

Université du Québec  
Institut National de la Recherche Scientifique  
Centre eau, terre et environnement

# DÉSAGRÉGATION STATISTIQUE DE LA PRÉCIPITATION MÉSOÉCHELLE

Par  
**Patrick GAGNON**

Thèse présentée pour obtenir le grade de  
*Philosophiæ Doctor* (Ph.D.) en Sciences de l'eau

## Jury d'évaluation

Examineur externe et président du jury	M. Vincent FORTIN, chercheur scientifique Environnement Canada
Examineur externe	M. Robert LECONTE, professeur Université de Sherbrooke
Examineur interne	M. Richard TURCOTTE, chercheur invité INRS-ETE Université du Québec
Codirecteur de recherche	M. Daniel CAYA, professeur Université du Québec à Montréal
Codirecteur de recherche	M. Alain MAILHOT, professeur INRS-ETE Université du Québec
Directeur de recherche	M. Alain ROUSSEAU, professeur INRS-ETE Université du Québec

Février 2012



# Résumé

La modélisation hydrologique nécessite des données de précipitation à une échelle spatiale adaptée au bassin versant analysé. En passé récent, les données de précipitation proviennent généralement d'observations aux stations météorologiques qui sont interpolées directement (*e.g.* polygones de Thiessen, pondération inverse à la distance, krigeage). Pour les données futures, que ce soit en prévision météorologique ou en projection climatique, la résolution spatiale des champs de précipitation est déterminée par la grille du modèle numérique utilisé. Dans la majorité des cas, la taille des tuiles de ces grilles est de l'ordre de la dizaine de kilomètres, ce qui est trop grossier pour certaines applications hydrologiques sur de petits bassins versants. Bien que la résolution spatiale des modèles météorologiques et climatiques se raffine avec l'évolution des ressources informatiques et des connaissances physiques du climat, un écart entre la résolution disponible à partir des modèles physiques et la résolution nécessaire pour certaines applications demeure.

Pour palier ce problème, un modèle de *désagrégation* statistique de la précipitation, qui distribue spatialement la précipitation d'une grille mésoéchelle (*i.e.* résolution  $> 10$  km) vers une grille à plus fine échelle, peut être appliqué. Les modèles de désagrégation utilisés jusqu'à maintenant dans les applications hydrologiques ont l'avantage d'être simple à mettre en application et nécessitent peu de temps de calcul comparativement aux modèles physiques. La simplicité d'un modèle le rend accessible pour un plus grand nombre d'utilisateurs potentiels et le faible temps de calcul permet aussi de produire plusieurs champs désagrégés pour un même événement. Cependant, ces modèles ont comme principal inconvénient de donner des champs de précipitation comportant des discontinuités visibles ne reflétant pas la structure spatiale observée de la précipitation.

L'objectif principal de la thèse est de développer un modèle de désagrégation statistique de la précipitation mésoéchelle produisant des structures spatiales réalistes. L'échantillonnage de Gibbs, une méthode Monte Carlo par chaîne de Markov, est choisi pour produire les valeurs

désagrégées. Cette méthode permet de générer un ensemble de valeurs (ici, l'ensemble des valeurs désagrégées sur les pixels d'une tuile de la grille à résolution mésoéchelle) lorsque la distribution statistique de chaque élément (ici, chaque pixel qui subdivise la tuile) dépend des autres éléments de l'ensemble et qu'aucun élément n'est connu. La thèse comporte trois objectifs spécifiques : (i) le développement du modèle sur une région avec peu de relief, (ii) l'adaptation du modèle à une région montagneuse, et (iii) l'application du modèle pour la modélisation hydrologique d'un événement de pluie intense sur un petit bassin versant québécois.

La région avec peu de relief étudiée se situe dans les États de la Floride, de l'Alabama, de la Géorgie et de la Caroline de Sud, au sud-est des États-Unis, et couvre environ  $400 \times 400 \text{ km}^2$ . Des données journalières observées à fine résolution (environ 4 km) de 2002 à 2005 sont utilisées pour définir la structure du modèle et pour en estimer les paramètres. La précipitation sur un pixel est supposée issue de la distribution lognormale avec espérance égale à la moyenne des huit pixels les plus proches à laquelle s'ajoutent deux termes : un tenant compte de la distance des pixels les plus proches et un autre tenant compte de l'anisotropie, définie à partir de la vitesse et de la direction du vent à 700 hPa. L'écart-type de la précipitation sur un pixel croît en fonction de l'espérance et de l'énergie potentielle de convection disponible durant l'événement. Au total, le modèle compte cinq paramètres à estimer. Les données journalières observées de 2006 à 2008 sont comparées aux données désagrégées par le modèle. Dans l'ensemble, le modèle reproduit avec justesse la corrélation spatiale et l'anisotropie. Le modèle reproduit plus difficilement les valeurs extrêmes associées aux événements convectifs.

L'adaptation du modèle à une région montagneuse se fait en considérant un territoire d'environ  $300 \times 300 \text{ km}^2$  situé dans les États de Washington et de l'Oregon, au nord-ouest des États-Unis. L'analyse des données journalières observées à fine résolution de 2002 à 2005 a mené à l'ajout au modèle de la Partie I d'un seul terme pour tenir compte de la topographie. Ce terme dépend de l'*anomalie topographique*, qui correspond, *grosso modo*, à la différence entre l'altitude observée sur un pixel et l'altitude estimée à partir des pixels du voisinage. Contrairement au sud-est des États-Unis, la vitesse et la direction du vent n'ont pas permis d'expliquer l'anisotropie dans cette région à topographie complexe. Néanmoins, le modèle, appliqué à la précipitation journalière de 2006 à 2008, reproduit adéquatement les plus importantes pointes locales de précipitation causées par la topographie. Il a aussi

été observé que le modèle sans le terme tenant compte de la topographie est capable de reproduire l'impact de cette topographie lorsque les variations d'altitude sont peu importantes.

L'application en modélisation hydrologique du modèle développé au sud-est des États-Unis se fait sur un événement de pluie intense sur le bassin de la rivière des Anglais, à la frontière entre le Québec (Canada) et l'État de New-York (États-Unis). Le modèle hydrologique distribué HYDROTEL est utilisé. Les débits simulés varient selon la méthode utilisée pour répartir spatialement la pluie, illustrant l'importance d'un modèle de désagrégation adéquat. Les résultats montrent que le modèle amène une valeur ajoutée par rapport à l'interpolation puisqu'il peut produire un ensemble de champs de précipitation pour un seul événement. Pour un gestionnaire, un ensemble de valeurs facilite le processus décisionnel puisqu'il permet d'illustrer l'incertitude sur les résultats d'une simulation. Aussi, le modèle reproduit généralement bien l'emplacement de la pointe de précipitation de l'événement, mais la sous-estime. L'impact de la désagrégation diminue avec la superficie drainée, mais est toujours présent à l'exutoire du bassin (730 km<sup>2</sup>). Le débit simulé est fortement conditionné par la précipitation utilisée en entrée, mais dépend aussi beaucoup des types de sol et du calage des paramètres du modèle hydrologique.



# Remerciements

Je tiens d'abord à remercier mon superviseur Alain Rousseau pour son soutien constant au cours des dernières années. Alain a su me redonner confiance lorsque j'en avais besoin et a toujours trouvé une place dans son agenda bien rempli pour me donner des commentaires pertinents sur mes travaux. Je remercie aussi les membres de son équipe, particulièrement Alain Royer, Stéphane Savary et Sébastien Tremblay pour leur aide avec le logiciel HYDROTEL et le support technique.

Je remercie mes codirecteurs, messieurs Alain Mailhot et Daniel Caya. J'ai grandement apprécié travailler avec Alain et ses commentaires judicieux ont grandement amélioré la thèse. J'ai aussi grandement apprécié les discussions que j'ai eues avec Daniel et je le remercie d'avoir supervisé mon stage à l'Université de Victoria à l'hiver 2009. J'en profite pour remercier les membres du *Pacific Climate Impacts Consortium* (PCIC) de m'avoir accueilli dans leurs bureaux au cours de ce stage.

Je tiens aussi à remercier les examinateurs interne et externes, messieurs Richard Turcotte, Vincent Fortin et Robert Leconte pour leurs commentaires pertinents. Je remercie également Simon Ricard du Centre d'expertise hydrique du Québec (CEHQ) de m'avoir fourni des données météorologiques sur le territoire québécois.

De nombreuses personnes ont fait en sorte que l'INRS représente plus qu'un lieu de travail pour moi. Il m'est impossible de tous les nommer, mais je tiens à mentionner que je m'ennuierai des discussions plus ou moins importantes et plus ou moins pertinentes du midi.

Merci Valérie pour tout ton amour, ton soutien et ta patience au cours des trois dernières années et particulièrement lors de la rédaction de ma thèse. Merci à Émile et à toi de me rappeler l'essentiel.



# Table des matières

Résumé	i
Remerciements	v
Table des matières	vii
Liste des tableaux	xi
Table des figures	xiii
<b>1 INTRODUCTION</b>	<b>1</b>
1.1 Problématique . . . . .	1
1.1.1 Données de précipitation pour le passé récent . . . . .	1
1.1.2 Données de précipitation pour le futur . . . . .	3
1.1.3 Mise à l'échelle spatiale de la précipitation . . . . .	6
1.2 Objectifs de la thèse . . . . .	8
1.2.1 Premier objectif spécifique . . . . .	9
1.2.2 Deuxième objectif spécifique . . . . .	10
1.2.3 Troisième objectif spécifique . . . . .	10
1.3 Structure de la thèse . . . . .	11
<b>I DÉSAGRÉGATION STATISTIQUE DE LA PRÉCIPITA- TION SUR UNE RÉGION AVEC PEU DE RELIEF</b>	<b>13</b>
<b>2 REVUE DE LITTÉRATURE</b>	<b>15</b>
2.1 Cellules de précipitation . . . . .	16
2.2 Modèles multi-échelles . . . . .	17
2.3 Processus stochastiques . . . . .	20

2.4	Discussion . . . . .	22
<b>3</b>	<b>DÉVELOPPEMENT DU MODÈLE</b>	<b>25</b>
3.1	Région à l'étude . . . . .	25
3.1.1	Données de précipitation . . . . .	26
3.1.2	Données atmosphériques . . . . .	27
3.2	Structure mathématique du modèle . . . . .	28
3.2.1	Espérance de la précipitation sur un pixel . . . . .	30
3.2.2	Écart-type de la précipitation sur un pixel . . . . .	39
3.2.3	Distribution statistique de la précipitation sur un pixel . . . . .	42
3.3	Algorithme de désagrégation . . . . .	47
3.4	Discussion . . . . .	50
<b>4</b>	<b>RÉSULTATS DE LA DÉSAGRÉGATION</b>	<b>55</b>
4.1	Sensibilité à la taille du domaine . . . . .	56
4.2	Structure spatiale . . . . .	59
4.3	Valeurs extrêmes . . . . .	64
4.4	Discussion . . . . .	67
 <b>II DÉSAGRÉGATION STATISTIQUE DE LA PRÉCIPITA-</b>		
<b>TION SUR UNE RÉGION MONTAGNEUSE</b>		<b>73</b>
<b>5</b>	<b>REVUE DE LITTÉRATURE</b>	<b>75</b>
5.1	Modèles de précipitation orographique (MPO) . . . . .	76
5.2	Inclusion de la topographie en désagrégation . . . . .	78
5.3	Discussion . . . . .	79
<b>6</b>	<b>DÉVELOPPEMENT DU MODÈLE</b>	<b>81</b>
6.1	Région à l'étude . . . . .	81
6.1.1	Données de précipitation . . . . .	82
6.1.2	Données atmosphériques . . . . .	83
6.1.3	Données topographiques . . . . .	84
6.2	Structure mathématique du modèle . . . . .	84
6.2.1	Espérance de la précipitation sur un pixel . . . . .	85
6.2.2	Écart-type de la précipitation sur un pixel . . . . .	91
6.3	Discussion . . . . .	93

<b>7</b>	<b>RÉSULTATS DE LA DÉSAGRÉGATION</b>	<b>97</b>
7.1	Structure spatiale . . . . .	97
7.2	Valeurs extrêmes . . . . .	99
7.3	Discussion . . . . .	103
<b>III</b>	<b>IMPACT DE LA DÉSAGRÉGATION STATISTIQUE DE LA PRÉCIPITATION EN MODÉLISATION HYDROLO- GIQUE</b>	<b>105</b>
<b>8</b>	<b>REVUE DE LITTÉRATURE</b>	<b>107</b>
8.1	Impact de la résolution spatiale de la précipitation en modélisation hydrologique	107
8.2	Désagrégation appliquée à la modélisation hydrologique . . . . .	110
8.3	Discussion . . . . .	112
<b>9</b>	<b>PROTOCOLE EXPÉRIMENTAL</b>	<b>115</b>
9.1	Territoire à l'étude . . . . .	117
9.1.1	Choix de l'événement . . . . .	118
9.1.2	Données hydrométriques . . . . .	120
9.1.3	Données de précipitation . . . . .	120
9.1.4	Données atmosphériques . . . . .	123
9.2	Estimation des paramètres du modèle de désagrégation . . . . .	124
9.3	Modèle hydrologique . . . . .	127
9.3.1	Description sommaire d'HYDROTEL . . . . .	128
9.3.2	Application sur le bassin de la rivière des Anglais . . . . .	131
<b>10</b>	<b>RÉSULTATS DES SIMULATIONS HYDROLOGIQUES</b>	<b>137</b>
10.1	Exactitude . . . . .	138
10.1.1	Exactitude de la pluie désagrégée . . . . .	138
10.1.2	Exactitude de l'écoulement désagrégé . . . . .	140
10.2	Précision . . . . .	143
10.2.1	Précision de la pluie désagrégée . . . . .	146
10.2.2	Précision de l'écoulement désagrégé . . . . .	149
10.3	Discussion . . . . .	153
<b>11</b>	<b>CONCLUSION</b>	<b>157</b>
11.1	Résumé des constats . . . . .	157

11.1.1	Partie I . . . . .	158
11.1.2	Partie II . . . . .	159
11.1.3	Partie III . . . . .	160
11.2	Perspectives . . . . .	161
11.2.1	Analyse approfondie des propriétés du modèle de désagrégation . . .	161
11.2.2	Analyse de sensibilité de la simulation hydrologique à partir de la pluie désagrégée . . . . .	163
<b>BIBLIOGRAPHIE</b>		<b>165</b>
<b>A Liste des abréviations et des sigles</b>		<b>183</b>
<b>B Liste des variables utilisées</b>		<b>187</b>
<b>C Échantillonnage de Gibbs</b>		<b>191</b>
<b>D Résultats de la désagrégation sur les pixels de 7,6, 11,4 et 15,2 km sur une région avec peu de relief</b>		<b>195</b>
<b>E Résultats de la désagrégation sur les pixels de 8,8, 13,2, 17,6 et 26,4 km sur une région montagneuse</b>		<b>205</b>
<b>F Écoulement désagrégé sur les tronçons de tête</b>		<b>215</b>

# Liste des tableaux

3.1	Somme du carré des écarts entre la précipitation observée à 3,8 km et différentes équations d'espérance (2002-2005). . . . .	32
3.2	Composantes principales (CP) créées à partir des 5 136 976 valeurs de $\vec{A}^{(i,j)}$ sur les pixels de 3,8 km (2002-2005). . . . .	33
3.3	Somme du carré des écarts entre la précipitation observée à 3,8 km et différentes équations d'espérance tenant compte des variables atmosphériques (2002-2005). . . . .	36
3.4	Valeurs estimées des paramètres de l'équation (3.4) pour les différentes tailles de pixels (2002-2005). . . . .	37
3.5	Somme du carré des écarts entre les écart-types observés et estimés sur les 797 groupes selon les divers modèles considérés (2002-2005). . . . .	42
3.6	Valeurs estimées des paramètres de l'équation (3.7) pour les différentes tailles de pixels (2002-2005). . . . .	44
3.7	Distributions à deux paramètres couramment utilisées pour modéliser la précipitation. . . . .	45
4.1	Valeurs de $\phi(d_1, d_2, d_e)$ (%) pour différents nombre de tuiles $d_1$ , $d_2$ et $d_e$ calculées sur les champs désagrégés à résolution 3,8 et 22,8 km. . . . .	57
6.1	Somme du carré des écarts entre la précipitation observée à 4,4 km et deux équations d'espérance tenant compte du vent à 500 et à 700 hPa (2002-2005). . . . .	86
6.2	Somme du carré des écarts entre la précipitation observée à 4,4 km et certaines équations d'espérance tenant compte de la topographie. . . . .	87
6.3	Valeurs estimées des paramètres de l'équation (3.2) pour le Modèle 1 (ignorant la topographie) et de l'équation (6.6) pour le Modèle 2 (incluant la topographie) pour les différentes tailles de pixels (2002-2005). . . . .	91

6.4	Valeurs estimées du coefficient de variation $\theta_0$ (équation (6.8)) communes aux Modèles 1 et 2 pour les différentes tailles de pixels (2002-2005). . . . .	94
9.1	Pourcentages des différents types de sol sur les sept sous-bassins. . . . .	117
9.2	Pourcentages des différentes occupations du sol sur les sept sous-bassins. . .	120
9.3	Sources de précipitations journalières utilisées pour la désagrégation et/ou pour la simulation hydrologique. . . . .	124
9.4	Paramètres du modèle de désagrégation estimés sur la tuile couvrant le bassin de la rivière des Anglais et sur le sud-est des États-Unis (Partie I). . . . .	127
9.5	Valeurs estimées de certains paramètres d'HYDROTEL. . . . .	134
F.1	Conductivité hydraulique à saturation pour les types de sols présents sur le bassin. . . . .	216

# Table des figures

2.1	Exemple synthétique de précipitation désagrégée selon l'approche des cellules de précipitation. . . . .	16
2.2	Illustration de la désagrégation selon l'approche multifractale. . . . .	18
2.3	(a) Les 24 pixels voisins et les six paramètres considérés dans le modèle de Allcroft et Glasbey (2003). (b) Les huit pixels voisins et les deux types de voisins considérés dans le modèle de Mackay <i>et al.</i> (2001). . . . .	21
3.1	Région étudiée avec les $96 \times 96$ pixels de 3,8 km utilisés. . . . .	26
3.2	Pixels voisins retenus pour exprimer l'espérance de la précipitation sur le pixel $(i, j)$ . . . . .	33
3.3	Exemple synthétique d'une précipitation désagrégée lorsque le coefficient d'anisotropie est trop grand. . . . .	38
3.4	(a) Diagrammes en boîtes des résidus des données disponibles à 3,8 km regroupées selon 39 classes d'espérance (2002-2005). (b) Zoom sur les boîtes. . . . .	39
3.5	Estimation des paramètres (a) $\theta_1$ et (b) $\theta_2$ de l'équation (3.6) pour chacune des 20 classes de CAPE (2002-2005). . . . .	41
3.6	Écart-types observés et estimés à partir des équations $\sigma_{R_{i,j}} = (\theta_0 + \theta_1 C_{i,j}) \bar{R}_{i,j}^{\theta_2}$ (cercles) et $\sigma_{R_{i,j}} = (\theta_0 + \theta_1 C_{i,j}) \bar{R}_{i,j}^{\theta_2 + \theta_3 C_{i,j}}$ (carrés) pour les 797 groupes (2002-2005). . . . .	43
3.7	Histogrammes des valeurs de la fonction de répartition des distributions normale (gauche), lognormale (centre) et Gamma (droite) calculées sur les données disponibles à 3,8 km (2002-2005). . . . .	46
3.8	Illustration de l'algorithme de désagrégation. . . . .	48
3.9	Écart-type de la précipitation désagrégée des huit scénarios en fonction de l'itération. . . . .	50
3.10	Autocorrélation de l'écart-type des champs désagrégés pour les huit scénarios considérés. . . . .	51

4.1	Illustration des champs désagrégés nécessaires pour calculer $\phi(d_1, d_2, d_e)$ avec $d_e = 2$ (lignes plus épaisses), $d_1 = 4$ et $d_2 = 6$ (lignes minces) lorsque $N_p = 2$ .	58
4.2	Précipitation moyenne (mm/jour) sur les 801 jours (2006-2008) pour les observations (gauche) et pour un des huit champs désagrégés (droite) pour les pixels de taille (a-b) 3,8 et (c-d) 22,8 km. . . . .	59
4.3	Coefficient de corrélation moyen dans les quatre directions calculé sur les jours avec (a) CAPE < 1000 J/kg et (b) CAPE > 1000 J/kg pour les pixels de 3,8 km observés (lignes continues) et des huit séries désagrégées (lignes pointillées) : 45° (rouge), 90° (vert), 135° (bleu) et 0° (noir). . . . .	61
4.4	Coefficient de corrélation moyen dans les quatre directions calculé sur les jours avec (a) CAPE < 1000 J/kg et (b) CAPE > 1000 J/kg pour les pixels de 22,8 km observés (lignes continues) et des huit séries désagrégées (lignes pointillées) : 45° (rouge), 90° (vert), 135° (bleu) et 0° (noir). . . . .	62
4.5	Coefficient de corrélation moyen dans les quatre directions calculé sur les jours avec (a) CAPE < 1000 J/kg et (b) CAPE > 1000 J/kg pour les pixels de 3,8 km observés (lignes continues) et obtenus par la désagrégation en cascade (lignes pointillées) : 45° (rouge), 90° (vert), 135° (bleu) et 0° (noir). . . . .	63
4.6	Précipitation moyenne (mm/jour) sur tous les jours (2006-2008) avec CAPE < 1000 J/kg (haut) et CAPE > 1000 J/kg (bas) pour les observations (gauche), pour un des huit champs désagrégés directement à 3,8 km (centre) et désagrégés en cascade à 3,8 km (droite). . . . .	64
4.7	Précipitation journalière maximale (mm) sur les pixels de 3,8 (haut) et 22,8 km (bas) durant la période 2006-2008 : observée (gauche), plus petite (centre) et plus grande (droite) valeurs maximales journalières désagrégées. . . . .	65
4.8	Distributions cumulatives empiriques des précipitations journalières plus grandes que 30 mm observées (lignes continues) et désagrégées (lignes pointillées) pour quatre pixels à résolution 3,8 (noir) et 22,8 km (bleu). . . . .	66
4.9	Précipitations maximales observée et désagrégées sur les pixels de 3,8 km des 6 × 6 tuiles centrales pour chaque jour avec (a) CAPE < 1000 J/kg et (b) CAPE > 1000 J/kg. . . . .	67
4.10	Précipitations maximales observée et désagrégées sur les pixels de 22,8 km des 6 × 6 tuiles centrales pour chaque jour avec (a) CAPE < 1000 J/kg et (b) CAPE > 1000 J/kg. . . . .	68

4.11	Précipitations maximales observée et désagrégées en cascade sur les pixels de 3,8 km des $6 \times 6$ tuiles centrales pour chaque jour avec (a) CAPE $< 1000$ J/kg et (b) CAPE $> 1000$ J/kg. . . . .	69
6.1	Région étudiée avec l'altitude (m) des $72 \times 72$ pixels de 4,4 km utilisés. . . .	82
6.2	Précipitation journalière moyenne durant la période 2002-2005 pour les $72 \times 72$ pixels de 4,4 km en fonction de leur altitude. . . . .	88
6.3	Exemple synthétique en deux dimensions illustrant les trois variables topographiques : la pente $P_{i,j}^{(L)}$ , l'orientation $O_{i,j}^{(L)}$ , correspondant à la direction du gradient, et l'anomalie topographique $T_{i,j}^{(L)}$ avec $L = 3$ . . . . .	89
6.4	Écarts-types observés et estimés à partir de l'équation (6.8) pour les 54 classes d'espérance du Modèle 1 sur les données disponibles à 4,4 km. . . . .	93
7.1	Précipitation moyenne (mm/jour) sur les 750 jours (2006-2008) (a) observée à 4,4 km, (b) observée sur les tuiles de 52,8 km et désagrégée à 4,4 km par les Modèles (c) 1 et (d) 2. . . . .	98
7.2	Écart-type inter-journalier (mm) sur les 750 jours (2006-2008) (a) observé à 4,4 km, (b) observé sur les tuiles de 52,8 km et désagrégé à 4,4 km par les Modèles (c) 1 et (d) 2. . . . .	99
7.3	Précipitation journalière maximale (mm) sur les 750 jours (2006-2008) (a) observée à 4,4 km, (b) observée sur les tuiles de 52,8 km et désagrégée à 4,4 km par les Modèles (c) 1 et (d) 2. . . . .	100
7.4	Précipitations maximales observée et produites par les huit réalisations désagrégées des Modèles (a) 1 (cercles) et (b) 2 (triangles) sur les pixels de 4,4 km des $4 \times 4$ tuiles centrales pour chaque jour. . . . .	101
7.5	Queues des distributions empiriques de la précipitation journalière observée à 4,4 km (ligne continue noire), observée à 52,8 km (ligne continue rouge) et désagrégée à 4,4 km par les Modèles 1 (lignes pointillées bleues) et 2 (lignes pointillées vertes) pour six pixels sélectionnés. . . . .	102
9.1	Contexte de l'application de la désagrégation en simulation hydrologique. . .	116
9.2	Bassin versant de la rivière des Anglais avec les différentes sources de données hydrométéorologiques utilisées. . . . .	118
9.3	Exutoire et numéro de référence des sept sous-bassins de la rivière des Anglais analysés. . . . .	119

9.4	Précipitation moyenne (mm/jour) de mai à novembre (2002-2005) des données de <i>Stage IV</i> pour les $36 \times 36$ pixels étudiés. . . . .	121
9.5	Précipitation journalière de mai à novembre (2002-2005) provenant des données de <i>Stage IV</i> et de la grille du CEHQ agrégées sur chacune des $3 \times 3$ tuiles. . . . .	123
9.6	Précipitations (mm) de référence pour le 11 juin 2002 utilisées pour simuler les débits de référence à 4,4, 8,8, 13,2, 17,6, 26,4 et 52,8 km. . . . .	125
9.7	Schéma conceptuel du sous-modèle BV3C. . . . .	130
9.8	Bassin de la rivière des Anglais divisé en 182 UHRH. . . . .	131
9.9	Débits journaliers du 9 au 20 juin 2002 observés à la station hydrométrique (noir) et simulés à partir des pluies aux différentes résolutions : 4,4 (bleu), 8,8 (bleu pâle), 13,2 (vert), 17,6 (mauve) et 26,4 km (rouge). . . . .	133
9.10	Illustration de la méthode d'interpolation bilinéaire de la pluie agrégée utilisée. . . . .	135
10.1	Schéma illustrant la procédure utilisée pour l'évaluation de la valeur ajoutée de la désagrégation. . . . .	138
10.2	Moyennes spatiales des pluies désagrégées versus celles des pluies de référence à résolution (a) 4,4, (b) 8,8, (c) 13,2, (d) 17,6 et (e) 26,4 km pour le 11 juin 2002 sur les sept sous-bassins. . . . .	139
10.3	Moyennes spatiales des pluies aléatoires versus celles des pluies de référence à résolution (a) 4,4, (b) 8,8, (c) 13,2, (d) 17,6 et (e) 26,4 km pour le 11 juin 2002 sur les sept sous-bassins. . . . .	140
10.4	Biais de la pluie désagrégée par rapport à la pluie de référence du 11 juin 2002 pour les sept sous-bassins en fonction de la résolution du pixel. . . . .	141
10.5	Biais sur l'écoulement du 11 juin en fonction du biais sur la pluie du 11 juin 2002 à résolution (a) 4,4, (b) 8,8, (c) 13,2, (d) 17,6 et (e) 26,4 km pour les trois sources de pluie et les sept sous-bassins (Figure 9.3). . . . .	142
10.6	Biais sur l'écoulement des 11 et 12 juin en fonction du biais sur la pluie du 11 juin 2002 à résolution (a) 4,4, (b) 8,8, (c) 13,2, (d) 17,6 et (e) 26,4 km pour les trois sources de pluie et les sept sous-bassins (Figure 9.3). . . . .	143
10.7	Écoulement moyen du 11 juin simulé (a et d) à partir des paramètres calés à 4,4 km, (b et e) à partir d'une valeur de $Z_3$ de 4 m et (c et f) à partir d'une lame de l'HGM de 0,0005 m en fonction de la pluie moyenne du 11 juin 2002 à résolution 4,4 (haut) et 26,4 km (bas) pour les trois sources de pluie et les sept sous-bassins (Figure 9.3). . . . .	144

10.8	Seuil observé et largeur de l'intervalle de confiance des pluies désagrégée (cercles) et aléatoire (carrés) à résolution (a) 4,4, (b) 8,8, (c) 13,2, (d) 17,6 et (e) 26,4 km pour le 11 juin 2002 sur les sept sous-bassins. . . . .	145
10.9	Largeur de l'intervalle de confiance des pluies désagrégée (cercles) et aléatoire (carrés) à résolution (a) 4,4, (b) 8,8, (c) 13,2, (d) 17,6 et (e) 26,4 km pour le 11 juin 2002 en fonction de la superficie des sept sous-bassins. . . . .	146
10.10	Largeur de l'intervalle de confiance de la pluie désagrégée du 11 juin 2002 à résolution 4,4 km en fonction de la superficie drainée par les 167 tronçons. . . . .	147
10.11	Largeur de l'intervalle de confiance de la pluie désagrégée pour le 11 juin 2002 pour les sept sous-bassins en fonction de la résolution du pixel. . . . .	148
10.12	Seuil observé et largeur de l'intervalle de confiance des écoulements désagrégé (cercles) et aléatoire (carrés) à résolution (a) 4,4, (b) 8,8, (c) 13,2, (d) 17,6 et (e) 26,4 km pour le 11 juin 2002 sur les sept sous-bassins. . . . .	149
10.13	Largeur de l'intervalle de confiance des écoulements en fonction de la largeur de l'intervalle de confiance des pluies désagrégée (cercles) et aléatoire (carrés) à résolution (a) 4,4, (b) 8,8, (c) 13,2, (d) 17,6 et (e) 26,4 km pour le 11 juin 2002 sur les sept sous-bassins. . . . .	150
10.14	Largeur de l'intervalle de confiance de l'écoulement désagrégé du 11 juin 2002 à résolution 4,4 km en fonction de la superficie des 167 tronçons. . . . .	151
10.15	Largeur de l'intervalle de confiance de l'écoulement désagrégé du 11 juin 2002 à résolution 4,4 km en fonction de la superficie des 81 tronçons de tête. . . . .	152
10.16	Largeur de l'intervalle de confiance de l'écoulement du 11 juin simulé (a et d) à partir des paramètres calés à 4,4 km, (b et e) à partir d'une valeur de $Z_3$ de 4 m et (c et f) à partir d'une lame de l'HGM de 0,0005 m en fonction de la largeur de l'intervalle de confiance de la pluie du 11 juin 2002 à résolution 4,4 (haut) et 26,4 km (bas) pour les deux sources de pluie et les sept sous-bassins. . . . .	153
D.1	Précipitation moyenne (mm/jour) sur les 801 jours (2006-2008) pour les observations (gauche) et pour les huit champs désagrégés (droite) pour les pixels de taille (a-b) 7,6, (c-d) 11,4 et (e-f) 15,2 km. . . . .	196
D.2	Coefficient de corrélation moyen dans les quatre directions calculé sur les jours avec (a) CAPE < 1000 J/kg et (b) CAPE > 1000 J/kg pour les pixels de 7,6 km observés (lignes continues) et désagrégés (lignes pointillées) : 45° (rouge), 90° (vert), 135° (bleu) et 0° (noir). . . . .	197

D.3	Coefficient de corrélation moyen dans les quatre directions calculé sur les jours avec (a) CAPE < 1000 J/kg et (b) CAPE > 1000 J/kg pour les pixels de 11,4 km observés (lignes continues) et désagrégés (lignes pointillées) : 45° (rouge), 90° (vert), 135° (bleu) et 0° (noir). . . . .	198
D.4	Coefficient de corrélation moyen dans les quatre directions calculé sur les jours avec (a) CAPE < 1000 J/kg et (b) CAPE > 1000 J/kg pour les pixels de 15,2 km observés (lignes continues) et désagrégés (lignes pointillées) : 45° (rouge), 90° (vert), 135° (bleu) et 0° (noir). . . . .	199
D.5	Précipitation journalière maximale (mm) sur les pixels de 7,6 (haut), 11,4 (centre) et 15,2 km (bas) durant la période 2006-2008 : observée (gauche), plus petite (milieu) et plus grande (droite) valeurs maximales journalières désagrégées.	200
D.6	Précipitations maximales observée et désagrégées sur les pixels de 7,6 km des 6 × 6 tuiles centrales pour chaque jour avec (a) CAPE < 1000 J/kg et (b) CAPE > 1000 J/kg. . . . .	201
D.7	Précipitations maximales observée et désagrégées sur les pixels de 11,4 km des 6 × 6 tuiles centrales pour chaque jour avec (a) CAPE < 1000 J/kg et (b) CAPE > 1000 J/kg. . . . .	202
D.8	Précipitations maximales observée et désagrégées sur les pixels de 15,2 km des 6 × 6 tuiles centrales pour chaque jour avec (a) CAPE < 1000 J/kg et (b) CAPE > 1000 J/kg. . . . .	203
E.1	Précipitation moyenne journalière (mm; haut), écart-type inter-journalier (mm; centre) et précipitation journalière maximale (mm; bas) calculés sur les 750 jours (2006-2008) observés (gauche) et désagrégés par les Modèles 1 (milieu) et 2 (droite) à 8,8 km. . . . .	206
E.2	Précipitation moyenne journalière (mm; haut), écart-type inter-journalier (mm; centre) et précipitation journalière maximale (mm; bas) calculés sur les 750 jours (2006-2008) observés (gauche) et désagrégés par les Modèles 1 (milieu) et 2 (droite) à 8,8 km. . . . .	207
E.3	Précipitation moyenne journalière (mm; haut), écart-type inter-journalier (mm; centre) et précipitation journalière maximale (mm; bas) calculés sur les 750 jours (2006-2008) observés (gauche) et désagrégés par les Modèles 1 (milieu) et 2 (droite) à 8,8 km. . . . .	208

E.4	Précipitation moyenne journalière (mm ; haut), écart-type inter-journalier (mm ; centre) et précipitation journalière maximale (mm ; bas) calculés sur les 750 jours (2006-2008) observés (gauche) et désagrégés par les Modèles 1 (milieu) et 2 (droite) à 8,8 km. . . . .	209
E.5	Précipitations maximales observée et produites par les huit séries désagrégées des Modèles (a) 1 et (b) 2 sur les pixels de 8,8 km des 4 × 4 tuiles centrales pour chaque jour. . . . .	210
E.6	Précipitations maximales observée et produites par les huit séries désagrégées des Modèles (a) 1 et (b) 2 sur les pixels de 13,2 km des 4 × 4 tuiles centrales pour chaque jour. . . . .	211
E.7	Précipitations maximales observée et produites par les huit séries désagrégées des Modèles (a) 1 et (b) 2 sur les pixels de 17,6 km des 4 × 4 tuiles centrales pour chaque jour. . . . .	212
E.8	Précipitations maximales observée et produites par les huit séries désagrégées des Modèles (a) 1 et (b) 2 sur les pixels de 26,4 km des 4 × 4 tuiles centrales pour chaque jour. . . . .	213
F.1	Écoulement journalier simulé en fonction de la pluie du 11 juin 2002 pour les 1000 séries désagrégées à 4,4 km sur les tronçons (a) 2, (b) 15, (c) 43 et (d) 98.	216
F.2	Seuil de précipitation en fonction (a) de la teneur en eau du sol au début de l'événement, (b) du type de sol et (c) de l'occupation du sol sur le sous-bassin.	217
F.3	Teneur en eau du sol au début de l'événement en fonction (a) du type de sol et (b) de l'occupation du sol sur le sous-bassin. . . . .	218
F.4	Valeur de $b_0$ de l'équation (F.1) en fonction (a) du seuil de précipitation, (b) de la teneur en eau du sol au début de l'événement, (c) du type de sol et (d) de l'occupation du sol sur le sous-bassin. . . . .	219
F.5	Valeur de $b_1$ de l'équation (F.1) en fonction (a) du seuil de précipitation, (b) de la teneur en eau du sol au début de l'événement, (c) du type de sol et (d) de l'occupation du sol sur le sous-bassin. . . . .	220
F.6	Valeur de $b_2$ de l'équation (F.1) en fonction (a) du seuil de précipitation, (b) de la teneur en eau du sol au début de l'événement, (c) du type de sol et (d) de l'occupation du sol sur le sous-bassin. . . . .	221
F.7	Écoulement journalier simulé à partir des calages à 4,4 (noir) et à 26,4 km (rouge) en fonction de la pluie du 11 juin 2002 pour les 1000 séries désagrégées à 4,4 km sur les tronçons (a) 2, (b) 15, (c) 43 et (d) 98. . . . .	222



# Chapitre 1

## INTRODUCTION

Ce chapitre explique d’abord pourquoi la mise à l’échelle spatiale de la précipitation constitue un sujet de recherche pertinent ayant des retombées potentielles en hydrologie tant pour des impératifs de prévision que de gestion à court et moyen termes. Ensuite, les objectifs de recherche sont énoncés. La structure de la thèse est présentée à la dernière section de ce chapitre.

### 1.1 Problématique

La précipitation est une variable déterminante pour plusieurs applications en hydrologie. Que ce soit pour la gestion des prélèvements des eaux d’une rivière, d’un réseau d’égout pluvial ou d’un barrage, il est important d’avoir des mesures ou des estimations des distributions spatio-temporelles des précipitations. Étant donné leur grande variabilité spatiale (Wood *et al.*, 2000a; Jensen et Pederson, 2005), il faut des données de précipitation à une résolution assez fine pour répondre aux besoins de ces diverses applications.

#### 1.1.1 Données de précipitation pour le passé récent

La précipitation est mesurée par des stations météorologiques. Comme il s’agit de mesures directes, les données de précipitation aux stations météorologiques sont souvent utilisées comme valeurs de référence pour valider la précipitation estimée à partir de radars météorologiques (*e.g.* Creutin *et al.*, 1997; Wood *et al.*, 2000b; Golz *et al.*, 2006) ou simulée par un modèle de prévision météorologique ou encore un modèle régional de climat (MRC; *e.g.* Sanso et Guenni, 2004; Gagnon *et al.*, 2009; Konan *et al.*, 2010). Les données de stations météorologiques ne sont cependant pas exemptes de toute erreur. Par exemple, la

sous-captation causée par le vent peut être importante, atteignant jusqu'à 40 % pour la précipitation solide (Goodison *et al.*, 1998; Yang *et al.*, 1999). Pour les augets basculants, un type de pluviomètre couramment utilisé, la sous-captation peut être importante pour les événements de précipitation intenses.

La principale faiblesse des stations météorologiques est qu'elles ne fournissent que des valeurs ponctuelles. Pour avoir une couverture spatiale complète sur un territoire donné à partir de stations météorologiques, il faut appliquer une méthode d'interpolation, comme le krigeage (Baillargeon, 2005). C'est le cas dans le sud du Québec où le Centre d'Expertise Hydrique du Québec (CEHQ) a développé une grille météorologique d'une résolution de  $0,1^\circ$ , soit d'environ 10 km (Turcotte *et al.*, 2007; Bourdillon *et al.*, 2011). Il faut noter qu'en général, moins le réseau de stations météorologiques est dense, plus le choix de la méthode d'interpolation est important (Christensen *et al.*, 1998; Renard et Comby, 2006). Un réseau peu dense ne capte pas toutes les variations locales causées par exemple par la topographie ou par la proximité d'un plan d'eau.

Pour obtenir une estimation de la hauteur de précipitation tombée sur toute une région sans passer par l'interpolation de données ponctuelles, des outils qui mesurent indirectement la précipitation, comme les radars météorologiques, peuvent être utilisés. Un tel radar émet des impulsions électromagnétiques dans toutes les directions, impulsions en partie réfléchies par les gouttes d'eau (Lee *et al.*, 2007). Connaissant la direction de l'impulsion et le temps écoulé avant que celle-ci ne soit réfléchi vers le radar, il est possible de déterminer la position de la goutte. L'intensité de la précipitation est estimée à partir d'une relation entre la réflectivité ( $Z$ ) et la précipitation, appelée relation  $Z - R$  ( $R$  pour *rain*; Fulton *et al.*, 1998; Lee *et al.*, 2007) ou  $Z - S$  ( $S$  pour *snow*; Saltikoff *et al.*, 2000), de la forme  $Z = aR^b$ . Les radars météorologiques peuvent estimer l'intensité de la précipitation sur un domaine circulaire d'un rayon pouvant atteindre 250 km à une fine résolution ( $< 5$  km). Cependant, ces données sont sujettes à plusieurs sources d'erreurs (Joss et Germann, 2000), notamment les échos de terrain (problématiques surtout en régions montagneuses), la bande brillante (les hydrométéores passant de l'état solide à l'état liquide reflètent beaucoup plus d'énergie que des gouttes de pluie, surestimant ainsi la précipitation), la propagation anormale (surréfraction ou infraréfraction causées par des inversions de température) et l'atténuation (lorsque la précipitation éloignée du radar est masquée par la précipitation plus près du radar).

Il existe certaines cartes dites *composites*, qui allient la fiabilité des données de stations météorologiques et la couverture spatiale des données radars. Aux États-Unis, le *National Weather Service* (NWS) a créé un produit appelé *Stage II* qui donne des estimations de la précipitation horaire pour chaque radar en tenant compte des précipitations mesurées aux stations météorologiques (Fulton *et al.*, 1998). Ensuite, le NWS a créé *Stage III* qui regroupe les données de plusieurs radars pour former une grille pour chacun des 12 *River Forecast Centers* (RFC) américains. En plus de couvrir un plus large territoire que *Stage II*, les données aux six heures de *Stage III* sont soumises à un contrôle de qualité manuel (Fulton *et al.*, 1998). Finalement, *Stage IV* regroupe les données de chaque RFC pour former une grille couvrant toute la partie continentale des États-Unis. Les données de *Stage IV* depuis 2002 sont disponibles en ligne (<http://data.eol.ucar.edu>) avec contrôle de qualité manuel au pas de temps journalier ou de six heures et sans contrôle de qualité manuel au pas de temps horaire. Au Canada, Environnement Canada développe actuellement un modèle générant la précipitation au pas de temps de six heures sur une grille à résolution 15 km couvrant l'Amérique du Nord en fusionnant et interpolant différentes sources de données (pluviomètres, radars, satellites, modèles atmosphériques ; Mahfouf *et al.*, 2007).

Les données de satellites météorologiques peuvent aussi estimer les hauteurs de précipitation. La résolution spatiale varie d'environ 100 km pour les satellites géostationnaires à moins de 5 km pour les satellites circumpolaires (Forman *et al.*, 2008). Les satellites météorologiques peuvent estimer la précipitation à des endroits inaccessibles pour les radars météorologiques, comme les régions à topographie complexe ou au-dessus des océans. Cependant, il est difficile d'évaluer correctement l'incertitude associée à l'estimation des hauteurs de précipitation à partir de satellites météorologiques (Teo et Grimes, 2007) et, selon l'information consultée, il n'existe pas de jeux de données de précipitation fiables dérivées des satellites météorologiques qui couvrent de larges territoires sur une longue période de temps à fine résolution spatio-temporelle.

### 1.1.2 Données de précipitation pour le futur

Les données de précipitation pour le futur peuvent provenir soit de *prévisions météorologiques* ou de *projections climatiques*. La prévision météorologique est une extrapolation déterministe qui se veut le plus près possible de l'état de l'atmosphère à court terme (*i.e.* prochaines heures, prochains jours) alors qu'une projection climatique produit une réalisation possible

du climat évaluée sur une longue période (*i.e.* quelques décennies). Dans les deux cas, il n'est pas adéquat de simplement extrapoler statistiquement les données de précipitation passées. Pour les prévisions météorologiques, l'état actuel de l'atmosphère doit être pris en compte puisqu'il définit l'évolution à court terme. Pour les projections climatiques, dans un contexte de changements climatiques (CC) notamment, les caractéristiques statistiques des précipitations en climat futur pourraient être différentes de celles en climat historique.

Les prévisions météorologiques et les projections climatiques sont faites à l'aide de modèles numériques simulant les processus physiques gouvernant l'état de l'atmosphère à partir des lois de conservation de masse, d'énergie et de mouvement (Peixoto et Oort, 1992). Pour les prévisions météorologiques, les conditions initiales de la simulation sont déterminantes pour la prévision à court terme. À l'inverse, pour les projections climatiques, une période de rodage d'au moins un an, durant laquelle les sorties ne sont pas retenues, est spécifiée pour que les conditions initiales soient « oubliées » et que le climat atteigne son équilibre. Par contre, les projections climatiques doivent tenir compte de l'évolution dans le temps de la concentration de gaz à effet de serre (GES) dans l'atmosphère. Comme cette variable évolue en fonction de facteurs anthropiques, divers scénarios d'émissions de GES ont été élaborés (Nakicenovic et Swart, 2000).

Les modèles climatiques se divisent en deux groupes selon le domaine spatial sur lequel ils sont appliqués : il y a les modèles de climat global (MCG) et les modèles régionaux de climat (MRC). Les MCG simulent le climat sur une grille en trois dimensions couvrant toute la planète. Pour obtenir des résultats sur plusieurs décennies sur toute la planète, le temps de calcul nécessaire est considérable. C'est pourquoi la résolution horizontale des MCG est de l'ordre de  $10^2$  km (Fowler *et al.*, 2007). Les MRC simulent les mêmes processus physiques que les MCG, bien que certains éléments de leur paramétrisation diffèrent, mais leur grille tri-dimensionnelle ne couvre qu'une région du globe. Les conditions aux frontières du domaine proviennent des sorties de simulation d'un MCG ou, dans le cas d'une simulation en passé récent, de réanalyses d'observations (*e.g.* Mesinger *et al.*, 2004). Comme la région à simuler est plus petite, la résolution spatiale d'un MRC est plus fine que celle d'un MCG. Les termes grande échelle et mésoéchelle (*large scale* et *mesoscale* en anglais) sont souvent employés pour désigner l'échelle spatiale des MCG et des MRC respectivement (Perica et Foufoula-Georgiou, 1996). Il existe aussi des MCG à résolution variable, qui simulent à mésoéchelle sur une région d'intérêt et à grande échelle sur le

reste du globe (ARPEGE-Climat ; Déqué, 2007). La résolution horizontale d'un MRC varie typiquement de 25 à 50 km, mais certaines simulations sont maintenant disponibles à de plus fines résolutions (Maraun *et al.*, 2010). C'est le cas par exemple pour le Modèle Régional Canadien du Climat (MRCC) qui a produit quelques simulations à une résolution de 15 km (Daniel Caya, communication personnelle).

Il est aussi possible de produire des simulations à plus fine résolution spatiale en utilisant un MRC sur un domaine emboîté dans un autre domaine plus grand. Les données aux frontières sont alors fournies par la simulation faite à partir du MRC sur le plus grand domaine. Yu *et al.* (2002) ont utilisé un MRC sur trois domaines emboîtés pour des résolutions de 108, 36 et 12 km. Lakhtakia *et al.* (1999) avaient fait de même, mais pour des résolutions de 36, 12 et 4 km. Cependant, un modèle construit pour représenter les processus physiques à une résolution de 36 km n'est peut-être pas adéquat pour représenter les processus à 4 km. Aussi, cette approche demande un long temps de calcul puisqu'elle fait intervenir plusieurs simulations climatiques.

En prévision météorologique, une contrainte s'ajoute puisque les résultats doivent être produits rapidement. Par contre, la période à simuler est beaucoup plus courte qu'en modélisation climatique, ce qui permet de produire des simulations à l'échelle mondiale à une résolution spatiale comparable à celle des MRC. Par exemple, les prévisions opérationnelles d'Environnement Canada sont faites sur une grille d'environ 33 km de résolution (Velazquez *et al.*, 2009). Il faut noter qu'Environnement Canada produit aussi des prévisions sur une grille de 2,5 km pour les 24 prochaines heures, mais seulement sur certaines régions du pays. Néanmoins, dans la majorité des cas, la plus fine résolution disponible pour la précipitation future ( $\approx 15$  km) est généralement plus grossière que la plus fine résolution disponible pour la précipitation du passé récent issue de radars météorologiques et de cartes composites ( $< 5$  km). Pour obtenir des données de précipitation future à la même échelle que les données disponibles pour le passé récent, il faut appliquer une méthode de mise à l'échelle spatiale sur les données des modèles météorologiques ou climatiques. Une bonne méthode de mise à l'échelle de la précipitation pourrait par exemple améliorer l'estimation des débits de pointe. En prévision météorologique, cela peut se traduire par une meilleure réaction des autorités en cas de crue imminente simulée en intégrant les données de précipitation et de température à un modèle hydrologique distribué. En projection climatique, cela peut se traduire par une meilleure évaluation de l'impact des CC à l'échelle locale et à l'élaboration de stratégies

d'adaptation.

### 1.1.3 Mise à l'échelle spatiale de la précipitation

La mise à l'échelle d'une grille de précipitation peut se faire : (i) en ramenant la précipitation sur un ou des points d'intérêts (approche *grille vers points*), ou (ii) en ramenant la précipitation sur une grille à plus fine résolution (approche *grille vers grille*). Le choix de l'une ou l'autre des approches dépend du type de données de précipitation disponibles pour développer et valider la méthode. Si on a accès à une grille de précipitation à fine résolution spatiale (*e.g.* carte radar), il est possible d'appliquer l'approche grille vers grille. Si les seules données de précipitation fiables disponibles proviennent de stations météorologiques, il n'y a d'autre choix que de développer une méthode selon l'approche *grille vers points*.

Les méthodes d'interpolation linéaire, telles la pondération inverse à la distance (*inverse distance weighting* en anglais) et le krigeage, sont parfois utilisées en supposant que la précipitation sur une maille de la grille d'un modèle météorologique ou climatique a été observée sur un point situé au centre de la maille (Ludwig *et al.*, 2003; Konan *et al.*, 2010). Cette façon de faire peut générer des champs de précipitation trop lisses, surtout si l'interpolation est faite sur de courtes échelles temporelles (*e.g.* horaire, journalière). Les méthodes d'interpolation doivent être adaptées pour tenir compte du fait que les données à interpoler sont des moyennes spatiales sur une surface et non des données ponctuelles. Kyriakidis (2004) a adapté le krigeage pour en faire une approche *grille vers points*.

Pour la mise à l'échelle de la précipitation simulée par un modèle climatique, les méthodes *grille vers points* sont appelées *statistical downscaling* en anglais, et parfois raffinement ou désagrégation statistique en français (Herrera *et al.*, 2006). Ces méthodes peuvent être classées en deux catégories : (i) les fonctions de transfert et (ii) les générateurs météorologiques stochastiques. Les fonctions de transfert consistent à utiliser des variables autres que la précipitation simulée par le modèle climatique pour générer des champs de précipitation à l'échelle de la station. Cette méthode a été très utilisée puisqu'avant l'arrivée des MRC, les seules projections disponibles provenaient de MCG et donnaient des champs de précipitation sur une grille à résolution de quelques centaines de kilomètres. Avec une grille si grossière, les MCG ne peuvent simuler adéquatement les hauteurs de précipitation, surtout si l'événement est convectif (Entekhabi et Eagleson, 1989). Pour palier ce problème, d'autres champs atmosphériques simulés par le MCG (*e.g.* taux d'humidité, pression, température)

sont utilisés comme variables prédictrices de la hauteur de précipitation tombée localement. La relation statistique entre la précipitation locale et les champs atmosphériques est déterminée de manière empirique à partir de la précipitation de stations météorologiques et de variables atmosphériques provenant de réanalyses d'observations. Cette méthode est généralement utilisée pour mettre à l'échelle la précipitation d'un MCG (*e.g.* Wilby *et al.*, 1999; Salathé *et al.*, 2007), mais peut aussi être utilisée pour mettre à l'échelle la précipitation d'un MRC (Hellström et Chen, 2003). Bien que certaines de ces méthodes tiennent compte de la corrélation spatiale entre les sites considérés (*e.g.* Hughes *et al.*, 1999), elle ne donnent pas une couverture spatiale complète d'un territoire donné.

Les générateurs météorologiques stochastiques (*weather generators* en anglais, voir *e.g.* Yates *et al.*, 2003; Sharif et Burn, 2004; Brissette *et al.*, 2007; Caron *et al.*, 2008; Khalili *et al.*, 2011) simulent des séries synthétiques de variables météorologiques, le plus souvent la précipitation et la température, à partir des propriétés statistiques d'une série observée. La série observée peut être utilisée pour définir les paramètres nécessaires à la génération des séries synthétiques (*e.g.* probabilité d'occurrence de la précipitation au jour  $j$  sachant l'occurrence ou non de la précipitation au jour  $j - 1$ , intensité moyenne de la précipitation sachant qu'il y a occurrence; Brissette *et al.*, 2007). Il est également possible de rééchantillonner aléatoirement des valeurs de la série observée pour générer les séries synthétiques (Yates *et al.*, 2003; Sharif et Burn, 2004). Les données observées proviennent de stations météorologiques, mais des perturbations peuvent être ajoutées pour tenir compte des CC simulés par les modèles climatiques (Khalili *et al.*, 2011). Les générateurs météorologiques stochastiques peuvent aussi tenir compte de la corrélation spatiale et ainsi produire des champs en plusieurs points sur un territoire donné (*e.g.* Brissette *et al.*, 2007). Pour une revue plus complète des méthodes de *statistical downscaling*, le lecteur est invité à consulter Fowler *et al.* (2007) ou Herrera *et al.* (2006) en français.

L'approche grille vers grille distribue spatialement la précipitation d'une grille sur une autre grille à plus fine résolution. Certains auteurs utilisent le terme *downscaling* en anglais pour référer à cette approche (Ferraris *et al.*, 2003), mais le terme *disaggregation* est plus souvent employé. Dans la présente thèse, le terme **désagrégation** référerà à l'approche grille vers grille. À noter qu'il faut distinguer la définition du terme désagrégation utilisée ici de celle utilisée par Herrera *et al.* (2006), qui emploient le terme désagrégation statistique pour désigner la mise à l'échelle selon l'approche *grille vers points*. Contrairement aux méthodes

de *statistical downscaling* (Fowler *et al.*, 2007), la désagrégation considère le volume de précipitation sur la grille originale comme étant fiable et le volume de précipitation sur la grille désagrégée est égal au volume sur la grille originale.

Étant donné la faible quantité de données observées à fine résolution, les modèles de désagrégation ont souvent, pour ne pas dire toujours, été développés à partir de quelques événements seulement. Avec des jeux de données maintenant disponibles sur quelques années et sur un grand territoire, il est maintenant possible de développer un modèle de désagrégation à partir d'une grande quantité d'événements couvrant différentes conditions atmosphériques et différentes caractéristiques physiographiques. À notre connaissance, aucune étude de ce genre n'a été publiée à ce jour.

## 1.2 Objectifs de la thèse

L'objectif principal de la thèse est de développer un modèle statistique de désagrégation spatiale d'une grille mésoéchelle de précipitation capable de produire des champs de précipitation *réalistes* à fine résolution pour un événement journalier donné. Un champ de précipitation désagrégé est dit *réaliste* s'il reproduit les propriétés statistiques du champ de précipitation observé à fine résolution, notamment en matière de valeurs extrêmes (Pegram et Clothier, 2001), d'anisotropie et de la corrélation spatiale (Lovejoy et Schertzer, 2010b). Les données journalières de *Stage IV* sont utilisées dans le cadre de cette thèse comme précipitation de référence. La grande quantité de données disponibles permet de développer et de valider le modèle de désagrégation sur des périodes distinctes chacune couvrant un grand nombre d'événements journaliers. Les propriétés statistiques de la précipitation à fine résolution peuvent ainsi être évaluées pour diverses conditions météorologiques.

Lors du développement du modèle, les équations simples (*i.e.* faible nombre de paramètres à estimer, faciles à interpréter) seront privilégiées. Un modèle simple est souhaitable pour deux raisons. Premièrement, il permet de produire rapidement plusieurs séries de précipitation équiprobables à partir d'une série donnée ; un atout important surtout pour la désagrégation de la précipitation de prévisions météorologiques à court terme. Deuxièmement, un modèle ayant peu de paramètres à estimer peut plus facilement être adapté à différentes régions.

Trois objectifs spécifiques ont été définis et sont décrits aux sections suivantes. Le cadre de travail est le suivant :

- La résolution de la grille à désagréger, nommée grille *agrégée*, est d'environ 50 km ;
- La résolution de la grille désagrégée varie d'environ 4 à 25 km ;
- La désagrégation se fait au pas de temps journalier.

La résolution de la grille à désagréger est comparable à la résolution d'un MRC. Le pas de temps journalier a été choisi pour le développement du modèle notamment parce que les données à fine résolution utilisées sont soumises à un contrôle de qualité manuel au pas de temps journalier, mais pas au pas de temps horaire.

### 1.2.1 Premier objectif spécifique

Le premier objectif spécifique est de développer le modèle de désagrégation statistique sur une région avec peu de relief, c'est-à-dire sur une région sans variation importante de l'altitude. Une région avec peu de relief est sélectionnée parce qu'en première approche, il est souhaitable de développer le modèle dans des conditions aussi simples que possible. La région étudiée est centrée sur l'état de la Géorgie, mais touche également à la Floride, l'Alabama et la Caroline du Sud, dans le sud-est des États-Unis, et couvre environ  $400 \times 400$  km<sup>2</sup>. Cette région a été choisie notamment parce qu'il s'y produit beaucoup d'événements pluvieux, de stratiformes à très convectifs. L'évaluation des propriétés statistiques de la précipitation peut se faire sous diverses conditions météorologiques. Aussi, les précipitations sous forme de neige sont négligeables, ce qui permet de focaliser uniquement sur la pluie en première approche.

Quatre années de données sont utilisées pour développer le modèle, c'est-à-dire pour déterminer les variables pertinentes expliquant la structure spatiale de la précipitation à fine résolution et pour estimer les paramètres du modèle. Trois autres années sont utilisées pour valider le modèle, c'est-à-dire pour désagréger la précipitation. La précipitation désagrégée est comparée avec la précipitation observée à fine résolution pour s'assurer que le modèle reproduit adéquatement les propriétés statistiques de la précipitation et aussi pour évaluer la valeur ajoutée, en terme de réalisme, qu'apporte la désagrégation par rapport à la précipitation agrégée à 50 km.

## 1.2.2 Deuxième objectif spécifique

Les modèles météorologiques et climatiques peuvent tenir compte de la précipitation orographique à mésoéchelle, mais ne peuvent en tenir compte à des échelles plus fines que leur résolution spatiale. Le deuxième objectif spécifique est d'adapter le modèle de désagrégation développé dans le cadre du premier objectif spécifique afin de tenir compte de la topographie.

La région étudiée se situe dans les États de Washington et de l'Oregon, au nord-ouest des États-Unis. Quatre années sont utilisées pour estimer les paramètres du modèle et trois autres années sont utilisées pour valider le modèle. La valeur ajoutée qu'apporte la précipitation désagrégée de ce modèle est évaluée par rapport à la précipitation agrégée à 50 km, mais aussi par rapport à la précipitation désagrégée par le modèle développé sur une région avec peu de relief.

## 1.2.3 Troisième objectif spécifique

La désagrégation de la précipitation d'un modèle météorologique ou climatique a pour but de ramener la précipitation à une échelle adéquate pour des fins de modélisation hydrologique. D'un point de vue pratique, la désagrégation sera utile et pertinente seulement si la précipitation désagrégée permet une meilleure estimation des débits que la précipitation agrégée. Le troisième objectif spécifique est d'évaluer les conditions pour lesquelles la précipitation désagrégée est susceptible de donner une meilleure information sur le débit que la précipitation agrégée.

La désagrégation peut améliorer les résultats d'une simulation hydrologique de deux façons :

- En donnant des débits plus près de la valeur observée ;
- En donnant, à partir de plusieurs champs désagrégés, un intervalle de confiance du débit permettant une meilleure prise de décision en plusieurs points d'un bassin versant dans le cas d'un événement intense qui pourrait avoir des conséquences plus importantes localement (exutoire d'un sous-bassin de quelques km<sup>2</sup>) que régionalement (exutoire d'un bassin de plus de 10<sup>3</sup> km<sup>2</sup>).

La désagrégation est faite pour un événement intense survenu en 2002 sur le bassin versant de la rivière des Anglais, un sous-bassin de 730 km<sup>2</sup> du bassin de la rivière Châteauguay, situé

sur la rive sud du fleuve St-Laurent. Il est attendu que l'impact de la désagrégation devrait être plus fort pour les bassins de plus petites superficies (Carpenter et Georgakakos, 2006). Le modèle hydrologique distribué HYDROTEL (Fortin *et al.*, 2001a,b; Turcotte *et al.*, 2003) est utilisé pour estimer le débit à différentes échelles spatiales. Comme il n'y a qu'une seule station hydrométrique active sur le bassin en 2002, la comparaison entre le débit simulé à partir de la pluie désagrégée et le débit observé ne pourrait se faire qu'à un endroit. C'est pourquoi le débit simulé à partir de la précipitation observée est utilisé comme débit de référence.

### 1.3 Structure de la thèse

La thèse est divisée en trois parties, une pour chaque objectif spécifique. La première partie comprend une revue des modèles de désagrégation statistique qui ont été proposés jusqu'à maintenant pour des régions avec peu de relief (Chapitre 2), présente le modèle proposé (Chapitre 3) et les résultats de la désagrégation (Chapitre 4). Les travaux de la première partie ont fait l'objet d'une publication (Gagnon *et al.*, 2012a). Toutefois, la première partie de la présente thèse contient de nouveaux résultats par rapport à Gagnon *et al.* (2012a). La deuxième partie comprend une revue des méthodes pour quantifier la précipitation orographique à l'échelle locale (Chapitre 5), explique comment le modèle proposé au chapitre 3 est adapté pour tenir compte de la topographie (Chapitre 6) et présente les résultats de la désagrégation (Chapitre 7). L'essentiel des travaux de la deuxième partie se retrouve dans Gagnon *et al.* (2012b). La troisième partie présente une revue de littérature sur l'impact de la résolution spatiale de la précipitation en modélisation hydrologique (Chapitre 8), décrit la méthode utilisée pour produire les simulations hydrologiques (Chapitre 9) et présente les résultats des simulations hydrologiques réalisées à partir de la précipitation désagrégée (Chapitre 10). En conclusion, un résumé des constats des travaux de la thèse et des axes de recherche à poursuivre sont présentés (Chapitre 11).



## Première partie

# DÉSAGRÉGATION STATISTIQUE DE LA PRÉCIPITATION SUR UNE RÉGION AVEC PEU DE RELIEF



# Chapitre 2

## REVUE DE LITTÉRATURE

Ce chapitre présente les modèles qui ont été proposés pour produire des champs de précipitation sur une grille à fine échelle à partir d'une grille mésoéchelle. Pour le reste de la thèse, les termes *tuile* et *pixel* sont employés pour désigner une maille de la grille mésoéchelle à désagréger et une maille de la grille à fine résolution, respectivement. Rappelons que seuls les modèles qui distribuent spatialement la précipitation partout à l'intérieur de la tuile sont d'intérêt. Cela exclut les méthodes grille vers points (Chapitre 1). Cela exclut aussi les modèles de désagrégation qui utilisent la précipitation de stations météorologiques en entrée en plus de la précipitation sur la tuile (voir *e.g.* Fiorucci *et al.*, 2001).

Pour les modèles présentés dans ce chapitre, tous les pixels à l'intérieur d'une tuile présentent des caractéristiques géographiques homogènes. Cela implique qu'aucune variable permettant de différencier les pixels, comme la topographie ou la proximité d'un océan, ne peut être considérée. Les modèles considérant la topographie seront présentés au Chapitre 5. Cependant, des variables atmosphériques inhérentes à l'événement de précipitation peuvent être considérées (*e.g.* vitesse et direction du vent, énergie de convection).

Trois types d'approches ont été identifiés dans la littérature pour représenter la distribution spatiale de la précipitation à fine échelle : les cellules de précipitation, les modèles multi-échelles et les processus stochastiques.

## 2.1 Cellules de précipitation

Lors d'un événement convectif, il se forme des cellules de précipitation de quelques kilomètres de diamètre dans lesquelles la pluie est plus intense. Certains modèles de désagrégation reproduisent ce phénomène en distribuant spatialement les centres de cellules selon un processus de Poisson uniforme sur le domaine (Cowpertwait *et al.*, 2002; Ferraris *et al.*, 2003; Burton *et al.*, 2008). L'intensité de précipitation en un endroit donné est la somme des intensités de toutes les cellules évaluées à cet endroit (Figure 2.1). Ces modèles sont l'analogie spatiale des processus de Neyman-Scott et des processus de Bartlett-Lewis utilisés pour distribuer temporellement la précipitation en un point donné.

Les cellules peuvent être identiques (Ferraris *et al.*, 2003) ou de superficies aléatoires (Cowpertwait *et al.*, 2002; Burton *et al.*, 2008). L'intensité de pluie dans une cellule peut être uniforme dans toute la cellule ou décroître en s'éloignant du centre. Dans la littérature consultée, les cellules sont toujours circulaires, ce qui implique que les modèles supposent l'isotropie (*i.e.* les propriétés statistiques de la précipitation sont les mêmes dans toutes les directions), mais il serait possible d'inclure l'anisotropie en considérant des cellules de formes elliptiques.

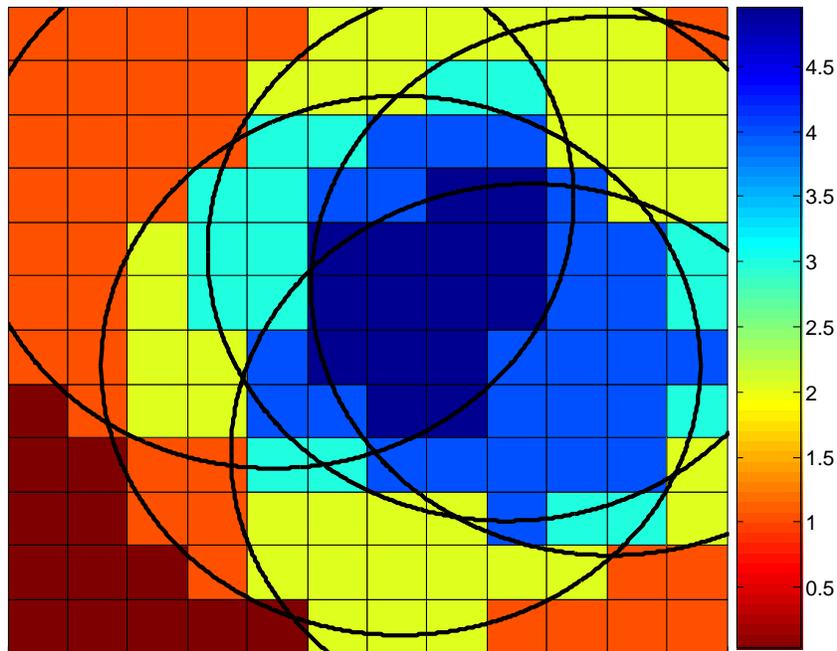


Figure 2.1 – Exemple synthétique de précipitation désagrégée selon l'approche des cellules de précipitation. Les cercles noirs sont les cellules à l'intérieur desquelles l'intensité est égale à 1 unité.

À l'exception de Ferraris *et al.* (2003), qui utilisent des données de radars météorologiques, ces modèles sont calés à partir de stations météorologiques. Ces modèles ont été conçus pour représenter simplement les cellules de précipitation, mais ils n'utilisent pas toute l'information spatiale contenue dans les données de radars météorologiques, comme par exemple la corrélation spatiale à fine échelle. De plus, il est difficile d'estimer les paramètres de ces modèles (Krajewski *et al.*, 1996). Cowpertwait *et al.* (2002) estiment les six paramètres de leur modèle en six étapes différentes tandis que Burton *et al.* (2008) utilisent cinq méthodes différentes pour estimer les paramètres de leur modèle.

## 2.2 Modèles multi-échelles

L'approche la plus répandue en désagrégation spatiale est de supposer que la précipitation a des propriétés *multi-échelles* spatiales, c'est-à-dire que la distribution statistique de la précipitation est identique d'une échelle spatiale à l'autre, à un paramètre d'échelle près. Cette approche est populaire parce qu'elle donne des modèles ayant peu de paramètres et faciles à estimer. Deux types de modèles multi-échelles sont répandues : (i) les modèles multifractals et (ii) les modèles d'ondelettes.

Les modèles multifractals se basent sur une généralisation de la théorie des fractals. Soit  $X_k(i, j)$  la précipitation à la résolution spatiale  $k$  sur le pixel  $(i, j)$  en un pas de temps donné et  $S_q(k) = \mathbb{E}[X_k^q]$ , le  $q$ -ième moment statistique des valeurs de  $X_k$ . La précipitation a des propriétés *multifractales* spatiales si

$$S_q(k) \sim k^{\zeta(q)} \quad (2.1)$$

avec  $\zeta(q)$  étant une fonction non-linéaire de  $q$  qui ne dépend pas de  $k$  (Deidda, 2000). Avant que la multifractalité spatiale soit utilisée en désagrégation spatiale, la multifractalité temporelle a été analysée en désagrégation temporelle pour la précipitation de stations météorologiques (Menabde *et al.*, 1997; Olsson, 1998; Güntner *et al.*, 2001; Licznar *et al.*, 2011). Certains modèles désagrègent dans le temps et dans l'espace à partir des propriétés multifractales spatio-temporelles de la précipitation (Over et Gupta, 1996; Deidda, 2000). Dans cette section, seule la multifractalité spatiale est considérée.

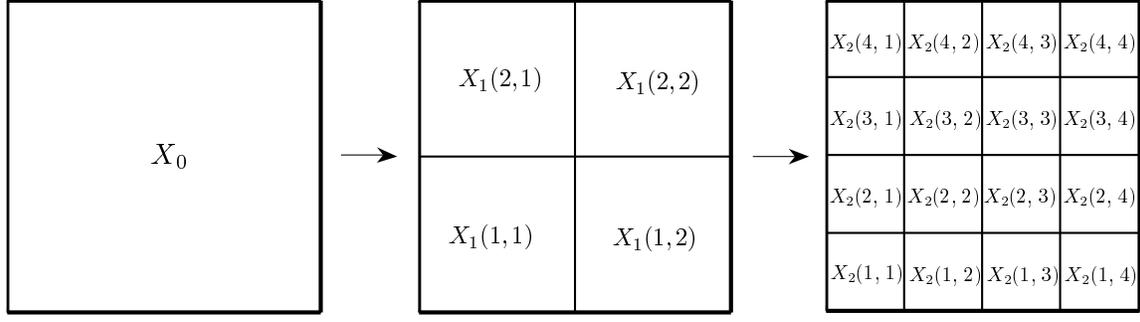


Figure 2.2 – Illustration de la désagrégation selon l'approche multifractale.

Bien que la théorie pour exprimer la multifractalité de la précipitation ne soit pas triviale (voir *e.g.* Over et Gupta, 1996 ; Deidda, 2000 ; Lovejoy et Schertzer, 2010a,b), les modèles de désagrégation multifractals sont simples à conceptualiser (Figure 2.2). Soit  $X_0$  la précipitation sur la tuile à désagréger. La première étape consiste à diviser la tuile en  $a^2$  sous-tuiles de tailles identiques. Pour le reste de la section,  $a = 2$ , comme pour la grande majorité des applications. Ensuite, une valeur de précipitation est générée sur chacune de ces  $2^2 = 4$  sous-tuiles à partir de quatre facteurs aléatoires multiplicatifs. La distribution statistique de ces facteurs est choisie de sorte que l'équation (2.1) soit respectée. Ces étapes sont répétées jusqu'à ce que les sous-tuiles aient la résolution spatiale désirée. Cette approche par étape est dite *en cascade multiplicative*. Notons  $X_k(i, j)$  la précipitation générée à l'échelle  $k$  sur la sous-tuile à la position  $(i, j) \{i, j = (1, \dots, 2^k), k \geq 1\}$ . La précipitation sur la sous-tuile parente est en fait  $X_{k-1}(\lceil i/2 \rceil, \lceil j/2 \rceil)$  où l'opérateur  $\lceil x \rceil$  extrait le plus petit entier supérieur ou égal à  $x$  (voir Figure 2.2). Par définition,  $X_0 \equiv X_0(1, 1)$ . La distribution statistique de  $X_k(i, j)$  peut dépendre uniquement de la précipitation sur la sous-tuile parente, mais dépend généralement aussi de l'échelle  $k$ .

Soit  $Y_k(i, j) = X_k(i, j)/X_{k-1}(\lceil i/2 \rceil, \lceil j/2 \rceil)$ , la perturbation aléatoire qui permet de générer la précipitation à la sous-tuile  $(i, j)$  à l'échelle  $k$ . Forman *et al.* (2008) supposent que, pour chaque  $k$ , les  $Y_k(i, j)$  sont indépendants et identiquement distribués (iid) selon la distribution lognormale d'espérance 1 et de variance  $\sigma_k^2 = \sigma_0^2 2^{-2kH}$  où les deux paramètres à estimer sont  $\sigma_0^2$ , la variance à l'échelle de la tuile, et  $H$ , appelé l'*exposant d'échelle*. Over et Gupta (1996) ont traité  $Y_k(i, j)$  à l'aide d'une distribution mixte appelée  $\beta$ -lognormale. Ce modèle est plus adapté lorsque la précipitation ne couvre pas toute la tuile. Le paramètre  $\beta$  définit la probabilité que la précipitation soit nulle et la distribution lognormale est utilisée pour générer les valeurs de  $Y_k(i, j)$  lorsqu'elles sont plus grandes que 0. Over et Gupta (1996)

supposent que la variance de la précipitation est constante en fonction de l'échelle, *i.e.*  $H = 0$ . Ils n'ont que deux paramètres à estimer,  $\beta$  et la variance de la lognormale, qui dépendent tous deux de l'intensité de précipitation sur la tuile. Le modèle plus général qui n'utilise pas nécessairement la distribution lognormale et pour lequel les paramètres peuvent dépendre de l'échelle est nommé  $\beta$ -logstable (Veneziano *et al.*, 2006). Avec ce type de modèles, le volume total d'eau désagrégé converge asymptotiquement vers le volume d'eau agrégé, puisque l'espérance de l'intensité de la précipitation sur une sous-tuile est égale à celle de la sous-tuile parente, mais les deux volumes ne sont pas exactement les mêmes. Un facteur multiplicatif peut être appliqué à chaque échelle pour s'assurer de la conservation du volume d'eau tombé.

Une autre approche populaire en désagrégation multi-échelle consiste à utiliser les ondelettes de Haar pour générer les fluctuations d'une échelle à l'autre (*e.g.* Perica et Foufoula-Georgiou, 1996; Harris et Foufoula-Georgiou, 2001; Bacchi et Ranzi, 2003). Au lieu de générer des perturbations multiplicatives par rapport à la sous-tuile parente, ce sont les fluctuations latitudinales ( $\xi_k^{(1)}$ ), longitudinales ( $\xi_k^{(2)}$ ) et diagonales ( $\xi_k^{(3)}$ ) qui sont simulées :

$$\xi_k^{(1)}(i, j) = \frac{1}{4} \begin{bmatrix} \{X_{k+1}(2i-1, 2j-1) + X_{k+1}(2i-1, 2j)\} \\ - \{X_{k+1}(2i, 2j-1) + X_{k+1}(2i, 2j)\} \end{bmatrix}; \quad (2.2)$$

$$\xi_k^{(2)}(i, j) = \frac{1}{4} \begin{bmatrix} \{X_{k+1}(2i-1, 2j-1) + X_{k+1}(2i, 2j-1)\} \\ - \{X_{k+1}(2i-1, 2j) + X_{k+1}(2i, 2j)\} \end{bmatrix}; \text{ et} \quad (2.3)$$

$$\xi_k^{(3)}(i, j) = \frac{1}{4} \begin{bmatrix} \{X_{k+1}(2i-1, 2j-1) - X_{k+1}(2i, 2j-1)\} \\ - \{X_{k+1}(2i-1, 2j) - X_{k+1}(2i, 2j)\} \end{bmatrix}. \quad (2.4)$$

Comme pour les modèles multifractals, les modèles à ondelettes utilisent aussi une approche *en cascade* puisque la désagrégation se fait de  $x$  km à  $x/2$  à  $x/4$  jusqu'à la résolution voulue. Pour les trois fluctuations, il est supposé que  $z_k^{(d)}(i, j) = \xi_k^{(d)}(i, j)/X_k(i, j)$  suit une normale d'espérance nulle et de variance  $\sigma_k^{2(d)} = \sigma_0^{2(d)} 2^{-2kH_d}$  où  $d = 1, 2, 3$ . Une fois les trois  $z_k^{(d)}(i, j)$  générés et connaissant la valeur de  $X_k(i, j)$ , les quatre variables permettant de calculer les quatre valeurs de précipitation sur la sous-tuile à l'échelle  $k+1$  sont connues. Cela implique que le volume total de précipitation est conservé sans avoir à appliquer un facteur multiplicatif. Les seuls cas où il faut appliquer une correction sont lorsque les valeurs de  $z_k^{(d)}(i, j)$  mènent à des valeurs de précipitation négatives. Dans ce cas, la ou les valeurs négatives sont ramenées à 0 et un facteur multiplicatif est appliqué pour ramener le volume

d'eau désagrégé au volume agrégé.

L'avantage des ondelettes de Haar est qu'elles peuvent tenir compte de l'anisotropie par les fluctuations directionnelles. Elles comptent cependant six paramètres comparativement à deux pour le modèle  $\beta$ -lognormale. Cependant, aucun des modèles de désagrégation multi-échelles existants ne tient compte de la corrélation spatiale inter-tuile. Harris et Foufoula-Georgiou (2001) en ont tenu compte indirectement en remplaçant les plus grandes valeurs de précipitation d'une sous-tuile plus près des grandes valeurs des sous-tuiles voisines. Groppelli *et al.* (2011) ont quant à eux ajouté une étape à leur modèle  $\beta$ -lognormale pour que les valeurs nulles de précipitation soient regroupées. Néanmoins, il reste qu'une carte de précipitation désagrégée à partir d'un modèle multi-échelle montre des discontinuités très visibles entre les tuiles (voir *e.g.* Ahrens, 2003; Bacchi et Ranzi, 2003; Sharma *et al.*, 2007). Lovejoy et Schertzer (2010a,b) attribuent ce manque de réalisme à la discrétisation de l'espace. Ils ont développé une méthode continue dans l'espace pour améliorer les propriétés statistiques des champs désagrégés à partir de la théorie multifractale. Cependant, le modèle qui en résulte est plus complexe, alors que le principal avantage des modèles multi-échelles résidait jusqu'ici dans leur relative simplicité. Selon notre connaissance de la littérature existante, aucun modèle de désagrégation continu dans l'espace n'a été appliqué à des données de précipitation.

## 2.3 Processus stochastiques

La précipitation sur tous les pixels d'une tuile peut être vue comme un vecteur aléatoire dont il faut estimer la distribution statistique. Compte tenu de la corrélation spatiale, l'estimation de la matrice de variance-covariance de ce vecteur n'est pas un exercice trivial (voir *e.g.* Huang *et al.*, 2002). Par exemple, une tuile contenant  $12 \times 12$  pixels produit un vecteur aléatoire de longueur  $n = 144$  et une matrice de variance-covariance contenant  $n(n + 1)/2 = 10440$  termes différents. Pour simplifier les calculs, seuls les termes les plus importants sont conservés, c'est-à-dire les termes associés aux voisins les plus proches, et une *marque aléatoire* (MRF pour *Markov random field* en anglais) est appliquée pour générer les valeurs à chaque pixel. La méthode utilisée par les modèles mentionnés dans cette section est l'*échantillonnage de Gibbs* et peut être résumée comme suit : (i) une valeur initiale est allouée sur chaque pixel ; (ii) des valeurs de précipitation sont générées sur un petit groupe de pixels (ou sur un seul pixel) en tenant compte des voisins les plus proches ; (iii) l'étape

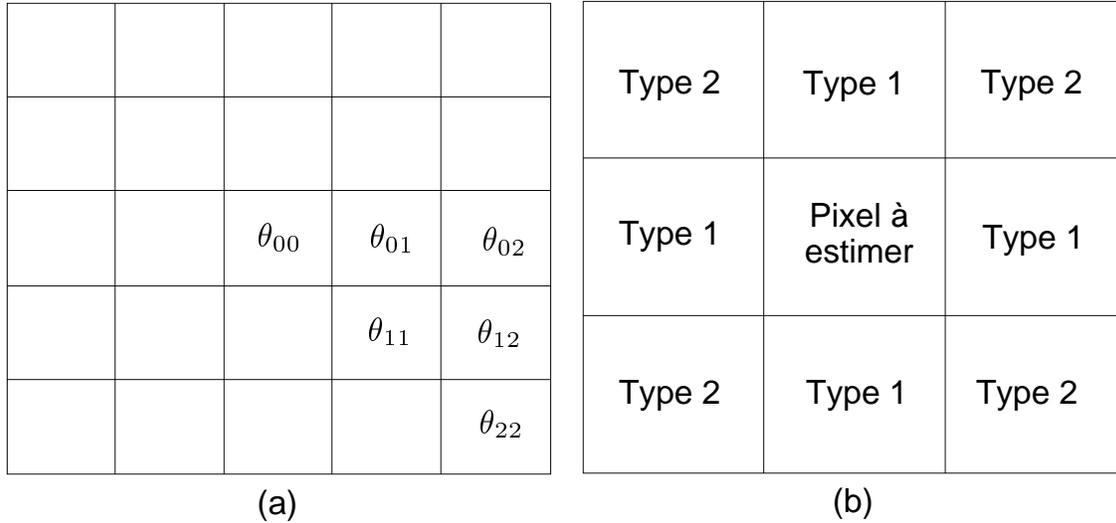


Figure 2.3 – (a) Les 24 pixels voisins et les six paramètres considérés dans le modèle de Allcroft et Glasbey (2003). (b) Les huit pixels voisins et les deux types de voisins considérés dans le modèle de Mackay *et al.* (2001). Dans les deux cas, la précipitation est supposée isotropique.

