont légèrement moins bons. En considérant les EA, la ME affiche une plus faible valeur, bien que la RMSE et le coefficient de corrélation soient de moins bonnes qualités.

<sup>⊕</sup> Ce constat est biaisé puisque nous comparons des observations pluviométriques avec leurs interpolations certes par LOO.

Les résultats obtenus par KED et MERG montrent pour le cas HIVER des différences prononcées. La technique du MERG semble la plus adaptée pour ce domaine d'étude puisqu'elle permet de prendre en compte des stations extérieures. Quantitativement, les résultats obtenus sont de moindre qualité que ceux obtenus par la technique OK\_RG mais meilleurs que ceux des mesures RADAR ou des estimations par KED. Il s'avère que la correction des EA (effets considérés comme indésirables) altère ici la reconstruction du champ de pluie.

## IV-5 CAS ÉTÉ : 31 Août 2011

### IV-5.1 Situation météorologique

Dans la journée du 31 août 2011, le positionnement d'un anticyclone sur l'Angleterre (croix rouge – Figure IV-16a) perturbe la propagation du flux d'ouest. Une dépression, située sur la péninsule Ibérique, favorise la mise en place d'un flux de sud (flèche orange) accompagnant un front chaud. Cette remontée d'air chaud et humide en provenance de la Méditerranée générera dans l'après-midi de l'instabilité sous la forme d'orage (*cf.* ligne de grain représentée par la ligne verte). Développées dans le Massif central et entraînées dans un flux de sud-ouest, des cellules convectives (Figure IV-16b) accompagnées de fortes précipitations s'abattront sur tout le département en fin d'après-midi et début de soirée.

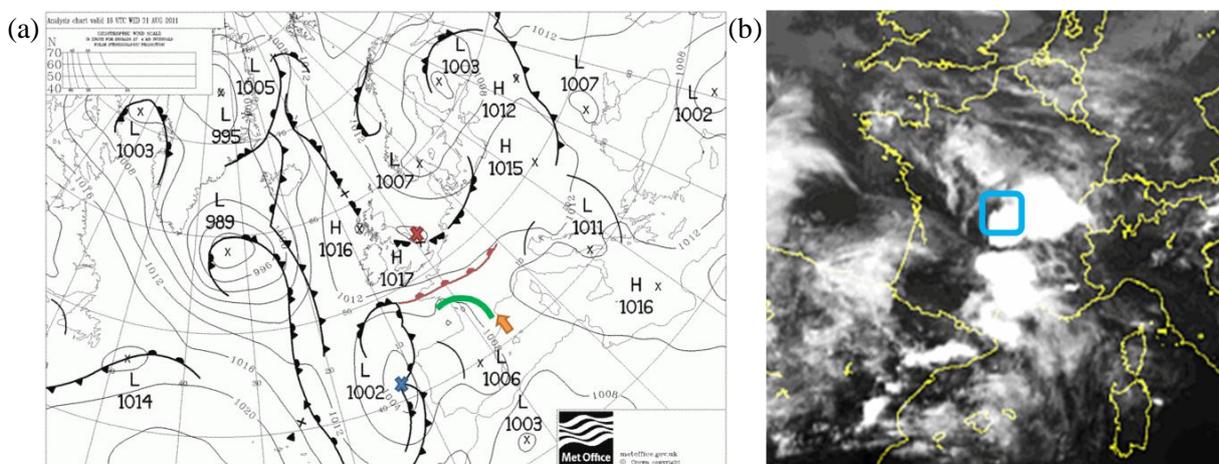


Figure IV-16 : CAS ÉTÉ = Situation météorologique du 31 août 2011 à 18 h UTC : (a) carte d'analyse ; (b) image satellite infrarouge.

Sources : Met Office et EUMETSAT (archive : <http://www.meteo60.fr/archives.php> | a - coloriée | b - zoomée).

## IV-5.2 Les données

Durant 8 heures, nous observerons le passage de trois événements précipitant successifs : le premier est survenu entre 16 et 17 h UTC, il fut suivi par un second épisode plus long (entre 18 h 30 et 21 h UTC), enfin les dernières pluies s'abattent sur l'agglomération clermontoise après 22 h UTC. La Figure IV-17 montre les évolutions des cumuls de pluie (bruts ou contrôlés) observées par trois stations respectivement localisées au nord-est, au centre-est et en plein centre de notre domaine (*cf.* Figure IV-19a). Les mesures RADAR correspondantes sont reportées, en distinguant celles corrigées, ou non, des EA.

Pour les trois stations présentées, le contrôle de qualité appliqué ne suspecte aucune donnée (*i.e.* les courbes en rouge clair et rouge foncé se superposent). Pour les 12 autres stations, le QC détectera deux observations issues de deux stations différentes qui ont été supprimées de notre base de données<sup>Ⓢ</sup>.

La station Bédât (Figure IV-17a) affiche un petit cumul total d'environ 14 mm. Nous constatons que la correction des EA sur les mesures RADAR a ici un impact non négligeable : le cumul sur l'événement montre une différence de 2.1 mm entre les deux mesures RADAR. La prise en compte des EA permet toutefois de se rapprocher davantage des observations pluviométriques (différence finale de 3 mm). Nous noterons que les deux instruments (pluviomètres et RADAR) présentent toutefois des dissemblances temporelles importantes. En effet, les événements mesurés par le RADAR semblent plus courts en durée (*cf.* lignes en pointillé). Pour le second épisode, outre le fait que les mesures RADAR semblent débiter 10 minutes avant les observations au sol, nous observons un palier à valeur nulle (*cf.* flèche) après une demi-heure d'enregistrement (et qui perdurera 25 minutes). Ce comportement n'est pourtant pas observé par le pluviomètre. Enfin, les faibles précipitations observées après 20 h UTC ne sont pas mesurées par le RADAR. La concordance entre les deux types d'instruments n'est pas satisfaisante pourtant nous avons conservé cette station pour notre analyse.

Plus à l'ouest, la station d'Herbert (Figure IV-17b) enregistre un cumul de pluie sur l'événement plus important. Ici, la correction des EA sur les mesures RADAR augmente de près de 80% le cumul total. Cette différence est principalement concentrée en une seule mesure enregistrée durant le premier épisode (*cf.* flèche). De ce fait, ce sont les mesures RADAR sans correction des EA qui s'accordent le plus avec les observations du pluviomètre. Contrairement à la station précédente, les événements semblent correctement enregistrés.

<sup>Ⓢ</sup> À titre d'information : à 15 h 55 à la station CEMAGREF (valeur mesurée = 2.2 mm, valeur attendue = 0 mm) ; à 16 h 40 à la station Blanzat (valeur mesurée = 4.4 mm, valeur attendue = 2.3 mm).

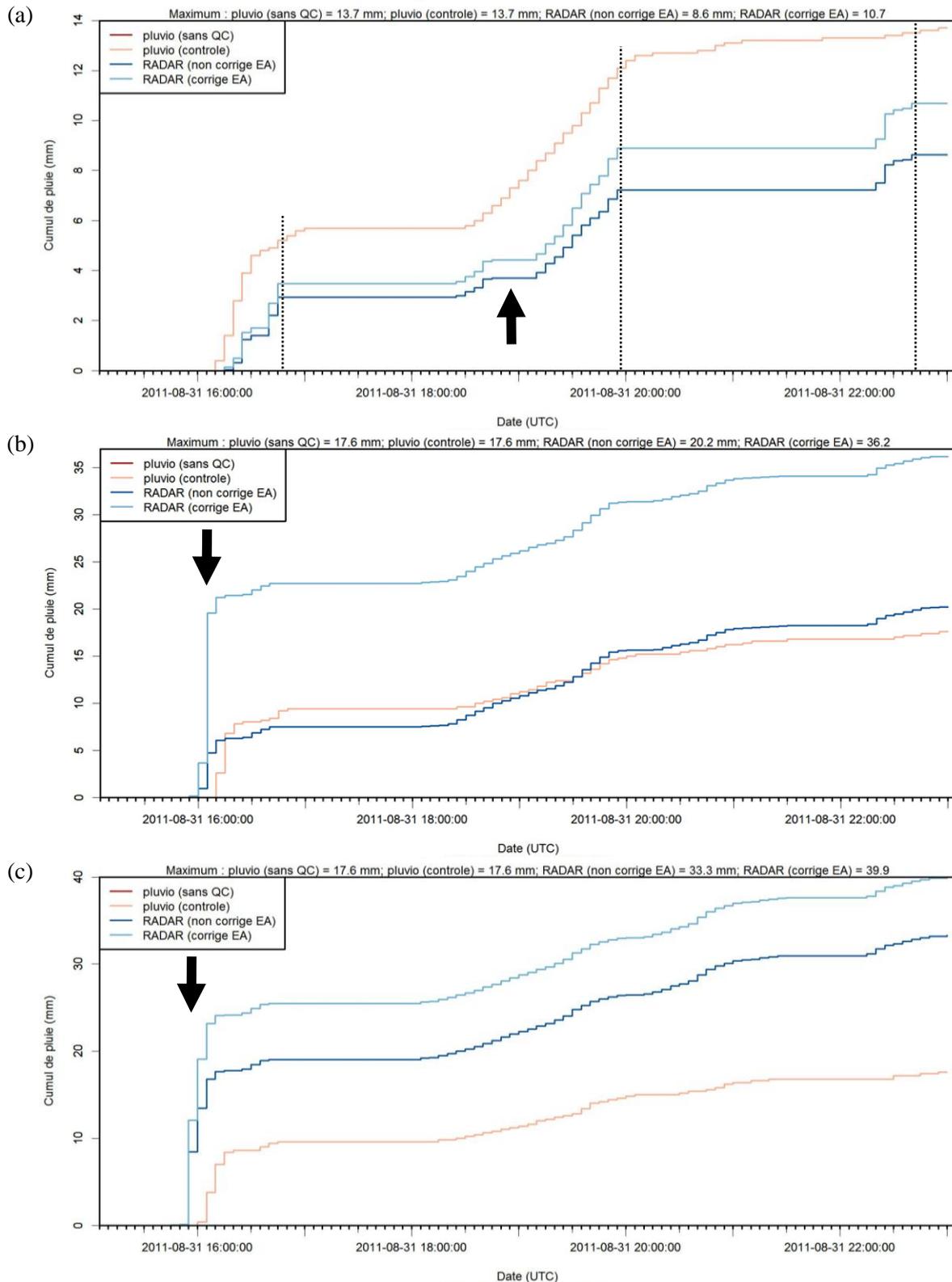


Figure IV-17 : CAS ÉTÉ = Évolution temporelle des cumuls à 5 minutes (en mm) mesurés à : (a) Bédât (secteur NE), (b) Herbert (secteur E) et (c) St-Jacques (centre). La courbe en rouge foncé (clair) correspond aux observations pluviométriques brutes (soumises au QC). La courbe en bleu foncé (clair) donne la valeur du pixel RADAR le plus proche de la localisation du pluviomètre non corrigées (corrigées) des EA. Les cumuls totaux enregistrés sont reportées en haut de chaque image.

La dernière station, St Jacques (Figure IV-17c), située à moins de 800 mètres à vol d'oiseau du RADAR, montre également une bonne synchronisation des données hormis quelques mesures RADAR anormales enregistrées durant le premier épisode (*cf.* flèche). Au final, ces erreurs d'estimations amènent à une différence de +15.7 mm (+22.3 mm avec correction des EA) avec les observations au sol. La correction des EA n'est ici encore pas justifiée.

Concernant les autres stations (non montré), la correspondance entre les mesures RADAR et les observations pluviométriques sont variées. Pour 12 stations sur 17, nous avons constaté que la correction des EA affiche des écarts positifs et importants par rapport aux mesures RADAR non corrigées. Sinon, la synchronisation des données semble compromise pour l'ensemble des stations. C'est pourquoi nous travaillerons *comme pour le cas HIVER* uniquement sur le pas de temps de l'événement.

Pour ce second cas d'étude, les résultats qui sont présentés dans ce manuscrit ont été obtenus à partir des variogrammes paramétriques (Figure IV-11) pour les observations (technique du krigeage ordinaire) et les résidus (KED – selon l'approche de Velasco-Forero *et al.* (2009)) définis par un modèle **sphérique**.

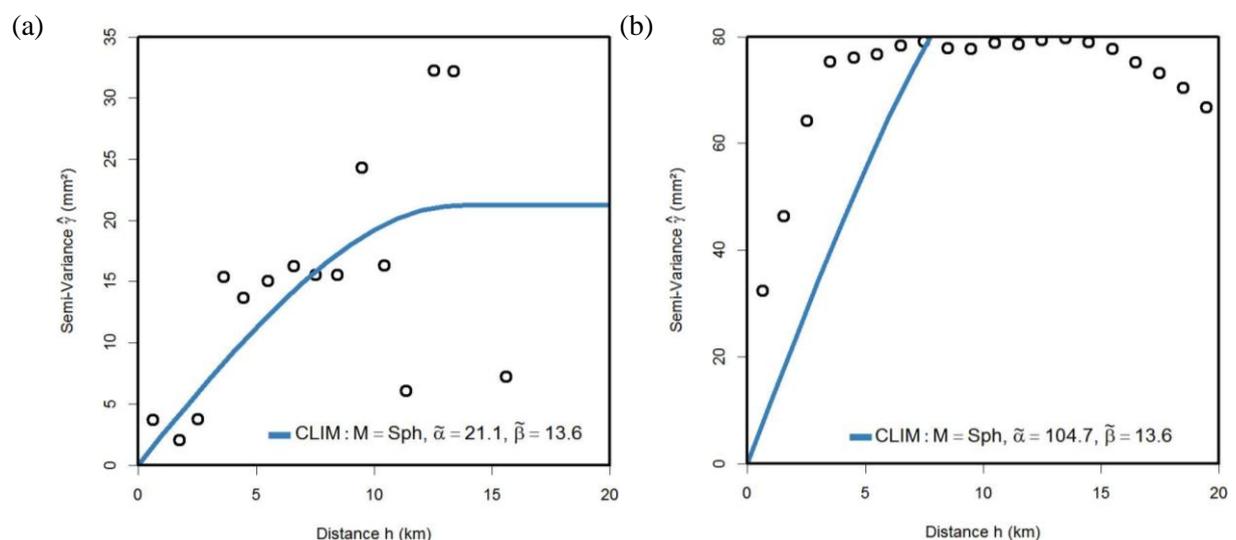


Figure IV-18 : CAS ÉTÉ = Semi-variogramme expérimental (a) des observations pluviométriques et (b) des résidus obtenus à partir des mesures RADAR. (T = 8 h). Les variogrammes paramétriques considérés sont représentés par les courbes en bleu.

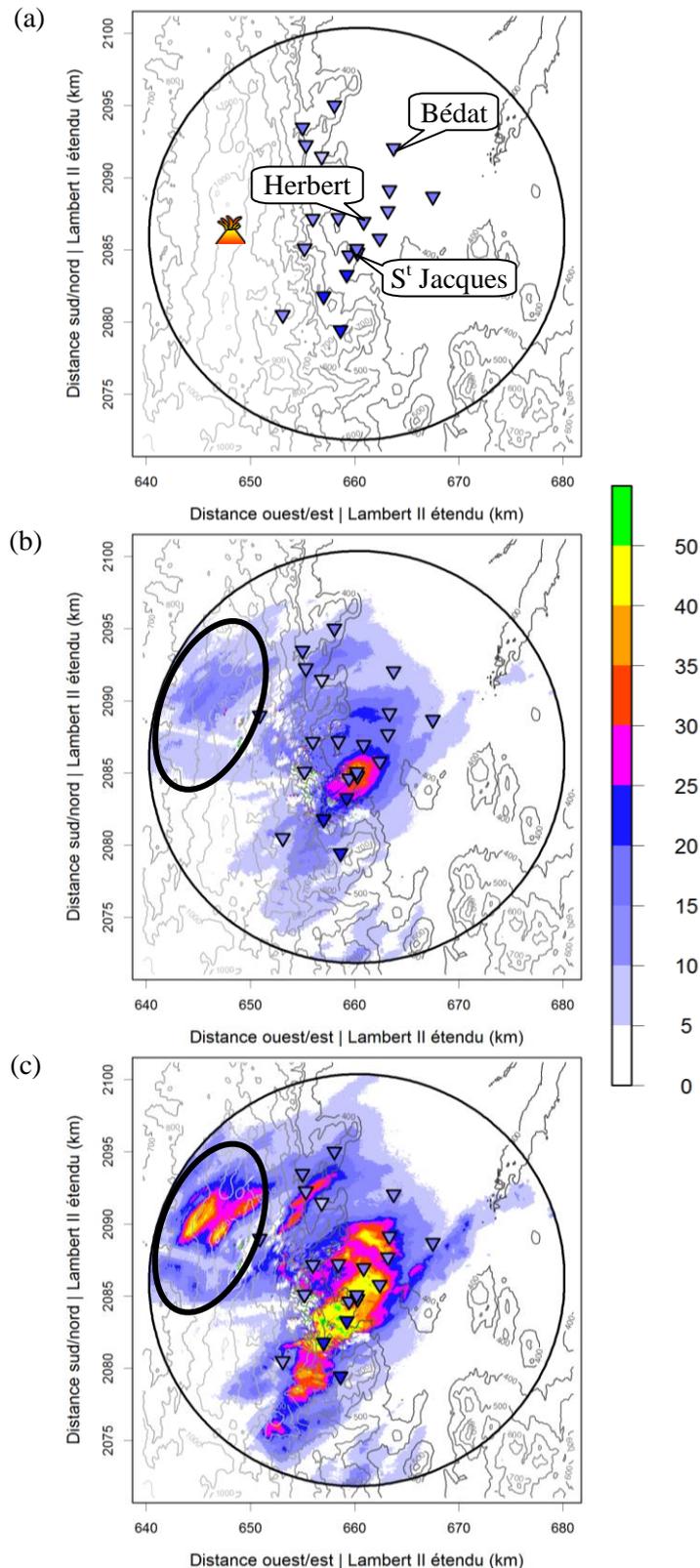


Figure IV-19 : CAS ÉTÉ = Champ de pluie (en mm) observé par (a) les pluviomètres et le RADAR (b) en excluant et (c) en corrigeant les EA. (T = 8 h).  
 La topographie de la région est rajoutée sous forme d'iso-contours en nuances de gris, du plus foncé pour les altitudes proches de 100 m, au plus clair pour les altitudes atteignant plus de 1 400 m (marqueur volcan = puy de Dôme). Sur les figures (b-c), les observations pluviométriques sont reportées par les valeurs contenues dans les triangles pointant vers le bas.

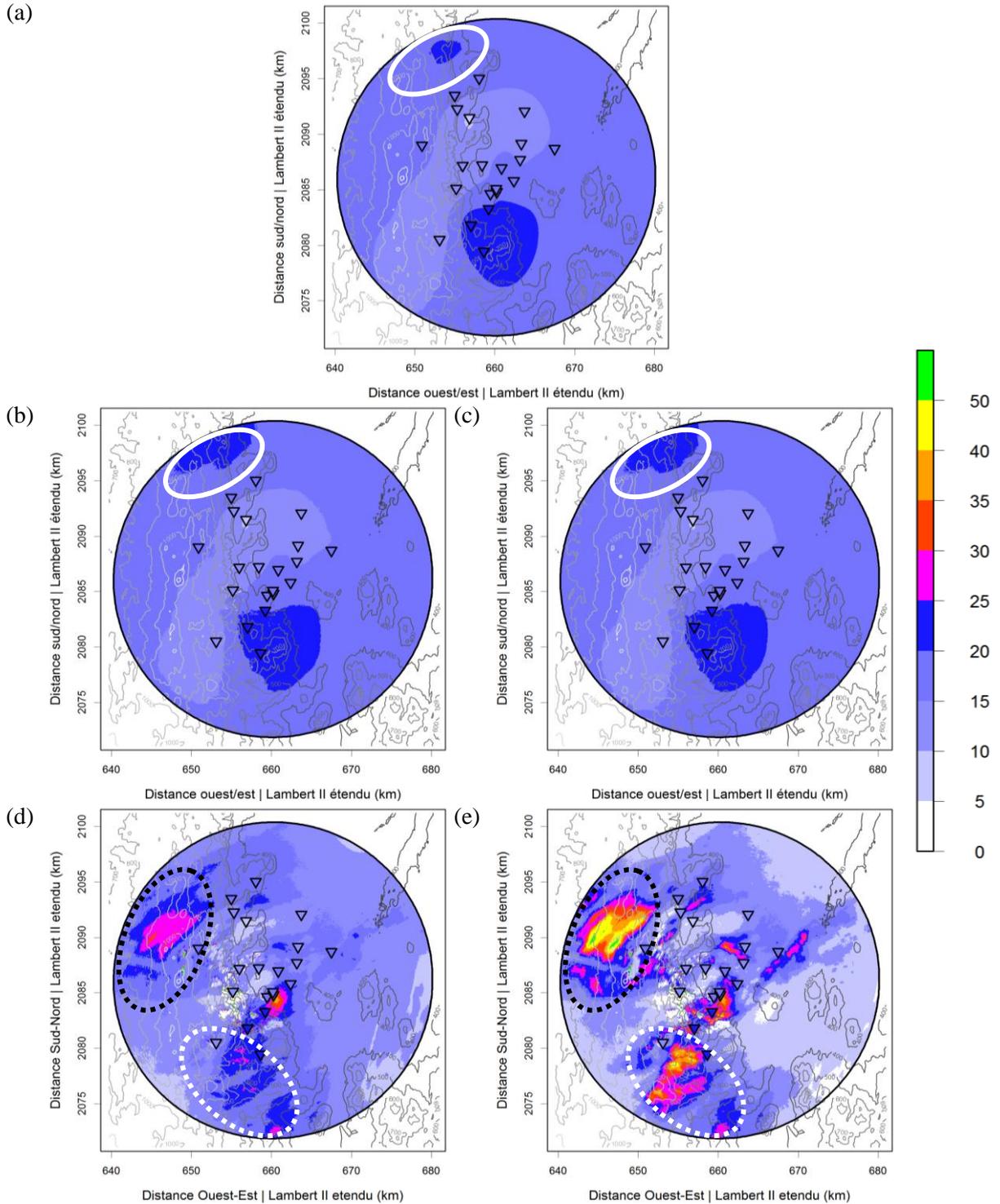


Figure IV-20 : CAS ÉTÉ = Champ de pluie (en mm) interpolé par (a) OK\_RG, (b et c) KED et (d et e) MERG. (T = 8 h).  
 Les figures (c et e) sont issues les mesures RADAR corrigées des effets d'atténuations, contrairement aux figures (b et d).

La Figure IV-19 présente le champ de pluie cumulée sur  $T = 8$  heures observé par les observations pluviométriques et le RADAR. Les précipitations s'organisent selon un axe sud-ouest nord-est, avec une cellule importante au centre du domaine. Les maximums sont enregistrés au sud du réseau de pluviomètres.

Première constatation : toutes les stations du réseau enregistrent de la pluie en quantité équivalente. Le champ RADAR non corrigé des EA (Figure IV-19b) semble s'accorder à ces observations hormis au centre du domaine. Les cumuls en bordure de domaine sont très faibles à nuls (*i.e.* inférieurs à 5 mm en 8 heures). Une cellule en bordure ouest-nord-ouest (*cf.* ellipse) est détectée par le RADAR. La correction des EA (Figure IV-19c) semble intensifier les estimations par le RADAR et étendre spatialement le champ de pluie. Enfin, nous notons la présence de pixels affichant de très fortes valeurs de cumul.

## IV-5.3 Reconstruction d'un champ de précipitation à l'échelle de l'événement

### IV-5.3.1 Validation qualitative

Le champ de pluie reconstruit par la technique du OK\_RG, du KED et du MERG est représenté par la Figure IV-20. Pour ce cas, nous présentons les résultats obtenus avec le variogramme paramétrique défini par le modèle **sphérique**.

L'absence de pluviomètres en limite de domaine fait ici défaut. En effet, l'interpolation par OK\_RG (Figure IV-20a) donne des précipitations sur tout le domaine. Le champ estimé ne présente alors aucune ressemblance avec celui du RADAR. L'intégration des mesures RADAR dans la technique du KED (Figure IV-20 b et c) n'apporte pas d'améliorations particulières. Sur le champ du OK\_RG, nous notons l'apparition d'une petite cellule de pluie en bordure nord-nord-ouest, plus intense pour le KED (*cf.* ellipses blanches). Enfin, la correction des EA n'a ici que peu d'effet sur le champ du KED.

Seule la technique du MERG (Figure IV-20 d et e) permet de retrouver un semblant de caractère spatial des précipitations. Malgré un cumul de pluie surestimé sur tout le domaine, quelques cellules de pluie détectées par le RADAR sont redessinées (*cf.* ellipses en trait pointillé). En bordure du domaine, les cumuls de pluie sont moindres (que ceux du KED par exemple) mais ne correspondent pas encore avec les cumuls très faibles à nuls observés par le RADAR. Contrairement au champ obtenu par KED, nous notons la présence de pixels résiduels constatés sur les mesures RADAR. La correction des EA intensifie les cumuls de toutes les cellules précipitantes mais favorise leur diminution en bordure de domaine.

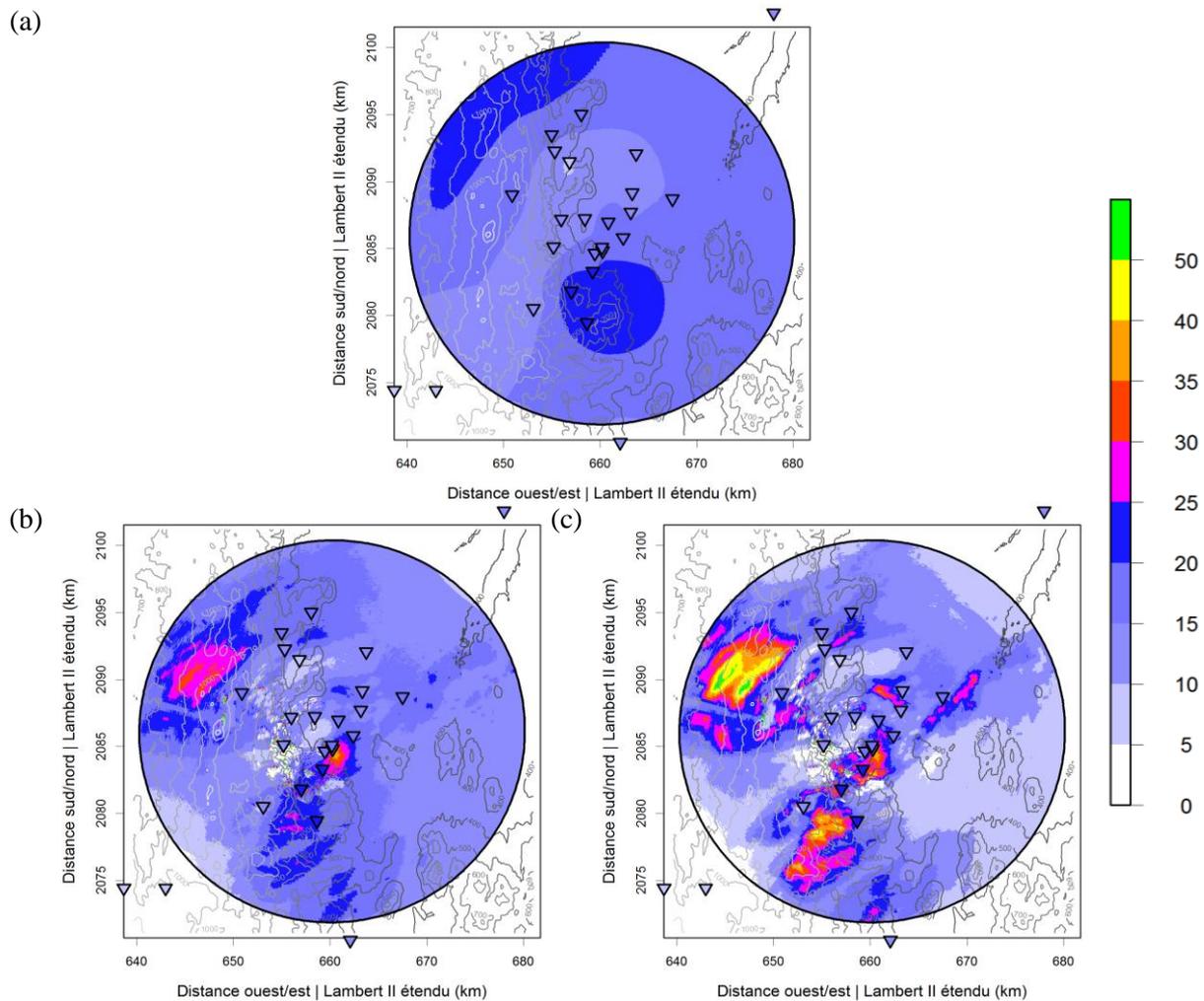


Figure IV-21 : CAS ÉTÉ = Champ de pluie (en mm) interpolé par (a) OK\_RG et (b et c) MERG en intégrant les stations situées hors du domaine d'étude. ( $T = 8h$ ).  
 Les figures (b et c) sont issues des mesures RADAR sans et avec correction des EA. Seules 4 stations cirvoisines (sur 8) apparaissent sur la figure.

La Figure IV-21 montre les modifications apportées lors de la prise en compte des observations situées en dehors du domaine d'observation par la technique du OK\_RG et du MERG. Seulement, ces stations étant également touchées par des précipitations, la minimisation des effets associés aux extrapolations n'a donc pas les résultats escomptés (à savoir la suppression des pluies hors du réseau de pluviomètres).

La variance de krigeage (VK) associée à la technique du krigeage ordinaire ( $\text{Var}_{OK} - \text{OK\_RG}$  et MERG – Figure IV-22 a et b) illustre bien cet atout. En effet,  $\text{Var}_{OK}$  est améliorée aux abords des limites du domaine. En intégrant les stations extérieures, la VK moyenne est réduite de près de 11% ( $\overline{\text{Var}_{OK}} = 56.9$  et  $50.9 \text{ mm}^2$  respectivement).

La grande ressemblance du champ de la VK associé à la technique du KED ( $\text{Var}_{KED}$  – Figure IV-22 c et d) avec  $\text{Var}_{OK}$  (sans prise en compte des données externes), nous a conduit à la représenter sous forme de différence. Les zones de dissemblances nous assurent la prise en compte des mesures RADAR (que nous pouvons douter sur la Figure IV-20). L'influence de cette seconde variable est minimale au regard des valeurs des différences qui n'excèdent guère plus de 5  $\text{mm}^2$ . De plus, nous retrouvons la présence des pixels résiduels mentionnée précédemment (ex : cf. ellipse).

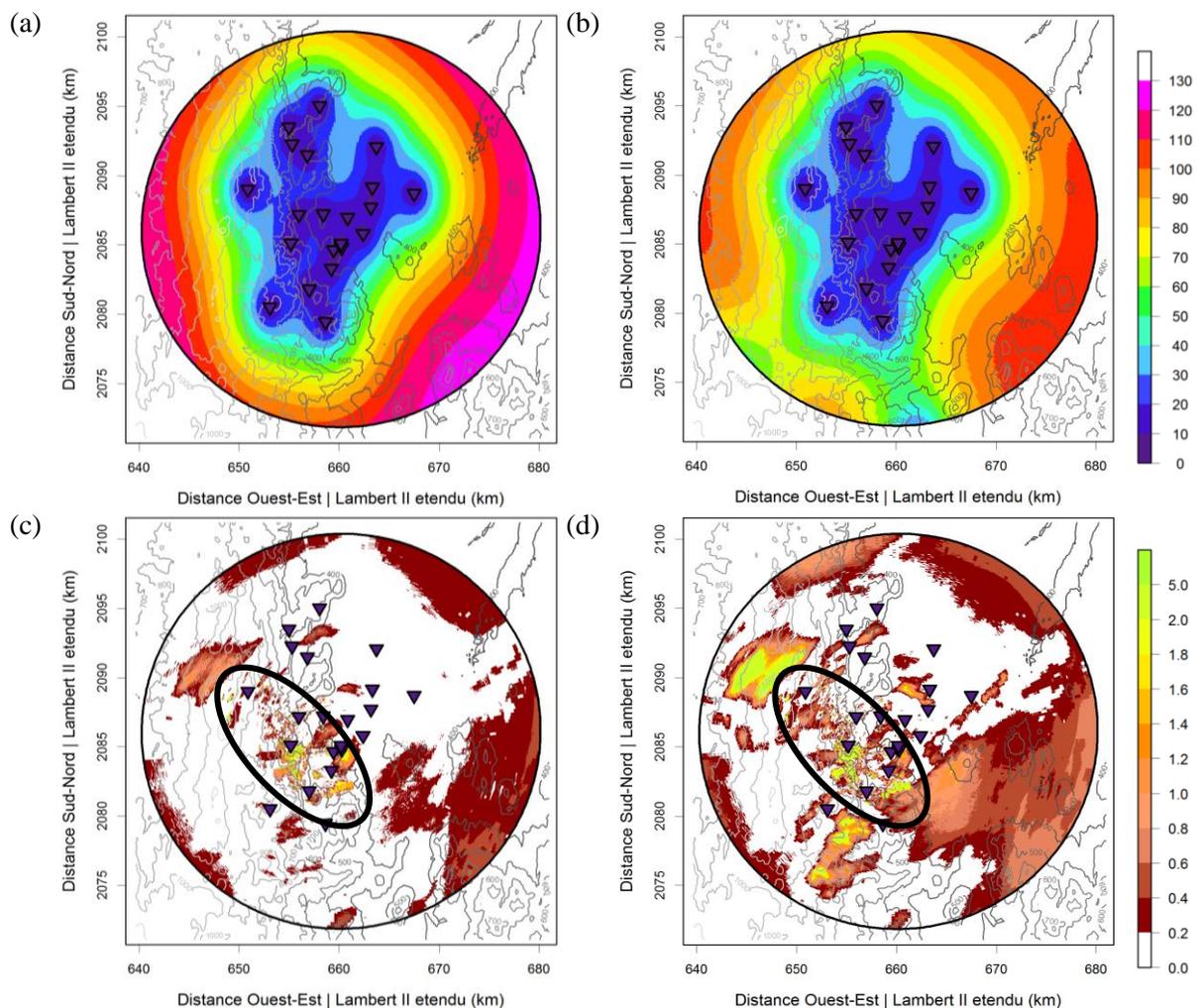


Figure IV-22 : CAS ÉTÉ = (a et b) Champ de la variance de krigeage (en  $\text{mm}^2$ ) associée au krigeage ordinaire sans et avec inclusion des stations extérieures au domaine d'étude ; (c et d) Différence (en  $\text{mm}^2$ ) entre la variance de krigeage associée au KED et celle associée au krigeage ordinaire, sans ou avec correction des EA. ( $T = 8$  h). La variance de krigeage à la position des pluviomètres est reportée dans les triangles pointant vers le bas.

Au travers de ce deuxième cas d'étude (cas ÉTÉ), nous avons voulu tester à l'extrême les performances des deux techniques étudiées (KED et MERG). Qualitativement, aucune ne montre un champ de pluie final réellement convenable. Pour ce cas, la petite surface couverte par le réseau de pluviomètres n'est favorable ni à la technique du OK\_RG ni au KED. Toutefois, nous retrouvons une ébauche de structure spatiale (en lien avec celle observée par le RADAR) grâce à la technique du MERG. De plus, la prise en compte des stations hors du domaine permet de diminuer sensiblement la variance de krigeage en bordure de domaine mais ce comportement apparaît comme minimal aux vues des faibles modifications apportées sur le champ reconstruit par MERG.

Finalement, il apparaît donc que les poids attribués à la variable auxiliaire dans la technique du KED sont plus faibles que pour celle du MERG. Toutefois, nous avons soulevé au travers de ce cas d'ÉTÉ le problème que posent les estimations du RADAR.

### IV-5.3.2 Validation quantitative

Une première comparaison quantitative entre les différents champs de précipitation obtenus est donnée dans le Tableau IV-7. Nous avons reporté la valeur maximale, la proportion de pixels ayant un cumul de pluie supérieur à 5 mm (sur T = 8 heures) et la quantité totale de pluie pour lesquels nous avons distingué les cas avec ou sans la correction des EA sur les mesures RADAR.

Tableau IV-7 : CAS ÉTÉ = Comparatif des différents champs spatiaux. (T = 8 h).  
 Les valeurs des premier et dernier centiles sont indiquées dans la première ligne. La surface du champ de pluie correspond au pourcentage du nombre de pixels ayant une valeur supérieur ou égale à 5 mm. L'astérisque (\*) indique les cas où les stations extérieures au domaine d'étude ont été prises en compte.

	sans correction EA [0.1 - 35.4] mm			avec correction EA [0 - 43.7] mm		
	Valeur maximale	Surface du champ	Quantité tombée	Valeur maximale	Surface du champ	Quantité tombée
RADAR	35.4 mm	62.1 %	5.272 Mt	43.7 mm	66.6 %	8.233 Mt
OK_RG	23.9 mm	100 %	14.530 Mt	23.7 mm	100 %	15.542 Mt
OK_RG*	23.9 mm	100 %	15.022 Mt	23.7 mm	100 %	16.105 Mt
KED	24.6 mm	100 %	14.675 Mt	24.5 mm	100 %	15.689 Mt
MERG	35.3 mm	99.2 %	13.159 Mt	43.7 mm	95.7 %	12.506 Mt
MERG*	35.4 mm	99.2 %	13.650 Mt	43.7 mm	96.1 %	13.070 Mt

Nous notons que la correction des EA intensifie les cellules précipitantes vues par le RADAR tant sur le cumul maximal (+23 %) que la surface (+7%) ou la quantité totale d'eau (+2.9 mégatonnes). Pour les deux cas considérés (équivalent à des intervalles différents de valeurs entre le premier et dernier centile), les pluies estimées par OK\_RG touchent tout le domaine d'étude. Les estimations de la valeur maximale sont certes fortement diminuées par rapport aux mesures RADAR (facteur = 1.5 et 1.8 sans et avec correction des EA) mais les quantités de pluie sont presque doublées (pour l'intervalle [0 - 43.7]) voire triplées (pour l'intervalle [0.1 - 35.4]). L'inclusion des stations situées hors du domaine d'étude (OK\_RG\* – observations également affectées par des précipitations) augmente le cumul total de pluie de près de 500 kilotonnes. La prise en compte d'une plus large gamme de valeurs ([0 - 43.7]) augmente également cette quantité de pluie de plus de 1 mégatonne.

Sur ces trois critères (valeur maximale, surface du champ et quantité tombée), le champ issu par la technique du KED est très proche de celui obtenu par OK\_RG (quel que soit l'intervalle de valeurs considéré). Les principales différences apparaissent uniquement sur la technique du MERG. En effet, en comparaison avec OK\_RG, nous notons une augmentation de l'estimation de la valeur maximale (alors plus proche des mesures RADAR) tout en affichant des structures précipitantes légèrement plus restreintes et plus faibles en intensité. Comme pour OK\_RG, la prise en compte des stations circonvoisines (MERG\*) augmente la quantité totale de près de 500 kilotonnes sans même affecter la superficie des cellules de pluie. La correction des EA augmente de presque 24% la valeur du cumul maximum mais réduit légèrement le champ de pluie en superficie (-3%) et en intensité (-600 kilotonnes).

Seul le MERG n'estime pas de la pluie sur tout le domaine. L'influence des mesures RADAR (affichant une surface de pluie couvrant 70 % du domaine) est donc plus importante que dans la technique du MERG. En comparant le champ du MERG avec celui du OK\_RG, la correction des EA intensifie les cellules précipitantes mais restreint légèrement leur superficie et leur intensité. L'inclusion des stations hors du domaine (OK\_RG\* et MERG\*) ne semble pas être un atout pour ce cas étant donné que toute la région est touchée par cette situation météorologique.

Une seconde étude comparative est donnée dans le Tableau IV-8 qui présente trois scores statistiques résultants de la comparaison des observations au sol (considérées comme valeurs de référence) avec, d'une part la valeur du pixel du champ RADAR le plus proche des

pluviomètres, et d'autre part les estimations par LOO aux points des pluviomètres par OK\_RG, KED et MERG. Nous distinguerons les cas où la correction des EA a été, ou non, appliquée. Enfin, nous avons également ajouté les scores associés aux deux techniques, intégrant les observations extérieures au domaine d'étude (marquées d'un astérisque).

La large valeur de la RMSE sur les estimations RADAR nous indique la présence de larges écarts qui se compensent (petite ME positive). La correction des EA n'améliore ni la ME (-8.3 mm) ni la RMSE (+55 %) par contre, le coefficient de corrélation est meilleur d'un facteur 2.6. Quelle que soit la technique d'interpolation et les options choisies, les mesures RADAR affichent les moins bons scores.

Les estimations par OK\_RG sont sommairement meilleures que les mesures RADAR avec de plus faibles ME et RMSE. Le coefficient de corrélation est tout juste satisfaisant (*i.e.* égale à 0.5%) mais est le meilleur de toutes les estimations faites. La prise en compte des stations extérieures (OK\_RG\*) améliore très légèrement les scores obtenus (*cf.* coefficient de corrélation). Nous noterons que les estimations associées sont les seules à présenter une légère sous-estimation (ME devenue négative).

Concernant les deux techniques combinant les observations, nous constatons que le KED (sans correction des EA) affiche des scores bien supérieurs à ceux du MERG : meilleures ME (-0.6 mm) et RMSE (-64 %) et coefficient de corrélation (plus grand d'un facteur 2.7). Pour le KED, la prise en compte des EA dégrade tous les scores. Pour le MERG, l'intégration des observations situées hors du domaine d'étude améliore très légèrement l'ensemble des scores associés (à l'exception du coefficient de corrélation avec l'utilisation des mesures RADAR non corrigées des EA).

Tableau IV-8 : CAS ÉTÉ = Comparaison entre les observations du réseau principal de pluviométriques et les estimations des différents champs spatiaux. (T = 8 h). L'astérisque (\*) indique les cas où les stations extérieures au domaine d'étude ont été prises en compte. Les indications (-) et (+) indiquent respectivement le moins bon et le meilleur score.

		Erreur moyenne		RMSE		Coef. corrélation		
		ME (mm)		(mm <sup>2</sup> )				
RADAR	sans	avec correction EA	0.8 (-)	-7.5 (-)	9.2 (-)	14.3 (-)	0.124 (-)	0.321
OK_RG			0.1 (+)		3.0		0.504	
OK_RG*			-0.1 mm		3.0 (+)		0.524 (+)	
KED	sans	avec correction EA	-0.1 (+)	0.1	3.1	3.5	0.490	0.356
MERG	sans	avec correction EA	0.5	0.6	8.7	12.6	0.181	0.269 (-)
MERG*	sans	avec correction EA	0.3	0.4	8.6	12.5	0.135	0.273

Finalement, les résultats obtenus pour la technique du KED et du MERG sont ambigus : la technique du MERG semble visuellement meilleure pourtant, c'est la technique du KED qui présente les meilleurs résultats statistiquement parlant même si la validation reste tout juste convaincante (coefficient de corrélation inférieur à 0.5% dans le meilleur cas). Pour le MERG, la prise en compte des stations extérieures au domaine apportent une amélioration minime sur les scores qui en résultent.

## IV-6 Conclusions

L'étude quantitative des événements précipitants touchant l'agglomération clermontoise que nous proposons ici est une démarche nouvelle pour le LaMP. L'utilisation des techniques géostatistiques (OK\_RG, MERG ou KED) a donc nécessité en premier lieu la détermination d'un variogramme paramétrique qui représente au mieux les données du bassin clermontois. Pour ce faire, nous avons opté pour une approche climatologique qui nous permet de le définir une seule et unique fois. Les résultats obtenus nous ont conduit vers deux modèles mathématiques (**sphérique** et **exponentiel**) dont les paramètres (portée, effet de pépite et palier) affichent un lien certain avec la fenêtre temporelle.

Dès lors, nous avons pu appliquer les deux techniques combinant les observations à travers deux cas d'étude. *Contrairement au domaine d'étude des Cévennes-Vivarais*, le réseau d'observation est plus petit spatialement et surtout n'est pas entièrement couvert par le réseau de pluviomètres. Au travers de ce chapitre, nous avons donc testé à l'extrême la technique du KED et du MERG.

Pour les deux événements, nos études préliminaires ont montré quelques problèmes avec les mesures RADAR : atténuation du signal, mauvaise synchronisation des données, présence de pixels résiduels (assimilés à des échos de sol) qui affichent des cumuls très forts, etc. La validation quantitative des techniques de reconstruction spatiale de la pluie s'est donc limitée à l'échelle de l'événement et pour un nombre défini de localisation (*i.e.* pixels compris entre le premier et le dernier centile)

Pour le cas HIVER, nous avons mis en exergue le caractère prometteur des deux techniques avec la retranscription du caractère hétérogène des structures de pluie observé par le RADAR et le réajustement des quantités de pluie pour s'accorder aux observations pluviométriques. La petitesse de la zone couverte par le réseau principal occasionne hélas des estimations en bordure de domaine faussées par le manque de données. Ce problème peut être contourné par l'utilisation notamment de la technique du MERG qui permet d'intégrer les

observations situées hors du domaine. Au final, le champ reconstruit est visuellement amélioré et la variance de krigeage associée est considérablement réduite en dehors du réseau de pluviomètres. La correction des effets d'atténuation sur les mesures RADAR avait tendance à intensifier les cellules de pluie et à étaler spatialement les précipitations (*i.e.* une surface plus grande).

Pour le cas ÉTÉ, seule la technique du MERG semble reproduire une structure de pluie qui s'apparente quelque peu avec celle mesurée par le RADAR. La particularité de ce cas d'étude a permis de montrer que la technique du KED accorde moins d'importance à la seconde variable. En effet, le champ final n'affiche aucune ressemblance avec le champ RADAR alors que le champ de variance associé nous assure cette prise en compte. Enfin, comme tout le département a été touché par le système précipitant, l'intégration des observations extérieures au domaine pour la méthode du MERG n'a eu d'influence que sur la variance de krigeage.

Dans le chapitre précédent, nous avons conclu que la technique du KED et du MERG permettaient de considérer le caractère spatial observé sur les mesures RADAR en faisant abstraction de la qualité des estimations quantitatives. Avec les mesures RADAR du bassin clermontois, nous avons été confrontés à un problème d'un autre ordre que nous n'avons pas cherché à déterminer ni à corriger (car ce n'était pas le sujet de cette thèse). N'ayant pas d'autres données, nous avons toutefois fait le choix d'exploiter celles-ci en tout état de connaissance.

Au travers des applications réalisées dans la région Cévennes-Vivarais, nous avons montré les bonnes performances de la technique du MERG notamment. Nous espérons que les travaux menés au LaMP conduiront peut être un jour à montrer l'apport des techniques combinant les observations pluviométriques et mesures RADAR pour des petits domaines tels que celui du bassin clermontois.



# Chapitre V

## Modélisation de la pluie avec WRF

Les précipitations sont le résultat de nombreux processus météorologiques complexes et extrêmement variables. Pour être le plus réaliste possible, ces phénomènes devraient pouvoir être étudiés sur de larges gammes d'échelles spatiale et temporelle. Il existe actuellement plusieurs modèles météorologiques capables de reproduire les pluies. Quasiment tous sont construits sur la même base : un module dynamique qui permet d'obtenir l'évolution spatiale des paramètres atmosphériques (*i.e.* vent, pression, température et humidité) et un module microphysique qui représente l'évolution temporelle des variables décrivant les nuages et les précipitations. Pour faciliter la résolution des équations, des simplifications sur les comportements dynamique et physique sont nécessaires. Ces compromis sont inévitables puisqu'ils conditionnent le temps de calcul des prévisions.

Dans ce chapitre, nous reviendrons tout d'abord sur l'origine des pluies (succinctement évoquée au Chapitre I). Cette introduction démontrera alors combien il est difficile de modéliser la totalité des processus responsables de la formation des précipitations. Les deux paragraphes suivant s'axeront sur la description généralisée des modèles météorologiques en mettant l'accent sur les modules de microphysique. Enfin, nous détaillerons les spécificités du modèle avec lequel nous avons été amenés à travailler et nous exposerons ses performances sur la prévision des pluies dans le cadre d'une étude réalisée au travers d'une simulation idéalisée.

### V-1 Origine de la pluie – partie II

Dans le paragraphe I-1.1, nous avons brièvement décrit l'origine de la pluie d'un point de vue dynamique. Dans cette partie, nous allons davantage zoomer à l'intérieur d'un nuage et s'intéresser aux processus responsables de la formation des précipitations à l'échelle microscopique (Figure V-1).

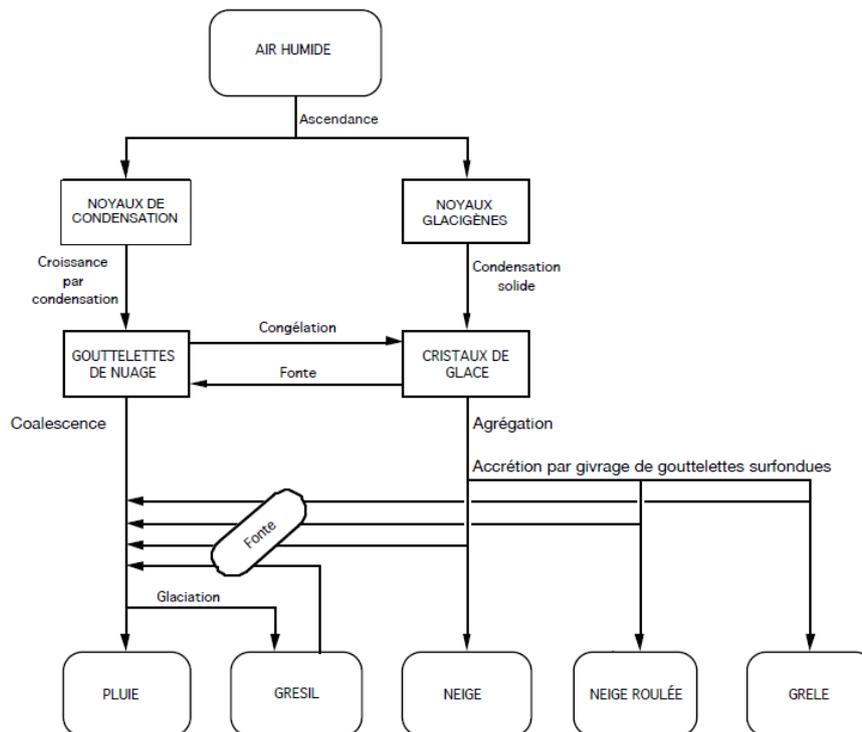


Figure V-1 : Principaux processus responsables de la formation de la pluie.

Source : <http://pascal.sca.uqam.ca/~sca2626/Presentations/06-Precipitation51.pdf> (modifiée).

Un nuage se forme lorsque l'air excède sa capacité à contenir l'eau sous forme gazeuse. De taille, d'origine et de composition chimique diverses, des particules d'aérosols<sup>⊕</sup> jouent le rôle de noyaux de condensation de nuage (ou CCN pour l'acronyme anglais « Cloud Condensation Nucleii ») sur lesquels l'excès de vapeur d'eau se condense.

Au cœur des nuages dits « chauds » ( $T > 0^{\circ}\text{C}$ ), les gouttelettes nuageuses croissent principalement par diffusion de la vapeur d'eau (processus de condensation). La collision avec d'autres gouttelettes crée des particules plus grosses (processus de coalescence). À des températures bien inférieures à  $0^{\circ}\text{C}$ , des gouttelettes d'eau surfondue (*i.e.* encore à état liquide) peuvent coexister avec les particules formant ainsi des nuages dits « mixtes ». La collision entre une goutte liquide et un cristal de glace peut être à l'origine des cristaux plus grande taille (processus de givrage). Le phénomène de congélation instantanée des gouttes liquides ou des aérosols humides se produit dans l'intervalle de températures de  $-5$  à  $-28^{\circ}\text{C}$  ;

<sup>⊕</sup> Ensemble de particules solides ou liquides, d'origine naturelle (érosion des sols, embruns marin, cendre volcaniques, etc.) ou anthropique (usines, chauffages, catalyseurs sur véhicules légers, etc.) en suspension dans un milieu gazeux. Ces petites molécules ont des propriétés physiques telles qu'elles autorisent le dépôt d'eau liquide sur leur surface.

cette étape est appelée « nucléation hétérogène ». Pour des températures inférieures, la formation instantanée des cristaux à partir des noyaux glacogènes (IN pour l'acronyme anglais « Ice-Nucleii ») est également possible ; cette étape est appelée « nucléation homogène ». Dans ces nuages dits « froids » (ou de glace), les cristaux peuvent croître par diffusion de vapeur d'eau (processus de déposition) et collision entre eux (processus d'agrégation). Notons que l'ensemble de ces processus de croissance est fortement dépendant des conditions intra-nuageuses, mais également de la concentration en CCN et IN.

Lorsque les hydrométéores atteignent une certaine taille (ou plus exactement une masse), ces particules tombent par gravité. En franchissant l'isotherme 0°C, les cristaux de glace fondent progressivement. Les gouttes d'eau qui en résultent peuvent continuer de grossir par collision avec d'autres gouttes ou bien s'éclater par contrainte dynamique (processus de « break-up ») pour former des gouttes plus petites (voire suffisamment petites pour être à nouveau entraînées dans les flux nuageux). Les plus petites gouttes d'eau pourront s'évaporer avant de toucher le sol.

La nature et les hétérogénéités (spatiale et temporelle) des nuages font des précipitations une grandeur physique difficile à prévoir. Malgré l'amélioration des connaissances et l'avancée des puissances de calculs, les scientifiques cherchent encore à décrire la pluie de la façon la plus réaliste possible tant les enjeux liés à leur prévision sont considérables.

## V-2 Les modèles météorologiques

### V-2.1 Généralités

Les modèles météorologiques sont des programmes informatiques qui, au travers de millions de lignes de codes, calculent une prévision de l'état de l'atmosphère à partir d'un état initial. Le concept consiste en une résolution d'un jeu d'équations décrivant les processus atmosphériques sur un maillage tridimensionnel de l'atmosphère.

La description des principaux phénomènes météorologiques (Figure V-2) nécessite de considérer plusieurs échelles spatiales (de plusieurs centimètres à quelques dizaines de milliers de kilomètres) et temporelles (de quelques secondes à plusieurs années). À l'heure actuelle, aucune machine de calcul ne peut intégrer l'ensemble de ces échelles, les résolutions d'un modèle doivent donc être définies en adéquation avec le phénomène étudié (*i.e.* étendue spatiale, cycle de vie, etc.). Il existe alors différents types de modèles numériques :

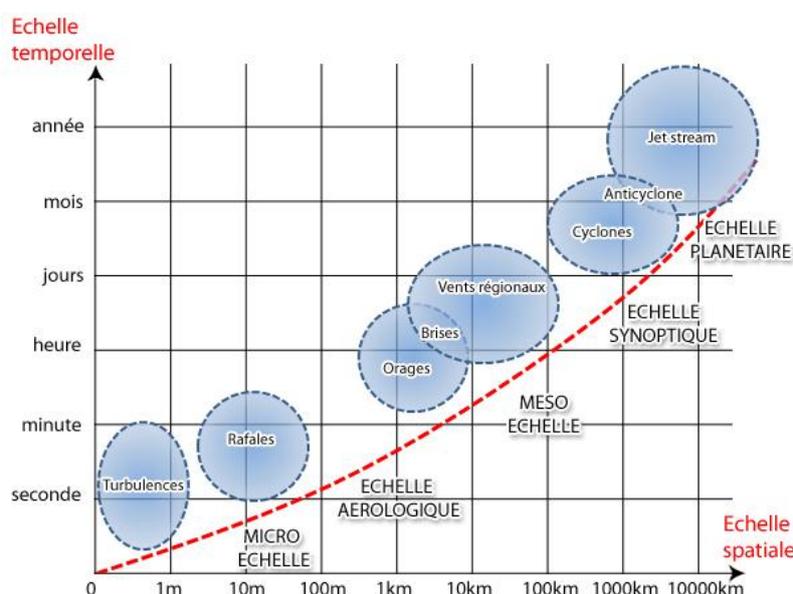


Figure V-2 : Échelles spatio-temporelles considérées en modélisation.  
 Source : Planche (2011, p. 24) d'après Orlanski (1975).

- les modèles globaux qui simulent les phénomènes à l'échelle de la planète ;
- les modèles de prévision qui se focalisent généralement sur l'échelle synoptique et qui se nourrissent parfois d'observations aérologiques grâce à des techniques d'assimilations ;
- les modèles à méso-échelles qui étudient les nuages et l'atmosphère à une échelle régionale qui sont alimentés par les sorties des modèles synoptiques.

Toute prévision commence par la détermination aussi précise que possible de l'état initial de l'atmosphère à un instant donné. Cette étape est rendue possible grâce aux observations effectuées en permanence et dans le monde entier. Chaque point de grille du modèle reçoit une valeur (observée ou adaptée par assimilation) pour chaque paramètre atmosphérique.

Tous les modèles numériques se basent sur l'évolution de quatre paramètres : le vent, la pression, la température et l'humidité (ou teneur de l'air en vapeur d'eau). Les interactions entre ces paramètres sont décrites mathématiquement par les équations i) du bilan de la quantité de mouvement, ii) de la conservation de la masse, iii) du bilan de l'énergie thermodynamique et iv) de la conservation de l'eau. Cette dernière équation est un élément essentiel dans la prévision des nuages et des précipitations éventuellement associées.

L'ensemble de ces interactions font intervenir des phénomènes de très petites échelles qui ne peuvent pas être décrits explicitement. Pour faciliter leurs prises en compte au sein des modèles, il est commun d'avoir recours à des équations paramétrisées. Les paramétrisations de ces phénomènes peuvent être regroupées comme suit :

- les schémas de rayonnement qui décrivent les processus de réchauffement de l'atmosphère *via* les différents flux radiatifs à courtes et longues longueurs d'ondes ;
- les schémas de surface qui récupèrent les variables d'état et caractéristiques associées au terrain considéré (terre, eau, glace de mer, etc.) pour prédire les flux de chaleur et d'humidité en chaque point de la surface du sol ;
- les schémas de couche limite qui fournissent les tendances de température, de l'humidité et de la quantité de mouvement (transports turbulents) au sein de la couche limite planétaire ;
- les schémas de convection qui renseignent sur les profils verticaux de chaleur et d'humidité générés par les mouvements (verticaux et horizontaux) qui sévissent au cœur des nuages (module uniquement activé pour des simulations de très grandes échelles) ;
- les schémas de microphysique qui recensent l'ensemble des processus à l'origine des nuages et des précipitations.

L'intérêt des travaux présentés dans ce chapitre se porte uniquement sur la prévision de la pluie. C'est pourquoi nous ne développerons que le dernier groupe de paramétrisations. Notons que dans certains modèles la résolution des phénomènes microphysiques peut être explicite (donc non paramétrisée).

## V-2.2 Modélisation des nuages et des précipitations

La structure micro-échelle d'un nuage (caractérisée par la concentration en nombre des particules CCN ou IN, de leur composition physico-chimique et de leur dimension) influence directement sa capacité à produire des précipitations (pluie, neige, grêle, etc.) et sa durée de vie (Planche *et al.* 2010). Près du sol, l'humidité est difficilement prévisible de par les fortes interactions avec la surface qui compliquent la prévision du brouillard et autres nuages bas. Outre les processus complexes de la formation des nuages et des précipitations, de nombreuses interactions avec l'environnement sont donc à considérées.

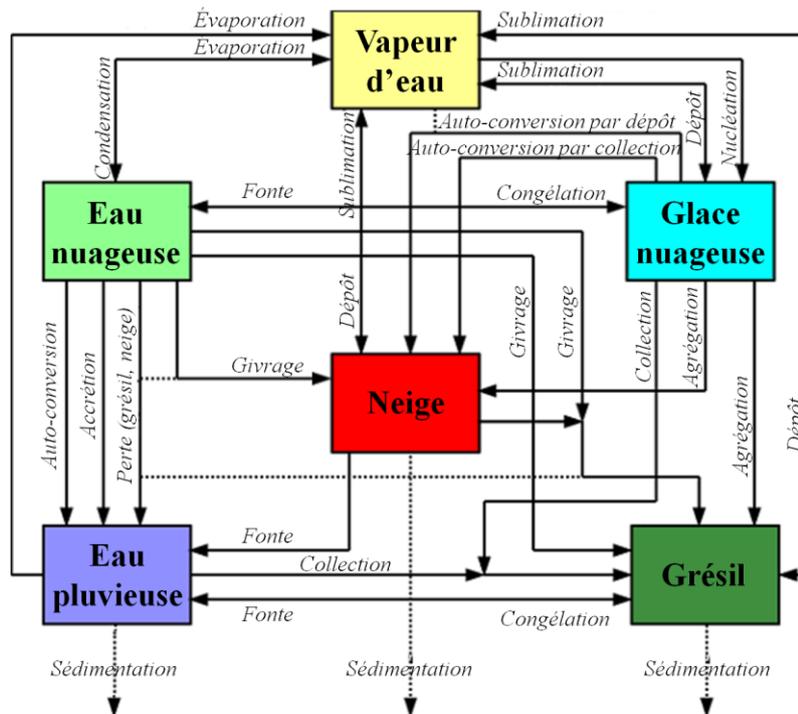


Figure V-3 : Exemple d'interactions entre 6 catégories d'hydrométéores.  
 Source : <http://www.dwd.de/> [section : grid-scale precipitation] (traduite).

Les schémas de microphysiques permettent la résolution des processus d'échanges de masse entre les différents réservoirs de l'eau atmosphérique (*i.e.* vapeur, eau liquide et glace). Selon le schéma utilisé, les réservoirs « liquide » et « solide » sont partitionnés en plusieurs catégories. Leur nombre varie entre deux (eau nuageuse et pluie) et cinq (ex : eau nuageuse, pluie, neige, glace nuageuse et grésil – Figure V-3). Les schémas microphysiques sont classés en deux classes : les modèles à un ou plusieurs moments (dits « bulk ») et les modèles à microphysique détaillée ou spectrale (dits « bin »).

Dans les modèles « bulk » à un moment, les propriétés des nuages sont décrites par la masse totale de chaque hydrométéore considéré (Kessler 1969; Hong *et al.* 2004). Les modèles à deux moments intègrent également le nombre de particules de chaque catégorie (Thompson *et al.* 2008; Morrison *et al.* 2009). Dans les deux cas, la distribution en taille des hydrométéores est fixée selon une fonction prédéterminée définie par des paramètres de forme (exponentielle, log-normale ou gamma). Les processus microphysiques assimilés, fortement dépendant de la taille des hydrométéores, représentent alors le comportement moyen de la masse (et du nombre de particule pour les schémas à deux moments) de chaque catégorie. Ces hypothèses permettent de diminuer considérablement le nombre de variables pronostiques à une (pour les schémas à 1 moment) ou deux (pour les schémas à 2 moments).

Pour considérer le développement des tailles individuelles de chaque particule, l'utilisation des modèles « bin » est à préférer. Ce type de modèle pronostique de nombreuses variables associées aux distributions dimensionnelles de chacun des hydrométéores considérés (*i.e.* nombre en fonction de la taille). Le nombre de variables utilisées pour représenter la microphysique est étroitement lié au nombre de catégories (ou « bin » en anglais) intégrant les spectres des hydrométéores. Ces modèles sont capables de décrire tous les processus microphysiques en fonction de la taille des hydrométéores (Khain *et al.* 2004; Leroy *et al.* 2009). En contre partie, ces modèles sont extrêmement coûteux en mémoire et temps de calcul, ils ne peuvent donc pas encore être utilisés pour la prévision opérationnelle.

## V-3 Le modèle à méso-échelle WRF

Dans le cadre de nos recherches, nous avons été amenés à travailler avec le modèle météorologique « Weather Research et Forecasting » (WRF) que nous allons décrire dans ce qui suit.

### V-3.1 Généralités

Le développement du modèle WRF est le fruit d'un effort commun de diverses agences et centres de recherches universitaires à travers le monde. Leur objectif : construire la nouvelle génération de modèle de prévision à méso-échelle totalement flexible. Depuis son apparition en 2000, WRF est mis à jour presque tous les 6 mois. Les applications de ce modèle s'échelonnent depuis l'étude des tourbillons jusqu'aux simulations globales, sur Terre comme sur une autre planète.