(ii) est répétée sur tous les groupes de pixels pour mettre à jour tous les pixels; (iv) et l'étape (iii) est répétée plusieurs fois afin d'éliminer l'impact de la condition initiale. Cette approche itérative permet de tenir compte explicitement de la corrélation spatiale, sans avoir à travailler directement avec la matrice de variance-covariance. L'échantillonnage de Gibbs est expliqué plus en détail à l'Annexe C.

En modélisation statistique, il est commun de supposer un vecteur aléatoire comme étant normal (ou gaussien), quitte à devoir procéder à une transformation pour que ce soit le cas. Allcroft et Glasbey (2003) ont transformé la précipitation à l'aide d'un polynôme d'ordre 2 pour la rendre gaussienne et ainsi appliquer une *marche aléatoire gaussienne* (GMRF pour *Gaussian Markov random field* en anglais) pour désagréger. Leur modèle tient compte des 24 voisins les plus proches (2 voisins dans chaque direction) au pas de temps présent et au pas de temps précédent. Il compte au total 12 termes de covariance, soit six pour chacun des pas de temps, et ne tient pas compte de l'anisotropie (Figure 2.3a). Il faut considérer que plus il y a de termes de covariance dans le modèle, plus leur estimation risque d'être instable et leur interprétation difficile (Allcroft et Glasbey, 2003).

Wheater *et al.* (1999), Chandler *et al.* (2000) et Mackay *et al.* (2001) ont développé un modèle de désagrégation ne tenant compte que des huit voisins les plus proches (1 pixel dans chaque

direction, Figure 2.3b) et de la précipitation au pas de temps précédent. La distribution spatiale de la précipitation se fait en trois étapes : (i) génération d'un nombre de pixels mouillés; (ii) distribution spatiale des pixels mouillés; et (iii) allocation des hauteurs de précipitation sur les pixels mouillés. À l'étape (i), le logarithme du ratio du nombre de pixels mouillés sur le nombre de pixels secs est une fonction linéaire du logarithme de la précipitation sur la tuile. Pour l'étape (ii), la régression logistique est utilisée pour déterminer la probabilité qu'un pixel soit mouillé. Deux variables explicatives sont retenues : la proportion des quatre pixels voisins de type 1, c'est-à-dire les pixels partageant un côté avec le pixel à estimer, et des quatre voisins de type 2, c'est-à-dire les pixels partageant un coin avec le pixel à estimer, qui sont mouillés (Figure 2.3b). Leur modèle tient donc compte de la distance des voisins, mais ne tient pas compte de l'anisotropie. À l'étape (iii), seulement présente dans Mackay *et al.* (2001), autant de valeurs de hauteur de précipitation qu'il y a de pixels mouillés sont générées selon la distribution Gamma et les plus grandes valeurs vont aux pixels ayant la plus grande distance avec le pixel sec le plus proche. Wheeler *et al.* (1999) et Chandler *et al.* (2000) ne s'étaient intéressés qu'à l'occurrence de la précipitation.

## 2.4 Discussion

Tel que mentionné au chapitre introductif, il y a peu de sources de données présentant la précipitation sur une grille à fine résolution spatiale comparativement aux données de stations météorologiques. Cela explique pourquoi les propriétés temporelles de la précipitation, évaluées à partir de stations météorologiques, sont souvent mieux connues que les propriétés des distributions spatiales de la précipitation à fine échelle spatiale. Cela explique aussi pourquoi la majorité des modèles de désagrégation spatiale, en occurrence les cellules de précipitation et les modèles multi-échelles, repose sur des propriétés de la précipitation d'abord utilisées en désagrégation temporelle. À défaut d'avoir des données de qualité à fine résolution spatiale, les propriétés spatiales de la précipitation sont parfois évaluées à partir d'un réseau de stations météorologiques (Guillot et Lebel, 1999; Segond *et al.*, 2006; Groppelli *et al.*, 2011).

Cette difficulté à obtenir des données de qualité à fine résolution spatiale fait en sorte que les modèles de désagrégation proposés, peu importe l'approche utilisée, ont été développés à partir de très peu d'événements, parfois même un seul. Au mieux, les paramètres sont estimés à partir d'un an de données (voir *e.g.* Wheeler *et al.*, 1999; Chandler *et al.*, 2000;

Mackay *et al.*, 2001; Forman *et al.*, 2008). Un modèle de désagrégation ne peut tenir compte du type d'événement s'il est développé sur une trop courte période. Néanmoins, Perica et Foufoula-Georgiou (1996) ont réussi à exprimer les paramètres de leur modèle multi-échelle utilisant les ondelettes de Haar en fonction de la variable atmosphérique « énergie potentielle de convection disponible », ou *Convective Available Potential Energy* (CAPE), mesurée juste avant le début de l'événement. La variable CAPE est une mesure de l'instabilité de l'atmosphère (Blanchard, 1998), une grande valeur de CAPE indiquant un risque d'orages violents. Dans Perica et Foufoula-Georgiou (1996), l'exposant d'échelle  $H$  est une fonction linéaire positive de CAPE alors que le paramètre multiplicatif  $\sigma_0$  est une fonction linéaire négative de CAPE. Forman *et al.* (2008) n'ont pas inclus la variable CAPE dans leur modèle puisque son impact n'était pas significatif. Ils ont préféré utiliser des valeurs différentes de paramètres pour chaque mois, comme le font plusieurs auteurs (*e.g.* Wheeler *et al.*, 1999; Chandler *et al.*, 2000; Mackay *et al.*, 2001; Cowpertwait *et al.*, 2002; Burton *et al.*, 2008; Groppelli *et al.*, 2011). Ils tiennent compte ainsi indirectement du type d'événement. Cependant, l'estimation de paramètres mensuels ne tient pas compte des propriétés physiques de l'événement et n'est donc pas robuste aux perturbations causées par les CC, problématique importante en projections climatiques, ou par des phénomènes atmosphériques tels *El Niño* et l'oscillation nord-Atlantique (ONA). Avec la grande quantité de données maintenant disponibles, il est dorénavant envisageable de développer un modèle de désagrégation dont les paramètres sont exprimés en fonction du type d'événement.

L'approche utilisée pour le développement du modèle de désagrégation doit permettre de répondre à l'objectif de recherche principal défini à la Section 1.2, c'est-à-dire qu'elle doit mener à des champs de précipitation réalistes. Les modèles multi-échelles sont de loin les plus répandus grâce à leur simplicité, mais ils ont de la difficulté à bien représenter la corrélation spatiale. Il est possible de rendre les modèles multifractals plus réalistes (Lovejoy et Schertzer, 2010a,b), mais le modèle devient conceptuellement plus complexe. Les modèles de cellules de précipitation tentent de reproduire les cellules convectives qui se produisent lors d'averses. Ces modèles peuvent être intéressants pour la désagrégation d'événements convectifs sur de courts pas de temps. Cependant, ils n'utilisent pas toute l'information que les champs à fine résolution spatiale peuvent fournir, comme la corrélation entre les pixels voisins. Dans le cadre de cette thèse, l'approche stochastique a été choisie. Avec cette approche, il est possible d'estimer la hauteur de précipitation sur un pixel à partir des hauteurs sur les pixels voisins. Cela permet de tenir compte des propriétés essentielles

pour que la structure spatiale soit réaliste, comme la corrélation spatiale et l'anisotropie. Le modèle doit utiliser toute l'information nécessaire pour bien représenter la précipitation à fine échelle tout en demeurant simple. La simplicité du modèle dépend de la rapidité de mise en œuvre et du temps de calcul nécessaire, mais aussi du nombre de paramètres à estimer et de la relation entre la hauteur de précipitation sur un pixel et les variables explicatives (*e.g.* hauteurs de précipitation sur les pixels voisins, variables atmosphériques). Le Chapitre 3 présente le modèle développé.

# Chapitre 3

## DÉVELOPPEMENT DU MODÈLE

Le modèle de désagrégation statistique est, dans un premier temps, développé sur une région avec peu de relief pour évaluer le potentiel de la méthode dans un contexte d'application simple. Ce chapitre décrit la région d'étude, la structure mathématique du modèle et présente l'algorithme basé sur l'échantillonnage de Gibbs utilisé pour désagréger la précipitation mésoéchelle.

### 3.1 Région à l'étude

Le modèle est développé sur une région d'environ  $400 \times 400 \text{ km}^2$  touchant l'Alabama, la Géorgie, la Caroline du Sud et la Floride, dans le sud-est des États-Unis (Figure 3.1). Les variations d'altitude sont peu importantes pour cette région, l'altitude maximale étant d'environ 350 m dans l'extrême nord-ouest de la région. Cette région a été choisie principalement parce qu'il y a une grande quantité de données disponibles (sept années) à fine résolution spatiale et parce qu'il s'y produit beaucoup d'événements dans un large éventail de types (de stratiformes à très convectifs).

Le climat de cette région est subtropical humide, caractérisé par des hivers frais, des étés chauds et humides et l'occurrence d'ouragans et de tornades. Pratiquement toute la précipitation tombe sous forme liquide. Pour les sept années étudiées (2002-2008), il est tombé en moyenne de 2,5 à 4,2 mm de pluie par jour selon le pixel. La précipitation annuelle moyenne sur la région étudiée varie de 850 mm (2006) à 1550 mm (2002). Durant cette période, le mois le plus pluvieux a été le mois d'août (4,8 mm/jour) et le moins pluvieux le mois de mai (2,2 mm/jour).

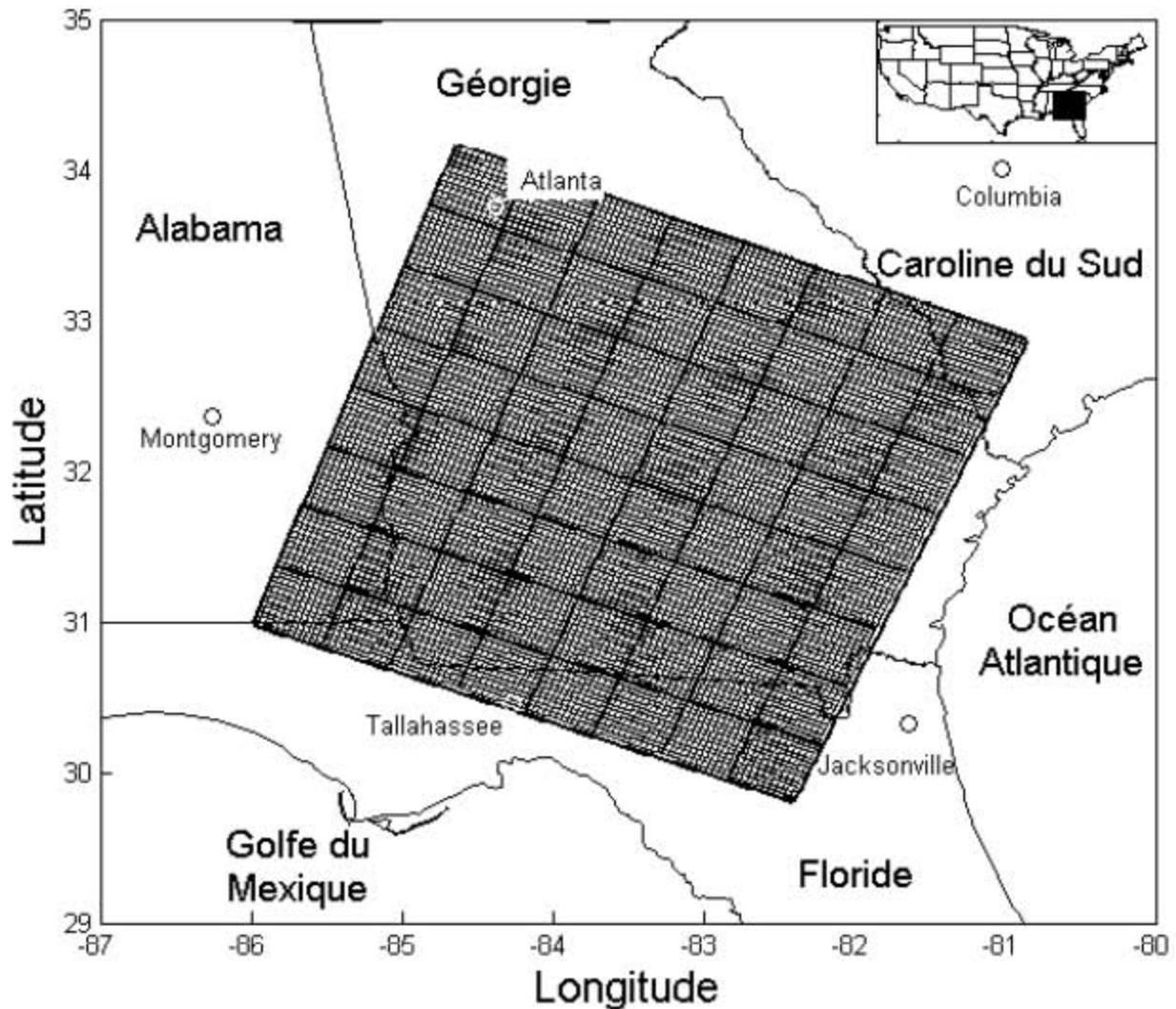


Figure 3.1 – Région étudiée avec les  $96 \times 96$  pixels de 3,8 km utilisés. Les lignes plus épaisses montrent les  $8 \times 8$  tuiles de 45,6 km de résolution.

### 3.1.1 Données de précipitation

La précipitation provient de la grille de *Stage IV*, produite par le *National Centers for Environmental Prediction* (NCEP), faisant partie du NWS. Cette grille à résolution d'environ 4 km et couvrant toute la partie continentale des États-Unis est créée à partir des grilles *Stage III*, produites à chacun des 12 RFC continentaux américains. La région étudiée est située dans le *Southeast RFC* (SERFC). Les données journalières, aux six heures et horaires sont disponibles en ligne à l'adresse <http://data.eol.ucar.edu>. Un contrôle de qualité manuel est effectué sur les données journalières et aux six heures. Pour plus de détails sur *Stage IV*, le lecteur est invité à consulter Lin et Mitchell (2005).

Un total de  $96 \times 96$  pixels de la grille *Stage IV* ont été choisis pour les analyses (Figure 3.1). Seules les données journalières de 2002 à 2008 sont utilisées. La résolution spatiale d'un pixel est de 3,8 km sur la région étudiée. Les accumulations journalières sont arrondies au millimètre près. Les données de la période 2002-2005 servent à l'évaluation des propriétés statistiques de la précipitation à fine échelle permettant le développement du modèle de désagrégation. Les données aux pixels sont agrégées pour que les paramètres du modèle soient estimés à différentes résolutions spatiales : 7,6 ( $2 \times 2$  pixels), 11,4 ( $3 \times 3$  pixels), 15,2 ( $4 \times 4$  pixels) et 22,8 km ( $6 \times 6$  pixels). Les données de la période 2006-2008 servent à l'application de la désagrégation pour sa validation (Chapitre 4). Les  $8 \times 8$  tuiles à désagréger à résolution 45,6 km sont formées en agrégeant chaque bloc de  $12 \times 12$  pixels (Figure 3.1). La dimension d'une tuile correspond *grosso modo* à la dimension d'une tuile d'un MRC. La précipitation est désagrégée à différentes résolutions spatiales (de 3,8 à 22,8 km). Les données aux pixels de la période 2006-2008 sont aussi utilisées comme référence pour la validation des champs de précipitation désagrégés.

### 3.1.2 Données atmosphériques

La structure spatiale d'un événement de précipitation journalier peut dépendre de variables atmosphériques. Pour approximer la valeur de ces champs, des *réanalyses d'observations*, qui sont en fait des données de différentes sources assimilées sur une grille régulière par un modèle atmosphérique, sont utilisées.

Dans la présente thèse, les données du *North American Regional Reanalysis* (NARR), produites par le NCEP, sont utilisées. Cette grille couvre l'Amérique du Nord à une résolution horizontale d'environ 32 km et sur 29 niveaux de pression. Les données sont disponibles aux trois heures de 1979 à aujourd'hui à l'adresse <http://www.esrl.noaa.gov/psd/data/gridded/data.narr.html>. Pour plus de détails sur les données NARR, le lecteur est invité à consulter Mesinger *et al.* (2004).

Les variables atmosphériques considérées sont l'énergie convective potentielle disponible (CAPE), une mesure de l'instabilité de l'atmosphère évaluée sur toute la colonne d'air (en J/kg), et la vitesse et la direction du vent évaluées aux pressions 700 et 850 hPa. Pour chaque variable, une valeur par jour par tuile est retenue en calculant la moyenne des valeurs des

huit pas de temps de trois heures couvrant la journée et de tous les points de la grille NARR situés à l'intérieur de la tuile.

## 3.2 Structure mathématique du modèle

Certaines hypothèses sont posées afin d'orienter la construction du modèle de désagrégation. Premièrement, la précipitation accumulée sur les tuiles utilisées en entrée dans le modèle est supposée précise et non-biaisée, c'est-à-dire que la précipitation en entrée est fidèle à ce qui est réellement tombé. En pratique, ce n'est pas nécessairement le cas lorsque les tuiles proviennent d'un modèle météorologique ou d'un modèle climatique. Il est reconnu que les sorties d'un MRC peuvent être biaisées et certains auteurs ont proposé des méthodes pour corriger ces biais (*e.g.* Déqué, 2007). Allcroft et Glasbey (2003) tiennent compte de l'incertitude sur les tuiles utilisées en entrée en tolérant une différence de 10 % avec la précipitation sur la tuile produite en sortie par le modèle de désagrégation. Cependant, l'évaluation du biais ou de l'imprécision des données d'un MRC ou d'un modèle météorologique va au-delà des objectifs de la thèse et ne sera pas pris en compte.

Deuxièmement, les accumulations de précipitation sur les pixels *Stage IV* utilisés sont supposées exactes. Puisque les données de radars météorologiques sont sujettes à plusieurs sources d'incertitudes (Joss et Germann, 2000), elles sont souvent considérées adéquates pour représenter la distribution spatiale de la précipitation, mais inadéquates pour représenter les intensités de précipitation. Ainsi, Wheeler *et al.* (1999) et Chandler *et al.* (2000) ont développé leur modèle de désagrégation stochastique en considérant uniquement l'occurrence de la précipitation sur les pixels. Les accumulations ont été simulées à une étape subséquente (Mackay *et al.*, 2001). Dans le cas présent, les données *Stage IV* proviennent d'observations de radars et de stations météorologiques auxquelles des contrôles de qualité automatique et manuel ont été appliqués. Il est donc raisonnable de faire confiance aux accumulations journalières fournies par *Stage IV*. C'est une des raisons pour laquelle le modèle de désagrégation développé dans le présent chapitre ne procède pas par étape comme dans Wheeler *et al.* (1999), Chandler *et al.* (2000) et Mackay *et al.* (2001). L'autre raison est que des précipitations journalières sont désagrégées ici alors que Wheeler *et al.* (1999), Chandler *et al.* (2000) et Mackay *et al.* (2001) ont désagrégé des données horaires. À l'échelle journalière, lorsqu'une tuile reçoit une accumulation non négligeable, il est fort probable que tous les pixels de la tuile reçoivent de la précipitation. Il n'est donc

pas intéressant de désagréger en considérant uniquement l'occurrence ou non de précipitation.

Comme troisième hypothèse, il est supposé que le modèle n'a pas à tenir compte du pas de temps précédent. En termes mathématiques, soit  $\vec{R}(t)$  et  $\vec{\mathbf{R}}(t)$  les vecteurs des accumulations de précipitation au pas de temps  $t > 1$  sur tous les pixels et sur toutes les tuiles, respectivement. Il est supposé que pour tout ensemble de valeurs  $W$ ,  $\mathbb{P}[\vec{R}(t) \in W \mid \vec{R}(t-1), \vec{\mathbf{R}}(t)] = \mathbb{P}[\vec{R}(t) \in W \mid \vec{\mathbf{R}}(t)]$ . Cette hypothèse n'implique pas nécessairement que les champs désagrégés de deux pas de temps consécutifs ne sont pas corrélés. Cela implique plutôt que la hauteur de précipitation désagrégée sur un pixel au pas de temps précédent n'apporte pas d'information pertinente supplémentaire lorsque les hauteurs de précipitation sur les tuiles au temps présent sont connues. Cette hypothèse est vérifiée plus loin. Il faut garder en tête que la désagrégation est faite sur des données journalières; si le pas de temps était plus court, cette hypothèse ne serait peut-être plus valable.

Le modèle est développé en analysant les propriétés statistiques des données de précipitations journalières à 3,8 km. Cette analyse permet notamment de trouver la forme la plus simple possible pour exprimer les deux premiers moments (espérance et écart-type) de la précipitation sur un pixel pour une journée donnée. Si possible, les paramètres sont exprimés en fonction des variables atmosphériques considérées. Avec des paramètres tenant compte des propriétés physiques de l'événement, il est permis de supposer que les valeurs estimées des paramètres sont valides peu importe le moment de l'année, dans un contexte de projections influencées par les CC ou non.

Pour estimer les paramètres d'un modèle, la méthode du maximum de vraisemblance (MV) est souvent utilisée. Cependant, considérant la grande quantité de données (plus de cinq millions de données disponibles à 3,8 km), il a été impossible d'obtenir des estimations des paramètres à partir des distributions statistiques considérées. Cela s'explique par le fait qu'un modèle, aussi bon soit-il, ne représente jamais exactement le processus qu'il estime. Lorsque peu de données sont disponibles, la disparité entre le processus réel représenté dans les données observées et le processus vu par le modèle peut se confondre avec la variabilité naturelle du processus. Lorsque la quantité de données disponibles est gigantesque, comme c'est le cas ici, il y a beaucoup d'observations pour lesquelles le modèle n'est pas adéquat, même si en proportion, ces observations ne représentent qu'un très faible pourcentage des

données. Ces observations ont des vraisemblances pratiquement nulles et cela empêche l'utilisation du MV pour l'estimation des paramètres.

La stratégie adoptée pour l'estimation de la distribution statistique de la précipitation sur un pixel procède en trois étapes. D'abord, une équation adéquate pour l'espérance est développée ; ensuite, une équation adéquate pour l'écart-type est développée ; et finalement, une distribution classique appropriée est ajustée. À chaque étape, le modèle qui fait le meilleur compromis entre la performance selon des critères objectifs et visuels et la parcimonie est retenu. Aucun test statistique n'est utilisé pour le développement du modèle. Avec la quantité gigantesque de données disponibles, l'impact aussi minime soit-il de chaque nouvelle variable serait statistiquement significatif. Ainsi, les modèles les plus complexes seraient toujours retenus par les tests statistiques. Le choix des équations et de la distribution statistique est fait sur les données disponibles à 3,8 km. Le modèle résultant est utilisé pour les autres tailles de pixels, mais avec des valeurs de paramètres différentes.

### 3.2.1 Espérance de la précipitation sur un pixel

Les modèles stochastiques mentionnés au Chapitre 2 expriment la précipitation sur un pixel en fonction des pixels voisins. Le présent modèle fait de même, en supposant que l'espérance de la précipitation sur un pixel est une fonction linéaire des valeurs aux pixels voisins. Afin de trouver la bonne formulation pour l'espérance, une régression linéaire est appliquée avec la hauteur de précipitation sur un pixel central comme variable expliquée et les hauteurs de précipitation sur les pixels voisins comme variables explicatives. Diverses équations tenant compte de différents nombres de voisins sont testées. Une équation avec peu de paramètres et qui donne une petite somme du carré des écarts entre la précipitation observée sur les pixels et l'espérance calculée est souhaitée. La somme du carré des écarts est effectuée sur les 5 136 976 données disponibles sur les pixels de 3,8 km provenant des 38 882 tuiles ayant reçues plus de 0,1 mm de précipitation journalière (2002-2005). Si la variabilité de la précipitation augmente avec l'espérance (Onibon *et al.*, 2004), cette façon de faire donne plus de poids aux données ayant les plus grandes valeurs de précipitation.

Soit  $R_{i,j}(t)$  la hauteur de précipitation sur le pixel  $(i, j)$  au jour  $t$ . Le Tableau 3.1 présente la somme du carré des écarts pour différentes équations utilisées. Pour alléger le texte, posons  $R_{i,j} \equiv R_{i,j}(t)$  dans tous les cas où le pas de temps précédent n'est pas considéré. La somme du carré des écarts de chaque équation est exprimée en pourcentage de la somme

du carré des écarts pour le modèle naïf qui suppose que  $\mathbb{E} [R_{i,j} | \vec{R}_{-(i,j)}] = \mathbf{R}_{i,j}$  où  $\mathbf{R}_{i,j}$  est la hauteur de précipitation sur la tuile contenant le pixel  $(i, j)$  et  $\vec{R}_{-(i,j)}$  est le vecteur contenant la précipitation sur tous les pixels, sauf le pixel  $(i, j)$ . Dans tous les cas, la somme des paramètres estimés  $\beta$  est imposée égale à 1.

Tel qu'attendu, le fait de tenir compte de la précipitation sur le pixel au jour précédent ne permet pas de diminuer la somme du carré des écarts par rapport au modèle qui suppose que l'espérance au pixel est égale à la hauteur de précipitation sur la tuile. Le simple fait de tenir compte de la précipitation moyenne sur les huit voisins les plus proches réduit la somme du carré des écarts de plus de 80 %. Différencier les quatre pixels les plus proches des quatre pixels plus éloignés diminue la somme du carré des écarts d'un 5 % supplémentaire avec un seul paramètre à estimer. L'ajout de deux autres paramètres pour tenir compte des quatre directions, soit 45°, 90°, 135°, et 0° par rapport à l'axe ouest-est de la grille, ne réduit pas significativement la somme du carré des écarts, mais permet de tenir compte d'une éventuelle anisotropie. Ces deux paramètres seront donc conservés. L'équation intégrant les 24 voisins donne une somme du carré des écarts légèrement plus faible que les autres modèles, mais ne sera pas retenu par souci de parcimonie. Comme les autres équations ayant plus de paramètres ne donnent pas des résultats significativement meilleurs, l'équation à trois paramètres est retenue. Pour simplifier la notation, posons  $\vec{A}^{(i,j)} = (A_{/}^{(i,j)}, A_{|}^{(i,j)}, A_{\setminus}^{(i,j)}, A_{-}^{(i,j)})'$  où  $A_{/}^{(i,j)} = \frac{R_{i-1,j-1} + R_{i+1,j+1}}{2}$ ,  $A_{|}^{(i,j)} = \frac{R_{i-1,j} + R_{i+1,j}}{2}$ ,  $A_{\setminus}^{(i,j)} = \frac{R_{i-1,j+1} + R_{i+1,j-1}}{2}$  et  $A_{-}^{(i,j)} = \frac{R_{i,j-1} + R_{i,j+1}}{2}$  (Figure 3.2). L'équation provisoirement retenue s'écrit donc comme suit :

$$\mathbb{E} [R_{i,j} | \vec{R}_{-(i,j)}] = \beta_1 A_{/}^{(i,j)} + \beta_2 A_{|}^{(i,j)} + \beta_3 A_{\setminus}^{(i,j)} + \beta_4 A_{-}^{(i,j)} \quad (3.1)$$

avec  $\sum_{k=1}^4 \beta_k = 1$ . Le problème avec cette équation est que les quatre variables explicatives sont très corrélées. Les coefficients de corrélation entre les paires d'éléments de  $\vec{A}^{(i,j)}$  calculés sur toutes les données de 2002 à 2005 varient de 0.957 à 0.977. Lorsque les variables explicatives sont corrélées, les paramètres estimés sont difficiles à interpréter (Allcroft et Glasbey, 2003).

Pour éliminer la corrélation entre les variables, une analyse en composantes principales (ACP) est faite. L'ACP crée quatre variables indépendantes, appelées *composantes principales* (CP) à partir de quatre combinaisons linéaires des éléments de  $\vec{A}^{(i,j)}$ . Le Tableau 3.2 présente les résultats de l'ACP, soit le coefficient de chacun des quatre éléments de  $\vec{A}^{(i,j)}$  définissant

Tableau 3.1 – Somme du carré des écarts entre la précipitation observée à 3,8 km et différentes équations d'espérance (2002-2005).

$\mathbb{E} [R_{i,j}   \vec{R}_{-(i,j)}] =$	Nombre de paramètres*	Nombre de voisins	Somme de carrés ( % par rapport à $\mathbb{E} [R_{i,j}   \vec{R}_{-(i,j)}] = \mathbf{R}_{i,j}$ )
$\mathbf{R}_{i,j}$	0	0	100
$\beta_1 \mathbf{R}_{i,j}(t) +$ $\beta_2 (R_{i,j}(t-1) - \mathbf{R}_{i,j}(t-1))$	1	0	100,0
$\frac{1}{8} \begin{pmatrix} R_{i-1,j} + R_{i,j-1} + \\ R_{i,j+1} + R_{i+1,j} + \\ R_{i-1,j-1} + R_{i-1,j+1} + \\ R_{i+1,j-1} + R_{i+1,j} \end{pmatrix}$	0	8	19,2
$\frac{\beta_1}{4} \begin{pmatrix} R_{i-1,j} + R_{i,j-1} + \\ R_{i,j+1} + R_{i+1,j} \end{pmatrix} +$ $\frac{\beta_2}{4} \begin{pmatrix} R_{i-1,j-1} + R_{i-1,j+1} + \\ R_{i+1,j-1} + R_{i+1,j} \end{pmatrix}$	1	8	14,2
$\beta_1 \left( \frac{R_{i-1,j-1} + R_{i+1,j+1}}{2} \right) +$ $\beta_2 \left( \frac{R_{i-1,j} + R_{i+1,j}}{2} \right) +$ $\beta_3 \left( \frac{R_{i-1,j+1} + R_{i+1,j-1}}{2} \right) +$ $\beta_4 \left( \frac{R_{i,j-1} + R_{i,j+1}}{2} \right)$	3	8	14,1
$\beta_1 R_{i-1,j-1} + \beta_2 R_{i-1,j} +$ $\beta_3 R_{i-1,j+1} + \beta_4 R_{i,j-1} +$ $\beta_5 R_{i,j+1} + \beta_6 R_{i+1,j-1} +$ $\beta_7 R_{i+1,j} + \beta_8 R_{i+1,j+1}$	7	8	14,1
$\sum_{\substack{k_i, k_j = \{-2, \dots, 2\} \\ (k_i, k_j) \neq (0, 0)}} \beta_{k_i, k_j} R_{i+k_i, j+k_j}$	23	24	13,5

\*Le nombre de paramètres à estimer est le nombre de  $\beta_i - 1$  puisque la somme des  $\beta_i$  doit être égale à 1.

$R_{i+1,j-1}$	$R_{i+1,j}$	$R_{i+1,j+1}$
$R_{i,j-1}$	$R_{i,j}$	$R_{i,j+1}$
$R_{i-1,j-1}$	$R_{i-1,j}$	$R_{i-1,j+1}$

Figure 3.2 – Pixels voisins retenus pour exprimer l’espérance de la précipitation sur le pixel  $(i, j)$ .

Tableau 3.2 – Composantes principales (CP) créées à partir des 5 136 976 valeurs de  $\vec{A}^{(i,j)}$  sur les pixels de 3,8 km (2002-2005).

CP = $p_1 A_{\nearrow}^{(i,j)} + p_2 A_{\mid}^{(i,j)} + p_3 A_{\searrow}^{(i,j)} + p_4 A_{\_}^{(i,j)}$					
CP	$p_1$	$p_2$	$p_3$	$p_4$	Variance expliquée (%)
1	0,499	0,504	0,494	0,504	97,9
2	0,666	0,041	-0,744	0,029	1,1
3	-0,042	0,738	-0,024	-0,673	0,6
4	-0,553	0,447	-0,450	0,541	0,4

chacune des CP et aussi la variance expliquée par chacune des composantes.

Pour la première CP, les poids des éléments de  $\vec{A}^{(i,j)}$  sont presque identiques. Cette CP peut être vue comme la moyenne arithmétique des quatre éléments de  $\vec{A}^{(i,j)}$ , qui est en fait la moyenne arithmétique des huit voisins les plus proches. Pour la quatrième CP, le poids des quatre pixels les plus proches est approximativement égale à  $-1 \times$  le poids des quatre voisins plus éloignés. Pour les deux autres CP, le poids dans une direction ( $45^\circ$  pour la deuxième CP,  $90^\circ$  pour la troisième) est approximativement égale à  $-1 \times$  le poids dans la direction perpendiculaire ( $135^\circ$  et  $0^\circ$  pour les deuxième et troisième CP, respectivement). L’équation (3.1) peut donc être réécrite en utilisant ces quatre variables quasi-indépendantes :

$$\mathbb{E} \left[ R_{i,j} \mid \vec{R}_{-(i,j)} \right] = \bar{A}^{(i,j)} + \beta_d \left( \frac{A_{\uparrow}^{(i,j)} + A_{\downarrow}^{(i,j)}}{2} - \frac{A_{\nearrow}^{(i,j)} + A_{\searrow}^{(i,j)}}{2} \right) + \beta_{\times} \left( A_{\nearrow}^{(i,j)} - A_{\searrow}^{(i,j)} \right) + \beta_{+} \left( A_{\uparrow}^{(i,j)} - A_{\downarrow}^{(i,j)} \right) \quad (3.2)$$

où  $\bar{A}^{(i,j)} = \frac{A_{\nearrow}^{(i,j)} + A_{\downarrow}^{(i,j)} + A_{\searrow}^{(i,j)} + A_{\uparrow}^{(i,j)}}{4}$  correspond à la moyenne des huit voisins les plus proches. Aucun paramètre n'est assigné à  $\bar{A}^{(i,j)}$ ; son coefficient est égal à 1 pour s'assurer que l'espérance ne soit pas biaisée. Le paramètre  $\beta_d$  tient compte de la distance des voisins. Il représente l'augmentation de l'espérance en proportion de la différence entre la précipitation moyenne sur les quatre voisins les plus proches et sur les quatre voisins plus éloignés. Ce paramètre doit être positif puisque la corrélation est supposée plus forte avec les pixels les plus proches. Les deux autres paramètres tiennent compte de l'anisotropie. Une valeur positive (négative) de  $\beta_{\times}$  indique une plus forte corrélation dans la direction  $45^\circ$  ( $135^\circ$ ) que dans la direction  $135^\circ$  ( $45^\circ$ ). Une valeur positive (négative) de  $\beta_{+}$  indique une plus forte corrélation dans la direction  $90^\circ$  ( $0^\circ$ ) que dans la direction  $0^\circ$  ( $90^\circ$ ). Il y a isotropie si  $(\beta_{\times}, \beta_{+}) = (0, 0)$ .

Un simple développement mathématique montre que l'équation (3.1) mène aux mêmes résultats que l'équation (3.2). Les  $\beta_k$   $k \in \{1, 2, 3, 4\}$  de l'équation (3.1) s'expriment comme suit :

$$\begin{aligned} \beta_1 &= \frac{1}{4} - \frac{1}{2}\beta_d + \beta_{\times}; \\ \beta_2 &= \frac{1}{4} + \frac{1}{2}\beta_d + \beta_{+}; \\ \beta_3 &= \frac{1}{4} - \frac{1}{2}\beta_d - \beta_{\times}; \\ \beta_4 &= \frac{1}{4} + \frac{1}{2}\beta_d - \beta_{+}. \end{aligned}$$

Peu importe les valeurs de  $\beta_d$ ,  $\beta_{\times}$  et  $\beta_{+}$ , la condition  $\sum_{k=1}^4 \beta_k = 1$  est respectée. L'équation (3.2) donne donc les mêmes résultats que l'équation (3.1), mais ses paramètres sont physiquement interprétables. Cela implique que les paramètres de l'équation (3.2) sont plus aptes à inclure des variables explicatives comme CAPE et la vitesse du vent. Un modèle exprimant ses paramètres en fonction de la physique de l'événement est plus facile à adapter

pour une utilisation dans différentes conditions.

Le paramètre  $\beta_d$  peut dépendre de la variable CAPE. En effet, pour un événement stratiforme (CAPE près de 0), il ne devrait pas y avoir de grandes différences entre la précipitation journalière des quatre voisins les plus proches et des quatre voisins plus éloignés. Par contre, si l'événement est très convectif (grandes valeurs de CAPE ; *e.g.*  $2 - 4 \times 10^3$  J/kg), les hauteurs de précipitation devraient varier de manière plus importante d'un pixel à l'autre. Dans ce cas, les pixels plus proches auraient plus de poids que les pixels plus éloignés et  $\beta_d$  devrait ainsi être plus grand. Pour vérifier l'impact de CAPE sur  $\beta_d$ , posons  $\beta_d = \beta_d^{(0)} + \beta_d^{(1)} C_{i,j}$  avec  $\beta_d^{(0)}$  et  $\beta_d^{(1)} > 0$  et où  $C_{i,j}$  est la valeur moyenne journalière de CAPE (J/kg) pour la tuile contenant le pixel  $(i, j)$ .

Les paramètres d'anisotropie  $\beta_\times$  et  $\beta_+$  peuvent dépendre du vent. Il peut être supposé en effet que plus le vent est fort dans une direction, plus la corrélation spatiale augmente dans cette direction. Cela implique que la valeur de  $\beta_\times$  ( $\beta_+$ ) est maximale si le vent souffle dans la direction  $45^\circ$  ( $90^\circ$ ) et minimale s'il souffle dans la direction  $135^\circ$  ( $0^\circ$ ). Ces conditions sont respectées si  $\beta_\times = \beta_\times^{(1)} V_{i,j} \cos \{2(\omega_{i,j} - 45^\circ)\}$  et  $\beta_+ = \beta_+^{(1)} V_{i,j} \cos \{2(\omega_{i,j} - 90^\circ)\}$  où  $V_{i,j}$  et  $\omega_{i,j}$  sont respectivement la vitesse en valeur absolue (m/s) et la direction (en  $^\circ$ ) du vent moyen journalier sur la tuile contenant le pixel  $(i, j)$ . L'axe de référence pour la direction du vent est l'axe ouest-est de la grille de précipitation. Il n'y a pas de  $\beta_\times^{(0)}$  et  $\beta_+^{(0)}$  parce que l'isotropie est supposée si  $V_{i,j} = 0$ . L'espérance de  $R_{i,j}$  incluant CAPE et le vent s'écrit :

$$\begin{aligned} \mathbb{E} \left[ R_{i,j} \mid \vec{R}_{-(i,j)} \right] &= \bar{A}^{(i,j)} + \left( \beta_d^{(0)} + \beta_d^{(1)} C_{i,j} \right) \left( \frac{A_{|}^{(i,j)} + A_{-}^{(i,j)}}{2} - \frac{A_{/}^{(i,j)} + A_{\backslash}^{(i,j)}}{2} \right) + \\ &\quad \beta_\times^{(1)} V_{i,j} \cos \{2(\omega_{i,j} - 45^\circ)\} \left( A_{/}^{(i,j)} - A_{\backslash}^{(i,j)} \right) + \\ &\quad \beta_+^{(1)} V_{i,j} \cos \{2(\omega_{i,j} - 90^\circ)\} \left( A_{|}^{(i,j)} - A_{-}^{(i,j)} \right). \end{aligned} \quad (3.3)$$

La vitesse et la direction du vent dépendent de l'altitude à laquelle le vent est mesuré. Deux niveaux de pression ont déjà été considérés comme variables explicatives de la précipitation : 700 hPa (Rodríguez-Puebla *et al.*, 2001 ; Bürger et Chen, 2005 ; Tolika *et al.*, 2007) et 850 hPa (Bürger et Chen, 2005 ; Min *et al.*, 2011). Ces niveaux de pression correspondent à des altitudes d'environ 3000 et 1500 m respectivement.

Tableau 3.3 – Somme du carré des écarts entre la précipitation observée à 3,8 km et différentes équations d'espérance tenant compte des variables atmosphériques (2002-2005).

Équation (3.2) avec :	Nombre de paramètres	Somme de carrés (% p/r à l'équation (3.2) )
$\beta_d = \left( \beta_d^{(0)} + \beta_d^{(1)} C_{i,j} \right)$	4	99,6
$\beta_d = \left( \beta_d^{(0)} + \beta_d^{(1)} C_{i,j} \right) ;$ $\beta_{\times} = \beta_{\times}^{(1)} V_{i,j}^{(700\text{hPa})} \cos \left\{ 2 \left( \omega_{i,j}^{(700\text{hPa})} - 45^\circ \right) \right\} ;$ $\beta_{+} = \beta_{+}^{(1)} V_{i,j}^{(700\text{hPa})} \cos \left\{ 2 \left( \omega_{i,j}^{(700\text{hPa})} - 90^\circ \right) \right\} .$	4	96,2
$\beta_d = \left( \beta_d^{(0)} + \beta_d^{(1)} C_{i,j} \right) ;$ $\beta_{\times} = \beta_{\times}^{(1)} V_{i,j}^{(850\text{hPa})} \cos \left\{ 2 \left( \omega_{i,j}^{(850\text{hPa})} - 45^\circ \right) \right\} ;$ $\beta_{+} = \beta_{+}^{(1)} V_{i,j}^{(850\text{hPa})} \cos \left\{ 2 \left( \omega_{i,j}^{(850\text{hPa})} - 90^\circ \right) \right\} .$	4	97,5
$\beta_{\times} = \beta_{\times}^{(1)} V_{i,j}^{(700\text{hPa})} \cos \left\{ 2 \left( \omega_{i,j}^{(700\text{hPa})} - 45^\circ \right) \right\} ;$ $\beta_{+} = \beta_{+}^{(1)} V_{i,j}^{(700\text{hPa})} \cos \left\{ 2 \left( \omega_{i,j}^{(700\text{hPa})} - 90^\circ \right) \right\} .$	3	96,9
$\beta_{\times} = \beta_{\times}^{(1)} V_{i,j}^{(850\text{hPa})} \cos \left\{ 2 \left( \omega_{i,j}^{(850\text{hPa})} - 45^\circ \right) \right\} ;$ $\beta_{+} = \beta_{+}^{(1)} V_{i,j}^{(850\text{hPa})} \cos \left\{ 2 \left( \omega_{i,j}^{(850\text{hPa})} - 90^\circ \right) \right\} .$	3	98,1
$\beta_{\times} = \beta_v V_{i,j}^{(700\text{hPa})} \cos \left\{ 2 \left( \omega_{i,j}^{(700\text{hPa})} - 45^\circ \right) \right\} ;$ $\beta_{+} = \beta_v V_{i,j}^{(700\text{hPa})} \cos \left\{ 2 \left( \omega_{i,j}^{(700\text{hPa})} - 90^\circ \right) \right\} .$	2	96,9
$\beta_{\times} = \beta_v V_{i,j}^{(850\text{hPa})} \cos \left\{ 2 \left( \omega_{i,j}^{(850\text{hPa})} - 45^\circ \right) \right\} ;$ $\beta_{+} = \beta_v V_{i,j}^{(850\text{hPa})} \cos \left\{ 2 \left( \omega_{i,j}^{(850\text{hPa})} - 90^\circ \right) \right\} .$	2	98,2

Tableau 3.4 – Valeurs estimées des paramètres de l'équation (3.4) pour les différentes tailles de pixels (2002-2005).

Taille des pixels	$\beta_d$	$\beta_v$ (s/m)
3,8 km	0,835	0,0146
7,6 km	0,885	0,0188
11,4 km	0,806	0,0198
15,2 km	0,769	0,0196
22,8 km	0,724	0,0204

Le Tableau 3.3 permet l'analyse des écarts entre la précipitation observée sur les données disponibles à 3,8 km et l'espérance pour chaque donnée calculée selon diverses équations basées sur l'équation (3.2), mais tenant compte des variables atmosphériques. L'ajout d'un paramètre pour CAPE ne réduit que très légèrement la somme du carré des écarts. Par souci de parcimonie, il est préférable de ne pas inclure CAPE dans l'équation de l'espérance. Cette variable ne sera considérée que pour l'écart-type. La vitesse et la direction du vent permettent de diminuer la somme du carré des écarts sans ajouter de paramètres. Le vent à 700 hPa donne de meilleurs résultats que le vent à 850 hPa. En théorie, si le vent est la seule cause de l'anisotropie, les valeurs de  $\beta_x^{(1)}$  et de  $\beta_+^{(1)}$  devraient être égales. Si ces deux valeurs sont imposées égales à  $\beta_v$ , des résultats quasi-identiques sont obtenus avec un paramètre de moins. L'équation retenue pour la désagrégation est donc

$$\begin{aligned}
 \mathbb{E} [R_{i,j} | \vec{R}_{-(i,j)}] &= \bar{A}^{(i,j)} + \beta_d \left( \frac{A_{|}^{(i,j)} + A_{-}^{(i,j)}}{2} - \frac{A_{/}^{(i,j)} + A_{\backslash}^{(i,j)}}{2} \right) + \\
 &\quad \beta_v V_{i,j}^{(700\text{hPa})} \cos \left\{ 2 \left( \omega_{i,j}^{(700\text{hPa})} - 45^\circ \right) \right\} \left( A_{/}^{(i,j)} - A_{\backslash}^{(i,j)} \right) + \\
 &\quad \beta_v V_{i,j}^{(700\text{hPa})} \cos \left\{ 2 \left( \omega_{i,j}^{(700\text{hPa})} - 90^\circ \right) \right\} \left( A_{|}^{(i,j)} - A_{-}^{(i,j)} \right) . \quad (3.4)
 \end{aligned}$$

L'équation (3.4) est aussi utilisée pour les autres tailles de pixels. Le Tableau 3.4 présente les valeurs estimées des paramètres. Les valeurs du paramètre  $\beta_d$  diminuent lorsque la taille du pixel augmente. C'est possiblement parce que la corrélation entre les pixels diminue lorsque leur taille augmente. Pour le paramètre  $\beta_v$ , il y a une grande augmentation de 3,8 à 7,6 km, suivi d'une légère tendance positive pour les autres tailles de pixels.

Avant de pouvoir utiliser l'équation (3.4) pour la désagrégation, il y a deux éléments à ajuster. Premièrement, si les coefficients d'anisotropie  $\beta_v V_{i,j}^{(700\text{hPa})} \cos \{2(\omega_{i,j}^{(700\text{hPa})} - 45^\circ)\}$  ou  $\beta_v V_{i,j}^{(700\text{hPa})} \cos \{2(\omega_{i,j}^{(700\text{hPa})} - 90^\circ)\}$  sont trop grands en valeur absolue, les champs désagrégés sont totalement irréalistes (Figure 3.3). Pour éviter que cela ne se produise, les coefficients d'anisotropie sont limités à l'intervalle -0,15 et 0,15. Deuxièmement, l'équation (3.4) peut produire des valeurs d'espérance négatives pour certains cas particuliers. Lorsque cela se produit, les valeurs sont généralement très près de 0. Pour éliminer les valeurs négatives, la valeur minimale de l'espérance est fixée à 0,1 mm.

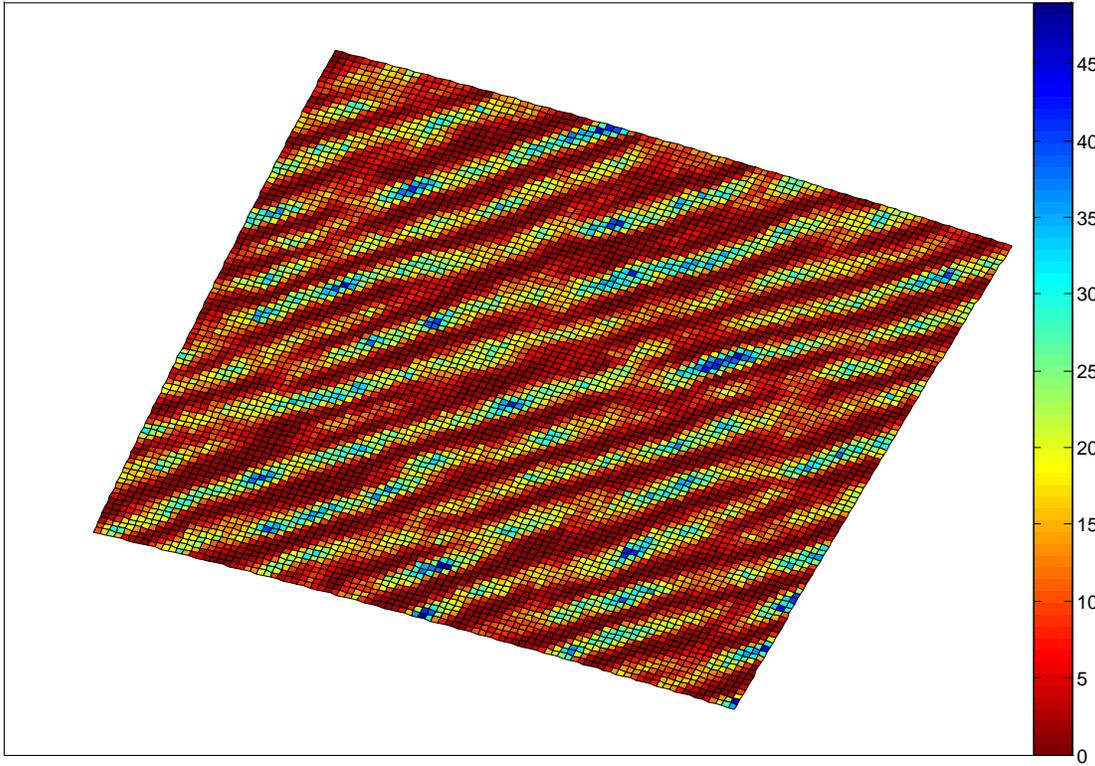


Figure 3.3 – Exemple synthétique d'une précipitation désagrégée lorsque le coefficient d'anisotropie est trop grand ( $\beta_v V_{i,j}^{(700\text{hPa})} \cos \{2(\omega_{i,j}^{(700\text{hPa})} - 45^\circ)\} = 0,25$ ).

La Figure 3.4 présente des diagrammes en boîtes des résidus  $R_{i,j} - \mathbb{E} [R_{i,j} | \vec{R}_{-(i,j)}]$ , soit la différence entre la précipitation sur le pixel de 3,8 km et son espérance selon l'équation (3.4), pour différentes classes d'espérance. Les résidus sont centrés à 0, ce qui montre que l'équation (3.4) s'ajuste bien aux données. Ce n'est pas le cas pour la première classe (espérance près de 0) parce que les données observées sont arrondies au millimètre. La largeur des diagrammes en boîtes augmente avec l'espérance, ce qui illustre que l'écart-type augmente avec l'espérance.

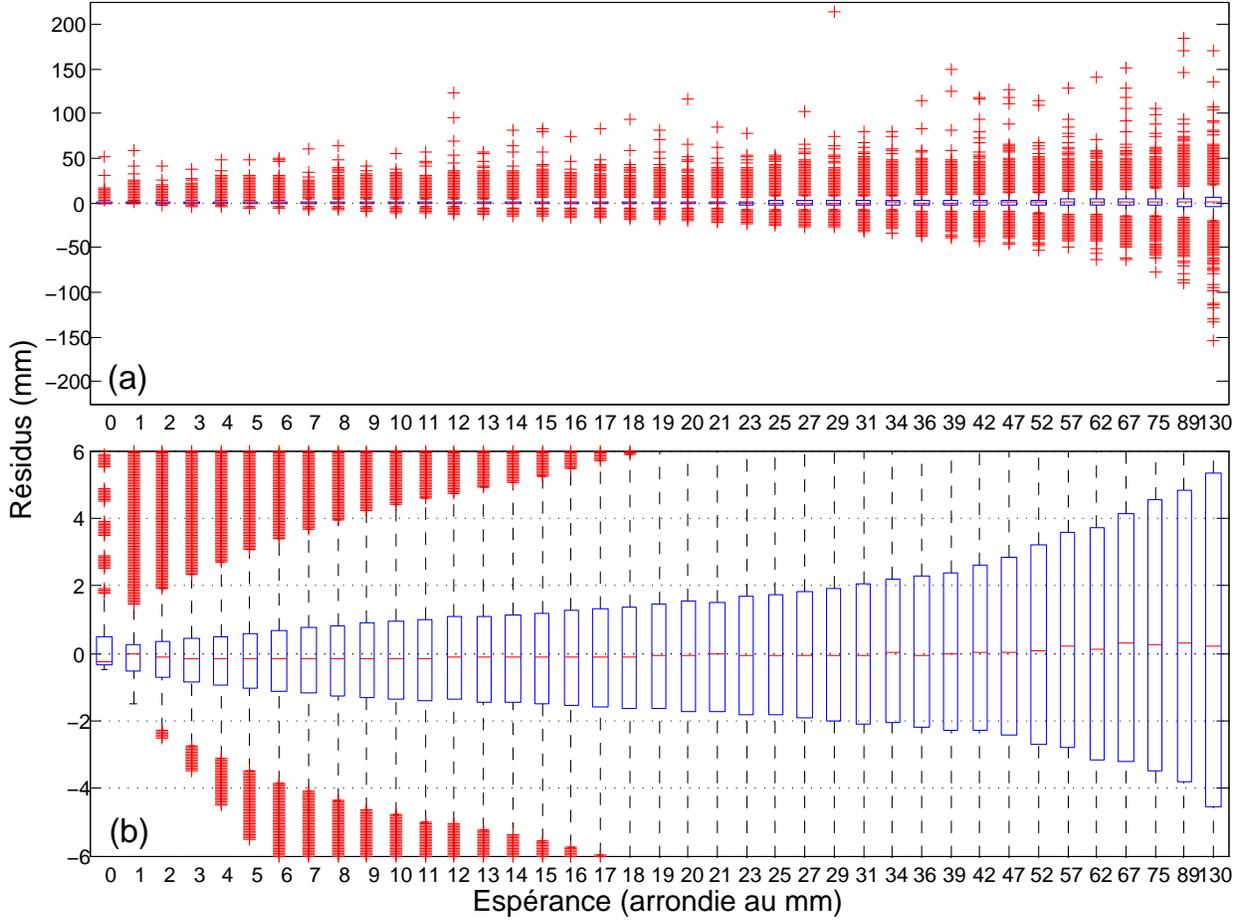


Figure 3.4 – (a) Diagrammes en boîtes des résidus des données disponibles à 3,8 km regroupées selon 39 classes d’espérance (2002-2005). (b) Zoom sur les boîtes.

### 3.2.2 Écart-type de la précipitation sur un pixel

Deux variables sont considérées pour l’estimation de l’écart-type sur un pixel. Pour une variable représentant une quantité aléatoire, l’écart-type augmente généralement avec l’espérance. La précipitation ne fait pas exception (Figure 3.4). L’approche la plus simple est de supposer que le coefficient de variation de la précipitation sur un pixel est constant, c’est-à-dire que l’écart-type est directement proportionnel à l’espérance :

$$\sqrt{\text{Var} [R_{i,j} | \vec{R}_{-(i,j)}]} = \theta \bar{R}_{i,j} \quad (3.5)$$

où  $\bar{R}_{i,j} = \mathbb{E} [R_{i,j} | \vec{R}_{-(i,j)}]$  tel que défini à l’équation (3.4) et  $\theta$  est le coefficient de variation.

Pour simplifier la notation, posons  $\sigma_{R_{i,j}} = \sqrt{\text{Var} [R_{i,j} | \vec{R}_{-(i,j)}]}$ . Onibon *et al.* (2004) ont quant à eux ajouté un degré de liberté et ont utilisé l’équation suivante :

$$\sigma_{R_{i,j}} = \theta_1 \bar{R}_{i,j}^{\theta_2} . \quad (3.6)$$

L'autre variable considérée est CAPE. Il est supposé que plus CAPE est grand (*i.e.* plus l'événement est convectif), plus la variabilité spatiale devrait être grande, donc plus l'écart-type à l'échelle du pixel devrait être grand. Cependant, la fonction qui relie CAPE à l'écart-type est inconnue. En désagrégation statistique, seuls Perica et Foufoula-Georgiou (1996) ont inclus CAPE dans leur modèle multi-échelle. Ils ont exprimé les deux paramètres liés à l'écart-type, soit l'exposant d'échelle et l'écart-type à l'échelle de la tuile, par une fonction linéaire de CAPE.

Afin d'identifier le modèle le plus approprié pour l'écart-type, des groupes de données ayant à peu près les mêmes valeurs d'espérance et de CAPE sont formés. Pour chaque groupe, l'écart-type des résidus est calculé. Il y a donc autant d'estimations d'écart-type que de groupes. L'idée est de former le plus de groupes possibles pour avoir une bonne évaluation de la relation entre CAPE et l'écart-type, tout en ayant assez de données dans chaque groupe pour que l'estimation de l'écart-type à l'intérieur d'un groupe soit robuste. Au total, 800 groupes de données à 3,8 km ont été formés à partir de 40 classes d'espérance  $\times$  20 classes de CAPE. Après avoir retiré les groupes ayant moins de 100 données de 3,8 km et ceux dont les données proviennent de moins de cinq jours différents, il reste 797 groupes. Il y a donc 797 valeurs d'écart-type associées à autant de valeurs d'espérance et de CAPE permettant de développer l'équation de l'écart-type sur un pixel.

Afin d'évaluer l'impact de CAPE sur l'écart-type, l'équation (3.6) est appliquée séparément pour chacune des 20 classes de CAPE. Les 20 estimations de  $\theta_1$  et  $\theta_2$  sont présentées à la Figure 3.5. Pour  $\theta_1$ , il se produit une augmentation rapide pour les faibles valeurs de CAPE. Pour les plus grandes valeurs de CAPE, l'augmentation est moins prononcée, presque constante. Pour  $\theta_2$ , les valeurs sont près de 1 (coefficient de variation constant) pour CAPE près de 0, mais diminuent rapidement lorsque CAPE augmente et se stabilisent ensuite pour de grandes valeurs de CAPE. La relation entre les paramètres et CAPE s'apparente à une relation logarithmique (lignes vertes de la Figure 3.5). En première approche, les paramètres  $\theta_1$  et  $\theta_2$  sont donc exprimés en fonction du logarithme de CAPE. La relation linéaire proposée par Perica et Fofoula-Georgiou (1996) est aussi testée ainsi que les relations constante et puissance.

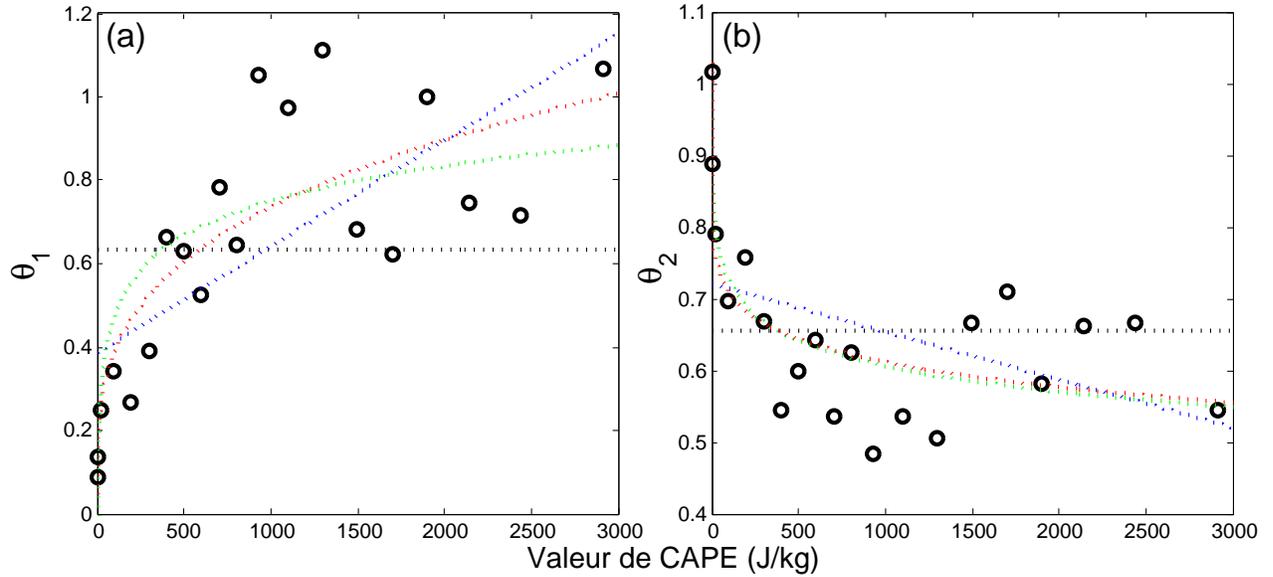


Figure 3.5 – Estimation des paramètres (a)  $\theta_1$  et (b)  $\theta_2$  de l'équation (3.6) pour chacune des 20 classes de CAPE (2002-2005). Les lignes pointillées représentent les diverses relations testées au Tableau 3.5 : constante (noir), linéaire (bleu), logarithmique (vert) et puissance (rouge).

Le Tableau 3.5 présente la somme du carré des écarts entre les écart-types observés et estimés pour chaque groupe selon les divers modèles considérés. L'équation utilisée par Onibon *et al.* (2004) donne de meilleurs résultats que l'équation du coefficient de variation constant. Aussi, il est à noter que tenir compte de CAPE diminue substantiellement la somme du carré des écarts. Les fonctions linéaires de CAPE performant mieux que les fonctions logarithmiques. Ce résultat peut sembler étonnant considérant les constations faites à la Figure 3.5. Il faut noter que tous les points de cette figure ne sont pas estimés à partir du même nombre de données. Pour l'estimation de l'écart-type, les points sur la Figure 3.5 n'ont donc pas tous le même poids.

Il n'est pas évident de déterminer à partir des résultats présentés au Tableau 3.5 quelle équation est la meilleure entre  $\sigma_{R_{i,j}} = (\theta_0 + \theta_1 C_{i,j}) \bar{R}_{i,j}^{\theta_2}$  et  $\sigma_{R_{i,j}} = (\theta_0 + \theta_1 C_{i,j}) \bar{R}_{i,j}^{\theta_2 + \theta_3 C_{i,j}}$ . La somme du carré des écarts dans ce dernier cas est 17 % plus petite, mais l'équation compte un paramètre de plus.

La Figure 3.6 présente les écart-types observés et estimés à partir des équations  $\sigma_{R_{i,j}} = (\theta_0 + \theta_1 C_{i,j}) \bar{R}_{i,j}^{\theta_2}$  et  $\sigma_{R_{i,j}} = (\theta_0 + \theta_1 C_{i,j}) \bar{R}_{i,j}^{\theta_2 + \theta_3 C_{i,j}}$  pour les 797 groupes. Les différences sont dans l'ensemble peu importantes entre les deux équations. Par souci de parcimonie, l'équation

Tableau 3.5 – Somme du carré des écarts entre les écart-types observés et estimés sur les 797 groupes selon les divers modèles considérés (2002-2005).

$\sigma_{R_{i,j}} =$	Nombre de paramètres	Somme de carrés (% p/r à l'équation (3.5) )
$\theta \bar{R}_{i,j}$	1	100
$\theta_1 \bar{R}_{i,j}^{\theta_2}$	2	60,1
$(\theta_0 + \theta_1 C_{i,j}) \bar{R}_{i,j}^{\theta_2}$	3	17,9
$\{\theta_0 + \theta_1 \ln(C_{i,j} + 1)\} \bar{R}_{i,j}^{\theta_2}$	3	28,0
$(\theta_0 + \theta_1 C_{i,j}^{\theta_2}) \bar{R}_{i,j}^{\theta_3}$	4	15,6
$\theta_1 \bar{R}_{i,j}^{\theta_2 + \theta_3 C_{i,j}}$	3	23,1
$\theta_1 \bar{R}_{i,j}^{\theta_2 + \theta_3 \ln(C_{i,j} + 1)}$	3	24,0
$\theta_1 \bar{R}_{i,j}^{\theta_2 + \theta_3 C_{i,j}^{\theta_4}}$	4	18,2
$(\theta_0 + \theta_1 C_{i,j}) \bar{R}_{i,j}^{\theta_2 + \theta_3 C_{i,j}}$	4	14,8
$\{\theta_0 + \theta_1 \ln(C_{i,j} + 1)\} \bar{R}_{i,j}^{\theta_2 + \theta_3 \ln(C_{i,j} + 1)}$	4	23,7

ayant le moins de paramètres à estimer est retenue. L'équation exprimant l'écart-type de la précipitation sur un pixel qui est retenue pour la désagrégation est donc :

$$\sigma_{R_{i,j}} = (\theta_0 + \theta_1 C_{i,j}) \bar{R}_{i,j}^{\theta_2} . \quad (3.7)$$

Les valeurs estimées des paramètres de l'équation (3.7) sur les différentes tailles de pixels sont présentées au Tableau 3.6. Les valeurs du paramètre  $\theta_1$  augmentent avec la taille du pixel. Par contre, il n'y a pas de tendance claire pour les paramètres  $\theta_0$  et  $\theta_2$ . Cela indique que plus la taille du pixel est grande plus le caractère convectif joue un rôle sur la variabilité de la précipitation.

### 3.2.3 Distribution statistique de la précipitation sur un pixel

Il aurait été intéressant, avec la quantité gigantesque de données disponibles à fine résolution, d'estimer la distribution de manière empirique. Ce n'est toutefois pas possible puisque ces données ne sont pas iid. En effet, chacune de ces données a sa propre espérance (équation (3.4)), son propre écart-type (équation (3.7)) et il y a de la corrélation entre deux valeurs

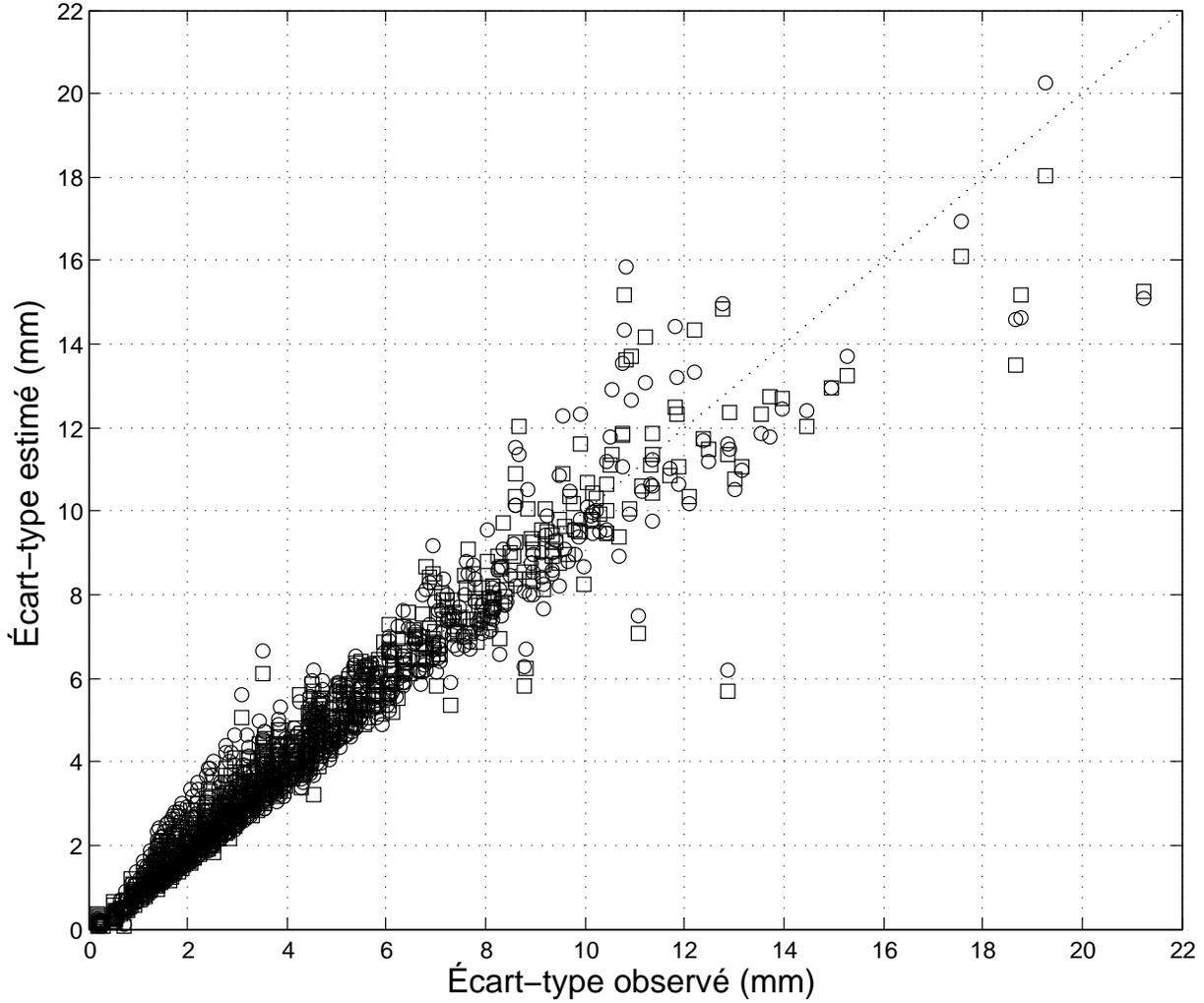


Figure 3.6 – Écart-types observés et estimés à partir des équations  $\sigma_{R_{i,j}} = (\theta_0 + \theta_1 C_{i,j}) \bar{R}_{i,j}^{\theta_2}$  (cercles) et  $\sigma_{R_{i,j}} = (\theta_0 + \theta_1 C_{i,j}) \bar{R}_{i,j}^{\theta_2 + \theta_3 C_{i,j}}$  (carrés) pour les 797 groupes (2002-2005).

provenant d'une même journée. Il a ainsi été décidé d'utiliser une distribution classique.

Avant de choisir la distribution, il faut décider si la précipitation journalière sur un pixel est considérée comme étant une variable strictement positive ou non-négative. En théorie, la précipitation journalière sur un pixel est une variable aléatoire non-négative dont la fonction de répartition  $F_{R_{i,j}}(x) = \mathbb{P}[R_{i,j} \leq x]$  est nulle pour  $x < 0$ , fait un saut de  $\mathbb{P}[R_{i,j} = 0]$  à  $x = 0$  et devient continue et croissante pour  $x > 0$ . Une telle variable aléatoire à la fois discrète et continue possède une distribution dite *mixte*. Pour les applications où il est important de bien représenter  $\mathbb{P}[R_{i,j} = 0]$ , une distribution mixte est utilisée pour simuler la précipitation (Over et Gupta, 1996 ; Guillot et Lebel, 1999 ; Veneziano *et al.*, 2006 ; Van Den Berg *et al.*, 2011).

Tableau 3.6 – Valeurs estimées des paramètres de l'équation (3.7) pour les différentes tailles de pixels (2002-2005).

Taille des pixels	$\theta_0$ (mm)	$\theta_1$ (mm kg/J)	$\theta_2$
3,8 km	0,442	$1,79 \times 10^{-4}$	0,628
7,6 km	0,533	$3,50 \times 10^{-4}$	0,564
11,4 km	0,514	$3,79 \times 10^{-4}$	0,598
15,2 km	0,522	$3,94 \times 10^{-4}$	0,621
22,8 km	0,534	$4,57 \times 10^{-4}$	0,601

Pour les applications hydrologiques à l'échelle journalière, les plus fortes valeurs de précipitation sont généralement plus intéressantes (*e.g.* pour l'estimation d'un débit de pointe). C'est principalement pour cette raison et aussi par souci de simplicité que la précipitation est représentée par une distribution continue. Trois des distributions classiques utilisées pour représenter la précipitation à l'échelle événementielle ont deux paramètres qui peuvent être exprimés en fonction de l'espérance et de l'écart-type (Tableau 3.7). La distribution normale est utilisée après transformation des données de précipitation (Allcroft et Glasbey, 2003) ou pour générer les fluctuations pour un modèle multi-échelle utilisant les ondelettes de Haar (Perica et Foufoula-Georgiou, 1996; Harris et Foufoula-Georgiou, 2001; Bacchi et Ranzi, 2003). Cependant, une distribution symétrique et pouvant générer des valeurs négatives comme la distribution normale n'est pas adéquate pour représenter la précipitation journalière non-transformée. Les deux distributions à deux paramètres les plus utilisées sont la distribution lognormale (Over et Gupta, 1996; Fiorucci *et al.*, 2001; Forman *et al.*, 2008; Groppelli *et al.*, 2011) et la distribution Gamma (Guillot et Lebel, 1999; Mackay *et al.*, 2001; Van Den Berg *et al.*, 2011).

Il y a deux autres distributions qui ont souvent été utilisées pour générer des valeurs de précipitation à l'échelle événementielle. La première, la distribution exponentielle est un cas particulier de la distribution Gamma et est intéressante puisqu'elle ne comporte qu'un seul paramètre (Entekhabi et Eagleson, 1989; Lebel *et al.*, 1998; Yu *et al.*, 2002). Cependant, cette distribution impose que l'espérance soit égale à l'écart-type. Les équations (3.4) et (3.7) montrent que cette condition n'est pas vérifiée dans le cas présent. L'autre distribution souvent utilisée est la distribution des valeurs extrêmes généralisée, mieux connue sous

Tableau 3.7 – Distributions à deux paramètres couramment utilisées pour modéliser la précipitation.

Distribution (paramètres)	Densité $\left(\frac{dF_{R_{i,j}}(x)}{dx}\right)$	Paramètres en fonction de l'espérance et de l'écart-type	
Normale $(\mu, \sigma)$	$\frac{\exp\left\{\frac{(x-\mu)^2}{2\sigma^2}\right\}}{\sigma\sqrt{2\pi}}$	$\mu = \bar{R}_{i,j}$	$\sigma = \sigma_{R_{i,j}}$
Lognormale $(\mu, \sigma)$	$\frac{\exp\left\{\frac{(\ln x - \mu)^2}{2\sigma^2}\right\}}{x\sigma\sqrt{2\pi}}$	$\mu = \frac{-\ln\left(\frac{\sigma_{R_{i,j}}^2}{\bar{R}_{i,j}^4} + \frac{1}{\bar{R}_{i,j}^2}\right)}{2}$	$\sigma = \sqrt{2\left(\ln \bar{R}_{i,j} - \mu\right)}$
Gamma $(\alpha, \beta)$	$\frac{x^{\alpha-1} \exp\left(-\frac{x}{\beta}\right)}{\beta^\alpha \Gamma(\alpha)}$	$\alpha = \frac{\bar{R}_{i,j}^2}{\sigma_{R_{i,j}}^2}$	$\beta = \frac{\sigma_{R_{i,j}}^2}{\bar{R}_{i,j}}$

le nom de GEV pour *Generalized Extreme Values* en anglais. La GEV est la distribution limite du maximum d'un échantillon de variables aléatoires iid. Elle est souvent utilisée en pratique pour modéliser des valeurs maximales annuelles de précipitation. En désagrégation, les valeurs de précipitation sont générées sur tous les pixels d'une tuile, pas seulement sur le pixel ayant la plus grande valeur, ce qui disqualifie la distribution GEV. C'est pourquoi seules les distributions lognormale et Gamma sont considérées dans la présente étude.

Afin de vérifier l'ajustement de la distribution, il n'est pas possible de comparer la fonction de densité théorique à la fonction de densité empirique puisque chaque donnée analysé a ses propres paramètres (*i.e.* sa propre espérance et son propre écart-type). Pour obtenir les paramètres d'une distribution pour un pixel, son espérance et son écart-type sont calculés selon les équations (3.4) et (3.7) et les équations du Tableau 3.7 sont appliquées. Il est ainsi possible de calculer la valeur de la fonction de répartition pour chaque donnée disponible. Si la précipitation sur un pixel est réellement distribuée selon la distribution choisie, les valeurs de la fonction de répartition devraient être distribuées uniformément. La Figure 3.7 présente l'histogramme des valeurs de la fonction de répartition calculée sur les données disponibles à 3,8 km (2002-2005) pour les distributions lognormale, Gamma et normale. Les 3 406 403 cas où espérance > 1 mm et les 1 389 173 cas où espérance > 10 mm sont présentés.

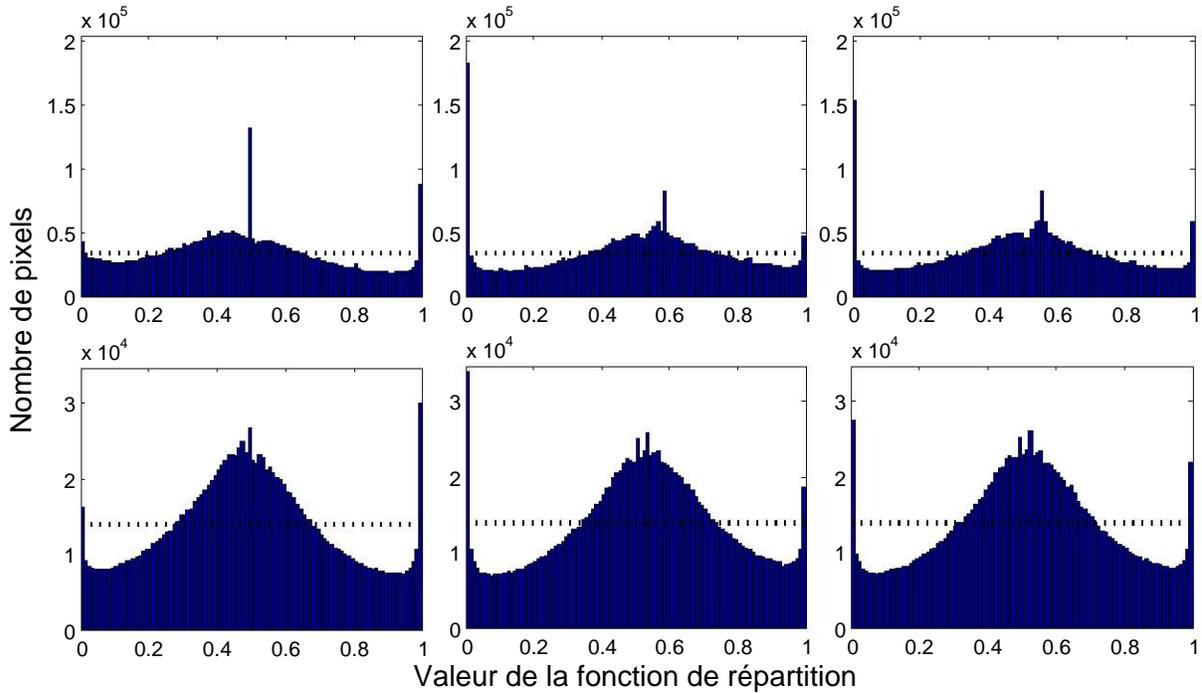


Figure 3.7 – Histogrammes des valeurs de la fonction de répartition des distributions normale (gauche), lognormale (centre) et Gamma (droite) calculées sur les données disponibles à 3,8 km (2002-2005). En haut : Sur les 3 406 403 cas où  $\bar{R}_{i,j} > 1$  mm. En bas : Sur les 1 389 173 cas où  $\bar{R}_{i,j} > 10$  mm. La ligne pointillée est ce qui serait obtenu si la distribution s’ajustait parfaitement aux données.

Les distributions lognormale et Gamma donnent des résultats comparables. Il y a trop de valeurs près de 0 et près de 1, ce qui signifie que les distributions utilisées sont sous-dispersées. Il y a aussi trop de valeurs près de 0,5. Elles sont néanmoins toutes deux meilleures que la distribution normale qui donne trop souvent des valeurs trop près de 1, ce qui signifie que la distribution normale sous-estime trop souvent la précipitation. La distribution lognormale est choisie puisqu’une distribution ne sous-estimant pas trop souvent la précipitation est désirée. La distribution Gamma aurait aussi pu faire l’affaire. Lorsque tous les cas où  $\bar{R}_{i,j} > 1$  mm sont considérés (Figure 3.7 en haut), il y a une bande au centre du spectre contenant plus de valeurs que ses voisines. C’est le cas pour les trois distributions utilisées, surtout pour la distribution normale. Ce phénomène arrive parce que les valeurs observées sont arrondies au millimètre près.