Les prévisions s'appuient sur la formulation non-hydrostatique des équations de la mécanique des fluides (en utilisant une approche eulérienne) et de la thermodynamique. Le modèle propose deux noyaux dynamiques. Celui que nous avons choisi d'utiliser est le noyau ARW (pour l'acronyme anglais : « Advanced Research WRF ») développé par le National Center for Atmospheric Research (USA).

Les paramétrisations du modèle physique sont codées de façon modulaire, où tous les schémas interagissent entre eux (Figure V-4). Le premier schéma appelé est le schéma de rayonnement, suivi du schéma de surface qui requiert la valeur des flux radiatifs calculés. Les flux de chaleur et d'humidité déduits, le schéma de couche limite planétaire s'insère ensuite. En dernier lieu arrivent les schémas de microphysique et de convection (si activé).

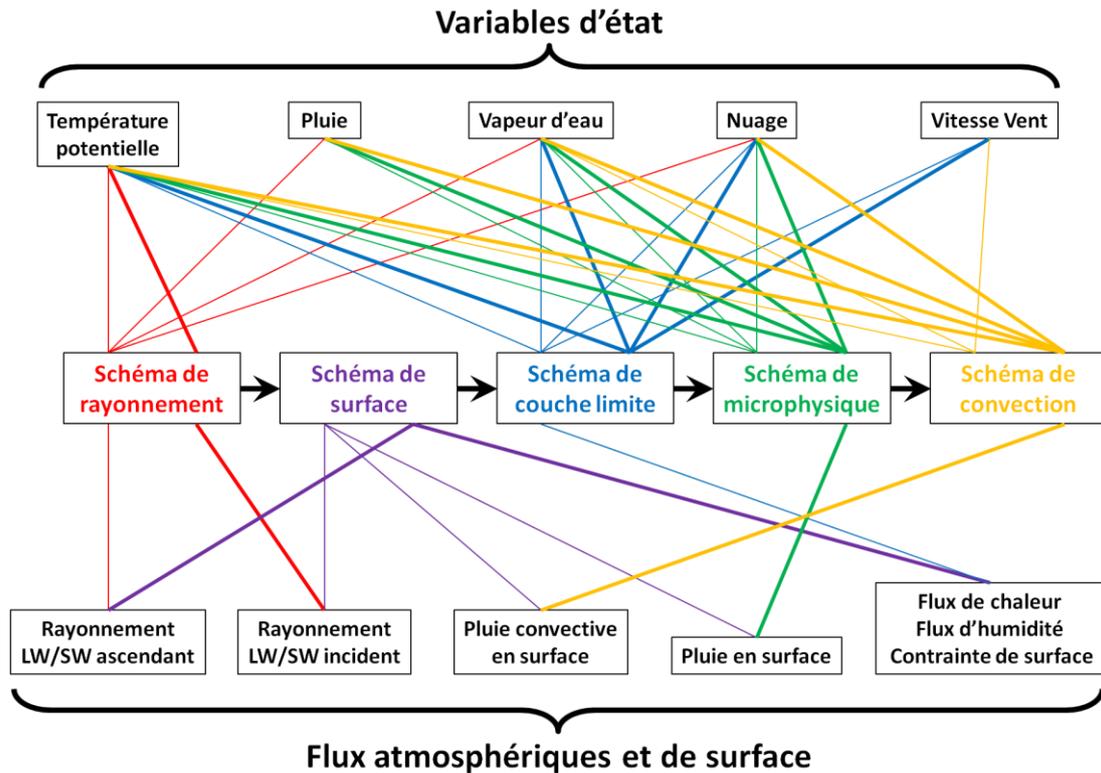


Figure V-4 : Interactions entre les différentes paramétrisations de WRF (centre) liants les variables d'état (haut) et les flux atmosphériques et de surface (bas). Les traits fins (gras) représentent les variables d'entrée (de sortie). Acronymes : LW/SW = longue/courte longueur d'onde. Figure réalisée à partir du tableau 8.6 de Skamarock et al. (2008, p. 81).

La philosophie des créateurs du modèle WRF-ARW est de permettre l'étude de tous les phénomènes météorologiques, du plus simple au plus complexe, du plus grand au plus petit. Nous présentons ici une liste non exhaustive des différentes options proposées par le modèle dans sa version 3.2.1 (Skamarock et al. 2008; Wang et al. 2011).

### V-3.1.1 Module dynamique (ARW)

Les évolutions des variables thermodynamiques sont principalement définies par :

- le système de coordonnées employé sur l'horizontal (Arakawa-C<sup>⊕</sup> étendue) et la verticale (niveaux variables définis selon la pression, suivant le terrain et dont le dernier niveau est une surface à valeur constante), 4 projections disponibles (polaire stéréographique, Lambert-conforme, Mercator, Latitude/Longitude) ;

<sup>⊕</sup> Les variables thermodynamiques (pression, température et humidité) sont calculées au centre des mailles, tandis que les composantes du vent sont données sur les 2 faces des mailles selon la direction correspondante et donc décentrées d'un demi-point de grille (et de dimension +1).

- les conditions initiales fixées par l'intégration de données réelles/idéales ou analyses (sur une à trois dimensions spatiales), les conditions aux limites (périodiques<sup>⊕</sup> ou ouvertes pour un cas idéalisé, ou encore spécifique avec une zone de relaxation newtonienne pour un cas réel) et la communication entre les domaines définis (descente d'échelle de type « 2-way nesting »<sup>⊕⊕</sup> par défaut) ;
- les approximations faites sur les équations (compressibles, eulériennes, non-hydrostatiques, conservation des variables scalaires pronostiques, inclusion de tous les termes de Coriolis et des effets de rotondité de la Terre) ;
- le type d'intégration sur le temps (Runge-Kutta de 2<sup>ième</sup> ou 3<sup>ième</sup> ordre avec possibilité de désintégration du pas de temps pour les ondes acoustiques et ondes de gravité) et l'espace (discrétisation des termes de diffusion et d'advection de 2<sup>ième</sup> à 6<sup>ième</sup> ordre).

### V-3.1.2 Module physique

Les évolutions des variables décrivant les nuages et les précipitations sont caractérisées par :

- le schéma de rayonnement de longues (6 schémas disponibles) et courtes (10 schémas) longueurs d'onde qui incluent ou non les effets des nuages et des flux de surface ;
- le schéma de surface qui permet la résolution des flux turbulents (8 schémas) et des interactions atmosphère/surface (10 possibilités, du simple modèle thermique aux modèles plus complets qui considèrent les effets liés à la végétation, l'humidité du sol, la présence de neige et glace de mer ou encore l'urbanisation) ;
- le schéma de couche limite planétaire (10 schémas) nécessaire pour la prédiction de l'énergie cinétique turbulente ;
- le schéma de microphysique (11 schémas de type « bulk » à 1 ou 2 moments qui prennent en charge ou non les phases glace et mixte, si activé) ;
- le schéma de convection (12 schémas, si activé) basé sur l'ajustement de l'humidité sur la colonne d'air et la description complète des flux de masse (incluant la convection profonde et/ou peu profonde et les vitesses de transports).

<sup>⊕</sup> Les conditions limites périodiques correspondent à une symétrie « cyclique » (termes parfois employé par abus de langage). Ces conditions suppriment les frontières artificielles de l'écoulement ce qui se traduit par une continuité des flux. Schématiquement, tout flux sortant du domaine de simulation par une face donnée est réintroduit par l'extrémité opposée. Généralement, une condition de périodicité est imposée pour les limites du domaine parallèles à l'écoulement moyen où l'on ne s'attend pas à trouver des effets de bords dominants.

<sup>⊕⊕</sup> Contrairement aux modèles globaux, les modèles régionaux permettent de focaliser la simulation sur un espace géographique limité *via* un raffinement horizontal (ou « nesting » en anglais). Cette descente d'échelle se fait par emboîtements successifs de domaines dont la résolution spatiale horizontale est à chaque fois plus fine. Ce processus s'accompagne d'échanges d'information à chaque échéance soit uniquement du domaine à maille large vers le domaine à maille plus fine (« 1-way nesting »), soit dans les deux sens (« 2-way nesting »).

Avec le temps, le nombre de contributeurs augmente, tout comme le nombre de paramétrisations. Il devient alors parfois difficile de faire un choix. Dans la partie suivante nous nous proposons de tester les performances du modèle WRF dans la simulation des pluies. Nous nous focaliserons sur une situation idéalisée dont les processus générant les précipitations sont bien connus. Nous porterons alors particulièrement attention au choix du schéma de microphysique.

## **V-4 Simulation idéalisée de la pluie avec le modèle WRF**

Ce dernier paragraphe est décomposé en trois paragraphes. La première revient sur la description du phénomène que nous avons cherché à simuler. Le second paragraphe décrira les conditions initiales mises en place pour réaliser nos simulations réalisées avec le modèle WRF. Enfin, nous comparerons les résultats obtenus avec les différents schémas de microphysiques disponibles.

### **V-4.1 Contexte de l'étude : les pluies orographiques**

En région montagneuse, les flux atmosphériques sont fortement perturbés. La présence du relief joue un rôle considérable dans le cycle de vie des phénomènes précipitants. Leurs caractéristiques dépendent principalement de la barrière topographique rencontrée (*i.e.* de l'altitude, de la pente et de l'orientation de la ou les montagnes).

L'image classique des précipitations dites « orographiques » se dessine par une chaîne de montagne dont l'axe principal est orienté perpendiculairement à la direction du vent dominant. Les pluies se déversent alors principalement sur les versants (exposés) au vent, résultant en une délimitation marquée du climat, de la flore et de la faune par rapport aux versants sous le vent (Roe 2011).

Il existe différents types de mécanismes liés à ces précipitations (Roe 2011) :

- cas d'une atmosphère stable (Figure V-5a) : l'ascension forcée de la masse d'air implique un refroidissement de la parcelle d'air. Les processus de condensation associés conduisent à d'importantes précipitations sur les versants au vent et donc une diminution du rapport de mélange de la vapeur d'eau en aval de la montagne. Après avoir franchi le sommet, la descente sur les versants sous le vent s'accompagne par un réchauffement de la masse d'air ;

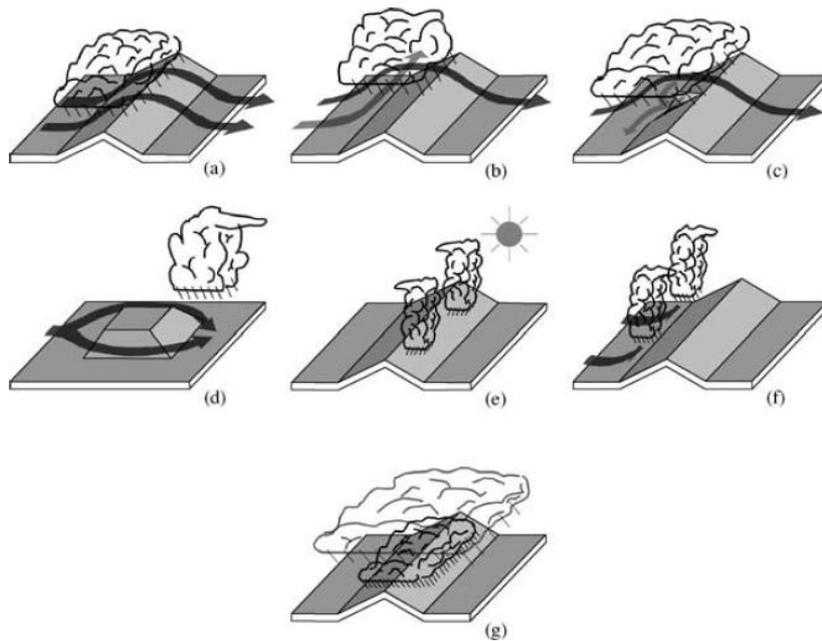


Figure V-5 : Illustration schématique des différents mécanismes à l'origine des précipitations orographiques.

(a) perturbation stable en amont, (b) blocage partiel, (c) retour de flux en vallée induit par le refroidissement par évaporation, (d) convergence sous le vent, (e) convection déclenchée par le rayonnement solaire, (f) convection entraînée par l'élévation mécanique au dessus du niveau de convection libre et (g) mécanisme de « seeder-feeder ». Source : Roe (2011).

- cas d'une atmosphère extrêmement stable ou d'un faible champ de vent : ne pouvant pas s'élever, la masse d'air est bloquée au pied de l'obstacle. Plusieurs scénarios sont alors possibles :
  - le flux est dévié le long des versants au vent. L'arrière de la masse d'air peut alors s'élever au dessus de l'avant de la masse d'air qui joue le rôle de tremplin. Ce phénomène peut provoquer d'importantes précipitations sur les versants au vent (Figure V-5b) ;
  - sous un nuage déjà précipitant, les processus de fonte et/ou d'évaporation des précipitations refroidissent la masse d'air pouvant entraîner un fort flux de retour vers la vallée (Figure V-5c) ;
  - si sa forme est propice, l'obstacle peut être contourné et la convergence des flux en aval favorise la formation de pluies sur les versants sous le vent (Figure V-5d).

- cas d'une atmosphère conditionnellement instable (*i.e.* convective) : en s'élevant, la masse d'air peut passer le niveau de convection libre (la parcelle d'air devient moins dense que son environnement et continue donc de s'élever). Des cellules convectives isolées peuvent ainsi se former en amont de l'obstacle (Figure V-5f). Les phénomènes de réchauffement des versants exposés au soleil responsables de la convection dans les orages d'été sont également répertoriés dans cette catégorie (Figure V-5e).
- le mécanisme de « seeder-feeder » (proposé par Bergeron (1965) – Figure V-5g) : ce processus permet d'expliquer l'intensification des précipitations au dessus de petites collines et pour lesquelles le flux d'air est trop rapide pour que les hydrométéores se forment uniquement par coalescence/collision des gouttelettes nuageuses (= auto-conversion). Néanmoins, en présence de gouttes de pluie ou flocons de neige présents dans les couches supérieures, le processus d'accrétion est plus efficace pour générer de la pluie et surtout plus rapide. Ainsi, si l'apport d'hydrométéores par un nuage de plus haute altitude (dit « seeder ») est suffisant, il peut permettre la formation de précipitations par coalescence ou givrage au sein d'un nuage généré par le passage de la masse d'air par dessus une colline (dit « feeder »).

L'étude que nous nous proposons de réaliser repose sur la simulation de ce dernier mécanisme particulier. Notre objectif est de tester l'ensemble des schémas de microphysiques disponible dans le modèle WRF-ARW (v.3.2.1) dans l'espoir de trouver celui qui permet de reproduire le plus réalistement les précipitations. Avant d'analyser les résultats obtenus, nous allons décrire les conditions météorologiques utilisées pour l'initialisation du modèle.

## V-4.2 Initialisation du modèle

La situation idéalisée étudiée est très proche de celle utilisée par Zängl *et al.* (2010). Les simulations ont été réalisées sur un unique domaine de 201×101 points de grille, avec une résolution horizontalement de 750 m. La grille verticale comprend 51 points non-équidistants (Figure V-6) dont les niveaux varient entre plus de 110 m près de la surface à près de 700 m au sommet du domaine (= 20 km) tout en épousant le relief.

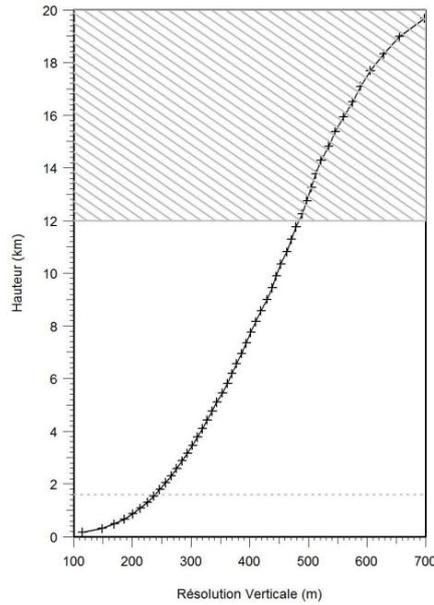


Figure V-6 : Variation de la résolution verticale entre le sol et le sommet du domaine. La zone hachurée représente la couche de Rayleigh de 8 km appliquée de façon à prévenir des réflexions non-physiques de la propagation des ondes de gravité sur la verticale. La hauteur maximale du relief considéré est donnée par la ligne grise en tirets.

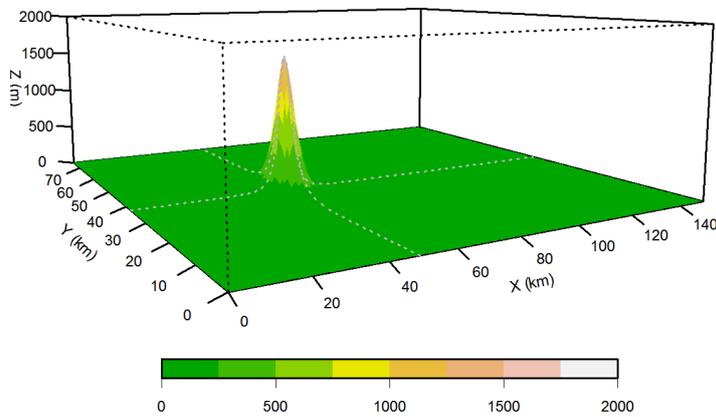


Figure V-7 : Représentation de la topographie de la situation d'étude. Les lignes grises en pointillé indiquent la position de la hauteur maximum du relief considéré (i.e.  $x \approx 49$  km,  $y \approx 38$  km).

Au premier tiers du domaine se dessine une montagne isolée (Figure V-7) dont la forme  $h_m$  est définie par :

$$h_m(x, y) = \frac{h_0}{\left(1 + \left(\frac{x}{L_x}\right)^2 + \left(\frac{y}{L_y}\right)^2\right)^2}$$

Son sommet atteint  $h_0 = 1\,600$  m d'altitude et sa largeur est donnée par  $L_x = 4$  km et  $L_y = 1.25 \cdot L_x$ .

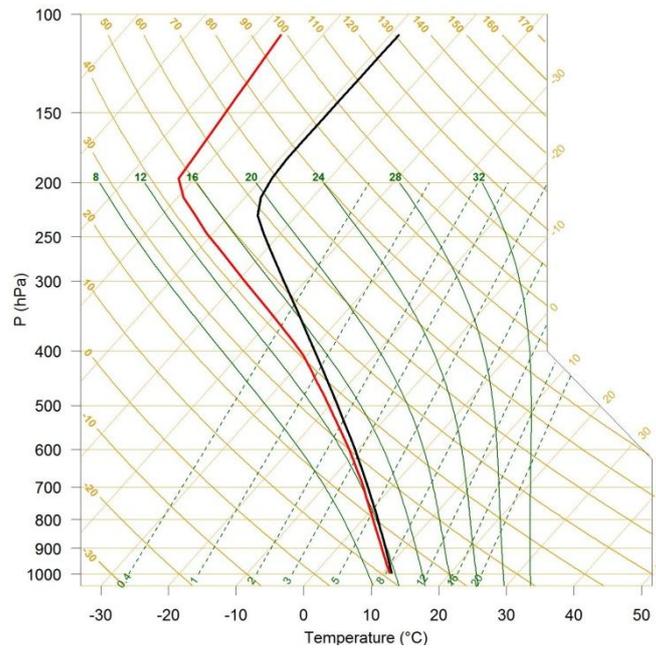


Figure V-8 : Profils verticaux de la température (en °C – noir) et de la température du point de rosé (en °C – rouge) pour un cas été.

Les conditions de l'atmosphère sont initialisées par l'implémentation d'un sondage thermodynamique représentatif d'une saison d'été. Les profils de température et d'humidité sont représentés sur la Figure V-8. Le champ de vent est supposé unidirectionnel selon la direction des  $x$  avec une vitesse horizontale augmentant de  $10 \text{ m s}^{-1}$  au niveau du sol à  $30 \text{ m s}^{-1}$  à 11 km d'altitude pour rester constant au-delà. Les conditions aux limites sont ouvertes selon la direction des  $x$  (ouest-est) et périodiques selon la direction des  $y$  (sud-nord).

Chaque simulation est intégrée sur 6 heures (avec un pas de temps de 2 secondes) durant laquelle nous estimons 2 à 3 heures de « spin-up » (*i.e.* temps nécessaire pour que les prévisions se stabilisent numériquement).

Pour tous les cas, l'aspect macroscopique (*i.e.* développement et morphologie) du nuage précipitant devrait être le même. Le vent horizontal provoque, à son arrivée sur l'obstacle, une ascendance de la masse d'air. Un nuage composé d'hydrométéores liquides se forme au sommet de la montagne. Au cours de la simulation, l'onde de gravité formée s'intensifie en aval de la montagne (Figure V-9). À haute altitude, la phase glace peut apparaître et par des phénomènes de fonte, intensifier les précipitations (Planche 2011, p. 122). Les pluies au sol sont localisées à proximité immédiate du sommet de la montagne (Zängl *et al.* 2010). Au niveau des pentes sous le vent, le rendement des précipitations est réduit par des processus d'évaporation qui s'organisent sous le flux d'air descendant.

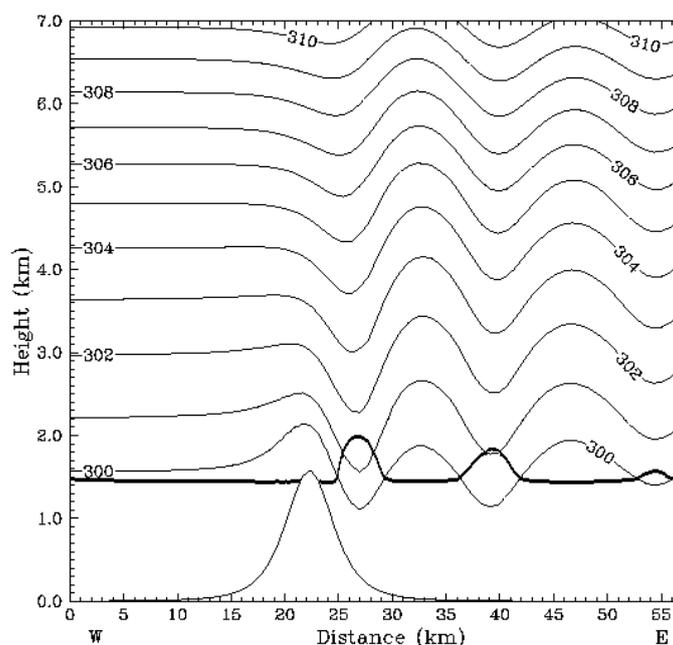


Figure V-9 : Coupe verticale de la température potentielle équivalente (degrés Kelvin). La ligne en gras indique le niveau de l'isotherme 0°C. Source : Zängl et Seifert (2009) (modifiée).

### V-4.3 Les schémas de microphysique

Dans le cadre du projet collaboratif dans lequel s'inscrit cette thèse, il avait été envisagé une confrontation de nos résultats à fines échelles avec les simulations météorologiques à l'échelle régionale réalisées avec le modèle WRF par l'un des partenaires impliqués. Les travaux ont été orientés pour une application opérationnelle.

Dans sa version 3.2.1 (août 2010), le modèle WRF-ARW proposait 11 schémas de microphysiques de type « bulk » que nous avons listé dans le Tableau I. Dès lors, nous avons rapidement écarté les 5 schémas à deux moments car leur lourd poids numérique ne peut être envisagé dans une application en temps réel. Sur les 6 schémas à 1 moment restants, le schéma de Kessler et celui de Purdue-Lin ont été considérés comme trop simplifiés et le schéma WRF Single-Moment (WSM) 3-class n'intègre pas la phase glace. Le schéma de Goddard n'a pas été retenu car il s'avère être principalement destiné pour la modélisation des nuages de glace à très haute altitude (cirrus) et non des précipitations. Notre choix s'est finalement porté sur les schémas WRF Single-Moment 6-class, sachant que le schéma WSM5 s'axe sur un même code à une catégorie d'hydrométéore près. Nous verrons que les conclusions de cette étude viennent conforter ce choix.

Tableau V-1 : Comparatif des schémas de microphysique implémentés dans WRF v3.2.1.

Nom variables : Q = rapport de mélange ; N = concentration en nombre.

Indices associés : v = vapeur d'eau ; c = eau nuageuse ; r = eau pluvieuse ; i = glace nuageuse ; s = neige ; g = grésil ; sl = neige mouillée ; h = grêle ; pi = précipitation solide (ex : neige mouillée, grêle et grésil) ; CCN = noyaux de condensation de nuage.

	Schéma	Référence	Variables considérées <i>Remarque(s)</i>	Inclusion phase glace	Inclusion phase mixte
Modèle « bulk » mono-moment	Kessler	(Kessler 1969)	Qv, Qc, Qr ⊕ schéma simple de nuage chaud.		
	Purdue-Lin	(Lin <i>et al.</i> 1983)	Qv, Qc, Qr, Qi, Qs, Qg ⊕ l'ajustement de la saturation de la glace et de la vapeur d'eau est traité de façon combinée.	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	WRF Single-Moment 3-class (WSM3)	(Hong <i>et al.</i> 2004)	Qv, Qc+i, Qr+s ⊕ la concentration est déduite du contenu en masse (non de la température) ; ⊕ considère l'eau nuageuse et la pluie (la glace et la neige) pour des températures au dessus (au dessous) du point de congélation.	<input checked="" type="checkbox"/>	
	WRF Single-Moment 5-class (WSM5)		Qv, Qc, Qr, Qi, Qs ⊕ l'ajustement de la saturation de la glace et de la vapeur d'eau est traité séparément.	<input checked="" type="checkbox"/>	
	WRF Single-Moment 6-class (WSM6)	(Hong et Lim 2006)	Qv, Qc, Qr, Qi, Qs, Qg ⊕ l'ajustement de la saturation de la glace et de la vapeur d'eau est traité séparément ; ⊕ utilise une unique vitesse de chute pour la neige et grésil (pondérée par les rapports de mélange respectifs).	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	Goddard Cumulus Ensemble Model Scheme	(Tao et Simpson 1993)	Qv, Qc, Qr, Qi, Qs, Qg/h ⊕ tous les processus n'impliquant pas les processus de transfert d'un type d'hydrométéore à un autre sont calculés à partir d'un même état thermodynamique.	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	Modèle « bulk » multi-moments	Thompson	(Thompson <i>et al.</i> 2008)	Qv, Qc, Qr, Qi, Qs, Qg, Ni, Nr ⊕ distribution en nombre de forme gamma généralisée (à l'exception de Qn et Qc) ; ⊕ la collection/coalescence des hydrométéores dépend de leur rapport de mélange respectif.	<input checked="" type="checkbox"/>
Milbrandt 2-moment		(Milbrandt et Yau 2005)	Qv, Qc, Qr, Qi, Qs, Qg, Qh, Nc, Nr, Ni, Ns, Ng, Nh ⊕ la forme de la distribution en taille des hydrométéores est considérée comme une variable pronostique, déterminée au travers d'un troisième moment : la réflectivité RADAR.	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Morrison 2-moment		(Morrison <i>et al.</i> 2009)	Qv, Qc, Qr, Qi, Qs, Qg/h, Nr, Ni, Ns, Ng ⊕ développé pour l'étude des pluies stratiformes.	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
WRF Double-Moment 5-class (WDM5) 6-class (WSM6)		(Lim et Hong 2010)	Qv, Qc, Qr, Qi, Qs (Qg WDM6), NCCN, Nc, Nr ⊕ développé pour le traitement des pluies chaudes avec pour objectif la possibilité d'étudier les effets aérosols sur la propriété des nuages et les processus de précipitations ; ⊕ inclusion de la concentration en nombre des CCN, basée sur la relation du nombre de CCN activés et la valeur de super saturation du milieu environnant.	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/> WDM 6-class)

Au cours des simulations, les précipitations étaient généralement les plus abondantes pour 180 min (= 3 heures) d'intégration. C'est pourquoi nous avons décidé de montrer principalement les résultats obtenus pour ce pas de temps.

### V-4.3.1 Le schéma WSM6

Sous l'action du vent horizontal (selon l'axe des  $x$  croissants), la masse d'air s'élève le long des pentes au vent. Au passage de la montagne, une onde de gravité est générée en altitude, instaurant une succession de subsidences ( $w$  négatives) et d'ascendances ( $w$  positives) en aval de la montagne. Après 3 heures d'intégration, nous dénombrons trois alternances successives (Figure V-10a). Les phases d'ascendances s'associent à la création de particules nuageuses (par condensation) et de particules de glace (par congélation des gouttes ou activation des IN) (Figure V-10b). À l'inverse, les phases de subsidence sont corrélées à la précipitation des hydrométéores (Figure V-10c).

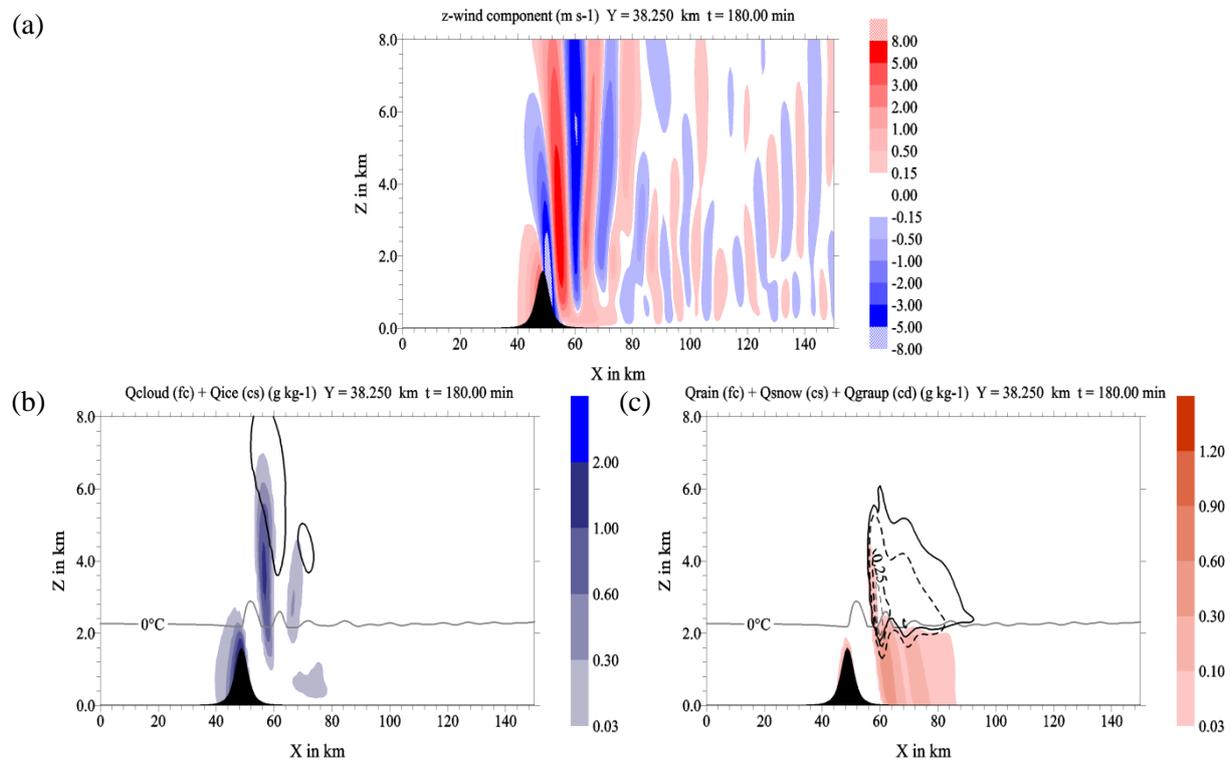


Figure V-10 : Schéma WSM6 = coupe verticale (a) du vent vertical  $w$  (en  $m s^{-1}$ ), (b) des contenus en eau nuageuse ( $Q_{cloud}$  en  $g kg^{-1}$ ) et glace ( $Q_{ice}$  en  $g kg^{-1}$ , iso-contours en trait plein) et (c) des contenus en pluie ( $Q_{rain}$  en  $g kg^{-1}$ ), neige ( $Q_{snow}$  en  $g kg^{-1}$ , iso-contours en trait plein) et graupel ( $Q_{graup}$  en  $g kg^{-1}$ , iso-contours en pointillé) obtenue après 3 heures d'intégration.