### 3.3 Algorithme de désagrégation

La distribution statistique de la précipitation journalière sur un pixel est maintenant définie conditionnellement à la précipitation sur les pixels voisins. Cependant, seule la précipitation sur les tuiles est connue au départ. Il existe une méthode appelée *échantillonnage de Gibbs* qui permet de simuler un vecteur aléatoire dont la distribution est inconnue, mais dont les distributions conditionnelles de chaque élément est connue. Dans le cas présent, le vecteur aléatoire est la précipitation sur tous les pixels pour une journée donnée et chaque élément représente la précipitation sur un pixel. À la section précédente, la distribution conditionnelle de la précipitation sur un pixel a été définie comme étant une lognormale d'espérance donnée par l'équation (3.4) qui dépend seulement des huit voisins les plus proches et dont l'écart-type est donné par l'équation (3.7). L'Annexe C présente la théorie de l'échantillonnage de Gibbs. Dans la présente section, l'algorithme basé sur l'échantillonnage de Gibbs utilisé pour la désagrégation est décrit.

La précipitation de  $N_{t_1} \times N_{t_2}$  tuiles pour une journée donnée est désagrégée en  $N_{p_1} \times N_{p_2}$  pixels par tuile pour former des champs de  $N_{P_1} \times N_{P_2}$  où  $N_{P_1} = (N_{t_1} \times N_{p_1})$  et  $N_{P_2} = (N_{t_2} \times N_{p_2})$  pixels. Dans la présente application, sauf dans la Section 4.1,  $N_t \equiv N_{t_1} = N_{t_2} = 8$  (Figure 3.1) et  $N_p \equiv N_{p_1} = N_{p_2}$  et varie de 2 à 12 pour les tailles de pixels de 22,8 à 3,8 km respectivement. Posons  $N_P \equiv N_{P_1} = N_{P_2}$ . Rappelons qu'au départ, seules les valeurs de précipitation journalières moyennées sur les tuiles, de CAPE, de direction et de vitesse du vent à 700 hPa sont connues. La désagrégation se fait selon l'algorithme suivant :

1. Comme valeurs initiales sur les pixels, posons pour tout  $(i, j) \in (\{1, \dots, N_P\}^2)$ ,  $R_{i,j} = \mathbf{R}_{i,j}$  (Figure 3.8a).
2. La valeur de  $\vec{A}^{(1,1)}$  est calculée. Comme seulement trois des huit pixels voisins sont disponibles, ce qui arrive aussi aux pixels  $(1, N_P)$ ,  $(N_P, 1)$  et  $(N_P, N_P)$ , il faut redéfinir  $\vec{A}^{(1,1)}$  et les composantes de  $\vec{A}^{(1,1)}$  (Figure 3.8b). Une fois que c'est fait, une nouvelle valeur de précipitation pour le pixel  $(1,1)$  est générée à partir de la distribution lognormale d'espérance et d'écart-type donnés par les équations (3.4) et (3.7), respectivement.
3. La valeur de  $\vec{A}^{(1,2)}$  est calculée. Comme seulement cinq des huit pixels voisins sont disponibles, ce qui arrive à tous les pixels touchant à la frontière sans être sur un des quatre coins de la région étudiée, il faut redéfinir  $\vec{A}^{(1,2)}$  et les composantes de  $\vec{A}^{(1,2)}$

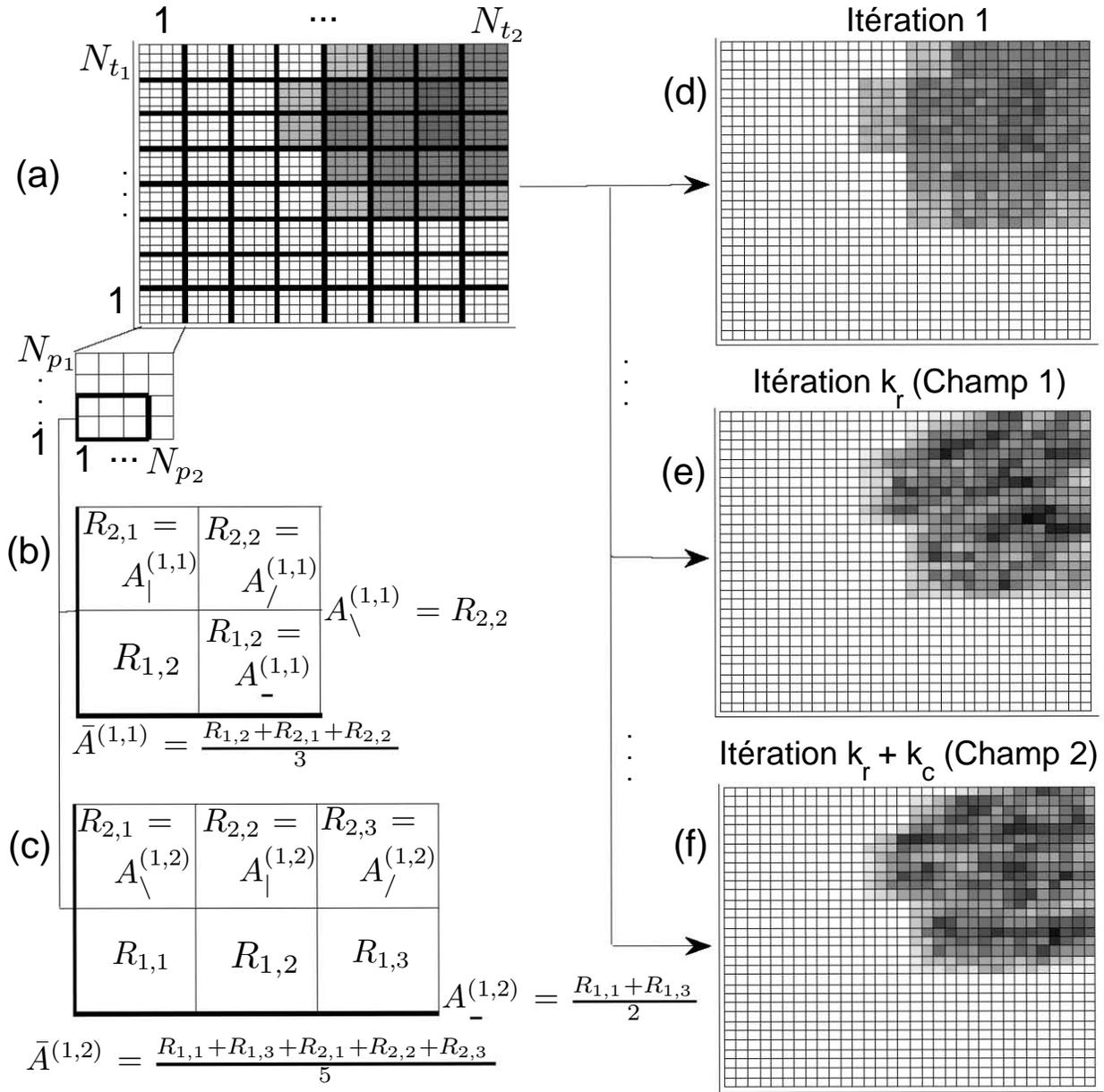


Figure 3.8 – Illustration de l’algorithme de désagrégation : (a) Valeurs initiales sur les pixels égales aux valeurs sur la tuile ; (b) Définition de  $\vec{A}$  pour le pixel  $(1,1)$  ; (c) Définition de  $\vec{A}$  pour le pixel  $(1,2)$  ; (d) Champ de précipitation après une itération de Gibbs ; (e) Premier champ de précipitation retenu après  $k_r$  itérations de Gibbs ; (f) Deuxième champ de précipitation retenu après  $k_r + k_c$  itérations de Gibbs.

(Figure 3.8c). Une fois que c’est fait, une nouvelle valeur de précipitation pour le pixel  $(1,2)$  est générée à partir de la distribution lognormale d’espérance et d’écart-type donnés par les équations (3.4) et (3.7), respectivement.

4. L’étape 3 est répétée pour tous les pixels de  $(1,3)$  à  $(N_P, N_P)$ . Un facteur multiplicatif

est appliqué pour s'assurer que le volume de précipitation par tuile soit conservé. La première itération de Gibbs est alors complétée (Figure 3.8d).

5. Les étapes 2, 3 et 4 sont répétées  $k_r$  fois pour éliminer l'impact de la distribution initiale (Figure 3.8e). Un premier champ de précipitation désagrégée est conservé.
6. Les étapes 2, 3 et 4 sont répétées  $k_c$  fois pour éliminer l'autocorrélation avec le dernier champ retenu (Figure 3.8f). Un autre champ de précipitation désagrégée est conservée.
7. L'étape 6 est répétée aussi souvent que voulu. À la fin,  $S$  champs de précipitation désagrégée ont été retenus pour la journée.

Il reste deux variables à estimer :  $k_r$  et  $k_c$ . Pour ce faire, la précipitation de huit champs synthétiques de  $8 \times 8$  tuiles créés à partir de huit scénarios est désagrégée : (précipitation sur toutes les tuiles égale à 10 ou 50 mm)  $\times$  événement stratiforme (CAPE = 0 J/kg sur toutes les tuiles) ou convectif (CAPE = 2500 J/kg sur toutes les tuiles)  $\times$  aucune anisotropie (aucun vent) ou anisotropie maximale dans la direction  $45^\circ$  ( $\beta_v V_{i,j}^{(700\text{hPa})} \cos \{2(\omega_{i,j}^{(700\text{hPa})} - 45^\circ)\} = 0,15$  pour toutes les tuiles). À chaque itération de chaque scénario, l'écart-type des valeurs désagrégées sur toutes les tuiles est calculé. Au départ, l'écart-type est nul (étape 1) puis augmente et au final, il se stabilise, c'est-à-dire qu'il oscille autour d'une constante, après l'itération  $k_r$ . À la Figure 3.9, la stabilité est atteinte après environ 100 itérations, sauf pour le cas stratiforme avec anisotropie maximale et précipitation sur les tuiles égale à 50 mm. Comme c'est un cas particulier et qu'il y a beaucoup de jours à désagrégier (trois ans), il n'est pas souhaitable de prendre une valeur inutilement grande de  $k_r$ . C'est pourquoi la valeur de  $k_r$  est fixée à 100, quitte à devoir revoir cette valeur pour certains cas particuliers. D'autres conditions initiales ont été testées, mais n'ont pas donné des résultats nettement différents.

Pour estimer  $k_c$ , l'autocorrélation de l'écart-type pour les itérations de  $k_r$  à 1000 est calculée pour chaque scénario (Figure 3.10). Même si elle ne s'est pas encore stabilisée autour de 0 pour tous les scénarios, l'autocorrélation est près de 0 après environ 50 itérations pour tous les scénarios, sauf pour le même scénario mentionné ci-haut. Il faut considérer que ce scénario est particulièrement intense. La valeur de  $k_c$  est donc fixée à 50 itérations pour tous les jours, même si dans certains cas, il se pourrait que l'autocorrélation soit non-négligeable après 50 itérations.

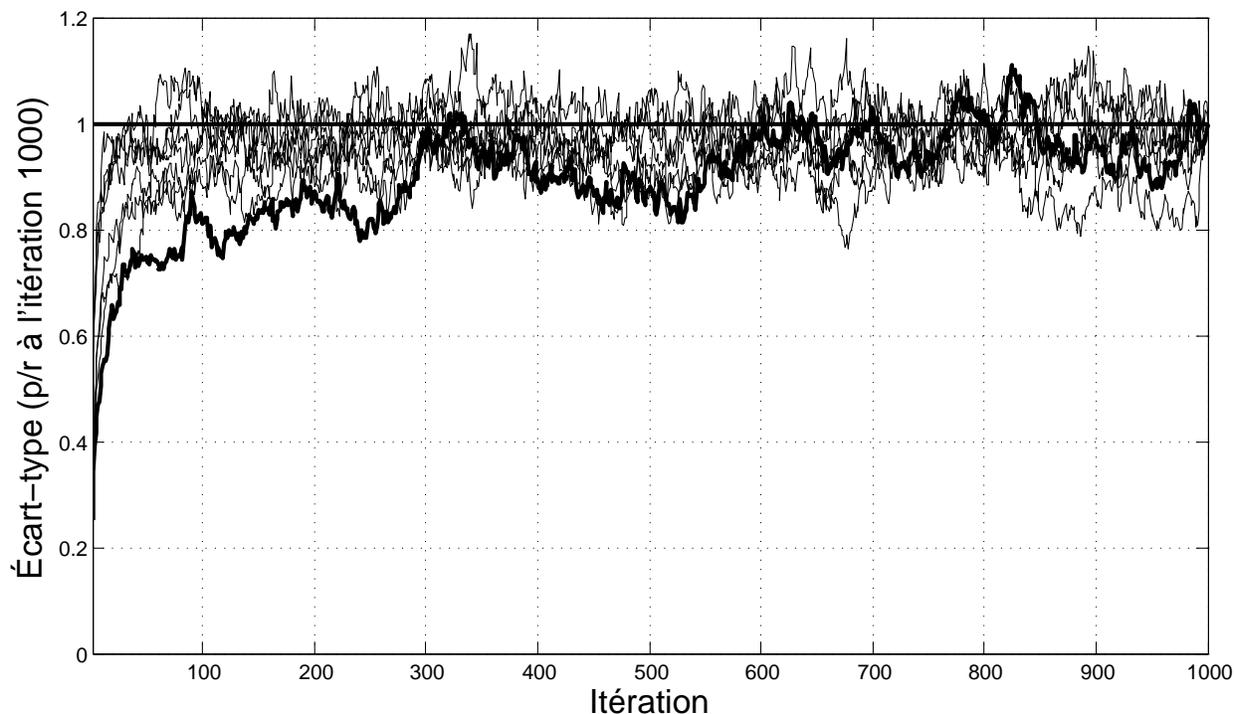


Figure 3.9 – Écart-type de la précipitation désagrégée des huit scénarios en fonction de l’itération. La ligne en gras représente le scénario stratiforme avec anisotropie maximale et précipitation sur les tuiles égale à 50 mm.

### 3.4 Discussion

Le modèle de désagrégation est développé à partir des propriétés statistiques de la précipitation journalière sur des pixels de 3,8 km de résolution sur une région du sud-est des États-Unis durant une période de quatre ans (2002-2005). L’énorme quantité d’observations disponibles rend difficile l’estimation des paramètres, mais en même temps, rend les estimations plus représentatives de l’ensemble des événements. L’espérance et l’écart-type de la distribution ont été estimés séparément. Les modèles offrant le meilleur compromis entre la performance (réalisme) et la parcimonie ont été retenus.

Pour l’espérance, les résultats du Tableau 3.1 ont montré qu’il ne valait pas la peine de tenir compte de plus de huit pixels voisins (1 pixel dans chaque direction), comme l’ont d’ailleurs proposé Wheeler *et al.* (1999), Chandler *et al.* (2000) et Mackay *et al.* (2001). Les données du jour précédent n’apportent pas d’information supplémentaire. L’ACP a permis de former, à partir des quatre directions, quatre variables presque indépendantes : la moyenne des huit voisins, la différence entre les quatre voisins les plus proches et les quatre voisins plus éloignés, la différence entre les directions 45° et 135° et la différence entre les

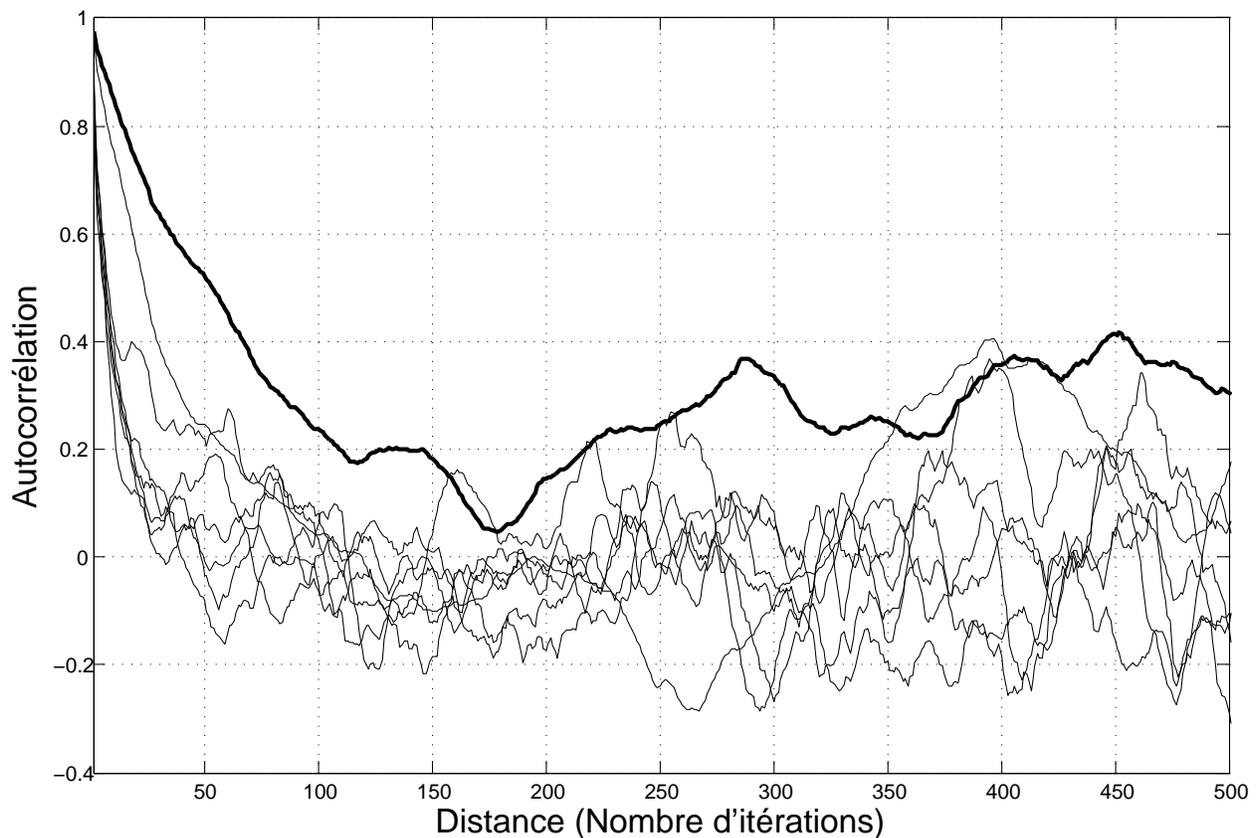


Figure 3.10 – Autocorrélation de l'écart-type des champs désagrégés pour les huit scénarios considérés. La ligne en gras représente le scénario stratiforme avec anisotropie maximale et précipitation sur les tuiles égale à 50 mm.

directions  $90^\circ$  et  $0^\circ$ . L'espérance est donnée par une fonction linéaire de ces quatre variables (équation (3.4)). Le coefficient de la moyenne des huit voisins est fixé à 1, afin de préserver asymptotiquement le volume de précipitation sur la tuile. Les coefficients des autres variables ont été estimés en minimisant la somme du carré des écarts entre la précipitation observée et l'espérance calculée pour chaque pixel. Le coefficient de la différence entre les quatre voisins les plus proches et les quatre voisins plus éloignés aurait pu être exprimé linéairement en fonction de CAPE, mais l'amélioration de la somme du carré des écarts n'a pas été jugée assez importante. Les coefficients des deux autres variables s'expriment en fonction de la vitesse et de la direction du vent à 700 hPa. Un seul paramètre nécessite d'être estimé pour ces deux variables. Ainsi, seulement deux paramètres ont à être estimés pour l'espérance. À cela peut s'ajouter le seuil minimal de 0,1 mm fixé à l'espérance pour éviter qu'elle ne soit négative. Ce seuil a un impact hydrologique limité pour les grands événements de précipitation, qui sont particulièrement intéressants en désagrégation. Des bornes de  $\pm 0,15$  ont aussi été fixées pour les coefficients d'anisotropie afin d'éviter que le modèle produise des

champs irréalistes (Figure 3.3). Le modèle ainsi développé peut tenir compte de l'anisotropie jusqu'à un certain degré ; la valeur maximale exacte d'anisotropie que le modèle est capable de reproduire n'est pas connue. En fait, elle peut varier d'une journée à l'autre. Comme il y a beaucoup de journées à désagréger, les bornes fixées sont conservatrices (*i.e.* pas trop grandes) afin de s'assurer qu'en aucun temps, des scénarios comme celui exposé à la Figure 3.3 ne se produisent.

Pour l'écart-type, deux variables ont un impact certain et ont été retenues : l'espérance de la précipitation et CAPE. Il est cependant difficile de trouver intuitivement la relation entre l'écart-type et ces deux variables. Certaines fonctions utilisées ont déjà été proposées par d'autres auteurs pour inclure l'espérance (Onibon *et al.*, 2004) et CAPE (Perica et Fofoula-Georgiou, 1996). Au final, l'équation (3.7) a été retenue. Le terme multiplicatif de l'écart-type en fonction de l'espérance est exprimé par une fonction linéaire de CAPE. Cette équation est simple et donne de bons résultats (Tableau 3.5), mais elle ne décrit pas parfaitement l'impact du type d'événement sur l'écart-type à l'échelle du pixel. La Figure 3.5 montre des différences entre les paramètres estimés sur des événements stratiformes (CAPE  $\approx 0$ ) et convectifs (CAPE  $> 1000$  J/kg). Par exemple, pour les événements stratiformes,  $\theta_2 \approx 1$ , c'est-à-dire que le coefficient de variation est presque constant. L'exposant  $\theta_2$  a ainsi été exprimé en fonction de CAPE, mais les résultats n'ont pas été jugés concluants (Tableau 3.5).

Trois distributions classiques ont été testées pour représenter la précipitation, soit les distributions normale, lognormale et Gamma. Seules les distributions à deux paramètres ont été considérées afin de pouvoir exprimer les paramètres en fonction des deux premiers moments (espérance et écart-type). Malgré une sous-dispersion, la distribution lognormale a été retenue. La distribution Gamma donnait des résultats comparables à la lognormale. Une distribution ayant un coefficient d'applatissage plus élevé pourrait être considérée afin de réduire cette sous-dispersion. Cependant, il faut comprendre que la précipitation à l'échelle du pixel ne suit pas exactement une lognormale, ni une Gamma, ni une GEV, ni aucune distribution classique. Les distributions classiques sont utilisées pour approcher la distribution empirique. Aucun processus naturel ne suit exactement une distribution statistique classique. Cela peut expliquer pourquoi il n'y a pas de consensus à savoir qu'elle est la meilleure distribution pour représenter la précipitation (Kundu et Siddani, 2007).

Pour être capable de reproduire les propriétés statistiques de la précipitation à fine résolution spatiale à partir de la précipitation sur les tuiles, l'échantillonnage de Gibbs est appliqué avec 100 itérations pour éliminer l'impact de la condition initiale (Figure 3.9) et 50 itérations pour éliminer l'autocorrélation avec le dernier champ retenu (Figure 3.10). Le prochain chapitre montre les résultats de la désagrégation réalisée à partir de l'algorithme décrit à la section précédente.



## Chapitre 4

# RÉSULTATS DE LA DÉSAGRÉGATION

Les 801 jours ayant reçu de la précipitation durant la période 2006-2008 sont désagrégés selon l'algorithme décrit à la Section 3.3. Pour chaque jour et chaque taille de pixels (3,8, 7,6, 11,4, 15,2 et 22,8 km),  $S = 8$  champs désagrégés (après 100, 150, . . . , 450 itérations) sont retenus. Les résultats sont uniquement présentés pour les pixels de 3,8 et 22,8 km. Les résultats pour les autres tailles de pixels sont présentés à l'Annexe D.

Il n'y a pas vraiment de convention sur la méthode de validation d'un modèle de désagrégation. Parmi les comparaisons possibles, il y a par exemple la comparaison visuelle entre la précipitation observée et désagrégée (Perica et Foufoula-Georgiou, 1996; Wheeler *et al.*, 1999; Chandler *et al.*, 2000; Mackay *et al.*, 2001; Allcroft et Glasbey, 2003), la comparaison entre leurs moments centrés (variance, coefficient d'asymétrie et kurtose; Deidda, 2000; Ferraris *et al.*, 2003) ou leurs distributions empiriques (Perica et Foufoula-Georgiou, 1996; Deidda, 2000; Kundu et Siddani, 2007). Ferraris *et al.* (2003) considèrent trois facteurs à prendre en compte pour évaluer un modèle de désagrégation :

- La facilité et la rapidité de son implantation numérique;
- Sa capacité à reproduire les propriétés statistiques des données à partir desquelles il a été développé;
- Sa capacité à reproduire les propriétés statistiques de données autres que celles utilisées pour son développement.

Le premier facteur est vérifié pour tous les modèles de désagrégation statistique. C'est leur avantage sur les modèles dynamiques. Le deuxième est nécessaire et plus facilement vérifié que le troisième. C'est ce dernier facteur qui demeure le plus important. C'est pourquoi le modèle est validé sur une période différente (2006-2008) de celle qui a servi à son développement (2002-2005).

Les statistiques utilisées pour évaluer le modèle doivent dépendre des propriétés qu'il est supposé reproduire. Dans le cas présent, l'objectif est de produire des champs de précipitation *réalistes*, soit des champs représentant adéquatement la corrélation spatiale, l'anisotropie et les valeurs extrêmes à l'échelle journalière. C'est pourquoi les résultats du modèle sont analysés considérant sa capacité à reproduire la structure spatiale d'un événement journalier, ce qui inclut la corrélation spatiale et l'anisotropie (Section 4.2), et les valeurs extrêmes (Section 4.3). Mais avant, une analyse de sensibilité par rapport à la taille de la région analysée est effectuée. Comme au chapitre précédent, les tests statistiques ne sont pas utilisés pour vérifier si le modèle représente bien la réalité ou non. Un modèle statistique simple ne peut représenter exactement la précipitation observée ; étant donnée la grande quantité de données utilisées, il est clair qu'un test statistique détecterait des différences statistiquement significatives entre les observations et les champs désagrégés (Chandler *et al.*, 2000). Pour les résultats qui suivent, l'attention est donnée aux différences en valeurs absolues plutôt qu'aux différences statistiquement significatives.

## 4.1 Sensibilité à la taille du domaine

Le modèle de désagrégation utilise l'information des pixels voisins et indirectement des tuiles voisines. Or, les résultats de la désagrégation sur une tuile dépendent du nombre de tuiles qui l'entourent. Cette section vise à déterminer s'il est nécessaire de conserver une *zone tampon* autour de la région d'intérêt pour la désagrégation. Une zone tampon sera jugée nécessaire si, pour une région d'intérêt donnée, les résultats de la désagrégation diffèrent selon qu'il y ait ou non une zone tampon. Pour tester la sensibilité des résultats à la taille du domaine désagrégé, le modèle est appliqué sur les  $8 \times 8$  tuiles, et aussi sur les  $6 \times 6$ ,  $4 \times 4$  et  $2 \times 2$  tuiles centrales. Notons  $R_{i,j}^{(s,d)}(t)$  la précipitation du  $s$ -ième champ désagrégé ( $s \in \{1, \dots, 8\}$ ) du pixel  $(i, j)$  du domaine formé de  $d \times d$  tuiles ( $d \in \{2, 4, 6, 8\}$ ) pour le jour  $t$  ( $t \in \{1, \dots, 801\}$ ). Pour une région d'intérêt de  $d_e \times d_e$  tuiles, les résultats de la désagrégation effectuée sur des domaines de  $d_1 \times d_1$  et  $d_2 \times d_2$  tuiles sont comparés à

Tableau 4.1 – Valeurs de  $\phi(d_1, d_2, d_e)$  (%) pour différents nombre de tuiles  $d_1$ ,  $d_2$  et  $d_e$  calculées sur les champs désagrégés à résolution 3,8 et 22,8 km.

$(d_1, d_2)$	Taille pixel = 3,8 km			Taille pixel = 22,8 km		
	$d_e = 2$	$d_e = 4$	$d_e = 6$	$d_e = 2$	$d_e = 4$	$d_e = 6$
(2,4)	87,9	-	-	49,4	-	-
(2,6)	85,7	-	-	34,7	-	-
(2,8)	78,2	-	-	36,9	-	-
(4,6)	14,6	36,4	-	1,6	35,5	-
(4,8)	14,1	37,2	-	-0,2	29,6	-
(6,8)	11,9	10,7	22,2	-7,9	-2,3	17,4

partir de la variable  $\phi(d_1, d_2, d_e)$ , soit le ratio de la somme du carré des écarts entre deux simulations effectuées sur des domaines différents sur la somme du carré des écarts entre deux simulations effectuées sur le même domaine :

$$\phi(d_1, d_2, d_e) = 100\% \times \left[ \frac{\left\{ \frac{\sum_{t=1}^{801} \sum_{i=1}^{N_P(d_e)} \sum_{j=1}^{N_P(d_e)} \left( R_{i,j}^{(1,d_1)}(t) - R_{i,j}^{(1,d_2)}(t) \right)^2 + \dots}{\sum_{t=1}^{801} \sum_{i=1}^{N_P(d_e)} \sum_{j=1}^{N_P(d_e)} \left( R_{i,j}^{(2,d_1)}(t) - R_{i,j}^{(2,d_2)}(t) \right)^2} \right\}}{\left\{ \frac{\sum_{t=1}^{801} \sum_{i=1}^{N_P(d_e)} \sum_{j=1}^{N_P(d_e)} \left( R_{i,j}^{(1,d_1)}(t) - R_{i,j}^{(2,d_1)}(t) \right)^2 + \dots}{\sum_{t=1}^{801} \sum_{i=1}^{N_P(d_e)} \sum_{j=1}^{N_P(d_e)} \left( R_{i,j}^{(1,d_2)}(t) - R_{i,j}^{(2,d_2)}(t) \right)^2} \right\}} - 1 \right] \quad (4.1)$$

où les indices des pixels  $(i, j)$  se réfèrent aux pixels de la région formée de  $d_e \times d_e$  tuiles sur lesquelles la somme du carré des écarts est effectuée et  $N_P(d_e) = d_e \times N_p$  (Figure 4.1). Les deux premiers champs désagrégés sont utilisés, mais n'importe quelle paire de champs désagrégés aurait pu être utilisée. Noter qu'il faut que  $d_e \leq \min(d_1, d_2)$ . Si les champs désagrégés sur  $d_1 \times d_1$  et sur  $d_2 \times d_2$  tuiles sont équivalents pour les  $d_e \times d_e$  tuiles analysées,  $\phi(d_1, d_2, d_e)$  devrait être près de 0 ; sinon,  $\phi(d_1, d_2, d_e)$  sera plus grand que 0. Le Tableau 4.1 illustre les valeurs de  $\phi(d_1, d_2, d_e)$  pour les pixels de tailles 3,8 et 22,8 km.

Il ressort que, peu importe les tailles  $d_e \times d_e$  de régions d'intérêt analysées, les valeurs de  $\phi(d_1, d_2, d_e)$  sont de l'ordre de 10 % pour les pixels de 3,8 km et autour de 0 % pour les pixels de 22,8 km lorsque les deux domaines donnent une zone tampon ( $d_e < \min(d_1, d_2)$ ).

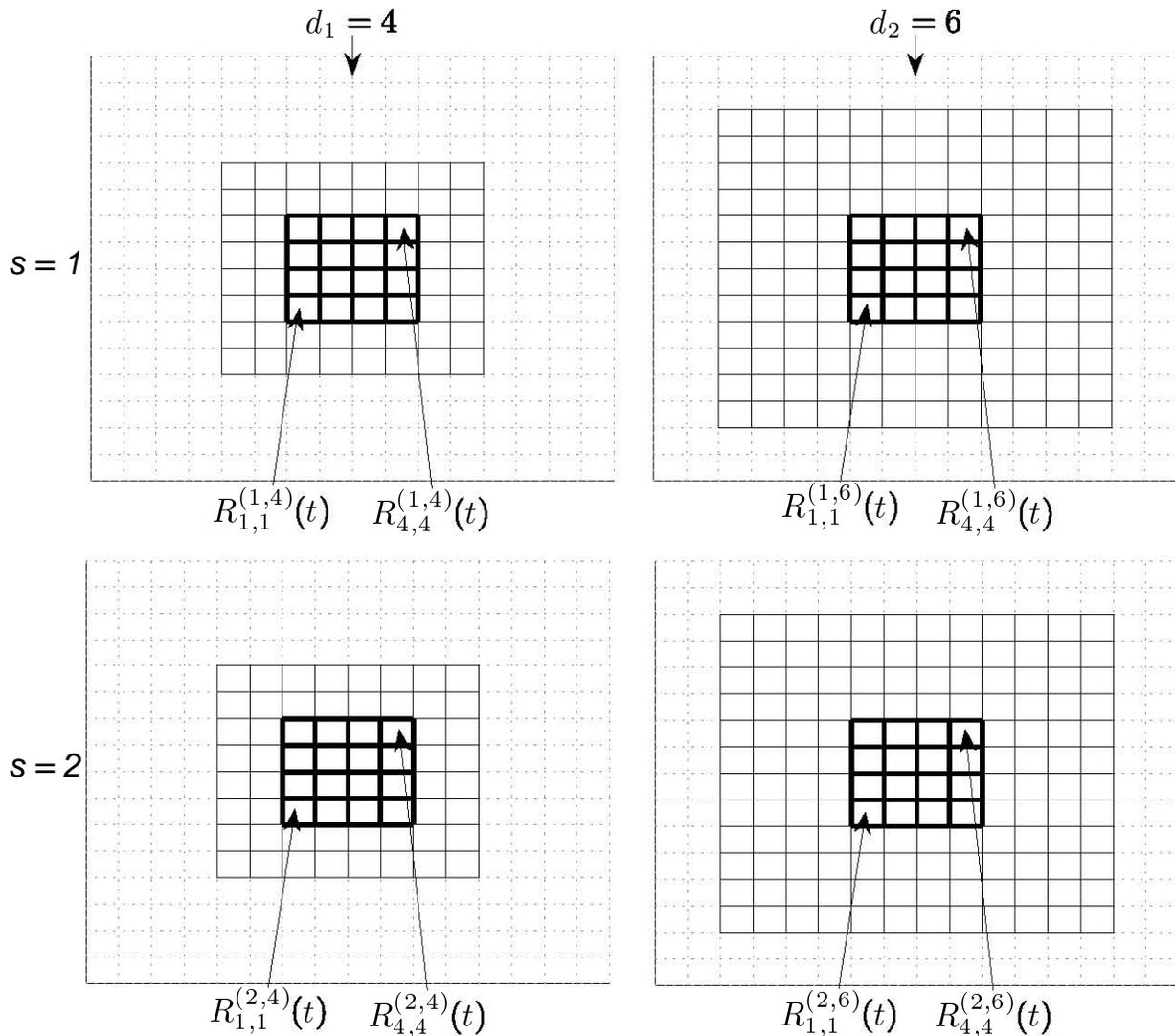


Figure 4.1 – Illustration des champs désagrégés nécessaires pour calculer  $\phi(d_1, d_2, d_e)$  avec  $d_e = 2$  (lignes plus épaisses),  $d_1 = 4$  et  $d_2 = 6$  (lignes minces) lorsque  $N_p = 2$ .

Lorsqu'un des deux domaines est égal à la région d'intérêt, les valeurs de  $\phi$  sont plus élevées. Cela illustre que la conservation d'une zone tampon d'une tuile de chaque côté a un impact sur les résultats. Le tableau montre aussi qu'il n'y a pas de différence significative entre conserver une zone tampon d'une ou deux tuiles de chaque côté. Pour le reste du chapitre, les résultats produits à partir de la désagrégation sur  $8 \times 8$  tuiles sont présentés, mais seulement pour les  $6 \times 6$  tuiles centrales.

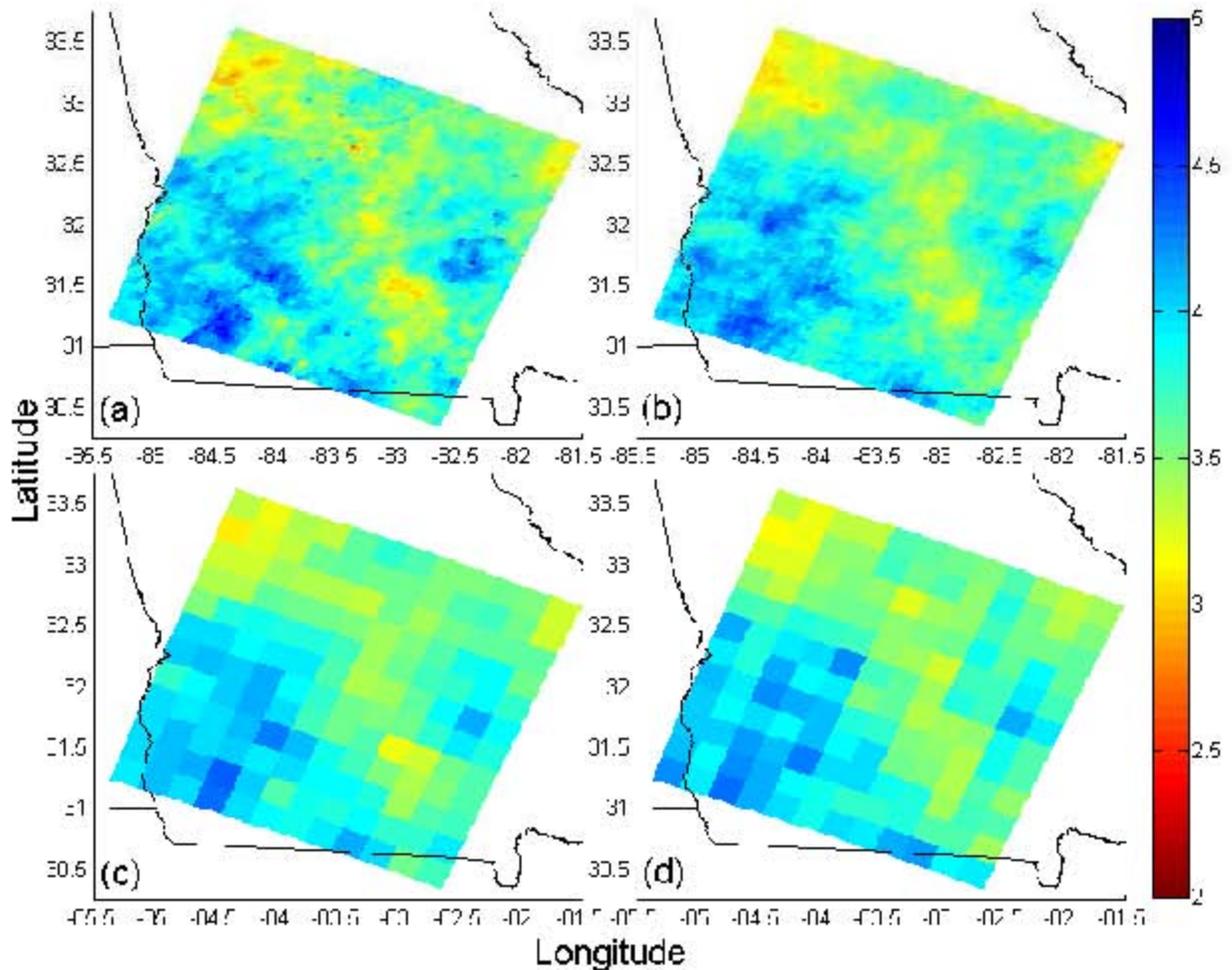


Figure 4.2 – Précipitation moyenne (mm/jour) sur les 801 jours (2006-2008) pour les observations (gauche) et pour un des huit champs désagrégés (droite) pour les pixels de taille (a-b) 3,8 et (c-d) 22,8 km.

## 4.2 Structure spatiale

Il est difficile de trouver une statistique assez complète pour exprimer la structure spatiale de la précipitation. Comme première illustration, la moyenne de la précipitation observée et désagrégée sur les 801 jours est présentée (2006-2008 ; Figure 4.2). La structure spatiale de la précipitation moyenne de 2006 à 2008 est quand même bien représentée, surtout pour la désagrégation à 3,8 km. Cependant, le modèle doit être évalué à l'échelle temporelle à laquelle il devrait être utilisé, soit l'échelle journalière.

Posons  $\rho[(i, j), (i', j')]$ , le coefficient de corrélation non-paramétrique de Kendall (Lehmann, 1998) pour chaque paire de pixels  $(i, j)$  et  $(i', j')$  calculé sur les jours où  $R_{i,j}$  et  $R_{i',j'}$  sont tous les deux plus grands que 0,1 mm. Afin d'évaluer la structure spatiale à l'échelle journalière,

une corrélation moyenne en fonction de la distance  $d$  dans chaque direction à partir des paires de pixels des  $6 \times 6$  tuiles centrales est calculée :

$$\rho_{45^\circ}(d) = \frac{1}{(N_P - 2N_p - d)^2} \sum_{i=N_p}^{(N_P - N_p - d)} \sum_{j=N_p}^{(N_P - N_p - d)} \rho[(i, j), (i + d, j + d)] ; \quad (4.2)$$

$$\rho_{90^\circ}(d) = \frac{1}{(N_P - 2N_p - d)(N_P - 2N_p)} \sum_{i=N_p}^{(N_P - N_p - d)} \sum_{j=N_p}^{(N_P - N_p)} \rho[(i, j), (i + d, j)] ; \quad (4.3)$$

$$\rho_{135^\circ}(d) = \frac{1}{(N_P - 2N_p - d)^2} \sum_{i=N_p + d}^{(N_P - N_p)} \sum_{j=N_p}^{(N_P - N_p - d)} \rho[(i, j), (i - d, j + d)] ; \quad (4.4)$$

$$\rho_{0^\circ}(d) = \frac{1}{(N_P - 2N_p)(N_P - 2N_p - d)} \sum_{i=N_p}^{(N_P - N_p)} \sum_{j=N_p}^{(N_P - N_p - d)} \rho[(i, j), (i, j + d)] . \quad (4.5)$$

Les Figures 4.3 et 4.4 présentent les valeurs moyennes de corrélation dans les quatre directions pour les pixels de tailles 3,8 et 22,8 km. Les jours stratiformes (CAPE moyen sur les tuiles plus petit que 1000 J/kg) sont séparés des jours convectifs (CAPE moyen sur les tuiles plus grand que 1000 J/kg).

Cette figure illustre que même si le modèle est construit en ne tenant compte que des voisins immédiats (un pixel dans chaque direction), il reproduit assez bien la corrélation spatiale sur de longues distances. Que ce soit sur les précipitations observées ou désagrégées, la corrélation spatiale est plus faible pour les événements convectifs (CAPE > 1000 J/kg) que pour les événements stratiformes (CAPE < 1000 J/kg). Bien que faibles, les différences entre les lignes des quatre directions indiquent la présence d'anisotropie, autant pour les données observées que désagrégées. L'anisotropie est plus importante pour les événements stratiformes que convectifs. Dans les deux figures, les huit lignes pointillées d'une même couleur ne peuvent être distinguées, ce qui veut dire que dans une direction donnée, les corrélations moyennes désagrégées sont pratiquement identiques pour les huit séries produites.

Pour les pixels de 3,8 km, le modèle sous-estime légèrement la corrélation spatiale des événements stratiformes (Figure 4.3a). Pour les événements convectifs (Figure 4.3b), la corrélation spatiale est surestimée dans les quatre directions utilisées pour développer le modèle. Pour les pixels de 22,8 km, le modèle reproduit remarquablement bien la corrélation spatiale dans les quatre directions pour les événements stratiformes (Figure 4.4a). Pour les

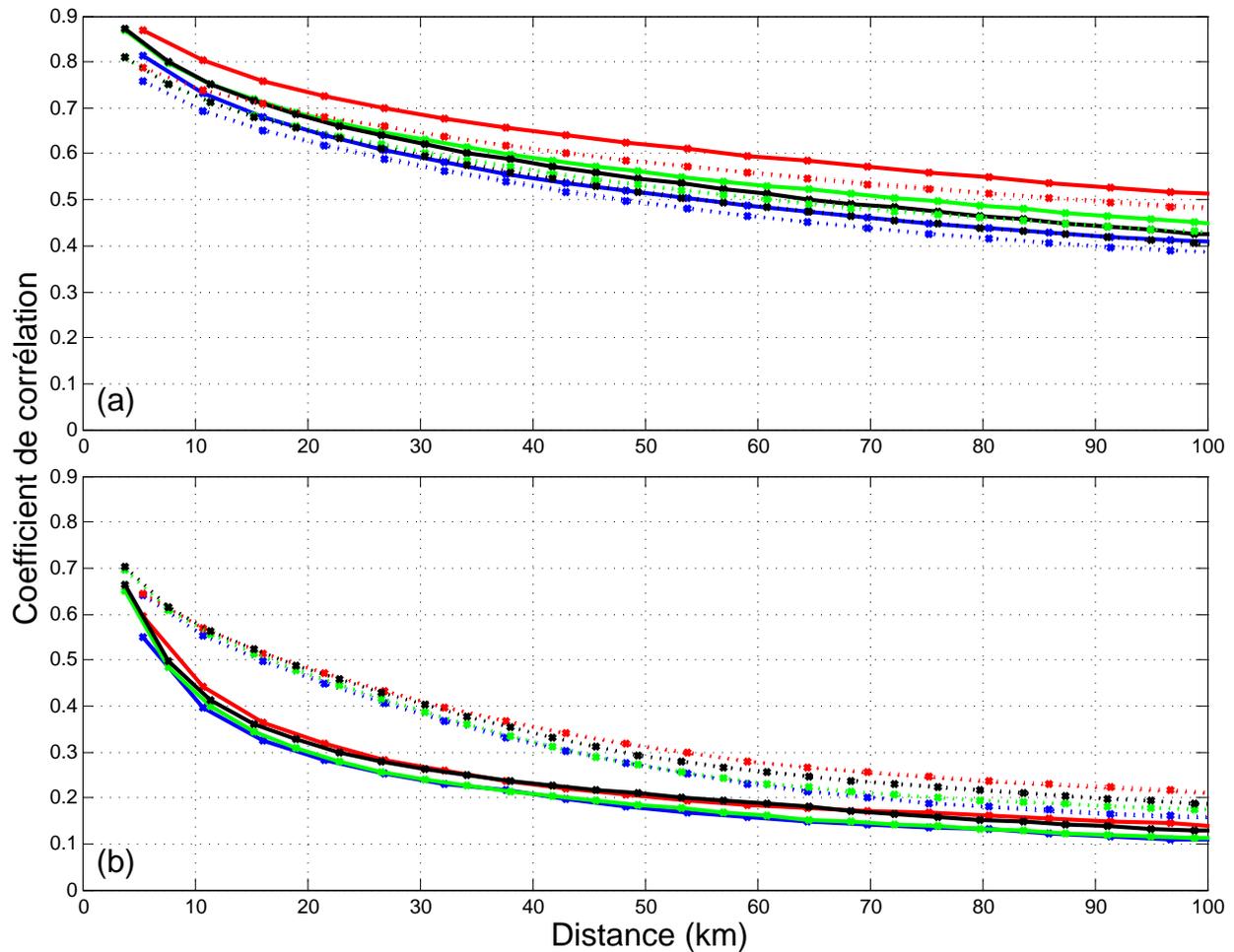


Figure 4.3 – Coefficient de corrélation moyen dans les quatre directions calculé sur les jours avec (a)  $CAPE < 1000 \text{ J/kg}$  et (b)  $CAPE > 1000 \text{ J/kg}$  pour les pixels de 3,8 km observés (lignes continues) et des huit séries désagrégées (lignes pointillées) :  $45^\circ$  (rouge),  $90^\circ$  (vert),  $135^\circ$  (bleu) et  $0^\circ$  (noir). Les corrélations des huit séries désagrégées sont quasi-identiques et impossible à distinguer.

événements convectifs, le modèle ne surestime que très légèrement la corrélation spatiale pour l'ensemble des distances (Figure 4.4b).

Les résultats pour les pixels de 3,8 km indiquent que même si l'écart-type augmente avec CAPE, le modèle crée des précipitations convectives trop corrélées spatialement. Une autre approche a dès lors été testée pour produire des champs désagrégés à résolution 3,8 km : la désagrégation en cascade. Cette approche, utilisée pour les modèles multi-échelles, consiste à désagréger d'une résolution  $x$  (ici,  $x = 45,6 \text{ km}$ ) vers une résolution  $x/a$  (généralement,  $a = 2$ ), puis de la résolution  $x/a$  à la résolution  $x/a^2$  et ainsi de suite jusqu'à la

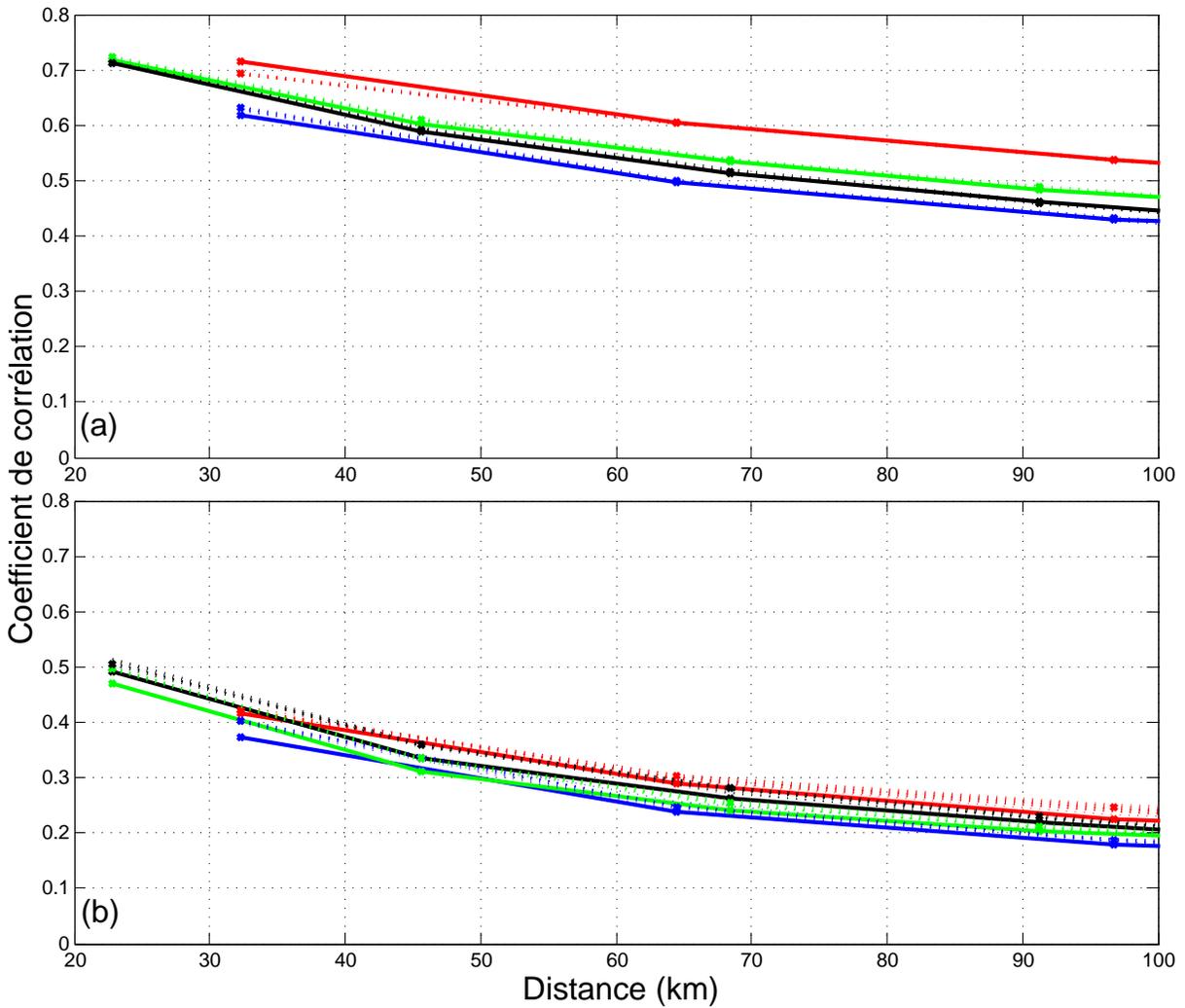


Figure 4.4 – Coefficient de corrélation moyen dans les quatre directions calculé sur les jours avec (a) CAPE < 1000 J/kg et (b) CAPE > 1000 J/kg pour les pixels de 22,8 km observés (lignes continues) et des huit séries désagrégées (lignes pointillées) : 45° (rouge), 90° (vert), 135° (bleu) et 0° (noir).

résolution désirée (ici, 3,8 km). Cette approche appliquée aux modèles multi-échelles donne généralement des précipitations avec des discontinuités visibles entre les tuiles (Ahrens, 2003; Bacchi et Ranzi, 2003; Sharma *et al.*, 2007), mais appliquée au présent modèle, elle pourrait permettre de diminuer la corrélation spatiale pour les événements convectifs. Pour chaque jour, le modèle désagrège de 45,6 à 22,8, de 22,8 à 11,4 et de 11,4 à 3,8 km à partir des paramètres estimés au Chapitre 3 pour chaque résolution utilisée.

La Figure 4.5 présente les valeurs moyennes de corrélation dans les quatre directions pour les pixels de 3,8 km obtenues par la désagrégation en cascade. Pour la précipitation stratiforme, la corrélation est légèrement plus sous-estimée avec la désagrégation en cascade

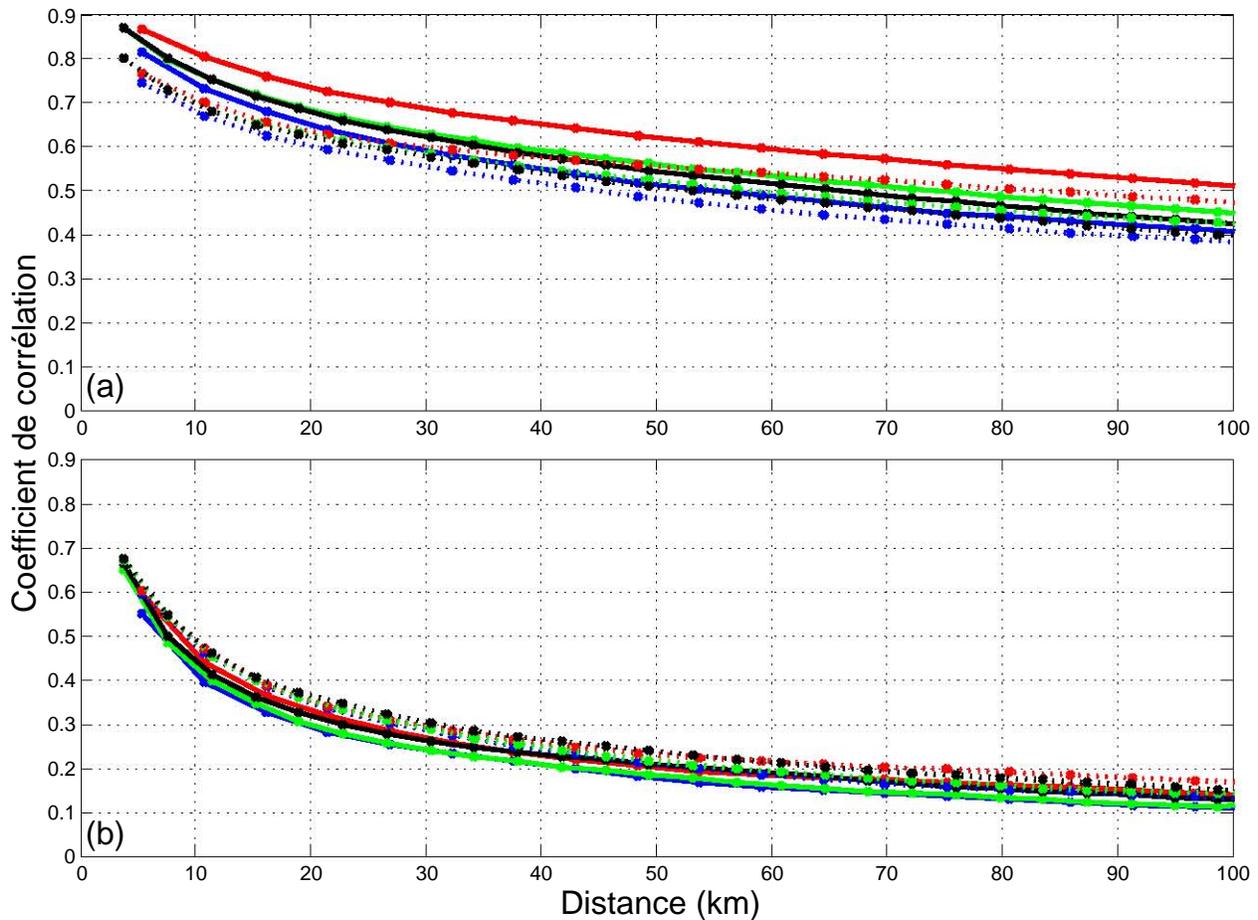


Figure 4.5 – Coefficient de corrélation moyen dans les quatre directions calculé sur les jours avec (a)  $CAPE < 1000$  J/kg et (b)  $CAPE > 1000$  J/kg pour les pixels de 3,8 km observés (lignes continues) et obtenus par la désagrégation en cascade (lignes pointillées) :  $45^\circ$  (rouge),  $90^\circ$  (vert),  $135^\circ$  (bleu) et  $0^\circ$  (noir).

(Figure 4.5a) qu’avec la désagrégation directement à 3,8 km (Figure 4.3a). Par contre, les valeurs de corrélation dans les quatre directions des données désagrégées en cascade sont beaucoup plus près de celles des données observées pour les jours convectifs (Figure 4.5b).

Cependant, même si la désagrégation en cascade représente mieux, en moyenne, la corrélation spatiale des événements convectifs, les champs qu’elle produit présentent le même genre de discontinuités que celles associées aux modèles multi-échelles (Lovejoy et Schertzer, 2010a,b). La Figure 4.6, qui présente la précipitation moyenne journalière (2006-2008) observée, désagrégée directement à 3,8 km et désagrégée en cascade à 3,8 km à chaque pixel, montre que les discontinuités sont même visibles à long terme, surtout sur les jours convectifs. La figure montre aussi que la désagrégation directement à 3,8 km donne des champs trop lisses pour les jours convectifs.

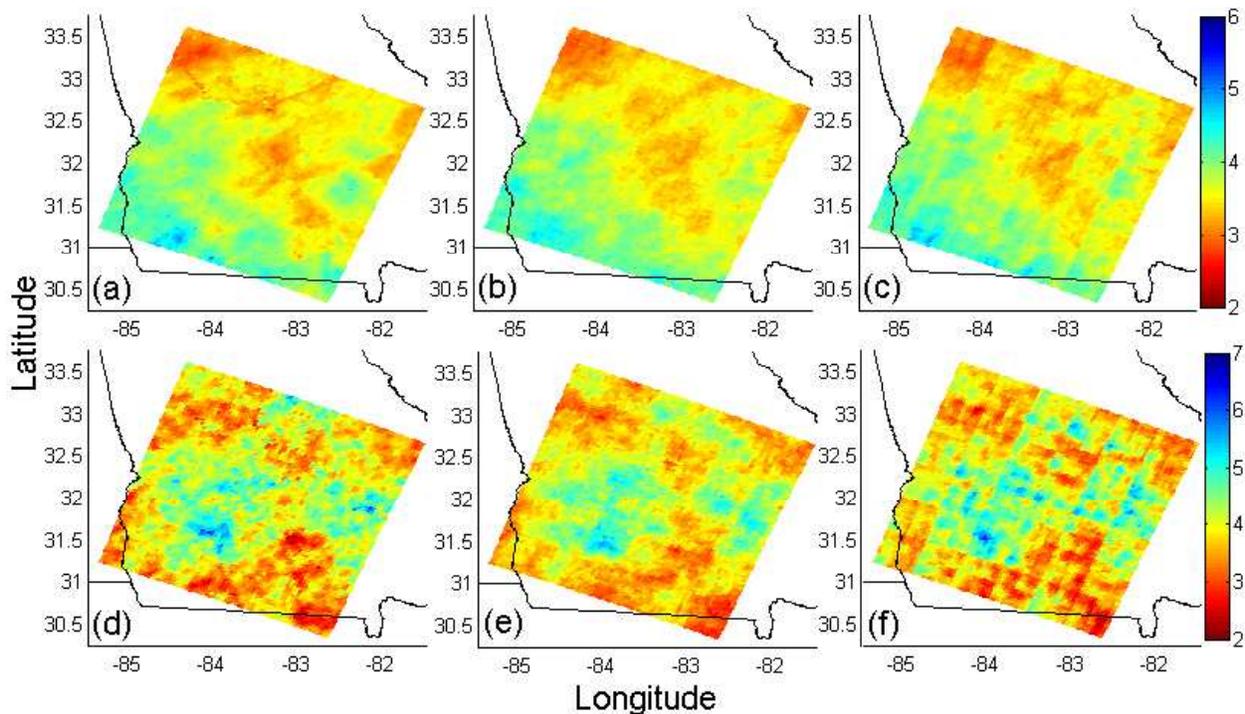


Figure 4.6 – Précipitation moyenne (mm/jour) sur tous les jours (2006-2008) avec  $CAPE < 1000 \text{ J/kg}$  (haut) et  $CAPE > 1000 \text{ J/kg}$  (bas) pour les observations (gauche), pour un des huit champs désagrégés directement à 3,8 km (centre) et désagrégés en cascade à 3,8 km (droite).

### 4.3 Valeurs extrêmes

Les valeurs extrêmes à l'échelle du pixel sont difficiles à reproduire lorsque seules les valeurs de précipitation à l'échelle de la tuile sont connues. Afin d'illustrer si le modèle permet de reproduire les plus grandes valeurs observées sur la période 2006-2008, la Figure 4.7 présente pour chaque pixel, la plus grande valeur journalière observée et désagrégée. Comme il y a huit séries désagrégées, il y a huit valeurs maximales pour chaque pixel ; la plus petite et la plus grande sont présentées.

Les valeurs maximales observées sont généralement situées entre les plus petites et les plus grandes valeurs maximales désagrégées. La Figure 4.7 présente seulement la plus grande valeur pour chaque pixel, mais il est aussi intéressant de voir comment sont distribuées les autres grandes valeurs de précipitation. La Figure 4.8 présente les queues des distributions empiriques observées et désagrégées des précipitations journalières de quatre pixels de 3,8 km situés au centre de la région. Les distributions des pixels de 22,8 km et de la tuile recouvrant les quatre pixels de 3,8 km sont aussi présentées.

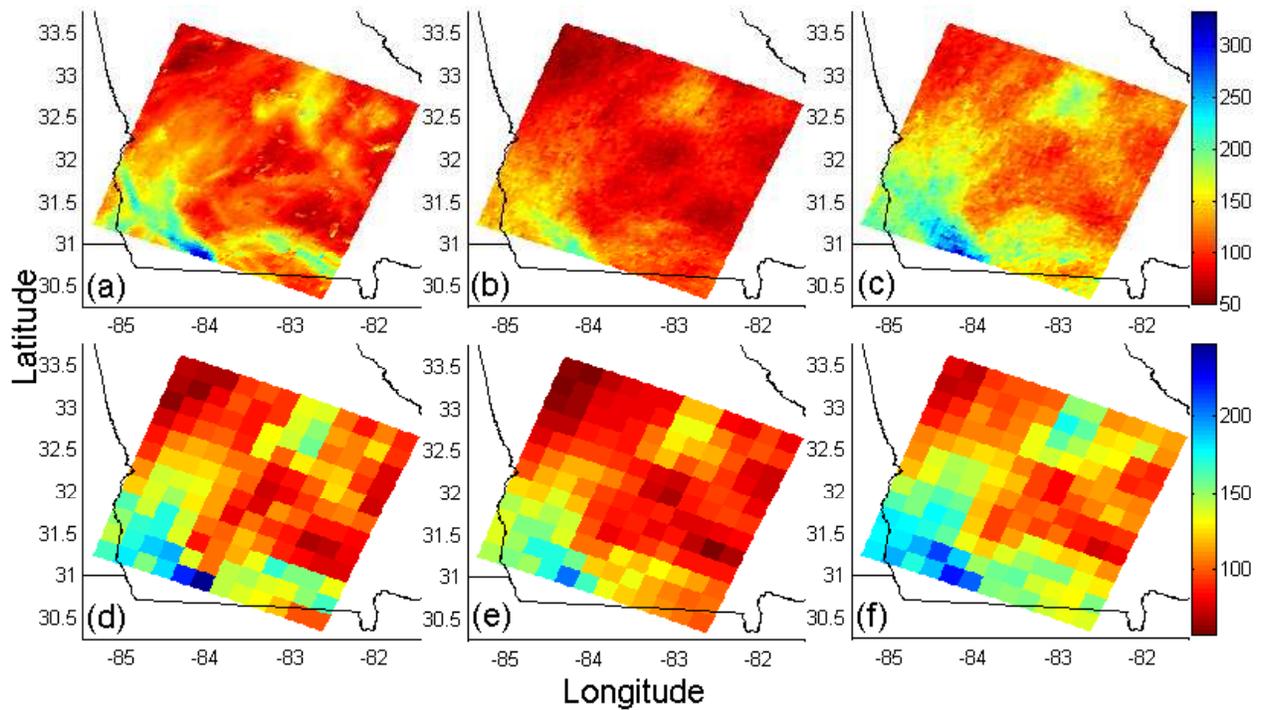


Figure 4.7 – Précipitation journalière maximale (mm) sur les pixels de 3,8 (haut) et 22,8 km (bas) durant la période 2006-2008 : observée (gauche), plus petite (centre) et plus grande (droite) valeurs maximales journalières désagrégées.

Dans l'ensemble, la distribution est bien représentée. Mise à part la précipitation journalière maximale, il n'y a pas de grande différence entre la distribution des extrêmes des trois résolutions spatiales analysées. Les différences entre les résolutions sont plus substantielles lorsque les événements journaliers sont analysés séparément.

Pour avoir une meilleure idée de ce qui se passe à l'échelle journalière, les Figures 4.9 et 4.10 présentent pour chaque jour les valeurs maximales désagrégées et observées sur les  $6 \times 6$  tuiles centrales pour les pixels de tailles 3,8 et 22,8 km. Les jours stratiformes sont séparés des jours convectifs.

Pour les pixels de 3,8 km, les différences entre les valeurs maximales observées à l'échelle du pixel et à l'échelle de la tuile sont généralement plus grandes pour les événements convectifs (Figure 4.9b) que pour les événements stratiformes (Figure 4.9a). Le modèle n'est pas capable de fournir des valeurs maximales aussi élevées que celles observées pour la plupart des événements forts convectifs et aussi pour certains événements forts stratiformes. Pour les pixels de 22,8 km, les différences entre les valeurs maximales observées et désagrégées

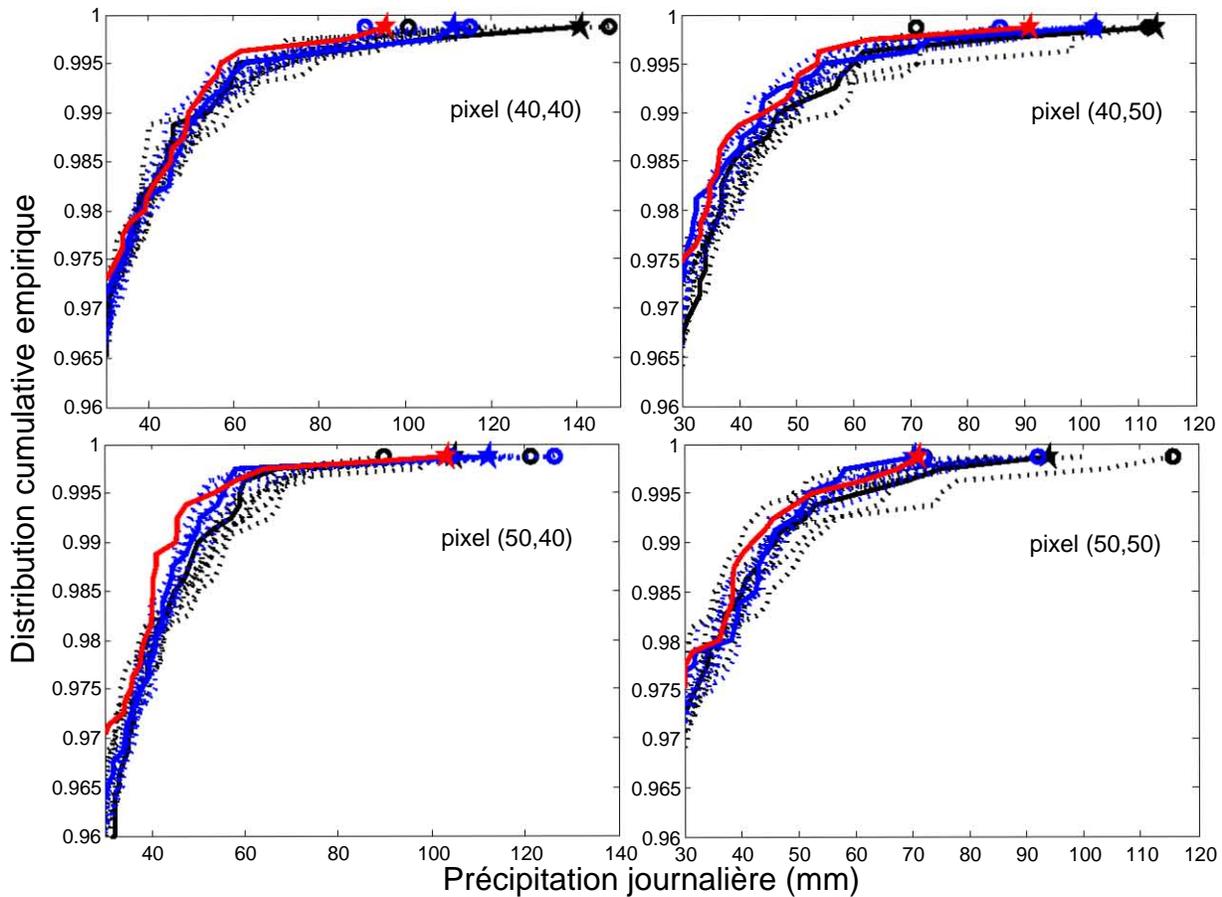


Figure 4.8 – Distributions cumulatives empiriques des précipitations journalières plus grandes que 30 mm observées (lignes continues) et désagrégées (lignes pointillées) pour quatre pixels à résolution 3,8 (noir) et 22,8 km (bleu). La distribution observée sur la tuile est donnée par la ligne rouge. Les étoiles indiquent la précipitation journalière maximale observée et les cercles la plus petite et la plus grande des huit valeurs journalières maximales désagrégées durant la période (2006-2008).

sont très faibles (Figure 4.10).

La Figure 4.11 présente pour chaque jour les valeurs maximales observées et désagrégées en cascade à 3,8 km. Les valeurs maximales sont mieux représentées par la désagrégation en cascade (Figure 4.11b) que par la désagrégation directement à 3,8 km (Figure 4.9b) pour certains événements convectifs. Mais pour la plupart des plus grandes valeurs maximales observées sur les jours convectifs, la désagrégation en cascade donne des valeurs maximales trop faibles.

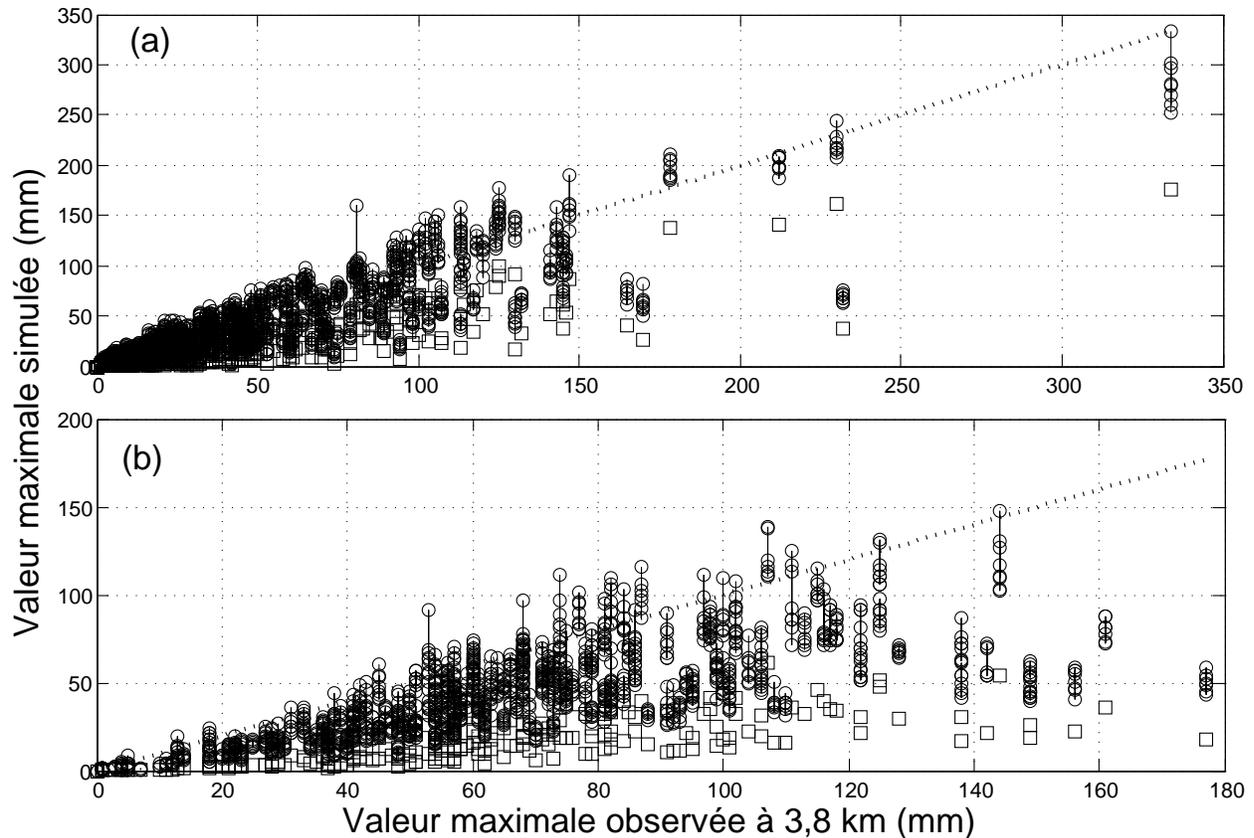


Figure 4.9 – Précipitations maximales observée et désagrégées sur les pixels de 3,8 km des  $6 \times 6$  tuiles centrales pour chaque jour avec (a)  $\text{CAPE} < 1000 \text{ J/kg}$  et (b)  $\text{CAPE} > 1000 \text{ J/kg}$ . Les cercles sont les valeurs maximales désagrégées des huit réalisations et les carrés les valeurs maximales sur les tuiles de 45,6 km utilisées en entrée dans le modèle.

## 4.4 Discussion

Le modèle de désagrégation proposé a été évalué selon sa capacité à reproduire la structure spatiale et les valeurs extrêmes de la précipitation observée. La période de validation du modèle (2006-2008) est différente de la période qui a été utilisée pour son développement (2002-2005).

Dans l'ensemble, le modèle produit des champs de précipitation réalistes. Les Figures 4.2 et 4.7 permettent une appréciation visuelle des champs désagrégés, alors que les graphiques de corrélation (Figures 4.3 et 4.4) illustrent d'une manière plus formelle la capacité du modèle à reproduire la corrélation spatiale et l'anisotropie. Les graphiques de corrélation montrent également que le modèle peut reproduire la corrélation spatiale sur de grandes distances, même s'il prend en considération seulement les huit pixels les plus proches. Cela implique

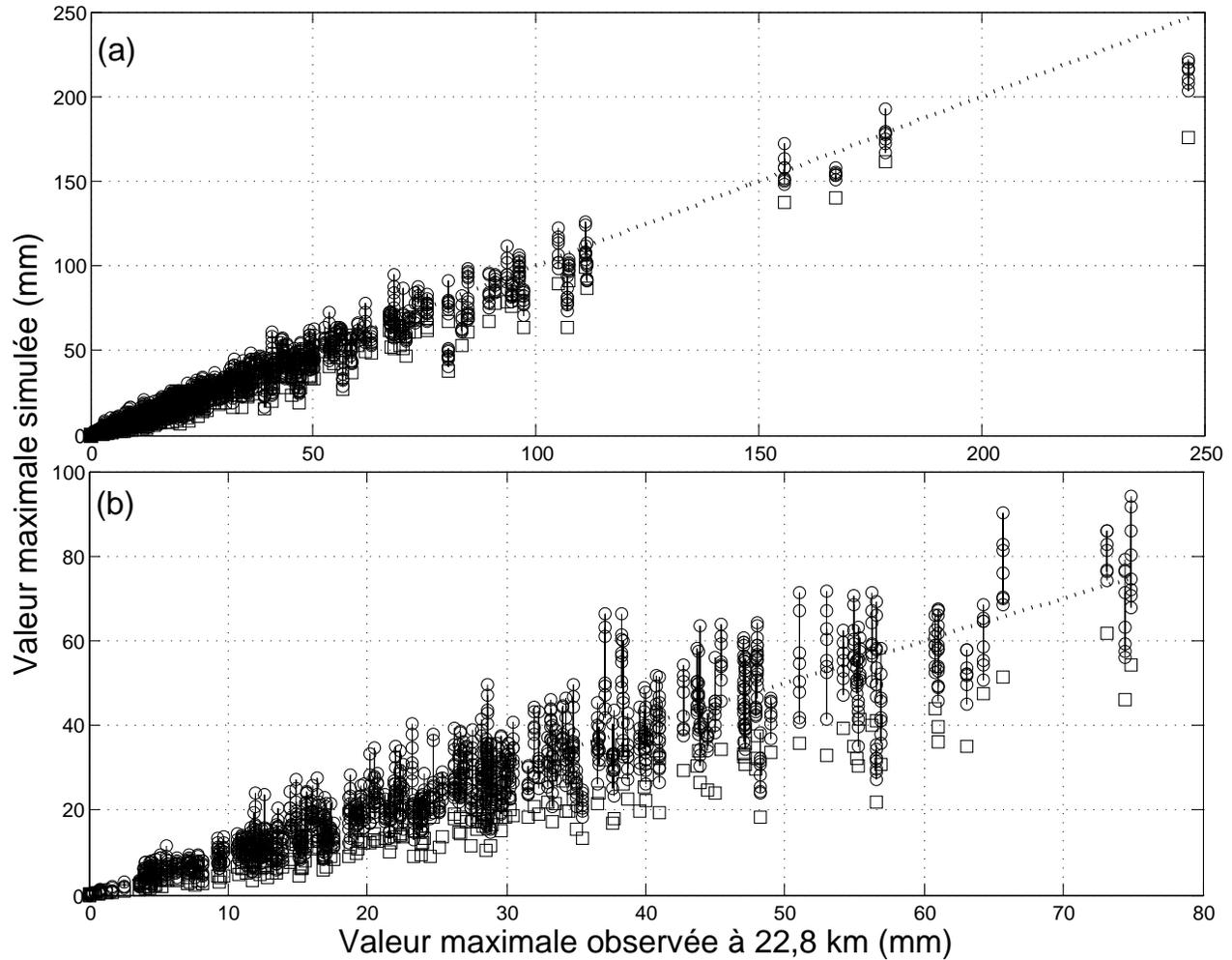


Figure 4.10 – Précipitations maximales observée et désagrégées sur les pixels de 22,8 km des  $6 \times 6$  tuiles centrales pour chaque jour avec (a)  $\text{CAPE} < 1000 \text{ J/kg}$  et (b)  $\text{CAPE} > 1000 \text{ J/kg}$ . Les cercles sont les valeurs maximales désagrégées des huit réalisations et les carrés les valeurs maximales sur les tuiles de 45,6 km utilisées en entrée dans le modèle.

que pour désagréger la précipitation d’une tuile, le modèle utilise l’information contenue sur les tuiles environnantes. C’est pourquoi il est préférable de conserver une zone tampon de largeur d’une tuile de chaque côté pour que la région d’intérêt soit bien simulée. L’impact de cette zone tampon est plus important sur les pixels de 3,8 km que sur ceux de 22,8 km (Tableau 4.1).