Valeurs des premiers iso-contour =  $0.03 g kg^{-1}$ , les iso-contours suivants sont dessinés tous les  $0.25 g kg^{-1}$  en traits plus fins. L'isotherme  $0^{\circ}C$  est donné par la ligne grise.

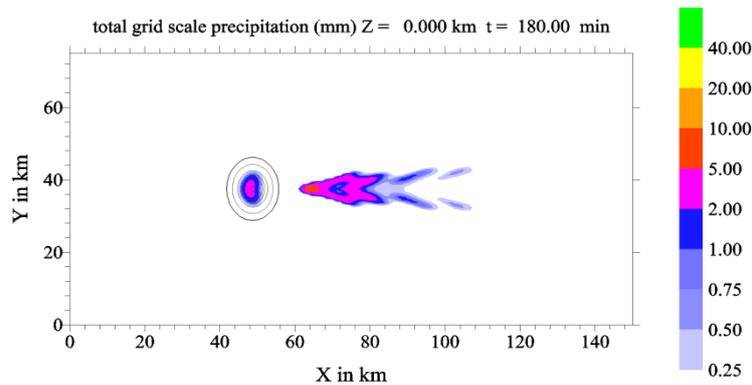


Figure V-11 : Schéma WSM6 = répartition de la pluie au sol cumulée sur une heure (mm) entre la 2<sup>ème</sup> et 3<sup>ème</sup> heure d'intégration.

La Figure V-11 présente la répartition spatiale de la pluie au sol cumulée entre la deuxième et troisième heure d'intégration. Nous retrouvons bien i) l'ascension de l'air par l'orographie qui déclenche les précipitations sur la montagne avec un maximum enregistré sur les versants au vent et à proximité du sommet et (ii) la génération d'une onde de gravité en aval où les flux ascendants sont le siège de phénomènes d'évaporation limitant la précipitation des plus petites gouttes de pluie ( $x$  autour de 60 ou 70 km) et où les flux descendants sont source de précipitations ( $x$  autour de 64 ou 78 km).

#### V-4.3.2 Les (dix) autres schémas

Dans la suite nous présentons quelques résultats obtenus avec les dix autres schémas de microphysique que propose le modèle. Dans la majorité des simulations réalisées, nous observons la présence de phénomènes singuliers qui n'apparaissent pas (ou du moins de façon moins prononcée) avec le schéma WSM6.

##### a | Les schémas de Kessler et de Thompson

Même si le schéma de Kessler n'intègre pas les processus liés à la phase glace, les résultats obtenus sont très similaires avec ceux issus d'un schéma plus complexe qu'est le schéma de Thompson.

Les deux schémas montrent une onde de gravité légèrement plus développée (5 successions) et bien plus structurée au-delà de  $x = 70$  km (Figure V-12a et Figure V-13a). Les effets associés aux flux ascendants sont bien représentés (Figure V-12b et Figure V-13b). Nous noterons toutefois la présence de particules nuageuse (Qcloud) à très basses altitudes (à peine au-dessus de 200 m). *A contrario*, la précipitation des hydrométéores favorisée par les flux descendants est fortement limitée comme nous pouvons le constater sur la Figure V-12c et Figure V-13c.

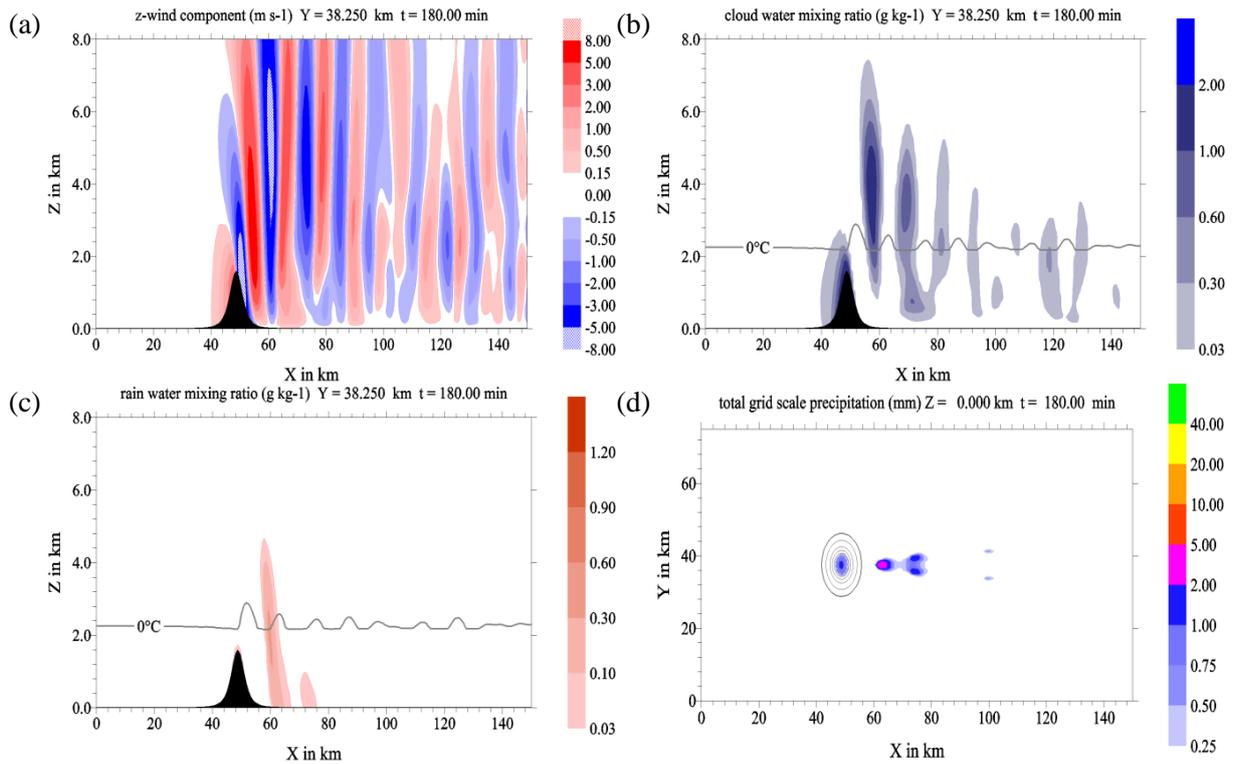


Figure V-12 : Schéma de Kessler = (a-c) *idem* Figure V-10 et (d) *idem* Figure V-11.

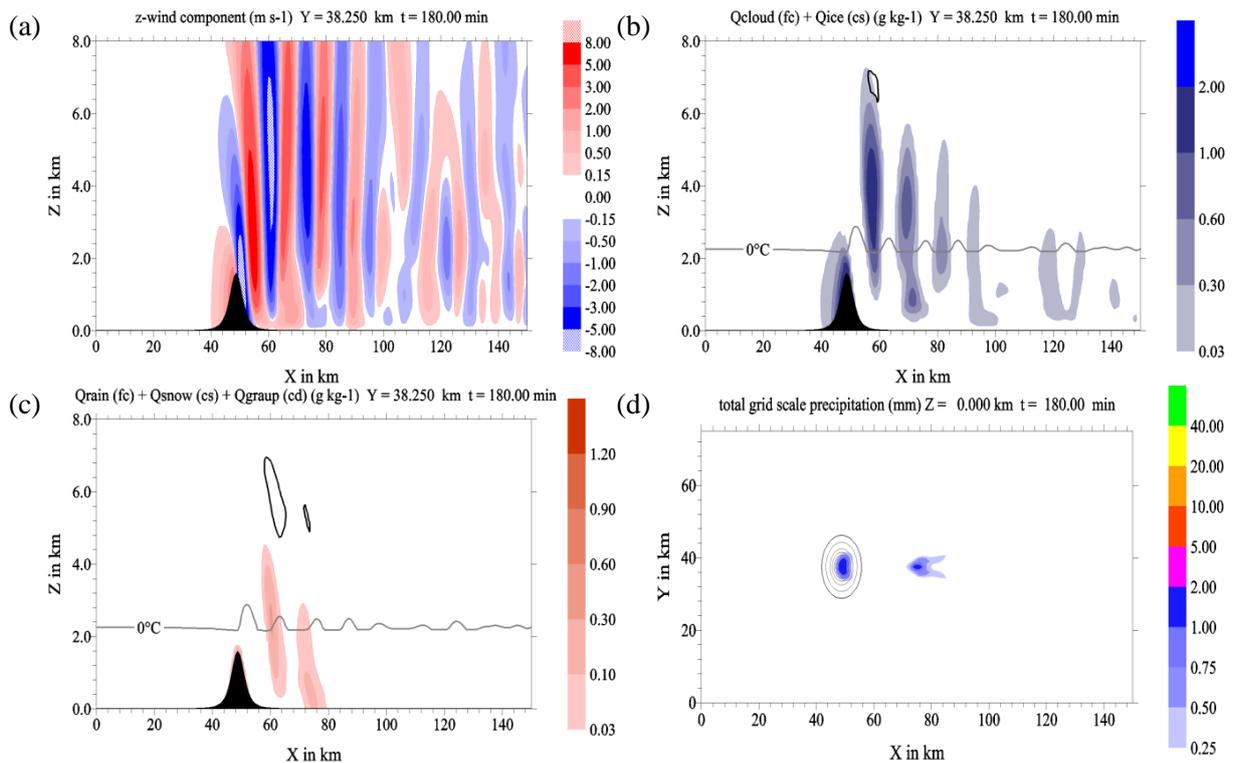


Figure V-13 : Schéma de Thompson = (a-c) *idem* Figure V-10 et (d) *idem* Figure V-11.

Les champs de pluie obtenus avec ces deux schémas (Figure V-12d et Figure V-13d) sont parmi les plus faibles de toutes les simulations réalisées. Le schéma de Thompson simule finalement très peu de glace ( $Q_{ice} + Q_{snow}$ ). Ces particules n'ont pas d'influence majeure sur les précipitations au sol étant donné qu'elles sont de plus localisées à trop haute altitude.

Nous avons pu constater que les précipitations s'affaiblissent jusqu'à quasiment disparaître après 6 heures de simulation (non montré).

### b | Le schéma Purdue-Lin

Avec le schéma de Purdue-Lin, les taux du champ de précipitation sont nettement amplifiés (Figure V-14a). Nous observons notamment un important maximum (supérieur à 20 mm) en aval de l'obstacle ( $x \approx 62\text{km}$ ). Ce comportement s'explique par l'imposante présence de glace derrière la montagne ( $Q_{snow} + Q_{graup}$  – Figure V-14b). Ces résultats s'apparentent plus à l'étude d'un cas de convection profonde que ceux attendus pour le cas d'une simple situation orographique. Des résultats similaires ont été observés dans l'étude d'un cas réel présentée par Lin et Colle (2009).

### c | Les schémas WSM3 et WSM5

Pour la simulation avec le schéma WSM3, la dynamique est quelque peu changée comme par exemple sur l'onde de gravité. Initiée par le franchissement de la montagne, elle est plus faible *que celle obtenue par le schéma WSM6* (Figure V-10a) et les effets associés sont disparates : la formation des particules nuageuses est dépréciée de par l'absence de la phase glace ( $Q_{cloud}$  – Figure V-15a) alors que la présence d'hydrométéores liquides est fortement accentuée notamment au dessus de l'isotherme  $0^\circ\text{C}$  ( $Q_{rain}$  – Figure V-15b). Nous noterons également une intensité des pluies au sommet de la montagne bien supérieure à celles observées dans toutes nos autres simulations.

La prise en compte d'une partie de la phase glace ( $Q_{ice}$  et  $Q_{snow}$ ) dans le schéma WSM5 permet de corriger ces effets. En comparaison avec WSM6, les précipitations sont légèrement plus faibles (Figure V-16a à comparer avec la Figure V-11). Ce constat s'explique par l'absence des plus grands cristaux de glace ( $Q_{graup}$ ). Les précipitations associées à la fonte de ces hydrométéores ne sont donc pas intégrées. C'est ce que nous pouvons par exemple observer sur la Figure V-16b avec une présence limitée d'hydrométéores liquides ( $Q_{rain}$ ) à  $x \approx 74$  ou  $82$  km (à comparer avec la Figure V-10c).

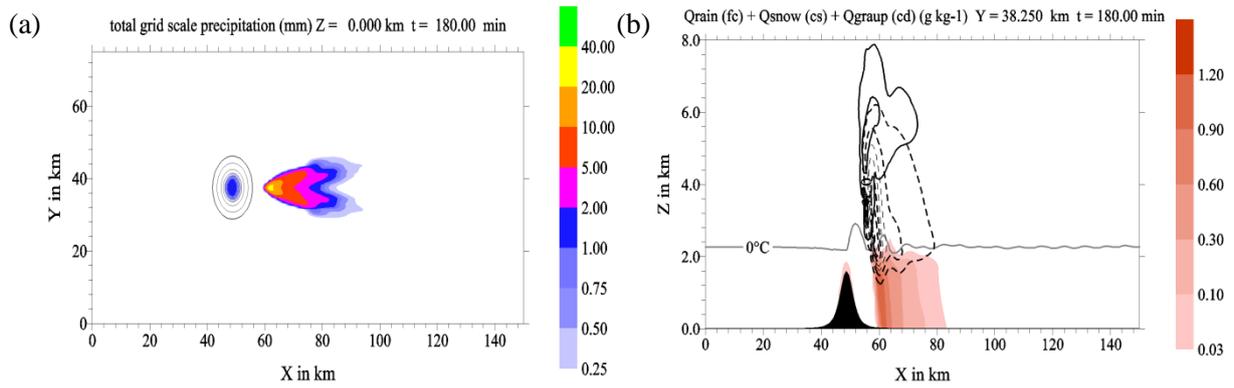


Figure V-14 : Schéma de Purdue-Lin = (a) *idem* Figure V-11 et (b) *idem* Figure V-10c.

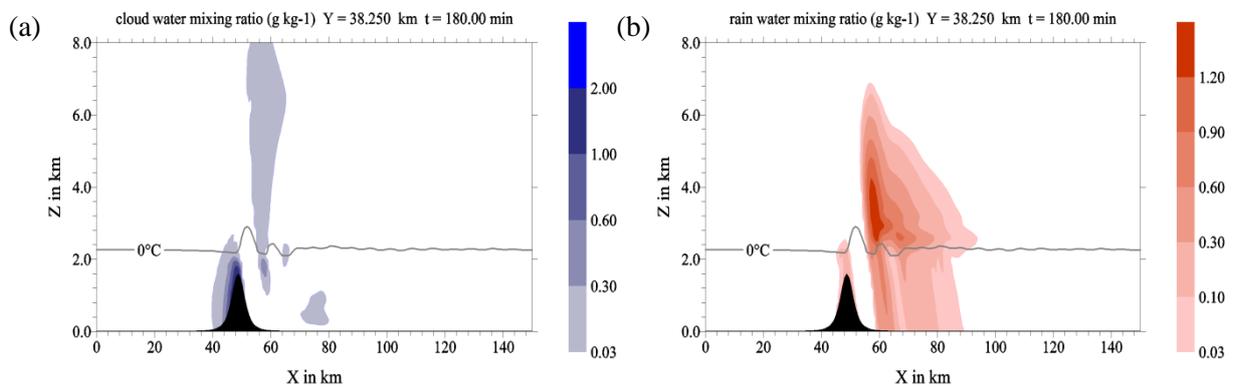


Figure V-15 : Schéma WSM3 = (a-b) *idem* Figure V-10(b-c).

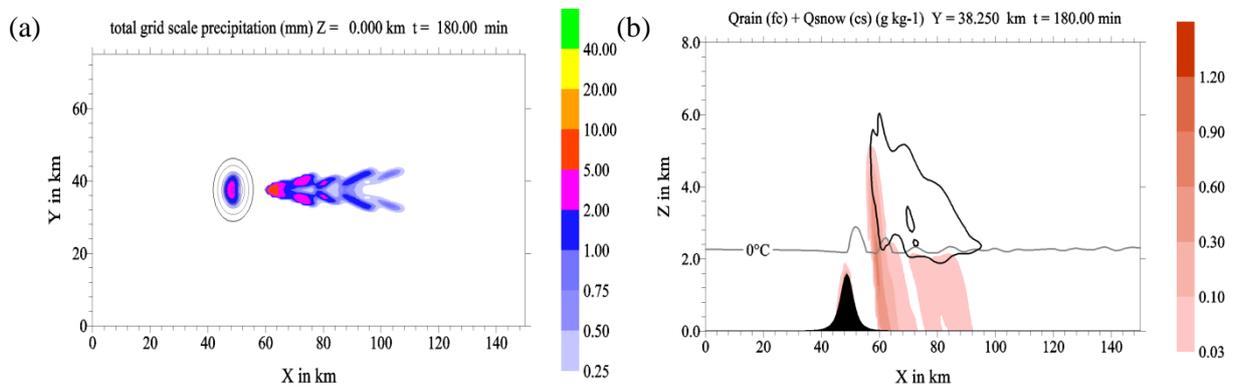


Figure V-16 : Schéma WSM5 = (a) *idem* Figure V-11 et (b) *idem* Figure V-10c.

### d | Le schéma de Goddard

Le schéma de Goddard, initialement conçu pour l'étude des nuages de glace, accorde trop d'importance à la phase glace pour notre cas. En effet, à haute altitude, les fortes concentrations en particules de glace ( $Q_{ice}$  – Figure V-17a) sont des plus importantes en comparaison avec les autres schémas testés. Nous notons également que la forte présence de neige entre 2 et 6 km d'altitude ( $Q_{snow}$  – Figure V-17b) qui, sous l'isotherme  $0^{\circ}\text{C}$ , fond pour donner d'importantes pluies. Enfin, même si nous observons la formation d'un nuage sur les versants au vent, aucune précipitation ne sera générée aux abords de l'obstacle.

### e | Les schémas de Milbrandt-Yau et Morrison

Toutes les simulations ont été réalisées avec les mêmes conditions, pourtant celles réalisées avec les schémas à double moment de Milbrandt-Yau et Morrison présentent une particularité étonnante. Premièrement, l'onde de gravité est particulièrement puissante et bien structurée (Figure V-18a et Figure V-19a – avec 9 successions flux ascendant/descendant, *contre 3 pour WSM6*). Comme pour les schémas de Kessler et Thompson (Figure V-12b et Figure V-13b), les effets associés aux flux ascendants sont fortement accentués avec notamment la présence de particules nuageuse à très basse altitude ( $Q_{cloud}$  – Figure V-18b et Figure V-19b). Les particules de glace ( $Q_{ice}$ ) ne sont que très peu voire pas du tout simulées. Il en va de même pour les hydrométéores précipitants : les phénomènes liés aux flux descendants ne sont pas correctement traduits ( $Q_{cloud}$  – Figure V-18c et Figure V-19c). Ces deux schémas sont les seuls à afficher un champ de pluie au sol nul (non montré).

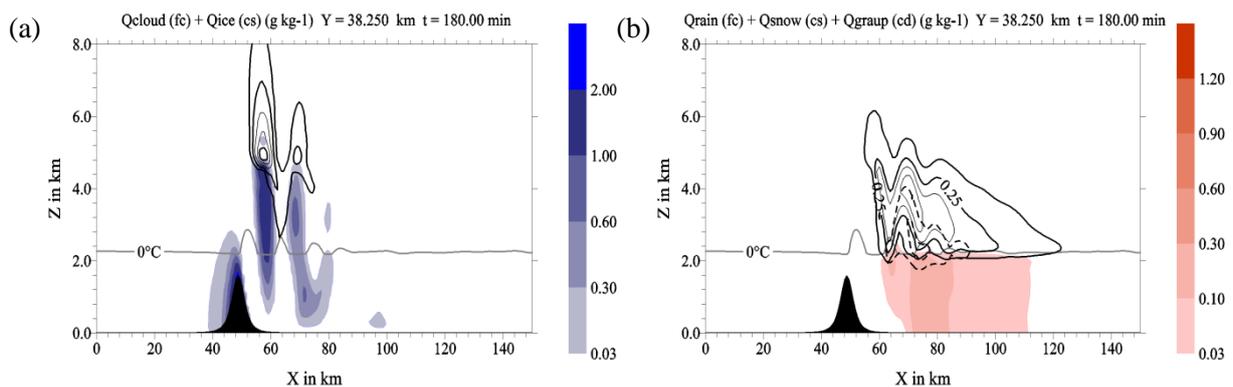


Figure V-17 : Schéma de Goddard = (a-b) *idem* Figure V-10(b-c).

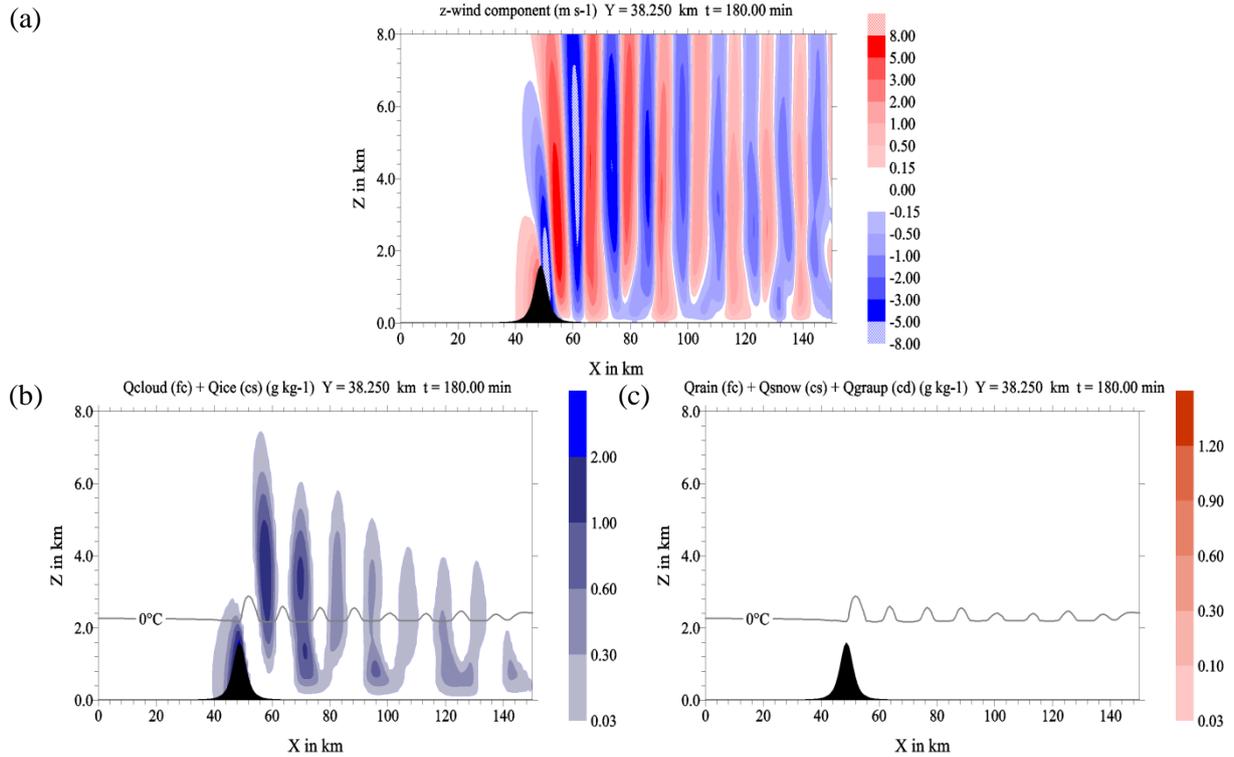


Figure V-18 : Schéma de Milbrandt-Yau = (a-c) *idem* Figure V-10.

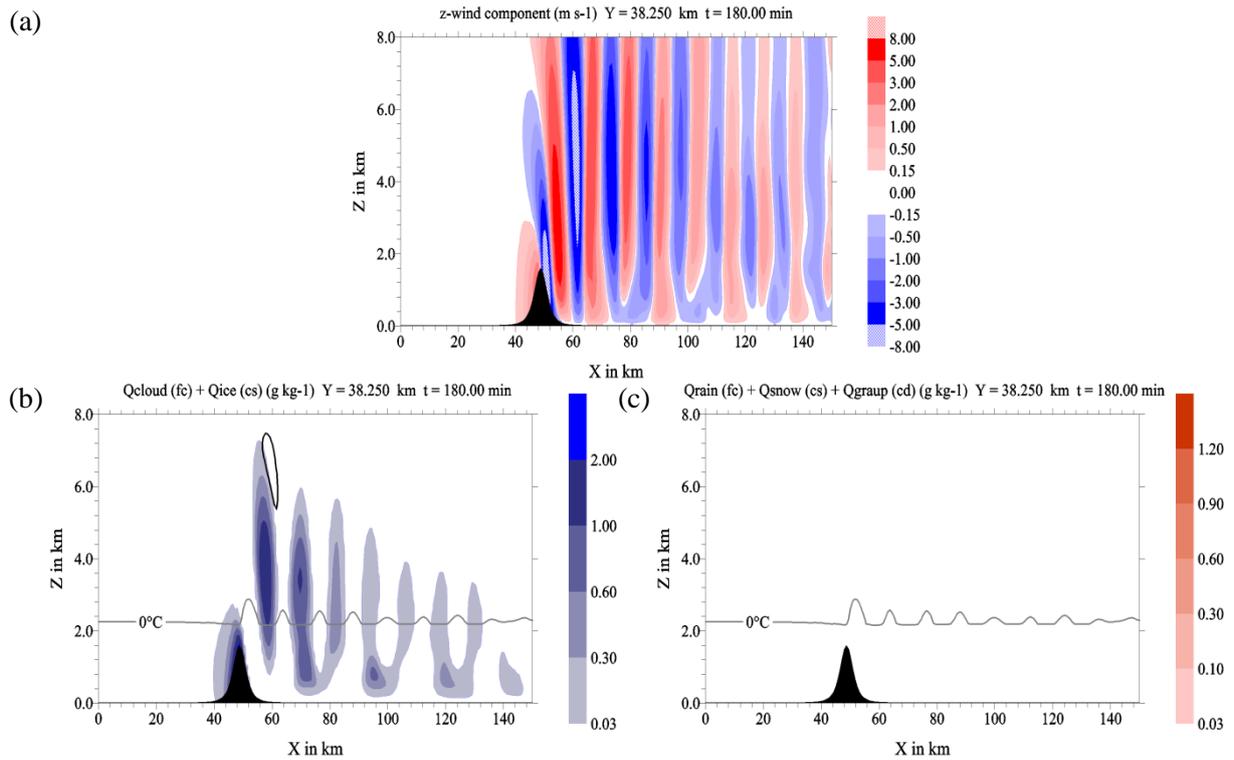


Figure V-19 : Schéma de Morrison = (a-c) *idem* Figure V-10.

## f | Les schémas WDM5 et WDM6

Malgré leur caractère détaillé, les schémas WRF-Double Moment (WDM) 5-class et 6-class simulent peu de précipitation au sol. Pourtant l'onde de gravité formée est semblable à celle obtenue avec le schéma WSM6 avec moins de structures sporadiques après  $x = 80$  km (Figure V-20a et Figure V-21a). Les nuages formés par les flux ascendants sont cependant moins bien dessinés lorsque nous nous approchons de l'isotherme  $0^{\circ}\text{C}$  (Qcloud – Figure V-20b et Figure V-21b) Les précipitations associées (Qrain – Figure V-20c et Figure V-20c) sont limitées par d'importants phénomènes d'évaporation (*cf.*  $x$  autour de 88 km). Nous noterons que la fonte des hydrométéores glacés (Qsnow & Qgraup) sous l'isotherme  $0^{\circ}\text{C}$  n'est pas très bien représentée comme en témoigne la zone sans pluie autour de  $x = 64$  km.

Le développement constant du modèle WRF ne facilite pas les choix qui doivent être faits pour obtenir des simulations réalistes. Les travaux réalisés dans ce chapitre nous ont permis de sélectionner un schéma de microphysique parmi la longue liste des paramétrisations alors proposées. Notre choix c'est porté sur le schéma WSM6 puisque i) il intègre la phase glace ( $\neq$  Kessler et WSM3) de façon modérée ( $\neq$  Purdue-Lin et Goddard), ii) il reproduit convenablement les précipitations orographiques ( $\neq$  Thompson, Milbrandt-Yau et Morrison) et iii) les effets associés à l'onde de gravité sont équilibrés ( $\neq$  WDM5 et WDM6).

Ce choix est conforté par les indications apportées par Hong *et al.* (2010). En effet, les auteurs indiquent que le schéma WSM6 a permis notamment d'obtenir des informations utiles pour la simulation à haute résolution de phénomènes météorologiques notamment dans les Grandes Plaines aux États-Unis. Ils précisent également que le Korean Meteorological Administration et la Korean Air Force ont choisi ce schéma de microphysique pour leurs prévisions en temps réel sur l'Asie de l'Est. Enfin, ils notent que les performances du schéma WSM6 ont été évaluées au travers de différentes situations météorologiques incluant les phénomènes de convection profonde au dessus des États-Unis, des cyclones extratropicaux et ouragans formés dans l'Atlantique ou des événements de pluie intenses en Asie de l'Est.

Au travers de ce chapitre, nous avons montré que la prévision de la pluie reste une tâche difficile. En optant pour le schéma WSM6, nous nous assurons d'une convenable représentation des précipitations. L'étape suivante est de quantifier la qualité des prévisions du modèle au travers d'un cas réel. La mise en place d'une technique de comparaison avec le champ des observations fait l'objet du chapitre suivant.

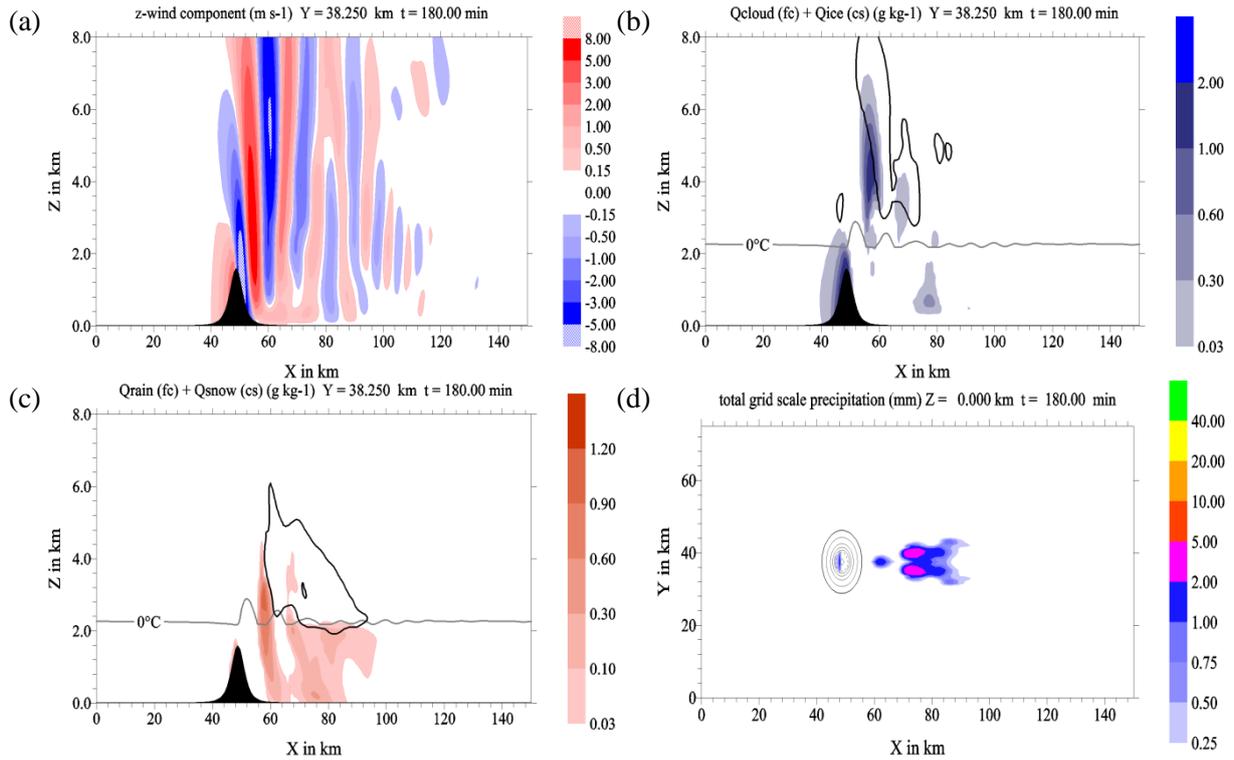


Figure V-20 : Schéma WDM5 = (a-c) *idem* Figure V-10 et (d) *idem* Figure V-11.