Les faiblesses identifiées touchent principalement la précipitation convective ( $\text{CAPE} > 1000 \text{ J/kg}$ ) désagrégée à 3,8 km, pour laquelle le modèle surestime la corrélation spatiale (Figure 4.3b) et sous-estime certaines valeurs extrêmes (Figure 4.9b). Pour le

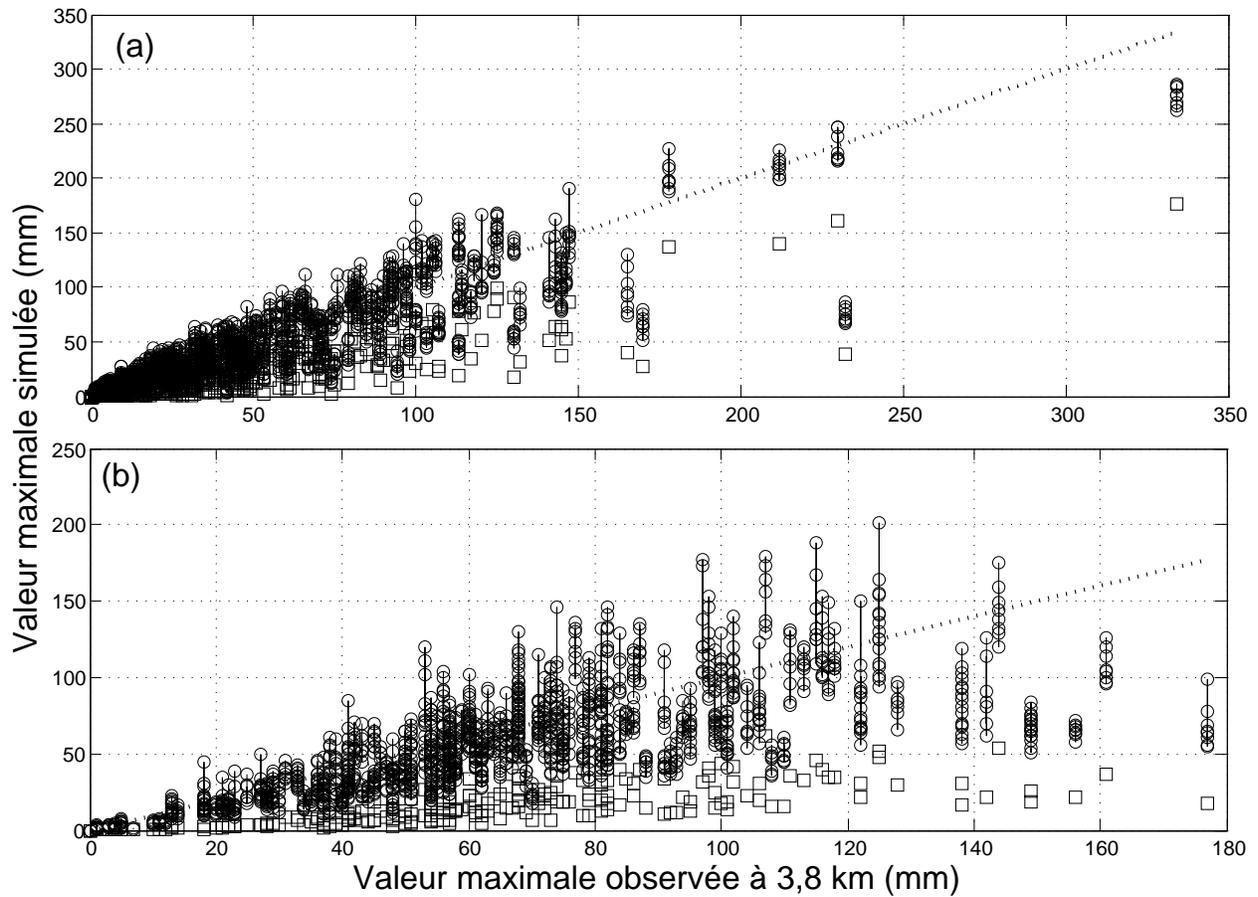


Figure 4.11 – Précipitations maximales observée et désagrégées en cascade sur les pixels de 3,8 km des  $6 \times 6$  tuiles centrales pour chaque jour avec (a) CAPE < 1000 J/kg et (b) CAPE > 1000 J/kg. Les cercles sont les valeurs maximales désagrégées des huit réalisations et les carrés les valeurs maximales sur les tuiles de 45,6 km utilisées en entrée dans le modèle.

modèle, plus un événement est convectif (*i.e.* plus CAPE est grand), plus l'écart-type de la précipitation sur un pixel est grand, la relation entre CAPE et l'écart-type étant linéaire. L'inclusion d'une variable comme CAPE pour tenir compte du type d'événement n'a été que très rarement considérée en désagrégation statistique. Bien que le présent modèle en tienne compte et que les champs désagrégés soient dans l'ensemble comparables aux champs observés, il y a place à amélioration pour tenir compte adéquatement du type d'événement.

La désagrégation en cascade, de 45,6 à 22,8 à 11,4 à 3,8 km, a donné, pour les jours convectifs, des corrélations spatiales moyennes (Figure 4.5b) et des valeurs maximales (Figure 4.11b) plus près des observations que la désagrégation directement à 3,8 km. Cependant, la désagrégation en cascade génère des champs discontinus dans l'espace, contrairement aux champs observés qui peuvent être très variables, mais pas discontinus (Figure 4.6). D'un

point de vue pratique, il n'est pas possible de dire uniquement à partir de ces constatations laquelle des deux approches est la meilleure pour une application sur un événement convectif. D'un point de vue théorique, l'amélioration de la structure spatiale des champs convectifs désagrégés passe vraisemblablement par une meilleure connaissance du rôle de CAPE sur la corrélation spatiale à fine échelle. Il faut aussi mentionner que l'utilisation d'une moyenne journalière de CAPE est due au fait que le moment exact où la pluie est tombée est inconnu. Cela n'aurait pas permis d'éliminer totalement les imperfections du modèle, mais les résultats auraient pu être améliorés si la valeur de CAPE juste avant l'événement de précipitation avait pu être utilisée, comme dans Perica et Foufoula-Georgiou (1996).

Les valeurs désagrégées sont généralement plus près des observations pour les pixels de 22,8 km que pour les pixels de 3,8 km. Dans le deux cas, la précipitation originale est à résolution 45,6 km. Il faut ainsi générer quatre valeurs aléatoires pour les pixels de 22,8 km et 144 valeurs aléatoires pour les pixels de 3,8 km. Il est donc plus facile de reproduire les propriétés statistiques de la précipitation sur des plus grands pixels puisque les valeurs cibles (*i.e.* les valeurs observées sur les pixels) sont plus près des valeurs utilisées en entrée. C'est pourquoi il serait incorrect de dire que le modèle à 22,8 km performe mieux que le modèle à 3,8 km. Pour une comparaison juste, il faudrait analyser les résultats sur une taille de pixel commune. Le modèle à 22,8 km est bon lorsqu'évalué sur des pixels de 22,8 km, mais il ne serait pas capable de représenter la corrélation spatiale et les valeurs extrêmes pour une résolution de 3,8 km.

La résolution à laquelle le modèle doit être utilisé dépend des besoins de l'application. Le modèle développé et appliqué à la plus fine résolution disponible pour son calage est plus apte à représenter les extrêmes à l'échelle locale, mais nécessite plus de temps de calcul. Par exemple, les champs désagrégés à 3,8 km comptent 36 fois plus de données à générer et à stocker que les champs à 22,8 km. Dans plusieurs cas, le temps de calcul ne sera pas une contrainte puisque le modèle génère des résultats assez rapidement. Dans la présente application, réalisée à partir d'un Intel(R) Core(TM) i7 CPU avec 2,67 gigahertz de fréquence et 6,0 gigabits de mémoire vive, il n'a fallu que quelques heures pour générer les huit champs de  $(96 \times 96)$  pixels de 3,8 km pour les 801 jours ; pour les pixels de 22,8 km, il n'a fallu que quelques minutes. Cependant, le temps peut être un facteur déterminant pour certaines applications, par exemple si un modèle météorologique prévoit l'imminence d'un événement extrême de précipitation sur un territoire et qu'un grand nombre de champs désagrégés est

désiré pour estimer l'ampleur des crues extrêmes à venir. Dans un tel cas, au temps nécessaire pour désagréger s'ajoute le temps nécessaire à la modélisation hydrologique de tous les champs désagrégés produits. Le temps nécessaire pour produire une simulation hydrologique peut augmenter lorsque la résolution spatiale de la précipitation utilisée en entrée augmente. Dans un tel cas, il sera important de déterminer quelle résolution spatiale est nécessaire pour répondre adéquatement au besoin de l'application. Cet aspect est analysé à la troisième partie de la thèse (Chapitres 8, 9 et 10).



## Deuxième partie

# DÉSAGRÉGATION STATISTIQUE DE LA PRÉCIPITATION SUR UNE RÉGION MONTAGNEUSE



# Chapitre 5

## REVUE DE LITTÉRATURE

La désagrégation de la précipitation en régions montagneuses est encore plus importante que sur les régions avec peu de relief puisque les modèles météorologiques et climatiques, de par leurs résolutions spatiales, ne représentent pas adéquatement l'impact à fine échelle de la topographie sur la précipitation (Maurer et Hidalgo, 2007). Bien qu'il soit évident que la topographie influence la précipitation, il est difficile de définir avec précision son impact sur la distribution spatiale de la précipitation à fine échelle. Une manière simple et couramment utilisée en interpolation pour tenir compte de la topographie est de supposer que la précipitation augmente linéairement avec l'altitude (Diodato, 2005; Clark et Slater, 2006; Neilson *et al.*, 2010). Bindlish et Barros (2000) sont allés plus loin en développant une méthode d'interpolation qui exprimait la précipitation sur un pixel comme une moyenne pondérée de la précipitation sur les tuiles environnantes. Le nombre de tuiles environnantes considérées dépend de la vitesse horizontale du vent et le poids des tuiles dépend de la vitesse verticale du vent et de la dimension fractale de la topographie.

Autrement, il existe des modèles stochastiques appelés *modèles de précipitation orographique* (MPO) qui ont pour but de mimer les processus physiques qui mènent à la précipitation en montagnes. Ce chapitre décrit d'abord ces modèles (Section 5.1) et explique par la suite comment la topographie a été incorporée aux modèles de désagrégation existants (Section 5.2). Une discussion conclut ce chapitre (Section 5.3).

## 5.1 Modèles de précipitation orographique (MPO)

Les MPO simulent la précipitation orographique à fine résolution à partir de l'information de champs atmosphériques tels la vitesse des vents, le taux d'humidité, la température et la pression. Pour développer ces modèles, les champs atmosphériques proviennent de réanalyses d'observations. Ces modèles sont utiles lorsque la précipitation est inconnue, même à mésoéchelle. Le principe de base de ces modèles est qu'une masse d'air humide qui se déplace vers une montagne s'élève en altitude, ce qui entraîne une diminution de la température et la précipitation de l'eau présente dans la colonne d'air sous forme d'humidité.

Sinclair (1994) exprime  $v_z$ , la vitesse verticale du vent près de la surface, proportionnellement au produit scalaire entre  $v_{x,y}$ , la vitesse horizontale des vents à basse altitude (moyenne des niveaux de pression 850 et 1000 hPa), et  $\Delta z$ , la différence d'altitude d'un pixel à un autre :

$$v_z \propto \vec{v}_{x,y} \cdot \vec{\Delta z} . \quad (5.1)$$

La vitesse verticale du vent au niveau de pression  $p$ , notée  $v_z^{(p)}$ , décroît avec l'altitude selon l'équation suivante :

$$v_z^{(p)} = \begin{cases} v_{z_0}^{(p)} + v_z \left( \frac{p-p_t}{p_s-p_t} \right)^{\tan(\gamma\pi/4)} & \text{pour } p_s > p > p_t \\ 0 & \text{autrement.} \end{cases} \quad (5.2)$$

La variable  $v_{z_0}^{(p)}$  est la vitesse verticale du vent à grande échelle simulée par les réanalyses d'observations au niveau de pression  $p$ ,  $p_s$  est la pression de surface,  $p_t$  une pression seuil ( $p_t = 200$  hPa dans Sinclair, 1994) et  $\gamma$  un paramètre à estimer dans l'intervalle (0,2). La hauteur de précipitation formée sur le niveau de pression  $p$  augmente linéairement par rapport à  $v_z^{(p)}$  et dépend aussi du rapport de mélange (*i.e.* masse de la vapeur d'eau sur la masse d'air sec), de la température et du taux d'humidité relative provenant des réanalyses. Pour plus de détails, le lecteur est invité à consulter Sinclair (1994).

Hay (1998), Hay et McCabe (1998), et Pandey *et al.* (2000) utilisent des concepts développés par Rhea (1978) pour simuler la précipitation montagneuse hivernale à fine résolution. Contrairement à Sinclair (1994), les champs atmosphériques des réanalyses d'observations sont seulement utilisés aux frontières du domaine. L'évolution des champs atmosphériques

sur la région analysée est entièrement simulée par le MPO. La grille à fine résolution est orientée dans la direction du vent à 700 hPa sur l'axe horizontal et est divisée en niveaux de pression sur l'axe vertical (Hay, 1998; Hay et McCabe, 1998; Pandey *et al.*, 2000). Le taux de précipitation à la frontière des pixels  $i$  et  $i + 1$  sur l'axe horizontal au niveau de pression  $p$ , noté  $r_{i,i+1}^{(p)}$ , est donné par l'équation suivante :

$$r_{i,i+1}^p = \frac{v_{x,y}^{(p)} \Delta h^{(p)}}{\rho_{\text{eau}} g l} E \left( q_i^{(p)} + \Delta c_{i,i+1}^{(p)} \right) \quad (5.3)$$

où  $v_{x,y}^{(p)}$  est la vitesse horizontale du vent à la frontière de la grille au niveau de pression  $p$ ,  $\Delta h^{(p)}$  est l'épaisseur (en unité de pression) du niveau de pression  $p$ ,  $\rho_{\text{eau}}$  est la densité de l'eau,  $g$  est l'accélération gravitationnelle,  $l$  est la résolution horizontale de la grille,  $E$  est le coefficient d'efficacité, c'est-à-dire la proportion de l'eau condensée qui tombe en précipitation,  $q_i^{(p)}$  est le contenu en vapeur d'eau sur le pixel  $i$  au niveau de pression  $p$  et  $\Delta c_{i,i+1}^{(p)}$  est la condensation (si  $\Delta c_{i,i+1}^{(p)} > 0$ ) ou l'évaporation (si  $\Delta c_{i,i+1}^{(p)} < 0$ ) causée par le déplacement vertical de l'air au niveau de pression  $p$  entre les pixels  $i$  et  $i + 1$  ( $r_{i,i+1}^p$  est ramené à 0 si  $(q_i^{(p)} + \Delta c_{i,i+1}^{(p)}) < 0$ ). Dans Pandey *et al.* (2000), les niveaux de pression entre 300 et 900 hPa sont utilisés et  $\Delta c_{i,i+1}^{(p)}$  est proportionnel à  $(p/900 \text{ hPa}) \Delta z_{i,i+1}$  où  $\Delta z_{i,i+1}$  est la différence d'altitude entre les pixels  $i$  et  $i + 1$ . Ils posent aussi  $q_i^{(p)} = (1 - E) (q_{i-1}^{(p)} + \Delta c_{i-1,i}^{(p)})$ . Hay (1998) pose  $E = -0,01 t_{\text{top}}$  où  $t_{\text{top}}$  est la température (en Celsius) pour le plus haut niveau de pression avec humidité relative évaluée à la frontière de la grille supérieure à 65 %. Dans Pandey *et al.* (2000),  $E$  est proportionnel à  $-0,01(t_{550\text{hPa}} - t_{700\text{hPa}})$ , mais est aussi fonction de la quantité d'humidité disponible, de la stabilité de l'atmosphère et de la direction des vents. Si les vents proviennent de l'océan, le coefficient  $E$  est plus élevé.

Smith et Barstad (2004) ont proposé un autre MPO, appliqué par Barstad et Smith (2005) et Schuler *et al.* (2008), qui distingue la vapeur d'eau des hydrométéores et qui tient compte du temps nécessaire pour transformer de la vapeur d'eau en hydrométéores et pour que l'hydrométéore tombe au sol. Ce modèle est uniquement applicable en régions montagneuses puisqu'il suppose que la précipitation est constante dans l'espace s'il n'y a pas de variation d'altitude; les Chapitres 3 et 4 ont montré que ce n'est pas le cas. Malgré les différences entre les MPO, il demeure que leur principale hypothèse commune est que la hauteur de précipitation est proportionnelle au produit scalaire entre le vecteur vent et le gradient d'altitude.

## 5.2 Inclusion de la topographie en désagrégation

Peu de modèles de désagrégation tiennent compte explicitement de la topographie. Selon notre connaissance de la littérature existante, la prise en compte de la topographie en désagrégation a été faite seulement à partir de modèles multi-échelles. Segond *et al.* (2006) mentionnaient que l'inclusion de la topographie à partir des modèles de cellules de précipitation restait à faire. Toujours en 2006, Clark et Slater (2006) mentionnaient que la topographie n'avait toujours pas été incorporée dans les modèles multifractals. Certains ont analysé les propriétés multi-échelles de la précipitation en régions montagneuses sans désagréger (Nykanen et Harris, 2003 ; Nykanen, 2008 ; Ebtehaj et Foufloula-Georgiou, 2010) ; d'autres ont désagrégué en montagnes à partir de modèles multi-échelles, mais sans tenir compte explicitement de la topographie (Ahrens, 2003 ; Deidda *et al.*, 2006 ; Groppelli *et al.*, 2011).

Avant 2006, les modèles de désagrégation en montagnes incorporaient l'impact de la topographie de manière indirecte à partir des moyennes observées à fine résolution sur de longues périodes. Jothityangkoon *et al.* (2000), Pathirana et Herath (2002), et Sharma *et al.* (2007) ont désagrégué à partir du modèle  $\beta$ -lognormal d'Over et Gupta (1996) et multiplié le champ résultant par le ratio de la précipitation interannuelle mensuelle moyenne à l'échelle du pixel sur la précipitation interannuelle mensuelle moyenne à l'échelle de la tuile (le volume d'eau sur la tuile est ainsi conservé). Badas *et al.* (2006) sont allés un peu plus loin en exprimant la précipitation moyenne à long-terme par une fonction linéaire de l'altitude du pixel.

Guan *et al.* (2009) mentionnent que les modèles statistiques existants qui créent des cartes de précipitations en régions montagneuses peuvent bien performer à l'échelle mensuelle, mais pas à l'échelle journalière ou horaire. Ils ajoutent qu'il faut inclure le vent pour bien reproduire la précipitation à fine résolution spatiale. Ils ont développé une méthode de désagrégation statistique qui calcule la moyenne pondérée de deux champs de précipitation. Le premier champ est estimé à partir d'une régression linéaire tenant compte de la direction du vent, de l'altitude, de la latitude, de la longitude, et de l'orientation de la pente à chaque pixel. Le deuxième est un champ désagrégué à partir d'un modèle multifractal homogène dans l'espace (Gupta et Waymire, 1993).

## 5.3 Discussion

La précipitation orographique à fine résolution spatiale est difficile à estimer. Lorsque la précipitation, même à mésoéchelle, est inconnue, les MPO peuvent être utilisés pour simuler la précipitation à fine résolution. Cependant, même s'il est supposé que les MPO reproduisent bien les processus physiques régissant la précipitation orographique, il faut considérer que les variables atmosphériques nécessaires à son estimation sont disponibles au mieux à mésoéchelle (*i.e.* résolution  $> 10$  km). Ceci étant dit, un fait demeure : un vent horizontal fort qui frappe une montagne crée un déplacement d'air vertical et mène ainsi à des conditions propices à la genèse de la précipitation (Sinclair, 1994 ; Hay, 1998 ; Hay et McCabe, 1998 ; Pandey *et al.*, 2000 ; Smith et Barstad, 2004). Ce processus peut-être utilisé en désagrégation.

Relativement peu de modèles de désagrégation existants incluent explicitement la topographie. Outre le fait que la précipitation orographique puisse être complexe à modéliser, il faut aussi considérer qu'il existe très peu de données de précipitation à fine résolution de qualité en montagnes. Les modèles de désagrégation existants tiennent compte de la topographie de manière très simple. Certains multiplient un champ homogène désagrégé par des moyennes à long-terme (Jothityangkoon *et al.*, 2000 ; Pathirana et Herath, 2002 ; Sharma *et al.*, 2007) ; d'autres supposent une augmentation linéaire de la précipitation en fonction de l'altitude (Badas *et al.*, 2006). Guan *et al.* (2009) sont allés plus loin en incorporant le vent et l'orientation de la pente à leur modèle. Cependant, même lorsqu'un modèle multi-échelle tient compte de la topographie, il a quand même de la difficulté à représenter la corrélation spatiale (Ahrens, 2003 ; Sharma *et al.*, 2007). Le Chapitre 4 a montré qu'un modèle stochastique pouvait représenter la corrélation spatiale. À notre connaissance, aucun modèle de désagrégation stochastique existant n'inclut explicitement la topographie. Dans les deux prochains chapitres, une version modifiée du modèle développé au Chapitre 3 qui tient compte de la topographie est développé et appliqué.



# Chapitre 6

## DÉVELOPPEMENT DU MODÈLE

Dans ce chapitre, le modèle développé pour une région avec peu de relief (Chapitre 3) est adapté afin qu'il puisse inclure l'impact de la topographie. Le modèle qui en résulte doit aussi produire des champs de précipitation réalistes. Le modèle qui offre le meilleur compromis entre la performance et la parcimonie est choisi. Les prochaines sections présentent la région et les données utilisées (Section 6.1), la structure mathématique du modèle (Section 6.2) et une discussion des résultats (Section 6.3).

### 6.1 Région à l'étude

Le modèle est développé sur une région du nord-ouest des États-Unis d'environ  $300 \times 300$  km<sup>2</sup> située principalement dans l'État de Washington et touchant aussi l'Oregon (Figure 6.1). Cette région a été choisie parce que les montagnes qui s'y trouvent sont parmi les plus hautes couvertes par les données de *Stage IV*. Il y a les montagnes Olympiques (*Olympic mountains*) au nord-ouest et la chaîne des Cascades (*Cascade Range*) qui traverse la région du sud-ouest au nord-est dont le point culminant est le mont Rainier (4392 m). Le mont St. Helens a été en éruption quelques jours durant la période 2004-2008, ce qui aurait pu influencer la précipitation dans les heures et les jours qui ont suivi. L'impact potentiel de ces petites éruptions est donc négligeable sur l'ensemble du jeu de données. Le climat y est océanique, sauf à l'est de la chaîne des Cascades, où le climat est beaucoup plus sec (*Western Regional Climate Center, 2002*). Durant la période d'étude (2002-2008), la précipitation annuelle moyenne sur la région a varié de 1069 (2008) à 1586 mm (2006). La précipitation journalière moyenne (2002-2008) varie de 0,3 à 14,7 mm selon le pixel, ce qui illustre l'impact important de la topographie sur les régimes pluviométriques dans cette région. Environ 80 % de la

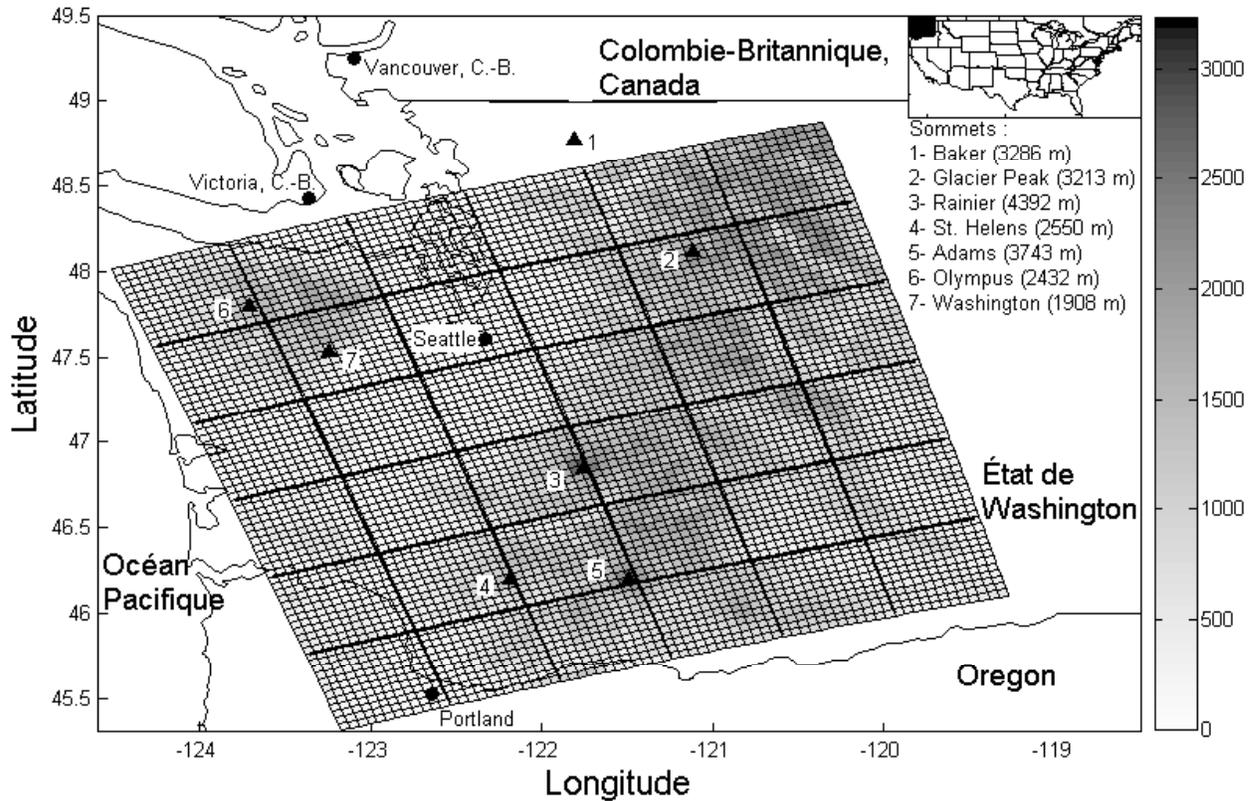


Figure 6.1 – Région étudiée avec l’altitude (m) des  $72 \times 72$  pixels, de 4,4 km utilisés. Les lignes plus épaisses forment les  $6 \times 6$  tuiles.

précipitation tombe entre novembre et avril, principalement sous forme de pluie pour les régions de faible altitude et sous forme de neige pour les régions de haute altitude.

### 6.1.1 Données de précipitation

Comme pour la première partie, les données de précipitation proviennent de la grille *Stage IV* (Lin et Mitchell, 2005), sauf que pour le nord-ouest des États-Unis, ces données sont traitées par le *Northwest RFC* (NWRFC). En montagnes, l’estimation de la précipitation à partir de radars météorologiques est problématique (Joss et Germann, 2000), notamment parce que les impulsions électromagnétiques émises frappent les montagnes et produisent de nombreux échos de terrain. À moins d’avoir un réseau de radars très dense ou de placer les radars aux sommets de montagnes, une estimation quantitative adéquate de la précipitation ne peut être obtenue pour ces régions. C’est pourquoi le NWRFC utilise le modèle *PRISM* (*Parameter-elevation Regressions on Independent Slopes Model*; Daly et al., 1994, 1997, 2002) pour estimer la précipitation à fine résolution. Dans PRISM,  $R_{i,j}$ , la précipitation au pixel  $(i, j)$ , est estimée par une régression linéaire simple par rapport à l’altitude  $Z_{i,j}$  :

$$R_{i,j} = a_0^{i,j} + a_1^{i,j} Z_{i,j} \quad (6.1)$$

où  $a_0^{i,j}$  et  $a_1^{i,j}$  sont des paramètres spécifiques au pixel  $(i, j)$  estimés à partir de la précipitation enregistrée aux stations météorologiques environnantes. Le poids d'une station dépend non seulement de sa distance par rapport au pixel, mais aussi de variables physiographiques déterminantes pour la précipitation, telles l'altitude, la proximité des côtes et l'orientation de la pente du terrain à la station. Les radars météorologiques ne fournissent pas d'estimations quantitatives de précipitation sur la région, mais ils sont quand même utilisés pour suivre le déplacement et caractériser la distribution spatiale de la précipitation (Alex Orr, NWRFC, communication personnelle).

La région est couverte par  $72 \times 72$  pixels de la grille de *Stage IV* à partir desquels sont formées  $6 \times 6$  tuiles pour la désagrégation (Figure 6.1). La précipitation journalière sur un pixel est arrondie au mm près. La résolution d'un pixel est de 4,4 km, ce qui donne des tuiles de 52,8 km de côté. De 2002 à 2005, les 2 996 076 données disponibles à 4,4 km pour lesquels la précipitation sur la tuile est supérieure à 0,1 mm servent à l'estimation des paramètres du modèle de désagrégation qui est appliqué sur les 750 jours ayant reçu de la précipitation de 2006 à 2008. Les propriétés statistiques de la précipitation journalière sont d'abord analysées à une résolution de 4,4 km, et ensuite aux résolutions 8,8, 13,2, 17,6 et 26,4 km à partir des pixels de *Stage IV* agrégés.

### 6.1.2 Données atmosphériques

Comme pour la Partie I, les données atmosphériques proviennent des réanalyses NARR (Mesinger *et al.*, 2004). La résolution horizontale de la grille est de 32 km et le pas de temps est de trois heures. Pour chaque variable atmosphérique d'intérêt, une valeur par jour par tuile est retenue en calculant la moyenne des valeurs sur les huit pas de temps et de tous les points de la grille NARR situés sur la tuile. Les variables d'intérêt sont CAPE, la vitesse et la direction du vent à 500 et 700 hPa.

La vitesse et la direction du vent à 700 hPa, correspondant à une altitude d'environ 3000 m, avaient été retenues pour expliquer l'anisotropie pour le modèle développé au sud-est des États-Unis. Comme l'altitude est plus élevée au nord-ouest des États-Unis, la vitesse et la

direction du vent à 500 hPa, correspondant à une altitude d'environ 5500 m, est également considérée. Pour ce qui est de CAPE, les valeurs sont généralement très faibles puisque la région est relativement froide comparée au sud-est des États-Unis et que la grande majorité des événements de précipitation ont lieu l'hiver. La plus grande valeur de CAPE observée sur une tuile ayant reçu de la précipitation est inférieure à 800 J/kg et inférieure à 500 J/kg pour les tuiles ayant reçu plus de trois millimètres de précipitation. Comme les événements convectifs sont faibles et marginaux, la variable CAPE n'est pas retenue pour le développement du modèle en montagnes.

### 6.1.3 Données topographiques

L'altitude sur la région est fournie à résolution d'une seconde d'arc (environ 30 m) par le *United States Geological Survey* (USGS). Les données sont disponibles gratuitement à l'adresse <http://seamless.usgs.gov>. Pour les analyses, seule la valeur moyenne d'altitude par pixel, calculée à partir de tous les points de la grille du USGS situés sur le pixel, est retenue. L'altitude sur la région évaluée à l'échelle des pixels de 4,4 km varie de 0 à 3236 m. La différence d'altitude maximale entre deux pixels voisins est de 1304 m.

## 6.2 Structure mathématique du modèle

Les hypothèses de base retenues initialement pour le développement du modèle applicable sur une région avec peu de relief sont reprises pour le modèle développé dans le présent chapitre. Premièrement, la précipitation initiale sur les tuiles est supposée précise et non-biaisée, ce qui implique que le volume de précipitation désagrégée est imposé égal au volume de précipitation sur la tuile. Si la précipitation mésoéchelle utilisée en entrée provient d'un modèle climatique ou météorologique, cette hypothèse est plus forte en régions montagneuses qu'en régions avec peu de relief. En effet, les modèles mésoéchelles ne représentent pas adéquatement les effets topographiques locaux (Maurer et Hidalgo, 2007), ce qui peut avoir comme effet de biaiser l'estimation de la précipitation. Dans le cas présent, les champs de précipitation à l'échelle des tuiles sont construits à partir de la précipitation à fine résolution. Pour une comparaison juste entre les champs à fine résolution observés et désagrégés, il faut que les volumes d'eau sur les deux champs soient égaux. Si le modèle désagrégait la précipitation d'un modèle climatique ou météorologique, il faudrait possiblement corriger le biais des volumes d'eau désagrégés. Cependant, la correction du

biais des modèles climatiques ou météorologiques va au-delà des objectifs de la présente thèse.

Deuxièmement, les accumulations de précipitation fournies par la grille à fine résolution de *Stage IV* sont considérées exactes et aucune correction de ces estimations n'est faite. Dans la région étudiée, l'estimation de la précipitation journalière à fine résolution est complexe de sorte que le NWRFC doit se servir d'un modèle statistique (PRISM) pour interpoler les données. L'incertitude sur la précipitation estimée dans cette région est selon toute vraisemblance plus grande que pour une région avec peu de relief. Néanmoins, les estimations du NWRFC fournies sur la grille de *Stage IV* correspondent à l'état de l'art pour l'estimation de la précipitation journalière en régions montagneuses. Troisièmement, il est supposé que connaissant la précipitation mésoéchelle au jour  $t$ , la précipitation à fine échelle au jour  $t-1$  n'amène pas d'information supplémentaire. Cette hypothèse a été vérifiée au Chapitre 3.

En l'absence de variation d'altitude, le modèle développé doit, pour des raisons de cohérence, être équivalent à celui développé sur une région avec peu de relief, à l'exception que CAPE n'est pas considéré. Pour adapter le modèle, l'approche utilisée est comparable à celle du Chapitre 3. L'impact de la topographie est d'abord évalué sur l'espérance et ensuite sur l'écart-type de la précipitation à l'échelle du pixel. La distribution de la précipitation sur un pixel est considérée lognormale (Chapitre 3). Dans un premier temps, le modèle est développé à partir des données disponibles sur les pixels de 4,4 km. Les paramètres sont ensuite estimés pour les autres tailles de pixels. Afin de vérifier l'impact de la prise en compte de la topographie dans le modèle, les résultats du modèle développé dans le présent chapitre sont comparés avec ceux du modèle qui ne tient pas compte de la topographie.

### 6.2.1 Espérance de la précipitation sur un pixel

Comme le modèle adapté doit être équivalent à celui de la Partie I en l'absence de variation d'altitude, il est supposé d'emblée que l'espérance sur un pixel est une fonction linéaire des huit voisins les plus proches comme à l'équation (3.2). Avant d'inclure l'impact de la topographie, l'impact du vent sur l'anisotropie est analysé de la même façon que sur une région avec peu de relief. Le Tableau 6.1 présente la somme du carré des écarts entre la précipitation observée et l'espérance calculée selon deux équations tenant compte du vent, une pour le vent à 700 hPa, l'autre à 500 hPa. Le vent à 700 hPa a été retenu pour le modèle de la Partie I. Comme l'altitude est plus élevée dans la présente région d'étude, la pression à la surface est plus faible et peut même être inférieure à 700 hPa. C'est pourquoi le vent à

Tableau 6.1 – Somme du carré des écarts entre la précipitation observée à 4,4 km et deux équations d’espérance tenant compte du vent à 500 et à 700 hPa (2002-2005).

Comme équation (3.2) (3 paramètres), sauf que :	Nombre de paramètres	Somme de carrés (% p/r à l’équation (3.2))
$\beta_{\times} = \beta_v V_{i,j}^{(700\text{hPa})} \cos \left\{ 2 \left( \omega_{i,j}^{(700\text{hPa})} - 45^\circ \right) \right\} ;$ $\beta_{+} = \beta_v V_{i,j}^{(700\text{hPa})} \cos \left\{ 2 \left( \omega_{i,j}^{(700\text{hPa})} - 90^\circ \right) \right\} .$	2	102,1
$\beta_{\times} = \beta_v V_{i,j}^{(500\text{hPa})} \cos \left\{ 2 \left( \omega_{i,j}^{(500\text{hPa})} - 45^\circ \right) \right\} ;$ $\beta_{+} = \beta_v V_{i,j}^{(500\text{hPa})} \cos \left\{ 2 \left( \omega_{i,j}^{(500\text{hPa})} - 90^\circ \right) \right\} .$	2	102,2

500 hPa est aussi testé. Pour une définition des variables du tableau, le lecteur est invité à consulter l’Annexe B ou le Chapitre 3.

Contrairement au sud-est des États-Unis, le vent ne permet pas d’expliquer l’anisotropie observée sur les huit voisins les plus proches pour le nord-ouest des États-Unis. Il y a deux explications potentielles : soit le vent à mésoéchelle est trop différent du vent à l’échelle locale, qui est modifié par la topographie ; soit le modèle PRISM, utilisé pour produire la précipitation à fine résolution, n’est pas capable de reproduire l’impact du vent à fine résolution, même s’il tient compte du positionnement des stations sur les montagnes. Quoiqu’il en soit, le vent à l’échelle de la tuile n’amène pas d’information pertinente pour la désagrégation et n’est pas utilisé dans l’équation de l’espérance pour le modèle ne tenant pas compte de la topographie.

La Figure 6.2 présente la précipitation moyenne journalière (2002-2005) de tous les pixels de 4,4 km en fonction de l’altitude. Même si en général la précipitation augmente en fonction de l’altitude, la relation n’est pas très forte ( $R^2 = 0,04$ ). Diverses manières d’inclure la topographie sont testées. Le Tableau 6.2 présente les principales. D’autres équations ont été testées, mais n’ont pas donné de meilleurs résultats. Dans tous les cas considérés, le coefficient tenant compte de la topographie est l’augmentation relative de précipitation par rapport à la précipitation moyenne sur les huit voisins les plus proches  $\bar{A}^{(i,j)}$ .

Les trois premières équations, tenant compte uniquement de l’altitude du pixel, mènent

Tableau 6.2 – Somme du carré des écarts entre la précipitation observée à 4,4 km et certaines équations d'espérance tenant compte de la topographie. Les variables sont définies à l'Annexe B. Tous les modèles testés ont un paramètre de plus que l'équation (3.2).

Comme équation (3.2) + ...	Somme de carrés (% p/r à l'équation (3.2))
$\beta (Z_{i,j} - \mathbf{Z}_{i,j}) \bar{A}^{(i,j)}$	97,6
$\beta (Z_{i,j} - \bar{Z}_{i,j}^{(1)}) \bar{A}^{(i,j)}$	99,5
$\beta (Z_{i,j} - \bar{Z}_{i,j}^{(2)}) \bar{A}^{(i,j)}$	99,7
$\beta P_{i,j}^{(1)} \bar{A}^{(i,j)}$	99,7
$\beta P_{i,j}^{(2)} \bar{A}^{(i,j)}$	99,8
$\beta P_{i,j}^{(1)} \cos(O_{i,j}^{(1)} - \omega_{i,j}^{(700\text{hPa})}) \bar{A}^{(i,j)}$	100,0
$\beta P_{i,j}^{(1)} V_{i,j}^{(700\text{hPa})} \cos(O_{i,j}^{(1)} - \omega_{i,j}^{(700\text{hPa})}) \bar{A}^{(i,j)}$	100,0
$\beta T_{i,j}^{(1)} \bar{A}^{(i,j)}$	96,9
$\beta T_{i,j}^{(2)} \bar{A}^{(i,j)}$	95,5
$\beta T_{i,j}^{(3)} \bar{A}^{(i,j)}$	95,2
$\beta T_{i,j}^{(4)} \bar{A}^{(i,j)}$	95,5
$\beta T_{i,j}^{(5)} \bar{A}^{(i,j)}$	95,9
$\beta T_{i,j}^{(3)} V_{i,j}^{(700\text{hPa})} \cos(O_{i,j}^{(3)} - \omega_{i,j}^{(700\text{hPa})}) \bar{A}^{(i,j)}$	98,7
$\beta T_{i,j}^{(3)} \cos(O_{i,j}^{(3)} - \omega_{i,j}^{(700\text{hPa})}) \bar{A}^{(i,j)}$	98,7

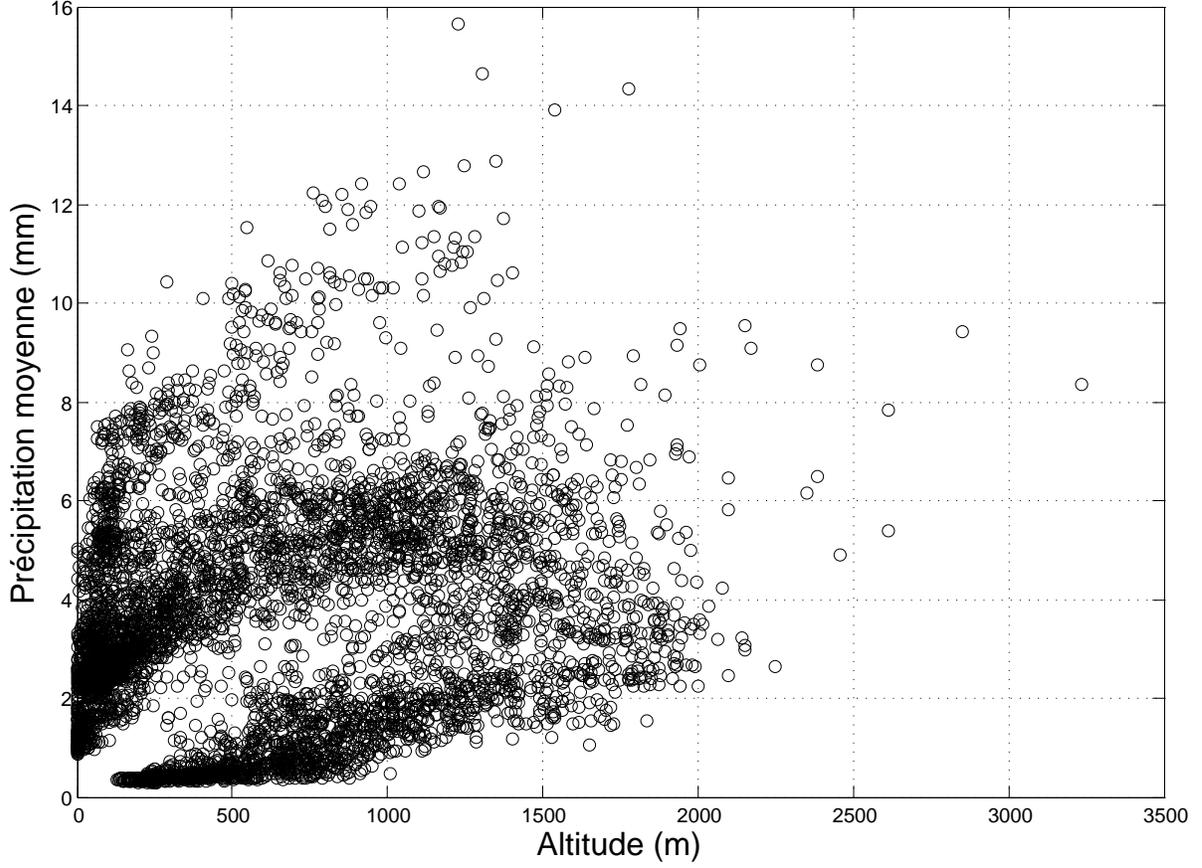


Figure 6.2 – Précipitation journalière moyenne durant la période 2002-2005 pour les  $72 \times 72$  pixels de 4,4 km en fonction de leur altitude.

à des améliorations très modestes par rapport à l'équation (3.2). En plus de l'altitude, certains auteurs (*e.g.* Bindlish et Barros, 2000 ; Badas *et al.*, 2006) mentionnent qu'il est important de tenir compte aussi de la pente et de son orientation. Afin d'estimer la pente sur un pixel, le plan qui représente le mieux la topographie dans le voisinage du pixel est calculé. Mathématiquement, posons  $L$  comme étant la distance, en nombre de pixels, dans chaque direction définissant la taille du voisinage d'un pixel. Ainsi, le pixel  $(i', j')$  est dans le voisinage du pixel  $(i, j) \Leftrightarrow \{|i' - i| \leq L \text{ et } |j' - j| \leq L\}$ . Une valeur d'altitude  $\hat{Z}$  est estimée pour chaque pixel  $(i', j')$  dans le voisinage du pixel  $(i, j)$  à partir de l'équation suivante :

$$\hat{Z}_{i',j'}^{(L)} = b_0^{(i,j)} + b_1^{(i,j)} (i' - i) r + b_2^{(i,j)} (j' - j) r \quad (6.2)$$

où  $r$  est la résolution (m) d'un pixel et les paramètres  $b_0^{(i,j)}$ ,  $b_1^{(i,j)}$  et  $b_2^{(i,j)}$  sont estimés de

sorte qu'ils minimisent la somme des  $(Z_{i',j'} - \hat{Z}_{i',j'})^2$ . Ainsi,  $b_1^{(i,j)}$  et  $b_2^{(i,j)}$  sont les variations d'altitude moyennes (m d'élévation / m horizontal) dans les directions nord et est de la grille respectivement et  $b_0^{(i,j)} = \hat{Z}_{i,j}^{(L)}$ , l'altitude estimée par le plan au pixel  $(i, j)$ . Trois variables topographiques peuvent ainsi être définies à partir des ces trois paramètres : la pente  $P_{i,j}^{(L)}$ , l'orientation  $O_{i,j}^{(L)}$  et l'anomalie topographique  $T_{i,j}^{(L)}$  (Figure 6.3) :

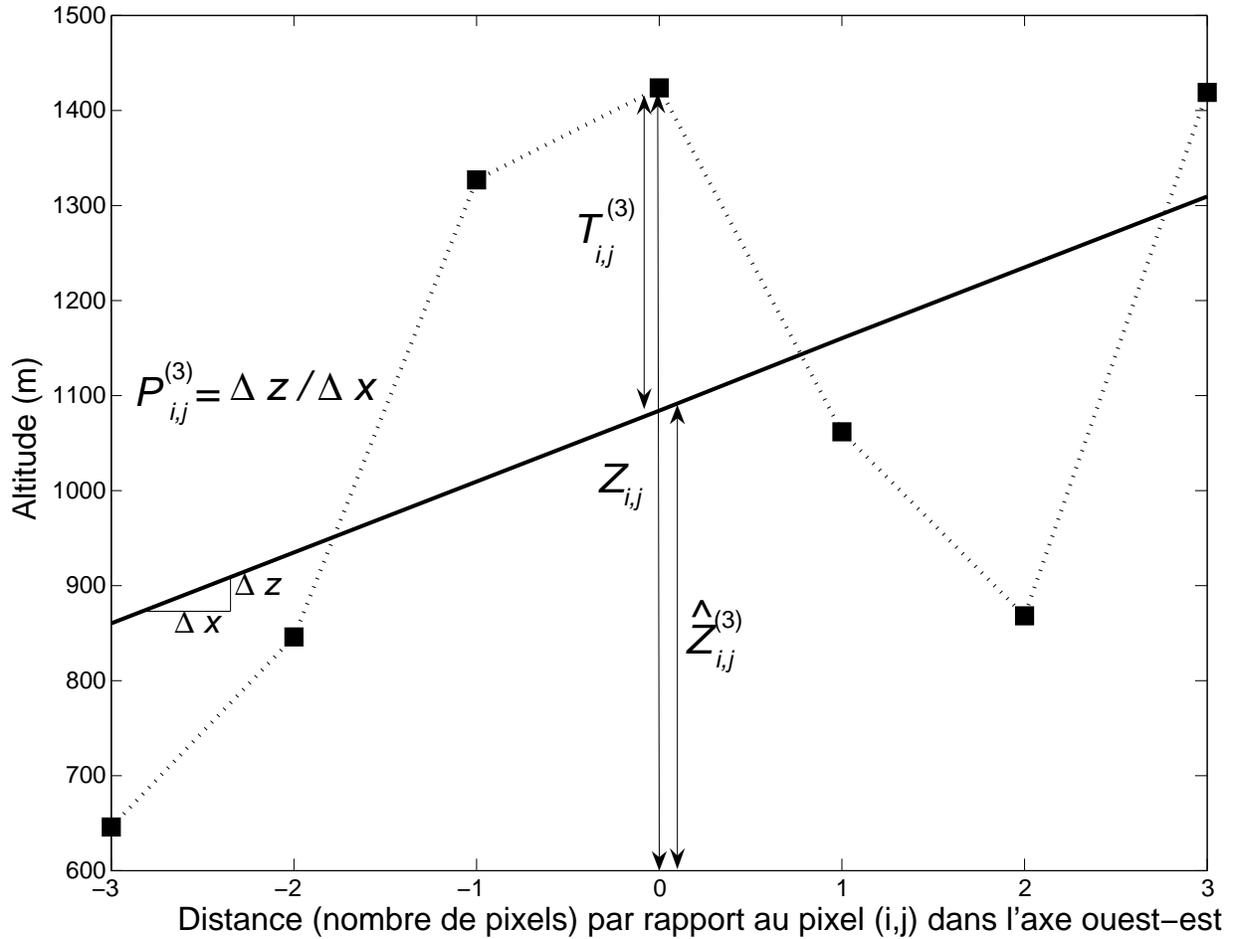


Figure 6.3 – Exemple synthétique en deux dimensions illustrant les trois variables topographiques : la pente  $P_{i,j}^{(L)}$ , l'orientation  $O_{i,j}^{(L)}$ , correspondant à la direction du gradient, et l'anomalie topographique  $T_{i,j}^{(L)}$  avec  $L = 3$ . Les carrés sont les altitudes aux pixels et la ligne continue est l'altitude estimée par le plan. Dans cet exemple, il n'y a pas de variation d'altitude dans l'axe nord-sud, ce qui implique que  $O_{i,j}^{(L)} = 0^\circ$ .

$$P_{i,j}^{(L)} = \sqrt{\left(b_1^{(i,j)}\right)^2 + \left(b_2^{(i,j)}\right)^2} ; \quad (6.3)$$

$$O_{i,j}^{(L)} = \begin{cases} \arctan\left(\frac{b_1^{(i,j)}}{b_2^{(i,j)}}\right) & \text{si } b_2^{(i,j)} \geq 0 \\ \arctan\left(\frac{b_1^{(i,j)}}{b_2^{(i,j)}}\right) + \pi & \text{si } b_2^{(i,j)} < 0 \text{ et } b_1^{(i,j)} \geq 0 \\ \arctan\left(\frac{b_1^{(i,j)}}{b_2^{(i,j)}}\right) - \pi & \text{si } b_2^{(i,j)} < 0 \text{ et } b_1^{(i,j)} < 0 \end{cases} ; \quad (6.4)$$

$$T_{i,j}^{(L)} = Z_{i,j} - b_0^{(i,j)} . \quad (6.5)$$

Différentes équations incluant la pente sont testées, mais peu importe la valeur de  $L$ , aucune n'apporte une amélioration significative (Tableau 6.2). Cela n'implique pas nécessairement que la pente n'a pas d'impact sur la distribution à fine résolution spatiale de la précipitation. Il est plutôt probable que l'impact de la pente soit déjà inclus dans les valeurs de précipitation des huit voisins les plus proches. La variable ayant le plus grand impact est l'anomalie topographique. Pour les pixels de 4,4 km, c'est lorsque  $L = 3$  pixels dans chaque direction pour définir le voisinage que l'impact de l'anomalie topographique est le plus important ; la somme du carré des écarts diminue de 4,8 % avec l'ajout du terme  $\beta T_{i,j}^{(3)} \bar{A}^{(i,j)}$ . Les ajouts de l'orientation de la pente, de la vitesse et de la direction du vent ne permettent pas d'améliorer le modèle (Tableau 6.2). L'équation de l'espérance du modèle tenant compte de la topographie est donc

$$\begin{aligned} \mathbb{E} \left[ R_{i,j} \mid \vec{R}_{-(i,j)} \right] &= \bar{A}^{(i,j)} + \beta_d \left( \frac{A_{|}^{(i,j)} + A_{-}^{(i,j)}}{2} - \frac{A_{/}^{(i,j)} + A_{\setminus}^{(i,j)}}{2} \right) \dots \\ &+ \beta_{\times} \left( A_{/}^{(i,j)} - A_{\setminus}^{(i,j)} \right) + \beta_{+} \left( A_{|}^{(i,j)} - A_{-}^{(i,j)} \right) + \beta_T T_{i,j}^{(L)} \bar{A}^{(i,j)} \end{aligned} \quad (6.6)$$

avec  $L = 3$  pour les pixels de 4,4 km. Le Tableau 6.3 présente les valeurs estimées des paramètres de l'équation (6.6) pour le modèle tenant compte de la topographie, ci-après appelé *Modèle 2*. Les valeurs estimées des paramètres de l'équation (3.2) pour le modèle ignorant la topographie, ci-après appelé *Modèle 1*, sont aussi présentées.

La valeur de  $\beta_T$  pour les pixels de 4,4 km est  $8,84 \times 10^{-5} \text{ m}^{-1}$ , ce qui implique que pour une anomalie topographique de 1705 m, soit l'anomalie maximale observée sur la région, l'augmentation espérée de la précipitation par rapport à la précipitation sur les huit voisins

Tableau 6.3 – Valeurs estimées des paramètres de l'équation (3.2) pour le Modèle 1 (ignorant la topographie) et de l'équation (6.6) pour le Modèle 2 (incluant la topographie) pour les différentes tailles de pixels (2002-2005).

Taille des pixels	Modèle 1			Modèle 2			
	$\beta_x$	$\beta_+$	$\beta_d$	$\beta_x$	$\beta_+$	$\beta_d$	$\beta_T^*$
4,4 km	0,09	0,08	0,67	0,07	0,09	0,58	$8,84 \times 10^{-5}$
8,8 km	-0,03	0,03	1,16	-0,03	0,04	0,92	$2,97 \times 10^{-4}$
13,2 km	-0,03	0,09	1,15	-0,03	0,12	0,90	$3,65 \times 10^{-4}$
17,6 km	-0,02	0,13	0,97	-0,02	0,17	0,72	$4,51 \times 10^{-4}$
26,4 km	0,07	0,10	1,19	0,05	0,15	0,88	$4,61 \times 10^{-4}$

\*Seul  $\beta_T$  a des unités ((mm d'augmentation)/(m d'anomalie

× mm de précipitation moyenne sur les huit voisins les plus proches))

les plus proches est de  $8,84 \times 10^{-5} \text{ m}^{-1} \times 1705 \text{ m} \times 100 \% = 15 \%$ . Seulement quatre pixels de 4,4 km ont des anomalies topographiques supérieures à 1131 m, menant à des augmentations espérées d'au moins 10 %. C'est pourquoi l'inclusion de la topographie dans l'équation de l'espérance n'a pas permis de diminuer de manière importante la somme du carré des écarts par rapport à l'équation (3.2).

Les valeurs des paramètres pour les Modèles 1 et 2 pour les autres tailles de pixels sont aussi présentées au Tableau 6.3. Il faut noter que pour les pixels plus grands que 4,4 km, l'anomalie topographique est définie à partir d'un voisinage de  $\pm L = 1$  pixel. Le tableau montre que les paramètres d'anisotropie  $\beta_x$  et  $\beta_+$  sont comparables pour les Modèles 1 et 2. Cependant, le Modèle 1 surestime le paramètre de distance  $\beta_d$ . Comme à la Partie I, il peut arriver que l'espérance calculée à partir des paramètres estimés donne une valeur très près de 0, mais négative. Dans ce cas, l'espérance est ramenée à 0,1 mm.

### 6.2.2 Écart-type de la précipitation sur un pixel

À la Partie I, l'écart-type de la précipitation sur le pixel  $(i, j)$ , noté  $\sigma_{R_{i,j}}$ , dépendait de  $\bar{R}_{i,j}$ , soit l'espérance sur le pixel, et de la valeur de la variable atmosphérique CAPE sur la tuile. Ici, comme pratiquement tous les événements sont stratiformes, la variable CAPE n'est pas considérée. L'équation de l'écart-type (3.7) retenue au Chapitre 3 devient donc :

$$\sigma_{R_{i,j}} = \theta_0 \bar{R}_{i,j}^{\theta_1} . \quad (6.7)$$

Cette équation tient compte de la topographie pour le Modèle 2 puisque l'espérance  $\bar{R}_{i,j}$  dépend de l'anomalie topographique. Avec seulement  $\bar{R}_{i,j}$  comme variable explicative de l'écart-type, les paramètres sont plus facilement estimables. Pour les Modèles 1 et 2 séparément, l'espérance à chaque données disponibles à 4,4 km (2002-2005) est calculée puis 54 classes d'espérance sont créées ; chacune d'elles contenant au moins 1500 données. L'écart-type observé est calculé pour chacune de ces classes. Ensuite, les valeurs des paramètres  $\theta_0$  et  $\theta_1$  qui minimisent la somme du carré des écarts entre les écarts-types observés et estimés selon l'équation (6.7) sont choisies. Pour les deux modèles, la valeur estimée de  $\theta_1$  sur les données disponibles à 4,4 km est 1,02. Comme elle est très près de 1, il a été décidé de poser  $\theta_1 = 1$  et d'estimer uniquement  $\theta_0$ , qui devient alors le coefficient de variation. L'équation de l'écart-type utilisée dans la présente région est donc :

$$\sigma_{R_{i,j}} = \theta_0 \bar{R}_{i,j} . \quad (6.8)$$

Les valeurs de  $\theta_0$  estimées sur les données disponibles à 4,4 km sont respectivement de 7,58 et de 7,63 % pour les Modèles 1 et 2. La très faible différence entre les deux estimations illustre le fait qu'il n'est pas nécessaire de tenir compte de la topographie dans l'estimation de l'écart-type. La Figure 6.4 compare les écarts-types observés et estimés à partir de l'équation (6.8) pour les 54 classes d'espérance sur les données disponibles à 4,4 km. Les valeurs présentées proviennent du Modèle 1 ; les valeurs du Modèle 2 sont quasi-identiques. La figure montre que l'équation (6.8) s'ajuste très bien aux données. Le Tableau 6.4 présente les valeurs estimées du coefficient de variation  $\theta_0$  communes aux deux modèles pour les différentes tailles de pixels.

Les paramètres essentiels à la désagrégation sont maintenant estimés. L'algorithme utilisé pour la désagrégation est identique à celui décrit à la Section 3.3, à l'exception que pour l'espérance, il faut maintenant se référer aux équations (3.2) et (6.6) pour les Modèles 1 et 2 respectivement, et que pour l'écart-type, il faut se référer à l'équation (6.8) pour les deux modèles.

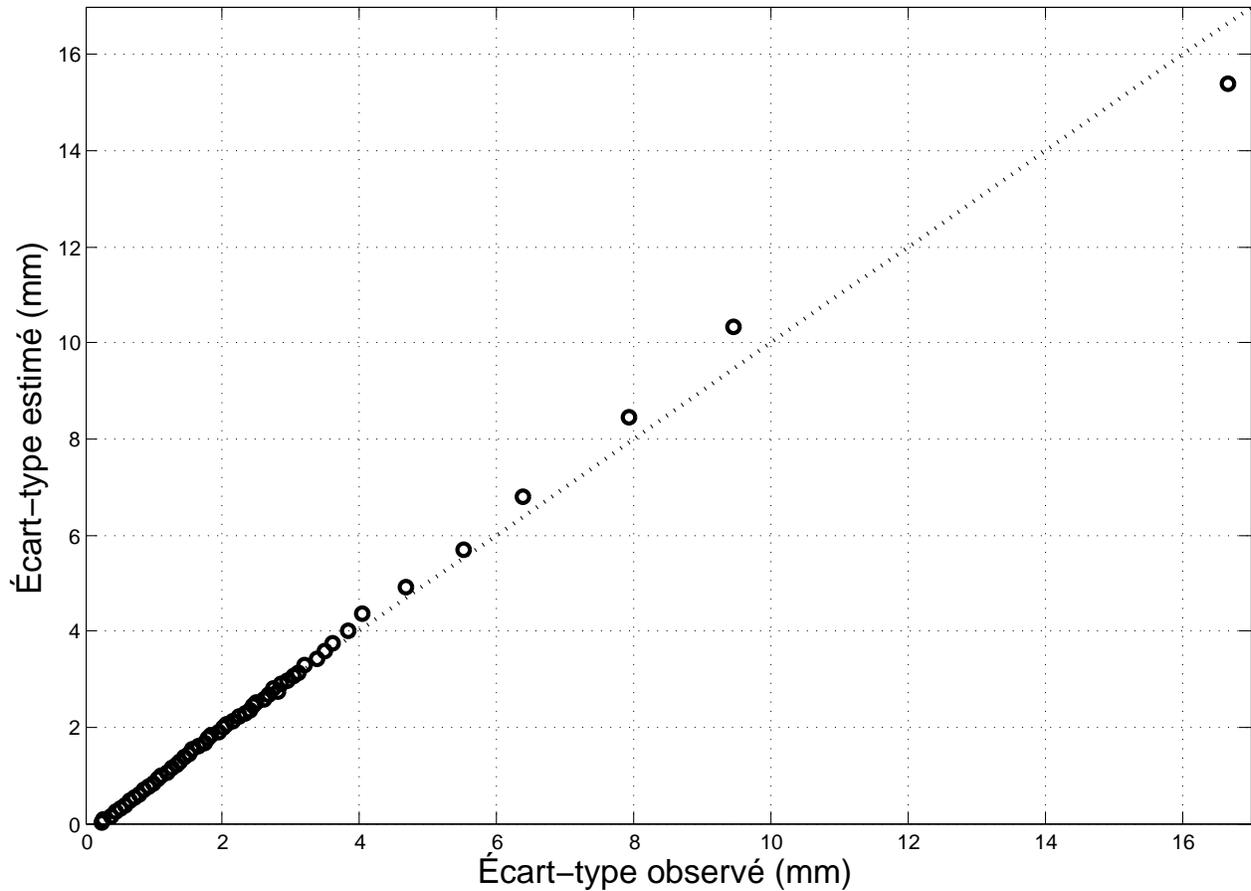


Figure 6.4 – Écart-types observés et estimés à partir de l’équation (6.8) pour les 54 classes d’espérance du Modèle 1 sur les données disponibles à 4,4 km. Pour le Modèle 2, les valeurs sont quasi-identiques.

### 6.3 Discussion

Le modèle développé au Chapitre 3 a été adapté pour tenir compte de la topographie à partir de seulement un paramètre et une variable supplémentaires (Modèle 2, équation (6.6)). L’anomalie topographique explique davantage les variations locales de la précipitation que les différences d’altitude ou que la pente. En l’absence de variation d’altitude, le Modèle 2 est équivalent au modèle développé au Chapitre 3, ce qui respecte la condition posée au départ, à l’exception que l’anisotropie n’est ici pas expliquée par le vent et que la variable CAPE n’est pas considérée.

Aussi, il est intéressant de noter que pour une région où la topographie est lisse, c’est-à-dire une région où il peut y avoir une pente prononcée, mais sans anomalie locale, le Modèle 2 ne considère pas la topographie. Dans un tel cas, l’espérance de la précipitation sur un pixel

Tableau 6.4 – Valeurs estimées du coefficient de variation  $\theta_0$  (équation (6.8)) communes aux Modèles 1 et 2 pour les différentes tailles de pixels (2002-2005).

Taille des pixels	$\theta_0$ (%)
4,4 km	7,6
8,8 km	8,8
13,2 km	10,0
17,6 km	12,9
26,4 km	14,5

n'est pas augmentée ou diminuée par rapport à la précipitation moyenne sur les voisins les plus proches. En fait, il faut une grande valeur d'anomalie topographique pour que l'impact soit substantiel. Pour les pixels de 4,4 km, l'espérance est augmentée de seulement 8,84 % si l'anomalie topographique du pixel est de 1000 m. Pour les autres tailles de pixels, les valeurs estimées du paramètre  $\beta_T$  sont plus élevées (Tableau 6.3), mais ne mènent pas nécessairement à de plus grandes variations de précipitation puisque les anomalies topographiques sont de plus en plus petites lorsque la taille des pixels augmente. Ces résultats suggèrent qu'à moins d'avoir de très grandes variations locales d'altitude, la topographie a peu d'impact sur la distribution spatiale de la précipitation à fine résolution. Cela corrobore les résultats de Deidda *et al.* (2006). Étudiant une région où les variations d'altitude sont inférieures à 600 m, ils ont constaté que les hétérogénéités spatiales de la précipitation à une résolution de 4 km étaient négligeables.

La taille du voisinage définissant l'anomalie topographique a été choisie en fonction de la performance de l'équation (6.6). Elle est de  $7 \times 7$  pixels (environ  $30 \times 30$  km<sup>2</sup>) pour les pixels de 4,4 km et du minimum possible de  $3 \times 3$  pixels pour les autres tailles de pixels. La taille du voisinage peut dépendre notamment de la taille des pixels, de la topographie sur la région étudiée et du pas de temps analysé. Il est difficile, à partir d'un seul jeu de données, d'expliquer pourquoi ce sont ces tailles qui ont donné de meilleurs résultats.

Concernant l'écart-type, un seul paramètre a été conservé en supposant que le coefficient de variation est constant. À la Partie I, il y avait deux paramètres de plus : un exposant pour l'espérance et un paramètre pour CAPE (équation (3.7)). Il faut noter que si uniquement les événements stratiformes avaient été conservés à la Partie I, le coefficient de variation aurait

pu être supposé constant (voir Figure 3.5b pour CAPE près de 0).

La vitesse et la direction du vent n'ont pas permis d'expliquer ni l'anisotropie, ni l'effet de la topographie sur la précipitation à fine résolution. Ces deux variables ont pourtant été retenues au Chapitre 3 pour expliquer l'anisotropie sur une région avec peu de relief. Étant modifié par la topographie, il est probable que le vent à fine échelle soit complètement différent du vent à mésoéchelle utilisé dans ce chapitre. Il est aussi probable que la méthode d'interpolation des observations utilisée par le NWRFC ne permette pas de bien représenter l'impact du vent. Il serait intéressant d'analyser un jeu de données à fine résolution spatiale n'ayant pas subi d'interpolation, mais c'est présentement impossible en travaillant sur des données observées en hautes montagnes.



# Chapitre 7

## RÉSULTATS DE LA DÉSAGRÉGATION

Ce chapitre présente les résultats de la désagrégation à 4,4 km des  $6 \times 6$  tuiles pour les 750 jours ayant reçu de la précipitation (2006-2008) à partir du Modèle 1, qui ne tient pas compte de la topographie, et du Modèle 2, qui tient compte de l'anomalie topographique (Chapitre 6). L'algorithme décrit à la Section 3.3 est utilisé avec  $N_t = 6$ , sauf que l'espérance est définie par les équations (3.2) et (6.6) pour les Modèles 1 et 2 respectivement et que l'écart-type est définie par l'équation (6.8) pour les deux modèles. Huit champs désagrégés sont retenus pour chaque jour, soit après 100, 150,  $\dots$ , 450 itérations de l'algorithme. Les résultats de la désagrégation sur les pixels de 8,8, 13,2, 17,6 et 26,4 km sont présentés à l'Annexe E.

### 7.1 Structure spatiale

La Figure 7.1 présente la précipitation moyenne observée et désagrégée sur les 750 jours ayant reçu de la précipitation (2006-2008). Un premier constat intéressant est que même le Modèle 1, qui ne tient pas compte de la topographie, est capable de reproduire l'impact mésoéchelle des montagnes. Cependant, il n'est pas capable de reproduire les variations à l'échelle locale, ce que permet le Modèle 2, notamment pour le mont Rainier (Figure 6.1), mais aussi pour d'autres variations moins importantes dans la chaîne des Cascades. Cependant, aucun des deux modèles ne reproduit adéquatement la structure spatiale de la précipitation moyenne dans les montagnes Olympiques. Il faut dire que ces montagnes sont situées aux frontières de la région étudiée et qu'en pratique, il serait préférable de ne pas

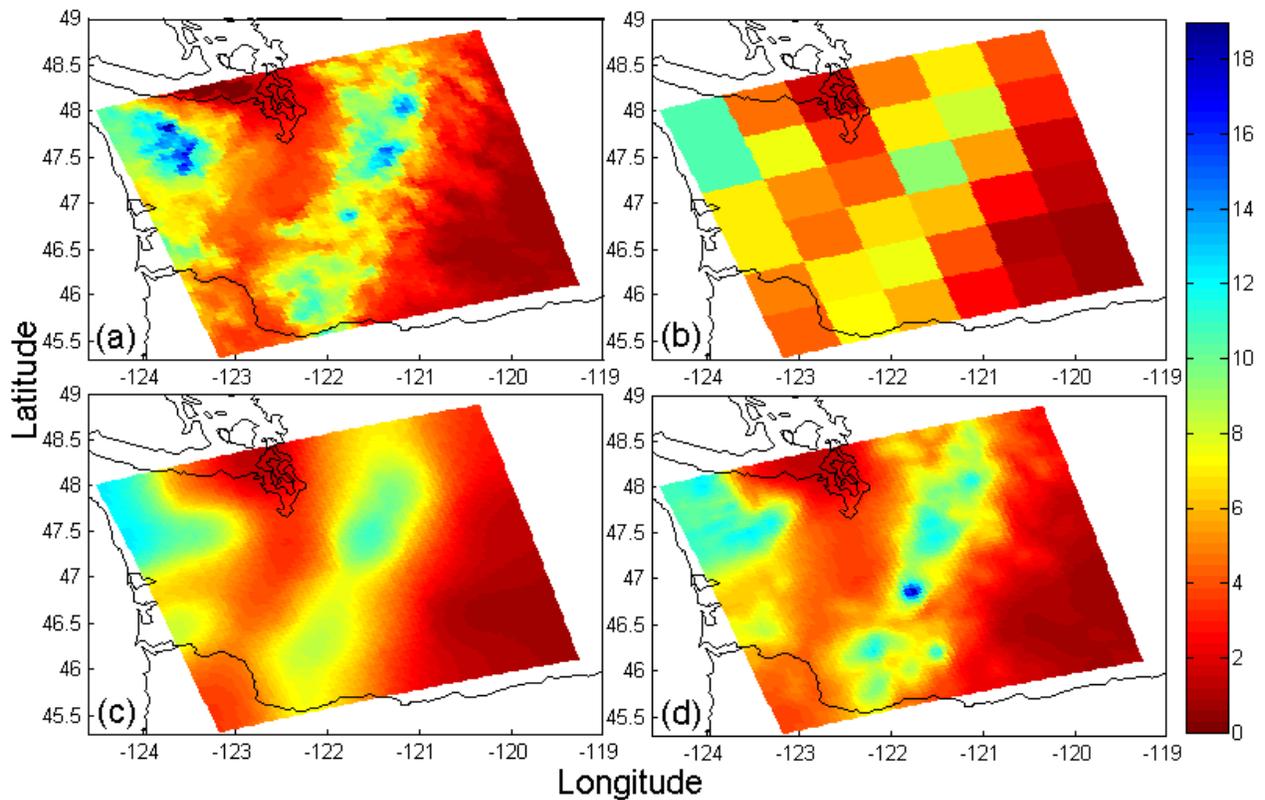


Figure 7.1 – Précipitation moyenne (mm/jour) sur les 750 jours (2006-2008) (a) observée à 4,4 km, (b) observée sur les tuiles de 52,8 km et désagrégée à 4,4 km par les Modèles (c) 1 et (d) 2. La précipitation moyenne désagrégée sur un pixel est la moyenne des valeurs produites par les huit réalisations.

analyser les tuiles aux frontières (Section 4.1).

Pour cette région, il n'est pas approprié de présenter la corrélation moyenne en fonction de la distance comme cela a été fait au Chapitre 4 (Figures 4.3, 4.4 et 4.5) puisqu'en plus de la distance, la topographie influence beaucoup la corrélation spatiale. Une autre façon d'illustrer la structure spatiale de la précipitation est de présenter l'écart-type inter-journalier pour chaque pixel (Figure 7.2). L'écart-type inter-journalier est quasi-identique, à une constante près, à la précipitation moyenne journalière pour chaque pixel (Figure 7.1), même pour la précipitation observée. Cela suggère que le coefficient de variation de la précipitation est constant, ce qui justifie le choix de l'équation (6.8) pour représenter l'écart-type des deux modèles.

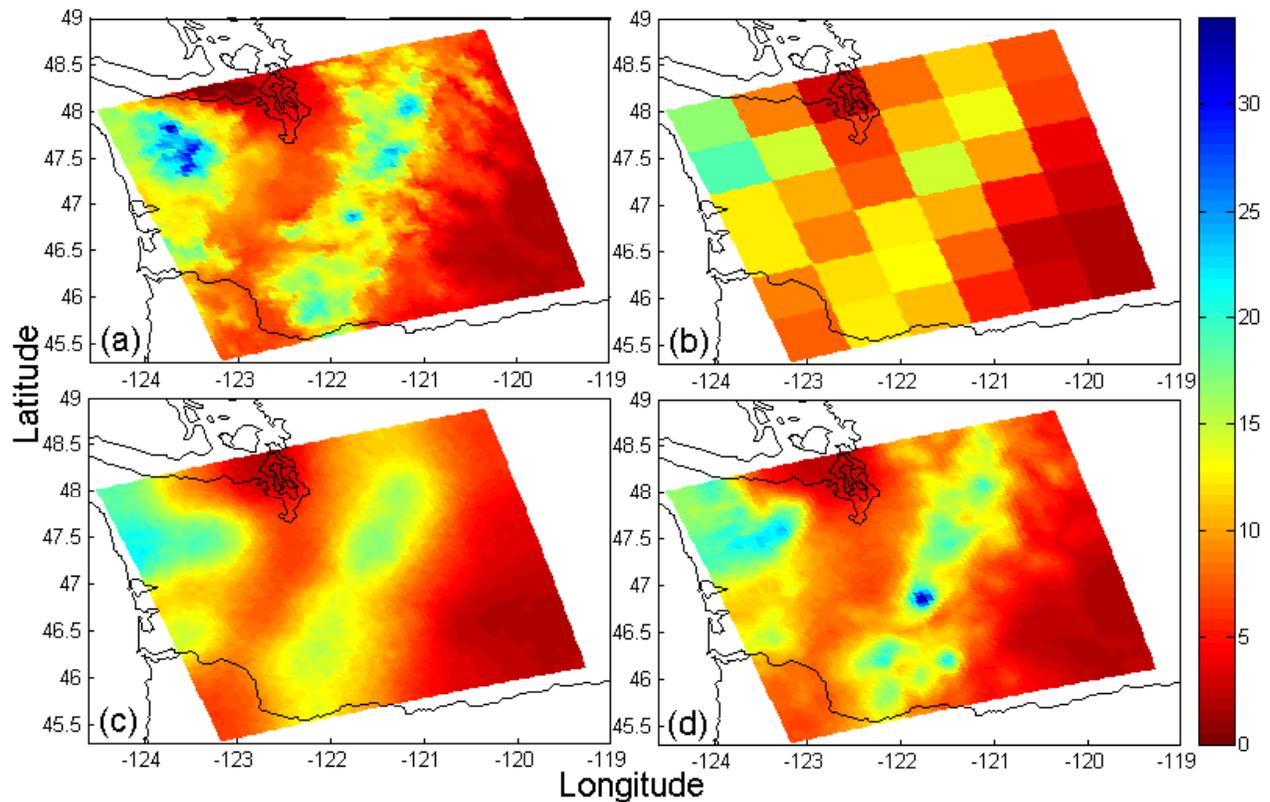


Figure 7.2 – Écart-type inter-journalier (mm) sur les 750 jours (2006-2008) (a) observé à 4,4 km, (b) observé sur les tuiles de 52,8 km et désagrégé à 4,4 km par les Modèles (c) 1 et (d) 2.

## 7.2 Valeurs extrêmes

La Figure 7.3 présente les valeurs de précipitation maximales observées et désagrégées sur chaque pixel durant la période 2006-2008. La valeur maximale désagrégée sur chaque pixel est la moyenne des plus grandes valeurs obtenues des huit réalisations. Les constats sont passablement les mêmes que pour les Figures 7.1 et 7.2. Le Modèle 1 est capable de reproduire l'impact mésoéchelle de la topographie sur les valeurs maximales, mais n'est pas capable de reproduire les variations locales comme le fait le Modèle 2. Même si le Modèle 2 manque certaines variations locales, il produit des valeurs maximales plus élevées dans les montagnes Olympiques que le Modèle 1 ; ce qu'il n'arrive pas à faire pour la moyenne journalière et l'écart-type inter-journalier.

Afin de voir si les modèles sont capables de reproduire la plus grande valeur de précipitation à l'échelle événementielle, la Figure 7.4 présente les plus grandes valeurs observées et désagrégées sur les pixels de 4,4 km à chaque jour. À noter que la valeur maximale des  $4 \times 4$

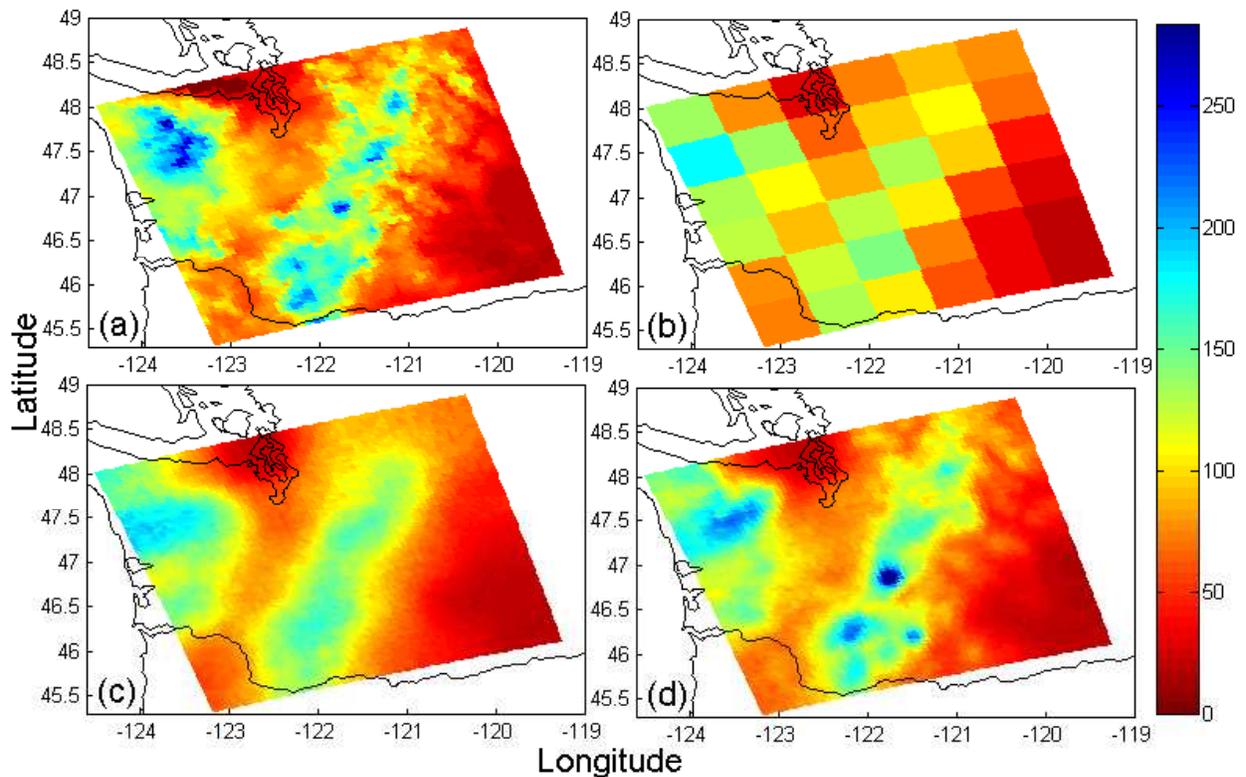


Figure 7.3 – Précipitation journalière maximale (mm) sur les 750 jours (2006-2008) (a) observée à 4,4 km, (b) observée sur les tuiles de 52,8 km et désagrégée à 4,4 km par les Modèles (c) 1 et (d) 2. La précipitation maximale désagrégée sur un pixel est la moyenne des valeurs maximales produites par les huit réalisations.

tuiles centrales est présentée ; les tuiles aux frontières forment la zone tampon (Section 4.1). La valeur maximale observée sur les tuiles de 52,8 km est en moyenne égale à environ la moitié de la valeur maximale observée sur les pixels de 4,4 km. La relation entre les valeurs maximales observées sur les tuiles et sur les pixels est comparable à celle obtenue dans la Partie I sur une région avec peu de relief pour les événements stratiformes (Figure 4.9a). Le Modèle 1 sous-estime d'environ 25 % la précipitation maximale observée (Figure 7.4a). Le Modèle 2 reproduit la valeur maximale observée avec une plus grande précision (Figure 7.4b).

La distribution des valeurs extrêmes de précipitation peut varier beaucoup d'un pixel à l'autre en régions montagneuses. La Figure 7.5 présente la queue de la distribution des valeurs journalières des précipitations observées et désagrégées pour six pixels de 4,4 km sélectionnés parmi les  $4 \times 4$  tuiles centrales : (a) pixel avec la plus haute altitude et aussi la plus grande anomalie topographique, (b) pixel voisin au nord de (a), (c) pixel avec la plus petite anomalie topographique, (d) pixel voisin au nord de (c), (e) pixel avec la plus petite

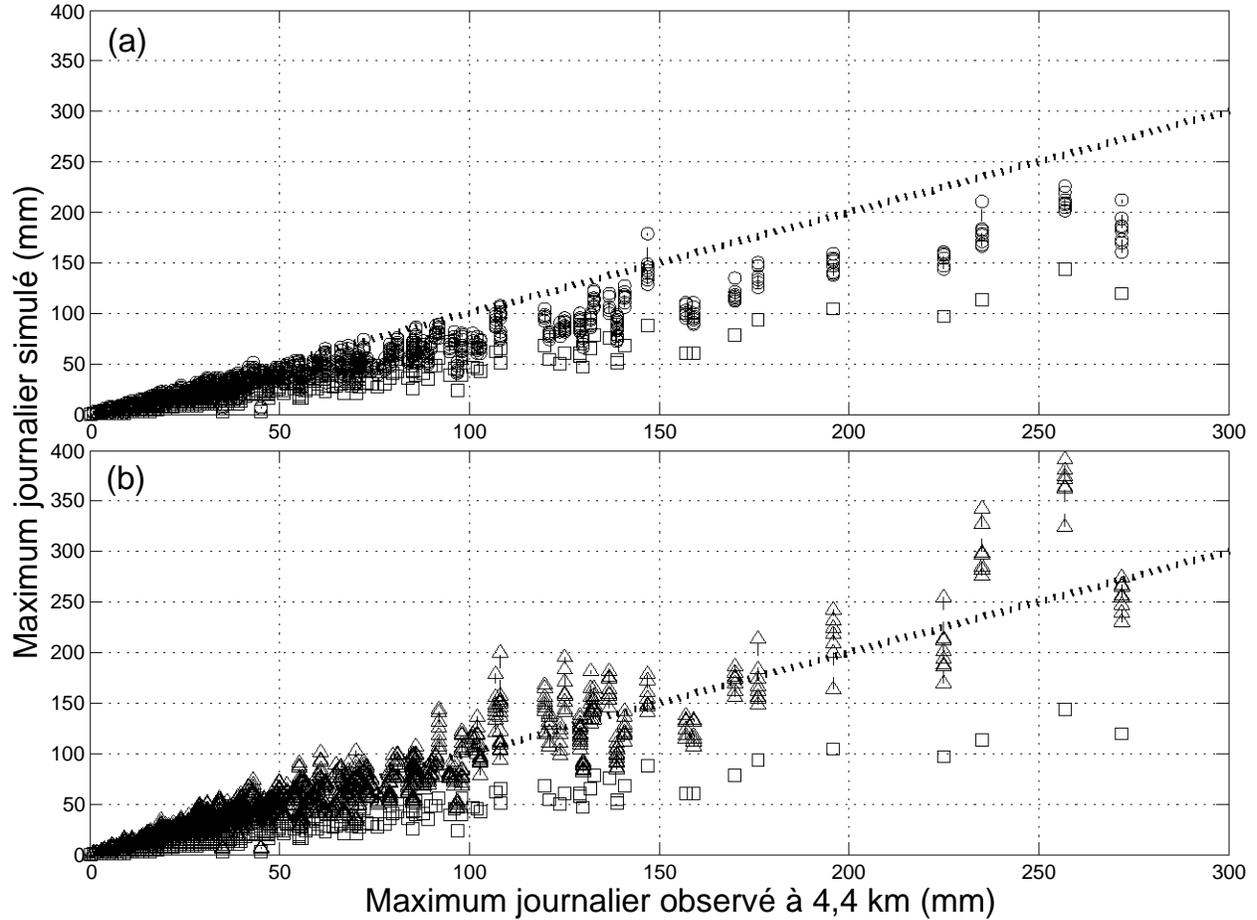


Figure 7.4 – Précipitations maximales observée et produites par les Modèles (a) 1 (cercles) et (b) 2 (triangles) sur les pixels de 4,4 km des  $4 \times 4$  tuiles centrales pour chaque jour. Les carrés sont les valeurs maximales sur les tuiles de 52,8 km utilisées en entrée dans le modèle.

altitude (excluant ceux situés au-dessus de l'eau) et (f) pixel avec l'anomalie topographique le plus près de 0.

Cette figure illustre la grande hétérogénéité spatiale dans la région : pour les six pixels sélectionnés, la plus grande valeur journalière observée (étoile noire) varie d'environ 40 à 260 mm. Cela contraste avec les queues de distribution examinées pour les pixels sur une région avec peu de relief (Figure 4.8). Il est difficile pour un modèle de désagrégation de reproduire avec exactitude ces variations locales. La Figure 7.5a montre que le Modèle 1 (lignes pointillées bleues) sous-estime la distribution des grandes valeurs pour le pixel à plus haute altitude (3236 m), qui est aussi le pixel avec la plus grande anomalie topographique (1705 m). En fait, il retourne à peu près les mêmes valeurs que celles de la tuile (ligne continue rouge) utilisée en entrée. Le Modèle 2 tient compte de l'anomalie topographique,

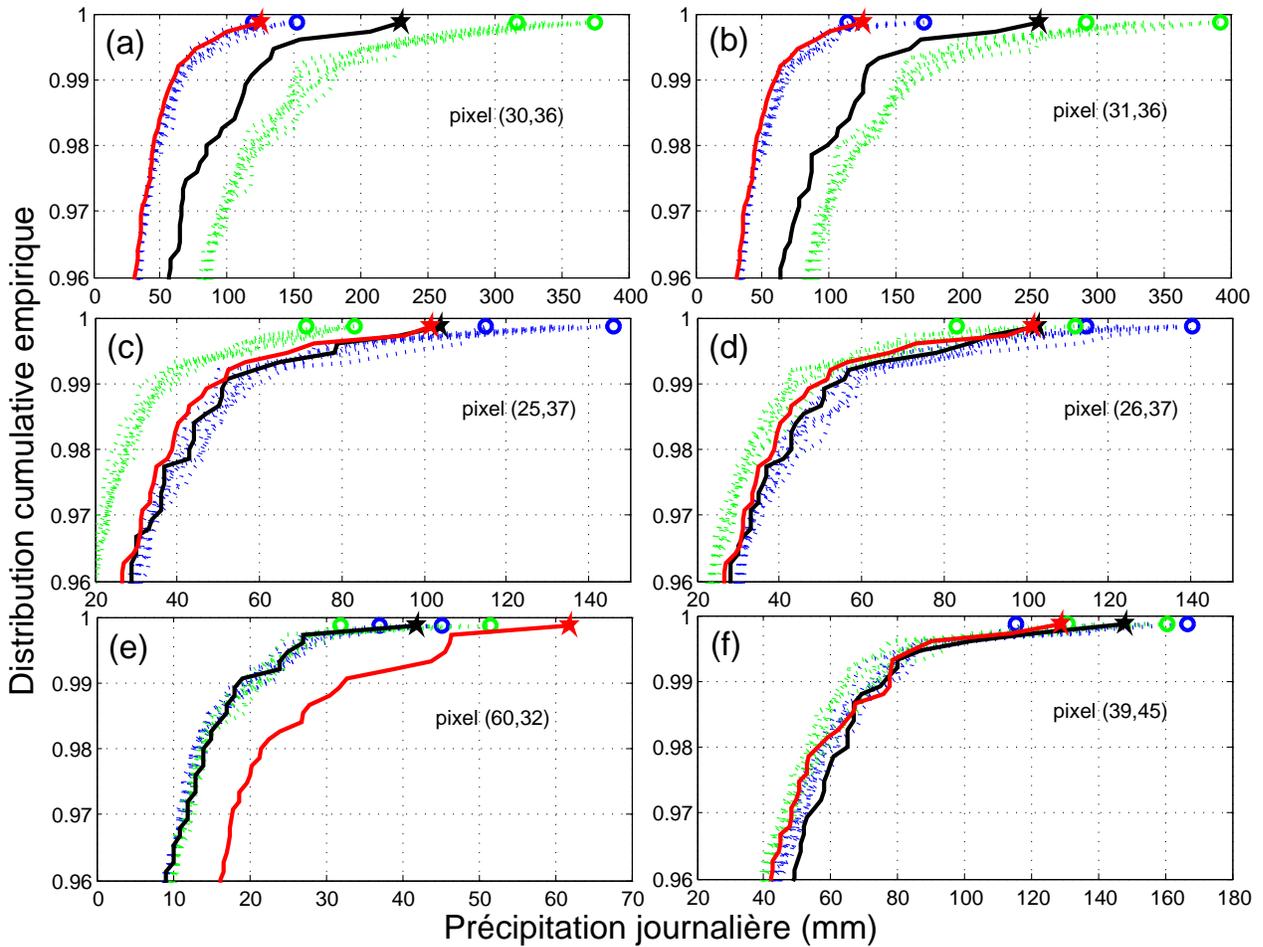


Figure 7.5 – Queues des distributions empiriques de la précipitation journalière observée à 4,4 km (ligne continue noire), observée à 52,8 km (ligne continue rouge) et désagrégée à 4,4 km par les Modèles 1 (lignes pointillées bleues) et 2 (lignes pointillées vertes) pour six pixels sélectionnés : (a) pixel avec la plus grande altitude et anomalie topographique, (b) pixel voisin au nord de (a), (c) pixel avec la plus petite anomalie topographique, (d) pixel voisin au nord de (c), (e) pixel avec la plus petite altitude et (f) pixel avec une anomalie topographique près de 0. Les étoiles indiquent la précipitation journalière maximale observée et les cercles la plus petite et la plus grande des huit valeurs journalières maximales désagrégées durant la période (2006-2008).

mais surestime son impact pour ce pixel (lignes pointillées vertes). Un examen des valeurs journalières de précipitation a montré que le pixel voisin immédiatement au nord reçoit plus de précipitation (Figure 7.5b), même si son anomalie topographique est inférieure (1322 m). Pour ce pixel, le Modèle 2 surestime encore les valeurs de précipitation, mais dans une plus faible mesure. Le Modèle 1 sous-estime encore plus les valeurs de précipitation.

Le Modèle 2 surestime aussi l'impact de l'anomalie topographique pour le pixel ayant la plus petite valeur d'anomalie (-659 m; Figure 7.5c). Pour ce pixel, la queue de la distribution de la précipitation observée à 4,4 km est comparable à celle observée sur la tuile de 52,8 km. Le Modèle 1 reproduit bien les valeurs observées, sauf pour la valeur maximale durant la période 2006-2008. La Figure 7.5d montre que le Modèle 2 performe beaucoup mieux sur le pixel voisin immédiatement au nord, présentant une anomalie topographique plus élevée (-263 m).

Pour le pixel avec la plus faible altitude (Figure 7.5e), les deux modèles reproduisent très bien la queue de la distribution observée. La précipitation sur la tuile surestime la précipitation à l'échelle du pixel. Pour le pixel avec une anomalie topographique très près de 0, les deux modèles ainsi que la précipitation sur la tuile sont comparables à la précipitation observée à 4,4 km (Figure 7.5f).

### 7.3 Discussion

Il est intéressant de constater que le Modèle 1, basé sur des résultats de la Partie I et qui ne tient pas compte de la topographie, peut représenter l'impact mésoéchelle de la topographie. Cela s'explique par le fait que la précipitation sur les tuiles utilisées en entrée est influencée par la topographie (Figures 7.1b, 7.2b et 7.3b). Comme le Modèle 1 tient compte des pixels voisins et indirectement des tuiles voisines, il peut produire des champs qui reproduisent les propriétés statistiques de la précipitation sur la majorité des pixels des deux chaînes de montagnes (Figures 7.1c, 7.2c et 7.3c).

Le Modèle 1 ne peut cependant pas représenter les variations locales de la précipitation causées par la topographie. Le Modèle 2, qui est basé sur le Modèle 1, mais adapté pour tenir compte de l'anomalie topographique, est capable de reproduire des variations locales (Figures 7.1d, 7.2d et 7.3d). Cela fait en sorte qu'il est plus apte à représenter les valeurs maximales à l'échelle journalière (Figure 7.4). Cependant, il y a quelques variations locales que le Modèle 2 n'est pas capable de reproduire, notamment celles dans les montagnes Olympiques, qui sont situées en partie aux frontières de la région étudiée.

Les variations locales pouvant être très prononcées, la performance des modèles peut varier significativement d'un pixel à l'autre (Figure 7.5). Par exemple, les pointes de précipitation causées par le mont Rainier (Figure 6.1) sont produites directement sur le pixel couvrant le mont par le Modèle 2 (Figure 7.5a), alors qu'elles sont généralement observées sur le pixel immédiatement au nord (Figure 7.5b). Le Modèle 2 est trop simple pour représenter les pointes de précipitation avec précision sur les pixels de 4,4 km. Pour leur MPO, Smith et Barstad (2004) tiennent compte du temps nécessaire pour transformer la vapeur d'eau en hydrométéores et du temps de chute de l'hydrométéore au sol. Ces valeurs sont de l'ordre de quelques minutes jusqu'à environ une demi-heure. Dépendamment de la vitesse du vent, il peut y avoir quelques kilomètres de distance entre l'endroit où la vapeur d'eau se condense et où l'hydrométéore tombe.

En somme, les résultats montrent que la structure du modèle de base développé à la Partie I peut tenir compte adéquatement de la topographie si les variations ne sont pas trop brusques. Il faut noter que cela est vrai seulement si la précipitation mésoéchelle utilisée en entrée est capable de bien représenter la topographie à mésoéchelle. Dans le cas présent, la précipitation utilisée en entrée provient d'observations interpolées, mais pour les applications, elle pourrait provenir de modèles météorologiques ou climatiques. Le Modèle 2 permet de représenter la plupart des variations locales, mais n'arrive pas à toutes les représenter. Des analyses sur d'autres régions montagneuses ayant des données de précipitation à fine résolution spatiale pourraient permettre de mieux comprendre la précipitation orographique à fine échelle.

Nonobstant les imperfections du Modèle 2, seules les anomalies topographiques de plusieurs centaines de mètres ont nécessité d'être représentées dans la région étudiée. L'anomalie topographique maximale sur un pixel est de 1705 m. Ce pixel contient le mont Rainier qui s'élève à 4392 m. À titre de comparaison, le plus haut sommet québécois au sud du 58<sup>e</sup> parallèle est le mont Jacques-Cartier, en Gaspésie, qui s'élève à 1268 m. Les résultats obtenus suggèrent que les endroits sur le territoire québécois où il vaudrait la peine de tenir compte de l'anomalie topographique sont très localisés et très peu nombreux. Cependant, considérant les incertitudes associées aux données de précipitation utilisées dans le nord-ouest des États-Unis, une analyse de la précipitation à fine résolution serait nécessaire pour confirmer que le Modèle 1 puisse être utilisé sur l'ensemble du Québec méridional. Il n'existe cependant pas de données observées à fine résolution couvrant l'ensemble du Québec méridional.

## Troisième partie

# IMPACT DE LA DÉSAGRÉGATION STATISTIQUE DE LA PRÉCIPITATION EN MODÉLISATION HYDROLOGIQUE



# Chapitre 8

## REVUE DE LITTÉRATURE

Les méthodes de mise à l'échelle statistique, telles la désagrégation ou les approches *grille vers points* (voir Chapitre 1), sont généralement développées pour fournir des champs de précipitation à une résolution spatiale suffisante pour la modélisation hydrologique tant dans un contexte de gestion opérationnelle que de planification stratégique. En pratique, il sera intéressant d'appliquer la désagrégation seulement si la précipitation désagrégée apporte de l'information supplémentaire par rapport à la précipitation agrégée. Dans ce chapitre, une revue d'analyses récentes sur l'impact de la résolution spatiale de la précipitation en modélisation hydrologique est présentée (Section 8.1) suivie d'une revue de modèles de désagrégation statistique de la précipitation qui ont été utilisés en modélisation hydrologique (Section 8.2). Une discussion conclut ce chapitre (Section 8.3).

### 8.1 Impact de la résolution spatiale de la précipitation en modélisation hydrologique

Bien qu'il soit reconnu que la résolution spatiale de la précipitation utilisée en entrée dans un modèle hydrologique influe sur l'écoulement simulé, le lien entre ces deux variables n'est pas totalement élucidé (Collier, 2009), ce qui mène parfois à des résultats contradictoires dans la littérature (Segond *et al.*, 2007). La relation entre la résolution de la précipitation et l'écoulement est difficile à évaluer puisqu'elle dépend des interactions complexes entre le type d'événement (convectif ou stratiforme), les caractéristiques (*e.g.* type de sols, occupation du sol) et la superficie du bassin versant (Segond *et al.*, 2007). Ces variables sont différentes d'une étude à une autre. C'est pourquoi il est très fréquemment mentionné dans les analyses de sensibilité en modélisation hydrologique que les résultats sont spécifiques aux conditions

expérimentales étudiées.

La majorité des études récentes sur la sensibilité de la modélisation hydrologique à la résolution spatiale de la précipitation utilisent des modèles hydrologiques distribués, qui sont capables de représenter l'hétérogénéité spatiale à l'intérieur d'un bassin versant. Il faut aussi considérer que chaque modèle hydrologique a sa propre représentation des processus physiques en jeu. Les résultats varient donc d'un modèle hydrologique à un autre (Arnaud *et al.*, 2002). Pour un modèle hydrologique donné, les résultats varient aussi selon la taille des unités hydrologiques définies par l'utilisateur (Bell et Moore, 2000; Indarto, 2002), les paramètres estimés (Arnaud *et al.*, 2002; Indarto, 2002; Liang *et al.*, 2004; Starks et Moriasi, 2009) et dépend aussi du rapport entre le pas de temps de simulation et le temps de réponse du bassin (Haan *et al.*, 1993).

La source de précipitation utilisée en entrée influe également sur les conclusions qui peuvent être tirées d'une analyse de sensibilité. Dans plusieurs travaux, seule la précipitation de stations météorologiques est utilisée pour analyser l'impact de la résolution spatiale de la précipitation en modélisation hydrologique, soit en faisant varier la densité du réseau de stations (Arnaud *et al.*, 2002), soit en gardant la même densité, mais en déplaçant les stations (Indarto, 2002). D'autres études utilisent des données de radars météorologiques (Bell et Moore, 2000; Liang *et al.*, 2004; Carpenter et Georgakakos, 2006; Morin *et al.*, 2006; Schuurmans et Bierkens, 2007; Segond *et al.*, 2007; Starks et Moriasi, 2009). Les radars météorologiques sont particulièrement intéressants pour l'analyse de la variabilité spatiale de la précipitation puisqu'ils offrent une couverture complète du bassin versant.

Malgré tous les facteurs pouvant masquer l'impact de la variabilité spatio-temporelle de la précipitation sur la réponse hydrologique, la précipitation demeure dans bon nombre de cas la source d'incertitude dominante (Ambroise, 1998; Indarto, 2002) et certains constats généraux émergent des analyses de sensibilité. Arnaud *et al.* (2002) mentionnent que la résolution spatiale de la précipitation affecte le volume total d'eau écoulee et dans une plus forte mesure le débit de pointe. Dans Segond *et al.* (2007), il est montré que l'impact sur le débit de pointe est deux fois plus important que l'impact sur le volume total d'eau de l'hydrogramme. Morin *et al.* (2006) ont montré, pour un événement convectif sur un petit bassin versant (148 km<sup>2</sup>), que le débit de pointe simulé peut doubler si l'angle de déplacement d'une cellule orageuse est légèrement modifié. La variabilité spatiale de la précipitation peut

aussi être importante pour les événements stratiformes. Schuurmans et Bierkens (2007) ont montré que sur un petit bassin versant plat (135 km<sup>2</sup>) recevant essentiellement de la pluie stratiforme, l'écart sur la précipitation moyenne pour une période de huit mois entre deux pixels d'un radar météorologique pouvait atteindre jusqu'à 20 %. Une variation de 20 % peut sembler faible, mais il faut noter qu'elle est calculée à partir d'une moyenne sur huit mois. De plus grandes variations sont à prévoir à l'échelle journalière. Cela démontre que la précipitation stratiforme, bien que plus structurée que la précipitation convective, n'est pas distribuée uniformément sur un territoire donné.

La résolution spatiale de la précipitation affecte aussi l'estimation des paramètres du modèle hydrologique (Bell et Moore, 2000). Arnaud *et al.* (2002) ont abaissé le seuil de ruissellement lorsque la précipitation en entrée était uniforme sur le bassin versant (1500 km<sup>2</sup>) pour palier le manque d'intensité locale de la précipitation résultant de l'agrégation spatiale. Dans Starks et Moriasi (2009), le paramètre gouvernant la recharge de l'aquifère s'est montré très sensible à la source de précipitation (trois réseaux de stations météorologiques ayant des densités différentes et une grille de radar météorologique) lors du calage du modèle hydrologique.

L'impact de la résolution spatiale de la précipitation varie en fonction de la taille du bassin versant. Arnaud *et al.* (2002) ont montré que la différence moyenne d'écoulement entre une simulation réalisée à partir de la précipitation à fine résolution et une simulation réalisée à partir de la valeur de précipitation moyenne sur le bassin est plus grande pour les plus grands bassins versants. Cela s'explique par le fait que lorsqu'une seule valeur de précipitation est prise sur un bassin, la résolution spatiale de la précipitation diminue lorsque la taille du bassin augmente.

En plus de comparer les différences moyennes d'écoulement entre diverses sources de précipitation, il est intéressant d'évaluer, pour une résolution spatiale donnée, l'impact sur l'écoulement des incertitudes sur la distribution spatiale de la précipitation. Carpenter et Georgakakos (2006) ont généré plusieurs champs de précipitation à fine résolution à partir de données de radars météorologiques perturbées pour tenir compte de l'incertitude sur ces données. Ils obtiennent alors un ensemble d'écoulements simulés pour chaque événement, pour les deux bassins versants étudiés et leurs sous-bassins. Cet ensemble permet de définir une mesure relative de l'incertitude pour chaque événement avec la variable  $\frac{E_{90\%} - E_{10\%}}{E_{50\%}}$ , où

$E_{10\%}$ ,  $E_{50\%}$  et  $E_{90\%}$  sont respectivement les percentiles 10, 50 et 90 % de l'écoulement simulé. Carpenter et Georgakakos (2006) ont montré qu'en moyenne, cette mesure d'incertitude diminue lorsque la superficie drainée augmente. Selon toute vraisemblance, lorsque la superficie drainée est petite, les anomalies de la précipitation (précipitation générée - précipitation radar originale) sur les quelques pixels recouvrant le sous-bassin peuvent être soient toutes négatives ou toutes positives ; lorsque la superficie drainée est grande, il est peu probable que les anomalies soient toutes positives ou toutes négatives. La forte incertitude à l'échelle du pixel se retrouve donc lissée lorsque le sous-bassin est couvert par plusieurs pixels. Carpenter et Georgakakos (2006) ont de plus montré qu'une fonction linéaire à pente négative exprime adéquatement la relation entre  $\frac{E_{90\%}-E_{10\%}}{E_{50\%}}$  et le logarithme de la superficie drainée.

## 8.2 Désagrégation appliquée à la modélisation hydrologique

Les études présentées à la section précédente comparent souvent la précipitation moyenne sur le bassin avec une précipitation à plus fine résolution. En désagrégation, seule la précipitation à mésoéchelle est connue, qu'elle provienne d'un modèle météorologique ou climatique. Dans ce contexte, l'analyse de l'impact sur l'écoulement de la résolution spatiale de la précipitation se fait en comparant les résultats obtenus en utilisant la précipitation sur la grille mésoéchelle (au lieu de la précipitation moyenne sur le bassin) aux résultats obtenus à partir d'une précipitation à plus fine résolution.

La plupart des études présentant un modèle de désagrégation statistique de la précipitation mentionnent que le modèle développé est conçu pour être appliqué en modélisation hydrologique. Cependant, très peu d'études ont utilisé de fait un modèle de désagrégation dans un contexte d'application hydrologique. À titre de comparaison, les méthodes *grille vers points* ont été beaucoup plus employées en modélisation hydrologique (voir *e.g.* Bürger et Chen, 2005 ; Salathé, 2005 ; Boé *et al.*, 2007). C'est possiblement dû au fait que ces approches ramènent la précipitation à l'échelle des stations et qu'en modélisation hydrologique, la précipitation utilisée en entrée provient généralement de stations météorologiques. Au Québec, notons les travaux de Quilbé *et al.* (2008) qui ont testé différentes méthodes *grille vers points* avec le modèle hydrologique distribué HYDROTEL (Fortin *et al.*, 2001a,b ; Turcotte *et al.*, 2003) sur le bassin versant de la rivière Chaudière et ceux Caron *et al.*

(2008) qui ont évalué un générateur météorologique stochastique afin de générer les données d'entrée de précipitation pour le modèle hydrologique global HSAMI (Bisson et Roberge, 1983; Fortin, 2000) sur le bassin versant de la rivière Châteauguay. Les applications de la désagrégation en modélisation hydrologique sont beaucoup plus limitées et à notre connaissance, aucune d'entre elles n'a été réalisée au Québec.

Les quelques études ayant appliqué la désagrégation en modélisation hydrologique ont utilisé diverses sources de précipitation en entrée. Sharma *et al.* (2007) ont utilisé un modèle de désagrégation multifractal, incluant un facteur multiplicatif pour tenir compte des hétérogénéités spatiales causées par la topographie, pour ramener la précipitation de deux tuiles d'un MCG à résolution  $2,8^\circ$  ( $\approx 300$  km) vers une résolution d'environ 2 km. Ils ont utilisé la pluie désagrégée comme donnée d'entrée dans un modèle hydrologique pour simuler le débit sur un bassin thaïlandais de  $34\,000$  km<sup>2</sup>. Ils ont montré, à partir de résultats mensuels, que la désagrégation appliquée après une correction du biais du MCG permettait d'obtenir des débits simulés plus près des débits observés. Leur résultats suggèrent que l'erreur associée à la faible résolution spatiale est aussi importante que le biais de ce MCG et qu'il est aussi important de désagréger que de corriger le biais du modèle.

Forman *et al.* (2008) ont appliqué un modèle spatio-temporel multi-échelle pour désagréger la précipitation de satellites météorologiques en Oklahoma, États-Unis. Ils ont simulé les débits sur un bassin versant de  $808$  km<sup>2</sup> sur une période d'un an. La précipitation a été désagrégée d'une résolution spatio-temporelle de  $1^\circ$  (environ 100 km) et 24 heures vers une résolution d'environ 4 km et une heure. Le débit simulé à partir de la pluie désagrégée est plus réaliste que le débit simulé à partir de la pluie de satellites météorologiques originale. En moyenne, l'eau s'écoulant par excès d'infiltration était environ deux fois plus important à partir de la précipitation désagrégée. Forman *et al.* (2008) ont aussi estimé l'incertitude sur le débit causée par la variabilité spatiale de la précipitation en fonction de la superficie drainée. Comme Carpenter et Georgakakos (2006), ils ont obtenu une relation négative approximativement linéaire entre  $\frac{E_{90\%}-E_{10\%}}{E_{50\%}}$  et le logarithme de la superficie drainée, sauf pour les superficies drainées inférieures à la superficie d'un pixel de 4 km. Pour ces petites superficies, ils ont observé que les valeurs de  $\frac{E_{90\%}-E_{10\%}}{E_{50\%}}$  sont plus élevées pour les régions à faible pente que sur les régions à forte pente.

Les études mentionnées ci-haut désagrègent une précipitation originalement disponible à

faible résolution spatiale. Dans de telles situations, il est attendu que les différences entre les débits simulés à partir des précipitations agrégées et désagrégées à fine résolution soient significatives. Bindlish et Barros (2000) ont, quant à eux, utilisé une précipitation d'un modèle atmosphérique à une échelle beaucoup plus fine (4 km) en entrée dans leur modèle de désagrégation. Cette résolution a été obtenue en emboîtant trois domaines (Yu *et al.*, 1999). Leur étude montre que désagréger de 4 à 1 km avait peu d'impact sur le volume total d'eau de l'hydrogramme simulé sur leur grand bassin versant montagneux (14 710 km<sup>2</sup>), mais qu'elle pouvait modifier le débit de pointe horaire simulé de 20 à 30 %.

Bacchi et Ranzi (2003) ont désagrégé la précipitation aux 10 minutes de 14 à 0,9 km à partir d'un modèle multi-échelle à ondelettes de Haar sur deux bassins versants de moins de 1500 km<sup>2</sup> situés dans les Alpes. Pour deux événements intenses, ils ont calculé les différences entre les débits de pointe simulés à partir des précipitations agrégées et de 10 séries désagrégées et ce pour 19 sous-bassins. Les différences moyennes n'augmentent que très légèrement lorsque la superficie drainée diminue ; pour une superficie drainée de 2,6 km<sup>2</sup>, la différence moyenne est d'environ 5 %. Bacchi et Ranzi (2003) mentionnent que ce faible impact peut être dû au fait que leur modèle de désagrégation multi-échelle en cascade ne reproduit pas la corrélation spatiale de la précipitation observée. Les variations obtenues par Bacchi et Ranzi (2003) d'une série désagrégée à une autre semblent être plus importantes lorsque la superficie drainée diminue, mais il y a trop peu de séries désagrégées (10) pour caractériser l'incertitude comme dans Forman *et al.* (2008).

### 8.3 Discussion

La désagrégation a très peu été utilisée en modélisation hydrologique comparativement à d'autres méthodes de mise à l'échelle de la précipitation. La désagrégation peut s'avérer intéressante en modélisation hydrologique pour deux raisons. Premièrement, la distribution spatiale de la précipitation désagrégée est plus réaliste que la précipitation agrégée, permettant en principe un écoulement simulé plus près de la réalité. L'impact devrait être plus fort sur les petites superficies drainées. Deuxièmement, un modèle de désagrégation génère plusieurs séries en peu de temps, ce qui permet de produire plusieurs débits simulés pour un seul événement de précipitation. Cet ensemble de débits simulés permet de représenter l'incertitude associée à la résolution spatiale de la précipitation. Dans un contexte de prévision météorologique ou de projection climatique, la gestion d'une situation

peut être améliorée lorsqu'un ensemble de débits est simulé puisqu'il permet aux décideurs d'intégrer les incertitudes sur les résultats au processus décisionnel (Boucher *et al.*, 2011). L'incertitude devrait être plus forte pour les petits bassins versants (Bacchi et Ranzi, 2003; Carpenter et Georgakakos, 2006; Forman *et al.*, 2008).

Cependant, même pour un modèle hydrologique tenant compte des processus physiques liés à l'écoulement, la représentation des macro-processus (*e.g.* les écoulements préférentiels dans les sols forestiers) demeure souvent approximative. Il est impossible de simuler ce qui se passe avec exactitude à fine échelle puisque la physique d'un modèle hydrologique ne permet pas de le faire et aussi parce que les données physiographiques (*e.g.* types de sols, épaisseurs des couches de sols) d'un bassin versant ne sont pas connues avec exactitude. Ainsi, le raffinement de la résolution spatiale de la précipitation et/ou de la taille des unités hydrologiques définies par l'utilisateur ne mène pas nécessairement à de meilleurs résultats. Par exemple, Bell et Moore (2000) ont obtenu de meilleurs résultats en discrétisant le bassin versant sur une grille à résolution 5 km plutôt que sur une grille à 2 km.

La simplification des processus physiques par le modèle hydrologique fait en sorte que certains paramètres doivent être *calés*, c'est-à-dire estimés empiriquement. Le jeu de paramètres estimés dépend notamment des données de précipitation utilisées (Liang *et al.*, 2004) et aussi des débits observés disponibles pour le calage. Ainsi, pour une juste comparaison entre deux sources de précipitation (*e.g.* précipitations agrégées et désagrégées) utilisées en entrée, il faut un calage à partir de chacune des sources (Bell et Moore, 2000) ou encore un seul calage effectué à partir d'une source de précipitation indépendante (Bacchi et Ranzi, 2003).

L'objectif général de la Partie III de la thèse est d'évaluer l'impact de l'utilisation du modèle de désagrégation développé dans cette thèse en modélisation hydrologique. Il est attendu que l'impact dépend de la superficie drainée. Le présent chapitre a montré qu'il y a plusieurs facteurs pouvant masquer la relation entre la résolution spatiale de la précipitation et la superficie drainée. Outre le modèle hydrologique utilisé et le calage de ses paramètres, il y a les conditions initiales du bassin (*e.g.* les teneurs en eau dans le sol), l'intensité de l'événement, le type d'événement (convectif ou stratiforme), la topographie, les types de sols (argileux, sablonneux, etc.) et l'occupation du sol (urbain, agricole, forestier, etc.). Il est impossible de tenir compte de tous ces facteurs dans une seule analyse d'impact.

Dans les deux prochains chapitres, l'impact de la désagrégation sur le débit simulé est évalué à partir d'une précipitation intense survenue en juin 2002 sur le bassin de la rivière des Anglais (730 km<sup>2</sup>), tributaire de la rivière Châteauguay. L'impact est analysé sur différents sous-bassins à l'aide du modèle hydrologique distribué HYDROTEL. Le Chapitre 9 présente le protocole expérimental, soit les raisons expliquant le choix de cet événement et de ce bassin versant, les descriptions du bassin et des données utilisées, les détails sur l'application de la désagrégation et la mise en place de la modélisation hydrologique. Le Chapitre 10 présente l'impact de la désagrégation en fonction de la superficie drainée. D'abord la capacité de la précipitation désagrégée à reproduire avec exactitude le débit simulé à partir de la précipitation de référence est évaluée. Ensuite, la valeur informative de l'intervalle des débits simulés à partir de la précipitation désagrégée est analysée. Ces résultats sont présentés en fonction de la résolution de la précipitation désagrégée et de quelques facteurs mentionnés plus haut (calage d'HYDROTEL, type de sols et teneur en eau du sol avant l'événement). Ces analyses ont pour but de déterminer les conditions où la désagrégation peut s'avérer pertinente.

## Chapitre 9

# PROTOCOLE EXPÉRIMENTAL

Considérant que la désagrégation représente une étape à ajouter à la modélisation hydrologique, il sera intéressant de l'appliquer seulement si les séries de précipitation désagrégées amènent une information supplémentaire par rapport à la série à résolution mésoéchelle (agrégée). Il est donc important d'évaluer les conditions pour lesquelles la désagrégation vaut la peine d'être utilisée. Le Chapitre 8 a montré que l'impact de la désagrégation augmente lorsque la superficie d'un bassin diminue (Bacchi et Ranzi, 2003; Forman *et al.*, 2008), mais qu'il y a plusieurs autres facteurs qui entrent en jeu. Ces facteurs peuvent être classés comme suit : (i) les caractéristiques de l'événement (intensité, durée, localisation); (ii) les caractéristiques physiographiques du bassin versant (topographie, densité du réseau hydrographique, couverture des sols, types de sol), de même que (iii) le modèle hydrologique utilisé (distribué ou global, processus physiques simulés menant à l'écoulement, pas de temps). À cela s'ajoute l'impact du modèle de désagrégation qui génère les champs de précipitation.

Dans le présent chapitre, le modèle de désagrégation spatiale de la précipitation développé à la Partie I est appliqué dans un contexte d'estimation du débit obtenu suite à un événement de pluie intense survenu sur un bassin versant du sud du Québec. Peu d'études ont utilisé la précipitation désagrégée en modélisation hydrologique et à notre connaissance de la littérature, aucun modèle de désagrégation statistique n'a été appliqué sur le territoire québécois. Aussi, le modèle de désagrégation développé à la Partie I tient compte explicitement de la corrélation spatiale contrairement aux modèles de désagrégation utilisés jusqu'à maintenant en modélisation hydrologique (*e.g.* Bacchi et Ranzi, 2003). L'impact de la désagrégation est évalué en fonction de la superficie drainée et de la résolution spatiale des pixels désagrégés. Il faut être conscient que les constats tirés de ces analyses dépendent

du contexte de la présente application et ne peuvent être généralisés à tous les autres bassins versants ou types d'événements. La Figure 9.1 illustre le contexte de la présente application. L'objectif de la Partie III de la thèse est d'évaluer la valeur ajoutée de la désagrégation sur la simulation hydrologique d'un événement intense. Dans la mesure où le contexte du cas d'étude le permet, des recommandations sur l'utilisation de la désagrégation dans un mode opérationnel de prévision hydrologique seront émises.

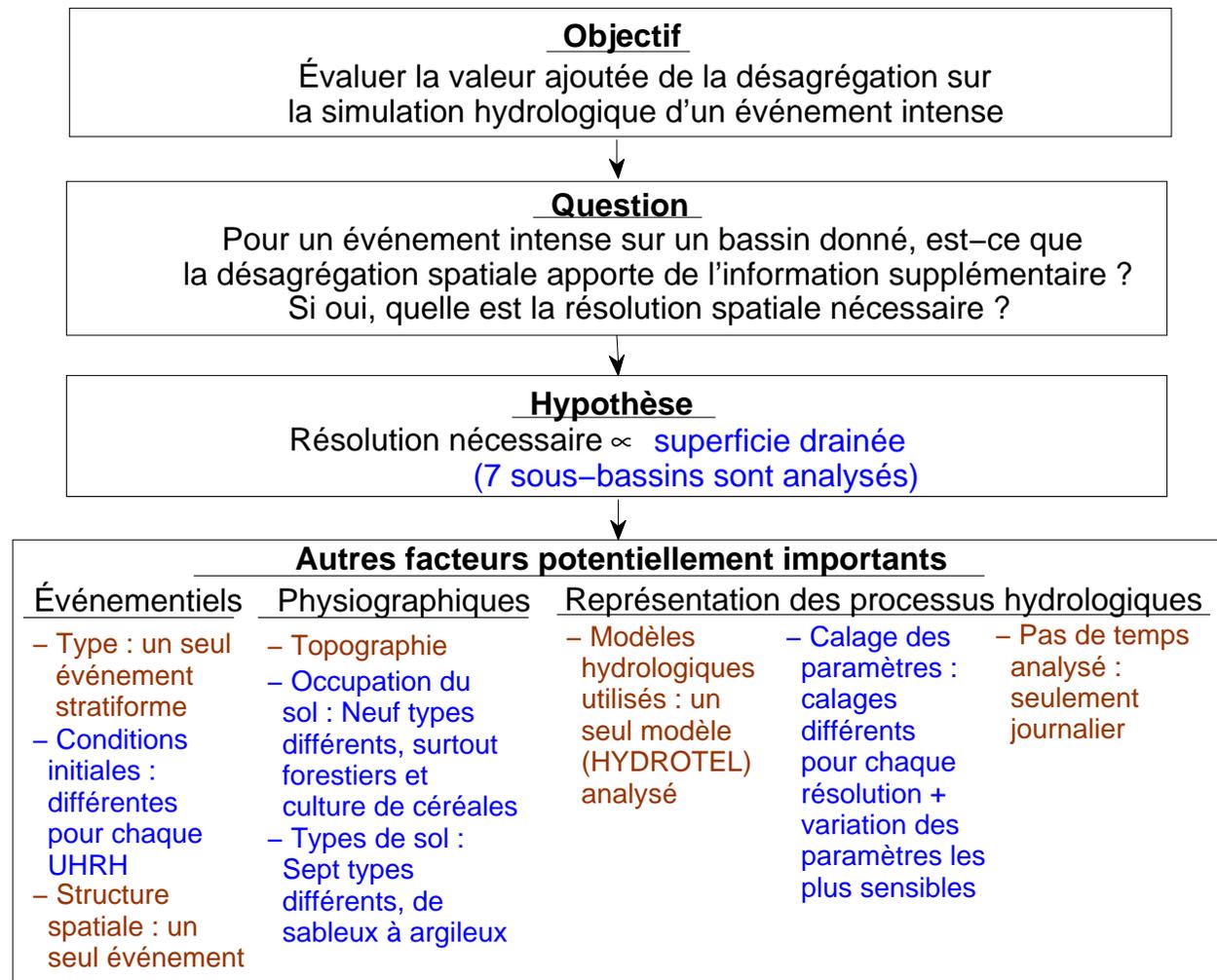


Figure 9.1 – Contexte de l'application de la désagrégation en simulation hydrologique. Les résultats sont exprimés en fonction des facteurs en bleu au Chapitre 10 et à l'Annexe F. L'impact des facteurs en marron n'a pu être considéré dans la présente étude de cas.

La suite de ce chapitre décrit le protocole expérimental, soit le territoire à l'étude, incluant le choix de l'événement et les données hydrométriques, météorologiques et atmosphériques utilisées (Section 9.1), l'estimation des paramètres du modèle de désagrégation (Section 9.2) et

Tableau 9.1 – Pourcentages des différents types de sol sur les sept sous-bassins. Le sous-bassin 1 est en fait le bassin de la rivière des Anglais en entier. Le type de sol dominant dans chaque sous-bassin est en caractère gras.

Sous-bassin	1	2	3	4	5	6	7
Superficie (km <sup>2</sup> )	729,7	297,0	58,9	140,0	60,4	170,3	111,3
Type de sol							
Sable	4,8	2,1	0,0	2,3	0,0	7,6	14,0
Sable loameux	14,7	8,6	16,8	3,0	<b>60,1</b>	5,1	<b>26,8</b>
Loam sableux	<b>24,2</b>	15,7	8,3	18,4	30,4	<b>64,0</b>	0,0
Loam	17,3	22,7	0,0	<b>46,7</b>	5,8	20,2	19,1
Loam-sableux argileux	5,7	7,8	0,1	12,3	1,0	0,0	16,0
Argile loameuse	15,8	3,1	0,0	0,1	2,7	3,1	23,5
Argile	17,4	<b>40,0</b>	<b>74,9</b>	17,3	0,0	0,0	0,6

le modèle hydrologique utilisé (Section 9.3). Les résultats de l'application suivent au Chapitre 10.

## 9.1 Territoire à l'étude

Le territoire à l'étude est le bassin versant de la rivière des Anglais (730 km<sup>2</sup>; Figure 9.2). La plus grande partie du bassin est situé au sud du territoire québécois, à l'exception de l'extrême sud du bassin qui est situé dans l'État de New-York. Il est caractérisé par un climat de type modéré et sub-humide, avec des étés chauds et des hivers froids (Côté *et al.*, 2006). La rivière des Anglais se jette dans la rivière Châteauguay qui à son tour se jette dans le fleuve Saint-Laurent, au sud de Montréal. Il aurait été intéressant d'étudier plusieurs bassins, mais le bassin de la rivière Des Anglais est le seul bassin québécois étant couvert par les données de précipitation à fine résolution *Stage IV*, permettant l'estimation des paramètres du modèle de désagrégation. Les résultats seront analysés à l'exutoire du bassin ainsi qu'à l'échelle de six sous-bassins (Figure 9.3).

Le relief du bassin est relativement plat. L'altitude atteint 320 m au sud, c'est-à-dire à l'amont, et elle varie de 30 à 100 m sur la majeure partie du bassin. Les sols y sont majoritairement sablonneux, sauf dans la partie nord du bassin, c'est-à-dire à l'aval, où

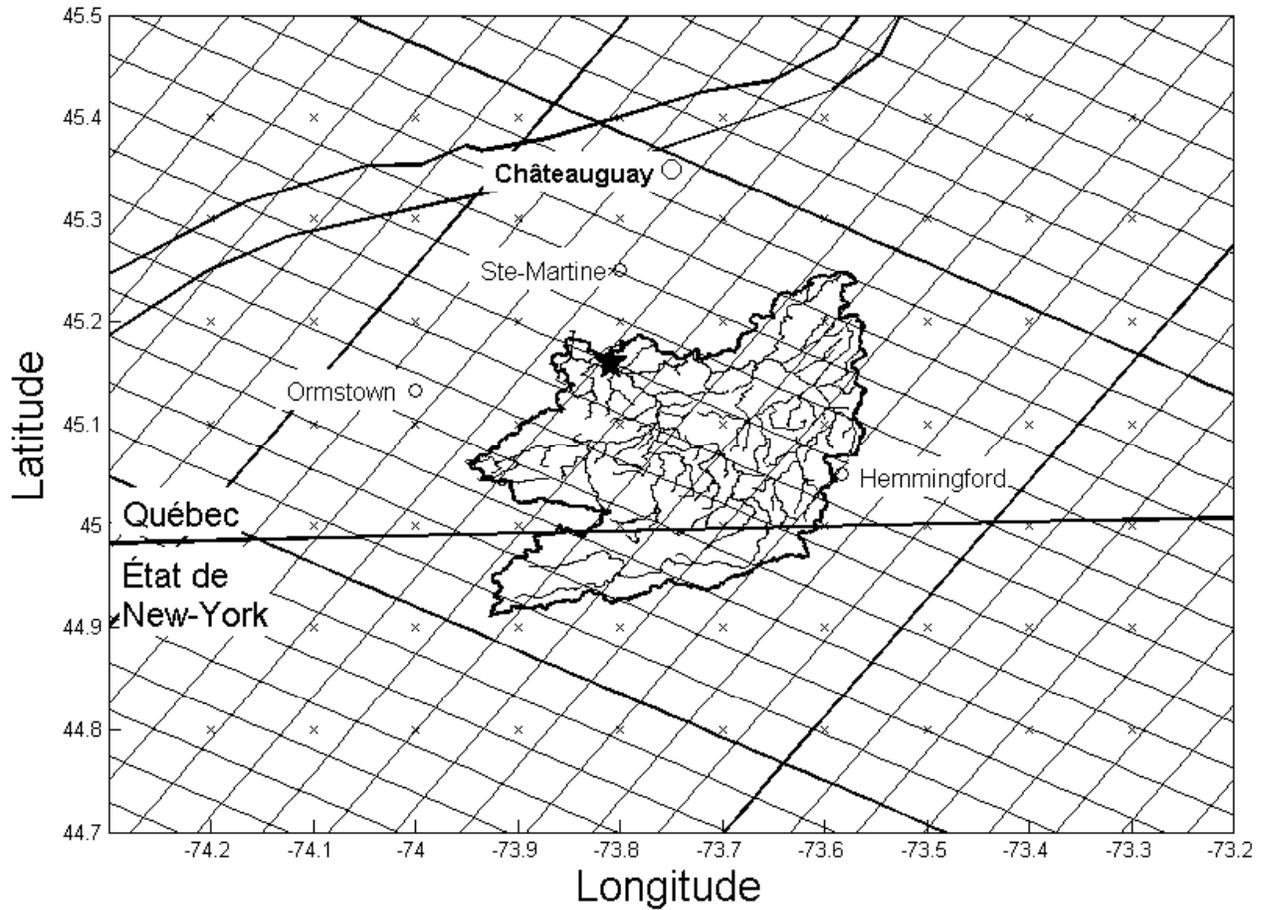


Figure 9.2 – Bassin versant de la rivière des Anglais avec les différentes sources de données hydrométéorologiques utilisées. La grille représente les données de *Stage IV* pour la tuile recouvrant le bassin, les « x » sont les points de la grille météorologique du CEHQ et l'étoile est la station hydrométrique disponible.

les sols sont plutôt argileux (Tableau 9.1). L'occupation du sol varie aussi du nord au sud (Tableau 9.2). L'activité agricole, soit les cultures de maïs et de céréales ainsi que l'élevage, est très présente dans la partie nord alors que la partie sud est plutôt couverte de forêt. Les zones urbaines occupent une fraction négligeable du bassin.

### 9.1.1 Choix de l'événement

L'intérêt pratique de la désagrégation réside surtout dans la simulation d'événements générant de forts débits localisés. Contrairement aux Parties I et II, où seulement huit champs désagrégés ont été produits pour des centaines d'événements, la désagrégation produit 1000 champs pour un seul événement pluvieux survenu le 11 juin 2002. Cet événement est de



Figure 9.3 – Exutoire et numéro de référence des sept sous-bassins de la rivière des Anglais analysés. Le sous-bassin 2 inclut les sous-bassins 3 et 4 et le sous-bassin 1 est en fait le bassin de la rivière des Anglais en entier.

loin le plus intense événement journalier à être survenu durant la période où les diverses sources de données étaient toutes disponibles (2002-2010). Les cartes synoptiques, disponibles à l'adresse [http://docs.lib.noaa.gov/rescue/dwm/data\\_rescue\\_daily\\_weather\\_maps.html](http://docs.lib.noaa.gov/rescue/dwm/data_rescue_daily_weather_maps.html), montrent qu'un front chaud s'est déplacé du mid-ouest américain pour se diriger vers la côte atlantique dans les jours précédant l'événement. Un front stationnaire s'est ensuite installé dans le nord-est américain pour quelques jours. Pour le 11 juin 2002, la précipitation sur les points de la grille météorologique du CEHQ (Section 9.1.3) recouvrant le bassin varie de 70,8 à 91,8 mm. Les débits moyens observés à la station hydrométrique du CEHQ située près de l'exutoire pour les deux journées suivantes sont de 190 et 194 m<sup>3</sup>/s ce qui correspond à un écoulement (*i.e.* débit par unité de superficie drainée) d'environ 25 mm/jour. À titre de comparaison, il s'est écoulé en moyenne moins de 2 mm/jour en juin de 2002 à 2008. Cette quantité considérable d'eau a provoqué des inondations localisées sur quelques centaines d'hectares (Van Der Heyden, 2008). Les sous-sections suivantes présentent les données utilisées.

Tableau 9.2 – Pourcentages des différentes occupations du sol sur les sept sous-bassins. L’occupation du sol dominante dans chaque sous-bassin est en caractère gras.

Sous-bassin	1	2	3	4	5	6	7
Superficie (km <sup>2</sup> )	729,7	297,0	58,9	140,0	60,4	170,3	111,3
Occupation du sol							
Urbain	0,4	0,6	0,0	0,5	0,0	0,0	0,1
Pâturage	6,3	6,7	4,6	5,5	2,3	1,0	6,0
Culture de céréales	28,0	<b>39,0</b>	<b>39,7</b>	30,1	14,4	20,7	19,1
Culture du maïs	7,3	4,7	2,9	4,6	2,5	1,1	13,8
Milieu humide	0,6	0,0	0,0	0,0	0,2	0,3	3,5
Sol nu	1,0	1,1	0,6	1,3	0,0	1,2	2,0
Friches	2,2	1,0	1,9	0,7	0,2	5,7	1,2
Forêt de feuillus	<b>36,8</b>	25,7	37,7	24,4	<b>61,0</b>	<b>49,5</b>	<b>45,0</b>
Forêt de conifères	17,3	21,3	12,7	<b>32,8</b>	19,5	20,6	9,4

### 9.1.2 Données hydrométriques

Les données de débits proviennent des stations hydrométriques du CEHQ et sont disponibles sur le site internet du CEHQ ([www.cehq.gouv.qc.ca](http://www.cehq.gouv.qc.ca)). Cinq stations sont situées sur le bassin, mais une seule, située près de l’exutoire, était opérationnelle en juin 2002 (Figure 9.2). Les données à la station sont des moyennes journalières, de minuit à minuit heure locale, du débit exprimé en m<sup>3</sup>/s. Seules les données de mai à novembre 2002 ont été utilisées pour le calage du modèle hydrologique (plus de détails à la Section 9.3.2).

### 9.1.3 Données de précipitation

Deux sources de données de précipitation ont été utilisées dans la présente application. Encore une fois, les données journalières de *Stage IV* (Lin et Mitchell, 2005), provenant cette fois-ci du *Northeast RFC* (NERFC), sont utilisées. Dans cette région, les pixels ont une résolution de 4,4 km. Ces données permettent l’estimation des paramètres du modèle de désagrégation et la formation de tuiles à désagréger. Les tuiles à résolution 52,8 km sont formées en agrégeant 12 × 12 pixels. Le bassin versant est couvert entièrement par une seule tuile, mais la Section 4.1 a montré qu’il est préférable de conserver une zone tampon d’une tuile de chaque côté de la ou des tuiles d’intérêt. C’est pourquoi la désagrégation sera faite

sur  $3 \times 3$  tuiles.

La Figure 9.4 présente la précipitation journalière moyenne des données de *Stage IV* de mai à novembre de 2002 à 2005 sur les  $36 \times 36$  pixels étudiés. La période de mai à novembre est utilisée pour pouvoir se concentrer uniquement sur la précipitation liquide. La figure montre qu'il y a des zones anormalement moins pluvieuses au sud et au nord. Au sud, elles ont vraisemblablement été causées par les montagnes Adirondack, qui diminuent la portée des radars. Au nord, les pixels sont probablement situés au-delà de la portée des radars utilisés, tous localisés aux États-Unis.

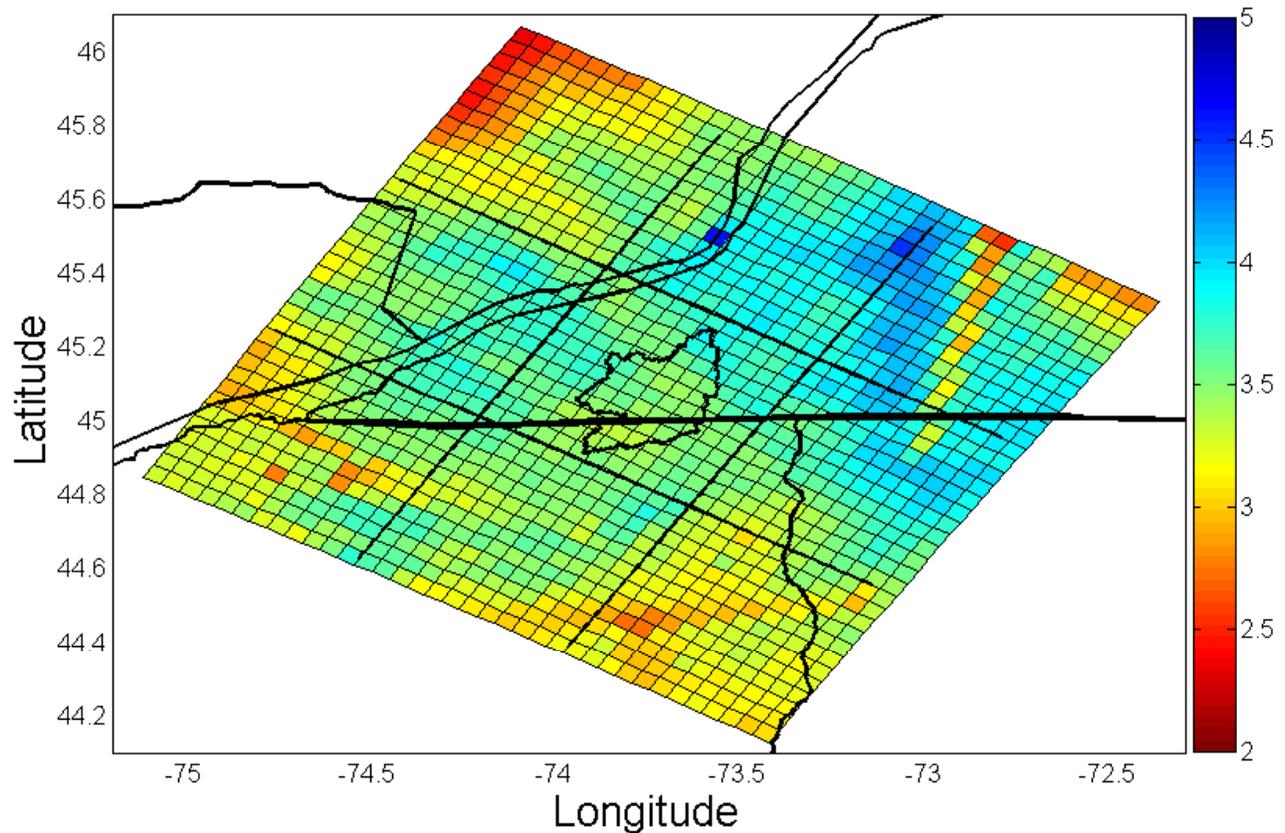


Figure 9.4 – Précipitation moyenne (mm/jour) de mai à novembre (2002-2005) des données de *Stage IV* pour les  $36 \times 36$  pixels étudiés.

Afin de vérifier la validité des données *Stage IV* sur le territoire étudié, les données de la grille météorologique du CEHQ sont aussi utilisées. Cette grille à résolution  $0,1^\circ$  fournit pour tout le sud du Québec la précipitation journalière, ainsi que les températures minimales

et maximales journalières, à partir de 1973. Pour la précipitation journalière, les valeurs aux points de grille sont obtenues par un krigeage simple des stations climatologiques et météorologiques opérées par le ministère du développement durable, de l'environnement et des parcs du Québec (MDDEP), à partir d'un variogramme sphérique (Bourdillon *et al.*, 2011).

Pour pouvoir comparer les données de *Stage IV* avec celles de la grille du CEHQ, il faut mettre les deux sources à la même échelle spatiale. Ainsi, pour chacune des neuf tuiles et chaque jour de mai à novembre de 2002 à 2005, la précipitation moyenne des  $12 \times 12$  pixels *Stage IV* et des points de la grille CEHQ inclus dans une tuile a été calculée (Figure 9.5). L'intervalle de 24h utilisé par le CEHQ s'étend de 8h à 8h heure locale et celui utilisé par *Stage IV* s'étend de midi à midi UTC (Temps Universel Coordonné), ce qui revient également de 8h à 8h heure locale à l'heure avancée de l'est de l'Amérique du Nord. Les résultats montrent que la tuile au centre (celle recouvrant le bassin) présente la meilleure adéquation entre les deux sources de précipitation ( $R^2 = 0,79$ ). La relation est beaucoup moins bonne pour certaines tuiles du nord et du sud.

Comme les différences entre les données de *Stage IV* et celles du CEHQ proviennent vraisemblablement des difficultés des radars à couvrir la région, les neuf tuiles à désagréger pour l'événement du 11 juin 2002 sont construites à partir des données de la grille du CEHQ. À noter que la pluie est réellement tombée entre le 11 juin à 8h et le 12 juin à 8h heure locale, mais pour la modélisation hydrologique, il sera supposé qu'elle est tombée le 11 juin (*i.e.* du 11 juin à 0h au 12 juin à 0h). Les données originales de la grille du CEHQ à résolution  $0,1^\circ$  sont aussi utilisées pour les autres jours de la simulation hydrologique (Section 9.3.2). Les  $12 \times 12$  pixels *Stage IV* sur la tuile couvrant le bassin versant sont utilisés de mai à novembre de 2002 à 2005 pour estimer les paramètres du modèle de désagrégation (Section 9.2). Pour le 11 juin 2002, ces mêmes pixels sont utilisés comme pluie de référence pour le calage du modèle hydrologique (Section 9.3.2). Dans ce cas, un facteur multiplicatif est appliqué pour que la précipitation moyenne sur les  $12 \times 12$  pixels soit égale à la précipitation moyenne sur la tuile formée à partir des points de la grille du CEHQ (83,2 mm). Le Tableau 9.3 résume les différentes sources de précipitation utilisées. La Figure 9.6 présente la précipitation de référence à diverses résolutions spatiales pour l'événement du 11 juin 2002. La portion de la tuile recevant le plus de précipitation est située à l'intérieur des limites du bassin. Un examen de la précipitation de *Stage IV* au pas de temps de six heures révèle que la pluie a

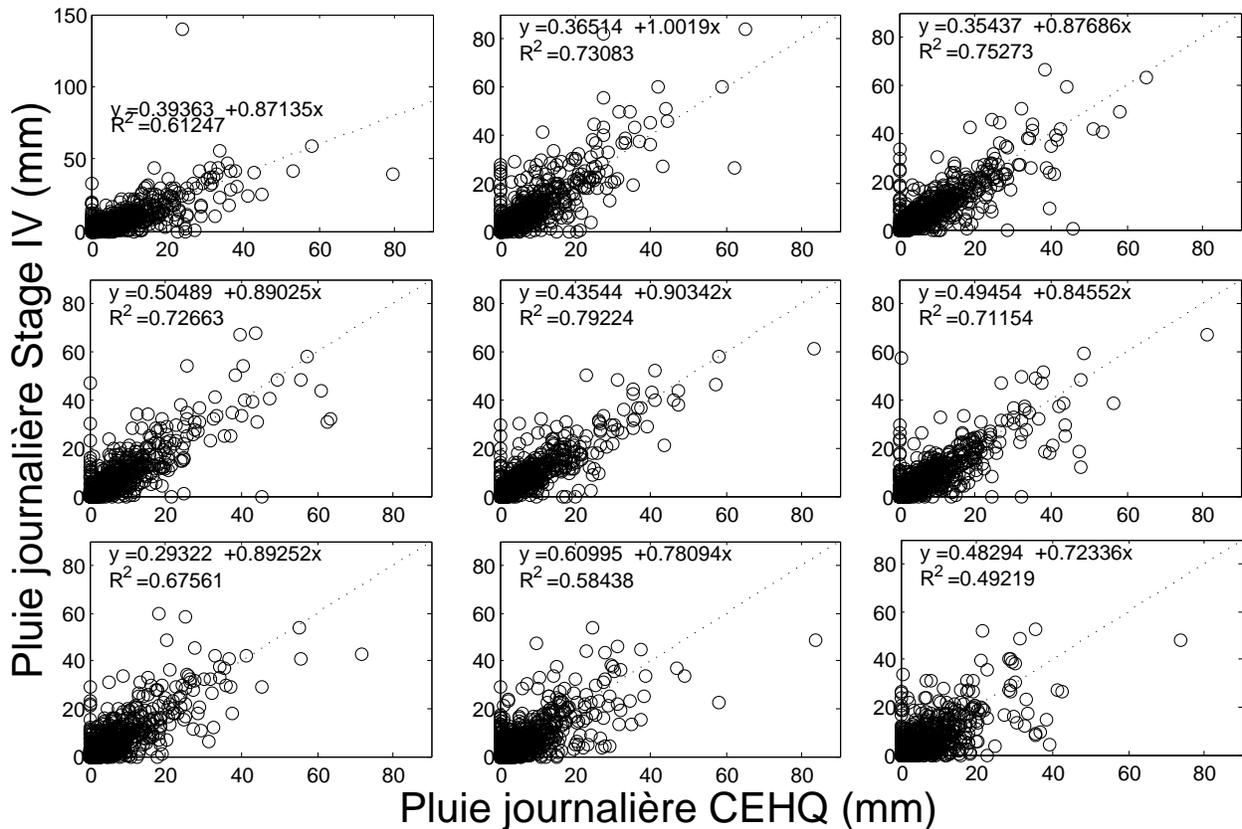


Figure 9.5 – Précipitation journalière de mai à novembre (2002-2005) provenant des données de *Stage IV* et de la grille du CEHQ agrégées sur chacune des  $3 \times 3$  tuiles. La figure centrale représente la tuile qui contient le bassin. En haut : tuiles du nord ; en bas : tuiles du sud ; à gauche : tuiles de l'ouest ; et à droite : tuiles de l'est.

débutée le 10 juin en soirée pour atteindre son intensité maximale le 11 juin en soirée et se terminer le 12 juin 2002 en avant-midi.

### 9.1.4 Données atmosphériques

Comme pour la Partie I, les valeurs de CAPE, de vitesse et de direction du vent à 700 hPa produites par le NARR (Mesinger *et al.*, 2004) sont utilisées. La moyenne des données des deux points de la grille NARR situés au-dessus de la tuile d'intérêt est calculée afin d'obtenir une seule valeur par jour sur la tuile pour chaque variable. Les données de mai à novembre de 2002 à 2005 sont utilisées pour l'estimation des paramètres du modèle de désagrégation (Section 9.2). Les données du 11 juin 2002 sont aussi utilisées en entrée dans le modèle

Tableau 9.3 – Sources de précipitations journalières utilisées pour la désagrégation et/ou pour la simulation hydrologique.

Source	Utilisation
Pixels <i>Stage IV</i> originaux	Calage du modèle de désagrégation (mai à novembre 2002-2005). Seulement les $12 \times 12$ pixels recouvrant le bassin sont utilisés.
Grille du CEHQ originale	Données d'entrée dans HYDROTEL pour le calage, sauf pour le 11 juin 2002.
« Facteur de correction »	Précipitation moyenne des points de la grille du CEHQ divisée par la précipitation moyenne des pixels <i>Stage IV</i> sur la tuile de 52,8 km recouvrant le bassin de la rivière des Anglais (Figure 9.2) pour le 11 juin 2002.
Pixels <i>Stage IV</i> corrigés	Précipitation sur les pixels <i>Stage IV</i> originaux le 11 juin 2002 $\times$ « Facteur de correction ». Précipitation de référence utilisée aux résolutions 4,4, 8,8, 13,2, 17,6 et 26,4 km en entrée dans HYDROTEL pour produire l' <i>écoulement de référence</i> .
Neuf tuiles du CEHQ	Trois $\times$ trois tuiles à résolution 52,8 km (tuile sur le bassin de la rivière des Anglais plus une tuile dans chaque direction ; Figure 9.2). La valeur sur chaque tuile est la moyenne de la précipitation sur le 11 juin 2002 des points de la grille du CEHQ inclus dans la tuile. Elles sont utilisées pour la désagrégation et l'interpolation (Figure 9.10).

de désagrégation. Pour cette journée, le vent à 700 hPa soufflait en direction est, avec une moyenne de 16,6 m/s, ce qui est légèrement plus élevé que le vent moyen journalier sur la période 2002-2005, et la valeur moyenne de CAPE était de 2,5 J/kg. La valeur de CAPE indique que l'énergie convective pour cet événement était faible. L'examen des valeurs de CAPE au pas de temps de trois heures montre que l'énergie convective était tout aussi faible pour les jours précédents.

## 9.2 Estimation des paramètres du modèle de désagrégation

Comme le modèle de désagrégation est appliqué sur un seul événement intense, il aurait été souhaitable de n'utiliser que des événements forts pour estimer les paramètres du modèle.

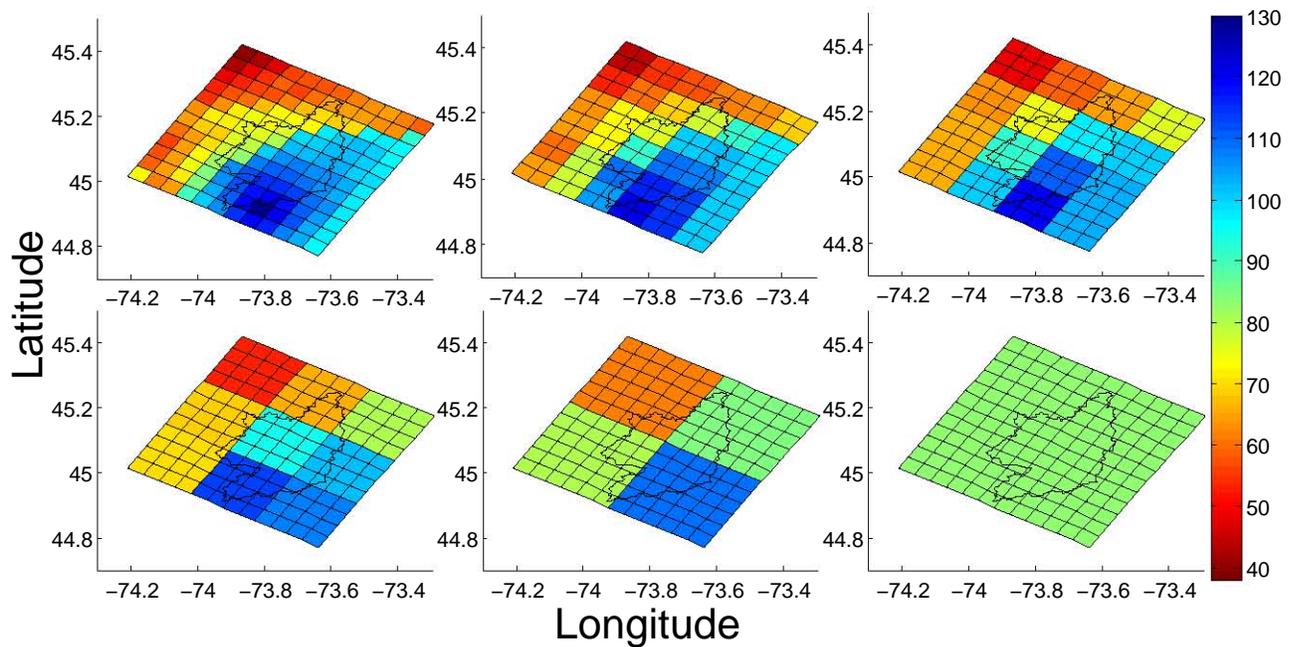


Figure 9.6 – Précipitations (mm) de référence pour le 11 juin 2002 utilisées pour simuler les débits de référence à 4,4, 8,8, 13,2, 17,6, 26,4 et 52,8 km.

Cependant, ces événements sont trop peu nombreux et ne fournissent pas assez de données pour permettre une estimation adéquate des paramètres. C'est pourquoi les paramètres sont estimés à partir de tous les événements de mai à novembre de 2002 à 2005. Pour les raisons expliquées à la Section 9.1.3, seuls les  $12 \times 12$  pixels de *Stage IV* de la tuile recouvrant le bassin sont retenus pour l'estimation des paramètres. Au total, cela représente environ 60 000 valeurs de précipitation disponibles sur les pixels de 4,4 km, ce qui est nettement inférieur au nombre de données disponibles à la Partie I (un peu plus de cinq millions). Comme le nombre de données est limité, les paramètres sont estimés seulement pour les pixels de 4,4 km. Si les résultats sont comparables à ceux obtenus au Chapitre 3 pour les données du sud-est des États-Unis, ces dernières valeurs de paramètres seront utilisées pour chacune des résolutions de pixels. Il est à noter que la majorité du territoire à désagréger présente peu de relief, sauf au sud, où se trouvent les montagnes Adirondack. Cependant, la Partie II a montré que le modèle de désagrégation doit être ajusté pour tenir compte de la topographie seulement s'il y a des anomalies topographiques de plusieurs centaines de mètres. Comme la topographie est peu variable sur la tuile couvrant le bassin, il n'est pas nécessaire de la considérer.

Les paramètres de l'espérance sont estimés comme à la Section 3.2.1. Dans un premier temps, il faut s'assurer que le vent a un impact. Dans le sud-est des États-Unis, le vent

à 700 hPa a permis d'expliquer l'anisotropie à partir d'une équation à deux paramètres (équation (3.4)). La somme du carré des écarts entre les valeurs observées et espérées a été 3,1 % plus petite pour l'équation (3.4) par rapport à l'équation (3.2), qui compte un paramètre de plus, mais qui ne tient pas compte du vent (Tableau 3.3). Par contre, dans les montagnes du nord-ouest des États-Unis (Partie II), le vent n'avait pu être incorporé au modèle. Pour estimer l'impact du vent sur la tuile du bassin de la rivière des Anglais, la même procédure qu'à la Section 3.2.1 est utilisée. Les analyses ont montré que l'équation (3.4) permet une réduction de 1,5 % par rapport à l'équation (3.2). Ce n'est qu'une légère diminution, mais le plus important est que l'équation (3.4) permet d'éliminer un paramètre. L'équation (3.4) est donc retenue pour calculer l'espérance sur un pixel, comme à la Partie I.

L'écart-type est estimé selon la méthode présentée à la Section 3.2.2. Le faible nombre de pixels rend un peu plus difficile l'estimation de l'écart-type, d'autant plus qu'il y a très peu d'événements convectifs forts comparativement au sud-est des États-Unis. Seulement 19 jours de 2002 à 2005 sont classés comme convectifs (*i.e.* CAPE > 1000 J/kg), ce qui est nettement inférieur à ce qui avait été obtenu pour le sud-est des États-Unis (plus de 200 jours convectifs). Aussi, les jours les plus convectifs ne sont pas ceux ayant reçus le plus de précipitation. La précipitation maximale observée sur la tuile durant une journée convective est de 18,5 mm. Pour les autres jours convectifs, la précipitation est inférieure à 10 mm. Quoi qu'il en soit, l'écart-type a été estimée à partir de la même équation qu'à la Partie I (équation (3.7)). Les paramètres d'espérance et d'écart-type estimés sur les données *Stage IV* sur le sud-est des États-Unis (Partie I) et sur la tuile du bassin de la rivière des Anglais sont présentés au Tableau 9.4. Les résultats sont comparables, malgré le fait que les deux régions soient climatiquement différentes. Les différences entre les paramètres d'écart-type des deux régions sont possiblement causées par le faible nombre d'événements convectifs important sur la tuile québécoise. L'impact de la différence de tailles des pixels (4,4 km pour le bassin des Anglais et 3,8 km pour le sud-est des États-Unis) est selon toute vraisemblance marginale.

Comme les résultats des deux régions sont comparables et comme il y a peu de données sur le bassin de la rivière des Anglais, les valeurs des paramètres estimées à la Partie I sont utilisées. Ces valeurs des paramètres pour toutes les tailles de pixels sont présentées au Tableau 3.4 pour l'espérance et au Tableau 3.6 pour l'écart-type. Comme pour les Parties I et II, la distribution lognormale est utilisée pour générer les valeurs aux pixels et les périodes de rodage et d'autocorrélation ont été fixées respectivement à 100 et 50 itérations.

Tableau 9.4 – Paramètres du modèle de désagrégation estimés sur la tuile couvrant le bassin de la rivière des Anglais et sur le sud-est des États-Unis (Partie I, Tableaux 3.4 et 3.6).

Région	Taille des pixels	Espérance			Écart-type	
		$\beta_d$	$\beta_v$ (s/m)	$\theta_0$ (mm)	$\theta_1$ (mm kg/J)	$\theta_2$
Bassin des Anglais	4,4 km	0,805	0,0167	0,343	$2,89 \times 10^{-4}$	0,526
Sud-est des États-Unis	3,8 km	0,835	0,0146	0,442	$1,79 \times 10^{-4}$	0,628

### 9.3 Modèle hydrologique

Pour simuler le débit, le modèle hydrologique HYDROTEL (Fortin *et al.*, 2001a,b; Turcotte *et al.*, 2003) est utilisé. HYDROTEL est un modèle *distribué*, c'est-à-dire qu'il tient compte explicitement de la distribution spatiale des caractéristiques du bassin et des données météorologiques utilisées en entrée (Indarto, 2002). Un modèle distribué comme HYDROTEL est donc adapté pour évaluer l'impact de la désagrégation d'un événement intense de pluie sur le débit résultant pour diverses superficies drainées.

Le CEHQ utilise HYDROTEL de manière opérationnelle pour prévoir les apports en eau aux réservoirs publics québécois (Turcotte *et al.*, 2004). Parmi les études récentes ayant utilisé HYDROTEL sur le sud du Québec, notons Ricard (2008, bassin de la rivière Châteauguay), Quilbé *et al.* (2008; bassin de la rivière Chaudière), Velazquez *et al.* (2009; cinq bassins), Turcotte (2010; trois bassins), Boucher *et al.* (2011; bassin de la rivière Gatineau) et Bourdillon *et al.* (2011; bassin versant intermédiaire du fleuve St-Laurent, 390 000 km<sup>2</sup>). Notons qu'HYDROTEL a aussi été utilisé sur d'autres régions, par exemple sur le bassin de la rivière Nécopastic, une petite rivière du Nord-du-Québec (Jutras *et al.*, 2009), sur le bassin de la rivière Ammer en Allemagne (Ludwig *et al.*, 2009), sur le bassin de la Chute-du-Diable au Saguenay-Lac-Saint-Jean (Khalili *et al.*, 2011), sur le bassin de la rivière Orb en France (Indarto, 2002) et sur le bassin de la rivière Sassandra en Côte-d'Ivoire (Bouda *et al.*, 2012). À notre connaissance, aucune application d'HYDROTEL n'a utilisé en entrée la précipitation d'un modèle de désagrégation. D'autres modèles hydrologiques distribués pourraient être utilisés afin d'éviter que l'impact de la désagrégation soit évalué à partir de la physique d'un seul modèle hydrologique. Néanmoins, HYDROTEL est largement utilisé pour la simulation hydrologique sur le territoire québécois. Même si certains constats qui

seront faits au Chapitre 10 dépendront de la structure d'HYDROTEL, ils seront pertinents pour tous les modélisateurs intéressés à la prévision hydrologique sur un petit bassin versant québécois.

Pour la suite de la présente section, une description sommaire du modèle HYDROTEL et une description plus détaillée de son application au bassin de la rivière des Anglais est présentée.

### 9.3.1 Description sommaire d'HYDROTEL

L'application d'HYDROTEL comprend deux grandes étapes : (i) l'acquisition des caractéristiques du bassin versant et (ii) la simulation hydrologique. La première étape est réalisée par le logiciel PHYSITEL (Turcotte *et al.*, 2001 ; Rousseau *et al.*, 2011). PHYSITEL intègre les données d'un modèle numérique d'altitude (MNA) et le réseau hydrographique (s'il n'est pas disponible, il est généré à partir du MNA) pour définir la superficie drainée à chaque tronçon, soit chaque segment de rivière. Chaque tronçon définit un sous-bassin versant élémentaire, appelé *unité hydrologique relativement homogène* (UHRH). L'UHRH est la plus petite unité de simulation dans HYDROTEL. PHYSITEL permet également de préciser les types de sol et les occupations du sol. Des types de sols et leurs propriétés peuvent être fournis à HYDROTEL en entrée. Dans la présente application, onze types de sol sont définis à partir de la classification de Rawls et Brakensiek (1989). Un seul type de sol est conservé par UHRH, soit le type dominant. Les types de sol influent sur les propriétés hydrodynamiques des écoulements et sur le calcul du bilan vertical dans HYDROTEL. Dix occupations du sol différentes peuvent être spécifiées. Pour chaque UHRH, le pourcentage représenté par chacune d'entre elles est retenu. L'occupation du sol influe sur l'évapotranspiration, le bilan vertical en eau, la fonte de la neige et l'écoulement sur la partie terrestre.

Avant de procéder à la simulation hydrologique, il faut également définir la précipitation et les températures minimale et maximale à chaque pas de temps (*e.g.* trois heures, journalier). La précipitation est répartie temporellement de manière uniforme sur le pas de temps. HYDROTEL compte six sous-modèles :

1. Interpolation spatiale des données météorologiques ;
2. Évolution du couvert nival ;
3. Évapotranspiration potentielle ;

4. Bilan d'eau vertical ;
5. Écoulement sur la partie terrestre du bassin ;
6. Écoulement par le réseau hydrographique.

Pour un événement intense de précipitation estivale, les deux sous-modèles les plus importants sont le bilan d'eau vertical, généré par BV3C (pour Bilan Vertical en Trois Couches), et l'écoulement sur la partie terrestre du bassin. BV3C divise le sol en trois couches et simule les échanges d'eau entre ces couches et l'atmosphère, tel que schématisés à la Figure 9.7. L'épaisseur de chaque couche est spécifiée par l'utilisateur. La première couche de sol est mince et lorsque sa capacité d'infiltration est dépassée ou lorsque le sol est saturé en eau, il y a ruissellement de surface. La première couche ne produit pas d'écoulement latéral. La deuxième couche est généralement un peu plus épaisse et l'écoulement latéral qu'elle produit est l'*écoulement retardé*. La hauteur d'écoulement retardé dépend de l'épaisseur de la deuxième couche, de la conductivité hydraulique, qui dépend de la teneur en eau et du type de sol, et de la pente de l'UHRH. La troisième couche est la plus épaisse et est supposée être toujours près de la saturation. L'écoulement latéral qu'elle produit est l'*écoulement de base* et est peu corrélé avec les précipitations des derniers jours. Cet écoulement de base dépend de l'épaisseur de la troisième couche, de la teneur en eau et du coefficient de récession, spécifié par l'utilisateur. Chaque couche contribue à l'évapotranspiration selon la profondeur et la distribution racinaires et les échanges entre les couches. Ainsi, la teneur en eau d'une couche de sol, de même que son écoulement latéral (couches 2 et 3) ou le ruissellement de surface (couche 1), dépendent des conditions de saturation et d'écoulement des deux autres couches. Le sous-modèle BV3C possède un pas de temps interne inférieur au pas de temps de la simulation spécifié par l'utilisateur. Pour plus de détails sur BV3C, le lecteur est référé à Ricard (2008) ou à la documentation HYDROTEL au <http://hydrotel.codeplex.com/documentation>.