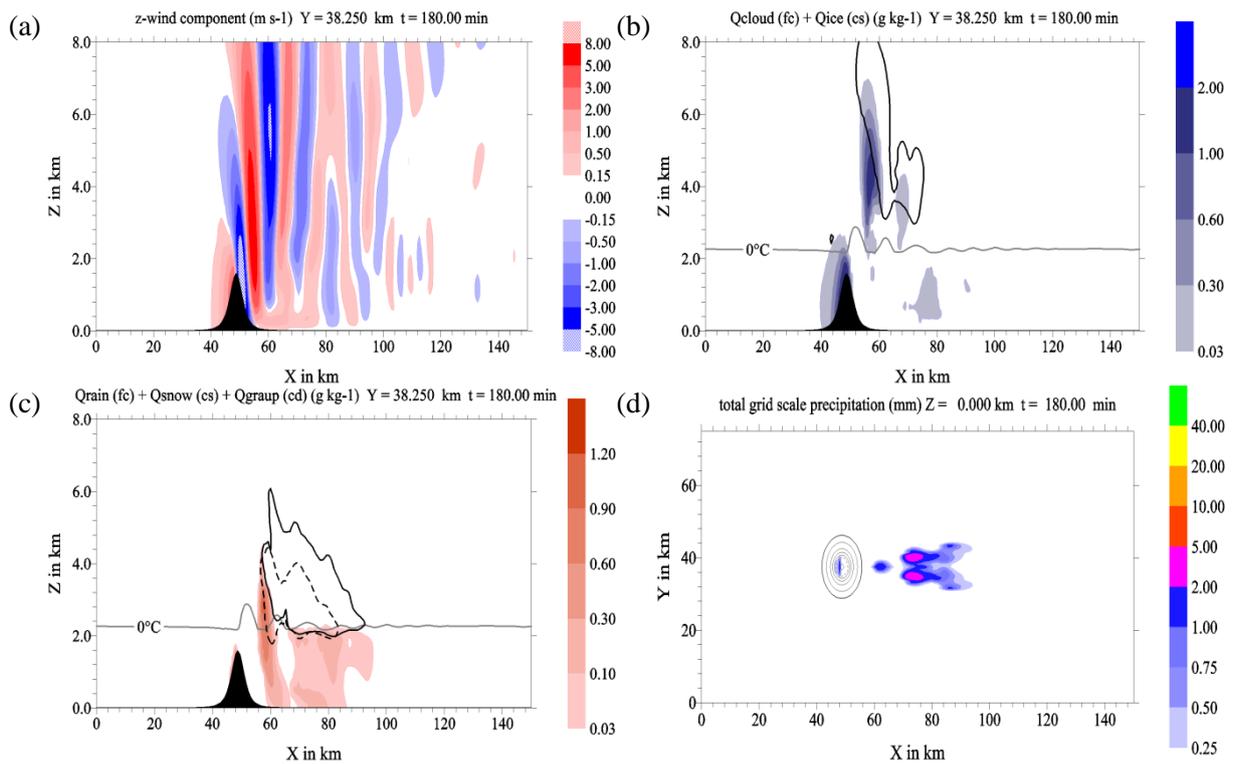


Figure V-21 : Schéma WDM6 = (a-c) *idem* Figure V-10 et (d) *idem* Figure V-11.



# Chapitre VI

## Évaluation du modèle WRF : application sur les précipitations

Dans le chapitre précédent, nous avons présenté le modèle météorologique WRF. Nous avons notamment mis en exergue sa capacité à simuler les pluies au travers d'une situation idéalisée. Nous proposons ici une étude sur la vérification de ces prévisions. Nous comprendrons alors l'intérêt porté sur la reconstruction du champ spatial à partir des observations pluviométriques et mesures RADAR (techniques présentées dans le 0 et testées dans les 0 et Chapitre IV). Ce chapitre se compose de trois parties. La première consiste en une revue sommaire des différentes méthodes de comparaison avec un focus particulier sur celles intégrant la variabilité spatiale des précipitations. Nous présenterons ensuite celle que nous avons retenue avant de l'appliquer dans un cas réel. Nous terminerons ce dernier chapitre par une discussion des atouts de la technique choisie en comparaison de méthodes plus traditionnelles.

### VI-1 Bonne ou mauvaise prévision : telle est la question

La vérification est une étape indispensable pour quantifier la correspondance entre prévisions et observations. Nous verrons au travers d'un rapide état de l'art que cette volonté ne date pas d'aujourd'hui. Puis nous développerons succinctement les différentes approches qui prennent en compte le caractère spatial du paramètre étudié (ici, le champ de pluie).

#### VI-1.1 Petit historique des méthodes de vérification

Dans le domaine de la météorologie, les premières techniques de vérification apparaissent au XIX siècle par les travaux de Cleveland ABBE (1838-1916). Il est le premier prévisionniste opérationnel à établir la vérification systématique de ses prévisions par des comparaisons simples avec des observations ponctuelles. Ces rapprochements lui permettent

alors de qualifier arbitrairement ses prévisions en fonction de leurs observations avérées. Un peu plus tard, John P. FINLEY (1854-1943) améliora la technique de vérification de son prédécesseur en créant les premières tables de contingence (qui sera expliqué dans le sous-paragraphe VI-4.1.1). Pour ses recherches, il disposait de plus de 800 stations dans 18 districts de l'Est des USA lui rapportant si oui ou non une tornade avait été observée. Ces observations lui permettaient de qualifier ses prévisions ainsi : lorsqu'une prévision annonçait une tornade dans un district, elle était alors considérée comme « correcte » si au moins un tube ou plus avait été observé et à l'inverse, considérée comme « fausse » si aucune tornade n'avait été observée. Il. Nombreux sont ceux qui ont dès lors cherché à définir des calculs statistiques qui en découlent pour obtenir l'information la plus représentative possible. C'est Allan H. MURPHY (1931-1997) qui révolutionna les études sur la vérification ; il est aujourd'hui considéré comme le « père de la vérification ». Dans son célèbre essai (Murphy 1993), il différencie trois catégories de vérification qui dépendent de ce que l'utilisateur souhaite mettre en avant à savoir :

- la correspondance entre une prévision et l'expertise de l'utilisateur (« Consistency ») ;
- la correspondance entre une prévision et l'(les)observation(s) associée(s) (« Quality ») ;
- l'impact d'une décision faite à partir des prévisions (« Value »).

Dans le domaine de la météorologie, les méthodes de vérification sont le plus souvent établies pour évaluer un modèle par rapport aux observations, ce qui correspond au second type « Quality ». Depuis, les techniques de comparaisons observation(s)/modèle(s) se sont complexifiées. Outre, le nombre d'observations météorologiques qui ne cesse d'augmenter et l'apparition de nouveaux instruments (ex : RADAR, satellites), de nombreux modèles météorologiques ont été développés, chacun étant de plus en plus complexes et ayant des résolutions spatiale et temporelle de plus en plus fines. À présent, un utilisateur ne cherche plus à faire des vérifications point-à-point mais sur des champs spatiaux-temporels (ex : précipitation, couverture nuageuse, etc.).

Il existe aujourd'hui un large choix de méthodes qui se différencient selon le type de variables considérées (variables nominales (ex : OUI il a plu / NON il n'a pas plu) ou continues (ex : valeurs de cumul de pluie)). Le choix d'une technique est variable selon les objectifs de l'utilisateur (ex : évaluation des performances d'un modèle, identification et correction des erreurs d'un modèle, développement d'outils d'aide à la décision, etc.) et de ce qu'il souhaite mettre en avant (Murphy 1993). Les résultats doivent être pour lui significatifs et faciles à interpréter.

Pour quantifier les erreurs de prévisions, nous avons choisi de mettre l'accent sur la vérification spatiale des champs de pluie. Cette procédure de vérification se limite rarement à un unique nombre (ou score). En effet, selon les approches choisies, les informations obtenues peuvent s'avérer différentes, quoique souvent complémentaires. Les scores dits « traditionnels » (ex : erreur moyenne, RMSE, coefficient de corrélation – cf. sous-paragraphe VI-4.2.1) considèrent les erreurs globales sur les champs moyens. Ces scores sont cependant fortement sensibles aux larges erreurs sur la quantité ou aux petites erreurs de localisation. Par exemple, Rossa *et al.* (2008) jugent préférable que la vérification soit faite au travers des scores catégoriels (ex : BIAIS, ETS – cf. sous-paragraphe VI-4.1.1). Comme nous le verrons dans le paragraphe suivant, il n'existe donc pas une unique technique de vérification spatiale.

## VI-1.2 Les méthodes de vérification spatiale

Il s'avère que la comparaison point-à-point ne peut donner une information complète sur la qualité d'une prévision car elle n'a pas de regard sur le caractère spatial intrinsèque au paramètre météorologique considéré (Gilleland *et al.* 2009; Casati *et al.* 2008). En effet, si par exemple un champ de pluie est correctement prédit en termes de surface ou d'intensité, il peut tout de même conduire à de très mauvais scores (traditionnel et/ou catégoriel) si le champ n'est pas localisé au bon endroit. Nous sommes alors pénalisés d'une part pour avoir manqué l'observation à un endroit précis et d'autre part pour l'avoir prévu ailleurs. Cette double pénalité peut rapidement devenir problématique notamment lorsque les prévisions sont réalisées à très fines échelles spatiales. En réponse à ces effets indésirables, les chercheurs proposent de nombreuses nouvelles techniques de vérification intégrant la notion d'espace voire de temps. Ces dernières sont regroupées en 4 catégories :

- les méthodes *neighborhood* ou *fuzzy* ou *upscaling*. Les grilles qui découpent le champ de prévision et d'observation sont redéfinies en considérant plusieurs points de grille voisins dans l'espace (voire dans le temps). Pour chaque nouveau pixel, une nouvelle valeur est alors calculée (à l'aide par exemple de la moyenne simple ou de la valeur maximale des pixels voisins). Les résultats sont des versions lissées des champs originaux où les valeurs extrêmes sont pour la plupart préservées. Le processus est répété en utilisant un voisinage plus large. Les champs obtenus peuvent alors être comparés par exemple sur leurs valeurs moyennes ou maximales ou en utilisant divers scores statistiques le plus couramment de type catégoriels. Ces techniques sont cependant fortement dépendantes de la définition des voisins et de la procédure d'agrégation des données. Il existe une dizaine de

techniques (Ebert 2008) dont : neighborhood (Ebert 2009), fractions skill score (Roberts et Lean 2008; Mittermaier et Roberts 2010) et multi-event contingency table (Atger 2001) ;

- les méthodes *scale-separation* ou *scale-decomposition*. Ces méthodes consistent à appliquer un filtre spatial à bande passante unique (ex : transformées de Fourier, ondelettes, transformées en cosinus discrètes, etc.) sur le champ de prévision et d'observation. Contrairement aux méthodes *neighborhood*, les champs résultants peuvent ne pas ressembler aux champs originaux. Les caractéristiques des champs obtenus sont ensuite analysées pour différentes échelles spatio-temporelle indépendantes ce qui permet d'isoler les phénomènes physiques associés à une échelle particulière (ex : système frontal, cellules convectives, etc.). Il existe quatre techniques dont : wavelets (Briggs et Levine 1997) et intensity-scale – IS (Casati *et al.* 2004; Casati 2010) ;
- les méthodes *objects-based*, *features-based* ou *cells-identification*. En définissant une valeur seuil, l'utilisateur identifie des « objets » (ex : cellules précipitantes). Des critères sur la localisation, l'intensité ou la structure (orientation et/ou taille et/ou forme) peuvent être envisagés pour quantifier les différences entre les objets prévu et observé. Une demi-douzaine de techniques existent dont : contiguous rain area (Ebert et McBride 2000; Ebert et Gallus 2009), method for object-based diagnostic evaluation – MODE (Davis *et al.* 2006, 2009), structure-amplitude-location (Wernli *et al.* 2008, 2009) ;
- les méthodes *field deformation* ou *morphing*. Cette dernière catégorie permet de visualiser sous forme de champ vectoriel les déplacements à appliquer au champ des prévisions pour le rendre le plus proche de celui des observations. Contrairement aux méthodes *objects-based*, celles-ci considèrent les champs dans leur globalité. L'analyse permet d'obtenir des informations sur l'erreur de localisation et/ou d'intensité. Il existe six techniques dont : binary image metrics (Gilleland *et al.* 2008; Gilleland 2011), forecast quality index (Venugopal *et al.* 2005), optical flow (Keil et Craig 2007, 2009) ou image warping (Alexander *et al.* 1998; Gilleland *et al.* 2010b).

Chaque méthode de vérification s'associe à un concept et une représentation des résultats particuliers. La Figure VI-1 illustre ce propos au travers d'un cas réel de vérification sur un champ de pluie. L'exemple provient d'un des neuf cas d'étude du Projet d'Inter-Comparaison (ICP<sup>Ⓢ</sup>) des méthodes de vérification spatiale.

<sup>Ⓢ</sup> Les méthodes de vérification spatiales ont été que très récemment introduites et beaucoup d'entre elles sont encore très peu répandues et utilisées. L'ICP (<http://www.ral.ucar.edu/projects/icp>) a été élaborée par plusieurs auteurs de ces techniques afin de comparer leurs performances respectives. Toutes les approches sont ainsi

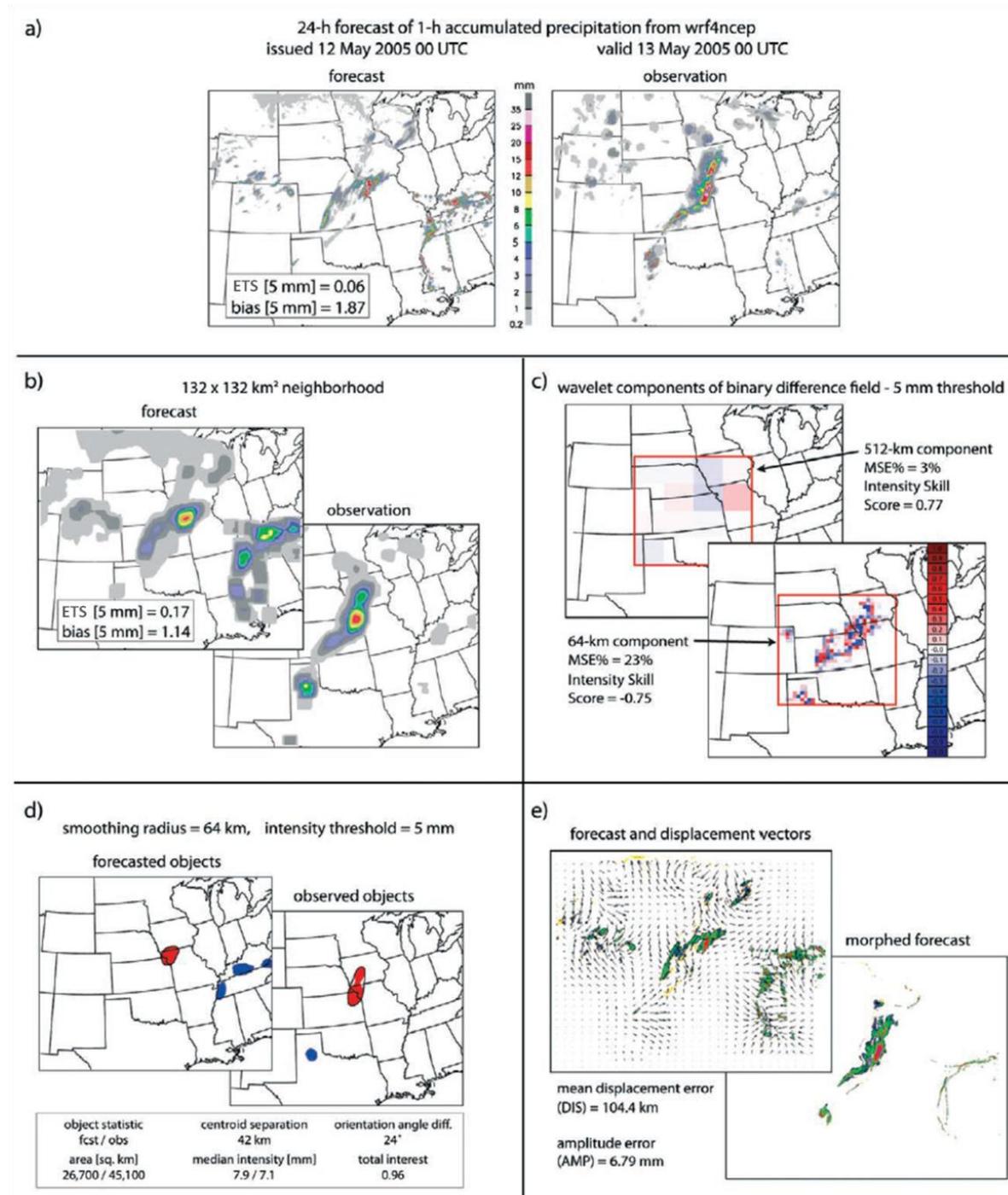


Figure VI-1 : (a) Champ de précipitation du 12-13 Mai 2005 prévu (à gauche) et observé (à droite). Les quatre catégories de méthodes de vérification spatiale sont présentées à travers des techniques : (b) neighborhood (utilisant une moyenne glissante pour les points de grille contenus dans un rayon de  $132 \times 132$  km) ; (c) IS (avec 2 décompositions à 64 et 512 km) ; (d) MODE (avec un seuillage de 5 mm et une décomposition à 64 km) et (e) optical flow (uniquement sur le champ des prévisions).  
Source : Gilleland *et al.* (2010a) (modifiée).

testées sur quelques cas d'étude (idéaux et réels) focalisés sur la prévision quantitative des précipitations (Ahijevych *et al.* 2009). L'objectif est de faire comprendre leurs principes, leurs caractéristiques et les informations récoltées pour qu'un utilisateur inexpérimenté puisse choisir la méthode qui conviendra le mieux à ses attentes.

La Figure VI-1a montre le champ de pluie prévu et observé sur une grille de  $4 \times 4$  km<sup>2</sup>. Les données horaires ont été cumulées sur 24 heures consécutives (du 12 au 13 Mai 2005). Une première comparaison entre les deux champs a été réalisée au travers de deux scores catégoriels (ETS et BIAIS) définis pour une valeur seuil de 5 mm (ce qui implique que la table de contingence a été établie telle que : OUI ou NON le cumul de pluie dépasse 5 mm).

Les champs de précipitation obtenus avec la technique neighborhood (Ebert 2009) sont présentés dans la Figure VI-1b. Les grilles ont été ici redéfinies en considérant un voisinage de  $132 \times 132$  km<sup>2</sup> et les nouvelles valeurs ont été calculées en utilisant une moyenne glissante. Nous observons alors une amélioration de l'ETS (augmentation de 0.06 à 0.17) et du BIAIS (diminution de 1.87 à 1.14). Ce type de filtrage permet de faire abstraction de petites erreurs de localisation de sorte qu'elles ne pénalisent pas le calcul des scores.

La Figure VI-1c illustre la différence entre le champ des prévisions et des observations qui ont été convertis en images binaires (*i.e.* 0 si le cumul est inférieur à 5 mm, 1 sinon) et décomposés selon l'approche IS (Casati *et al.* 2004; Casati 2010) pour deux échelles (= 512 et 64 km). La proportion de différences associée à la décomposition à 64 km est plus grande que celle à 512 km (MSE = 23 vs 3 %), ce qui suggère que les erreurs de prévisions sont importantes pour une échelle spatiale de 64 km que celles obtenus pour une échelle de 521 km.

L'exemple donné dans la Figure VI-1d est le résultat de la technique MODE (Davis *et al.* 2006, 2009). Il est alors estimé que l'objet rouge identifié dans le champ des prévisions présente une certaine correspondance avec l'objet rouge identifié dans le champ des observations. La comparaison quantitative de ces deux objets peut être donnée à travers le déplacement des barycentres, la différence entre les angles d'orientation ou l'intensité médiane des objets. Notons que les objets prévus qui ne trouvent pas de correspondance avec les observations sont représentés en bleu. Ces objets ne sont pas pris en compte dans l'analyse mais indiquent toutefois que le modèle affiche plusieurs fausses alarmes (*i.e.* cellules prévues mais non observées) à l'est du domaine et une cellule manquée (*i.e.* observée mais non prévue) au sud-ouest.

Enfin, la technique optical flow (Keil et Craig 2007, 2009) permet de déterminer le champ de déplacements à appliquer sur le champ des prévisions pour le rendre plus ressemblant au champ des observations (Figure VI-1e). En procédant de même pour le champ des observations, il est possible d'obtenir l'information sur l'erreur moyenne du déplacement (104.4 km) ou l'erreur sur l'intensité (6.79 mm).

Dans le cadre de nos travaux, nous nous sommes focalisés sur les méthodes *object-based* et plus particulièrement sur la technique « Structure – Amplitude – Location » (ou SAL) développée par Wernli *et al.* (2008).

## VI-2 La technique SAL

« Est-ce que le modèle est capable de prévoir le caractère spatial des précipitations ? Est-ce que la localisation des pluies correspond à celle observée ? Observons-nous un déplacement dans l'espace des cellules prédites par rapport aux observations ? Est-ce que les l'étendue spatiale des précipitations est plus large ou plus restreinte que celle observée ? Est-ce quantitativement les prévisions s'accordent avec les mesures ? »

Voilà le genre de questions aux quelles nous avons jugées pertinent de répondre pour quantifier les prévisions d'un modèle face aux observations. Pour satisfaire nos attentes, nous avons opté pour une approche de type *object-based* étant donné que nous ne souhaitons ni déterminer l'échelle spatiale à laquelle le modèle est le plus performant (*cf.* méthodes *neighborhood* ou *scale-decomposition*) ni de savoir quelles actions doivent être apportées pour faire correspondre les prévisions aux observations (*cf.* méthodes *field deformation*). Finalement, la technique retenue permet de quantifier trois aspects séparément : la structure (S), l'intensité (A pour « amplitude » en anglais) et la localisation (L) des précipitations survenues dans le domaine étudié. Dans ce qui suit, nous allons la décrire mathématiquement.

### VI-2.1 Définition mathématique

Développée par Wernli *et al.* (2008), la technique SAL se partitionne en trois étapes :

- la détermination du domaine d'étude : bassin versant, régions, pays, domaine d'un modèle numérique dans sa totalité ou en partie, etc. ;
- l'identification des objets : cette étape est réalisée par la détermination d'une valeur seuil  $R^*$  telle que :

$$R^* = \frac{1}{15} \cdot R^{max} \quad [VI-1]$$

où  $R^{max}$  est associé à la valeur maximale des précipitations enregistrée dans le domaine d'étude. Ainsi, tous les points de grilles voisins dont leur valeur excède  $R^*$  forment un « objet ». Chaque objet est numéroté  $m: \{m = 1, \dots, \mathcal{M}\}$ . Cette étape est appliquée indépendamment pour le champ des observations et des prévisions ;

- le calcul des trois composants (structure, intensité et localisation) comme décrit ci-après.

Afin de comprendre l'interprétation de chacun des trois composants de la technique SAL, la Figure VI-2 propose une vision schématisée de quelques situations idéalisées.

### VI-2.1.1 L'intensité $A$

L'intensité  $A$  correspond à la différence normalisée des quantités moyennes de précipitation sur le champ observé ( $\mathbb{R}_{obs}$ ) et prévu ( $\mathbb{R}_{mod}$ ) telle que :

$$A = \frac{\overline{\mathbb{R}_{mod}} - \overline{\mathbb{R}_{obs}}}{0,5 \cdot (\overline{\mathbb{R}_{mod}} + \overline{\mathbb{R}_{obs}})} \quad [\text{VI-2}]$$

Ce premier critère donne une mesure simple du biais sur la quantité de pluie prévue sur l'ensemble du domaine, abstraction faite de la variabilité spatiale des précipitations. Les valeurs de  $A$  s'échelonnent entre  $[-2 ; 2]$  où 0 indique une prévision « parfaite ». Des valeurs négatives indiquent une sous-estimation des prévisions (Figure VI-2a).

### VI-2.1.2 La localisation $L$

La localisation  $L$  est la somme de deux termes :  $L_1$  et  $L_2$ .

Le terme  $L_1$  mesure la distance normalisée entre les barycentres  $\mathbf{b}$  du champ global de précipitation observé et prévu tel que :

$$L_1 = \frac{|\mathbf{b}(\mathbb{R}_{mod}) - \mathbf{b}(\mathbb{R}_{obs})|}{d} \quad [\text{VI-3}]$$

où  $d$  est la distance maximale entre deux points de grille du domaine. Les valeurs de  $L_1$  sont comprises dans l'intervalle  $[0 ; 1]$  et  $L_1$  devient nul lorsque les barycentres des deux champs, observé et prévu, coïncident. Le terme  $L_1$  donne une indication sur le déplacement des pluies à l'intérieur du domaine (Figure VI-2b).

Le terme  $L_2$  est estimé au travers de la différence normalisée des distances moyennes entre le barycentre  $\mathbf{b}$  du champ prévu et observé et les ceux des objets  $\mathbf{b}_m$  associés tel que :

$$L_2 = 2 \cdot \frac{|r(\mathbb{R}_{mod}) - r(\mathbb{R}_{obs})|}{d} \quad [\text{VI-4}]$$

$$\text{avec} \quad r(\mathbb{R}) = \frac{\sum_{m=1}^{\mathcal{M}} \mathcal{R}_m \cdot |\mathbf{b}(\mathbb{R}) - \mathbf{b}_m|}{\sum_{m=1}^{\mathcal{M}} \mathcal{R}_m} \quad [\text{VI-5}]$$

où  $\mathcal{R}_m$  est la quantité de précipitation intégrée de l'objet  $m$ . Les valeurs de  $L_2$  s'étendent entre  $[0 ; 1]$ . Le terme  $L_2$  permet de différencier les cas où  $L_1$  est nul mais compensé par la présence d'un nombre différent d'objets entre les deux champs (Figure VI-2c).

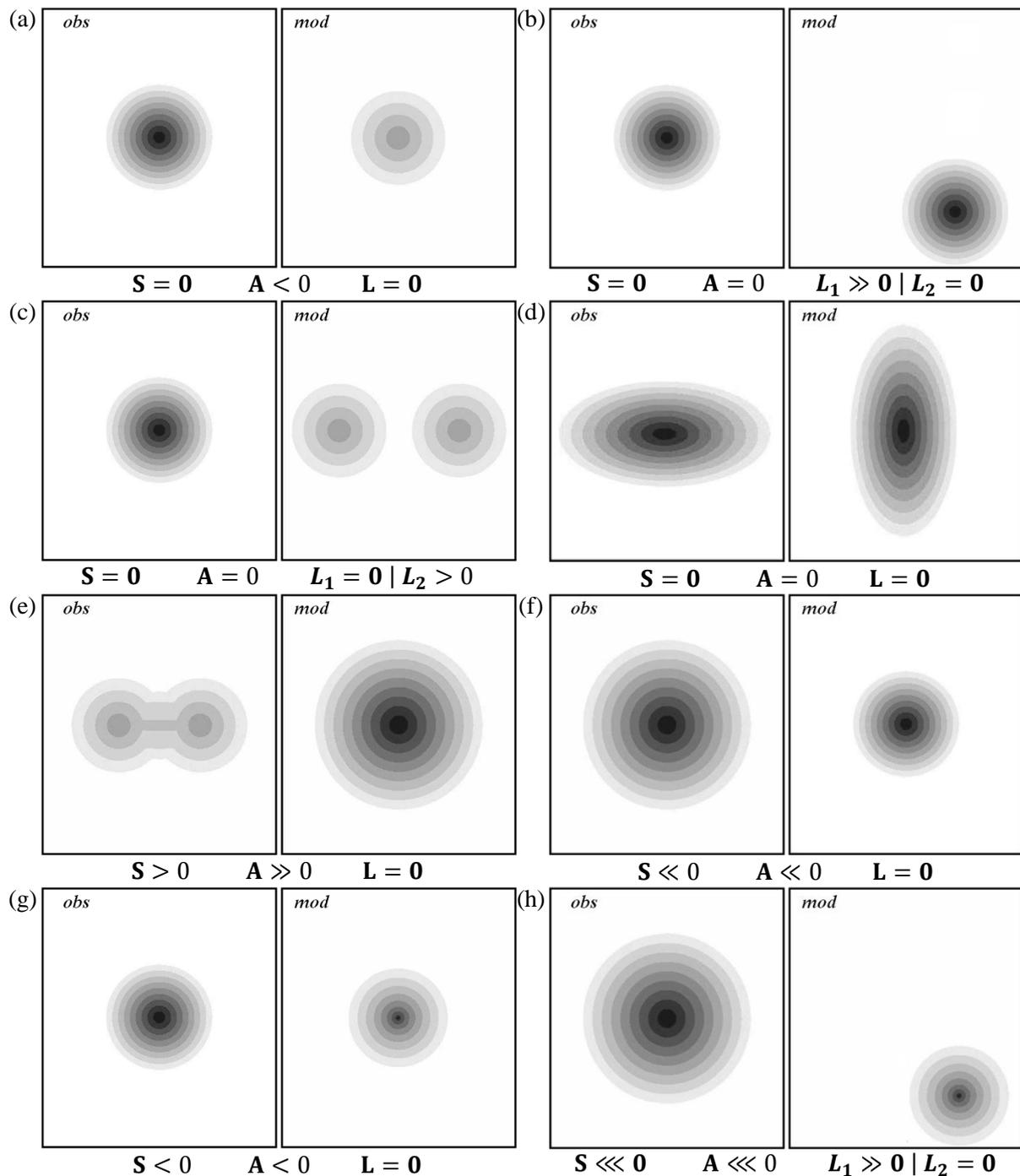


Figure VI-2 : Champ idéalisé de pluie. En comparaison du champ observé (gauche), le champ modélisé (droite) : (a) est sous-estimé, (b) est déplacé, (c) affiche deux objets de faible intensité distincts, (d) a subi une rotation de  $90^\circ$  autour du barycentre, (e) affiche une surface plus grande pour un même barycentre, (f) affiche une surface plus petite pour un même gradient d'intensité, (g) concentre les précipitations et (h) affiche une surface plus petite déplacée et qui concentre les précipitation.

Les contours en nuances de gris, du plus clair au plus foncé, indique le gradient de l'intensité des précipitations du plus faible au plus intense. Sources : Wernli et al. (2008) (images (b, d et h) modifiées).

Les valeurs de ce second critère  $L (= L_1 + L_2)$  sont comprises entre  $[0 ; 2]$ . La valeur 0 est obtenue uniquement pour les cas où les prévisions ont i) un barycentre  $\mathbf{b}$  et ii) une distance moyenne entre le barycentre  $\mathbf{b}$  et ceux des objets  $\mathbf{b}_m$  identiques à ceux des observations. Il faut savoir que  $L$  n'est pas sensible à la rotation des objets autour du barycentre  $\mathbf{b}$  (Figure VI-2d).

### VI-2.1.3 La structure $S$

La structure  $S$  est définie comme étant la différence normalisée des surfaces pondérées moyennes des objets  $\mathfrak{s}$  calculées pour le champ prévu et observé telle que :

$$S = \frac{\mathfrak{s}(\mathbb{R}_{mod}) - \mathfrak{s}(\mathbb{R}_{obs})}{0,5 \cdot (\mathfrak{s}(\mathbb{R}_{mod}) + \mathfrak{s}(\mathbb{R}_{obs}))} \quad [\text{VI-6}]$$

$$\text{avec} \quad \mathfrak{s}(\mathbb{R}) = \frac{\sum_{m=1}^{\mathcal{M}} \mathcal{R}_m \cdot \mathfrak{s}_m}{\sum_{m=1}^{\mathcal{M}} \mathcal{R}_m} \quad [\text{VI-7}]$$

$$\text{et} \quad \mathfrak{s}_m = \frac{\mathcal{R}_m}{\max(\mathcal{R}_m)}$$

Ce troisième critère permet de donner une information globale sur la taille et la forme des objets. Les valeurs prises par  $S$  se distribuent entre  $[-2 ; 2]$ .  $S$  peut devenir important lorsque par exemple le modèle prédit un champ de pluie étendu pour une situation où de petites cellules convectives sont observées (Figure VI-2e). Des valeurs négatives peuvent indiquer par exemple i) un objet prévu plus petit que celui observé (Figure VI-2f), ii) un objet observé et prévu de surfaces identiques mais dont les prévisions concentrent davantage les précipitations (Figure VI-2g) ou iii) une combinaison des possibilités i) et ii) (Figure VI-2h).