Le sous-modèle d'écoulement sur la partie terrestre du bassin apporte l'eau produite à chaque UHRH par BV3C vers le réseau d'hydrographique. Pour ce faire, l'algorithme utilise une approximation unidimensionnelle du système complet des équations de Saint-Venant, l'équation de l'onde cinématique. L'écoulement est représenté par la loi de Manning, avec un coefficient de rugosité dépendant de l'occupation du sol. Le sous-modèle d'écoulement terrestre distingue trois classes d'occupation du sol : les milieux forestiers, les milieux aqueux et les autres milieux. Il y a un coefficient de rugosité pour chaque classe à spécifier par

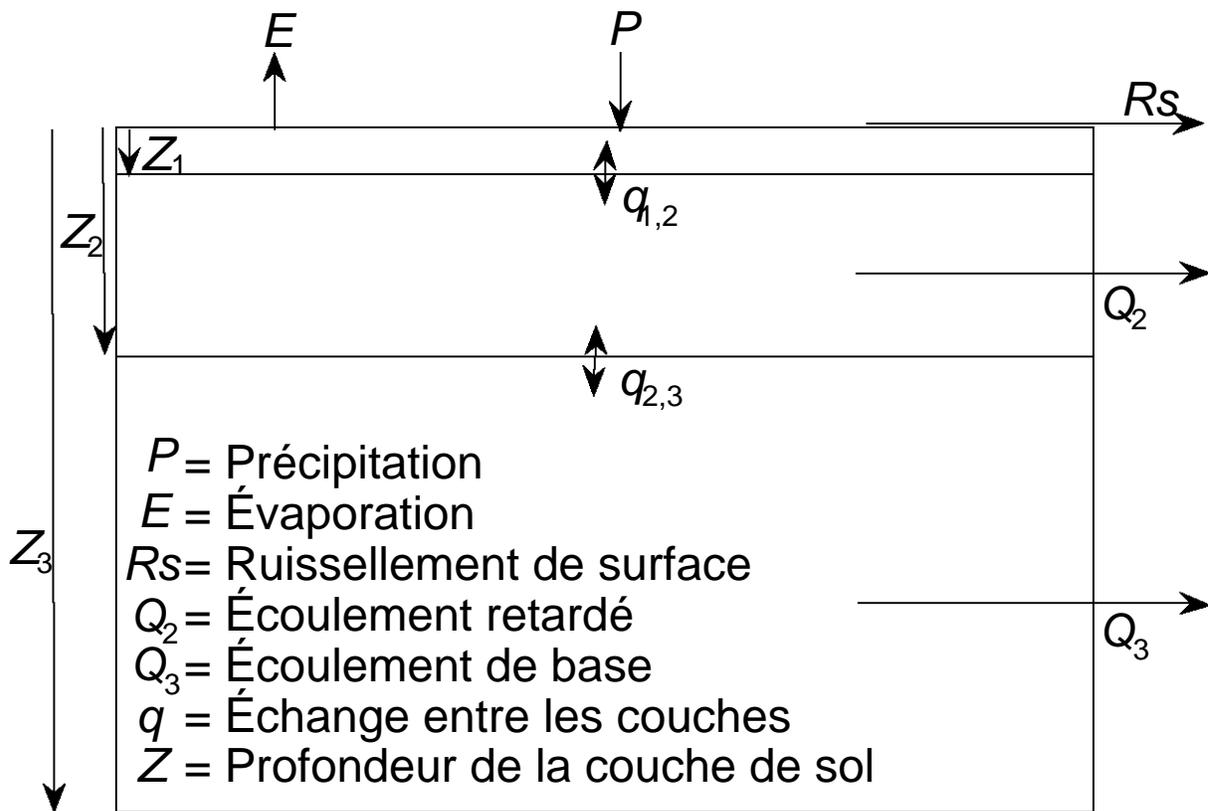


Figure 9.7 – Schéma conceptuel du sous-modèle BV3C.

l'utilisateur. Pour être plus réaliste, l'écoulement sur la partie terrestre pourrait être simulé à chaque pixel à la résolution du MNA. Cependant, cette approche nécessiterait un temps de calcul considérable. L'écoulement sur la partie terrestre est plutôt simulé à partir du bilan donné par BV3C à chaque UHRH et d'un hydrogramme géomorphologique (HGM). Un HGM est un hydrogramme fictif qui définit la distribution temporelle de l'écoulement dans le réseau hydrographique à partir d'un certain apport, uniforme sur le bassin. Cet apport, appelé *lame de référence de l'HGM* ou plus simplement *lame*, est précisé par l'utilisateur. Les HGM de tous les UHRH sont définis à partir d'une simulation préliminaire réalisée à fine échelle (*i.e.* à la résolution du MNA). Pour plus de détails sur ce sous-modèle, le lecteur est référé à la documentation d'HYDROTEL (<http://hydrotel.codeplex.com/documentation>).

Bien qu'HYDROTEL simule jusqu'à un certain point les processus physiques, il reste qu'il les représente à macro-échelle. Les équations des algorithmes des sous-modèles d'HYDROTEL ne sont qu'une approximation des processus. Également, même s'il était possible de représenter exactement les processus physiques, une simulation parfaite nécessiterait la connaissance des données physiographiques et hydrométéorologiques en tout point et en tout temps sur le bassin versant, ce qui n'est pas le cas en pratique. C'est pourquoi il est nécessaire de procéder

à un calage des paramètres du modèle. Pour les raisons mentionnées ci-haut, le calage est spécifique à la résolution spatio-temporelle des données d'entrées. Turcotte (2010) présente une méthode objective pour le calage d'HYDROTEL.

### 9.3.2 Application sur le bassin de la rivière des Anglais

Les données d'entrée dans PHYSITEL proviennent de Ricard (2008). Le bassin a une superficie de 730 km<sup>2</sup> et est divisé en 182 UHRH ayant une superficie moyenne d'environ 4 km<sup>2</sup> (Figure 9.8). Les simulations sont effectuées au pas de temps journalier. Les données de température proviennent de la grille du CEHQ. Les données de précipitation sont aussi fournies par la grille du CEHQ, sauf pour le 11 juin 2002, où diverses sources de précipitation sont utilisées (Tableau 9.3). Pour le 11 juin 2002, ces diverses sources de précipitation sont traitées par HYDROTEL comme étant des données radars. Dans HYDROTEL, les données radars doivent être sur une grille orientée nord-sud et est-ouest alors que la grille *Stage IV* est inclinée de  $-38^\circ$ . Afin de rendre les données compatibles aux critères d'HYDROTEL, une autre grille à résolution 2 km orientée nord-sud et est-ouest a été créée. La valeur de précipitation à chaque point de cette grille est la valeur de précipitation du pixel le recouvrant.

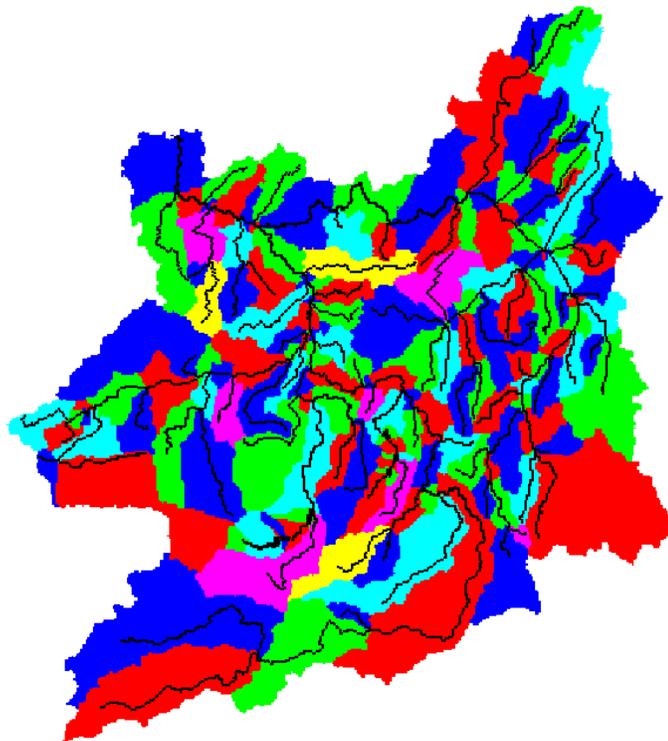


Figure 9.8 – Bassin de la rivière des Anglais divisé en 182 UHRH.

Les valeurs des paramètres estimées influent directement sur la façon dont le modèle représente les processus. C'est pourquoi il faut porter une attention particulière au calage d'HYDROTEL pour que l'impact de la désagrégation de la pluie soit évalué correctement. Pour une comparaison juste entre deux sources de précipitation en entrée dans un modèle hydrologique, il faut faire un calage approprié pour chaque source. Comme la pluie du 11 juin 2002 est donnée à HYDROTEL à cinq résolutions différentes, soit à 4,4, 8,8, 13,2, 17,6 et 26,4 km, cinq calages différents sont effectués de la façon suivante :

1. Un premier calage préliminaire est effectué de mai à novembre 2002, à partir de la station hydrométrique disponible (Figure 9.2) avec les données de précipitation de la grille du CEHQ. Le calage obtenu doit retourner des valeurs simulées et observées comparables pour le volume total d'eau sur toute la période et la pointe de débit résultant du 11 juin 2002.
2. À partir des paramètres estimés en 1, le calage est raffiné pour la période du 11 au 20 juin 2002 en prenant pour le 11 juin la précipitation *Stage IV* à 4,4 km à laquelle un facteur multiplicatif a été appliqué pour corriger par rapport à la tuile du CEHQ recouvrant le bassin (Tableau 9.3). Le calage est considéré adéquat si les volumes d'eau simulé et observé sur les 10 jours sont comparables, si les débits observés et simulés pour les 11 et 12 juin 2002 arrondis au  $m^3/s$  sont identiques et si les paramètres obtenus sont réalistes d'un point de vue physique. Le débit simulé à partir des paramètres obtenus est défini comme le *débit de référence à 4,4 km*.
3. L'étape 2 est répétée pour les autres tailles de pixels, mais en prenant comme référence le débit de référence à 4,4 km à l'exutoire au lieu du débit observé.

Les débits simulés aux différentes résolutions et observés à la station hydrométrique pour les jours suivant l'événement sont présentés à la Figure 9.9. Les valeurs des paramètres d'HYDROTEL calés aux différentes résolutions sont présentées au Tableau 9.5. Effectuer un calage à partir de différentes séries de précipitation permet aussi d'identifier les paramètres les plus sensibles. Le paramètre le plus sensible est l'épaisseur de la couche 3 ( $Z_3$ ). Indarto (2002) a déjà fait état que seul un ajustement des paramètres  $Z_3$  et de la lame de l'HGM permettait de compenser pour une faible résolution des données d'entrée dans HYDROTEL. Une diminution de l'épaisseur de la couche 3 permet d'augmenter le volume d'eau écoulée, compensant ainsi pour une hauteur de pluie trop faible et une augmentation de la lame de l'HGM permet de diminuer le temps de réponse suite à un événement pluvieux. Afin d'évaluer l'impact des valeurs attribuées à ces paramètres, d'autres simulations hydrologiques à partir

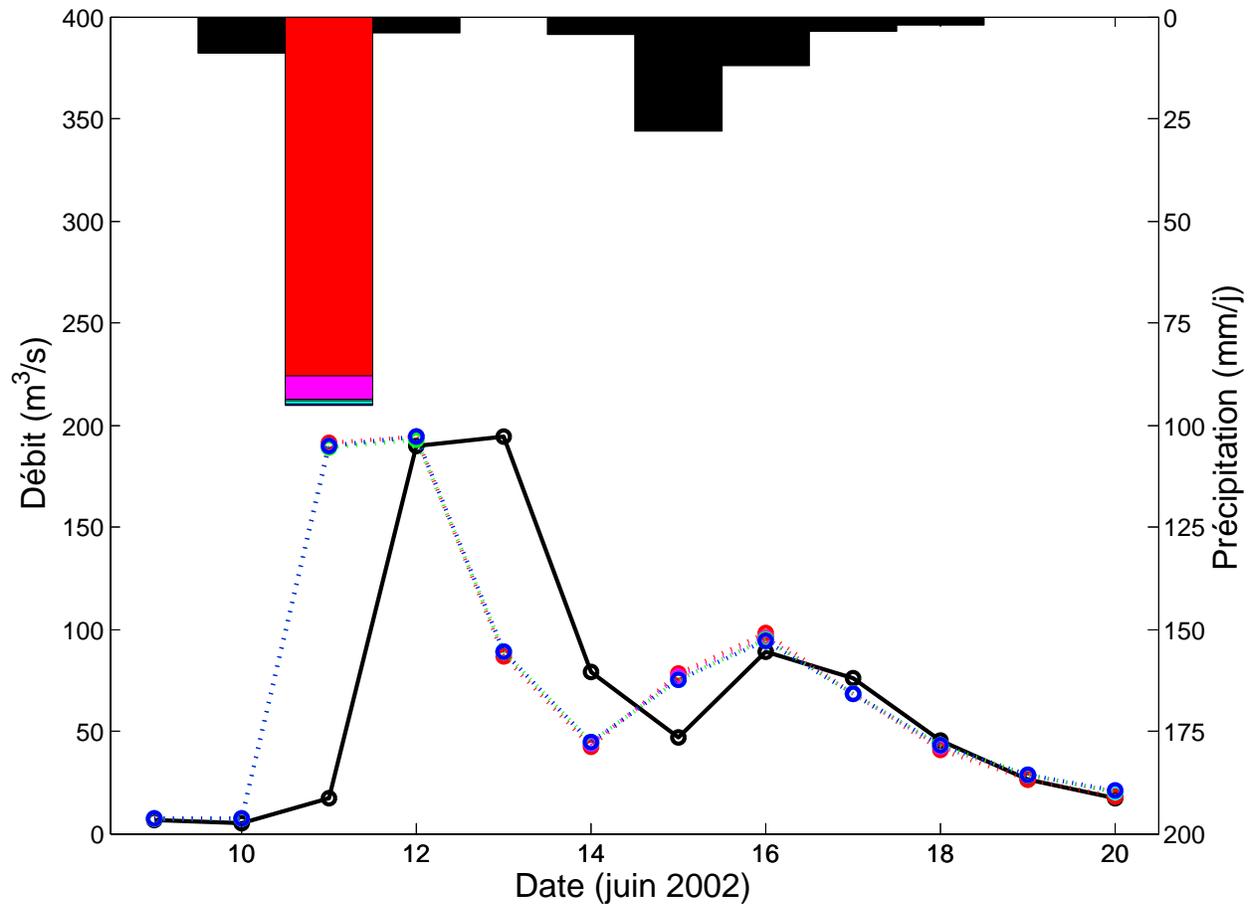


Figure 9.9 – Débits journaliers du 9 au 20 juin 2002 observés à la station hydrométrique (noir) et simulés à partir des pluies aux différentes résolutions : 4,4 (bleu), 8,8 (bleu pâle), 13,2 (vert), 17,6 (mauve) et 26,4 km (rouge). Sur l’axe inversé, la précipitation journalière durant cette période est présentée. Pour le 11 juin, la hauteur de pluie pour chaque résolution est présentée (mêmes couleurs que pour le débit). Pour les autres jours (noir), les mêmes hauteurs de précipitation pour toutes les résolutions sont utilisées.

de la précipitation désagrégée seront faites au Chapitre 10 avec des valeurs de  $Z_3$  et de la lame d’eau de l’HGM nettement différentes des valeurs de calage, mais quand même plausibles.

À l’étape 1, les tests effectués ont montré que la pointe du débit observé se produit une journée plus tard que la pointe du débit simulé, peu importe les paramètres utilisés (Figure 9.9). Ce résultat pourrait s’expliquer par le fait que la majeure partie de la pluie du 11 juin est tombée en fin de journée, alors que la pluie en entrée dans HYDROTEL est répartie uniformément sur toute la journée. Pour le calage, il a été décidé de décaler le débit observé d’une journée vers l’arrière (*e.g.* le débit observé le 12 juin devient le débit observé le 11

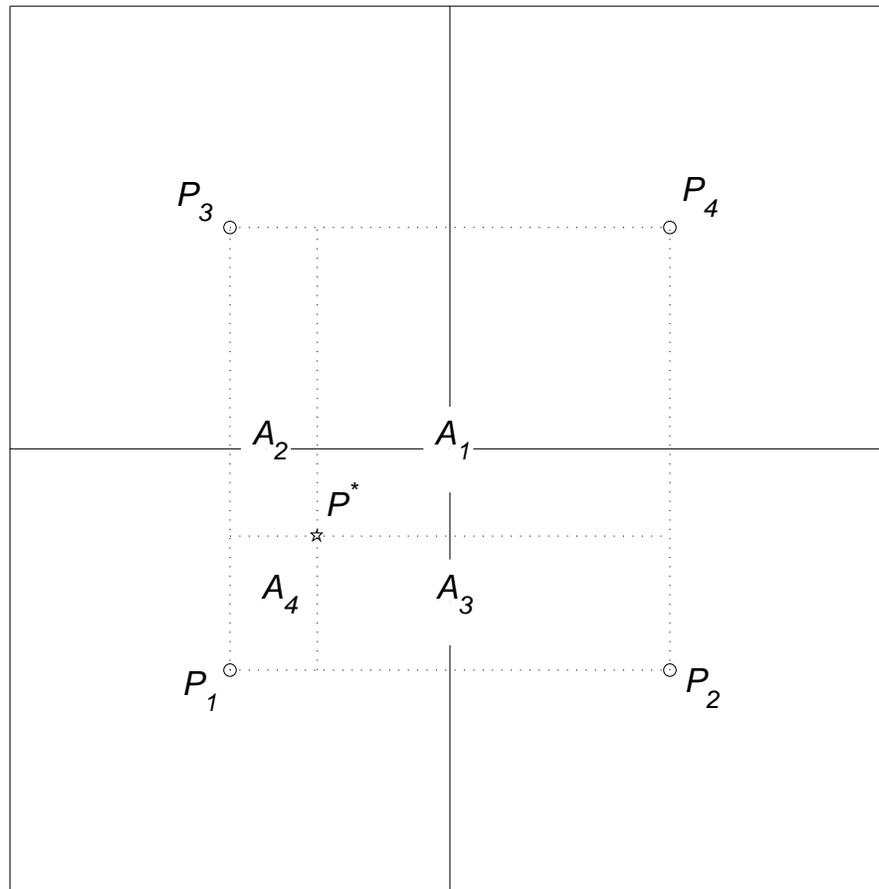
Tableau 9.5 – Valeurs estimées de certains paramètres d’HYDROTEL. Seuls les paramètres les plus déterminants pour un événement intense sont présentés. Les valeurs des paramètres qui ont été modifiées d’un calage à un autre sont en caractère gras.

Sous-modèle Résolution du calage	BV3C				Écoulement terrestre			
	$Z_1$ (m)	$Z_2$ (m)	$Z_3$ (m)	Coefficient de récession (m/h)	Lame (m)	Coefficient de Manning		
					Forêt	Eau	Autre	
4,4 km	0,1	<b>0,4</b>	<b>1,32</b>	$7,5 \times 10^{-5}$	0,005	0,30	0,03	0,10
8,8 km	0,1	0,4	<b>1,3</b>	$7,5 \times 10^{-5}$	0,005	0,30	0,03	0,10
13,2 km	0,1	0,4	<b>1,27</b>	$7,5 \times 10^{-5}$	0,005	0,30	0,03	0,10
17,6 km	0,1	0,4	<b>1,19</b>	$7,5 \times 10^{-5}$	0,005	0,30	0,03	0,10
26,4 km	0,1	<b>0,35</b>	<b>0,98</b>	$7,5 \times 10^{-5}$	0,005	0,30	0,03	0,10

juin). Ce biais temporel ne cause pas de souci puisque le débit observé n’est pas considéré comme le débit de référence pour le reste des analyses. L’important ici, c’est d’obtenir des valeurs de débits sur l’ensemble des UHRH à partir des valeurs de paramètres réalistes pour une application d’HYDROTEL, ce qui a été obtenu (Tableau 9.5). Il faut noter que dans le cadre d’une prévision hydrologique à court terme, il aurait été préférable de simuler sur un pas de temps plus petit que journalier pour une meilleure analyse du temps de réponse. Le pas de temps journalier a été choisi puisque c’est à ce pas de temps que le modèle de désagrégation a été développé.

Dans les applications classiques, le modèle est utilisé sur une période qui n’a pas servi au calage pour valider le modèle et le choix des paramètres. Dans le cas présent, c’est l’évaluation de la pluie désagrégée qui compte. Le calage des paramètres du modèle hydrologique n’a pas été fait dans le but de prédire les débits futurs, mais pour permettre une comparaison juste entre les différentes séries de précipitation. En principe, si la pluie désagrégée à une certaine résolution est semblable d’un point de vue statistique à la précipitation observée à la même résolution, le débit simulé à partir de la pluie désagrégée, qui sera nommé *débit désagrégé*, devrait être comparable au débit de référence (*i.e.* débit simulé à partir de la pluie de référence).

Pour chaque résolution, 1000 champs désagrégés sont produits et sont comparés avec deux autres sources de précipitation : (i) la *pluie interpolée* (ii) et la pluie désagrégée, mais répartie spatialement de manière totalement aléatoire, qui sera appelée *pluie aléatoire*. La



$$P^* = \frac{A_1 P_1 + A_2 P_2 + A_3 P_3 + A_4 P_4}{A_1 + A_2 + A_3 + A_4}$$

Figure 9.10 – Illustration de la méthode d’interpolation bilinéaire de la pluie agrégée utilisée. Les lignes pleines sont les frontières des tuiles, les cercles sont les centres des tuiles et l’étoile est un point à estimer. Les  $P_i$  ( $i \in \{1, 2, 3, 4\}$ ) sont les précipitations sur les tuiles et les  $A_i$  sont les superficies déterminées par les lignes pointillées. Un facteur multiplicatif est appliqué pour s’assurer que la précipitation sur la tuile soit conservée.

pluie interpolée est tout simplement la pluie agrégée ramenée à la résolution analysée en tenant compte des quatre points de grille les plus proches tel qu’illustré à la Figure 9.10. Ce type d’interpolation, appelé *interpolation bilinéaire*, ramène la pluie agrégée à la résolution voulue de manière simple, intuitive et économe, c’est-à-dire qu’elle nécessite un temps de calcul négligeable. Cependant, l’interpolation étant une méthode déterministe, un seul champ interpolé peut être produit. Comme la pluie interpolée sera analysée aux mêmes résolutions que la pluie désagrégée, les mêmes jeux de paramètres d’HYDROTEL peuvent être utilisés pour les deux sources de pluie.

La pluie aléatoire permet d'évaluer l'impact d'une pluie désagrégée à fine résolution de manière statistique, mais dont la structure spatiale est irréaliste. Un champ de pluie aléatoire est créé en prenant les  $x$  valeurs de pluie d'un champ désagrégé sur la tuile recouvrant le bassin ( $x = 4, 9, 16, 36$  ou  $144$  selon la résolution) et en les redistribuant une à une de manière totalement aléatoire. Cette façon de faire élimine la corrélation spatiale avec les pixels voisins. Cela implique l'élimination de la corrélation spatiale avec les tuiles voisines, comme pour les modèles de désagrégation multi-échelles ayant été appliqués jusqu'à maintenant en modélisation hydrologique. Il y aura 1000 champs aléatoires différents pour le 11 juin 2002 qui pourront être comparés aux 1000 champs désagrégés.

# Chapitre 10

## RÉSULTATS DES SIMULATIONS HYDROLOGIQUES

La pluie du 11 juin 2002 sur la tuile de 52,8 km couvrant le bassin de la rivière des Anglais est désagrégée aux résolutions 4,4, 8,8, 13,2, 17,6 et 26,4 km et utilisée en entrée dans le modèle hydrologique distribué HYDROTEL selon la procédure décrite au Chapitre 9. Tel que mentionné au chapitre précédent, l'objectif est d'évaluer la valeur ajoutée de la désagrégation spatiale de la pluie d'un événement intense sur la simulation hydrologique. Pour ce faire, les résultats obtenus à partir de la pluie désagrégée sont comparés avec ceux obtenus à partir des pluies interpolée et aléatoire (Chapitre 9). La valeur ajoutée peut se présenter sous deux aspects :

- L'exactitude : la capacité de bien reproduire en moyenne ce qui a été observé ;
- La précision : la valeur informative, du point de vue de gestion, de l'ensemble des valeurs possibles produites.

Chaque aspect est d'abord évalué à partir de la pluie du 11 juin 2002 seulement. Ensuite, l'écoulement, soit le débit redistribué sur l'ensemble de la superficie drainée, est analysé en fonction de la pluie pour les sept sous-bassins. L'analyse sur les sous-bassins permet d'exprimer l'écoulement en fonction de la superficie drainée et aussi d'évaluer l'impact d'autres facteurs, soit la teneur en eau initiale du sol, le type de sol et l'occupation du sol. L'impact sur l'écoulement du choix des valeurs des paramètres de calage les plus influents est aussi analysé (Figure 10.1).

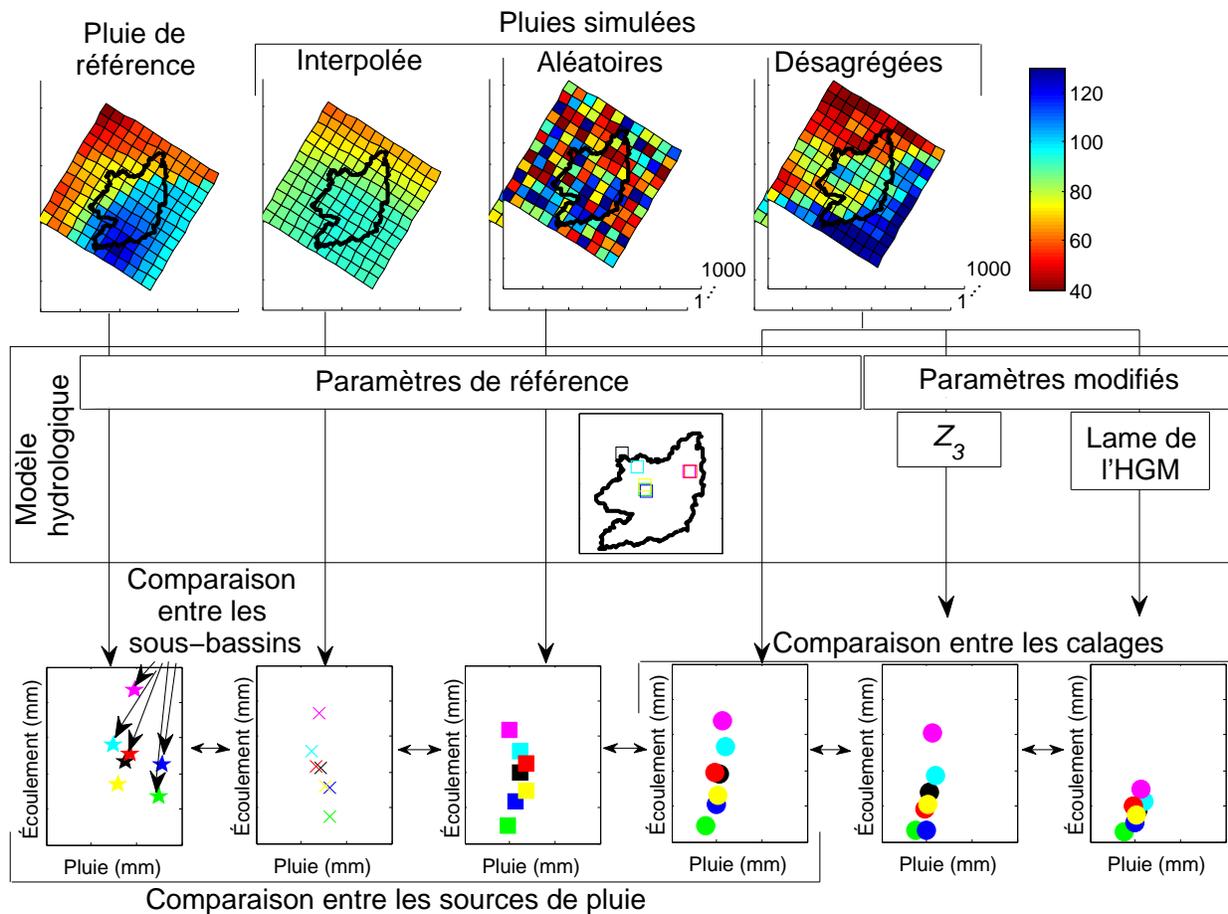


Figure 10.1 – Schéma illustrant la procédure utilisée pour l'évaluation de la valeur ajoutée de la désagrégation. La résolution présentée est de 4,4 km ; la procédure est la même pour les autres résolutions.

## 10.1 Exactitude

Dans cette section, la pluie de référence (observée) est d'abord comparée aux pluies simulées (désagrégée, aléatoire et interpolée). Par la suite, le même exercice est fait pour l'écoulement.

### 10.1.1 Exactitude de la pluie désagrégée

Les Figures 10.2 et 10.3 présentent les hauteurs de pluie du 11 juin 2002 observées comparées aux hauteurs de pluie désagrégées et aléatoires sur chaque sous-bassin. Ces deux figures contiennent aussi les hauteurs de pluie interpolées. Peu importe la résolution et la source de pluie simulée, la pluie de pointe, *i.e.* la plus grande valeur de pluie sur un sous-bassin, est sous-estimée. Néanmoins, l'intervalle 95 % de la pluie de pointe désagrégée couvre ou passe très près de la valeur de référence, sauf à 8,8 km (Figure 10.2).

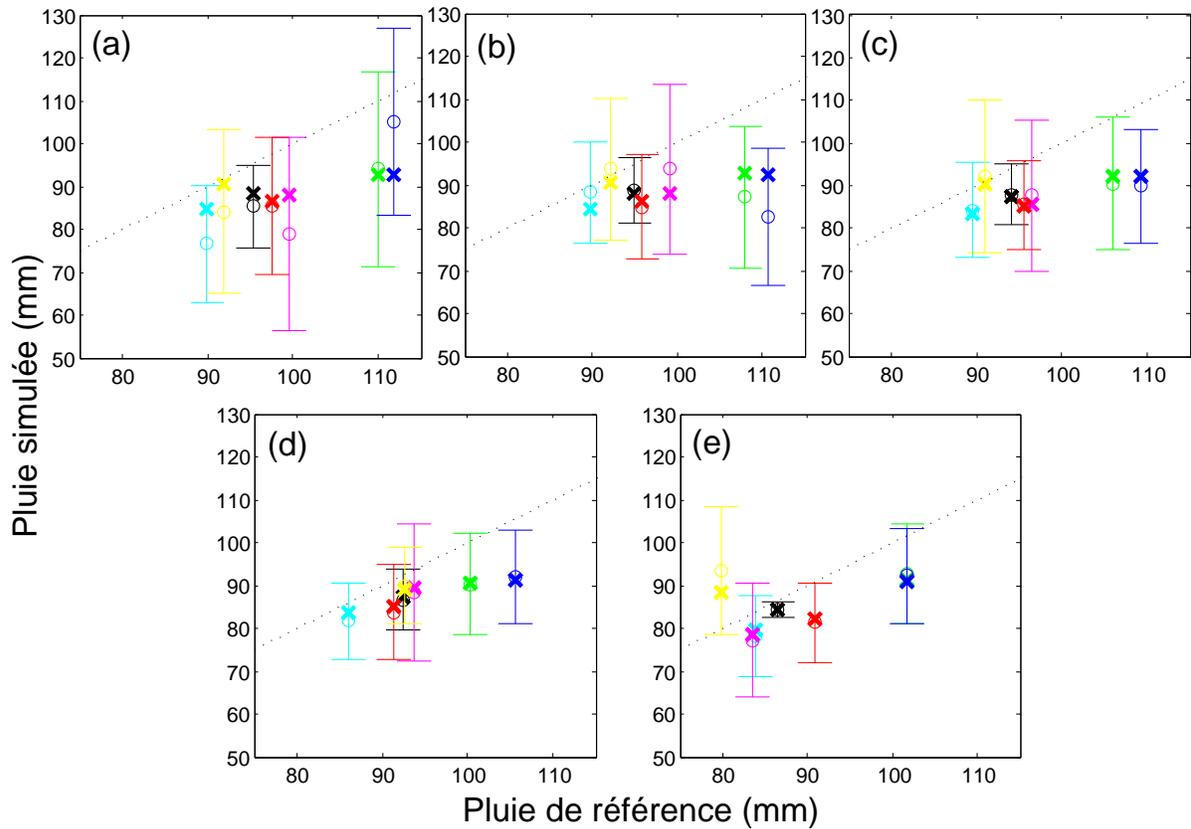


Figure 10.2 – Moyennes spatiales des pluies désagrégées versus celles des pluies de référence à résolution (a) 4,4, (b) 8,8, (c) 13,2, (d) 17,6 et (e) 26,4 km pour le 11 juin 2002 sur les sept sous-bassins. Les lignes couvrent du 2,5<sup>e</sup> au 97,5<sup>e</sup> percentiles des 1000 réalisations désagrégées, les cercles sont les médianes désagrégées et les « x » sont les pluies interpolées. Chaque couleur représente un sous-bassin (Figure 9.3). La ligne pointillée est la ligne pour laquelle la pluie simulée égale la pluie de référence.

En comparant les sources de pluie, il ressort qu'en général, la pluie interpolée (« x ») est très près de la pluie médiane désagrégée (Figure 10.2). La pluie aléatoire est celle qui performe le moins bien en moyenne (Figure 10.3). Comme la pluie aléatoire ne tient aucunement compte de la structure spatiale de l'événement, son espérance est identique à la moyenne sur la tuile (83,2 mm) pour tous les sous-bassins. Cela explique pourquoi la pluie désagrégée performe mieux pour la pluie de pointe. Il n'y a qu'à la résolution de 8,8 km où la pluie de pointe n'est pas mieux représentée par la pluie désagrégée que par la pluie aléatoire. Un examen des champs désagrégés montre qu'à 8,8 km, la pointe passe trop souvent à côté du bassin.

La Figure 10.4 montre le biais de la pluie désagrégée sur chaque sous-bassin en fonction de la taille du pixel. Il n'est pas possible d'affirmer à partir des résultats de cette figure que

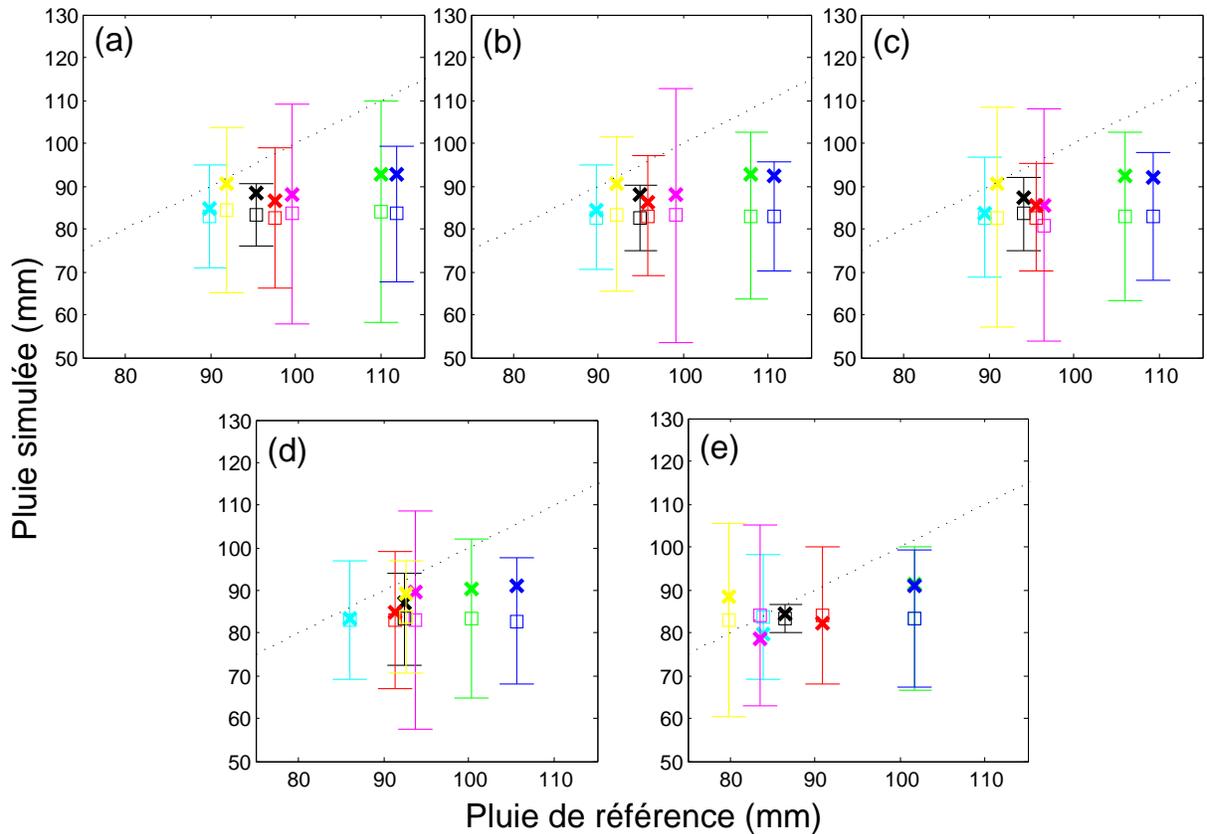


Figure 10.3 – Moyennes spatiales des pluies aléatoires versus celles des pluies de référence à résolution (a) 4,4, (b) 8,8, (c) 13,2, (d) 17,6 et (e) 26,4 km pour le 11 juin 2002 sur les sept sous-bassins. Les lignes couvrent du 2,5<sup>e</sup> au 97,5<sup>e</sup> percentiles des 1000 réalisations aléatoires, les carrés sont les médianes aléatoires et les « x » sont les pluies interpolées. Chaque couleur représente un sous-bassin (Figure 9.3). La ligne pointillée est la ligne pour laquelle la pluie simulée égale la pluie de référence.

la performance moyenne de la désagrégation s’améliore avec la résolution spatiale. Il faut considérer que la pluie désagrégée à une résolution  $x$  est comparée à la pluie de référence à résolution  $x$ . Ainsi, plus la taille du pixel analysée est grande, plus les valeurs comparées (pluie moyenne désagrégée et pluie de référence) se rapprochent de la valeur sur la tuile, commune aux deux sources. Ce constat est vrai aussi pour les pluies aléatoire et interpolée (pas présenté par souci de concision).

### 10.1.2 Exactitude de l’écoulement désagrégé

La Figure 10.5 permet de voir l’impact du biais de la pluie sur le biais de l’écoulement du 11 juin 2002. Rappelons que pour une résolution donnée, les simulations hydrologiques sont

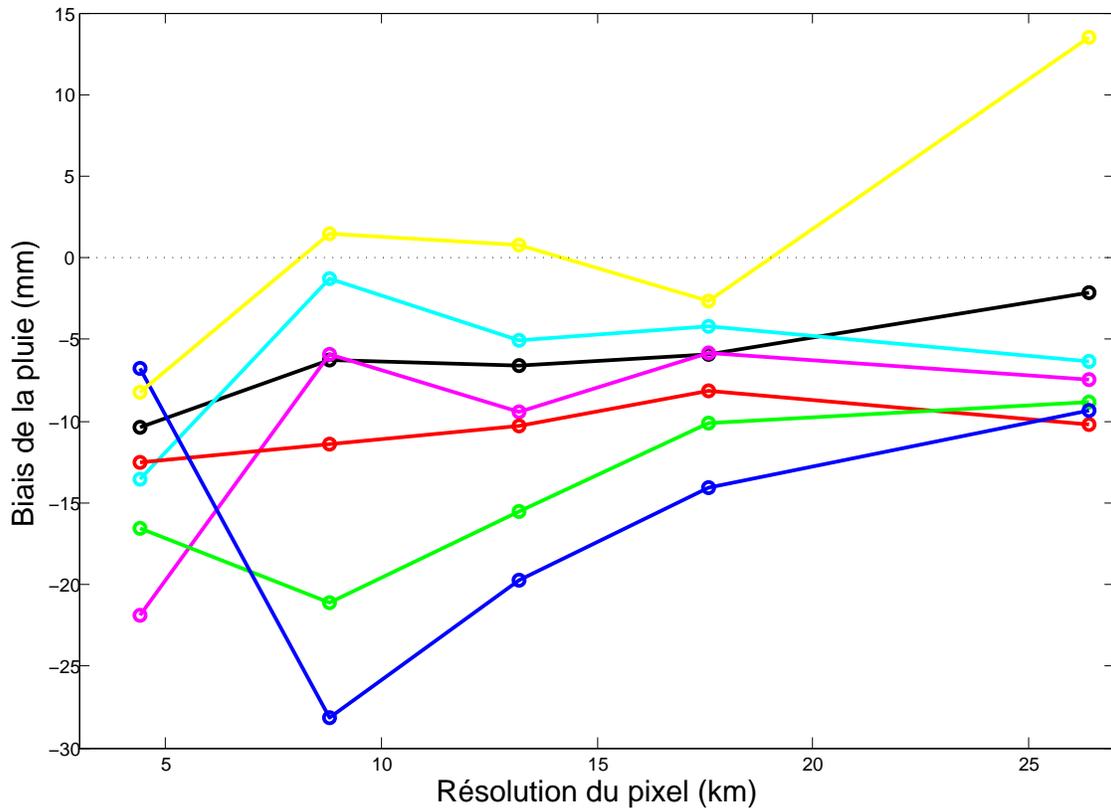


Figure 10.4 – Biais de la pluie désagrégée par rapport à la pluie de référence du 11 juin 2002 pour les sept sous-bassins en fonction de la résolution du pixel. Chaque couleur représente un sous-bassin (Figure 9.3).

faites à partir des mêmes valeurs de paramètres pour toutes les sources de pluie (Tableau 9.5). Pour un calage donné, le biais de l'écoulement est fortement lié au biais de la pluie. La relation entre les deux biais est linéaire et presque identique pour tous les sous-bassins, sauf pour les deux plus petits : le troisième (mauve) et le cinquième (vert). Le troisième sous-bassin est majoritairement argileux et le cinquième est essentiellement composé de sable et de loam. Une analyse de l'écoulement pour les tronçons de tête, présentée à l'Annexe F, a montré que pour cet événement, le type de sol influence fortement la vitesse à laquelle s'amorce le ruissellement. À pluie d'intensité égale, le ruissellement sur un sol argileux débutera plus tôt que sur un sol sableux. C'est la raison pour laquelle un biais négatif de la pluie a plus d'impact sur le biais de l'écoulement pour le troisième que pour le cinquième sous-bassin.

La Figure 10.6 présente aussi la relation entre le biais de la pluie et le biais de l'écoulement, mais cumulé pour les 11 et 12 juin 2002. Les résultats sont comparables à ceux de la Figure 10.5, à l'exception que la relation varie moins entre les sous-bassins. La différence de la

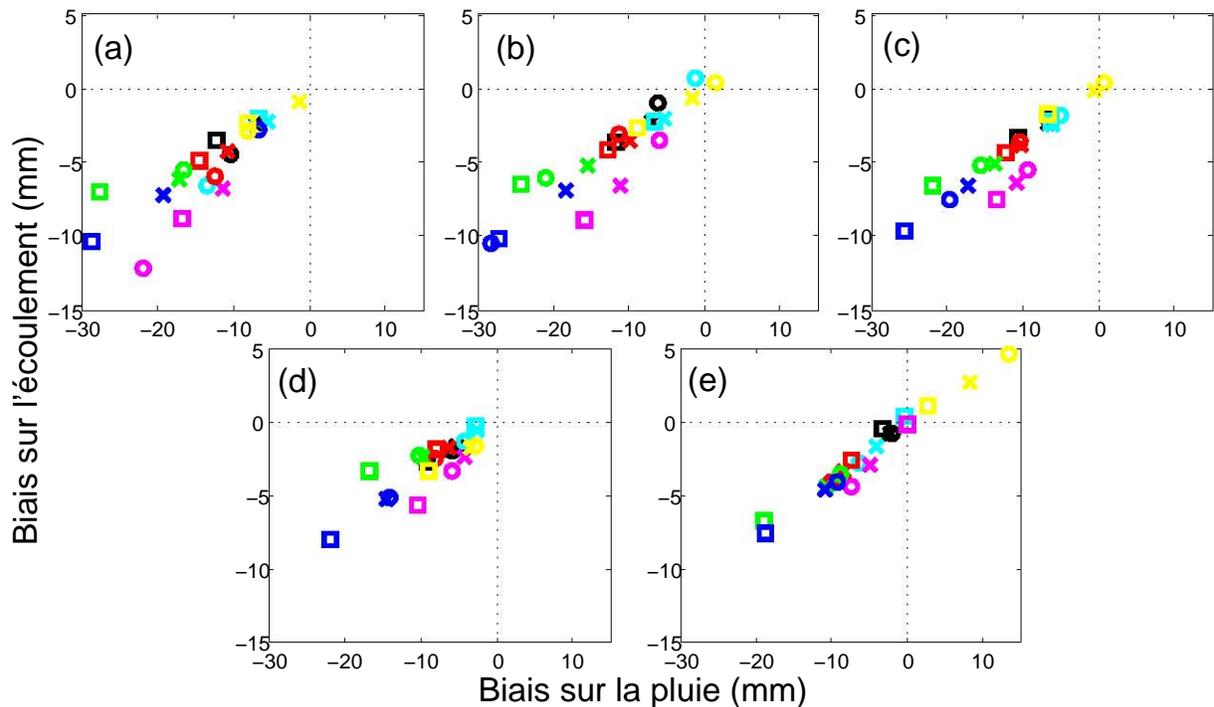


Figure 10.5 – Biais sur l'écoulement du 11 juin en fonction du biais sur la pluie du 11 juin 2002 à résolution (a) 4,4, (b) 8,8, (c) 13,2, (d) 17,6 et (e) 26,4 km pour les trois sources de pluie et les sept sous-bassins (Figure 9.3). Sources de pluie : interpolée (« x »), désagrégée (cercles) et aléatoire (carrés).

rapidité de la réponse entre les types de sol est moins influente lorsque l'échelle temporelle augmente.

L'étape suivante est l'analyse de l'impact des valeurs de paramètres du modèle HYDROTEL. Pour un événement de pluie intense, il est supposé que les deux paramètres d'HYDROTEL les plus sensibles sont la profondeur de la troisième couche de sol ( $Z_3$ ) et la lame d'eau de référence de l'HGM. Dans Indarto (2002), ces deux paramètres ont été les plus sensibles sur le bassin d'étude français. La Figure 10.7 présente l'écoulement simulé du 11 juin 2002 à partir de la pluie à résolution 4,4 et 26,4 km et de trois jeux de paramètres : (i) ceux utilisés pour la pluie à 4,4 km (Tableau 9.5), (ii) comme (i), mais avec  $Z_3 = 4$  m au lieu de 1,32 m et (iii) comme (i), mais avec lame = 0,0005 m au lieu de 0,005 m.

La Figure 10.7 illustre que pour un calage donné, la différence entre l'écoulement moyen simulé à partir de la pluie à 4,4 et à 26,4 km est faible, à l'exception du sous-bassin 6 (bleu), qui a reçu le plus de pluie. La variation de  $Z_3$  de 1,32 à 4 m donne une différence d'écoulement d'environ 10 mm en moyenne, mais qui varie substantiellement entre les sous-bassins. Pour le

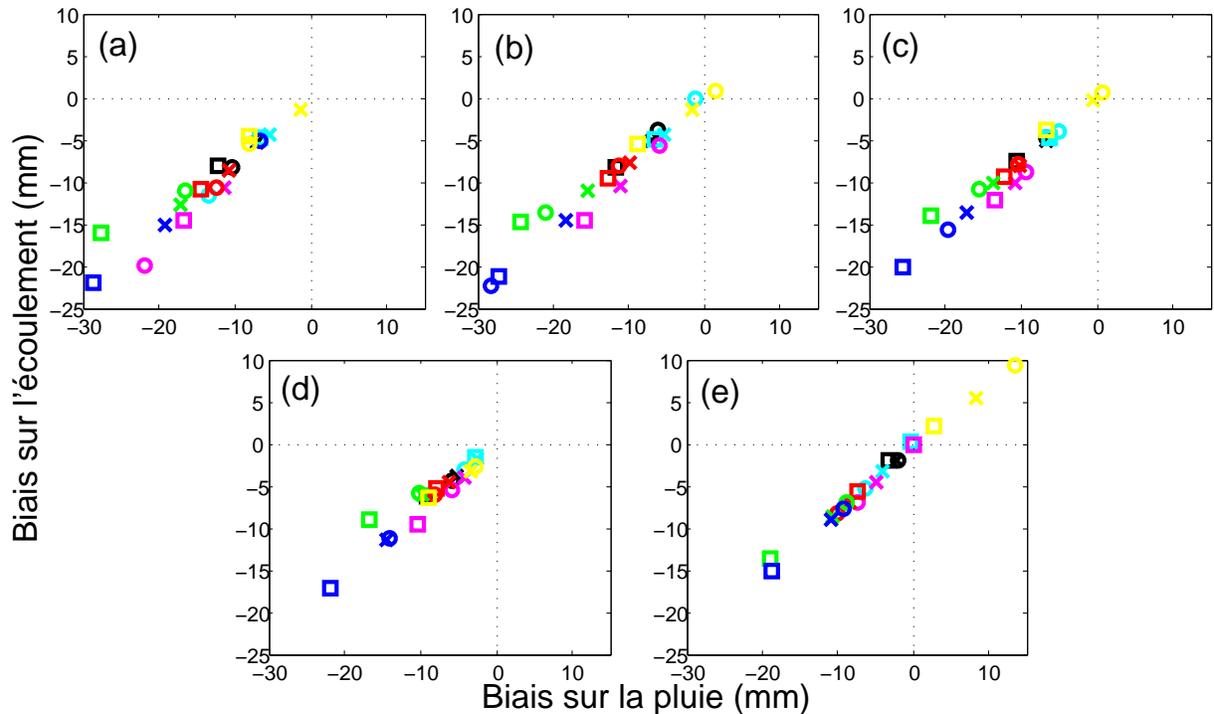


Figure 10.6 – Biais sur l'écoulement des 11 et 12 juin en fonction du biais sur la pluie du 11 juin 2002 à résolution (a) 4,4, (b) 8,8, (c) 13,2, (d) 17,6 et (e) 26,4 km pour les trois sources de pluie et les sept sous-bassins (Figure 9.3). Sources de pluie : interpolée (« x »), désagrégée (cercles) et aléatoire (carrés).

sous-bassin 6 (bleu) et la pluie désagrégée à 4,4 km, il y a une différence entre les écoulements simulés à partir des deux valeurs de  $Z_3$  d'environ 15 mm (75 %; Figures 10.7(a) et 10.7(b)). Cela est causé par le fait que lorsque la valeur de  $Z_3$  augmente, le ruissellement de surface prend plus de temps à se produire (Annexe F) et donc une plus grande fraction de la pluie se retrouve sous forme d'écoulements retardé et de base. C'est surtout vrai pour les UHRH sableux; pour les UHRH argileux, le ruissellement de surface prend aussi plus de temps à se produire si  $Z_3$  est grand, mais l'impact est moins important puisqu'ils saturent plus rapidement (Annexe F). La modification de la lame de référence de l'HGM a réduit d'environ de moitié l'écoulement, peu importe le sous-bassin (Figures 10.7(c) et 10.7(f)). Les constats sont les mêmes pour l'écoulement des 11 et 12 juin 2002 cumulé (pas présenté).

## 10.2 Précision

Dans cette section, la capacité de l'ensemble de valeurs produit par la désagrégation à fournir une information précise est analysée. Deux indicateurs sont utilisés pour définir la précision

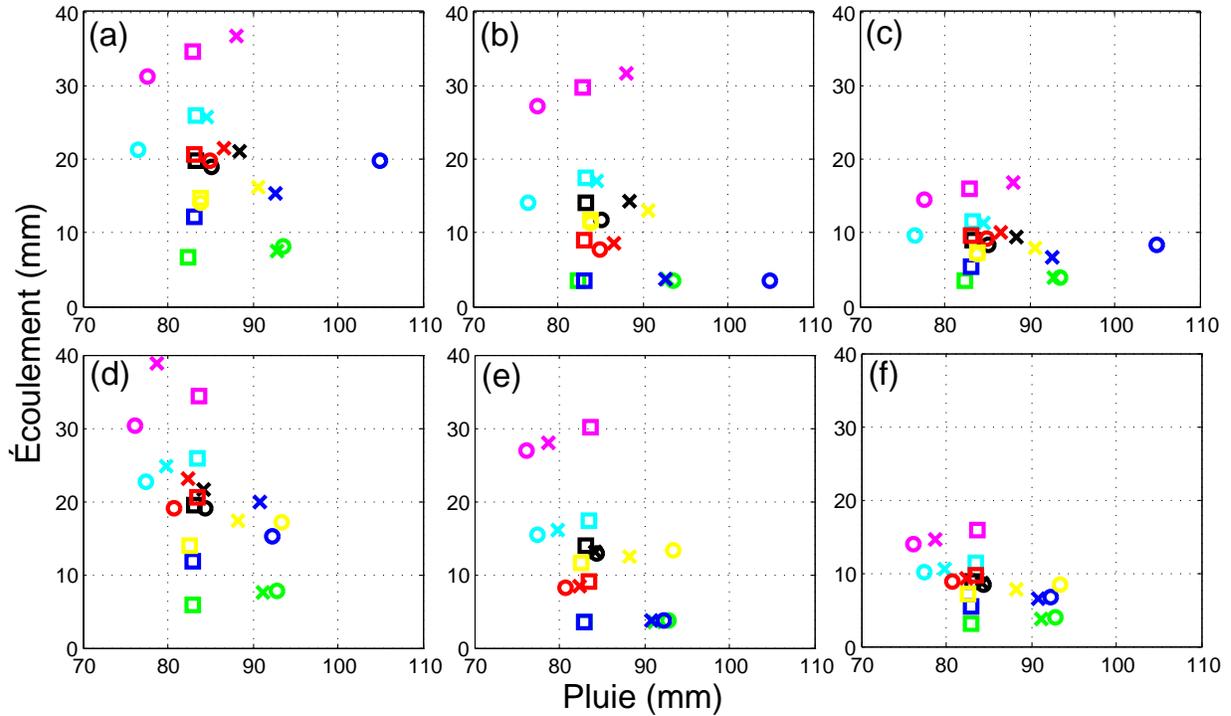


Figure 10.7 – Écoulement moyen du 11 juin simulé (a et d) à partir des paramètres calés à 4,4 km, (b et e) à partir d’une valeur de  $Z_3$  de 4 m et (c et f) à partir d’une lame de l’HGM de 0,0005 m en fonction de la pluie moyenne du 11 juin 2002 à résolution 4,4 (haut) et 26,4 km (bas) pour les trois sources de pluie et les sept sous-bassins (Figure 9.3). Sources de pluie : interpolée (« x »), désagrégée (cercles) et aléatoire (carrés).

d’un intervalle de valeur : (i) le seuil observé et (ii) la largeur de l’*intervalle de confiance* défini par les valeurs simulées. L’intervalle de confiance  $\alpha$  % est défini comme étant l’ensemble des valeurs contenues entre les  $\left(\frac{100-\alpha}{2}\right)^{\text{ième}}$  et  $\left(100 - \frac{100-\alpha}{2}\right)^{\text{ième}}$  percentiles simulés. Dans ce chapitre, des intervalles de 95 % sont utilisés. Le seuil observé (*p-value* en anglais) est la probabilité qu’une valeur simulée soit au moins aussi loin de la médiane simulée que l’est la valeur observée. Le seuil observé s’exprime de la façon suivante :

$$2 \min(\mathbb{P}[X_{\text{sim}} \leq x_{\text{obs}}], \mathbb{P}[X_{\text{sim}} \geq x_{\text{obs}}]) \quad (10.1)$$

où  $x_{\text{obs}}$  est la valeur observée et  $X_{\text{sim}}$  est la variable aléatoire représentant une valeur simulée. En pratique, les probabilités de l’équation (10.1) sont estimées en calculant la proportion des valeurs simulées qui sont plus petites et plus grandes que la valeur observée. Le seuil observé peut être vu comme un indicateur de l’exactitude d’un modèle tenant compte de la

variabilité des simulations. Il est usuel de considérer la différence entre les données simulées et la donnée de référence comme étant statistiquement significative si le seuil observé est plus petit que 0,05. Un seuil observé plus petit que 0,05 implique que la valeur de référence est en dehors de l'intervalle de confiance 95 %.

En plus de contenir la valeur de référence, un intervalle de confiance se doit de ne pas être inutilement trop large. Un très large intervalle a une plus grande probabilité de contenir la valeur de référence, mais n'amène pas une information précise sur la variable analysée. Il faut noter que plus l'ensemble des valeurs simulées est centré sur la valeur de référence, moins l'intervalle aura besoin d'être large pour la contenir.

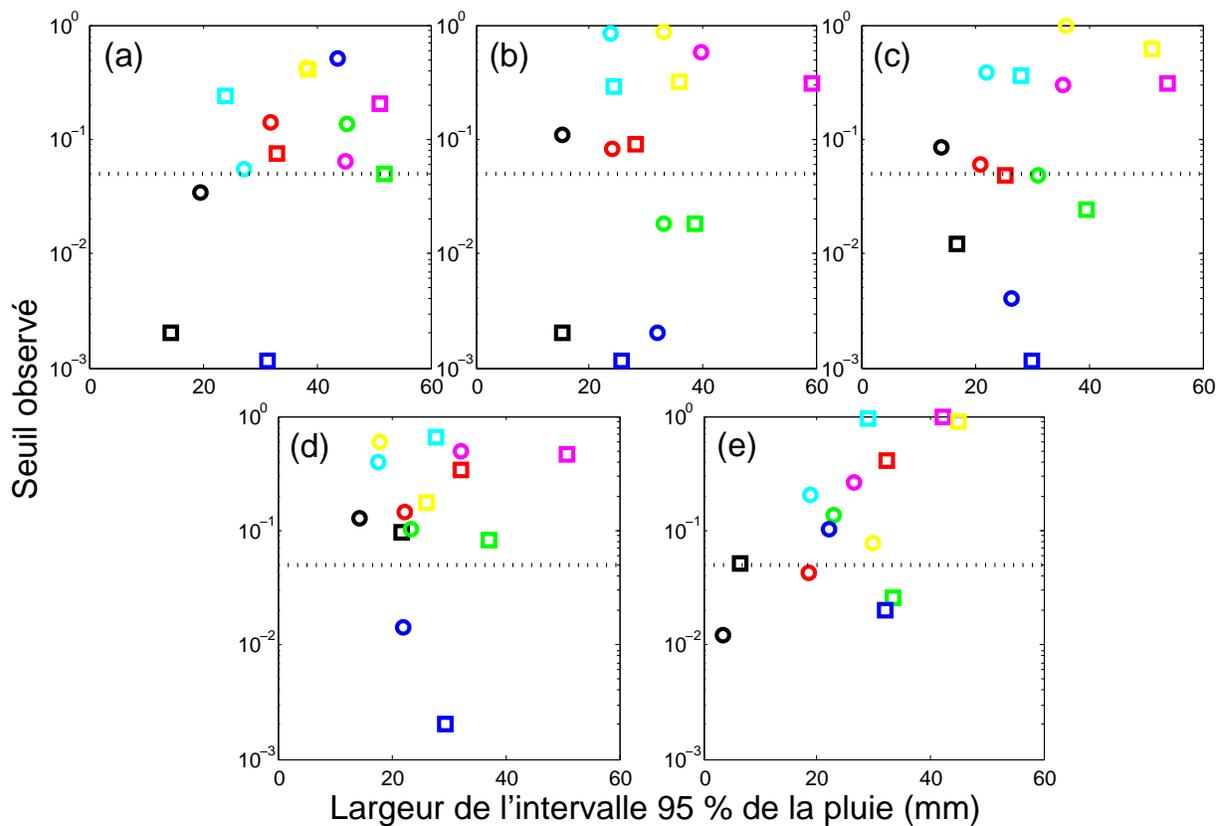


Figure 10.8 – Seuil observé et largeur de l'intervalle de confiance des pluies désagrégée (cercles) et aléatoire (carrés) à résolution (a) 4,4, (b) 8,8, (c) 13,2, (d) 17,6 et (e) 26,4 km pour le 11 juin 2002 sur les sept sous-bassins. La ligne pointillée représente un seuil observé de 0,05. Les valeurs de la pluie aléatoire (carrés) placées sur l'axe des abscisses ont des seuils observés inférieurs à 0,001.

Comme à la sous-section précédente, la pluie du 11 juin 2002 est d'abord analysée. L'écoulement est ensuite analysé en fonction de la pluie et par la suite, en fonction des autres facteurs. L'ensemble des valeurs désagrégées est comparé à l'ensemble des valeurs aléatoires. La pluie interpolée n'est pas utilisée dans cette section puisqu'elle ne produit qu'une seule valeur.

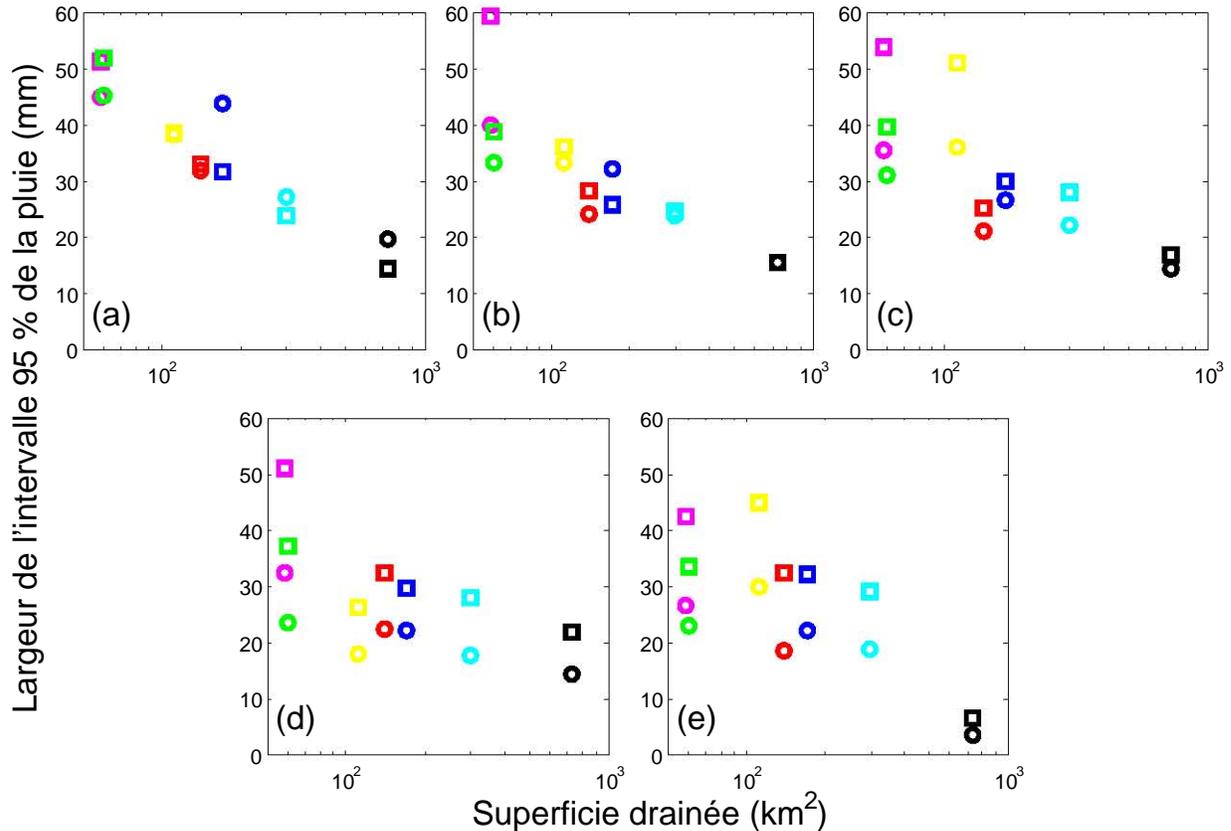


Figure 10.9 – Largeur de l'intervalle de confiance des pluies désagrégée (cercles) et aléatoire (carrés) à résolution (a) 4,4, (b) 8,8, (c) 13,2, (d) 17,6 et (e) 26,4 km pour le 11 juin 2002 en fonction de la superficie des sept sous-bassins.

### 10.2.1 Précision de la pluie désagrégée

La Figure 10.8 présente les seuils observés et la largeur des intervalles de confiance 95 % (différence entre le 97,5<sup>e</sup> et le 2,5<sup>e</sup> percentiles simulés) de la pluie du 11 juin 2002 désagrégée et aléatoire sur les sept sous-bassins pour les cinq résolutions. Pour ce type de graphes, il est souhaitable d'avoir des données le plus près du coin supérieur gauche, c'est-à-dire qu'il est souhaitable d'avoir de grands seuils observés (au moins plus grands que 0,05) et de petites largeurs d'intervalles. Cependant, il est très difficile d'obtenir de tels

résultats puisque, règle générale, plus l'intervalle de confiance est petit, plus le seuil observé risque d'être petit aussi. Ceci étant dit, la Figure 10.8 montre que la pluie désagrégée produit de meilleurs intervalles que la pluie aléatoire puisqu'elle donne des seuils observés moins souvent plus petits que 0,05 tout en produisant des intervalles généralement plus petits.

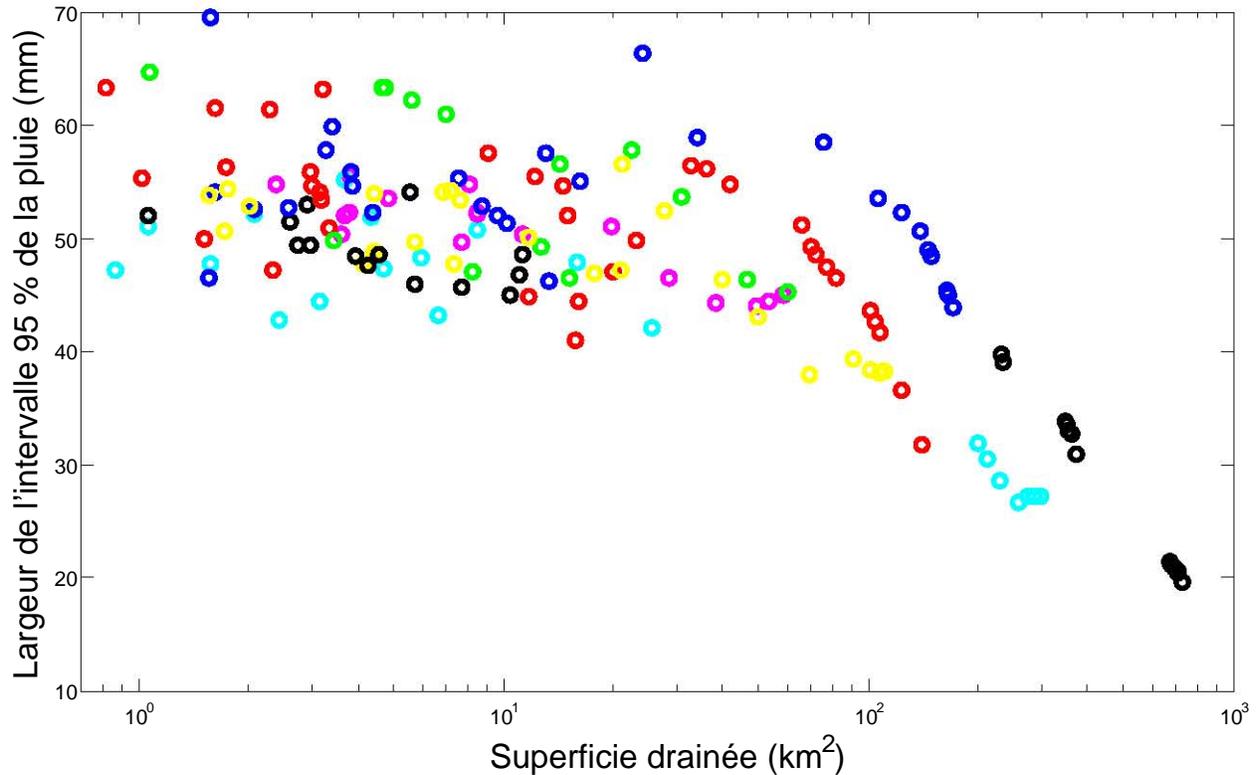


Figure 10.10 – Largeur de l'intervalle de confiance de la pluie désagrégée du 11 juin 2002 à résolution 4,4 km en fonction de la superficie drainée par les 167 tronçons. La couleur indique sur lequel des sept sous-bassins le tronçon est situé (Figure 9.3).

Cependant, il arrive trop souvent que l'intervalle de confiance 95 % de la pluie désagrégée ne contienne pas la valeur de référence, c'est-à-dire que le seuil observé de la pluie désagrégée soit plus petit que 0,05. Le sous-bassin 6 (bleu) est le moins bien simulé; le seuil observé est plus petit que 0,05 pour la pluie désagrégée aux résolutions 8,8, 13,2 et 17,6 km. Le seuil observé est aussi plus petit que 0,05 pour le sous-bassin 5 (vert) aux résolutions 8,8 et 13,2 km et pour le sous-bassin 4 (rouge) à la résolution 26,4 km. Pour tous ces sous-bassins, la faible valeur du seuil observé est causée par la sous-estimation de la pluie de pointe par le modèle de désagrégation (Figure 10.2).

Le seuil observé est aussi trop petit pour le sous-bassin 1 (exutoire du bassin des Anglais, noir) aux résolutions 4,4 et 26,4 km. Dans ce cas, la faible valeur du seuil observée est aussi causée par la faible largeur de l'intervalle de confiance, surtout à la résolution 26,4 km. La Figure 10.9 montre que la largeur de l'intervalle diminue avec la superficie drainée. Carpenter et Georgakakos (2006) ont suggéré qu'il existait une relation linéaire entre la variabilité à fine échelle de la précipitation et le logarithme de la superficie drainée. Comme sept sous-bassins ne sont pas suffisants pour caractériser précisément la relation entre la largeur de l'intervalle et le logarithme de la superficie drainée, la Figure 10.9 est reprise, mais pour tous les 167 tronçons du bassin (tous les UHRH, sauf les lacs) et seulement pour la pluie désagrégée à résolution 4,4 km (Figure 10.10). La relation est approximativement linéaire comme dans Carpenter et Georgakakos (2006) pour les superficies drainées supérieures à environ 50 km<sup>2</sup>, surtout à l'intérieur d'un sous-bassin. Pour les plus petites superficies, la largeur se stabilise. Le même phénomène est observé pour les autres tailles de pixels (résultats non présentés par souci de concision). À noter que même si l'impact de la désagrégation diminue avec la superficie drainée, il est toujours présent à l'exutoire du bassin de la rivière des Anglais (730 km<sup>2</sup>) où la largeur de l'intervalle de confiance 95 % de la pluie désagrégée atteint environ 20 mm (Figure 10.10), soit un peu plus de 20 % de la pluie de référence.

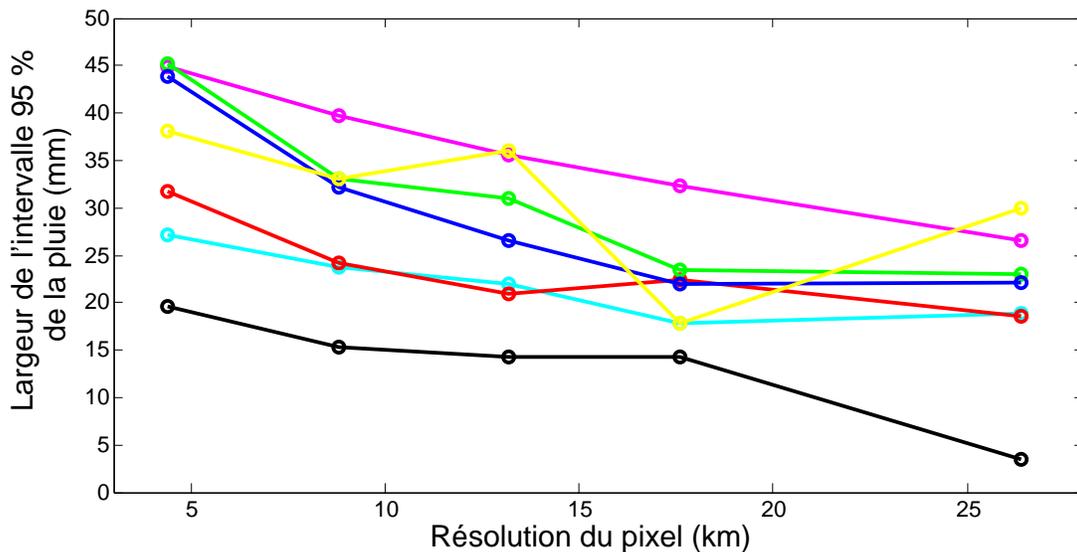


Figure 10.11 – Largeur de l'intervalle de confiance de la pluie désagrégée pour le 11 juin 2002 pour les sept sous-bassins en fonction de la résolution du pixel.

La Figure 10.11 montre que la largeur de l'intervalle de confiance diminue aussi avec une augmentation de la taille des pixels. Cela s'explique par le fait que plus la taille du pixel

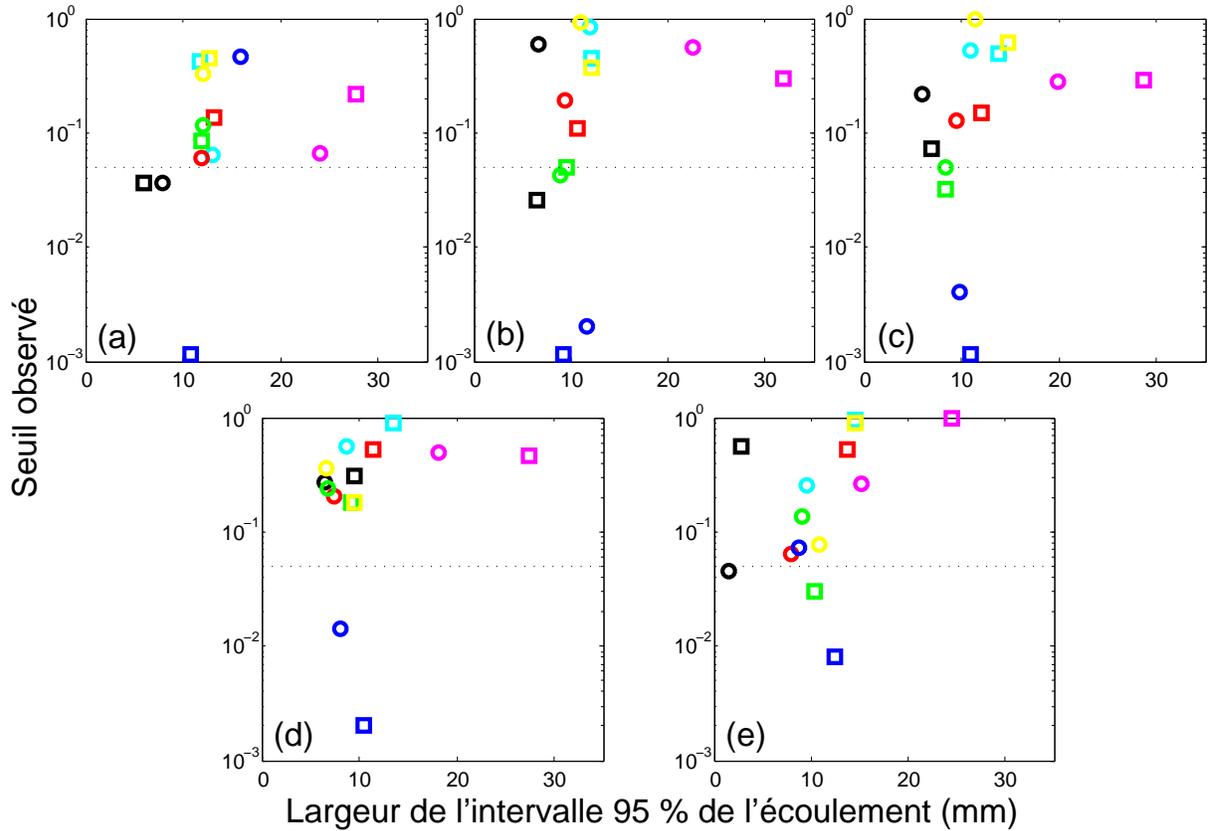


Figure 10.12 – Seuil observé et largeur de l’intervalle de confiance des écoulements désagrégé (cercles) et aléatoire (carrés) à résolution (a) 4,4, (b) 8,8, (c) 13,2, (d) 17,6 et (e) 26,4 km pour le 11 juin 2002 sur les sept sous-bassins. Les valeurs de l’écoulement aléatoire (carrés) placées sur l’axe des abscisses ont des seuils observés inférieurs à 0,001.

augmente, plus les données désagrégées sont contraintes. En effet, pour les plus petits pixels, chaque tuile compte  $12 \times 12 - 1 = 143$  degrés de liberté (*i.e.* le nombre de valeurs désagrégées moins la contrainte qu’il faut conserver la précipitation moyenne sur la tuile). Pour les autres résolutions, chaque tuile compte 35, 15, 8 et 3 degrés de liberté.

### 10.2.2 Précision de l’écoulement désagrégé

Dans cette sous-section, seul l’écoulement du 11 juin 2002, soit le jour de l’événement, est présenté. Les résultats pour l’écoulement cumulé des 11 et 12 juin sont très similaires. La Figure 10.12 présente la largeur de l’intervalle de confiance 95 % en fonction du seuil observé pour l’écoulement simulé du 11 juin 2002. Les seuils observés sont très semblables à ceux de la Figure 10.8, illustrant que la position de la valeur de l’écoulement de référence par rapport à l’intervalle des valeurs d’écoulement désagrégées est très fortement liée à la position de la valeur de pluie de référence par rapport à l’intervalle des valeurs de pluie désagrégées. Par

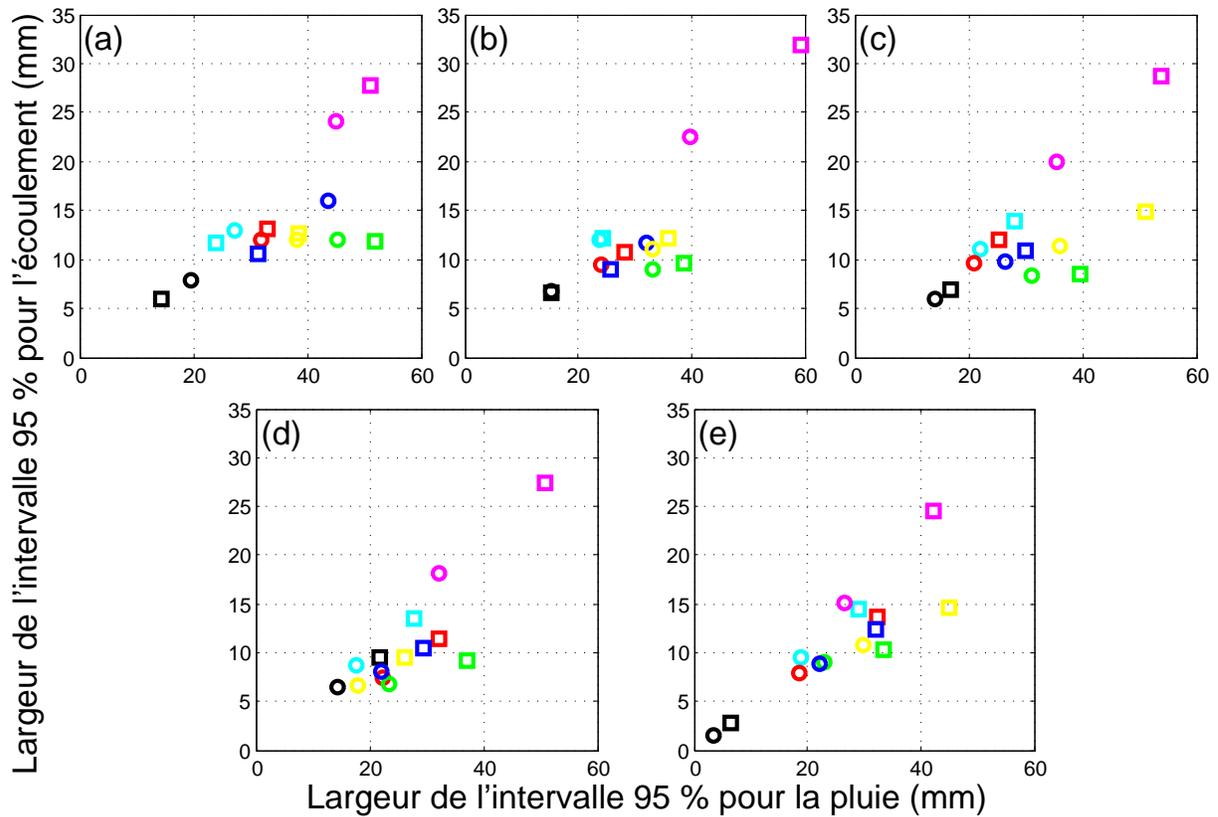


Figure 10.13 – Largeur de l'intervalle de confiance des écoulements en fonction de la largeur de l'intervalle de confiance des pluies désagrégée (cercles) et aléatoire (carrés) à résolution (a) 4,4, (b) 8,8, (c) 13,2, (d) 17,6 et (e) 26,4 km pour le 11 juin 2002 sur les sept sous-bassins.

exemple, si la pluie de référence est égale au 20<sup>e</sup> percentile de la pluie désagrégée, l'écoulement de référence devrait être approximativement égal au 20<sup>e</sup> percentile de l'écoulement désagrégé.

La Figure 10.13 illustre que la largeur de l'intervalle de l'écoulement dépend fortement de la largeur de l'intervalle de la pluie du 11 juin 2002, mais aussi du type de sol. Pour les sous-bassins sablonneux (5 et 7, vert et jaune), la réponse est moins rapide; l'intervalle des valeurs d'écoulement est donc plus petit. Il ressort aussi qu'en général, les largeurs des intervalles désagrégés sont plus petites que celles aléatoires, sauf sur les sous-bassins où la pluie de référence était plus forte.

La Figure 10.14, illustre la largeur de l'intervalle de l'écoulement désagrégé à 4,4 km sur les 167 tronçons du bassin en fonction de la superficie drainée. La variabilité entre les tronçons est beaucoup plus grande pour l'écoulement que pour la pluie (Figure 10.10); c'est vraisemblablement causé par les différents types de sol (Annexe F). Pour vérifier

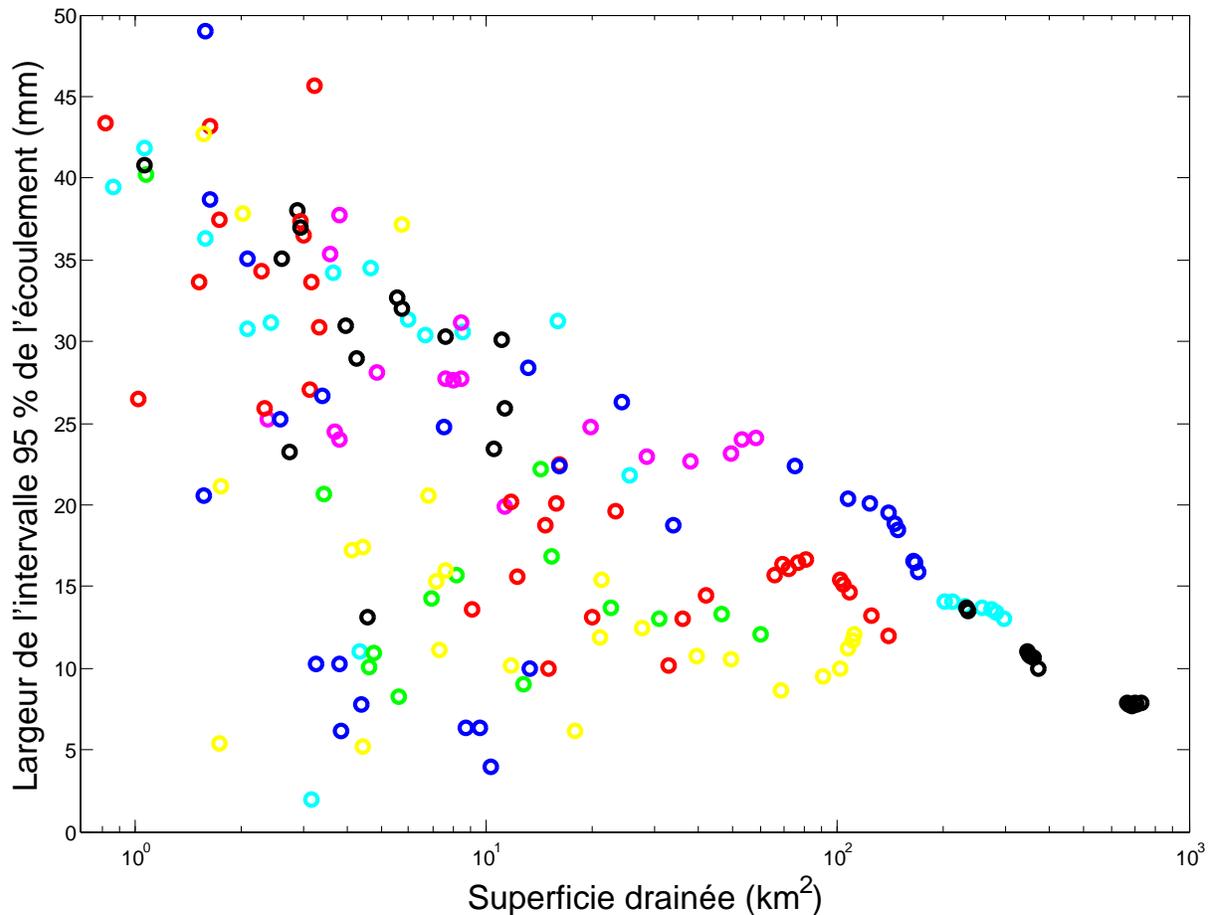


Figure 10.14 – Largeur de l’intervalle de confiance de l’écoulement désagrégé du 11 juin 2002 à résolution 4,4 km en fonction de la superficie des 167 tronçons. La couleur indique sur lequel des sept sous-bassins le tronçon est situé (Figure 9.3).

cela, la largeur de l’intervalle de l’écoulement désagrégé à 4,4 km est présentée pour les 81 tronçons de tête en fonction de la superficie drainée et du type de sol (Figure 10.15). Le type de sol explique une partie importante du bruit. Forman *et al.* (2008) ont aussi observé des différences pour l’incertitude de l’écoulement entre les plus petites superficies drainées analysées, mais ces différences étaient attribuées à la topographie. Il faut noter encore une fois que même si l’impact de la désagrégation diminue avec la superficie drainée, il est toujours présent à l’exutoire du bassin de la rivière des Anglais. La largeur de l’intervalle de confiance 95 % de l’écoulement désagrégé est d’environ 8 mm à l’exutoire (Figure 10.14), ce qui représente environ le tiers de l’écoulement de référence.

La Figure 10.16 présente les largeurs des intervalles d’écoulement en fonction des largeurs des intervalles de la pluie du 11 juin 2002 pour les sept sous-bassins et les résolutions 4,4 et 26,4 km à partir des mêmes valeurs des paramètres d’HYDROTEL utilisées à la Figure 10.7.

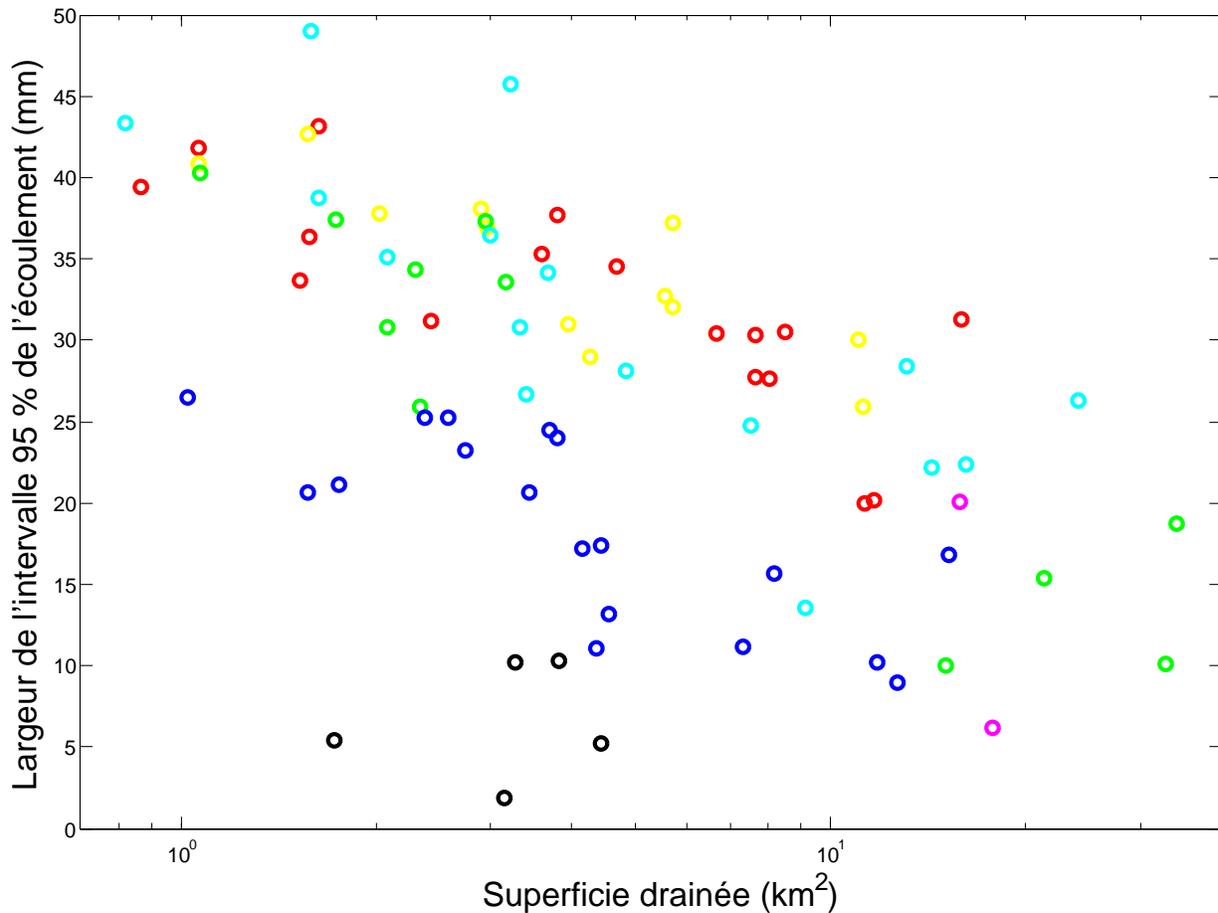


Figure 10.15 – Largeur de l’intervalle de confiance de l’écoulement désagrégé du 11 juin 2002 à résolution 4,4 km en fonction de la superficie des 81 tronçons de tête. La couleur indique le type de sol : sableux (noir), sable loameux (bleu), loam sableux (bleu pâle), loameux (vert), loam-sableux argileux (mauve), argile loameuse (jaune) et argileux (rouge).

Les impacts des variations de paramètres sur les largeurs des intervalles (Figure 10.16) sont les mêmes que sur l’écoulement moyen (Figure 10.7). La largeur de l’intervalle pour le sous-bassin le plus argileux (3, mauve) n’est que légèrement affectée par une variation substantielle de  $Z_3$  (de 1,32 à 4 m). Les largeurs des intervalles sur les sous-bassins plus sableux sont par contre nettement diminuées avec cette même augmentation de  $Z_3$ . La diminution de la lame de référence de l’HGM de 0,0005 m à 0,005 m a pour effet de réduire d’environ de moitié les largeurs des intervalles, peu importe le sous-bassin. Une différence par rapport à la Figure 10.7 est que, pour un même calage, les largeurs des intervalles de l’écoulement désagrégé à résolution 26,4 km peuvent être substantiellement plus petites que celles à résolution 4,4 km.

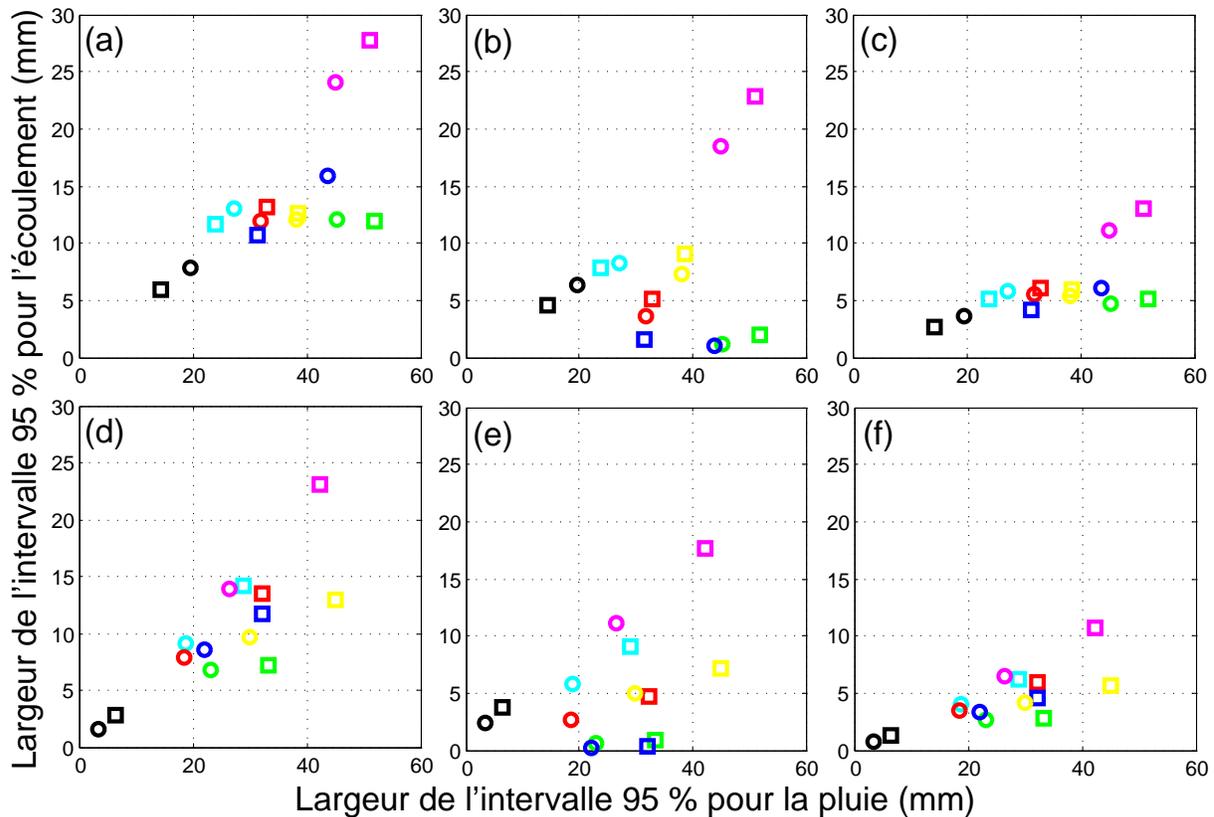


Figure 10.16 – Largeur de l’intervalle de confiance de l’écoulement du 11 juin simulé (a et d) à partir des paramètres calés à 4,4 km, (b et e) à partir d’une valeur de  $Z_3$  de 4 m et (c et f) à partir d’une lame de l’HGM de 0,0005 m en fonction de la largeur de l’intervalle de confiance de la pluie du 11 juin 2002 à résolution 4,4 (haut) et 26,4 km (bas) pour les deux sources de pluie et les sept sous-bassins. Sources de pluie : désagrégée (cercles) et aléatoire (carrés).

### 10.3 Discussion

Les résultats présentés dans ce chapitre proviennent d’un seul cas d’étude qui ne peut être transposé à toutes les situations de prévisions hydrologiques. Les simulations hydrologiques ont été effectuées au pas de temps journalier; une réponse et un calage différents auraient pu être obtenus à partir d’une résolution temporelle plus fine. Notons qu’un seul modèle hydrologique a été utilisé (HYDROTEL). Il n’a pas été jugé nécessaire d’en utiliser d’autres notamment parce qu’HYDROTEL est le modèle distribué utilisé par le CEHQ et Hydro-Québec, qui gèrent tous deux des réservoirs et des bassins versants québécois. Aussi, le bassin étudié compte très peu de zones urbaines et pas de grandes variations d’altitude. Malgré cela, certains constats ayant une portée plus générale que le présent cas d’étude

peuvent être faits. Les constats sont valides uniquement pour un événement de pluie intense. Il faut noter qu'à notre connaissance de la littérature existante, il s'agit d'une première application en modélisation hydrologique d'un modèle de désagrégation tenant compte de la corrélation spatiale avec les tuiles voisines et d'une première application de la désagrégation sur le territoire québécois.

Un premier résultat intéressant est qu'il a été possible d'utiliser les mêmes valeurs de paramètres du modèle de désagrégation au sud du Québec qu'au sud-est des États-Unis (Partie I). Cela suggère qu'il serait raisonnable d'utiliser les mêmes valeurs de paramètres sur des régions où l'estimation des paramètres n'a pu être faite, comme par exemple sur d'autres bassins du sud du Québec. D'autres analyses de la précipitation à fine résolution pourraient être faites sur d'autres régions afin d'évaluer plus précisément le comportement des paramètres du modèle de désagrégation.

Les résultats ont permis une illustration pour laquelle la désagrégation apporte une valeur ajoutée par rapport aux deux autres sources de pluie utilisées. L'interpolation donne des résultats comparables à la valeur moyenne désagrégée, mais ne tient pas compte de l'incertitude sur la distribution spatiale de la pluie puisqu'elle ne donne qu'une seule valeur. La pluie aléatoire, qui répartit aléatoirement les valeurs désagrégées sur la tuile, ne permet pas de reproduire la structure spatiale des précipitations. Elle est donc biaisée et donne des intervalles de confiance trop larges. Bien qu'un seul événement ait été analysé et que la structure spatiale soit différente d'un événement à un autre, la désagrégation a montré qu'elle pouvait mieux performer pour une application en hydrologie que les deux autres sources de pluies analysées puisqu'elle seule tient compte à la fois de la structure spatiale de l'événement et de son incertitude.

Il faut noter que pour l'événement étudié, la désagrégation, comme les deux autres sources, a sous-estimé la pluie de pointe sur le bassin. Deux causes sont possibles. Premièrement, le Chapitre 3 a montré que le modèle de désagrégation avait tendance à sous-estimer la valeur maximale de précipitation sur une tuile. Deuxièmement, bien que les 1000 réalisations désagrégées produisent toutes des champs différents, il se peut que le modèle produise des champs de précipitation désagrégés trop similaires. Par exemple, si la pointe de pluie se situe sur le bassin, mais que le modèle la place trop souvent en dehors du bassin, comme ce fut le cas avec la désagrégation à 8,8 km, l'écoulement sera sous-estimé. L'inverse, soit une

pointe observée à l'extérieur du bassin et des pointes désagrégées presque systématiquement sur le bassin, aurait aussi pu se produire et aurait causé une surestimation de l'écoulement. Dans ce cas, une façon d'éliminer, ou à tout le moins d'amoindrir ce manque de variabilité spatiale, serait de varier, d'une simulation à une autre, les données d'entrée (pluie sur les tuiles, CAPE, vitesse et direction du vent à 700 hPa) et les paramètres du modèle de désagrégation selon leurs incertitudes respectives.

La largeur des intervalles de confiance de la pluie désagrégée a tendance à diminuer lorsque la superficie drainée augmente, ce qui corrobore les résultats de Carpenter et Georgakakos (2006) et Forman *et al.* (2008). Ce résultat était attendu puisqu'en augmentant la superficie drainée, le nombre de pixels augmente et la moyenne sur le sous-bassin devient plus stable. La largeur des intervalles diminue aussi avec la résolution des pixels parce que plus les pixels sont grands, moins il y a de pixels sur la tuile et plus les valeurs sont contraintes à être près de la valeur sur la tuile.

La précipitation utilisée en entrée dans le modèle hydrologique influe de manière importante sur l'écoulement simulé. Si le calage du modèle est approprié, le biais de l'écoulement est directement lié au biais sur la pluie en entrée (Figure 10.5). La largeur de l'intervalle de confiance de l'écoulement est aussi directement liée à celle de la pluie (Figure 10.13).

Des facteurs autres que la précipitation influent sur l'écoulement simulé. La teneur en eau dans les sols avant l'événement doit aussi être considéré. L'Annexe F a montré que cette teneur en eau dépend notamment des types de sols et de la profondeur des couches. Pour une même hauteur de précipitation antérieure à l'événement, un sol argileux répond plus rapidement et l'écoulement est donc une fonction directe de la pluie. C'est pourquoi pour une même hauteur de pluie, l'écoulement moyen et la largeur des intervalles de confiance de l'écoulement du sous-bassin 3 (mauve) sont plus grands que pour les autres sous-bassins. S'il y avait eu un sous-bassin urbain, avec sol imperméable, il y aurait probablement eu une relation encore plus directe entre l'écoulement et la pluie. À l'inverse, un sol plus sableux est plus lent à saturer ; l'écoulement est donc plus sensible aux événements précédents et aux paramètres de couches de sols d'HYDROTEL. Ainsi, un mauvais calage du paramètre  $Z_3$  peut avoir un impact considérable pour les bassins sableux ou loameux (Figures 10.7 et 10.16 b-e). L'Annexe F a aussi montré que l'occupation du sol avait un impact négligeable pour cet événement intense ; c'est également ce qu'avaient obtenu Lavigne *et al.* (2004) sur

un autre bassin.

Un autre facteur à considérer est la vitesse de la réponse à un événement de pluie, donnée par l'HGM. En relatif, l'impact de la hauteur de la lame de l'HGM sur le biais et sur la largeur des intervalles de confiance est le même pour tous les tronçons (Figures 10.7 et 10.16 c-f). Il faut noter que l'analyse du temps de réponse aurait été plus complète si un pas de temps plus court avait été utilisé. Il aurait aussi été intéressant de procéder à une analyse de sensibilité plus complète des paramètres d'HYDROTEL afin d'évaluer précisément l'impact d'une modification du jeu de paramètres sur l'écoulement simulé et de comparer cet impact à celui de la source de précipitation utilisée en entrée, mais cela allait au-delà des objectifs de la thèse.

En somme, les résultats ont montré que la méthode utilisée pour distribuer spatialement la pluie d'un événement fort influe beaucoup sur l'écoulement, même dans le cas d'un événement stratiforme. Il est donc important d'en tenir compte. Les résultats ont aussi montré que la désagrégation avait un impact même à l'exutoire du bassin (730 km<sup>2</sup>). Il faudrait utiliser un bassin plus grand pour évaluer jusqu'à quelle superficie drainée la désagrégation a un impact. À noter qu'en plus de la superficie du bassin, l'impact de la désagrégation dépend de la taille de la tuile et du positionnement du bassin sur la grille mésoéchelle utilisée en entrée. Par exemple, pour un bassin ayant la même superficie qu'une tuile, la désagrégation aura moins d'impact si le bassin est couvert entièrement par une seule tuile que si le bassin est couvert par deux, trois ou quatre tuiles.

Les résultats ont montré que désagréger à plus fine résolution augmente la largeur des intervalles de confiance (Figure 10.11), mais ces mêmes résultats n'ont cependant pas permis d'évaluer clairement dans quelles circonstances une résolution spatiale est meilleure qu'une autre pour les applications. Néanmoins, il est recommandé de ne pas désagréger une tuile en  $2 \times 2$  pixels. Avec la contrainte qu'il faut que la précipitation sur la tuile soit conservée, il n'y a que trois degrés de liberté sur une tuile ; les valeurs désagrégées sont trop restreintes. Il n'est pas nécessaire non plus de désagréger à une résolution plus fine que la plus petite unité de simulation du modèle hydrologique. Pour HYDROTEL, cette unité est l'UHRH ; dans la présente application, les UHRH avaient une superficie moyenne d'environ 4 km<sup>2</sup>. Il est recommandé aussi de désagréger à une résolution qui n'est pas trop loin de la résolution utilisée pour le calage du modèle hydrologique et de la résolution des valeurs de précipitation disponibles pour estimer les paramètres du modèle de désagrégation.

# Chapitre 11

## CONCLUSION

Les données de précipitation future, qu'elles proviennent d'un MRC ou d'un modèle météorologique, sont généralement issues de grilles à résolution mésoéchelle. Le raffinement de l'échelle spatiale de ces modèles à base physique est soumis à deux contraintes. Premièrement, un modèle ayant une résolution plus fine doit tenir compte des processus physiques à plus fine échelle, donc généralement être plus complexe. Deuxièmement, un modèle à plus fine résolution demande un plus long temps de calcul. Avec l'évolution des connaissances sur le climat et des ressources informatiques, la résolution spatiale des MRC et des modèles météorologiques continuera de s'améliorer. Néanmoins, l'écart entre la résolution spatiale disponible à partir de ces modèles et la résolution spatiale de la précipitation nécessaire pour répondre aux besoins de certaines applications hydrologiques ne sera pas comblé à court terme.

L'objectif principal de la thèse était de développer un modèle statistique de désagrégation spatiale de la précipitation mésoéchelle capable de produire des champs réalistes. Les deux prochaines sections présentent un retour sur les résultats obtenus pour les trois objectifs spécifiques (Section 11.1) et une proposition d'axes de recherche à investiguer (Section 11.2).

### 11.1 Résumé des constats

Trois objectifs spécifiques ont été définis : *(i)* le développement du modèle de désagrégation sur une région avec peu de relief; *(ii)* l'ajustement du modèle pour tenir compte de la topographie; et *(iii)* l'application du modèle proposé en modélisation hydrologique. Afin

de répondre à ces objectifs, le modèle de désagrégation proposé a été développé sur trois régions géographiques différentes.

### 11.1.1 Partie I

Le modèle de désagrégation a été développé sur une région d'environ  $400 \times 400 \text{ km}^2$  du sud-est des États-Unis. Il est supposé que la précipitation sur un pixel est issue de la distribution lognormale avec espérance définie en fonction des huit pixels les plus proches et du vent à 700 hPa et dont l'écart-type dépend de l'espérance et de la variable atmosphérique CAPE. Le modèle utilise l'échantillonnage de Gibbs pour l'allocation des valeurs à chaque pixel. Le modèle produit des champs réalistes, c'est-à-dire qu'il est capable de produire une structure spatiale (corrélation spatiale, anisotropie) et des pointes de précipitation comparables à celles des champs observés. Le modèle est simple dans le sens qu'il compte peu de paramètres à estimer (deux pour l'espérance et trois pour l'écart-type). Aussi, aucune transformation n'a été faite sur les variables explicatives. Les paramètres du modèle sont ainsi plus facilement interprétables que pour un GMRF (Allcroft et Glasbey, 2003) ou un modèle utilisant les transformations de Fourier (Ferraris *et al.*, 2003). Les huit pixels voisins les plus proches permettent de définir quatre variables simples à conceptualiser. Grâce à l'échantillonnage de Gibbs, le modèle est capable de tenir compte de la corrélation spatiale sur de longue distance, même s'il ne tient compte que des voisins immédiats. Une des forces du modèle proposé est qu'il peut tenir compte de l'information sur les tuiles voisines, contrairement aux modèles multi-échelles et aux modèles simulant les cellules de précipitation (Chapitre 2).

Deux variables atmosphériques sont utilisées : le vent à 700 hPa permet de définir l'anisotropie et CAPE caractérise la variabilité spatiale de la précipitation. Même si le modèle n'est pas à base physique, définir les paramètres du modèle en fonction de variables atmosphériques permet de tenir compte indirectement des processus physiques régissant la précipitation. Les valeurs estimées des paramètres sont ainsi plus robustes aux variations des conditions atmosphériques de chaque événement. Cela implique qu'elles sont plus facilement transposables à d'autres régions et qu'elles sont plus robustes aux CC que des valeurs estimées sans tenir compte des variables atmosphériques.

Le modèle proposé a deux principales faiblesses : il produit des événements un peu trop lisses spatialement et il a tendance à sous-estimer les valeurs extrêmes. C'est surtout vrai

pour les événements convectifs. La désagrégation en cascade, telle qu'utilisée par les modèles de désagrégation multi-échelle (Chapitre 2), a permis d'éliminer ces faiblesses, mais en a créé une autre en produisant des champs ayant des structures spatiales invraisemblables. La désagrégation en cascade n'est donc pas conseillée, surtout pour la précipitation stratiforme.

### 11.1.2 Partie II

L'ajustement du modèle pour tenir compte des montagnes a été fait dans le nord-ouest des États-Unis, couvrant en partie les montagnes Olympiques et la chaîne des Cascades. La majeure partie de la précipitation tombe en hiver et contrairement au sud-est des États-Unis, pratiquement tous les événements sont stratiformes. C'est pourquoi la variable CAPE n'a pas été retenue. Le vent à l'échelle de la tuile n'a pas été retenu non plus puisqu'il n'apportait pas d'information pertinente pour la désagrégation. Il est impossible de dire si c'est parce que les montagnes perturbent le vent à mésoéchelle de sorte que le vent à fine échelle soit totalement différent ou si c'est parce que la précipitation de référence à fine échelle, produite à l'aide d'un modèle d'interpolation statistique (PRISM), n'est pas capable d'inclure l'impact du vent. Quoiqu'il en soit, aucune variable atmosphérique n'a été retenue pour cette partie.

Un des résultats les plus intéressants de cette partie est que le modèle original, qui ne tient pas compte des montagnes, est capable de reproduire l'impact mésoéchelle des deux chaînes de montagnes présentes (Chapitre 7). Le modèle est donc en mesure de bien utiliser l'information contenue par les tuiles utilisées en entrée. Les modèles qui n'utilisent pas l'information des tuiles voisines, comme les modèles multi-échelles et les modèles simulant les cellules de précipitation, ne pourraient reproduire l'impact d'une chaîne de montagnes sans que la topographie ait été explicitement spécifiée en entrée.

La pente et l'orientation de la topographie à l'échelle du pixel n'ont pas été retenues pour la désagrégation. Selon toute vraisemblance, l'impact de la pente est inclus dans la précipitation des huit pixels voisins. Seule l'anomalie topographique, soit la différence entre l'altitude du pixel et l'altitude estimée au pixel en supposant une pente constante, a été retenue. Cette variable a permis de reproduire la plupart des pointes de précipitation, mais n'explique pas entièrement la précipitation orographique à fine échelle. En somme, les résultats de la deuxième partie suggèrent que le modèle original tel que développé à la Partie I peut être

utilisé même en régions montagneuses, sauf si les anomalies topographiques sont de plusieurs centaines de mètres.

### 11.1.3 Partie III

Le modèle proposé à la Partie I est appliqué sur un événement intense survenu le 11 juin 2002 sur le bassin de la rivière des Anglais. Les champs désagrégés sont utilisés en entrée dans le modèle hydrologique distribué HYDROTEL. Comme les paramètres du modèle de désagrégation dépendent des variables atmosphériques, les mêmes valeurs de paramètres que celles estimées à la Partie I sur le sud des États-Unis ont pu être utilisées. À notre connaissance, c'est la première application de la désagrégation spatiale sur un bassin versant québécois et la première utilisation de champs de précipitations désagrégés dans HYDROTEL.

Les résultats ont montré que même pour un événement stratiforme, la variabilité spatiale pouvait être importante et qu'il fallait en tenir compte adéquatement. La pluie de référence a été comparée avec la pluie désagrégée, la pluie interpolée et la pluie désagrégée, mais répartie sans structure spatiale. Pour l'événement étudié, la médiane de la pluie désagrégée est comparable à la pluie interpolée. Ces deux sources de pluie placent la pointe de précipitation généralement au bon endroit, mais en sous-estiment la valeur. L'avantage de la désagrégation sur l'interpolation est qu'elle peut produire plusieurs champs simulés. Plusieurs séries peuvent aussi être produites à partir de la pluie sans structure spatiale, mais les intervalles de confiance sont biaisés et inutilement larges.

L'objectif était de déterminer si, pour un événement intense, la pluie désagrégée apportait une information supplémentaire par rapport à la pluie agrégée. Il était attendu que l'impact diminuerait si la superficie drainée augmentait. En fait, les résultats ont montré que la désagrégation avait un impact même à l'exutoire du bassin (superficie drainée de 730 km<sup>2</sup>). Par ailleurs, outre la pluie utilisée en entrée, le calage se doit d'être adéquat. L'impact varie aussi d'un type de sol à un autre. Un biais sur la précipitation aura plus d'impact sur un sol argileux, qui répond rapidement, que sur un sol sableux. Ces résultats suggèrent que l'impact serait encore plus grand en milieu urbain. La résolution spatiale de la précipitation désagrégée doit être adéquate par rapport à la résolution de la précipitation utilisée pour le calage du modèle hydrologique et pour l'estimation des paramètres du modèle de désagrégation. La désagrégation d'une tuile en 2 × 2 pixels est déconseillée puisque dans ce cas, la contrainte

(*i.e.* la précipitation sur la tuile doit être conservée) est trop forte par rapport au nombre de données désagrégées (4) sur une tuile.

## 11.2 Perspectives

Plusieurs éléments de cette thèse peuvent être approfondis dans de travaux futurs. Ils peuvent être classés selon deux axes principaux : (*i*) l’analyse approfondie des propriétés du modèle de désagrégation (Section 11.2.1) et (*ii*) l’analyse de sensibilité de la simulation hydrologique à partir de la pluie désagrégée (Section 11.2.2).

### 11.2.1 Analyse approfondie des propriétés du modèle de désagrégation

Les principales faiblesses du modèle sont qu’il sous-estime les pixels extrêmes et qu’il donne des champs de précipitation trop lisses, surtout pour les événements convectifs. Trois éléments pourraient être approfondis afin d’améliorer le modèle. Premièrement, dans la présente thèse, la distribution lognormale a été choisie puisqu’une distribution classique à deux paramètres qui puisse être exprimée en fonction de l’espérance et de l’écart-type était désirée. Une distribution à trois paramètres, dont le troisième paramètre permettrait d’ajuster le coefficient d’aplatissement de la distribution, pourrait être utilisée. Par exemple, en choisissant une distribution statistique à queue plus lourde, le modèle pourrait générer des valeurs extrêmes plus grandes, ce qui permettrait du même coup d’augmenter la variabilité spatiale du champ désagrégé. Cependant, une distribution à queue lourde comme la GEV peut détériorer l’estimation des pixels plus modérés (Caron *et al.*, 2008). Il faut considérer qu’aucune distribution classique et ayant peu de paramètres ne peut représenter exactement la précipitation à fine échelle. Pour une représentation plus précise de la distribution de la précipitation à fine échelle, il faut être prêt à sacrifier de la simplicité au modèle. Par exemple, Caron *et al.* (2008) ont appliqué des facteurs de correction à la distribution Gamma afin de mieux représenter les extrêmes. Ces facteurs de corrections sont différents d’une région à l’autre, ce qui augmente considérablement le nombre de paramètres si le modèle est appliqué sur un grand territoire.

Deuxièmement, le rôle de la variable atmosphérique CAPE pourrait être évalué avec plus de précision. Cette variable avait seulement été utilisée en désagrégation par Perica et Foufoula-Georgiou (1996). *A priori*, il était supposé que la variabilité augmenterait

avec CAPE, mais la relation exacte entre CAPE et l'écart-type n'était pas connue. La forme la plus simple qui donnait de meilleurs résultats a été retenue. Il a été supposé que l'écart-type augmentait linéairement en fonction de CAPE. Cependant, cette fonction est vraisemblablement trop simple pour bien représenter la variabilité spatiale de la précipitation. Il faudrait analyser d'autres régions ayant beaucoup d'événements convectifs afin de mieux évaluer le rôle de CAPE. Aussi, il serait conseillé, si possible, de prendre la valeur de CAPE juste avant le début de l'événement. Dans la présente thèse, la précipitation journalière a été utilisée et la valeur de CAPE moyenne journalière sur la tuile a été retenue. Si l'information à plus fine échelle temporelle est disponible, le rôle de CAPE peut être mieux évaluer. Il pourrait aussi être intéressant de considérer d'autres variables, comme le nombre de Richardson, pour expliquer la variabilité spatiale.

Troisièmement, la méthode utilisée pour estimer les paramètres du modèle a peut-être donné trop de poids aux valeurs plus modérées. Dans tous les cas, les valeurs qui minimisaient la somme du carré des écarts ont été choisies. Au lieu de produire des estimations ponctuelles des paramètres, une approche bayésienne pourrait être utilisée, c'est-à-dire que les paramètres pourraient être vus comme des variables aléatoires ayant une distribution statistique conjointe quelconque. Ainsi, pour chaque champ désagrégé, des valeurs des vecteurs  $\vec{\beta}$  et  $\vec{\theta}$  seraient sélectionnées à partir de leurs distributions respectives. Les distributions pourraient être estimées par rééchantillonnage :  $K$  valeurs du jeu de données seraient sélectionnées aléatoirement et une estimation ponctuelle des paramètres du modèle serait obtenue. Cette étape serait répétée  $N$  fois (*e.g.*  $N = 10\,000$ ,  $N = 100\,000$ ) afin d'avoir une estimation précise de la distribution des paramètres. Cette approche pourrait être utilisée sur plusieurs régions afin de voir si les distributions de paramètres sont comparables. Une autre façon d'ajouter de la variabilité au modèle serait de perturber les entrées (précipitation sur les tuiles, CAPE, vitesse et direction du vent) selon leurs incertitudes respectives. Cette approche n'est pas évidente à mettre en œuvre puisque les incertitudes de ces variables sont difficiles à estimer.

En hautes montagnes, le modèle proposé ne considère que l'anomalie topographique. Des analyses sur d'autres régions montagneuses devraient être effectuées afin de valider les résultats obtenus à la deuxième partie. Cependant, il est très difficile d'obtenir des données à fine résolution spatiale de qualité en hautes montagnes. Dans le nord-ouest des États-Unis, un modèle statistique est utilisé pour produire les champs journaliers de précipitation à fine échelle. Ces données sont considérées fiables, mais ne représentent pas

exactement la variabilité spatiale réelle. Une amélioration des modèles de désagrégation en hautes montagnes passe inévitablement par une amélioration de la qualité des données de précipitation à fine résolution.

Il faut aussi noter que la désagrégation n'a été effectuée qu'au pas de temps journalier et qu'à partir de tuiles d'environ 50 km (45,6 km à la Partie I, 52,8 km aux Parties II et III). Si la désagrégation est effectuée sur un plus petit pas de temps (*i.e.* horaire), il faut s'attendre à ce que la corrélation entre pixels voisins soit moins grande et il est possible que la précipitation à fine échelle du pas de temps précédent apporte de l'information pertinente. La taille de la tuile n'a pas d'impact sur les paramètres, qui sont estimés à l'échelle du pixel. La taille de la tuile pourrait avoir un impact si la tuile était beaucoup plus grande (*e.g.* tuile d'un MCG) ou si le pas de temps était beaucoup plus petit (*e.g.* horaire). Dans ces cas, la proportion de pixels secs pourrait être non-négligeable, même si la précipitation sur certains pixels est importante. Il vaudrait peut-être la peine dans ces situations d'utiliser une distribution mixte (Section 3.2.3) au lieu d'une distribution continue.

### **11.2.2 Analyse de sensibilité de la simulation hydrologique à partir de la pluie désagrégée**

Les analyses de la Partie III ont été faites pour un seul événement sur un seul bassin à partir d'un seul champ de précipitation mésoéchelle en entrée dans le modèle de désagrégation. Il serait intéressant d'étendre les analyses sur plusieurs événements et plusieurs bassins versants du sud du Québec à partir de différentes sources de précipitation mésoéchelle. Le contexte serait ainsi plus favorable au développement d'une relation plus générale entre l'écoulement et la superficie du bassin ou la résolution de la précipitation désagrégée. Notons que Bourdillon *et al.* (2011) ont déjà appliqué HYDROTEL sur l'ensemble du Québec méridional. Bien qu'il n'y ait pas nécessairement des données à fine résolution disponibles pour caler le modèle de désagrégation sur tous les bassins versants du sud du Québec, les analyses de la Partie III ont montré qu'il était raisonnable d'utiliser les mêmes valeurs de paramètres d'une région à une autre.

En prévision hydrologique, des ensembles de prévisions provenant de différents modèles météorologiques sont parfois utilisés (*e.g.* Velazquez *et al.*, 2009; Boucher *et al.*, 2011). En plus de ces diverses séries de précipitation en entrée, différents jeux de paramètres du modèle hydrologique peuvent être utilisés pour tenir compte de cette source d'incertitude.

Notons qu'il est possible avec HYDROTEL de procéder à un calage automatique qui permet une analyse objective des paramètres les plus sensibles. Il serait intéressant d'analyser l'écoulement simulé à partir de différentes valeurs des paramètres les plus sensibles du modèle hydrologique utilisé et de séquences de précipitation provenant de plusieurs modèles météorologiques. L'incertitude provenant des modèles météorologiques et celle provenant des paramètres du modèle hydrologique pourraient ainsi être isolées. L'incertitude associée à la structure du modèle hydrologique pourrait aussi être considérée en utilisant plusieurs modèles hydrologiques. À cela s'ajouterait l'incertitude sur la résolution spatiale de la précipitation en désagrégant chaque série météorologique plusieurs fois. Il serait ainsi possible de comparer l'incertitude sur la résolution spatiale avec l'incertitude provenant des autres sources. En particulier, cela permettrait de voir si la désagrégation permet d'éliminer la sous-dispersion des écoulements simulés à partir d'un ensemble de prévisions météorologiques qu'ont obtenue Velazquez *et al.* (2009).

Ce type d'analyse peut aussi se faire à partir d'un ensemble de projections climatiques. Dans ce cas, en plus des sources d'incertitudes mentionnées ci-haut, s'ajoutent les incertitudes sur le scénario d'émissions de GES et sur le calage du modèle hydrologique, qui aura peut-être besoin d'être modifié pour le futur à cause des CC ou d'une possible augmentation des zones urbaines sur le bassin.

L'impact de la désagrégation diminue lorsque la superficie du bassin augmente. Comme la désagrégation avait encore un impact à l'exutoire du bassin étudié à la Partie III, il faudrait étudier d'autres bassins plus grands afin d'évaluer la superficie à partir de laquelle il ne vaut plus la peine de désagréger. Cette superficie dépendra de la taille des tuiles et de la position du bassin par rapport aux tuiles de la grille.

# Bibliographie

Ahrens, B., 2003 : Rainfall downscaling in an alpine watershed applying a multiresolution approach. *Journal of Geophysical Research*, **108**(D8), 8388.

Allcroft, D.J., et C.A. Glasbey, 2003 : A latent Gaussian Markov random-field model for spatiotemporal rainfall disaggregation. *Applied Statistics*, **52**(4), 487-498.

Ambroise, B. 1998 : Genèse des débits dans les petits bassins versants ruraux en milieu tempéré : (1) - Processus et facteurs. *Revue des Sciences de l'Eau*, **4**, 471-495.

Arnaud, P., C. Bouvier, L. Cisneros, et R. Dominguez, 2002 : Influence of rainfall spatial variability on flood prediction. *Journal of Hydrology*, **260**, 216-230.

Bacchi, B., et R. Ranzi, 2003 : Hydrological and meteorological aspects of floods in the Alps : an overview. *Hydrology and Earth System Sciences*, **7**(6), 785-798.

Badas, M.G., R. Deidda, et E. Piga, 2006 : Modulation of homogeneous space-time rainfall cascades to account for orographic influences. *Natural Hazards and Earth System Science*, **6**(3), 427-437.

Baillargeon, S., 2005 : Le krigeage : revue de la théorie et application à l'interpolation spatiale des données de précipitations. Mémoire de maîtrise, Université Laval, Québec, 128 p.

Barstad, I., et R.B. Smith, 2005 : Evaluation of an Orographic Precipitation Model. *Journal of Hydrometeorology*, **6**, 85-99.

Bell, V.A., et R.J. Moore, 2000 : The sensitivity of catchment runoff models to rainfall

data at different spatial scales. *Hydrology and Earth System Sciences*, **4**(4), 653-667.

Bindlish, R., et A.P. Barros, 2000 : Disaggregation of rainfall for one-way coupling of atmospheric and hydrological models in regions of complex terrain. *Global and Planetary Change*, **25**(1-2), 111-132.

Bisson, J.L., et F. Roberge, 1983 : Prévision des apports naturels : Expérience d'Hydro-Québec. Dans *Workshop on flow predictions*, Toronto, ON, November 1983.

Blanchard, D.O., 1998 : Assessing the vertical distribution of convective available potential energy. *Weather and Forecasting*, **13**(3), 870-877.

Boé, J., L. Terray, F. Habets, et E. Martin, 2007 : Statistical and dynamical downscaling of the Seine basin climate for hydro-meteorological studies. *International Journal of Climatology*, **27**, 1643-1655.

Boucher, M.-A., F. Anctil, L. Perreault, et D. Tremblay, 2011 : A comparison between ensemble and deterministic hydrological forecasts in an operational context. *Advances in Geosciences*, **29**, 85-94.

Bouda, M., A.N. Rousseau, B. Konan, P. Gagnon, et S. Gumiere, 2012 : Case study : Bayesian uncertainty analysis of the distributed hydrological model HYDROTEL. Accepté pour publication dans *Journal of Hydrologic Engineering*. DOI : 10.1061/(ASCE)HE.1943-5584.0000550

Bourdillon, R., S. Ricard, D. Roussel, R. Turcotte, et J.-F. Cyr, 2011 : *Impact des changements climatiques sur le régime hydrique québécois : Évaluer, à l'horizon 2050, les impacts pour le Québec méridional sur les écoulements en eau, en excluant l'effet de la gestion des barrages, sur des indicateurs hydrologiques utilisés en gestion de l'eau*. Rapport interne, Centre d'Expertise Hydrique du Québec (CEHQ), Direction de l'expertise hydrique, Division de l'hydrologie et de l'hydraulique, Québec. 177 p.

Brissette, F.P, M. Khalili, et R. Leconte, 2007 : Efficient stochastic generation of multi-site synthetic precipitation data. *Journal of Hydrology*, **345**, 121-133.

Bürger, G., et Y. Chen, 2005 : Regression-based downscaling of spatial variability for hydrologic applications. *Journal of Hydrology*, **311**, 299-317.

Burton, A., C.G. Kilsby, H.J. Fowler, P.S.P. Cowpertwait, et P.E. O'Connell, 2008 : RainSim : A spatial-temporal stochastic rainfall modelling system. *Environmental Modelling & Software*, **23**, 1356-1369.

Caron, A., R. Leconte, et F. Brissette, 2008 : An Improved Stochastic Weather Generator for Hydrological Impact Studies. *Canadian Water Resources Journal*, **33**(3), 233-256.

Carpenter, T.M., et K.P. Georgakakos. 2006 : Discretization scale dependencies of the ensemble flow range versus catchment area relationship in distributed hydrologic modeling. *Journal of Hydrology*, **328**, 242-257.

Chandler, R.E., N. Mackay, H.S. Wheeler, et C. Onof, 2000 : Bayesian Image Analysis and the Disaggregation of Rainfall. *Journal of Atmospheric and Oceanic Technology*, **17**, 641-650.

Christensen, O.B., J.H. Christensen, B. Machenhauer, et M. Botzet, 1998 : Very High-Resolution Regional Climate Simulations over Scandinavia - Present Climate, *Journal of Climate*, **11**, 3204-3229.

Clark, M.P., et A.G. Slater, 2006 : Probabilistic quantitative precipitation estimation in complex terrain. *Journal of Hydrometeorology*, **7**(1), 3-22.

Collier, C.G., 2009 : On the propagation of uncertainty in weather radar estimates of rainfall through hydrological models. *Meteorological Applications*, **16** : 35-40.

Côté, M.-J., Y. Lachance, C. Lamontagne, M. Nastev, R. Plamondon, et N. Roy, 2006 : *Atlas du bassin versant de la rivière Châteauguay*. Collaboration étroite avec la Commission géologique du Canada et l'Institut national de la recherche scientifique - Centre eau, terre et environnement. Québec : Ministère du Développement durable, de l'Environnement et des Parcs. 64 p.

Cowpertwait, P.S.P, C.G. Kilsby, et P.E. O'Connell, 2002 : A space-time Neyman-Scott model of rainfall : Empirical analysis of extremes. *Water Resources Research*, **38**(8), 1131.

Creutin, J.D., H. Andrieu, et D. Faure, 1997 : Use of a weather radar for the hydrology of a mountainous area. Part II : radar measurement validation. *Journal of Hydrology*, **193**, 26-44.

Daly, C., R.P. Neilson, et D.L. Phillips. 1994 : A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *Journal of Applied Meteorology*, **33**(2), 140-158.

Daly, C., G.H. Taylor, et W.P. Gibson, 1997 : The PRISM approach to mapping precipitation and temperature. Preprints, *10th Conference on Applied Climatology*, Reno, NV, American Meteorological Society, 10-12.

Daly, C., W.P. Gibson, G.H. Taylor, G.L. Johnson, et P.A. Pasteris, 2002 : A knowledge-based approach to the statistical mapping of climate. *Climate Research*, **22**, 99-113.

Deidda, R., 2000 : Rainfall downscaling in a space-time multifractal framework. *Water Resources Research*, **36**(7), 1779-1794.

Deidda, R., M.G. Badas, et E. Piga, 2006 : Space-time multifractality of remotely sensed rainfall fields. *Journal of Hydrology*, **322**, 2-12.

Déqué, M., 2007 : Frequency of precipitation and temperature extremes over France in an anthropogenic scenario : Model results and statistical correction according to observed values. *Global and Planetary Change*, **57**, 16-26.

Diodato, N., 2005 : The influence of topographic co-variables on the spatial variability of precipitation over small regions of complex terrain. *International Journal of Climatology*, **25**(3), 351-363.

Ebtehaj, M., et E. Foufoula-Georgiou, 2010 : Orographic signature on multiscale statistics of extreme rainfall : A storm-scale study. *Journal of Geophysical Research D : Atmospheres*, **115**, D23112.

Entekhabi, D., et P.S. Eagleson, 1989 : Land surface hydrology parameterization for atmospheric general circulation models including sub-grid scale spatial variability. *Journal of Climate*, **2**, 816-831.

Ferraris, L., S. Gabellani, N. Rebora, et A. Provenzale, 2003 : A comparison of stochastic models for spatial rainfall downscaling. *Water Resources Research*, **39**(12), 1368.

Fiorucci, P., P. La Barbera, L.G. Lanza, et R. Minciardi, 2001 : A geostatistical approach to multisensor rain field reconstruction and downscaling. *Hydrology and Earth System Sciences*, **5**(2), 201-213.

Forman, B.A., E.R. Vivoni, et S.A. Margulis, 2008 : Evaluation of ensemble-based distributed hydrologic model response with disaggregated precipitation products. *Water Resources Research*, **44**, W12409.

Fortin, V., 2000 : *Le modèle météo-apport HSAMI : historique, théorie et application*. Rapport de recherche, révision 1,5. Institut de recherche d'Hydro-Québec (IREQ), Varennes, Canada, 68 p.

Fortin, J.-P., R. Turcotte, S. Massicotte, R. Moussa, J. Fitzback, et J.-P. Villeneuve, 2001a : A distributed watershed model compatible with remote sensing and GIS data. Part I : Description of the model. *Journal of Hydrologic Engineering*, **6**(2), 91-99.

Fortin, J.-P., R. Turcotte, S. Massicotte, R. Moussa, et J. Fitzback, 2001b : A distributed watershed model compatible with remote sensing and GIS data. Part 2 : Application to the Chaudière watershed. *Journal of Hydrologic Engineering*, **6**(2), 100-108.

Fowler, H.J., S. Blenkinsop, et C. Tebaldi, 2007 : Linking climate change modelling to impacts studies : recent advances in downscaling techniques for hydrological modelling. *International Journal of Climatology*, **27**, 1547-1578.

Fulton R.A., J.P. Breidenbach, D-J. Seo, et D.A. Miller, 1998 : The WSR-88D Rainfall Algorithm. *Weather and Forecasting*, **13**, 377-395.

Gagnon, P., B. Konan, A.N. Rousseau, et M. Slivitzky, 2009 : Hydrometeorological validation of a Canadian Regional Climate Model simulation within the Chaudière and Châteauguay watersheds (Québec, Canada). *Canadian Journal of Civil Engineering*, **36**, 253-266.

Gagnon, P., A.N. Rousseau, A. Mailhot, et D. Caya, 2012a : Spatial Disaggregation of Mean Areal Rainfall using Gibbs Sampling. Accepté pour publication dans *Journal of Hydrometeorology*, **13**(1), 324-337.

Gagnon, P., A.N. Rousseau, A. Mailhot, et D. Caya, 2012b : A Gibbs Sampling Disaggregation Model for Orographic Precipitation. Accepté pour publication dans *International Journal of Applied Earth Observation and Geoinformation*. DOI : 10.1016/j.jag.2011.11.002

Geman, S., et D. Geman, 1984 : Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis Machine and Intelligence*, **6**(6), 721-741.

Golz, C., T. Einfalt et S.C. Michaelides, 2006 : Quality control of rainfall measurements in Cyprus. *Meteorological Applications*, **13**, 197-201.

Goodison, B.E., P.Y.T. Louie et D. Yang, 1998 : *WMO solid precipitation measurement intercomparison : Instruments and observing methods*. Report No. 67. World Meteorological Organisation, Geneva.

Groppelli, B., D. Bocchiola, et R. Rosso, 2011 : Spatial downscaling of precipitation from GCMs for climate change projections using random cascades : A case study in Italy. *Water Resources Research*, **47**, W03519.

Guan, H., J.L. Wilson, et H. Xie, 2009 : A cluster-optimizing regression-based approach for

precipitation spatial downscaling in mountainous terrain. *Journal of Hydrology*, **375**(3-4), 578-588.

Guillot, G., et T. Lebel, 1999 : Disaggregation of Sahelian mesoscale convective system rain fields : Further developments and validation. *Journal of Geophysical Research D : Atmospheres*, **104**(D24), 31,533-31,551.

Güntner, A., J. Olsson, A. Calver, et B. Gannon, 2001 : Cascade-based disaggregation of continuous rainfall time series : the influence of climate. *Hydrology and Earth System Sciences*, **5**(2), 145-164.

Gupta, V.K., et E.C. Waymire, 1993 : A statistical analysis of mesoscale rainfall as a random cascade. *Journal of Applied Meteorology*, **32**(2), 251-267.

Haan, C.T., B.J. Barfield, et J.C. Hayes, 1993 : *Design hydrology and sedimentology for small catchments*. Academic Press, San Diego, CA, 588 p.

Harris, D., et E. Foufoula-Georgiou, 2001 : Subgrid variability and stochastic downscaling of modeled clouds : Effects on radiative transfer computations for rainfall retrieval. *Journal of Geophysical Research*, **106**(D10), 10,349-10,362.

Hay, L.E., 1998 : Stochastic calibration of an orographic precipitation model. *Hydrological Processes*, **12**(4), 613-634.

Hay, L.E., et G.J. McCabe, 1998 : Verification of the Rhea-orographic-precipitation model. *Journal of the American Water Resources Association*, **34**(1), 103-112.

Hellström, C., et D. Chen, 2003 : Statistical downscaling based on dynamically downscaled predictors : Application to monthly precipitation in Sweden. *Advances in Atmospheric Sciences*, **20**(6), 951-958.

Herrera, E., T.B.M.J. Ouarda, et B. Bobée, 2006 : Méthodes de désagrégation appliquées aux Modèles du Climat Global Atmosphère-Océan (MCGAO). *Revue des sciences de l'eau*, **19**(4), 297-312.

Huang, H., N. Cressie, et J. Gabrosek, 2002 : Fast, Resolution-Consistent Spatial Prediction of Global Processes From Satellite Data. *Journal of Computational and Graphical Statistics*, **11**(1), 63-88.

Hughes, J.P., P. Guttorp, et S.P. Charles, 1999 : A non-homogeneous hidden Markov model for precipitation occurrence. *Journal of the Royal Statistical Society. Series C : Applied Statistics*, **48**(1), 15-30.

Indarto, 2002 : *Découpages spatiaux et conséquences sur le bilan hydrologique : Application au bassin de l'Orb à travers une démarche de modélisation hydrologique distribuée*. Thèse de doctorat, École Nationale du Génie Rural, des Eaux et Forêts, Montpellier, 258 p.

Jensen, N.E., et L. Pederson, 2005 : Spatial variability of rainfall : Variations within a single radar pixel. *Atmospheric Research*, **77**, 269-277.

Jones, G.L., 2004 : On the Markov chain central limit theorem. *Probability Surveys*, **1**, 299-320.

Joss, J. et U. Germann, 2000 : Solutions and Problems when Applying Qualitative and Quantitative Information from Weather Radar. *Physics and Chemistry of the Earth (B)*, **25**(10-12), 837-841.

Jothityangkoon, C., M. Sivapalan, et N.R. Viney, 2000 : Tests of a space-time model of daily rainfall in southwestern Australia based on nonhomogeneous random cascades. *Water Resources Research*, **36**(1), 267-284.

Jutras, S., A.N. Rousseau, et C. Clerc, 2009 : Implementation of a peatland-specific water budget algorithm in HYDROTEL. *Canadian Water Resources Journal*, **34**(4), 349-364.

Khalili, M., F. Brissette, et R. Leconte, 2011 : Effectiveness of multi-site weather generator for hydrological modeling. *Journal of the American Water Resources Association*, **47**(2), 303-314.

Konan, B., M. Slivitzky, P. Gagnon et A.N. Rousseau, 2010 : Validation of the meteorological outputs of the Canadian Regional Climate Model using a kriging method : Application to southern Quebec. *Canadian Water Resources Journal*, **35**(3), 259-280.

Krajewski, W.F., E.N. Anagnostou, et G.J. Ciach, 1996 : Effects of the radar observation process on inferred rainfall statistics. *Journal of Geophysical Research*, **101**(D21), 26,493-26,502.

Kundu, P.K., et R.K. Siddani, 2007 : A new class of probability distributions for describing the spatial statistics of area-averaged rainfall. *Journal of Geophysical Research D : Atmospheres*, **112**(18), D18113.

Kyriakidis, P.C., 2004 : A Geostatistical framework for Area-to-Point Spatial Interpolation. *Geographical Analysis*, **36**(3), 259-289.

Lakhtakia, M.N., Z. Yu, B. Yarnal, R.A. White, et D.A. Miller, 1999 : Sensitivity of simulated surface runoff to mesoscale meteorological model resolution in a linked-model experiment. *Climate Resources*, **12**, 15-27.

Lavigne, M.-P., A.N. Rousseau, R. Turcotte, A.-M. Laroche, J.-P. Fortin, et J.-P. Villeneuve, 2004 : Validation and Use of a Semidistributed Hydrological Modeling System to Predict Short-Term Effects of Clear-Cutting on a Watershed Hydrological Regime. *Earth Interactions*, **8**(3), 1-19.

Lebel, T., I. Braud et J.-D. Creutin, 1998 : A space-time rainfall disaggregation model adapted to Sahelian mesoscale convective complexes. *Water Resources Research*, **34**(7), 1711-1726.

Lee, G.W., A.W. Seed, et I. Zawadzki, 2007 : Modeling the Variability of Drop Size Distributions in Space and Time. *Journal of Applied Meteorology and Climatology*, **46**, 742-756.

Lehmann, E.L., 1998 : *Nonparametrics : Statistical Methods Based on Ranks*. Prentice Hall, Upper Saddle River, New Jersey, 463 p.

Liang, X., J. Guo, et R.L. Leung, 2004 : Assessment of the effects of spatial resolutions on daily water flux simulations. *Journal of Hydrology*, **298**, 287-310.

Licznar, P., J. Lomotowski, et D.E. Rupp, 2011 : Random cascade driven rainfall disaggregation for urban hydrology : An evaluation of six models and a new generator. *Atmospheric Research*, **99**, 563-578.

Lin, Y., et K.E. Mitchell, 2005 : The NCEP Stage II/IV hourly precipitation analyses : development and applications. Preprints, 19th Conf. on Hydrology, *American Meteorological Society*, San Diego, CA, 9-13 January 2005, Paper 1.2.

Lovejoy, S., et D. Schertzer, 2010a : On the simulation of continuous in scale universal multifractals, part I : Spatially continuous processes. *Computers and Geosciences*, **36**(11), 1393-1403.

Lovejoy, S., et D. Schertzer, 2010b : On the simulation of continuous in scale universal multifractals, part II : Space-time processes and finite size corrections. *Computers and Geosciences*, **36**(11), 1404-1413.

Ludwig, R., S. Taschner, et W. Mauser, 2003 : Modelling floods in the Ammer catchment : limitations and challenges with a coupled meteo-hydrological model approach. *Hydrology and Earth System Sciences*, **7**(6), 833-847.

Ludwig, R., I. May, R. Turcotte, L. Vescovi, M. Braun, J.-F. Cyr, L.-G. Fortin, D. Chaumont, S. Biner, I. Chartier, D. Caya, et W. Mauser, 2009 : The role of hydrological model complexity and uncertainty in climate change impact assessment. *Advances in Geosciences*, **21**, 63-71.

Mackay, N.G., R.E. Chandler, C. Onof, et H.S. Wheater, 2001 : Disaggregation of spatial rainfall fields for hydrological modelling. *Hydrology and Earth System Sciences*, **5**, 165-173.

Mahfouf, J.-F., B. Brasnett, et S. Gagnon, 2007 : A Canadian precipitation analysis (CaPA) project : Description and preliminary results. *Atmosphere-Ocean*, **45**(1), 1-17.

Maraun, D., F. Wetterhall, A.M. Ireson, R.E. Chandler, E.J. Kendon, M. Widmann, S. Brienen, H.W. Rust, T. Sauter, M. Theme, V.K.C. Venema, K.P. Chun, C.M. Goodess, R.G. Jones, C. Onof, M. Vrac, et I. Thiele-Eich, 2010 : Precipitation downscaling under climate change : Recent developments to bridge the gap between dynamical models and the end user. *Reviews of Geophysics*, **48**(3), RG3003.

Maurer, E.P., et H.G. Hidalgo, 2007 : Utility of daily vs. monthly large-scale climate data : an intercomparison of two statistical downscaling methods. *Hydrology and Earth System Sciences Discussions*, **4**, 3413-3440.

Menabde, M., D. Harris, A. Seed, G. Austin, et D. Stow, 1997 : Multiscaling properties of rainfall and bounded random cascades. *Water Resources Research*, **33**(12), 2823-2830.

Mesinger, F., G. DiMego, E. Kalnay, P. Shafran, W. Ebisuzaki, D. Jovic, J. Woollen, K. Mitchell, E. Rogers, M. Ek, Y. Fan, R. Grumbine, W. Higgins, H. Li, Y. Lin, G. Manikin, D. Parrish, et W. Shi, 2004 : North American Regional Reanalysis. *15th Symp. on Global Change and Climate Variations*, paper P1.1, Combined Preprints CD-ROM, 84th AMS Annual Meeting, Seattle, WA.

Min, Y.-M., V.N. Kryjov, et J.-H. Oh, 2011 : Probabilistic interpretation of regression-based downscaled seasonal ensemble predictions with the estimation of uncertainty. *Journal of Geophysical Research D : Atmospheres*, **116**, D08101.

Morin, E., D.C. Goodrich, R.A. Maddox, X. Gao, H.V. Gupta, et S. Sorooshian, 2006 : Spatial patterns in thunderstorm rainfall events and their coupling with watershed hydrological response. *Advances in Water Resources*, **29**, 843-860.

Nakicenovic, N., et R. Swart, 2000 : *Special Report on Emissions Scenarios. A Special Report of Working Group III of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 599 pp.

Neilson, D., B. Taylor, G. Duke, J. Byrne, S. Kienzle, et T. Van der Gulik, 2010 :

Development and verification of daily gridded climate surfaces in the Okanagan basin of British Columbia. *Canadian Water Resources Journal*, **35**(2), 131-154.

Nykanen, D.K., 2008 : Linkages between Orographic Forcing and the Scaling Properties of Convective Rainfall in Mountainous Regions. *Journal of Hydrometeorology*, **9**, 327-347.

Nykanen, D.K., et D. Harris, 2003 : Orographic influences on the multiscale statistical properties of precipitation. *Journal of Geophysical Research D : Atmospheres*, **108**(D8), 8381.

Olsson, J., 1998 : Evaluation of a scaling cascade model for temporal rainfall disaggregation. *Hydrology and Earth System Sciences*, **2**(1), 19-30.

Onibon, H., T. Lebel, A. Afouda, and G. Guillot, 2004 : Gibbs sampling for conditional spatial disaggregation of rain fields. *Water Resources Research*, **40**(8), W084011-W0840112.

Over, T.M., et V.K. Gupta, 1996 : A space-time theory of mesoscale rainfall using random cascades. *Journal of Geophysical Research*, **101**(D21), 26,319-26,331.

Pandey, G.R., D.R. Cayan, M.D. Dettinger, K.P. Georgakakos, 2000 : A Hybrid Orographic plus Statistical Model for Downscaling Daily Precipitation in Northern California. *Journal of Hydrometeorology*, **1**(6), 491-506.

Pathirana, A., et S. Herath, 2002 : Multifractal modelling and simulation of rain fields exhibiting spatial heterogeneity. *Hydrology and Earth System Sciences*, **6**(4), 695-708.

Pegram, G.G.S., et A.N. Clothier, 2001 : Downscaling rainfields in space and time, using the String of Beads model in time series mode. *Hydrology and Earth System Sciences*, **5**(2), 175-186.

Peixoto, J.P., et A.H. Oort, 1992. Physics of climate. American Institute of Physics, New York, 520 p.

Perica, S., et E. Foufoula-Georgiou, 1996 : Model for multiscale disaggregation of spatial

rainfall based on coupling meteorological and scaling descriptions. *Journal of Geophysical Research*, **101**(D21), 26,347-26,361.

Quilbé, R., A.N. Rousseau, J.-S. Moquet, N.B. Trinh, Y. Dibike, P. Gachon, et D. Chaumont, 2008 : Assessing the Effect of Climate Change on River Flow Using General Circulation Models and Hydrological Modelling - Application to the Chaudière River, Québec, Canada. *Canadian Water Resources Journal*, **33**(1), 73-94.

Rawls, W.J., et D.L. Brakensiek, 1989 : Estimation of soil water retention and hydraulic properties. In *Unsaturated Flow in Hydrologic Modeling, Theory and Practice*, H.J. Morel-Seytoux, éd., 275-300. Norwell (MA-USA) : Kluwer Academic Publishers.

Renard, F., et J. Comby, 2006 : Evaluation de techniques d'interpolation spatiale de la pluie en milieu urbain pour une meilleure gestion d'événements extrêmes : le cas du Grand Lyon. *La Houille Blanche*, **6**, 73-78.

Rhea, J. O., 1978 : *Orographic Precipitation Model for Hydrometeorological Use*. Ph.D. Dissertation, Colorado State University, Department of Atmospheric Science, Fort Collins, Colorado, 198 p.

Ricard, S., 2008 : *Évaluation du modèle HYDROTEL à estimer la recharge des aquifères : Application sur le bassin versant de la rivière Châteauguay*. Mémoire de maîtrise, Institut National de la Recherche Scientifique - Centre eau, terre et environnement, Québec, 194 p.

Roberts, G.O., et A.F.M. Smith, 1994 : Simple conditions for the convergence of the Gibbs sampler and Metropolis-Hastings algorithms. *Stochastic Processes and Their Applications*, **49**, 207-216.

Rodríguez-Puebla, C., A.H. Encinas, et J. Sáenz, 2001 : Winter precipitation over the Iberian peninsula and its relationship to circulation indices. *Hydrology and Earth System Sciences*, **5**(2), 233-244.

Rousseau, A.N., J.-P. Fortin, R. Turcotte, A. Royer, S. Savary, F. Quévy, P. Noël, et

C. Paniconi, 2011 : PHYSITEL, a specialized GIS for supporting the implementation of distributed hydrological models. *Water News - Official Magazine of the Canadian Water Resources Association*, **31**(1), 18-20.

Salathé, E.P. Jr., 2005 : Downscaling simulations of future global climate with application to hydrologic modelling. *International Journal of Climatology*, **25**, 419-436.

Salathé, E.P. Jr., W.P. Mote, et M.W. Wiley, 2007 : Review of scenario selection and downscaling methods for the assessment of climate change impacts on hydrology in the United States pacific northwest. *International Journal of Climatology*, **27**, 1611-1621.

Saltikoff, E., J. Koistinen, et H. Hohti, 2000 : Experience of Real Time Spatial Adjustment of the Z-R Relation According to Water Phase of Hydrometeors. *Physics and Chemistry of the Earth (B)*, **25** (10-12), 1017-1020.

Sanso, B., et L. Guenni, 2004 : A Bayesian approach to compare observed rainfall data to deterministic simulations. *Environmetrics*, **15**, 597-612.

Schuler, T.V., P. Crochet, R. Hock, M. Jackson, I. Barstad, et T. Johannesson, 2008 : Distribution of snow accumulation on the Svartisen ice cap, Norway, assessed by a model of orographic precipitation. *Hydrological Processes*, **22**(19), 3998-4008.

Schuurmans, J.M., et M.F.P. Bierkens, 2007. Effect of spatial distribution of daily rainfall on interior catchment response of a distributed hydrological model. *Hydrology and Earth System Sciences*, **11**, 677-693.

Segond, M.-L., C. Onof, et H.S. Wheater, 2006 : Spatial-temporal disaggregation of daily rainfall from a generalized linear model. *Journal of Hydrology*, **331**, 674-689.

Segond, M.-L., H.S. Wheater, et C. Onof, 2007 : The significance of spatial rainfall representation for flood runoff estimation : A numerical evaluation based on the Lee catchment, UK. *Journal of Hydrology*, **347**, 116-131.

Sharif, M., et D.H. Burn, 2004 : Development and application of K-nearest neighbour

weather generating model. *57th Canadian Water Resources Association Annual Congress Water and Climate Change : Knowledge for Better Adaptation*, 6 pages, 16-18 juin 2004, Montreal, Québec, Canada.

Sharma, D., A.D. Gupta, et M.S. Babel, 2007 : Spatial disaggregation of bias-corrected GCM precipitation for improved hydrologic simulation : Ping River Basin, Thailand. *Hydrology and Earth System Sciences*, **11**(4), 1373-1390.

Sinclair, M.R., 1994 : A diagnostic model for estimating orographic precipitation. *Journal of Applied Meteorology*, **33**(10), 1163-1175.

Smith, R.B., et I. Barstad, 2004 : A linear theory of orographic precipitation. *Journal of the Atmospheric Sciences*, **61**(12), 1377-1391.

Starks, P.J., et D.N. Moriasi, 2009 : Spatial Resolution Effect of Precipitation Data on SWAT Calibration and Performance : Implications for CEAP. *Transactions of the ASABE*, **52**(4), 1171-1180.

Teo, C.K., et D.I.F. Grimes, 2007 : Stochastic modelling of rainfall from satellite data. *Journal of Hydrology*, **346**, 33-50.

Tolika, K., P. Maheras, M. Vafiadis, H.A. Flokas, et A. Arseni-Papadimitriou, 2007 : Simulation of seasonal precipitation and raindays over Greece : a statistical downscaling technique based on artificial neural networks (ANNs). *Journal International of Climatology*, **27**, 861-881.

Turcotte, R., 2010 : *Éléments du calage et de l'implantation d'un modèle hydrologique dans un perspective d'utilisation opérationnelle en prévision hydrologique*. Thèse de doctorat, Institut National de la Recherche Scientifique - Centre eau, terre et environnement, Québec, 194 p.

Turcotte, R., J.-P. Fortin, A.N. Rousseau, S. Massicotte, et J.-P. Villeneuve, 2001 : Determination of the drainage structure of a watershed using a digital elevation model and a digital river and lake network. *Journal of Hydrology*, **240**, 225-242.

Turcotte, R., A.N. Rousseau, J.-P. Fortin, et J.-P. Villeneuve, 2003 : Development of a process-oriented, multiple-objective, hydrological calibration strategy accounting for model structure. Dans Duan, Q., S. Sorooshian, H. Gupta, A.N. Rousseau, et R. Turcotte, *Advances in Calibration of Watershed Models*, Water Science & Application, **6**, AGU, 153-163.

Turcotte, R., P. Lacombe, C. Dimnik, et J.-P. Villeneuve, 2004 : Pr evision hydrologique distribu ee pour la gestion des barrages publics du Qu ebec. *Canadian Journal of Civil Engineering*, **31**, 308-320.

Turcotte, R., L.-G. Fortin, V. Fortin, J.-P. Fortin, et J.-P. Villeneuve, 2007 : Operational analysis of the spatial distribution and the temporal evolution of the snowpack water equivalent in southern Qu ebec, Canada. *Nordic Hydrology*, **38**(3), 211-234.

Van Den Berg, M.J., S. Vandenberghe, B. De Baets, et N.E.C. Verhoest, 2011 : Copula-based downscaling of spatial rainfall : a proof of concept. *Hydrology and Earth System Sciences*, **15**, 1445-1457.

Van Der Heyden, H., 2008 : * tude de faisabilit e visant l'am enagement int egr e de bassins de stockage, de bassins de s dimentation, de marais filtrants, de zones de d bordement et de digues afin de constituer des r serves d'eau destin es   l'irrigation et de r duire les impacts agricole et environnemental des d bordements des ruisseaux Norton, Cranberry, Gibeault Delisle et de la rivi re L'Acadie*. Compagnie de recherche Phytodata inc., Projet no. 5282. 146 p.

Velazquez, J.A., T. Petit, A. Lavoie, M.-A. Boucher, R. Turcotte, V. Fortin, et F. Anctil, 2009 : An evaluation of the Canadian global meteorological ensemble prediction system for short-term hydrological forecasting. *Hydrology and Earth System Sciences*, **13**(11), 2221-2231.

Veneziano, D., P. Furcolo, et V. Iacobellis, 2006 : Imperfect scaling of time and space-time rainfall. *Journal of Hydrology*, **322**, 105-119.

Western Regional Climate Center, 2002 : Climate of Washington. Site internet du Western Regional Climate Center, [En ligne]. <http://www.wrcc.dri.edu/narratives/WASHINGTON.htm> (Consulté en octobre 2010).

Wheater, H.S., T.J. Jolley, C. Onof, N. Mackay, et R.E. Chandler, 1999 : Analysis of aggregation and disaggregation effects for grid-based hydrological models and the development of improved precipitation disaggregation procedures for GCMs. *Hydrology and Earth System Sciences*, **3**, 95-108.

Wilby, R.L., L.E. Hay, et G.H. Leavesley, 1999 : A comparison of downscaled and raw GCM output : implications for climate change scenarios in the San Juan River basin, Colorado. *Journal of Hydrology*, **225**, 67-91.

Wood, S.J., D.A. Jones, et R.J. Moore, 2000a : Accuracy of rainfall measurement for scales of hydrological interest. *Hydrology and Earth System Sciences*, **4**(4), 531-543.

Wood, S.J., D.A. Jones, et R.J. Moore, 2000b : Static and dynamic calibration of radar data for hydrological use. *Hydrology and Earth System Sciences*, **4**(4), 545-554.

Yang, D., B.E. Goodison, J.R. Metcalfe, P.Y.T. Louie, G. Leavesley, D. Emerson, C.L. Hanson, V.S. Golubev, E. Elomaa, T. Gunther, T. Pangburn, E. Kang, et J. Milkovic, 1999 : Quantification of precipitation measurement discontinuity induced by wind shields on national gauges. *Water Resources Research*, **35**(2), 491-508.

Yates, D., S. Gangopadhyay, B. Rajagopalan, et K. Strzepek, 2003 : A technique for generating regional climate scenarios using a nearest-neighbor algorithm. *Water Resources Research*, **39**(7), 1199-1213.

Yu, Z., M. Lakhtakia, B. Yarnal, R.A. White, D.A. Miller, B. Frakes, E.J. Barron, C.J. Duffy, et F.W. Schwartz, 1999 : Simulating the river-basin response to atmospheric forcing : a mesoscale meteorological model and a hydrological model system. *Journal of Hydrology*, **218**, 72-91.

Yu, Z., E.J. Barron, B. Yarnal, M.N. Lakhtakia, R.A. White, D. Pollard, et D.A. Miller,

2002 : Evaluation of basin-scale hydrologic response to a multi-storm simulation. *Journal of Hydrology*, **257**, 212-225.

# Annexe A

## Liste des abréviations et des sigles

ACP : Analyse en Composantes Principales

BV3C : Bilan Vertical en Trois Couches

CAPE : *Convective Available Potential Energy* ou « Énergie potentielle de convection disponible »

CC : Changements Climatiques

CEHQ : Centre d'Expertise Hydrique du Québec

CP : Composante Principale

GES : Gaz à Effet de Serre

GEV : *Generalized Extreme Values distribution* ou distribution des valeurs extrêmes généralisée

GMRF : *Gaussian Markov Random Field* ou marche aléatoire gaussienne

HGM : Hydrogramme GéoMorphologique

iid : Indépendant et Identiquement Distribué

MCG : Modèle de Climat Global

MCMC : Monte Carlo par Chaîne de Markov ou *Markov Chain Monte Carlo*

MDDEP : Ministère du Développement Durable, de l'Environnement et des Parcs du Québec

MNA : Modèle Numérique d'Altitude

MPO : Modèle de Précipitation Orographique

MRC : Modèle Régional de Climat

MRCC : Modèle Régional Canadien du Climat

MRF : *Markov Random Field* ou marche aléatoire

MV : Maximum de vraisemblance

NARR : *North American Regional Reanalysis*

NCEP : *National Center for Environmental Prediction*

NERFC : *Northeast RFC*

NWRFC : *Northwest RFC*

NWS : *National Weather Service*

ONA : Oscillation Nord-Atlantique

PCIC : *Pacific Climate Impacts Consortium*

PRISM : *Parameter-elevation Regressions on Independent Slopes Model*

RFC : *River Forecast Center*

SERFC : *Southeast RFC*

UHRH : Unité Hydrologique Relativement Homogène

USGS : *United States Geological Survey*

UTC : Temps Universel Coordonné



# Annexe B

## Liste des variables utilisées

À moins d'avis contraire, toutes les variables mentionnées dans cette liste sont au pas de temps journalier.

$$A_{/}^{(i,j)} : \frac{R_{i-1,j-1} + R_{i+1,j+1}}{2}$$

$$A_{|}^{(i,j)} : \frac{R_{i-1,j} + R_{i+1,j}}{2}$$

$$A_{\setminus}^{(i,j)} : \frac{R_{i-1,j+1} + R_{i+1,j-1}}{2}$$

$$A_{-}^{(i,j)} : \frac{R_{i,j-1} + R_{i,j+1}}{2}$$

$$\vec{A}^{(i,j)} : \left( A_{/}^{(i,j)}, A_{|}^{(i,j)}, A_{\setminus}^{(i,j)}, A_{-}^{(i,j)} \right)'$$

$$\bar{A}^{(i,j)} : \frac{A_{/}^{(i,j)} + A_{|}^{(i,j)} + A_{\setminus}^{(i,j)} + A_{-}^{(i,j)}}{4} = \text{la moyenne des huit voisins les plus proches.}$$

$C_{i,j}$  : Valeur moyenne de CAPE (J/kg) sur la tuile contenant le pixel  $(i, j)$ .

$F_X(x) = \mathbb{P}[X \leq x]$  : Fonction de répartition de la variable  $X$  évaluée au point  $x$ .

$k_c$  : *Période d'autocorrélation*, soit le nombre d'itérations de l'échantillonnage de Gibbs nécessaires pour que l'autocorrélation soit négligeable.

$k_r$  : Itération de l'échantillonnage de Gibbs pour laquelle la *période de rodage* prend fin, soit le plus petit entier pour lequel la distribution des valeurs générées est jugée suffisamment

proche de la distribution cible inconnue.

$N_{p_1} \times N_{p_2} \equiv N_p^2$  : Nombre de pixels par tuile.

$N_{P_1} \times N_{P_2} \equiv N_P^2$  : Nombre de pixels sur la région étudiée.

$N_{t_1} \times N_{t_2} \equiv N_t^2$  : Nombre de tuiles sur la région étudiée.

$O_{i,j}^{(L)}$  : Orientation (radian) de la pente au pixel  $(i, j)$  calculée à partir des  $\pm L$  voisins les plus proches dans chaque direction (total de  $(2L + 1)^2 - 1$  les plus proches).

$P_{i,j}^{(L)}$  : Pente (m d'élévation / m horizontal) au pixel  $(i, j)$  calculée à partir des  $\pm L$  voisins les plus proches dans chaque direction (total de  $(2L + 1)^2 - 1$  les plus proches).

$R_{i,j}$  : Précipitation accumulée (mm) sur le pixel  $(i, j)$  pour une journée donnée.

$\vec{R}$  : Vecteur contenant les accumulations de précipitation (mm) de tous les pixels sur la région étudiée.

$\vec{R}_{-(i,j)}$  : Vecteur contenant les accumulations de précipitation (mm) de tous les pixels, sauf le pixel  $(i, j)$ .

$\bar{R}_{i,j}$  :  $\mathbb{E} [R_{i,j} | \vec{R}_{-(i,j)}]$  tel que défini aux équations (3.4) (Modèle des Parties I et III), (3.2) (Modèle 1 de la Partie II) et (6.6) (Modèle 2 de la Partie II).

$\mathbf{R}_{i,j}$  : Précipitation accumulée (mm) sur la tuile contenant le pixel  $(i, j)$ .

$\vec{\mathbf{R}}$  : Vecteur contenant les accumulations de précipitation (mm) de toutes les tuiles sur la région étudiée.

$S$  : Nombre de champs désagrégés par jour.

$T_{i,j}^{(L)}$  : Anomalie topographique (m) au pixel  $(i, j)$  calculée à partir des  $\pm L$  voisins les plus proches dans chaque direction (total de  $(2L + 1)^2 - 1$  les plus proches).

$V_{i,j}^{(p)}$  : Vitesse du vent (m/s) évaluée à la pression  $p$  au-dessus de la tuile contenant le pixel  $(i, j)$ .

$Z_{i,j}$  : Altitude (m) du pixel  $(i, j)$ .

$\mathbf{Z}_{i,j}$  : Altitude de la tuile contenant le pixel  $(i, j)$ .

$\bar{Z}_{i,j}^{(L)}$  : Altitude moyenne sur les  $\pm L$  voisins les plus proches du pixel  $(i, j)$  dans chaque direction (total de  $(2L + 1)^2 - 1$  les plus proches).

$\omega_{i,j}^{(p)}$  : Direction du vent (en  $^\circ$ ) par rapport à l'axe ouest-est de la grille de précipitation évaluée à la pression  $p$  au-dessus de la tuile contenant le pixel  $(i, j)$ .