## VI-2.2 Discussions

Dans ce qui suit, nous rapportons quelques remarques faites sur la technique SAL trouvées dans différentes études.

Premièrement, les trois critères ( $S$ ,  $A$  et  $L$ ) sont adimensionnées (Wernli *et al.* 2008). Contrairement à la majorité des scores de vérification de type *features-based*, la technique ne requiert pas une correspondance point-à-point du champ entre le champ des observations et celui des prévisions. Les composants ne sont donc pas affectés par la double pénalité (paragraphe VI-1.1). Nous noterons que la détermination de la valeur du seuillage  $R^*$  pour l'identification des objets ne requiert pas l'intervention d'un utilisateur. De plus, sa valeur varie d'une part entre observations et prévisions et d'autre part d'un événement à un autre, ce qui permet de comparer les performances d'un modèle à simuler les pluies de faibles comme

de fortes intensités. Wernli *et al.* (2008) pointent le fait que la définition de  $R^*$  influence significativement le nombre et la taille des objets (et donc sur les valeurs de  $S$  et  $L_2$ ). Un seuillage trop petit conduira à des objets surdimensionnés, à l'inverse un seuillage trop haut élèvera le nombre d'objets de plus petites tailles. Afin d'estimer au mieux cette valeur seuil  $R^*$  (équation [VI-1]), Wernli *et al.* (2009) remplacent  $R^{max}$  par  $R^{0.95}$  (= 95e centile des valeurs des champs) car il est fortement possible que la valeur maximale soit faussée par la présence de valeurs ponctuelles aberrantes (ex : pixel résiduel de très forte valeur).

Dans Wernli *et al.* (2008), la technique a été réalisée avec des modèles à fines échelles spatiales (ex : modèles régionaux) comme sur des modèles globaux. L'application de la technique sur l'ensemble des événements estivaux des années 2001 à 2004 sur le bassin versant de l'Elbe (République tchèque & Allemagne = 97 175 km<sup>2</sup>) a permis de mettre en avant les bonnes performances des deux de modèles. Wernli *et al.* (2009) avouent que les informations obtenues peuvent être ambiguës (par rapport à des scores plus traditionnels) lorsque le domaine d'étude considéré comporte différents systèmes météorologiques (ex : une zone frontale et une convection localisée d'une masse d'air). Les composants SAL résultants affichent alors des valeurs représentatives des champs moyens. De plus, les effets peuvent se compenser d'un système précipitant à un autre (ex : une zone convective sous-estimée et localisée dans une partie du domaine peut être compensée par une zone stratiforme surestimée présente dans une autre partie du domaine) faussant ainsi les interprétations. Les auteurs préconisent alors la découpe du domaine en zone spécifique relative à un phénomène particulier (ex : convectif vs stratiforme comme suggéré par Gilleland *et al.* (2010a)) de sorte de ne pas excéder un domaine supérieur à 500 × 500 km<sup>2</sup>.

En comparaison avec des scores traditionnels, Wernli *et al.* (2009) montrent que la technique affiche de meilleurs scores pour des simulations réalisées à hautes résolutions spatiales. Pour limiter les erreurs liées à des problèmes de synchronisation temporelle, des données journalières sont souvent utilisées. En effet, tout comme l'erreur sur la localisation des structures de pluie, la présence d'un décalage temporel entre observations et prévisions peut pénaliser la vérification. En effet, il est possible que le modèle prévoit correctement les cellules de pluie mais avec un certain retard. La comparaison se faisant sur les images de même date, les scores statistiques sont donc doublement faussés (prévision manquée puis fausse alarme). Pour contrecarrer ce problème, Zimmer et Wernli (2011) proposent d'ajouter une étape supplémentaire en amont du calcul des trois composants SAL. Cette modification permet de déterminer automatiquement, grâce au critère de la localisation  $L$ , un décalage temporel (de plus ou moins 3 heures) sur le champ des prévisions. La variabilité temporelle

des précipitations est ainsi intégrée dans la vérification. Notons que ce dernier point n'a pas été considéré dans les travaux présentés dans ce chapitre.

Finalement, la technique SAL permet de prendre en compte le caractère spatial (voire temporel) des précipitations et de quantifier les erreurs de prévisions en termes de quantité, de localisation et de structure. Répondant à toutes nos attentes, nous l'avons appliquée sur l'un de nos cas d'étude cévenol (cas A, cf. section III-3)<sup>Ⓢ</sup>.

## VI-3 Application de la technique SAL

### VI-3.1 Paramétrisation du modèle

Dans cet exemple, nous avons utilisé le modèle WRF (cf. section V-3) dans sa version v.3.3.1 (Août 2011). Nous listons ici les principales paramétrisations considérées :

- 3 domaines imbriqués (Figure VI-3) : le premier domaine comporte  $120 \times 120$  points de grille espacés de 27 km, le second domaine est caractérisé par  $190 \times 160$  points de grille avec une résolution de 9 km enfin le dernier domaine compte  $160 \times 190$  points de grille de 3 km de résolution (la communication entre les domaines est de type « 2-way »). La grille verticale comprend 52 points non-équidistants variant entre plus de 160 m proche de la surface à près de 1 700 m au sommet ( $\approx 20$  km). La projection utilisée est de type conique conforme de Lambert ;

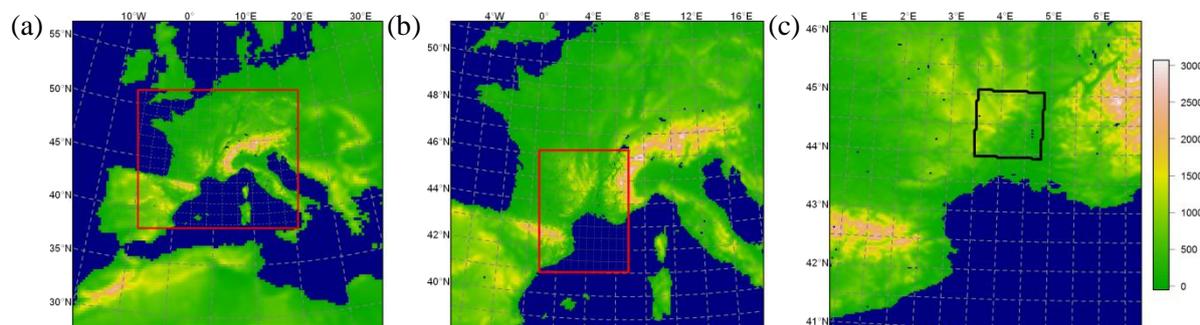


Figure VI-3 : Représentation des trois domaines imbriqués à (a) 27, (b) 9 et (c) 3 km de résolution.

*La topographie modélisée (en m) est représentée par le code couleur. Les rectangles rouges localisent les domaines « fils », le rectangle noir positionne la fenêtre d'observation considérée.*

<sup>Ⓢ</sup> Nous tenons à spécifier que nous avons également testé le modèle WRF sur le cas B (cf. section 0). Cependant, les prévisions obtenues ne nous ont pas donné entière satisfaction. Nous avons donc décidé de faire abstraction de ces tests dans ce chapitre.

- la simulation débute 6 heures avant la période d'observation considérée et est réalisée sur 78 heures consécutives. La résolution temporelle est de 120s avec une sortie toutes les heures ;
- le modèle est initialisé par les ré-analyses du modèle global ERA-INTERIM du Centre Européen de Prévision Météorologique à Moyen Terme (résolutions :  $0.25^\circ \times 0.25^\circ$ , 60 niveaux verticaux) ; la topographie est interpolée à partir de la base de données « United State Geological Survey » à 30 secondes d'arc de résolution ;
- schéma de microphysique = WSM6 (Hong et Lim 2006) ; schéma de convection = Kain-Fritsch (Kain 2004) pour les 2 premiers domaines seulement ; schémas de rayonnement = RRTM (Iacono *et al.* 2008) pour les longues longueurs d'onde et Dudhia (Dudhia 1989) pour les courtes longueurs d'onde ; schéma de surface = Noah (Mitchell 2005) avec 4 couches de sol ; schéma de couche limite planétaire = YSU ((Hong *et al.* 2006).

Les grilles entre le champ des observations et des prévisions devant être identiques, nous avons décidé d'interpoler (via la technique de spline Akima = technique polynomiale de troisième ordre) les prévisions du dernier domaine simulé. Ainsi nous limitons les erreurs d'estimation sur les observations (alors que les erreurs sur les prévisions sont déjà importantes).

### VI-3.2 CAS A : résultats de la technique SAL

Dans cette partie, nous ne redévelopperons pas la situation météorologique du cas d'étude ni le processus permettant de reconstruire le champ spatial des précipitations à partir des observations. Ces informations sont données dans la section III-3.

La Figure VI-4 (a-b) redonne les cartes de pluie cumulées sur le pas de temps de l'événement obtenues respectivement par la technique du KED et du MERG. Les prévisions du modèle WRF cumulées sur la même période sont illustrées dans la Figure VI-4c. Comme pour les observations, nous constatons que les pluies prévues s'organisent sur un axe sud-ouest/nord-est, principalement sur les contreforts du massif cévenol. Les prévisions sont cependant fortement sous-estimées. De plus, le maximum apparaît comme légèrement décalé vers le nord-est. Ces remarques ne sont que qualitatives, mais grâce à la technique SAL, nous espérons pouvoir les quantifier.

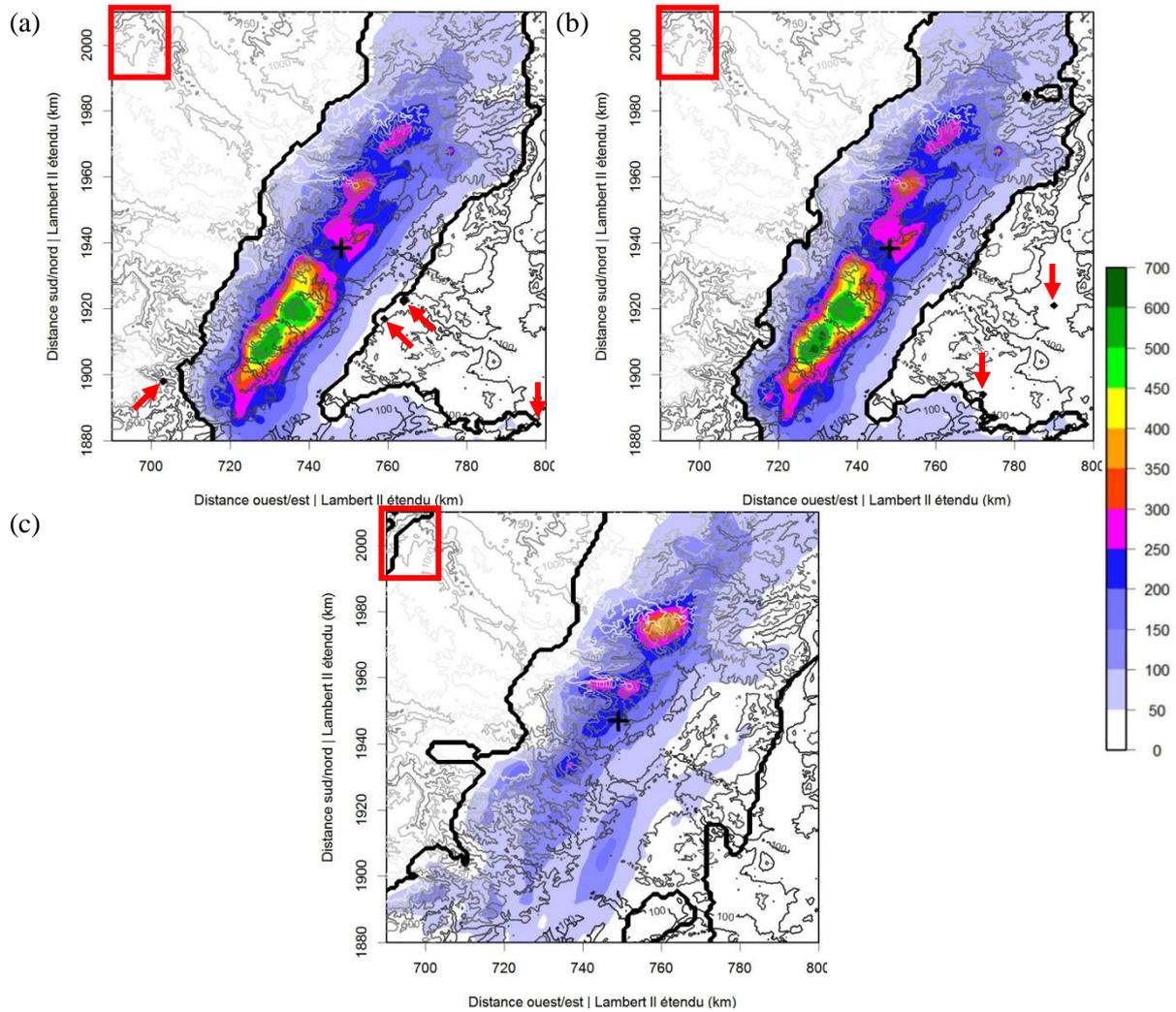


Figure VI-4 : CAS A = Champ de pluie (en mm) interpolé par (a) KED et (b) MERG et (c) prévu par le modèle WRF. (T = 72 h).

La topographie de la région est rajoutée sous forme d'iso-contours en nuances de gris, du plus foncé pour les altitudes proches de la surface de la mer, au plus clair pour les altitudes atteignant les 1 250 m. Les contours noirs délimitent les objets identifiés par la technique SAL, les croix noires indiquent la position des barycentres  $\mathbf{b}$  des champs totaux. Le rectangle rouge localise une cellule de pluie (ou objet) prédite mais non observée. Les flèches rouges positionnent des objets secondaires identifiés sur les deux champs des observations (KED et MERG).

Tableau VI-1 : CAS A = Comparatif des quantités calculées pour la technique SAL pour le champ des observations reconstruit par la technique du KED et du MERG et le champ des prévisions obtenu avec le modèle WRF. (T = 72 h).

	$R^*$	$\mathcal{M}$	$\overline{\mathbb{R}}$	$\mathbf{b}$	$r$	$s$
	valeur de seuillage	nombre d'objets	cumul moyen	position du barycentre	distance moyenne	surface pondérée
	équation [VI-1]				entre $\mathbf{b}$ et $\mathbf{b}_m$	moyenne des objets
					équation [VI-5]	équation [VI-7]
KED	40.9 mm	5	72.1 mm	748.22, 1938.35 km	0.646 km	1472.04
MERG	44.5 mm	3	72.0 mm	748.28, 1938.21 km	0.553 km	1331.10
WRF	25.4 mm	2	56.7 mm	749.04, 1946.88 km	1.681 km	1983.71

Le Tableau I donne certaines quantités calculées pour la technique SAL. Ces valeurs serviront de première comparaison entre le champ KED, MERG et WRF

La valeur seuil  $R^*$  utilisée pour la délimitation des objets dans chacun des champs est établie ici à partir du cumul maximum  $R^{max}$  (et non  $R^{0.95}$  comme suggéré par Wernli *et al.* (2009)). Déjà notifiée dans le sous-paragraphe III-3.4.1, la valeur maximale enregistrée est plus grande dans le champ du MERG que pour le KED. Cette différence se retrouve dans la valeur de  $R^*$  où nous notons une différence de +3.6 mm. Les sous-estimations des prévisions entraînent une importante diminution dans la valeur de seuillage de près de 43 % (38 %) par rapport au MERG (KED). L'application de ces valeurs seuil  $R^*$  permet de façonner les cellules de pluie qui seront alors considérées comme nos objets. Ces derniers sont représentés par les contours noirs dans la Figure VI-4. Le nombre total  $\mathcal{M}$  d'objets ainsi définis varie entre 2 et 5. Avec un seuillage légèrement plus faible, la valeur moyenne  $\bar{\mathbb{R}}$  des objets issus du champ KED est pourtant très similaire à celle du champ MERG. Nous notons encore ici la sous-estimation des prévisions WRF avec un  $\bar{\mathbb{R}}$  abaissé de plus de 15 mm.

Le champ total obtenu par KED et par MERG ont des barycentres  $\mathbf{b}$  (croix noires) très proches pourtant celui du KED apparaît comme légèrement décalé vers le nord-ouest par rapport au barycentre du MERG (plus semblables aux mesures RADAR – *cf.* paragraphe III-3.4). Concernant le champ de prévisions WRF, nous avons évoqué un déplacement manifeste du maximum de précipitation vers le nord-est. En effet, la position de  $\mathbf{b}$  apparaît à plus de 750 m à l'est et de 8.5 km au nord par rapport à ceux des champs des observations. Même si nous dénombrons davantage d'objets secondaires dans le champ des observations (flèches rouges), leurs petites surfaces et leurs faibles valeurs font que la distance moyenne  $r$  avec le barycentre  $\mathbf{b}$  est faible. *A contrario*, la taille non négligeable de la cellule de pluie prévue par WRF en bordure nord-ouest du domaine (rectangle rouge) élève considérablement cette distance moyenne d'un facteur 3.

Également soulevée dans le sous-paragraphe 0, la surface des précipitations observée dans le champ du KED est plus imposante que celle du MERG. Avec une valeur de  $\bar{\mathbb{R}}$  identique, la surface pondérée moyenne des objets  $s$  pour le MERG est plus grande de près de 10 % en comparaison du KED. Malgré les importantes sous-estimations des prévisions, le champ WRF possède une surface plus importante (soit +49 et +35 % par rapport au MERG et KED respectivement).

Tableau VI-2 : CAS A = Comparatif des composants SAL pour la comparaison le champ des observations (KED ou MERG) et celui des prévisions WRF. (T = 72 h).

	S	A	$L_1$	$L_2$
	<i>Structure</i>	<i>Intensité</i>	<i>Localisation globale</i>	<i>Localisation des objets</i>
	<i>équation [VI-6]</i>	<i>équation [VI-2]</i>	<i>équation [VI-3]</i>	<i>équation [VI-4]</i>
KED vs WRF	0.296	-0.239	0.050	0.012
MERG vs WRF	0.394	-0.238	0.051	0.013

Le Tableau VI-2 résume les valeurs des trois composants SAL pour la comparaison d'une part entre KED et WRF et d'autre part entre MERG et WRF. Dans les deux cas, nous notons i) une surestimation de la structure (S), ii) une sous-estimation de la quantité (A) tandis que ii) l'erreur sur la localisation ( $L = L_1 + L_2$ ) paraît minimale. Pour A et S, les erreurs semblent toutefois modérées. L'utilisation du champ MERG pour la vérification du modèle WRF montre des erreurs sur la quantité A (-0.4 %) et la localisation L (+3 %) très similaires à celles obtenues avec l'utilisation du champ KED mais implique des erreurs sur la structure S plus importantes (+33 %).

## VI-4 Autres méthodes de vérification

La technique SAL balaye un grand nombre de points de comparaison entre le champ observé et prévu. Les autres méthodes de vérification apportent des informations différentes quoique complémentaires. Dans ce qui suit nous illustrerons les résultats obtenus pour deux techniques : la vérification de type upscaling et la vérification par des scores statistiques plus traditionnels. Pour chacune, nous commencerons par un rapide rappel sur les estimateurs statistiques utilisés<sup>⊕</sup>.

### VI-4.1 Approche upscaling

#### VI-4.1.1 Les estimateurs statistiques utilisés

Par variable catégoriel, nous sous-entendons un référencement des valeurs continues par catégories. Ces classes sont souvent liées une question comme « est-ce que la quantité de pluie est supérieure ou égale à une valeur seuil  $R^*$  ? ». Ainsi, pour chaque couple de valeurs

<sup>⊕</sup> Une description détaillée peut être trouvée sur le site Internet : <http://www.cawcr.gov.au/projects/verification/>

(observation =  $V$  / prévision =  $\tilde{V}$ ), un comptage de chaque événement (*i.e.*  $V \geq R^*$  ou  $\tilde{V} \geq R^*$ ) est réalisé. Ce décompte est résumé dans une table de contingence telle que :

		Observations		
		OUI, $V \geq R^*$	NON, $V < R^*$	
Prévisions	OUI, $\tilde{V} \geq R^*$	Correct Positif	Fausse Alarme	Total OUI prévus
	NON, $\tilde{V} < R^*$	Manqué	Correct Négatif	Total NON prévus
		Total OUI observés	Total NON observés	Total de valeurs

Dès lors, il est possible d'évaluer :

- le BIAIS qui est défini par :

$$\text{BIAIS} = \frac{\text{Total OUI prévus}}{\text{Total OUI observés}} \quad [\text{VI-8}]$$

Ce score indique si le modèle a tendance à sur-prédire ( $\text{BIAIS} > 1$ ) ou sous-prédire ( $\text{BIAIS} < 1$ ) le nombre d'événements OUI. Ses valeurs sont comprises dans l'intervalle  $[0 ; +\infty[$ , 1 étant une valeur « parfaite ».

- l'ETS (« Equitable Threat Score »), exprimé par :

$$\text{ETS} = \frac{\text{Correct Positif} - \text{CP}^{\text{hasard}}}{\text{Correct Positif} + \text{Manqué} + \text{Fausse Alarme} - \text{CP}^{\text{hasard}}} \quad [\text{VI-9}]$$

$$\text{avec } \text{CP}^{\text{hasard}} = \frac{(\text{Total OUI observés}) \cdot (\text{Total OUI prévus})}{\text{Total de valeurs}}$$

Il permet de préciser dans quelle mesure les prévisions OUI correspondent aux observations OUI en considérant que certaines prévisions « Correctes Positives » peuvent être obtenues par « hasard » (*cf.* facteur  $\text{CP}^{\text{hasard}}$ ). Ses valeurs s'échelonnent entre  $-1/3$  et la valeur « parfaite » de 1, 0 indiquant qu'il n'y a pas de correspondance possible. Des valeurs négatives signifient que les prévisions sont les résultats du pur hasard.

#### VI-4.1.2 Application sur le cas A

La vérification upscaling permet de visualiser à quelles résolutions le modèle donne de meilleurs résultats. Elle considère différentes échelles spatiales de plus en plus larges. Dans l'exemple présenté ci-après nous avons examiné 8 classes pour lesquelles nous avons simplement moyenné les valeurs (observées et prévues) sur des surfaces allant de  $1 \times 1$  pixel à  $110 \times 110$  pixels (l'image RADAR ayant une résolution de  $110 \times 130$  pixels). Nous avons également distingué 10 classes de cumul de pluie (pour  $T = 72$  h) entre 20 et 200 mm. Ces

cumuls sont associés à différentes valeurs seuil  $R^*$  qui seront utilisées pour catégoriser les données. Dès lors, le BIAIS (équation [VI-8]) et l'ETS (équation [VI-9]) ont été établis en fonction du cumul de précipitation et de l'échelle spatiale. Les performances du modèle WRF sont ici comparées avec les deux techniques étudiées (KED et MERG).

Le BIAIS (Figure VI-5) nous indique que les meilleures correspondances entre les observations et les prévisions ( $\Leftrightarrow$  BIAIS = 1 = rouge foncé) s'établissent pour les plus petites valeurs de cumuls et les plus grandes échelles spatiales. Pour un seuillage à 20 ou 60 mm, les valeurs du BIAIS sont proches de cette perfection même pour les plus petites surfaces de voisinages. Pour un seuillage de 40 mm (semblable à celui considéré dans la technique SAL – Tableau VI-1) nous constatons que le modèle possède un nombre de pixel avec de la pluie supérieur à celui des champs des observations (BIAIS > 1) ; la surface couverte par de la pluie prévue par le modèle est plus imposante. Pour une valeur seuil de 80 mm, cette surface devient sous-estimée (BIAIS < 1). Enfin, pour des intensités supérieures à 140 mm, nous observons une dégradation importante du BIAIS. Le modèle n'a donc pas la capacité à prévoir les très forts cumuls. Entre la technique du KED et celle du MERG, nous notons des valeurs de BIAIS légèrement plus fortes pour le MERG (*i.e.* de moins bons scores).

L'ETS (Figure VI-6) permet d'exclure les prévisions considérées comme correctes positives et obtenues par hasard. *Comme pour le BIAIS*, les meilleurs scores ( $\Leftrightarrow$  ETS = 1 ou rouge foncé) sont obtenus pour les plus faibles cumuls et les plus grandes échelles spatiales. Nous noterons les bonnes performances du modèle (par comparaison avec KED ou MERG) pour un seuillage de 60 mm et sur presque toutes les tailles de voisinage. À partir d'un cumul de 140 mm, les valeurs de l'ETS montrent une qualité moindre dans les prévisions (*i.e.* ETS inférieur à 0.3) avec une dégradation progressive avec l'augmentation de l'échelle spatiale. Nous noterons que pour des seuillages inférieurs à 140 mm, les valeurs de l'ETS sont légèrement meilleures (soit des valeurs plus fortes) pour le KED que pour le MERG.

Nous retiendrons finalement que : i) le modèle n'est pas capable, même par hasard, de prévoir les forts cumuls et ii) les meilleures performances du modèle sont atteintes pour un seuillage de 20 mm ou 60 mm (en excluant un voisinage de 10 pixels qui présente une forte valeur de BIAIS).

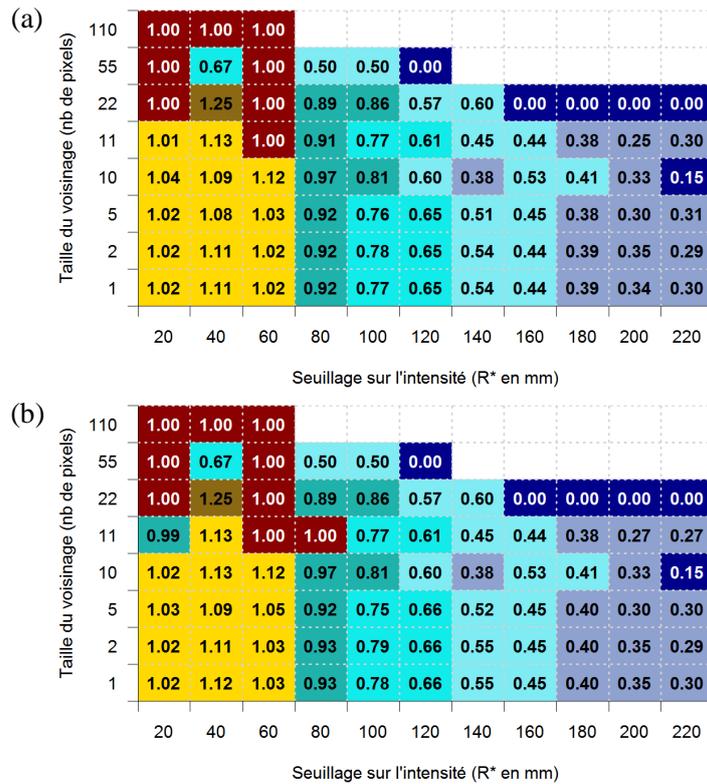


Figure VI-5 : CAS A = BIAIS en fonction de l'échelle spatiale (1 pixel = 1×1 km) et de la valeur de seuillage sur l'intensité pour la comparaison du champ des prévisions WRF avec le champ des observations obtenu avec la technique (a) du KED et (b) du MERG.

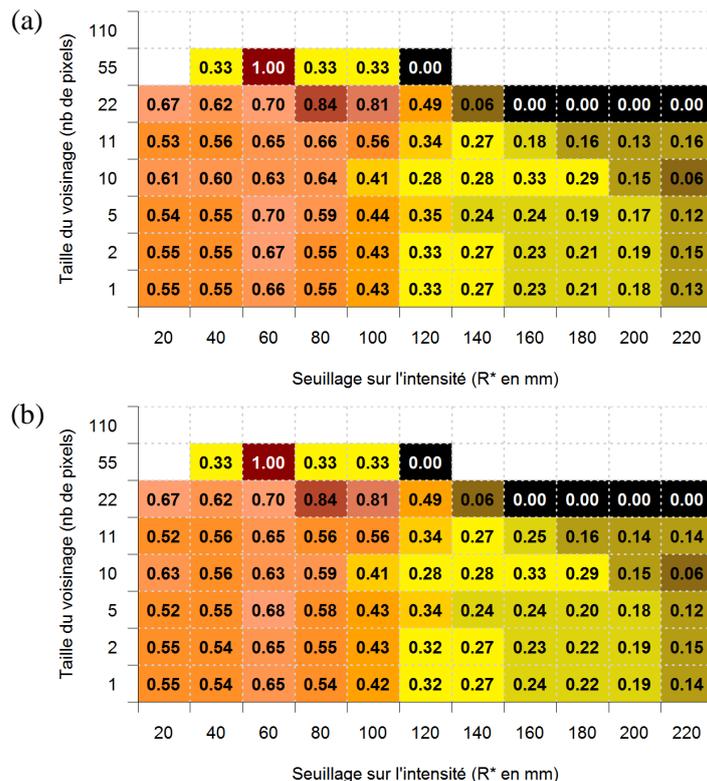


Figure VI-6 : CAS A = ETS en fonction de l'échelle spatiale (1 pixel = 1×1 km) et de la valeur de seuillage sur l'intensité pour la comparaison du champ des prévisions WRF avec le champ des observations obtenu avec la technique (a) du KED et (b) du MERG.

## VI-4.2 Approche traditionnelle

### VI-4.2.1 Les estimateurs statistiques utilisés

Notons  $V_i$  la valeur observée et  $\tilde{V}_i$  la valeur prévue par le modèle de la quantité de pluie au point de grille  $i$ . Considérons que le domaine d'étude comporte  $N$  points de grille, alors :

- l'erreur moyenne (ME) correspond à la moyenne des écarts entre les valeurs estimées et observées. Elle est définie par :

$$ME[\text{mm}] = \frac{1}{N} \sum_{i=1}^N (\tilde{V}_i - V_i) \quad [\text{VI-10}]$$

- l'écart type mesure de la dispersion des valeurs autour de la moyenne. Il est calculé indépendamment pour les observations et les prévisions, tel que :

$$\sigma_{obs}[\text{mm}] = \sqrt{\sum_{i=1}^N (V_i - \bar{V})^2} \quad \text{et} \quad \sigma_{mod}[\text{mm}] = \sqrt{\sum_{i=1}^N (\tilde{V}_i - \bar{\tilde{V}})^2} \quad [\text{VI-11}]$$

avec  $\bar{V}$  et  $\bar{\tilde{V}}$  les moyennes des valeurs observées et prévues.

- le coefficient de corrélation renseigne sur la phase (ou relation linéaire) entre le champ des prévisions et celui des observations. Il est donnée par :

$$r = \frac{\sum_{i=1}^N (\tilde{V}_i - \bar{\tilde{V}}) \cdot (V_i - \bar{V})}{\sigma_{obs} \cdot \sigma_{mod}} \quad [\text{VI-12}]$$

avec  $\bar{V}$  et  $\bar{\tilde{V}}$  les moyennes des valeurs observées et prévues.

- la différence centrée quadratique moyenne (CRMSD) est exprimée par :

$$CRMSD^2[\text{mm}^2] = \frac{1}{N} \sum_{i=1}^N \left( (\tilde{V}_i - \bar{\tilde{V}}) - (V_i - \bar{V}) \right)^2 \quad [\text{VI-13}]$$

### VI-4.2.2 Application sur le cas A

La vérification spatiale peut également être réalisée avec des scores statistiques plus traditionnels comme l'erreur moyenne (équation [VI-10]), l'écart-type (équation [VI-11]) ou le coefficient de corrélation (équation [VI-13]). Ceux-ci donnent une information sur la qualité des prévisions en considérant le champ global. Taylor (2001) a élaboré un nouveau critère qui permet de relier tous ces scores et ainsi les visualiser dans un unique diagramme. Les résultats obtenus pour une telle vérification sont présentés dans la Figure 7 (pour faciliter la lecture, les valeurs sont données dans le tableau de droite).

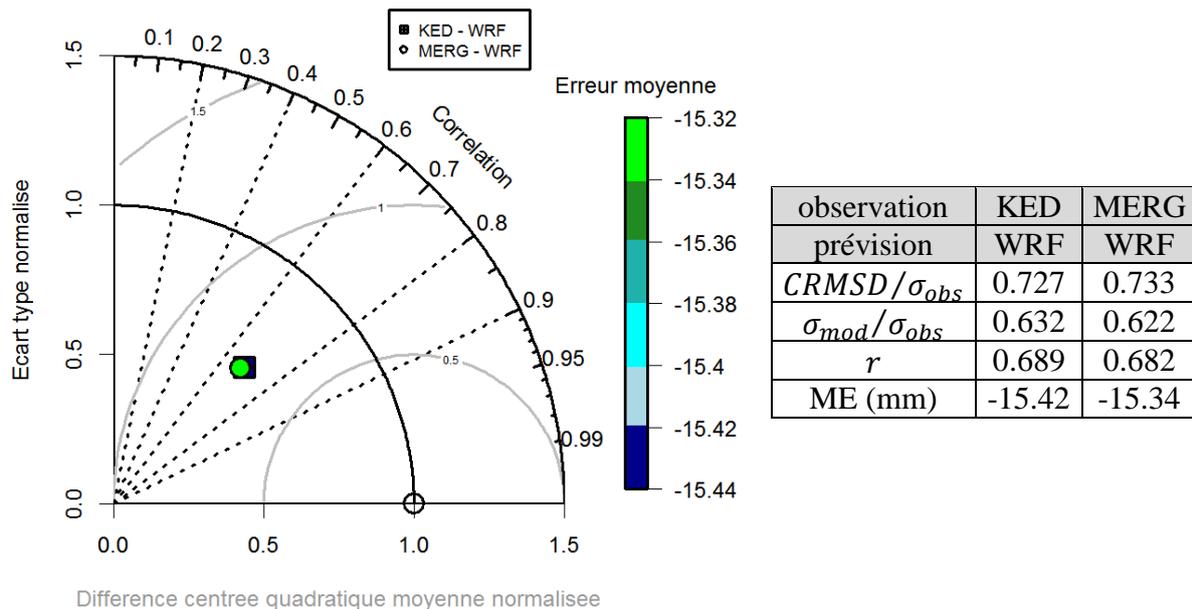


Figure VI-7 : Diagramme de Taylor pour le champ total obtenu par KED et par MERG comparé au champ WRF.

Ce type de diagramme permet de visualiser le coefficient de corrélation ( $r =$  lignes noires en pointillé), l'écart type des prévisions normalisé par l'écart type des observations ( $\sigma_{mod}/\sigma_{obs} =$  cercles noirs tracés pour 1 et 1.5), l'erreur moyenne (ME = code couleur) et la différence centrée quadratique moyenne (CRMSD) normalisée par l'écart type des observations ( $CRMSD/\sigma_{obs} =$  cercles gris dessinés pour 0.5, 1 et 1.5 mm). Plus ces scores se rapprochent du point « idéal » ( $\Leftrightarrow r = 1$ ,  $CRMSD/\sigma_{obs} = 0$  et  $\sigma_{mod}/\sigma_{obs} = 1$ ) avec une ME de 0, meilleures sont les prévisions.

Il en ressort que les valeurs des prévisions WRF sont presque 1.6 fois moins dispersées que les observations (*cf.* écart type) malgré une forte CRMSD (équation [VI-13]). La correspondance spatiale des champs semble correcte (coefficients de corrélation proches de 0.7). Enfin, nous noterons l'importante sous-estimation (*cf.* erreur moyenne) des prévisions. Ici encore, nous pouvons dire que globalement le champ de prévision WRF s'accorde d'autant plus avec le champ KED qu'avec celui du MERG. En effet, l'écart type normalisé et la ME sont certes plus importants (respectivement +1.6 et +0.5 %) pour le KED par rapport au MERG, la CRMSD normalisée et le coefficient de corrélation sont pourtant meilleurs (-0.8 et +1 % respectivement). Le point associé à la comparaison du champ du KED et de WRF (carré) est dès lors très légèrement plus proche du point idéal que celui de la comparaison MERG et WRF (cercle).

Les conclusions des diverses méthodes de vérification sont toutes différentes. L'utilisation de scores traditionnels nous renseigne sur les erreurs faites sur la globalité des champs, l'origine de ces écarts n'est donc pas connue. Grâce aux techniques de type upscaling, il est possible de déterminer à quelle échelle spatiale et pour quelle intensité le modèle affiche ses meilleures prévisions. De nombreux scores catégoriels peuvent être calculés comme i) le BIAIS qui nous indique indirectement les erreurs sur la surface du champ de pluie observé et prévu ou ii) l'ETS qui prend en compte le fait que les prévisions peuvent être considérées comme correctes par « hasard ». Nous estimons que la technique SAL est un outil de vérification simple à codifier (Rossa *et al.* 2008) qui permet d'obtenir des informations plus pertinentes notamment sur la structure (nombre et volume des cellules de pluie détectées), la quantité (maximale et moyenne) et la localisation des cellules de pluie (considérés dans leur totalité ou individuellement).

Gilleland *et al.* (2010a) indiquent que cette approche SAL est incluse dans l'outil de vérification développé par la communauté HIRLAM (« High-Resolution Limited-Area Model » ; <http://hirlam.org>). Les auteurs précisent également qu'elle est utilisée pour la vérification opérationnelle (interne) au Finnish Meteorological Institute. Wernli *et al.* (2009) complètent cette liste d'applications avec la vérification opérationnelle du modèle COSMO (« Consortium for Small-Scale Modeling » : <http://www.cosmo-model.org>) au German Weather Service, des recherches menées en Suisse ou en Allemagne, la vérification des modèles comme les estimations satellitaires. Enfin, Vincendon *et al.* (2011) ont employé cette technique pour tester les performances du modèle AROME (« Application de la Recherche à l'Opérationnel à Méso-Échelle » : <http://www.cnrm-game.fr/spip.php?article120>) à Météo-France.





# Conclusions générales

Depuis quelques années, les changements climatiques sont aux cœurs de nos préoccupations. En effet, il apparaît que les situations météorologiques extrêmes sont de plus en plus fréquentes. Les conséquences liées notamment aux précipitations sur l'environnement, les activités humaines voire sur les vies elles-mêmes deviennent peu à peu des enjeux prioritaires pour les décideurs (politiciens, grand groupes industriels, etc.). De nouveaux outils informatiques ont récemment été développés par les acteurs (chercheurs, ingénieurs, etc.) pour mieux connaître ces événements et espérer les anticiper et gérer les risques associés. Aujourd'hui, il est devenu essentiel de pouvoir déterminer avec précision l'étendue spatiale et temporelle des pluies. La demande d'estimations quantitatives des précipitations à hautes résolutions sur des zones locales est incontestablement en augmentation. C'est dans cette problématique que s'inscrivent les travaux présentés dans ce manuscrit. Cette thèse s'oriente autour de deux thématiques. La première se consacre à l'étude des moyens d'observations disponibles et la détermination de la meilleure cartographie possible de la pluie au sol. La seconde vise la mise en place d'une méthodologie fiable permettant de comparer observations et prévisions numériques.

Le Chapitre I présente un état de l'art sur les principales techniques de mesure de la pluie. Depuis plusieurs siècles, le pluviomètre est le standard des instruments de mesure. Aujourd'hui entièrement automatisé, il permet de donner avec une bonne précision la quantité de la pluie tombée pendant un intervalle de temps en un endroit donné. Avec l'avènement des technologies, des instruments permettant de localiser spatialement les précipitations ont vu le jour. Au sol ou embarqués, les RADAR météorologiques constituent la nouvelle référence dans la mesure des pluies, comblant ainsi les lacunes des mesures ponctuelles des pluviomètres. Basées sur le principe de télédétection, ces mesures sont toutefois entachées par de nombreuses sources d'erreurs (atténuation du signal, présence d'échos parasites, variabilités spatio-temporelles des précipitations, approximations employées, etc.) d'autant plus grandes en régions montagneuses.

Afin de tirer meilleur parti de ces deux principaux instruments, nous nous proposons de fusionner leurs données. Pour ce faire, nous avons opté pour une approche géostatistique. La méthode la plus répandue est le krigeage qui consiste en une estimation linéaire sans biais

et à variance minimale. Ce choix provient d'un souci de réalisme puisque c'est l'une des seules méthodes d'interpolation spatiale à prendre en compte la dépendance spatiale du phénomène observé. Comme énumérées dans le 0, il existe plusieurs variantes du krigeage, plus ou moins complexes dont certaines permettent d'utiliser diverses informations. Nous nous sommes focalisés sur deux techniques en particulier : le krigeage avec dérive externe (KED) et la fusion conditionnée (MERG). La principale différence repose sur le fait que le MERG utilise de façon indépendante les données RADAR pour conditionner l'interpolation des observations pluviométriques alors que les poids d'interpolation attribués aux pluviomètres lors du KED dépendent des mesures RADAR.

Ces deux techniques combinant les observations pluviométriques et mesures RADAR ont fait leur preuve à travers différentes études réalisées ces dernières années, ce quelle que soit la région géographique concernée, les événements considérés ou encore le réseau d'observations. Dans le cadre de nos travaux, nous avons investigué deux domaines d'étude distincts afin d'évaluer les performances de ces deux techniques pour mieux appréhender leurs avantages et inconvénients. L'objectif premier étant d'obtenir une estimation du champ de pluie la plus proche de la réalité.

Les premières études ont concerné la région Cévennes-Vivarais (0). Au travers de deux épisodes de pluies très intenses survenus en 2008, nous avons montré l'importance de prendre en compte les mesures RADAR dans les procédures d'interpolation des observations pluviométriques. Les estimations par la technique du KED et du MERG donnent des résultats très satisfaisants. Pour les deux études, nous observons une conservation des structures hétérogènes du champ de pluie observées par le RADAR et un réajustement des quantités pour être plus proches des observations pluviométriques. Quantitativement parlant, les erreurs d'estimations sont minimisées par rapport aux seules mesures RADAR ou aux interpolations des observations pluviométriques seules. Ce constat est d'autant plus vrai que la fenêtre temporelle considérée est grande. La très grande similarité entre les deux techniques ne nous a pas permis de vraiment les dissocier si ce n'est une légère supériorité (visuellement et quantitativement) du MERG face au KED. Nous avons surtout remarqué que l'influence des mesures RADAR est plus importante dans la technique du MERG que celle du KED. Nous avons également montré que la qualité des estimations par KED ou MERG était peu dépendante de la qualité des estimations RADAR ou de la densité de pluviomètres disponibles pour l'interpolation n'a que peu d'impact.

La seconde région étudiée se focalise sur le bassin de Clermont-Ferrand et de son agglomération de communes (Chapitre IV). Au travers de ce domaine bien moins

instrumenté, plus petit spatialement et surtout pour lequel le réseau principal de pluviomètre ne couvre pas la totalité de la surface d'étude, nous avons testé à l'extrême la technique du KED et du MERG. Avant cela, nous avons dû nous pencher sur la détermination du variogramme paramétrique nécessaire aux deux techniques géostatistiques. Nous avons alors opté pour une approche dite « climatologique » qui permet de le définir une seule et unique fois (basée sur Lebel et Bastin 1985). Il s'est avéré que les paramètres estimés (portée, effet de pépité et palier) affichent un lien certain avec les fenêtres temporelles considérées.

Nous avons choisi d'étudier deux systèmes précipitants survenus au cours de l'année 2011. Ici encore, les deux techniques (KED et MERG) présentaient un caractère prometteur. Cependant, la petitesse de la couverture du réseau d'observations n'a pas toujours permis de reproduire correctement la structure des précipitations en bordure de domaine. Ce problème peut toutefois être contourné grâce à la technique du MERG qui permet d'intégrer des observations situées hors de la zone d'étude. Au final, le champ reconstruit est visuellement amélioré. La particularité de ce second domaine d'étude nous a permis d'affirmer que la technique du KED accorde moins d'importance aux mesures RADAR. Outre les très hautes résolutions spatiale et temporelle du RADAR, nous avons pu pointer quelques soucis sur les mesures (atténuation du signal et présence de pixels résiduels qui affichent de très forts cumuls) entachant alors le processus de reconstruction du champ de pluie. De plus, pour limiter les effets liés à cette mauvaise synchronisation des données entre les deux instruments, nous avons dû limiter la procédure de validation à l'échelle de l'évènement.

Après s'être consacrés sur l'étude des moyens d'observation de la pluie et de leur combinaison, nous nous sommes intéressés à la modélisation de cette grandeur physique extrêmement variable spatialement et temporellement. Au cours de nos travaux, nous avons été amenés à travailler avec le modèle WRF (« Weather Research et Forecasting »), nouveau standard de simulation de l'atmosphère météorologique à méso-échelle. Le Chapitre V décrit brièvement les principales caractéristiques de ce modèle libre. Nous présentons également une étude préliminaire sur sa capacité à reproduire la pluie. Il faut savoir qu'il n'existe pas une unique description de la physique des nuages et des précipitations mais plusieurs paramétrisations qui permettent de simplifier la résolution des équations du modèle. Il est donc parfois difficile de faire un choix parmi un large nombre de possibilités. Au travers d'une simulation idéalisée, nous nous sommes focalisés sur la sélection d'un schéma de microphysique qui gouverne le cycle de vie du nuage et des précipitations. Cette étude permet de nous assurer une description réaliste des phénomènes précipitants dans un cas réel.

Le second enjeu de ces travaux de thèse s'est axé sur la vérification des prévisions de la pluie du modèle. Malgré un développement très récent (inférieur à 10 ans), de nombreux acteurs ont élaboré des outils permettant de quantifier spatialement les erreurs de prévisions. Ces méthodes sont revues dans le 0. C'est dans ce dernier chapitre que nous affichons alors l'importance accordée à la reconstruction d'un champ spatial de pluie à partir des observations car la nature et l'exactitude du champ de pluie observé ont un impact majeur sur les résultats de la vérification (Rossa *et al.* 2008; Casati *et al.* 2008). Après réflexion, notre intérêt s'est porté sur la technique développée par Wernli *et al.* (2008) qui renseigne sur la bonne ou mauvaise prévision de la structure, l'intensité et la localisation des systèmes précipitants. Nous avons mis en exergue les performances de la technique au travers d'une application sur un cas réel. La confrontation des résultats obtenus avec ceux de méthodes plus traditionnelles nous a permis de mettre en avant les atouts de la technique choisie.

## Perspectives

La méthode de vérification étant établie, il est maintenant possible de tester les performances du modèle WRF. Nous estimons qu'il serait judicieux de vérifier les prévisions pour différentes configurations dans l'espoir de trouver celle qui permettra de se rapprocher le plus des observations. Cette étape devra concerner la phase d'initialisation du modèle (choix des bases de données, définition des domaines et des résolutions des simulations) ou le choix des différentes paramétrisations proposées par la partie physique du modèle. Ces travaux pourront éventuellement s'orienter vers une étude sur l'apport des techniques d'assimilation de données. Enfin, pour compléter cette évaluation, nous pourrions établir une comparaison avec d'autres modèles météorologiques. Une étude de la qualité des prévisions sur de plus longues séries temporelles peut également être envisagée grâce à la base de données disponible pour la région clermontoise. Cette étude permettrait alors de distinguer les performances du modèle WRF selon les saisons ou bien le type de situation météorologique. Nous pourrions également comparer les résultats obtenus avec différentes fenêtres temporelles. Cette étape nécessitera alors la mise en place du critère sur les erreurs de timing comme celui proposé par exemple par Zimmer et Wernli (2011).

Avant cela, un important travail devra alors être fait en amont pour améliorer les estimations des champs de précipitation issus des observations. Concernant les données pluviométriques, bien que nous ayons mis en place un QC, celui-ci s'avère être quelque peu

sommaire. Nous proposons d'appliquer un test plus robuste (et donc plus compliqué à programmer) comme celui proposé par You *et al.* (2007). L'essentiel du travail portera sur les mesures RADAR du bassin clermontois. Des efforts devront être notamment faits sur l'élimination des pixels résiduels, des effets de masques, l'interprétation des mesures faites en altitude et ramenées à des estimations de pluie au sol. Cette étape pourra notamment s'appuyer sur les travaux de Tridon (2011) ou sur les algorithmes de QC proposés par Szturc *et al.* (2012). Dans le cadre du projet collaboratif auquel s'inscrit cette thèse, des travaux ont été menés en parallèle sur les estimations RADAR. Ils portent notamment sur l'optimisation de la loi  $Z-R$  avec une différenciation selon la saison, l'intensité des pluies ou encore la fenêtre temporelle (Mioche et Van Baelen 2013). Ces résultats pourront être intégrés dans le cadre de la reconstruction d'un champ de pluie à partir des observations.

Nous suggérons ensuite de revenir sur l'identification du variogramme paramétrique. En effet, il existe des techniques automatiques de modélisation du variogramme expérimental, décrites comme rapides et robustes (Velasco-Forero *et al.* 2009; Schiemann *et al.* 2011). Leur fonctionnement, basé sur des transformées de Fourier, permet de s'affranchir du choix des différents paramètres cruciaux (modèle, portée, effet de pépite et palier). Ces nouveaux variogrammes, dits non-paramétriques, semblent être particulièrement appréciés d'un point de vue opérationnel. Enfin, étant donnée la relation entre le relief et les précipitations, l'application de la technique du KED (et du MERG) pourra intégrer des données de topographie en guise de deuxième variable secondaire (en plus des mesures RADAR).

En mars 2012 et mai 2012, le LaMP a acquis deux RADAR en bande X (dont l'un permet de balayer l'atmosphère sur différents angles d'élévation). Ces deux RADAR ont fait leurs preuves durant la première Période d'Observation Spéciale (SOP1) de la campagne HyMeX (« HYdrological cycle in Mediterranean EXperiment », <http://www.hymex.org>) qui s'est tenu dans les Cévennes à l'automne 2012. Par ailleurs, dans le cadre du projet collaboratif, un partenaire privé a investi dans un RADAR en bande X qui est aujourd'hui installé à Chappes, une commune située au nord-est de l'agglomération clermontoise. L'acquisition et le traitement de ces données supplémentaires permettraient de couvrir une plus grande surface. Enfin, nous suggérons l'acquisition de nouveaux pluviomètres dont le choix de leurs emplacements devra être réfléchi en lien avec le variogramme expérimental.

Les résultats obtenus laissent entrevoir un avenir prometteur. Un suivi des précipitations aussi développé permettrait par exemple : d'élaborer un système local d'alerte aux crues et inondations pour les cours d'eau non surveillés par l'État (Ardilouze 2012),

d'apporter une aide soit aux professionnels dans l'organisation de leurs activités (ex : planification des traitements phytosanitaires ou des temps d'irrigation pour les agriculteurs, planification des travaux d'un chantier en BTP, etc.) soit aux collectivités locales dans la gestion de leurs activités quotidiennes (ex : arrosage des jardins publics, gestion des réseaux d'évacuation et d'assainissement, organisation de manifestations sportives ou culturelles, etc.). Les perspectives à ces travaux de thèse sont donc immenses.



# BIBLIOGRAPHIE

- Agence Nationale des Fréquences, 2005: *Perturbations du fonctionnement des radars météorologiques par les éoliennes*. Commission Consultative de la Compatibilité Electromagnétique, 50 pp. Rapport disponible à l'adresse : <http://www.anfr.fr/fileadmin/mediatheque/documents/etudes/Rapport%20perturbations%20fonctionnement%20radars%20meteorologiques%20par%20eoliennes.pdf>.
- Ahijevych, D., E. Gilleland, B. G. Brown, and E. E. Ebert, 2009: Application of spatial verification methods to idealized and NWP-gridded precipitation forecasts. *Weather and Forecasting*, **24**, 1485–1497.
- Alexander, G. D., J. A. Weinman, and J. L. Schols, 1998: The Use of Digital Warping of Microwave Integrated Water Vapor Imagery to Improve Forecasts of Marine Extratropical Cyclones. *Monthly Weather Review*, **126**, 1469–1496.
- Andrieu, H., and J. D. Creutin, 1995: Identification of Vertical Profiles of Radar Reflectivity for Hydrological Applications Using an Inverse Method. Part I: Formulation. *Journal of Applied Meteorology*, **34**, 225–239.
- Ardilouze, C., 2012: Discussion Personnelle du 02 Juillet 2012.
- Atger, F., 2001: Verification of intense precipitation forecasts from single models and ensemble prediction systems. *Nonlinear Processes in Geophysics*, **8**, 401–417.
- Audard-Vincendon, B., 2010: Modélisation hydrométéorologique couplée - Crues rapides - Cévennes-Vivarais - Prévision d'ensemble. Université Paul Sabatier - Centre National de Recherches Météorologiques (GAME), 254 pp.
- Baillargeon, S., 2005: *Le krigeage: revue de la theorie et application à l'interpolation spatiale de données de précipitations*. Université de Laval, Canada, 137 pp. [http://www.mat.ulaval.ca/fileadmin/Pages\\_personnelles\\_des\\_profs/lpr/EtudesGraduees/SBaillargeon.pdf](http://www.mat.ulaval.ca/fileadmin/Pages_personnelles_des_profs/lpr/EtudesGraduees/SBaillargeon.pdf) (Accessed August 24, 2010).
- Bastin, G., B. Lorent, C. Duqué, and M. Gevers, 1984: Optimal Estimation of the Average Areal Rainfall and Optimal Selection of Rain Gauge Locations. *Water Resources Research*, **20**, 463–470.
- Bergeron, T., 1965: On the Low-level Redistribution of Atmospheric Water Caused by Orography. *Proceedings - Supplement*, International Conference on Cloud Physics, Tokyo - Sapporo-shi, Japan, 96–100.
- Berne, A., G. Delrieu, J.-D. Creutin, and C. Obled, 2004: Temporal and spatial resolution of rainfall measurements required for urban hydrology. *Journal of Hydrology*, **299**, 166–179.
- Boudevillain, B., 2011: Discussion personnelle du 17 Mai 2011.
- Bouilloud, L., G. Delrieu, B. Boudevillain, and P.-E. Kirstetter, 2010: Radar rainfall estimation in the context of post-event analysis of flash-flood events. *Journal of Hydrology*, **394**, 17–27.

- Bousquet, O., 2009: Dynamical and microphysical properties of high impact orographic mesoscale convective systems from high resolution operational multiple-Doppler and polarimetric radar data. *Session 6B, Severe Weather II*, 34th Conference on Radar Meteorology, Williamsburg, VA, USA.
- Brandes, E. A., 1975: Optimizing Rainfall Estimates with the Aid of Radar. *Journal of Applied Meteorology*, **14**, 1339–1345.
- Briggs, W. M., and R. A. Levine, 1997: Wavelets and Field Forecast Verification. *Monthly Weather Review*, **125**, 1329–1341.
- Brocheton, F. and Coauthors, 2012: *Développement d'une plate-forme opérationnelle pour la gestion des risques environnementaux*. NUMTECH - WEATHER MEASURES - AGAETIS - Laboratoire de Météorologie Physique - Laboratoire de Mathématiques - Laboratoire Informatique, Modélisation et Optimisation des Systèmes, Aubière, France, 71 pp.
- Casati, B., 2010: New Developments of the Intensity-Scale Technique within the Spatial Verification Methods Intercomparison Project. *Weather and Forecasting*, **25**, 113–143.
- Casati, B., G. Ross, and D. B. Stephenson, 2004: A new intensity-scale approach for the verification of spatial precipitation forecasts. *Meteorological Applications*, **11**, 141–154.
- Casati, B. and Coauthors, 2008: Forecast verification: current status and future directions. *Meteorological Applications*, **15**, 3–18.
- Choisnel, E., and D. Payen, 1988: Les climats de la France. *La Recherche - Supplément*, **201**.
- Civiate, M., and F. Mandel, 2008: La mesure de la hauteur des précipitations. Fiche descriptive disponible à l'adresse : <http://comprendre.meteofrance.com/content/2009/1/20343-48.pdf>.
- Collombat, P.-Y., 2012: *Se donner les moyens de ses ambitions : les leçons des inondations du Var et du sud-est de la France*. Sénat, 388 pp. Rapport d'information disponible à l'adresse : <http://www.senat.fr/rap/r11-775/r11-775.html>.
- Cressie, N. A. C., 1993: *Statistics for spatial data*. John Wiley & Sons, Inc., New York, 928 pp.
- Davis, C., B. Brown, and R. Bullock, 2006: Object-Based Verification of Precipitation Forecasts. Part I: Methodology and Application to Mesoscale Rain Areas. *Monthly Weather Review*, **134**, 1772–1784.
- Davis, C. A., B. G. Brown, R. Bullock, and J. Halley-Gotway, 2009: The Method for Object-Based Diagnostic Evaluation (MODE) Applied to Numerical Forecasts from the 2005 NSSL/SPC Spring Program. *Weather and Forecasting*, **24**, 1252–1267.
- Delobbe, L., 2006: *Estimation des précipitations à l'aide d'un radar météorologique*. Institut Royal Météorologique de Belgique, Bruxelles, 48 pp.
- Delobbe, L., G. Bastin, P. Dierickx, E. Goudenhoofd, G. Leclercq, L. Moens, and M. Thunus, 2008: Evaluation of several radar-gauge merging techniques for operational use in the Walloon region of Belgium. Proceedings of WRaH 2008, Grenoble, France.

- Delrieu, G., L. Hucke, and J. D. Creutin, 1999: Attenuation in Rain for X- and C-Band Weather Radar Systems: Sensitivity with respect to the Drop Size Distribution. *Journal of Applied Meteorology*, **38**, 57–68.
- Delrieu, G., B. Boudevillain, J. Nicol, B. Chapon, P.-E. Kirstetter, H. Andrieu, and D. Faure, 2009: Bollène-2002 Experiment: Radar Quantitative Precipitation Estimation in the Cévennes–Vivarais Region, France. *Journal of Applied Meteorology and Climatology*, **48**, 1422–1447.
- DREAL - Auvergne, Serveur de Situation Hydrologique de la Direction Régionale de l'Environnement -AUVERGNE. <http://www.eaueauvergne.fr/> (Accessed January 28, 2013).
- Dubois, S., 2012: Un orage brutal met hors-service le Panoramique des Dômes : récit et témoignages en vidéo. *La Montagne*, May 27.
- Ducrocq, V., O. Nuissier, D. Ricard, C. Lebeaupin, and T. Thouvenin, 2008: A numerical study of three catastrophic precipitating events over southern France. II: Mesoscale triggering and stationarity factors. *Quarterly Journal of the Royal Meteorological Society*, **134**, 131–145.
- Dudhia, J., 1989: Numerical Study of Convection Observed during the Winter Monsoon Experiment Using a Mesoscale Two-Dimensional Model. *Journal of the Atmospheric Sciences*, **46**, 3077–3107.
- Ebert, E. E., 2008: Fuzzy verification of high-resolution gridded forecasts: a review and proposed framework. *Meteorological Applications*, **15**, 51–64.
- Ebert, E. E., 2009: Neighborhood Verification: A Strategy for Rewarding Close Forecasts. *Weather and Forecasting*, **24**, 1498–1510.
- Ebert, E. E., and J. L. McBride, 2000: Verification of precipitation in weather systems: determination of systematic errors. *Journal of Hydrology*, **239**, 179–202.
- Ebert, E. E., and W. A. Gallus, 2009: Toward Better Understanding of the Contiguous Rain Area (CRA) Method for Spatial Forecast Verification. *Weather and Forecasting*, **24**, 1401–1415.
- Ehret, U., 2003: Rainfall and flood nowcasting in small catchments using weather radar. Universität Stuttgart (Allemagne), 258 pp.
- Erdin, R., 2009: *Combining rain gauge and radar measurements of a heavy precipitation event over Switzerland: Comparison of geostatistical methods and investigation of important influencing factors*. ETH Zürich - MeteoSwiss, Zürich, 108 pp.
- Fabry, F., and I. Zawadzki, 1995: Long-Term Radar Observations of the Melting Layer of Precipitation and Their Interpretation. *J. Atmos. Sci.*, **52**, 838–851, doi:10.1175/1520-0469(1995)052<0838:LTROOT>2.0.CO;2.
- Feng, S., Q. Hu, and W. Qian, 2004: Quality control of daily meteorological data in China, 1951–2000: a new dataset. *International Journal of Climatology*, **24**, 853–870.
- Frezghi, M., and J. Smithers, 2008: Merged rainfall fields for continuous simulation modelling. *Water SA Manuscript*, **34**, 523–528.

- GIEC, 2007: *Bilan 2007 des changements climatiques : Rapport de synthèse. Contribution des Groupes de travail I, II et III au quatrième Rapport d'évaluation du Groupe d'experts intergouvernemental sur l'évolution du climat*. Groupe d'experts intergouvernemental sur l'évolution du climat, Genève, Suisse, 103 pp.
- Gilleland, E., 2011: Spatial Forecast Verification: Baddeley's Delta Metric Applied to the ICP Test Cases. *Weather and Forecasting*, **26**, 409–415.
- Gilleland, E., T. C. M. Lee, J. Halley Gotway, R. G. Bullock, and B. G. Brown, 2008: Computationally Efficient Spatial Forecast Verification Using Baddeley's Delta Image Metric. *Monthly Weather Review*, **136**, 1747–1757.
- Gilleland, E., D. Ahijevych, B. G. Brown, B. Casati, and E. E. Ebert, 2009: Intercomparison of Spatial Forecast Verification Methods. *Weather and Forecasting*, **24**, 1416–1430.
- Gilleland, E., D. A. Ahijevych, B. G. Brown, and E. E. Ebert, 2010a: Verifying Forecasts Spatially. *Bulletin of the American Meteorological Society*, **91**, 1365–1373.
- Gilleland, E., J. Lindström, and F. Lindgren, 2010b: Analyzing the Image Warp Forecast Verification Method on Precipitation Fields from the ICP. *Weather and Forecasting*, **25**, 1249–1262.
- Goovaerts, P., 2000: Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. *Journal of Hydrology*, **228**, 113–129.
- Gottardi, F., C. Obled, J. Gailhard, and E. Paquet, 2012: Statistical reanalysis of precipitation fields based on ground network data and weather patterns: Application over French mountains. *Journal of Hydrology*, **432–433**, 154–167.
- Goudenhoofd, E., and L. Delobbe, 2009: Evaluation of radar-gauge merging methods for quantitative precipitation estimates. *Hydrology and Earth System Sciences*, **13**, 195–203.
- Haberlandt, U., 2007: Geostatistical interpolation of hourly precipitation from rain gauges and radar for a large-scale extreme rainfall event. *Journal of Hydrology*, **332**, 144–157.
- Hitschfeld, W., and J. Bordan, 1954: Errors inherent in the radar measurement of rainfall at attenuating wavelengths. *Journal of Meteorology*, **11**, 58–67.
- Hong, S.-Y., and J.-O. J. Lim, 2006: The WRF Single-Moment 6-Class Microphysics Scheme (WSM6). *Journal of the Korean Meteorological Society*, **2**, 129–151.
- Hong, S.-Y., J. Dudhia, and S.-H. Chen, 2004: A Revised Approach to Ice Microphysical Processes for the Bulk Parameterization of Clouds and Precipitation. *Monthly Weather Review*, **132**, 103–120.
- Hong, S.-Y., Y. Noh, and J. Dudhia, 2006: A New Vertical Diffusion Package with an Explicit Treatment of Entrainment Processes. *Monthly Weather Review*, **134**, 2318–2341.
- Hong, S.-Y., K.-S. S. Lim, Y.-H. Lee, J.-C. Ha, H.-W. Kim, S.-J. Ham, and J. Dudhia, 2010: Evaluation of the WRF Double-Moment 6-Class Microphysics Scheme for Precipitating Convection. *Advances in Meteorology*, **10**, doi:10.1155/2010/707253.
- Houze, R. A. J., 1994: *Cloud Dynamics*. Academic Press, San Diego, 573 pp.
- Hubbard, K. G., 2000: Multiple Station Quality Control Procedures. *Automated Weather Stations for Applications in Agriculture and Water Resources Management: Current*

- Use and Future Perspectives*, Vol. AGM-3, WMO/TD No. 1074 of *Proceedings of an International Workshop*, Kenneth G. Hubbard and M.V.K. Sivakumar (eds), Lincoln, Nebraska, USA, 133–136.
- Hubbard, K. G., S. Goddard, W. D. Sorensen, N. Wells, and T. T. Osugi, 2005: Performance of Quality Assurance Procedures for an Applied Climate Information System. *Journal of Atmospheric and Oceanic Technology*, **22**, 105–112.
- Iacono, M. J., J. S. Delamere, E. J. Mlawer, M. W. Shephard, S. A. Clough, and W. D. Collins, 2008: Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. *Journal of Geophysical Research: Atmospheres*, **113**, n.c.
- IGN, Algorithmes Géodésiques. <http://geodesie.ign.fr/index.php?page=algorithmes> (Accessed June 2, 2011).
- Isaaks, E. H., and R. M. Srivastava, 1989: *An Introduction to Applied Geostatistics*. Oxford University Press, New York, 146 pp.
- Kain, J. S., 2004: The Kain–Fritsch Convective Parameterization: An Update. *Journal of Applied Meteorology*, **43**, 170–181.
- Kebaili Bargaoui, Z., and A. Chebbi, 2009: Comparison of two kriging interpolation methods applied to spatiotemporal rainfall. *Journal of Hydrology*, **365**, 56–73.
- Keil, C., and G. C. Craig, 2007: A Displacement-Based Error Measure Applied in a Regional Ensemble Forecasting System. *Monthly Weather Review*, **135**, 3248–3259.
- Keil, C., and G. C. Craig, 2009: A Displacement and Amplitude Score Employing an Optical Flow Technique. *Weather and Forecasting*, **24**, 1297–1308.
- Kessler, E., 1969: *On the distribution and continuity of water substance in atmospheric circulation*. American Meteorological Society, 84 pp.
- Khain, A., A. Pokrovsky, M. Pinsky, A. Seifert, and V. Phillips, 2004: Simulation of Effects of Atmospheric Aerosols on Deep Turbulent Convective Clouds Using a Spectral Microphysics Mixed-Phase Cumulus Cloud Model. Part I: Model Description and Possible Applications. *Journal of the Atmospheric Sciences*, **61**, 2963–2982.
- Kirstetter, P.-E., 2008: Estimation quantitative des précipitations par radar météorologique : inférence de la structure verticale des pluies, modélisation des erreurs radar-pluviomètres. Université Joseph Fourier - Laboratoire d'étude des Transferts en Hydrologie et Environnement, 289 pp.
- Kirstetter, P.-E., G. Delrieu, B. Boudevillain, and C. Obled, 2010: Toward an error model for radar quantitative precipitation estimation in the Cévennes-Vivarais region, France. *Journal of Hydrology*, **394**, 28–41.
- Lebel, T., and G. Bastin, 1985: Variogram identification by the mean-squared interpolation error method with application to hydrologic fields. *Journal of Hydrology*, **77**, 31–56.
- Lebel, T., G. Bastin, C. Obled, and J. D. Creutin, 1987: On the accuracy of areal rainfall estimation: A case study. *Water Resources Research*, **23**, 2123–2134.

- Leroy, D., W. Wobrock, and A. I. Flossmann, 2009: The role of boundary layer aerosol particles for the development of deep convective clouds: A high-resolution 3D model with detailed (bin) microphysics applied to CRYSTAL-FACE. *Atmospheric Research*, **91**, 62–78.
- Leroy, M., 2000: *Estimation de l'incertitude de mesure des précipitations*. Météo-France, 13 pp. Note interne disponible à l'adresse : <http://www.ccrom.org/ccrom/IMG/pdf/note42-3.pdf>.
- Li, D., and L. W. Lake, 1994: A moving window semivariance estimator. *Water Resources Research*, **30**, 1479–1489.
- Lim, K.-S. S., and S.-Y. Hong, 2010: Development of an Effective Double-Moment Cloud Microphysics Scheme with Prognostic Cloud Condensation Nuclei (CCN) for Weather and Climate Models. *Monthly Weather Review*, **138**, 1587–1612.
- Lin, Y., and B. A. Colle, 2009: The 4–5 December 2001 IMPROVE-2 Event: Observed Microphysics and Comparisons with the Weather Research and Forecasting Model. *Mon. Wea. Rev.*, **137**, 1372–1392, doi:10.1175/2008MWR2653.1.
- Lin, Y.-L., R. D. Farley, and H. D. Orville, 1983: Bulk Parameterization of the Snow Field in a Cloud Model. *Journal of Climate and Applied Meteorology*, **22**, 1065–1092.
- Marshall, J. S., and W. M. K. Palmer, 1948: The distribution of raindrops with size. *Journal of Meteorology*, **5**, 165–166.
- Météo-France, 2011: Le risque inondation. *Météo - Le magazine*, December.
- Météo-France, and Ministère de l'Écologie et du Développement Durable, Pluies Extrêmes en France métropolitaine. <http://pluiesextremes.meteo.fr/> (Accessed August 17, 2012).
- Michaelides, S., V. Levizzani, E. Anagnostou, P. Bauer, T. Kasparis, and J. E. Lane, 2009: Precipitation: Measurement, remote sensing, climatology and modeling. *Atmospheric Research*, **94**, 512–533.
- Milbrandt, J. A., and M. K. Yau, 2005: A Multimoment Bulk Microphysics Parameterization. Part II: A Proposed Three-Moment Closure and Scheme Description. *Journal of the Atmospheric Sciences*, **62**, 3065–3081.
- Mioche, G., and J. Van Baelen, 2013: One year of rainfall estimates from X band radar and rain gauges network in the urban area of Clermont-Ferrand, France. *Atmospheric Research*, **soumis**.
- Mitchell, K., 2005: The Community Noah Land-Surface Model (LSM) - User's Guide.
- Mittermaier, M., and N. Roberts, 2010: Intercomparison of Spatial Forecast Verification Methods: Identifying Skillful Spatial Scales Using the Fractions Skill Score. *Weather and Forecasting*, **25**, 343–354.
- Molga, P., 2012: Inondations : la moitié des communes françaises concernées par le risque. *Lesechos.fr*, October 10.
- Moral, F. J., 2010: Comparison of different geostatistical approaches to map climate variables: application to precipitation. *International Journal of Climatology*, **30**, 620–631.
- Morrison, H., G. Thompson, and V. Tatarskii, 2009: Impact of Cloud Microphysics on the Development of Trailing Stratiform Precipitation in a Simulated Squall Line:

- Comparison of One- and Two-Moment Schemes. *Monthly Weather Review*, **137**, 991–1007.
- Murphy, A. H., 1993: What Is a Good Forecast? An Essay on the Nature of Goodness in Weather Forecasting. *Weather and Forecasting*, **8**, 281–293.
- Nuissier, O., V. Ducrocq, D. Ricard, C. Lebeaupin, and S. Anquetin, 2008: A numerical study of three catastrophic precipitating events over southern France. I: Numerical framework and synoptic ingredients. *Quarterly Journal of the Royal Meteorological Society*, **134**, 111–130.
- OHMCV, Observatoire Hydro-météorologiques Méditerranéen Cévennes Vivarais. <http://www.ohmcv.fr/index.php?lang=> (Accessed May 17, 2011).
- Orlanski, I., 1975: A rational subdivision of scales for atmospheric processes. *Bulletin of the American Meteorological Society*, **56**, 527–530.
- Parent du Châtelet, J., 2003: Aramis, le réseau français de radars pour la surveillance des précipitations. *La Météorologie*, **8**, 44–52.
- Planche, C., 2011: Développement et évaluation d'un modèle tridimensionnel de nuage mixte à microphysique détaillée : application aux précipitations orographiques. Université Blaise Pascal - Laboratoire de météorologie physique, 180 pp.
- Planche, C., W. Wobrock, A. I. Flossmann, F. Tridon, J. Van Baelen, Y. Pointin, and M. Hagen, 2010: The influence of aerosol particle number and hygroscopicity on the evolution of convective cloud systems and their precipitation: A numerical study based on the COPS observations on 12 August 2007. *Atmospheric Research*, **98**, 40–56, doi:10.1016/j.atmosres.2010.05.003.
- Probert-Jones, J. R., 1962: The radar equation in meteorology. *Quarterly Journal of the Royal Meteorological Society*, **88**, 485–495.
- Roberts, N. M., and H. W. Lean, 2008: Scale-Selective Verification of Rainfall Accumulations from High-Resolution Forecasts of Convective Events. *Monthly Weather Review*, **136**, 78–97.
- Roe, G. H., 2011: Orographic Precipitation. *Annu. Rev. Earth Planet. Sci.*, **33**, 645–671.
- Rossa, A., P. Nurmi, and E. E. Ebert, 2008: Overview of methods for the verification of quantitative precipitation forecasts. *Precipitation: advances in measurement, estimation, and prediction By Michaelides, S.*, Heidelberg : Springer, Berlin, 419–452.
- Schiemann, R., R. Erdin, M. Willi, C. Frei, M. Berenguer, and D. Sempere-Torres, 2011: Geostatistical radar-raingauge combination with nonparametric correlograms: methodological considerations and application in Switzerland. *Hydrol. Earth Syst. Sci.*, **15**, 1515–1536.
- senat.fr, La mission « Inondations » interpelle sur la nécessité d'une approche globale du risque - Sénat.

- [http://www.senat.fr/espace\\_presse/actualites/201210/la\\_mission\\_inondations\\_interpelle\\_sur\\_la\\_necessite\\_dune\\_approche\\_globale\\_du\\_risque.html](http://www.senat.fr/espace_presse/actualites/201210/la_mission_inondations_interpelle_sur_la_necessite_dune_approche_globale_du_risque.html) (Accessed March 6, 2013).
- Sinclair, S., and G. Pegram, 2005: Combining radar and rain gauge rainfall estimates using conditional merging. *Atmospheric Science Letters*, **6**, 19–22.
- Skamarock, W. C. and Coauthors, 2008: *A description of the Advanced Research WRF Version 3*. Mesoscale and Microscale Meteorology Division, Boulder, Colorado, USA, 113 pp.
- Snyder, J. P., 1987: *Map Projections : A Working Manual*. Geological Survey (US), USA, 385 pp.
- Szturc, J., K. Ośródko, and A. Jurczyk, 2012: 12 - Quality Control Algorithms Applied on Weather Radar Reflectivity Data. *Doppler Radar Observations - Weather Radar, Wind Profiler, Ionospheric Radar, and Other Advanced Applications*, InTech, 289–306.
- Tao, W.-K., and J. Simpson, 1993: Goddard Cumulus Ensemble Model. Part I: Model Description. *Terrestrial, Atmospheric and Oceanic Sciences*, **4**, 35–72.
- Taylor, K. E., 2001: Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research: Atmospheres*, **106**, 7183–7192.
- Thompson, G., P. R. Field, R. M. Rasmussen, and W. D. Hall, 2008: Explicit Forecasts of Winter Precipitation Using an Improved Bulk Microphysics Scheme. Part II: Implementation of a New Snow Parameterization. *Monthly Weather Review*, **136**, 5095–5115.
- Tobin, C., L. Nicotina, M. B. Parlange, A. Berne, and A. Rinaldo, 2011: Improved interpolation of meteorological forcings for hydrologic applications in a Swiss Alpine region. *Journal of Hydrology*, **401**, 77–89.
- Tourasse, P., 1981: Analyses spatiales et temporelles de précipitations et utilisation opérationnelle dans un système de prévision des crues. Application aux régions Cévenoles. Université Joseph Fourier - Institut national polytechnique, n.c. pp.
- Tridon, F., 2011: Mesure des précipitations à l'aide d'un radar en bande X non-cohérent à haute résolution et d'un radar en bande K à visée verticale. Application à l'étude de la variabilité des précipitations lors de la campagne COPS. Université Blaise Pascal - Laboratoire de météorologie physique, 198 pp.
- Velasco-Forero, C. A., A. Seed, D. Sempere-Torres, and G. Pegram, 2008: Optimal estimation of rainfall fields merging radar and rain gauges data in an operational context. 5th European Conference on radar in Meteorology and Hydrology, Helsinki, Finland.
- Velasco-Forero, C. A., D. Sempere-Torres, E. F. Cassiraga, and J. Jaime Gomez-Hernandez, 2009: A non-parametric automatic blending methodology to estimate rainfall fields from rain gauge and radar data. *Advances in Water Resources*, **32**, 986–1002.
- Venugopal, V., S. Basu, and E. Foufoula-Georgiou, 2005: A new metric for comparing precipitation patterns with an application to ensemble forecasts. *Journal of Geophysical Research: Atmospheres*, **110**, n.c.

- Verworn, A., and U. Haberlandt, 2010: Spatial interpolation of hourly rainfall – effect of additional information, variogram inference and storm properties. *Hydrology and Earth System Sciences Discussions*, **7**, 6407–6446.
- Vincendon, B. and Coauthors, 2009: Flash flood forecasting within the PREVIEW project: value of high-resolution hydrometeorological coupled forecast. *Meteorology and Atmospheric Physics*, **103**, 115–125.
- Vincendon, B., V. Ducrocq, O. Nuissier, and B. Vié, 2011: Perturbation of convection-permitting NWP forecasts for flash-flood ensemble forecasting. *Natural Hazards and Earth System Sciences*, **11**, 1529–1544.
- Wang, W. and Coauthors, 2011: *ARW Version 3 Modeling System User's Guide (Version 3.2)*. Mesoscale & Microscale Meteorology Division - National Center for Atmospheric Research, Boulder, Colorado, USA,.
- Webster, R., and M. A. Oliver, 2008: *Geostatistics for Environmental Scientists - 2nd edition*. John Wiley & Sons, Ltd, 330 pp.
- Wernli, H., M. Paulat, M. Hagen, and C. Frei, 2008: SAL - A novel quality measure for the verification of quantitative precipitation forecasts. *Monthly Weather Review*, **136**, 4470–4487.
- Wernli, H., C. Hofmann, and M. Zimmer, 2009: Spatial Forecast Verification Methods Intercomparison Project: Application of the SAL Technique. *Weather and Forecasting*, **24**, 1472–1484.
- WMO, 2012: *Guide to Meteorological Instruments and Methods of Observation*. Edition de 2008 (mise à jour en 2010). Genève, Suisse, 716 pp. Rapport disponible à l'adresse : [ftp://ftp.wmo.int/Documents/MediaPublic/Publications/WMO8\\_CIMOguide/WMO8\\_Ed2008\\_Up2010\\_CORR1\\_en.pdf](ftp://ftp.wmo.int/Documents/MediaPublic/Publications/WMO8_CIMOguide/WMO8_Ed2008_Up2010_CORR1_en.pdf).
- Yates, E., 2006: Convection en région Cévennes - Vivarais: Etude de données pluviométriques, simulations numériques et validation multi-échelles. Univeristé Joseph Fourier - Laboratoire d'étude des Transferts en Hydrologie et Environnement, 242 pp.
- You, J., K. G. Hubbard, S. Nadarajah, and K. E. Kunkel, 2007: Performance of Quality Assurance Procedures on Daily Precipitation. *Journal of Atmospheric and Oceanic Technology*, **24**, 821–834.
- Zahumenský, I., 2004: *Guidelines on Quality Control Procedures for Data from Automatic Weather Stations Submitted by*. World Meteorological Organization, Genève, Suisse, 9 pp.
- Zängl, G., and A. Seifert, 2009: Misrepresentation of the seeder-feeder mechanism by Kessler-type autoconversion schemes. *Poster*, 30th International Conference on Alpine Meteorology (ICAM), Rastatt, Germany.
- Zängl, G., A. Seifert, and W. Wobrock, 2010: Modeling stable orographic precipitation at small scales: The impact of the autoconversion scheme. *Meteorologische Zeitschrift*, **19**, 405–416 (12).

- Zhang, X., F. W. Zwiers, G. C. Hegerl, F. H. Lambert, N. P. Gillett, S. Solomon, P. A. Stott, and T. Nozawa, 2007: Detection of human influence on twentieth-century precipitation trends. *Nature*, **448**, 461–465.
- Zimmer, M., and H. Wernli, 2011: Verification of quantitative precipitation forecasts on short time-scales: A fuzzy approach to handle timing errors with SAL. *Meteorologische Zeitschrift*, **20**, 95–105.

# LISTE DES SYMBOLES

## **B, b**

$b(\mathbb{R})$  : barycentre du champ  $\mathbb{R}$  (SAL)

$b_m$  : barycentre de l'objet  $m$  du champ  $\mathbb{R}$  (SAL)

## **C, c**

$c_0$  : effet de pépite (krigeage)

$\tilde{c}_0$  : estimation de  $c_0$  (variogramme paramétrique)

CCN : acronyme anglais de « Cloud Condensation Nucleii » (*ou noyau de condensation*) (origine de la pluie)

CRMSD : acronyme anglais de « Centered Root Mean Squared Difference » (*ou racine carrée de la différence centrée quadratique moyenne*) (vérification spatiale)

## **D, d**

$D$  : diamètre d'une particule (mesure RADAR)

$\mathcal{D}$  : domaine spatial d'étude (SAL)

## **E, e**

$E$  : espérance mathématique (krigeage)

EA : acronyme pour Effets d'Atténuation (mesure RADAR)

ETS : acronyme anglais de « Equitable Threat Score » (vérification spatiale)

## **H, h**

$h$  : vecteur de distance séparant deux stations  $i$  et  $j$  (krigeage)

## **I, i**

IC : acronyme pour Intervalle de Confiance (analyse statistique)

ICP : acronyme anglais de « Inter-Comparison Project » (vérification spatiale)

IN : acronyme anglais de « Ice-Nucleii » (*ou noyau glaçonène*) (origine de la pluie)

IS : acronyme anglais de « Intensity-Scale » (vérification spatiale)

**K, k**

$k$  : décalage temporel (QC)

$k'$  : décalage temporel optimal (QC)

KED : acronyme anglais de « Kriging with External Drift » (*ou krigeage avec dérive externe*)

**L, l**

$\ell$  : multiplicateur de Lagrange (krigeage)

LaMP : acronyme pour Laboratoire de Météorologie Physique

LOO : acronyme anglais de « Leave-One-Out » (analyse statistique)

**M, m**

$M$  : modèle mathématique (variogramme paramétrique)

$\mathcal{M}$  : nombre total d'objets  $m$  (SAL)

$m$  : facteur d'échelle pour le modèle de variogramme puissance (krigeage)

$\tilde{m}$  : estimation de  $m$  (variogramme paramétrique)

ME : acronyme anglais de « Mean Error » (*ou erreur moyenne*) (analyse statistique)

MERG : acronyme anglais de « conditionnal MERGing » (*ou fusion conditionnée*)

MODE : acronyme anglais de « Method for Object-based Diagnostic Evaluation » (vérification spatiale)

MSE : acronyme anglais de « Mean Squared Error » (*ou moyenne des erreurs quadratiques*) (analyse statistique)

MSIE : acronyme anglais de « Mean Squared Interpolation Error » (*ou moyenne des erreurs quadratiques d'interpolation*) (variogramme paramétrique)

**N, n**

$N(D)$  : distribution en taille des gouttes en fonction de leur diamètre  $D$  (variable microphysique)

NDS : acronyme pour Nombre De Stations (utilisées pour l'interpolation) (test des méthodes d'interpolation spatiale)

**O, o**

OHMCV : acronyme pour Observatoire Hydrométéorologique Méditerranéen Cévennes-Vivarais

OK\_RG : acronyme anglais de « Ordinary Kriging of Rain Gauge » (*ou krigeage ordinaire des observations pluviométriques*)

**P, p**

PPI : acronyme anglais de « Plan Position Indicator » (*ou mode de balayage panoramique*)  
(mesure RADAR)

**Q, q**

QC : acronyme anglais de « Quality Control » (*ou contrôle de qualité*)

Qc ou Qcloud : contenu en nombre de l'eau nuageuse (variable microphysique)

Qg ou Qgraup : contenu en nombre du grésil ou graupel (variable microphysique)

Qi ou Qice : contenu en nombre de la glace nuageuse (variable microphysique)

Qr ou Qrain : contenu en nombre de l'eau pluvieuse (variable microphysique)

Qs ou Qsnow : contenu en nombre de la neige (variable microphysique)

**R, r**

$\mathfrak{R}$  : taux de précipitation (mesure RADAR)

$\mathbb{R}$  : champ de précipitation (SAL)

$\overline{\mathbb{R}}$  : cumul moyen du champ  $\mathbb{R}$  (SAL)

$R^*$  : valeur de seuillage (SAL)

$R^{max}$  : valeur maximum du champ  $\mathbb{R}$  (SAL)

$R^{0.95}$  : valeur du 95e centile du champ  $\mathbb{R}$  (SAL)

$r$  : coefficient de corrélation (analyse statistique)

$r(\mathbb{R})$  : distance moyenne entre le barycentre  $\mathbf{b}(\mathbb{R})$  et les barycentres  $\mathbf{b}_m$  (SAL)

RADAR : acronyme anglais de « RADio Detection et Ranging » (*ou détection par ondes radio et mesure de distance*)

$r_{ij}^k$  : coefficient de corrélation entre les observations des stations  $i$  et  $j$  pour un décalage temporel  $k$  (QC)

$\mathcal{R}_m$  : quantité de précipitation intégrée de l'objet  $m$  (SAL)

RMSE : acronyme anglais de « Root Mean Squared Error » (ou racine carrée de l'erreur quadratique moyenne) (analyse statistique)

$RMSE_{ij}^k$  : RMSE entre  $V_i$  et  $\tilde{V}_{ij}^k$  (QC)

**S, s**

$s(\mathbb{R})$  : surfaces pondérées moyennes des  $\mathcal{M}$  objets (SAL)

$\mathbf{s}_0$  : vecteur de coordonnées du point où l'interpolation est souhaitée (krigeage)

SAL : acronyme anglais de « Structure – Amplitude – Location »

SEQ : acronyme pour Somme des Erreurs Quadratiques (variogramme paramétrique)

SIE : acronyme anglais de « Squared Interpolation Error » (*ou erreurs quadratique d'interpolation*) (variogramme paramétrique)

SMSE : acronyme anglais de « Standardized Mean Square Error » (*ou moyenne des erreurs quadratiques normalisées*) (variogramme paramétrique)

## **T, t**

T : fenêtre temporelle pour le calcul du cumul (test des techniques d'interpolation spatiale)

U,u

U : vent horizontal (modélisation)

## **V, v**

V : quantité observé (krigeage)

$\bar{V}$  : moyenne des observations  $V_i$  (analyse statistique)

$\tilde{\bar{V}}$  : moyenne des estimations  $\tilde{V}_i$  (analyse statistique)

Var : variance de krigeage (krigeage)

$V_i$  ou  $V(\mathbf{s})$  ou  $V(\mathbf{s}_i)$  : observation de  $V$  à la station  $i$  de vecteur de coordonnées  $\mathbf{s}$  (ou  $\mathbf{s}_i$ ) (QC, krigeage, analyse statistique)

$\tilde{V}_i$  : estimation de  $V_i$  (QC, krigeage, analyse statistique)

$V_i^{\text{attendue}}$  : valeur de  $V_i$  espérée si elle n'était pas considérée comme douteuse (QC)

$v_i$  : série temporelle des observations  $V$  de la station  $i$  (QC)

$\tilde{V}_{ij}^k$  : approximation linéaire de  $V_i$  par rapport  $V_j$  en tenant compte du décalage  $k$  (QC)

VK : acronyme pour Variance de Krigeage (krigeage)

$\tilde{V}_i^M(\mathbf{s}_i, \beta)$  : estimation de  $V(\mathbf{s}_i)$  pour l'événement  $l$  en considérant le modèle mathématique  $M$  et la portée  $\beta$  (variogramme paramétrique)

## **W, w**

w : vitesse du vent vertical (modélisation)

WDM5, WDM6 : acronyme anglais de « WRF Double-Moment 5/6 – class » (schéma de microphysique)

WRF : acronyme anglais de « Weather Research et Forecasting » (modélisation)

WSM3, WSM5, WSM6 : acronyme anglais de « WRF Single-Moment 3/5/6 – class » (schéma de microphysique)

**Y, y**

$Y_k$  : variable auxiliaire reliée linéairement à  $V$  (KED)

**Z, z**

$Z$  : facteur de réflectivité RADAR (mesure RADAR)

**A, a**

$\alpha(t)$  : facteur assimilé à un palier (variogramme paramétrique)

$\tilde{\alpha}$  : estimation de  $\alpha$  (variogramme paramétrique)

$\alpha_0$  : facteur associé à  $\alpha(t)$  (variogramme paramétrique)

$\tilde{\alpha}_0$  : estimation de  $\alpha_0$  (variogramme paramétrique)

**B,  $\beta$** 

$\beta$  : portée (krigeage)

$\tilde{\beta}$  : estimation de  $\beta$  (variogramme paramétrique)

 **$\Gamma$ ,  $\gamma$** 

$\Gamma(\mathbf{h}, \beta)$  : variogramme paramétrique dit « climatologique » (variogramme paramétrique)

$\gamma(\mathbf{h})$  : semi-variogramme expérimental (krigeage)

$\hat{\gamma}(\mathbf{h})$  : estimateur du semi-variogramme  $\gamma$  (krigeage)

$\tilde{\gamma}(\mathbf{h})$  : semi-variogramme paramétrique (krigeage)

 **$\Delta$ ,  $\delta$** 

$\delta(\mathbf{s})$  : fonction aléatoire stationnaire (krigeage)

 **$\Lambda$ ,  $\lambda$** 

$\lambda_{RADAR}$  : longueur d'onde du signal RADAR (mesure RADAR)

$\lambda_i$  : poids de krigeage associé à  $V_i$  (krigeage)

$\lambda$  : matrice des poids  $\lambda_i$  (krigeage)

**M,  $\mu$** 

$\mu(\mathbf{s})$  : structure déterministe pour l'espérance E de  $V(\mathbf{s})$  (krigeage)

**N, v**

$v$  : exposant pour le modèle de variogramme puissance (krigeage)

$\tilde{v}$  : estimation de  $v$  (variogramme paramétrique)

 **$\Sigma, \sigma$** 

$\sigma$  : écart type (analyse statistique)

**Caractères numériques**

$(\mathbf{0}_a)$  : vecteur des valeurs de  $V(\mathbf{s})$  (MERG)

$(\mathbf{0}_b)$  : vecteur des valeurs de la variable auxiliaire  $Y_1$  (MERG)

$(\mathbf{1})$  : vecteur des valeurs de  $(\mathbf{0}_a)$  interpolées (MERG)

$(\mathbf{2}_0)$  : vecteur des valeurs de  $(\mathbf{0}_b)$  aux points de coordonnées  $\mathbf{s}$  (MERG)

$(\mathbf{2})$  : vecteur des valeurs de  $(\mathbf{2}_0)$  interpolées (MERG)

$(\mathbf{3})$  : vecteur des erreurs entre  $(\mathbf{2})$  et  $(\mathbf{0}_b)$  (MERG)

$(\mathbf{4})$  : vecteur des valeurs  $(\mathbf{1})$  corrigées par  $(\mathbf{3})$  (MERG)



La majeure partie de ce travail de thèse a été réalisée grâce au **logiciel R** = R Development Core Team (2011). R : A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>.

Quelques packages ont été utilisés, dont :

- **gstat** = Pebesma, E.J., 2004. Multivariable geostatistics in S : the gstat package. *Computers & Geosciences*, 30: 683-691. <http://www.gstat.org/>
- **fields** = Reinhard Furrer, Douglas Nychka and Stephen Sain (2012). fields : Tools for spatial data. <http://CRAN.R-project.org/package=fields>
- **ncdf** = David Pierce (2011). ncdf : Interface to Unidata netCDF data files. <http://CRAN.R-project.org/package=ncdf>
- **raster** = Robert J. Hijmans & Jacob van Etten (2013). raster : Geographic data analysis and modeling. <http://CRAN.R-project.org/package=raster>
- **EBImage** = Oleg Sklyar, Gregoire Pau, Mike Smith and Wolfgang Huber (). EBImage : Image processing toolbox for R. <http://www.bioconductor.org/packages/2.12/bioc/html/EBImage.html>
- **verification** = NCAR - Research Application Program (2012). verification : Forecast verification utilities. <http://CRAN.R-project.org/package=verification>





---

## RÉSUMÉ

### **Amélioration des estimations quantitatives des précipitations à hautes résolutions : comparaison de deux techniques combinant les observations et application à la vérification spatiale des modèles météorologiques.**

Ces dernières années, de nombreux efforts ont été entrepris pour mieux comprendre les phénomènes précipitants parfois à l'origine de crues de cours d'eau et d'inondations ravageuses. Courant 2009, un consortium auvergnat a été mis en place pour notamment surveiller et prévoir ces événements. Les travaux menés dans cette thèse visent d'une part à améliorer les estimations quantitatives des précipitations (QPE) et d'autre part à vérifier les prévisions issues de modèles numériques sur de petites zones d'étude telles qu'une agglomération.

L'observation des précipitations peut être réalisée à l'aide soit d'un pluviomètre qui fournit une mesure directe et précise de la quantité de pluie tombée au sol mais ne renseigne pas sur la variabilité spatiale des pluies soit d'un RADAR météorologique qui donne une représentation détaillée de la structure spatiale des précipitations mais dont les estimations sont sujettes à diverses erreurs d'autant plus prononcées en régions montagneuses. Le premier déficit de cette thèse a été de trouver la meilleure façon de combiner ces deux informations complémentaires. Deux techniques géostatistiques ont été sélectionnées pour obtenir la meilleure QPE : le krigeage avec dérive externe (KED) et la fusion conditionnée (MERG). Les performances de ces deux méthodes ont été comparées au travers de deux domaines d'étude qui présentent des résolutions spatio-temporelles différentes.

La seconde partie de cette thèse est consacrée à la mise en place d'une méthodologie fiable permettant de comparer spatialement les champs de QPE alors reconstruits et les prévisions quantitatives des précipitations (QPF). L'effort fut porté sur le modèle « Weather Research et Forecasting » (WRF). Une étude préliminaire a été réalisée pour tester les capacités du modèle et plus particulièrement des schémas de microphysique à reproduire la pluie. Cette étude assure ainsi l'obtention de prévisions réalistes pour une application sur des cas réels. L'appréciation de la qualité des QPF s'est focalisée sur la quantification spatiale des erreurs de prévision en termes de structure, d'intensité et de localisation des systèmes précipitants (SAL : Wernli *et al.* 2008, 2009).

*Mots-clés : pluviomètres, RADAR météorologique, QPE, KED, MERG, QPF, WRF, vérification spatiale, SAL.*

---

## ABSTRACT

### **Improvement of quantitative precipitation estimations at high resolutions: comparison of two techniques combining observations and application to spatial forecast verification of numerical weather models.**

In the last decades, many efforts were made to better understand the origins of rain that sometimes lead to rivers runoff or devastating floods. In 2009, a consortium took place in Auvergne in order to observe and predict these events. These works were focused on the improvement of quantitative precipitation estimations (QPE) and the verification of numerical weather models over small areas such as urban environment.

Rainfall measurement could be operated either by rain gauges which provides direct and precise rainfall estimations but unfortunately cannot capture the spatial variability or by using weather RADAR which provides a detailed spatial representation of precipitation but estimates are derived indirectly and are subject to a combination of errors which are most pronounced over complex terrain. The main issue of these works was to find the best way to combine both observational systems which are complementary as well. In order to obtain the more truthful fields of QPE, two geostatistical techniques were selected: the kriging with external drift (KED) and the conditional merging (MERG). The performances of these two methods have been experienced on two catchments with different spatial and temporal resolutions.

The second part of these works is focused on a reliable method for QPE comparison and quantitative precipitation forecast (QPF). The main effort was focused on the "Weather Research and Forecasting" (WRF) model. A preliminary study was made to check the performances of the microphysics schemes of the model to ensure realistic forecasts for an application on real cases. The spatial verification of the model set up contains three distinct components that consider aspects of the structure, amplitude and location of the precipitation field (SAL : Wernli *et al.* 2008, 2009).

*Keywords: rain gauge, weather RADAR, QPE, KED, MERG, QPF, WRF, spatial verification, SAL.*

---