$\rho[(i, j), (i', j')]$  : Coefficient de corrélation non-paramétrique de Kendall pour la paire de pixels  $(i, j)$  et  $(i', j')$  calculé sur les jours où  $R_{i,j}$  et  $R_{i',j'}$  sont tous les deux plus grands que 0.

$$\sigma_{R_{i,j}} : \sqrt{\text{Var} [R_{i,j} | \vec{R}_{-(i,j)}]}$$



# Annexe C

## Échantillonnage de Gibbs

L'échantillonnage de Gibbs (Geman et Geman, 1984; Roberts et Smith, 1994) est une méthode Monte Carlo par chaîne de Markov (MCMC pour *Markov Chain Monte Carlo* en anglais) qui permet de générer les valeurs d'un vecteur aléatoire à distribution inconnue (ici, la précipitation sur tous les pixels de la région) à partir de la distribution conditionnelle de chaque élément du vecteur lorsque celle-ci est connue. Définissons d'abord ce qu'est une chaîne de Markov.

**Définition 1.** *La suite de variables aléatoires  $(X_k, k \geq 0)$  à valeurs dans  $\mathbb{R}_+$  est une chaîne de Markov à espace d'état discret si, pour tout  $k \geq 1$ , ensemble  $\mathcal{A} \subset \mathbb{R}_+$  et point  $x_i \in \mathbb{R}_+$  où  $i \in \{0, \dots, k-1\}$ , les deux conditions suivantes sont satisfaites :*

1.  $\mathbb{P}[X_k \in \mathcal{A} | (X_0, \dots, X_{k-1}) = (x_0, \dots, x_{k-1})] = \mathbb{P}[X_k \in \mathcal{A} | X_{k-1} = x_{k-1}]$
2.  $\mathbb{P}[X_k \in \mathcal{A} | X_{k-1} = x_0] = \mathbb{P}[X_1 \in \mathcal{A} | X_0 = x_0]$

La première condition implique que sachant la dernière valeur de la chaîne, les valeurs précédentes n'amènent pas d'information supplémentaire pour prédire la prochaine valeur. La deuxième condition implique que les transitions de probabilités sont invariantes dans le temps. Pour le reste de la présente annexe,  $\mathcal{A}$  est un ensemble inclus dans  $\mathbb{R}_+$  et  $x$  est une valeur dans  $\mathbb{R}_+$ . Le *noyau markovien* ou *noyau de transition*  $P$  de la chaîne de Markov est défini par :

$$P(x, \mathcal{A}) = \mathbb{P}[X_k \in \mathcal{A} | X_{k-1} = x] \quad \forall k \geq 1 . \tag{C.1}$$

Notons  $\nu$  la distribution initiale ( $k = 0$ ), à valeurs dans  $\mathbb{R}_+$ , de la chaîne de Markov. Ainsi :

$$\nu(\mathcal{A}) = \mathbb{P}[X_0 \in \mathcal{A}] . \quad (\text{C.2})$$

La distribution de  $X_1$ , la chaîne de Markov au temps 1 est donc donnée par :

$$\mathbb{P}[X_1 \in \mathcal{A}] = \int_{\mathbb{R}_+} \mathbb{P}[X_1 \in \mathcal{A} | X_0 = x] \mathbb{P}[X_0 \in dx] \quad (\text{C.3})$$

$$= \int_{\mathbb{R}_+} P(x, \mathcal{A}) \nu(dx) \quad (\text{C.4})$$

$$\equiv (\nu P)(\mathcal{A}) . \quad (\text{C.5})$$

De même, pour tout  $k \geq 1$ ,

$$\mathbb{P}[X_k \in \mathcal{A}] \equiv (\nu P^k)(\mathcal{A}) . \quad (\text{C.6})$$

Soit  $\nu_1$  et  $\nu_2$  deux distributions initiales pour la chaîne de Markov. Sous certaines conditions de régularités qui sont rencontrées dans le contexte d'application de la thèse (Jones, 2004) :

$$\lim_{k \rightarrow \infty} \|(\nu_1 P^k)(\mathcal{A}) - (\nu_2 P^k)(\mathcal{A})\| = 0 . \quad (\text{C.7})$$

En pratique, cela implique que peu importe la distribution initiale utilisée, la chaîne de Markov converge toujours vers la même distribution, appelée *distribution stationnaire*.

**Définition 2.** *La distribution  $\pi$  est une distribution stationnaire de la chaîne de Markov ( $X_k, k \geq 0$ ) avec noyau markovien  $P$  si :*

$$(\pi P)(\mathcal{A}) = \pi(\mathcal{A}) . \quad (\text{C.8})$$

Sous les conditions de régularités (Jones, 2004), la distribution  $\pi$  est unique. Pour montrer que la chaîne de Markov converge vers  $\pi$ , remplaçons  $\nu_2$  de l'équation (C.7) par  $\pi$ . Il suit que pour tout  $k \geq 2$ , la distribution de  $X_k$  est donnée par  $(\pi P^k) = (\pi P)(P^{k-1}) = \pi(P^{k-1}) = (\pi P)(P^{k-2}) = \dots = \pi$ . Ainsi,

$$\lim_{k \rightarrow \infty} \left\| (\nu_1 P^k)(\mathcal{A}) - \pi(\mathcal{A}) \right\| = 0 . \quad (\text{C.9})$$

Donc, peu importe la distribution initiale, la chaîne de Markov convergera vers sa distribution stationnaire. En pratique, cela implique qu'une valeur de la distribution inconnue  $\pi$  peut être générée approximativement par une chaîne de Markov si le noyau markovien  $P$  tel que  $\pi P = \pi$  est connu et si le nombre d'itérations  $k$  est suffisamment grand.

L'échantillonnage de Gibbs utilise la théorie des chaînes de Markov pour simuler un vecteur aléatoire  $\vec{X}$  dont la distribution  $\pi$  est inconnue, mais dont la distribution de chacun des  $M$  éléments conditionnellement aux valeurs de tous les autres éléments est connue. Notons  $\pi_{X^{(m)}|\vec{X}^{(-m)}}$  la distribution de l'élément  $X^{(m)}$  ( $m \in \{1, \dots, M\}$ ) conditionnelle à  $\vec{X}^{(-m)}$ , le vecteur aléatoire  $\vec{X}$  sans l'élément  $m$ . Le principe de l'échantillonnage de Gibbs est simple : comme condition initiale, une valeur arbitraire est assignée à chacun des éléments du vecteur et ensuite chacun des  $M$  éléments est mis à jour un à la fois à partir des distributions  $\pi_{X^{(m)}|\vec{X}^{(-m)}}$ . La chaîne de Markov ( $\vec{X}_k, k \geq 0$ ) est alors une suite de vecteurs aléatoires  $\vec{X}_k = (X_k^{(1)}, \dots, X_k^{(M)})$  et elle a  $M$  noyaux markoviens intermédiaires tels que pour tout  $k \geq 1$ ,  $m \in \{2, \dots, M-1\}$ , vecteur de longueur  $M$   $\vec{x} = (x_1, \dots, x_M)$  à valeur dans  $\mathbb{R}_+^M$  et sous-ensemble de  $\mathbb{R}_+^M$  ( $\mathcal{A}_1, \dots, \mathcal{A}_M$ ) :

$$P_1(\vec{x}, \mathcal{A}_1) = \mathbb{P} \left[ X_k^{(1)} \in \mathcal{A}_1 \mid X_{k-1}^{(2)} = x_2, \dots, X_{k-1}^{(M)} = x_M \right] ; \quad (\text{C.10})$$

$$P_m(\vec{x}, \mathcal{A}_m) = \mathbb{P} \left[ X_k^{(m)} \in \mathcal{A}_m \mid \dots \right. \\ \left. X_k^{(1)} = x_1, \dots, X_k^{(m-1)} = x_{m-1}, X_{k-1}^{(m+1)} = x_{m+1}, \dots, X_{k-1}^{(M)} = x_M \right]; \quad (\text{C.11})$$

$$P_M(\vec{x}, \mathcal{A}_M) = \mathbb{P} \left[ X_k^{(M)} \in \mathcal{A}_M \mid X_k^{(1)} = x_1, \dots, X_{k-1}^{(M-1)} = x_{M-1} \right] . \quad (\text{C.12})$$

Pour passer de  $\vec{X}_k$  à  $\vec{X}_{k+1}$ , il faut mettre à jour tous les éléments de 1 à  $M$ . Le noyau markovien de la chaîne de Markov ( $\vec{X}_k, k \geq 0$ ) est ainsi donné par  $P = (P_1 \dots P_m)$ . Il est possible de montrer que la distribution stationnaire de ( $\vec{X}_k, k \geq 0$ ) est la distribution conjointe inconnue  $\pi$  (Geman et Geman, 1984). Donc, si un nombre assez grand de mises à jour des éléments du vecteur aléatoire est effectué, la distribution de ( $\vec{X}_k, k \geq 0$ ) sera approximativement  $\pi$ . Le nombre de mises à jour nécessaires peut dépendre de  $M$  (plus  $M$  est grand, plus la convergence risque d'être lente), de la structure de dépendance entre les

éléments de  $\vec{X}$  et de la distribution initiale. Une distribution initiale près de  $\pi$  prendra peu de temps à converger. Notons  $k_r$  l'itération pour laquelle la *période de rodage* prend fin, soit le plus petit entier pour lequel la distribution de  $\vec{X}_{k_r}$  est jugée suffisamment proche de celle de  $\pi$ . Notons  $k_c$  la *période d'autocorrélation*, soit le nombre d'itérations nécessaires pour que la corrélation entre  $\vec{X}_k$  et  $\vec{X}_{k+k_c}$  soit négligeable. Ainsi, si des valeurs de  $(\vec{X}_k, k \geq 0)$  sont sélectionnées en commençant après la période de rodage (*i.e.* après l'itération  $k_r$ ) et toutes séparées par au moins  $k_c$  itérations, l'échantillon obtenu sera approximativement iid de la distribution conjointe  $\pi$ .

Pour la présente application, le vecteur aléatoire d'intérêt est  $\vec{R}$ , soit la précipitation journalière sur tous les pixels de la région étudiée. L'échantillonnage de Gibbs peut ainsi être utilisé pour simuler  $\vec{R}$  si la distribution statistique de  $R_{i,j} \mid \vec{R}_{-(i,j)}$ , soit la distribution de la précipitation sur le pixel  $(i, j)$  sachant la précipitation sur tous les autres pixels, est connue pour tout pixel  $(i, j)$  de la région. Pour la Partie I, la Section 3.2 a montré que cette distribution peut être estimée par une lognormale d'espérance et d'écart-type donnés par les équations (3.4) et (3.7), respectivement. Il est alors supposé que le vecteur fonction des huit voisins les plus proches  $\vec{A}^{(i,j)}$  contient toute l'information pertinente à la désagrégation contenue dans  $\vec{R}_{-(i,j)}$ . D'après le Tableau 3.1, cette hypothèse est réaliste.

## Annexe D

Résultats de la désagrégation sur les pixels de 7,6, 11,4 et 15,2 km sur une région avec peu de relief

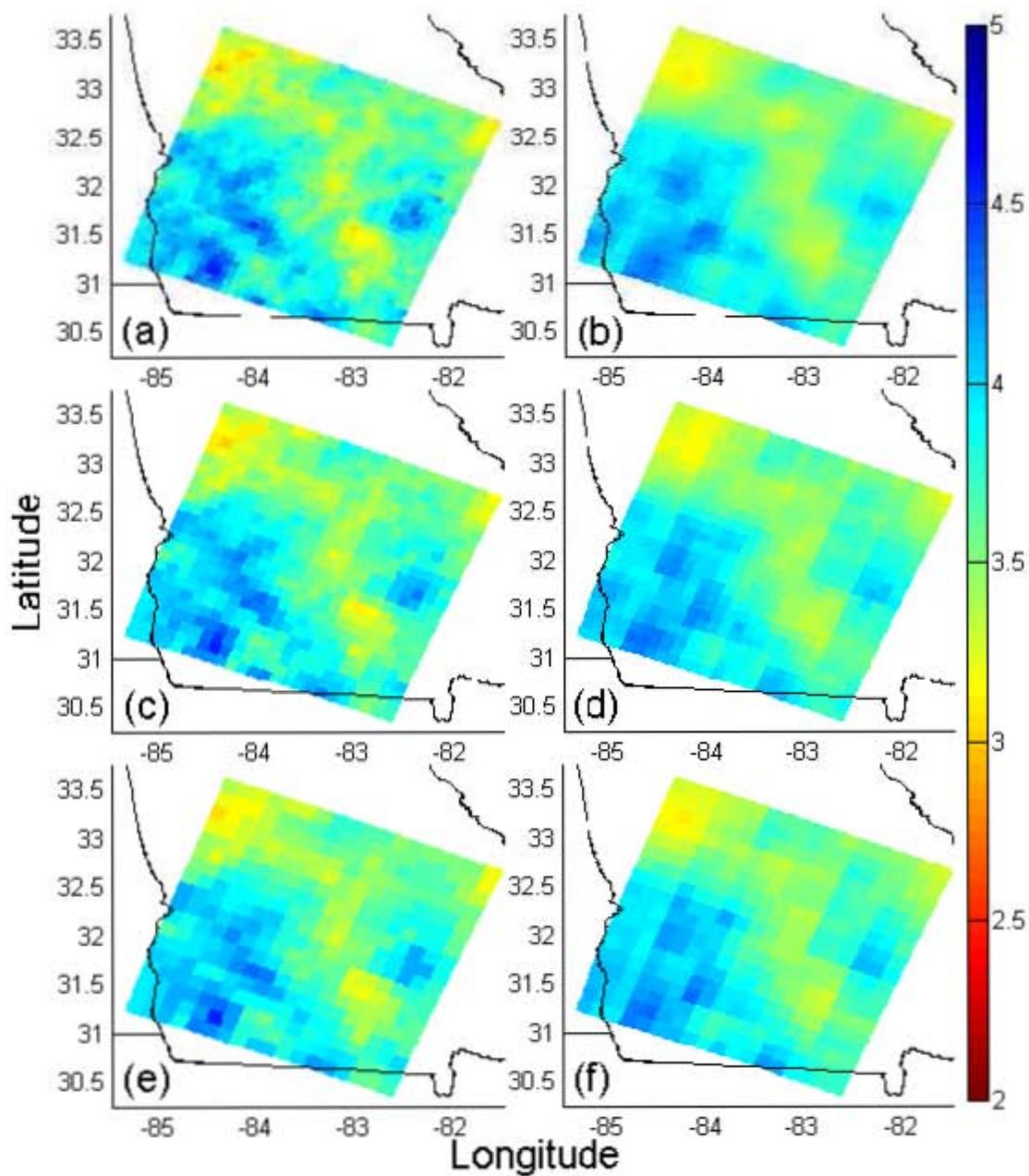


Figure D.1 – Précipitation moyenne (mm/jour) sur les 801 jours (2006-2008) pour les observations (gauche) et pour les huit champs désagrégés (droite) pour les pixels de taille (a-b) 7,6, (c-d) 11,4 et (e-f) 15,2 km.

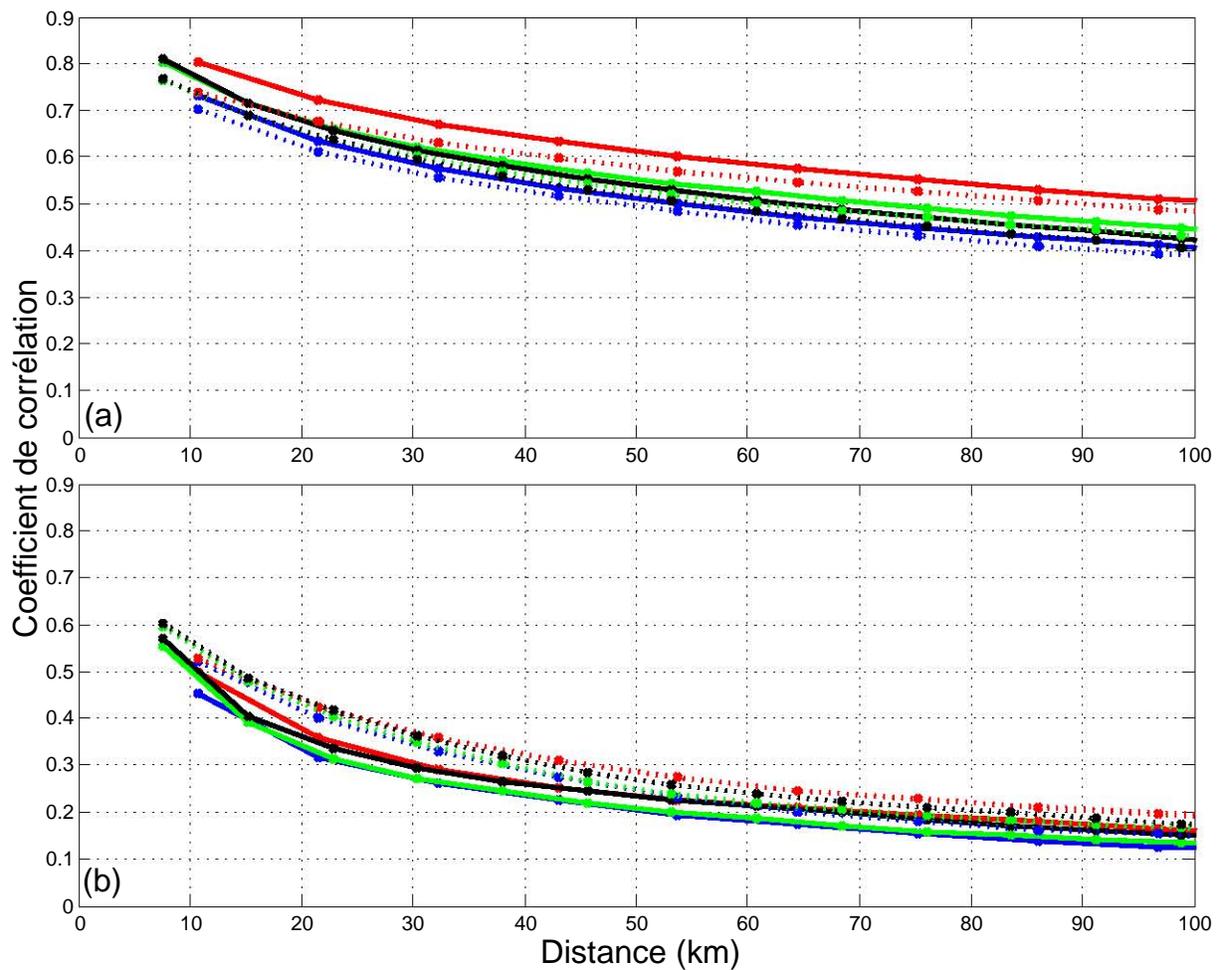


Figure D.2 – Coefficient de corrélation moyen dans les quatre directions calculé sur les jours avec (a) CAPE < 1000 J/kg et (b) CAPE > 1000 J/kg pour les pixels de 7,6 km observés (lignes continues) et désagrégés (lignes pointillées) : 45° (rouge), 90° (vert), 135° (bleu) et 0° (noir).

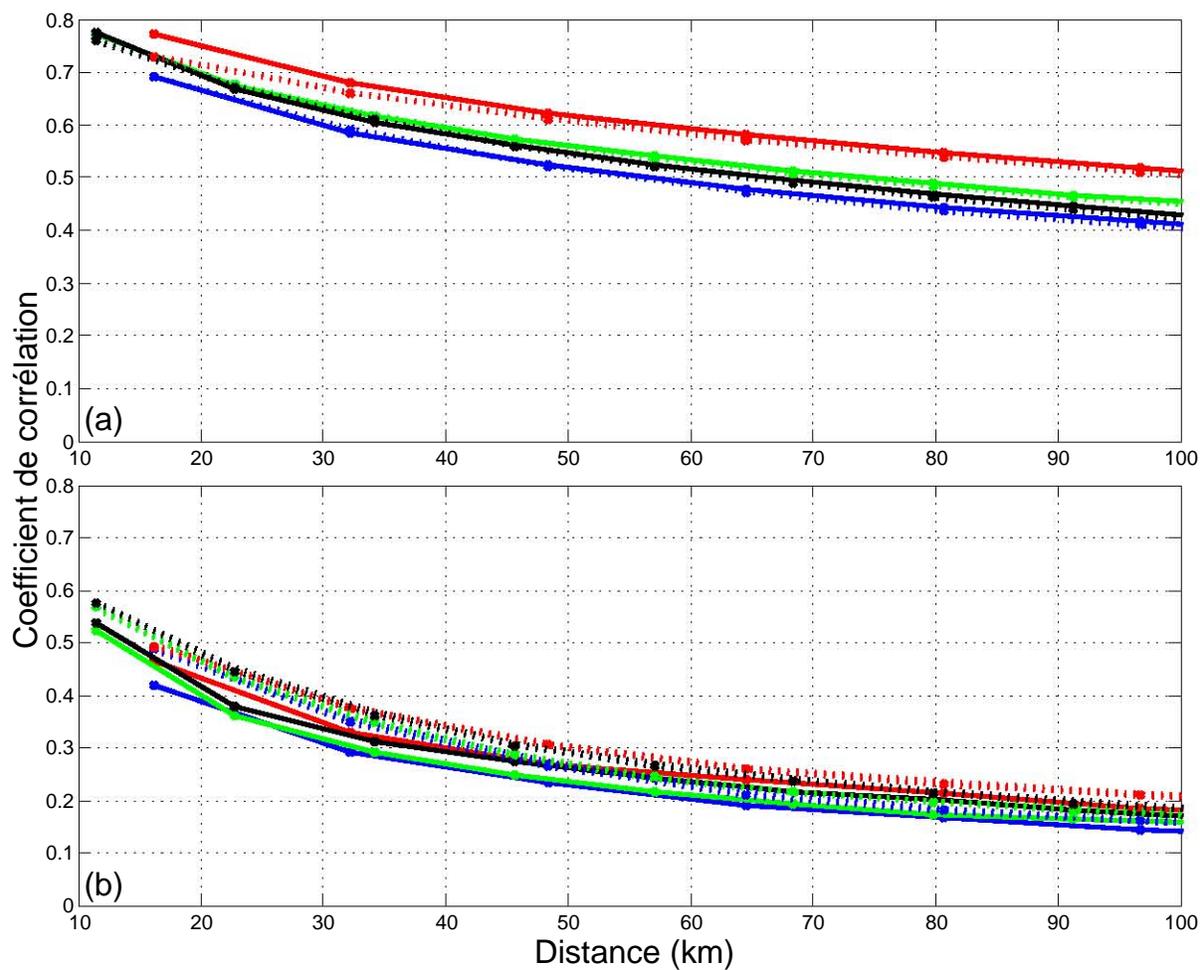


Figure D.3 – Coefficient de corrélation moyen dans les quatre directions calculé sur les jours avec (a)  $CAPE < 1000 \text{ J/kg}$  et (b)  $CAPE > 1000 \text{ J/kg}$  pour les pixels de 11,4 km observés (lignes continues) et désagrégés (lignes pointillées) :  $45^\circ$  (rouge),  $90^\circ$  (vert),  $135^\circ$  (bleu) et  $0^\circ$  (noir).

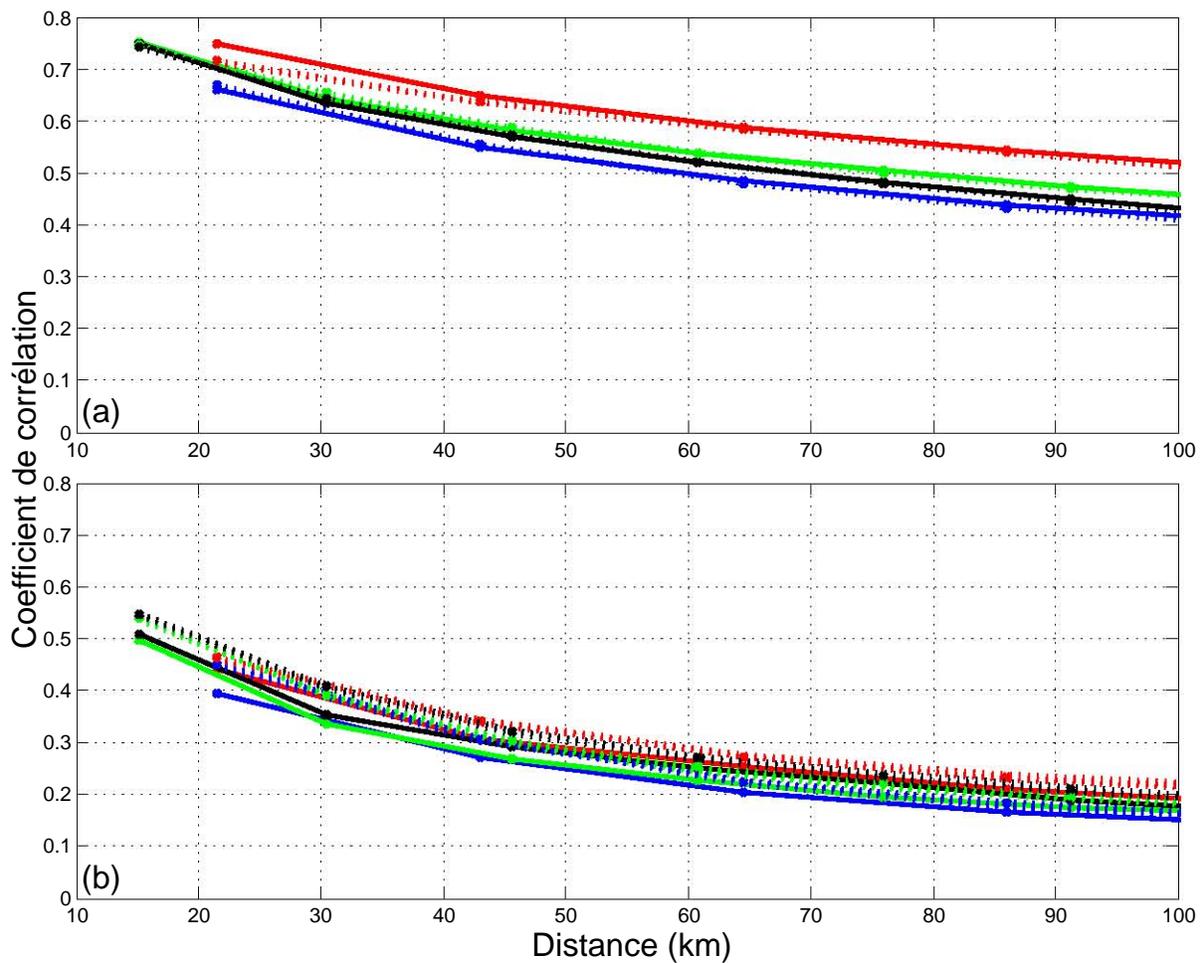


Figure D.4 – Coefficient de corrélation moyen dans les quatre directions calculé sur les jours avec (a)  $CAPE < 1000$  J/kg et (b)  $CAPE > 1000$  J/kg pour les pixels de 15,2 km observés (lignes continues) et désagrégés (lignes pointillées) : 45° (rouge), 90° (vert), 135° (bleu) et 0° (noir).

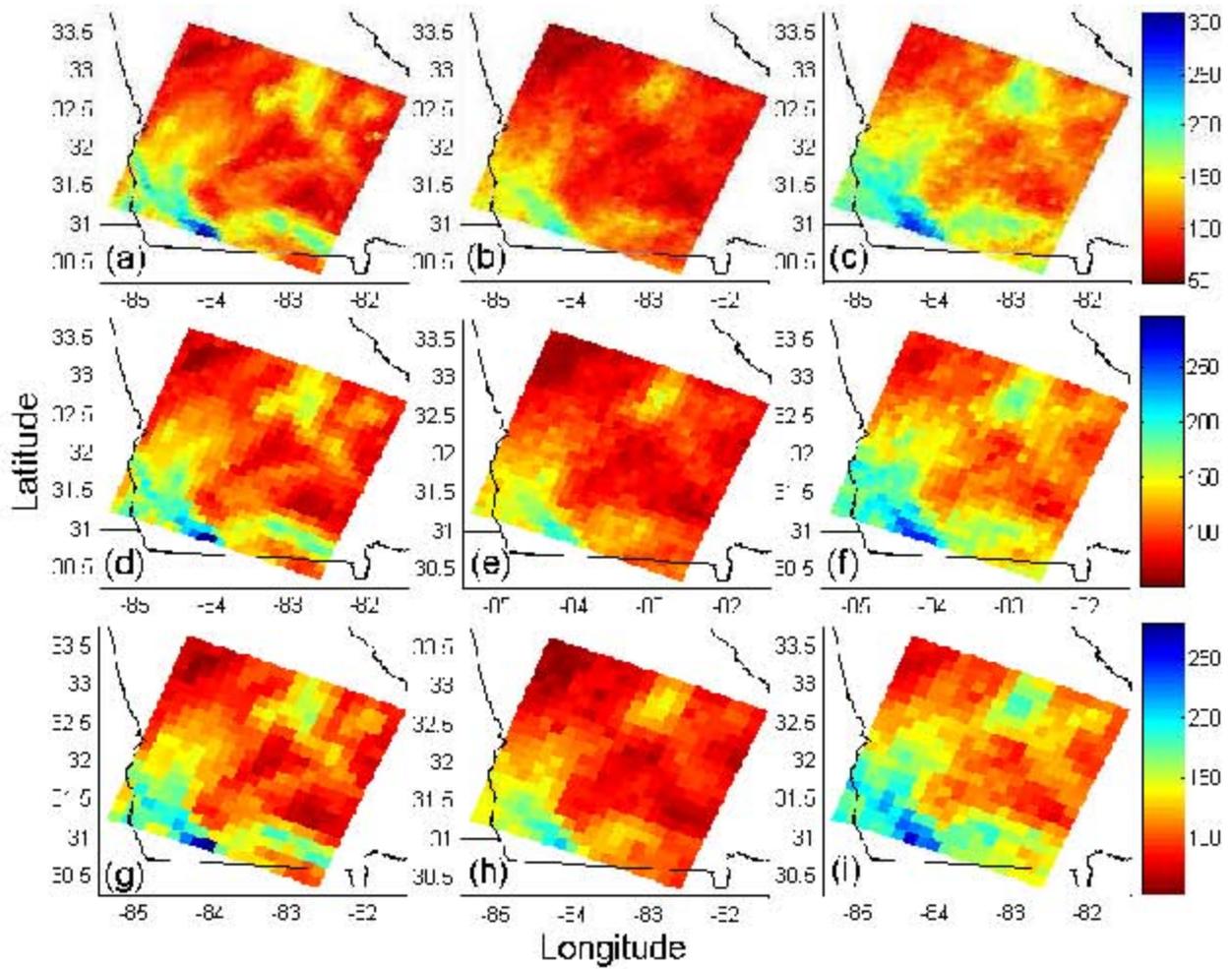


Figure D.5 – Précipitation journalière maximale (mm) sur les pixels de 7,6 (haut), 11,4 (centre) et 15,2 km (bas) durant la période 2006-2008 : observée (gauche), plus petite (milieu) et plus grande (droite) valeurs maximales journalières désagrégées.

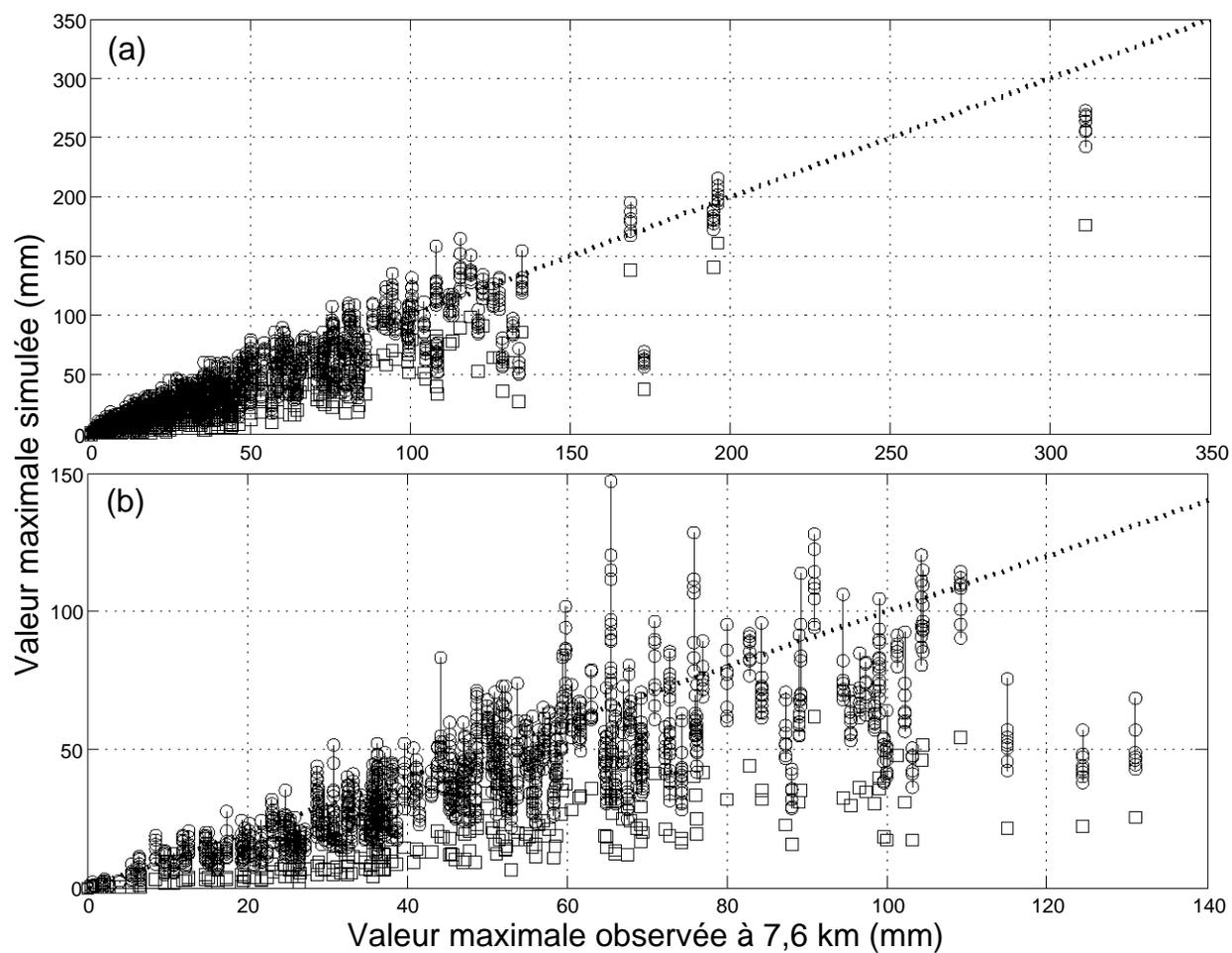


Figure D.6 – Précipitations maximales observée et désagrégées sur les pixels de 7,6 km des  $6 \times 6$  tuiles centrales pour chaque jour avec (a) CAPE < 1000 J/kg et (b) CAPE > 1000 J/kg. Les cercles sont les valeurs maximales désagrégées des huit réalisations et les carrés les valeurs maximales sur les tuiles de 45,6 km utilisées en entrée dans le modèle.

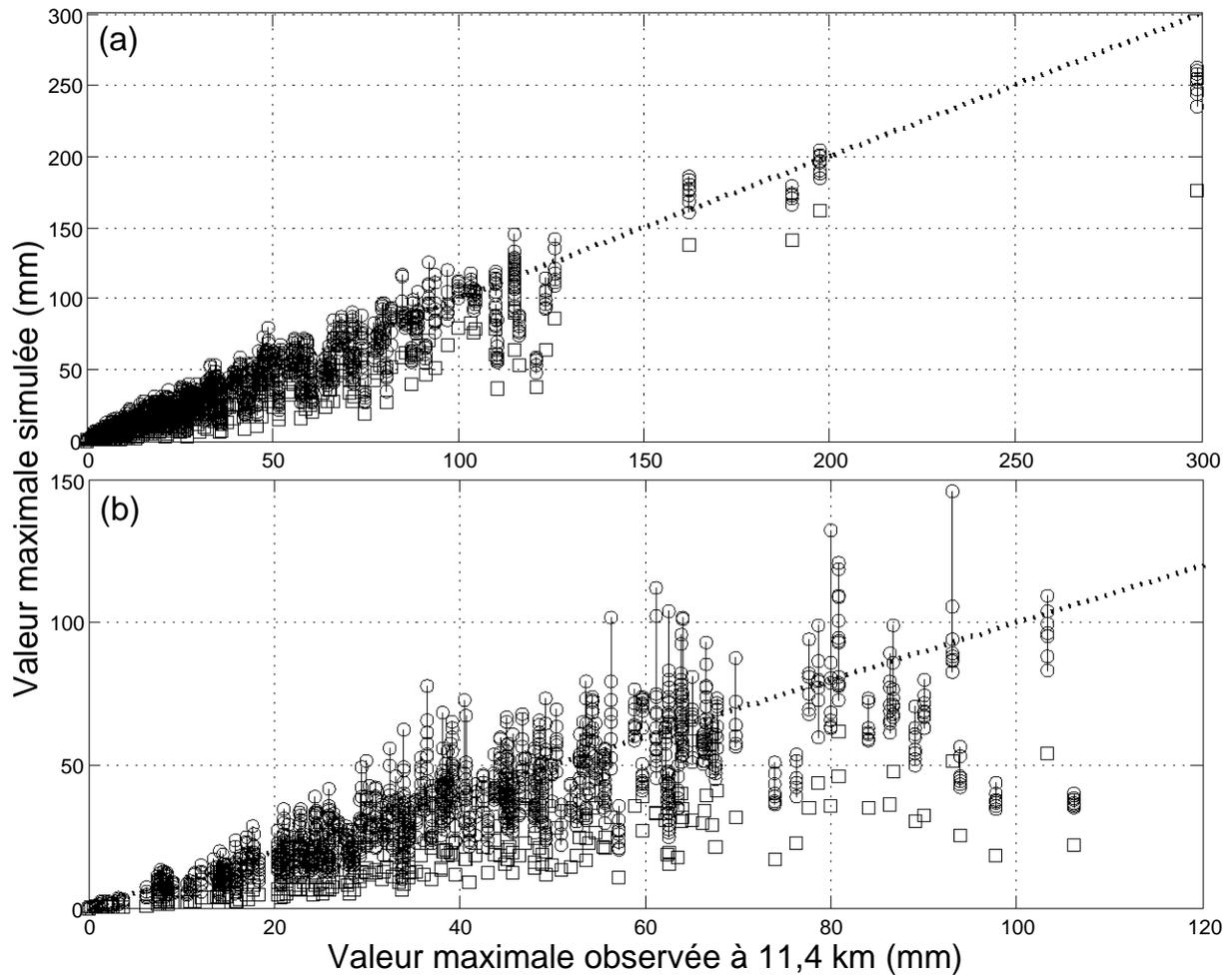


Figure D.7 – Précipitations maximales observée et désagrégées sur les pixels de 11,4 km des  $6 \times 6$  tuiles centrales pour chaque jour avec (a) CAPE  $< 1000$  J/kg et (b) CAPE  $> 1000$  J/kg. Les cercles sont les valeurs maximales désagrégées des huit réalisations et les carrés les valeurs maximales sur les tuiles de 45,6 km utilisées en entrée dans le modèle.

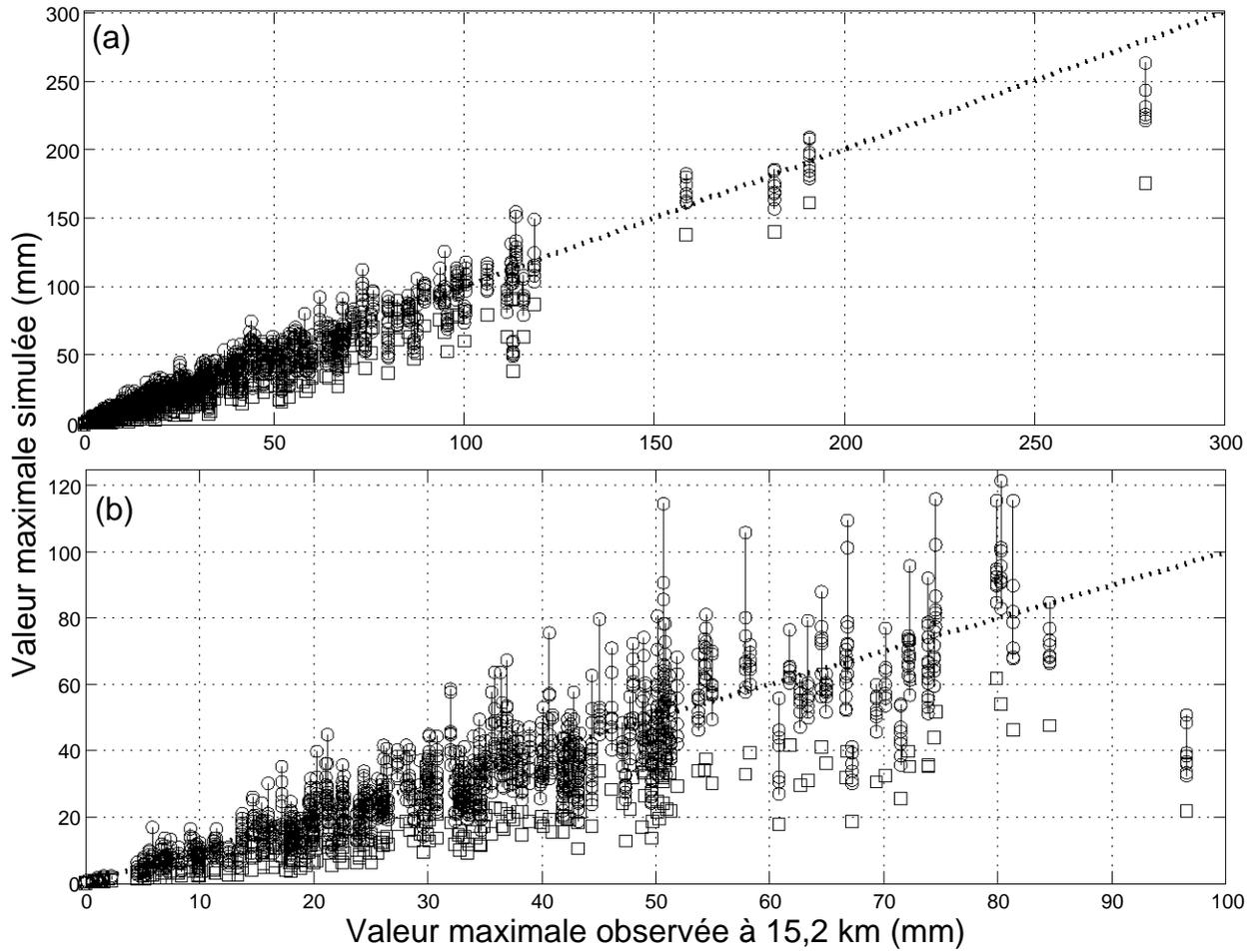


Figure D.8 – Précipitations maximales observée et désagrégées sur les pixels de 15,2 km des  $6 \times 6$  tuiles centrales pour chaque jour avec (a) CAPE  $< 1000$  J/kg et (b) CAPE  $> 1000$  J/kg. Les cercles sont les valeurs maximales désagrégées des huit réalisations et les carrés les valeurs maximales sur les tuiles de 45,6 km utilisées en entrée dans le modèle.



## **Annexe E**

**Résultats de la désagrégation sur les pixels de 8,8, 13,2, 17,6 et 26,4 km sur une région montagneuse**

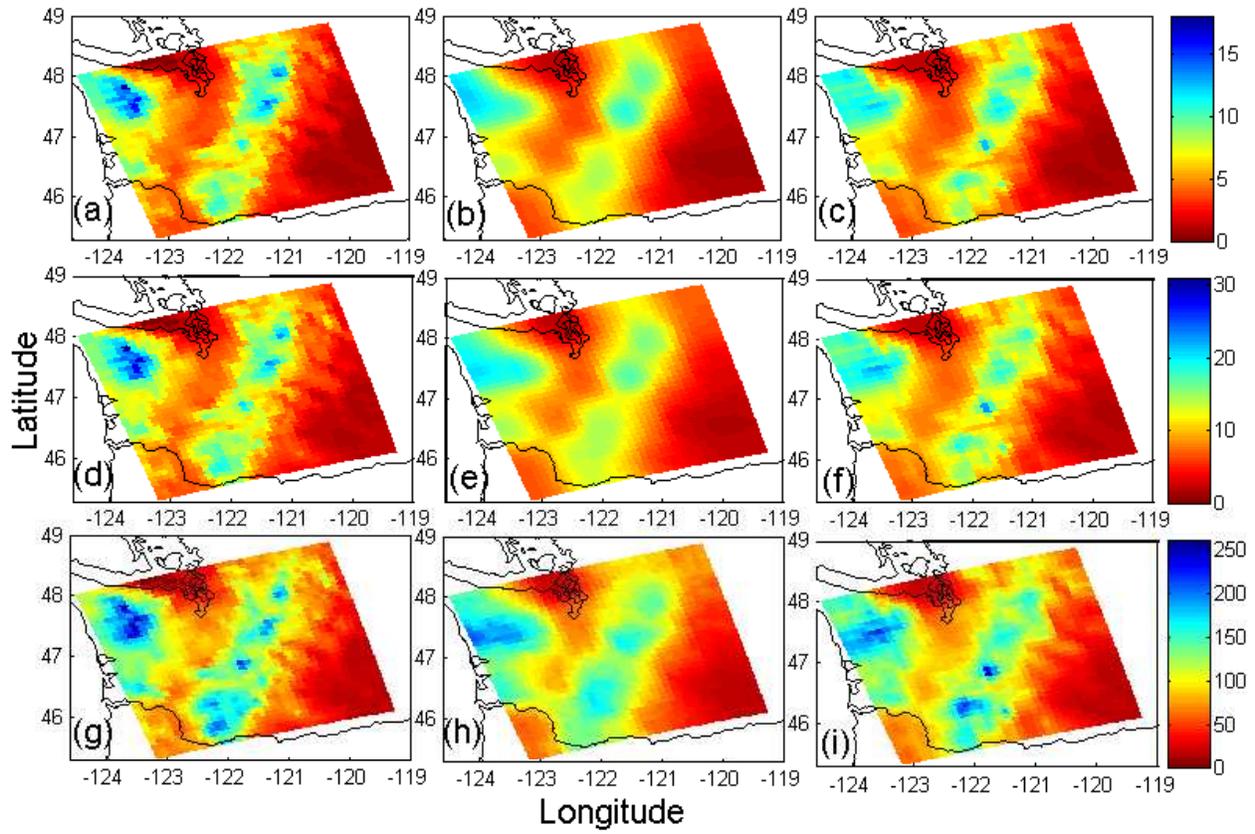


Figure E.1 – Précipitation moyenne journalière (mm ; haut), écart-type inter-journalier (mm ; centre) et précipitation journalière maximale (mm ; bas) calculés sur les 750 jours (2006-2008) observés (gauche) et désagrégés par les Modèles 1 (milieu) et 2 (droite) à 8,8 km. Pour les champs désagrégés (colonnes du milieu et de droite), les valeurs pour chaque pixel sont les moyennes des huit séries.

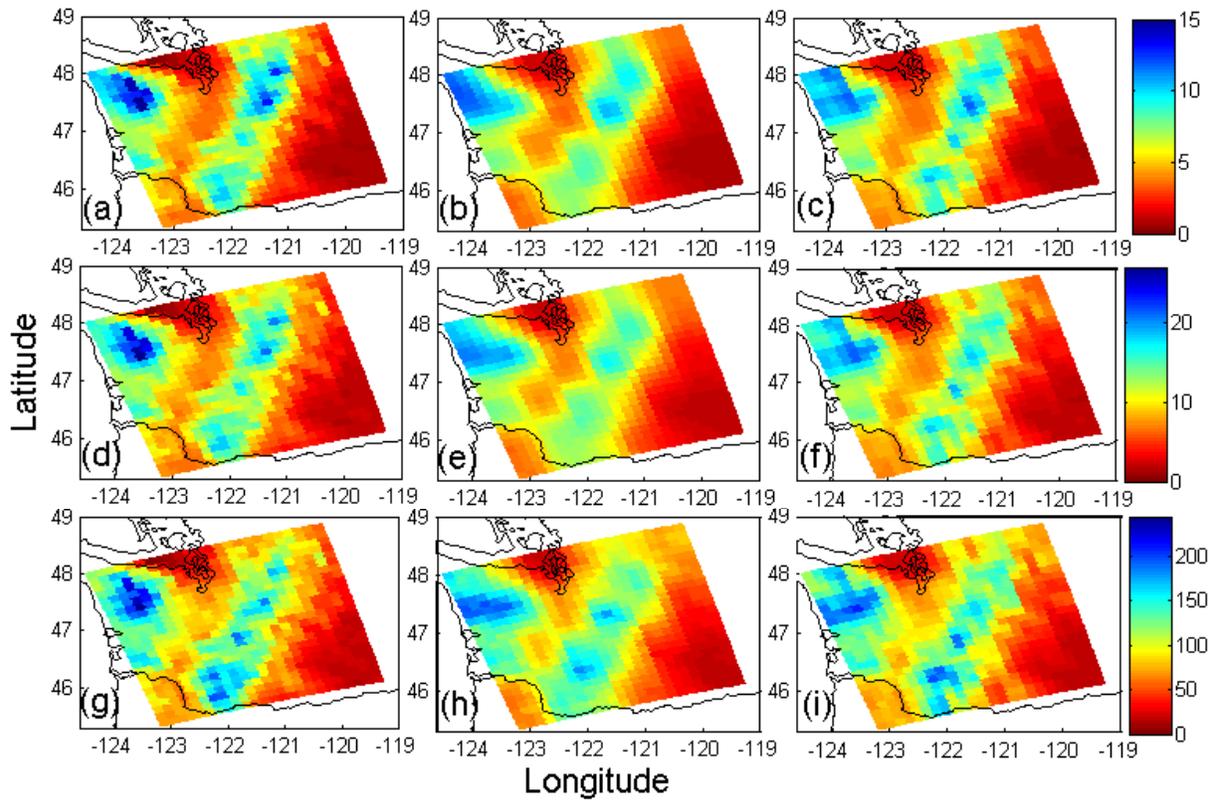


Figure E.2 – Précipitation moyenne journalière (mm ; haut), écart-type inter-journalier (mm ; centre) et précipitation journalière maximale (mm ; bas) calculés sur les 750 jours (2006-2008) observés (gauche) et désagrégés par les Modèles 1 (milieu) et 2 (droite) à 13,2 km. Pour les champs désagrégés (colonnes du milieu et de droite), les valeurs pour chaque pixel sont les moyennes des huit séries.

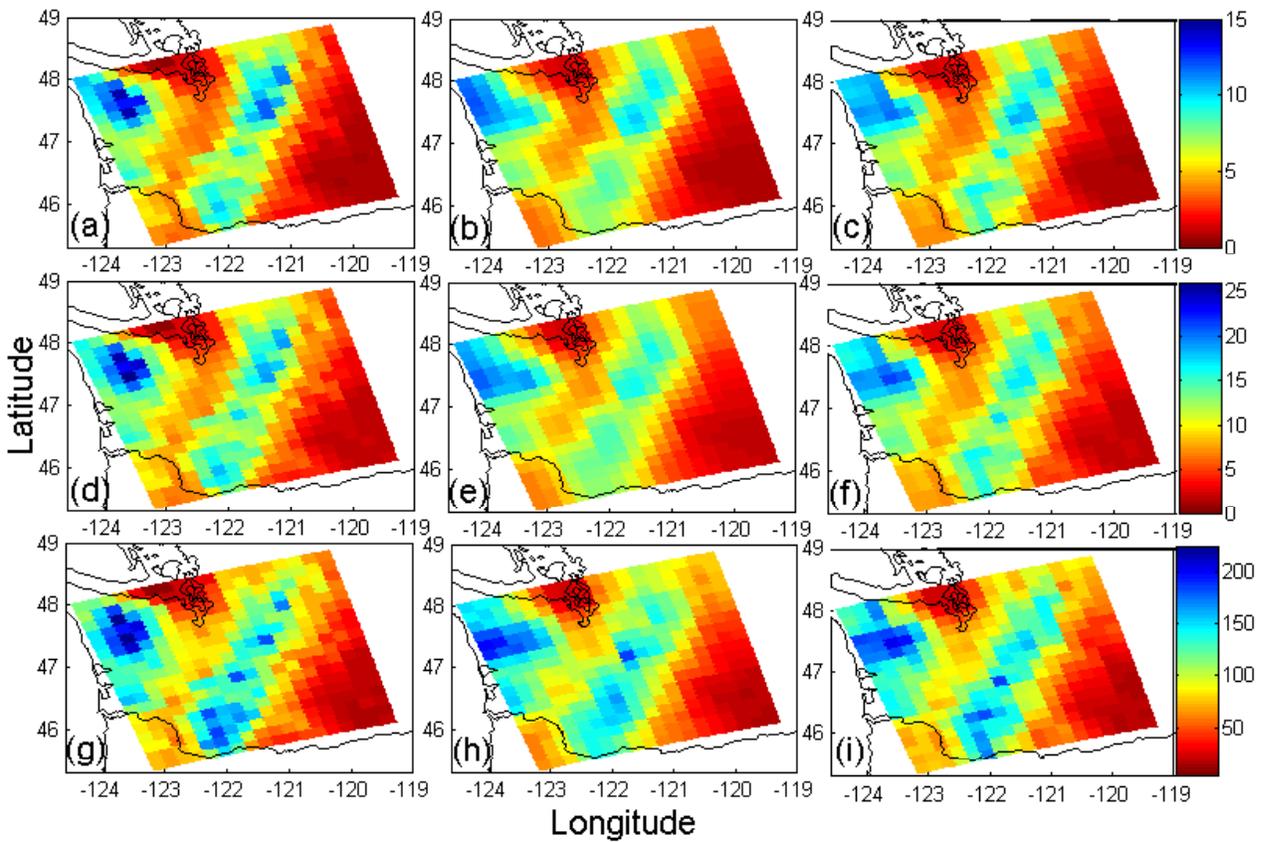


Figure E.3 – Précipitation moyenne journalière (mm ; haut), écart-type inter-journalier (mm ; centre) et précipitation journalière maximale (mm ; bas) calculés sur les 750 jours (2006-2008) observés (gauche) et désagrégés par les Modèles 1 (milieu) et 2 (droite) à 17,6 km. Pour les champs désagrégés (colonnes du milieu et de droite), les valeurs pour chaque pixel sont les moyennes des huit séries.

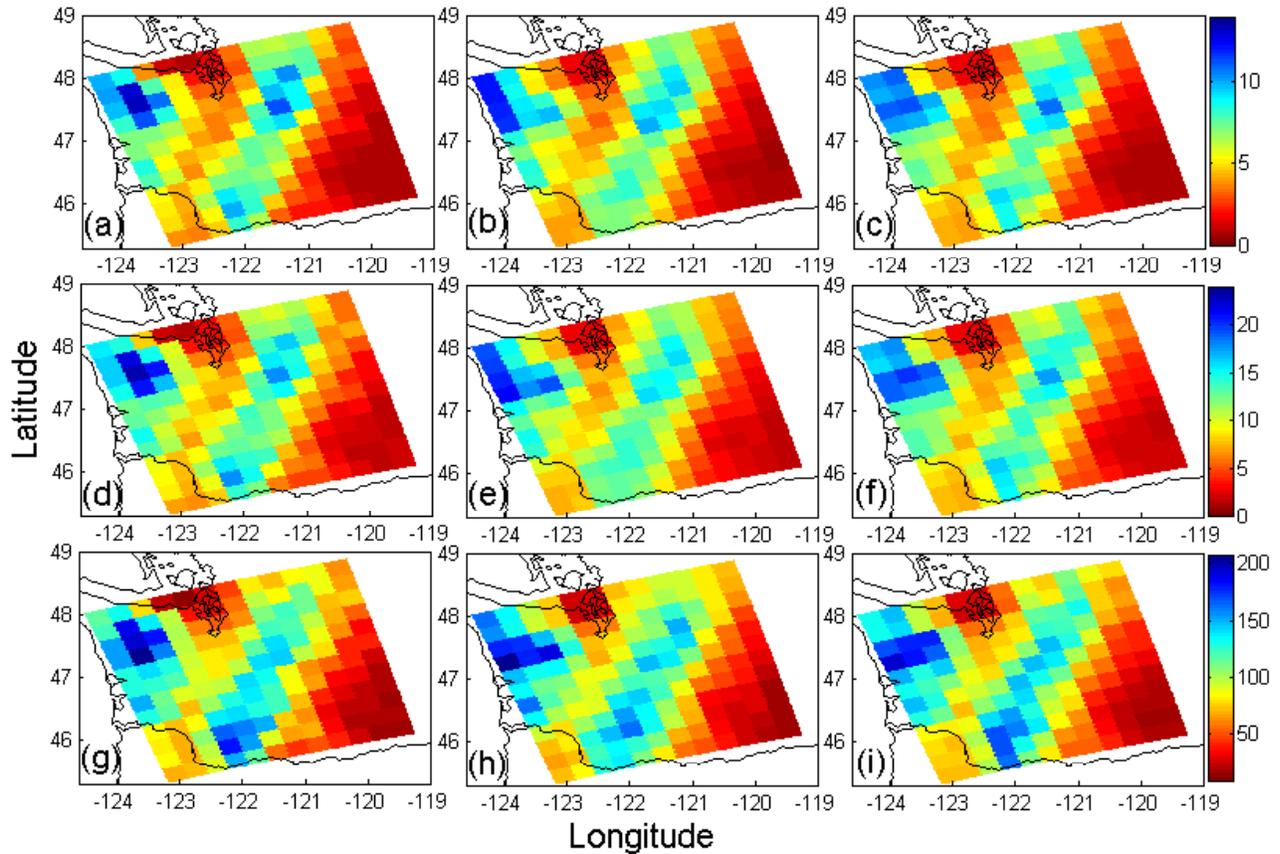


Figure E.4 – Précipitation moyenne journalière (mm ; haut), écart-type inter-journalier (mm ; centre) et précipitation journalière maximale (mm ; bas) calculés sur les 750 jours (2006-2008) observés (gauche) et désagrégés par les Modèles 1 (milieu) et 2 (droite) à 26,4 km. Pour les champs désagrégés (colonnes du milieu et de droite), les valeurs pour chaque pixel sont les moyennes des huit séries.

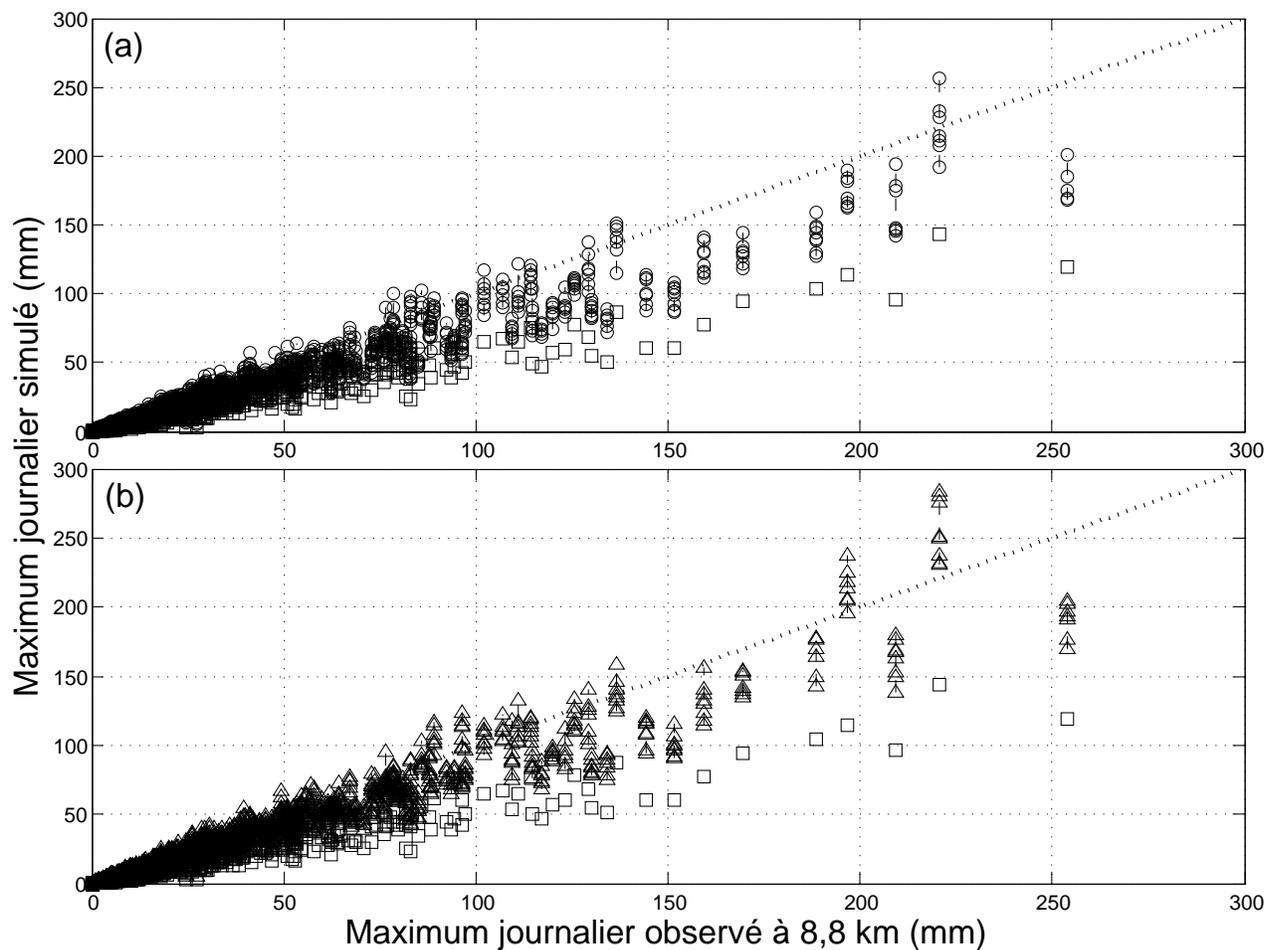


Figure E.5 – Précipitations maximales observée et produites par les huit séries désagrégées des Modèles (a) 1 et (b) 2 sur les pixels de 8,8 km des  $4 \times 4$  tuiles centrales pour chaque jour. Les carrés sont les valeurs maximales sur les tuiles de 52,8 km utilisées en entrée dans le modèle.

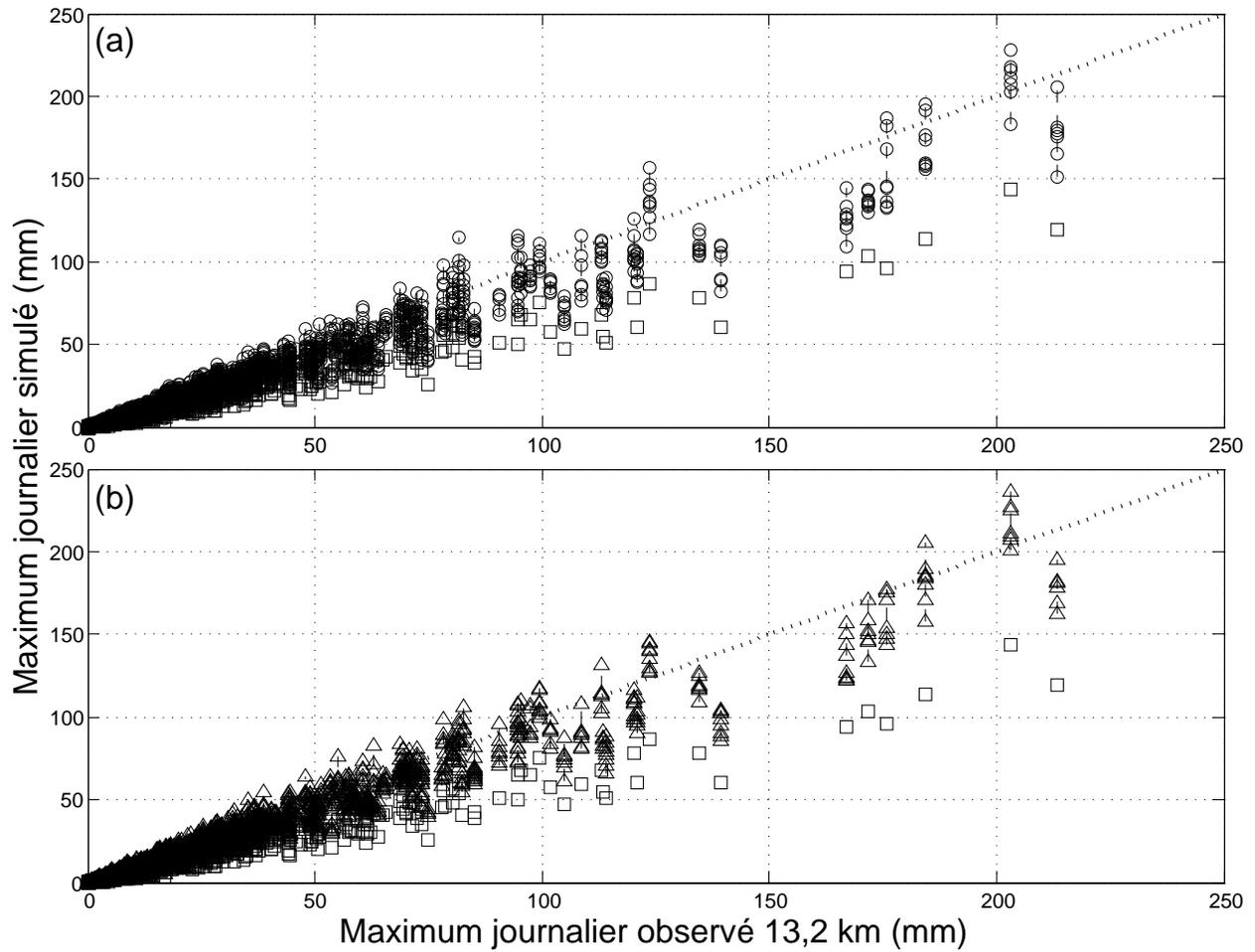


Figure E.6 – Précipitations maximales observée et produites par les huit séries désagrégées des Modèles (a) 1 et (b) 2 sur les pixels de 13,2 km des  $4 \times 4$  tuiles centrales pour chaque jour. Les carrés sont les valeurs maximales sur les tuiles de 52,8 km utilisées en entrée dans le modèle.

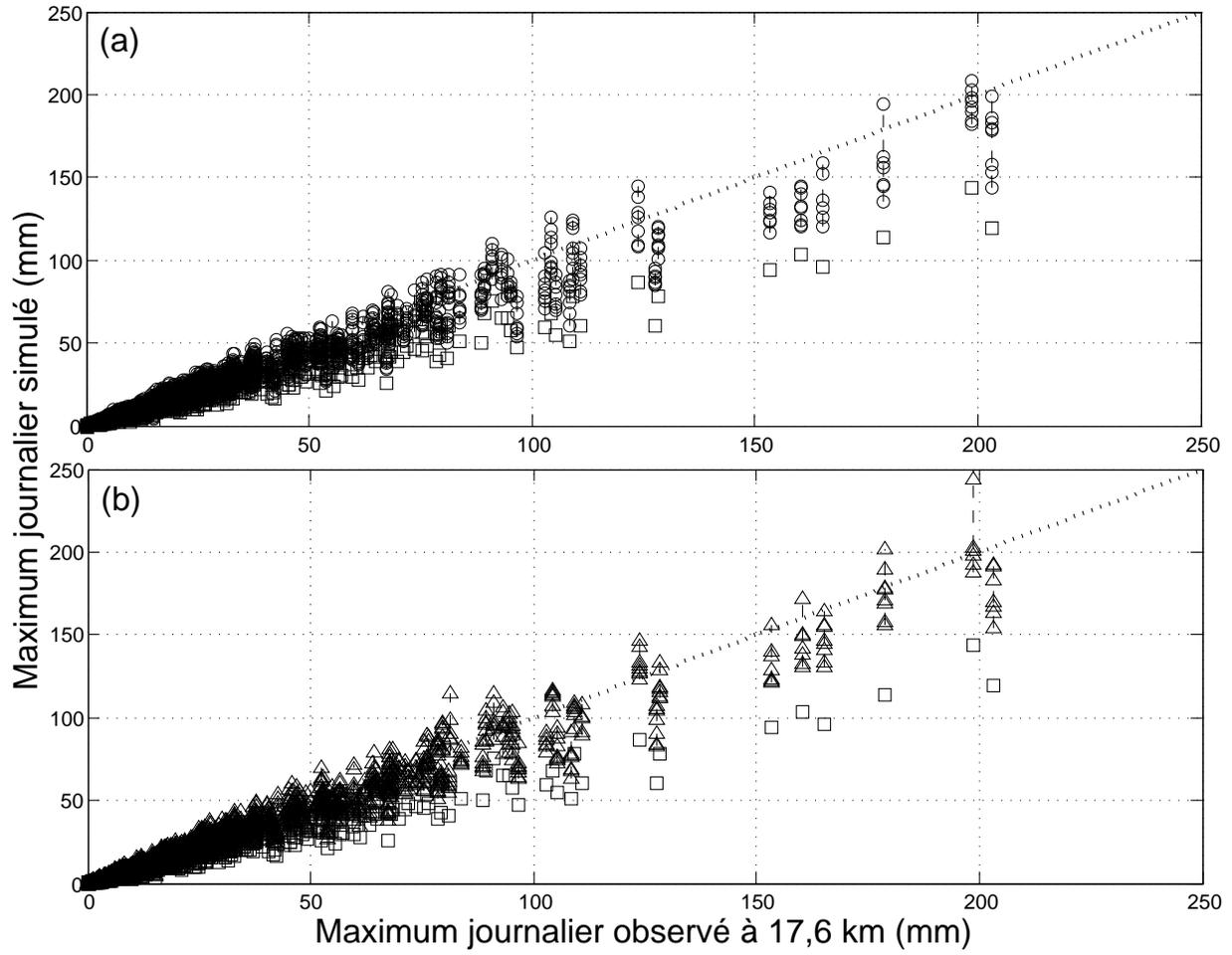


Figure E.7 – Précipitations maximales observée et produites par les huit séries désagrégées des Modèles (a) 1 et (b) 2 sur les pixels de 17,6 km des  $4 \times 4$  tuiles centrales pour chaque jour. Les carrés sont les valeurs maximales sur les tuiles de 52,8 km utilisées en entrée dans le modèle.

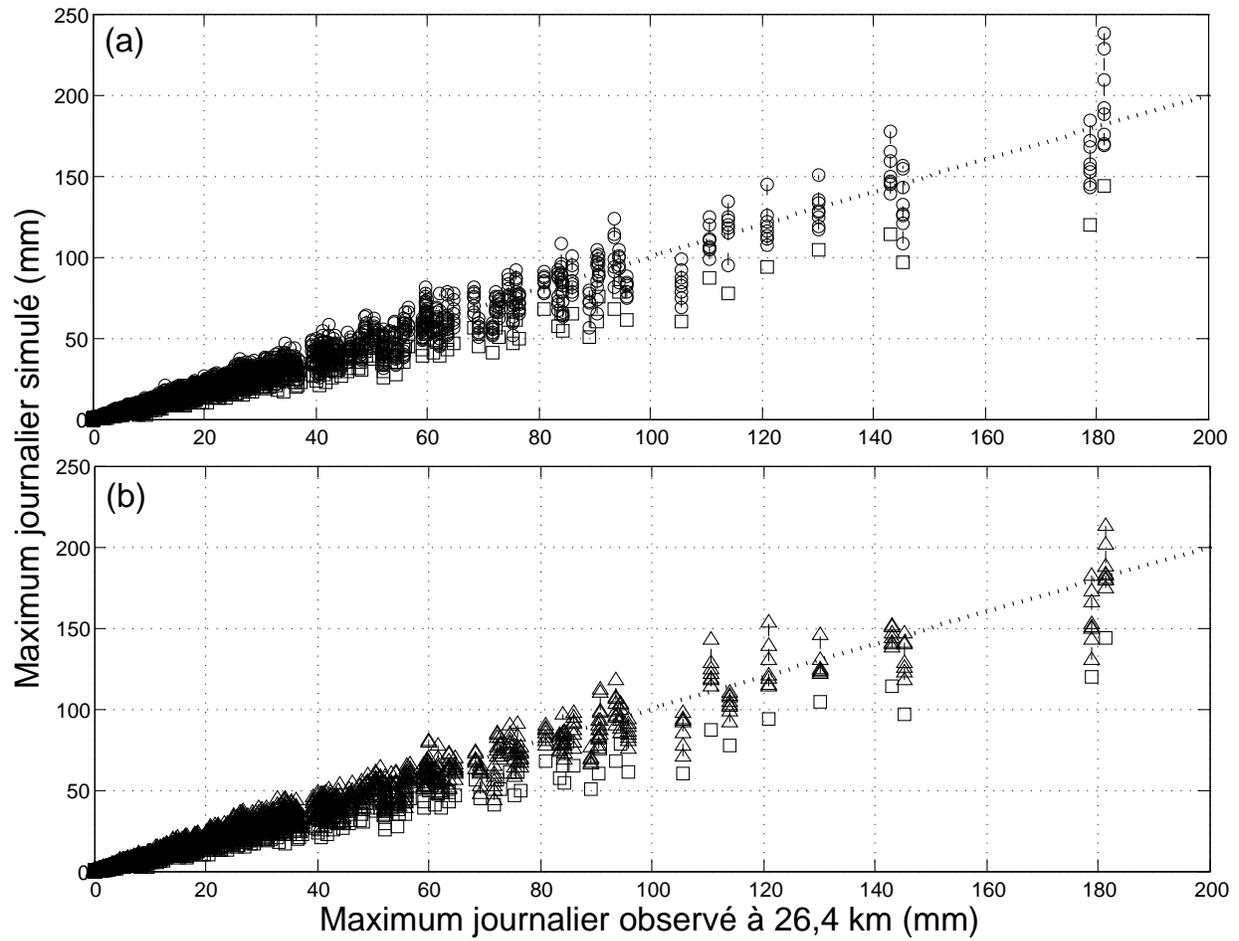


Figure E.8 – Précipitations maximales observée et produites par les huit séries désagrégées des Modèles (a) 1 et (b) 2 sur les pixels de 26,4 km des  $4 \times 4$  tuiles centrales pour chaque jour. Les carrés sont les valeurs maximales sur les tuiles de 52,8 km utilisées en entrée dans le modèle.



## Annexe F

# Écoulement désagrégé sur les tronçons de tête

Cette annexe a pour but d'illustrer la relation entre l'écoulement simulé et la précipitation du 11 juin 2002 en fonction des propriétés de la région drainée, soit la teneur en eau du sol au début de l'événement (le 11 juin à 0 h), le type de sol et l'occupation du sol. L'analyse est faite sur les 81 tronçons pour lesquelles un seul type de sol est considéré par HYDROTEL ; il s'agit essentiellement des tronçons de tête. La Figure F.1 présente 1000 valeurs d'écoulements simulés du 11 juin 2002 à partir des paramètres estimés à 4,4 km (Tableau 9.5) en fonction des précipitations désagrégées à 4,4 km pour quatre de ces tronçons.

La relation est très directe entre la précipitation et l'écoulement pour les tronçons de tête. Pour certains tronçons (*e.g.* tronçons 2 et 43, Figures F.1(a) et F.1(c)), la relation est linéaire sur l'ensemble des valeurs désagrégées. Pour d'autres (*e.g.* tronçons 15 et 98, Figures F.1(b) et F.1(d)), la pente est faible pour les plus petites valeurs de précipitation, mais à partir d'une certaine valeur seuil de précipitation, la pente augmente drastiquement. Selon toute vraisemblance, le seuil est la précipitation nécessaire pour que le sol soit saturé. Pour les tronçons où le seuil n'est pas présent (*e.g.* Figures F.1(a) et F.1(c)), le seuil est en fait plus petit que la plus petite valeur de précipitation désagrégée. L'écoulement est exprimé de la façon suivante :

$$Q = \begin{cases} b_0 + b_1P & \text{si } P \leq S \\ b_0 + b_1S + b_2(P - S) & \text{si } P > S \end{cases} \quad (\text{F.1})$$

Tableau F.1 – Conductivité hydraulique à saturation pour les types de sols présents sur le bassin.

Type de sol	Conductivité hydraulique à saturation (m/h)
Sable	$2,10 \times 10^{-1}$
Sable loameux	$6,11 \times 10^{-2}$
Loam sableux	$2,59 \times 10^{-2}$
Loam	$1,32 \times 10^{-2}$
Loam-sableux argileux	$4,30 \times 10^{-3}$
Argile loameuse	$9,00 \times 10^{-4}$
Argile	$6,00 \times 10^{-4}$

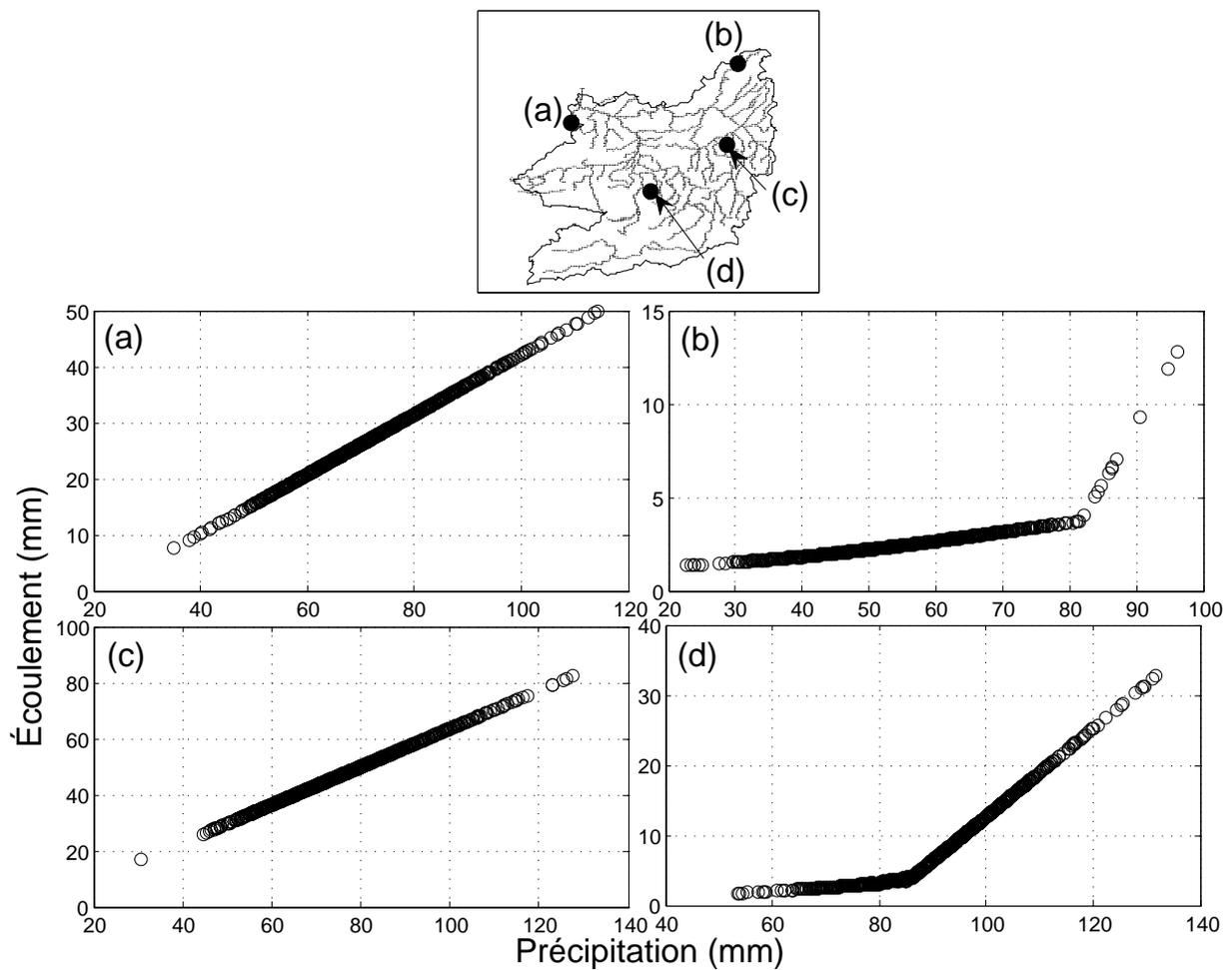


Figure F.1 – Écoulement journalier simulé en fonction de la pluie du 11 juin 2002 pour les 1000 séries désagrégées à 4,4 km sur les tronçons (a) 2, (b) 15, (c) 43 et (d) 98.

où  $Q$ ,  $P$  et  $S$  sont respectivement l'écoulement, la précipitation et le seuil de précipitation pour le 11 juin 2002 sur un sous-bassin donné. Il serait intéressant de trouver quelles variables peuvent expliquer  $S$  et les paramètres à estimer  $b_0$ ,  $b_1$  et  $b_2$ .

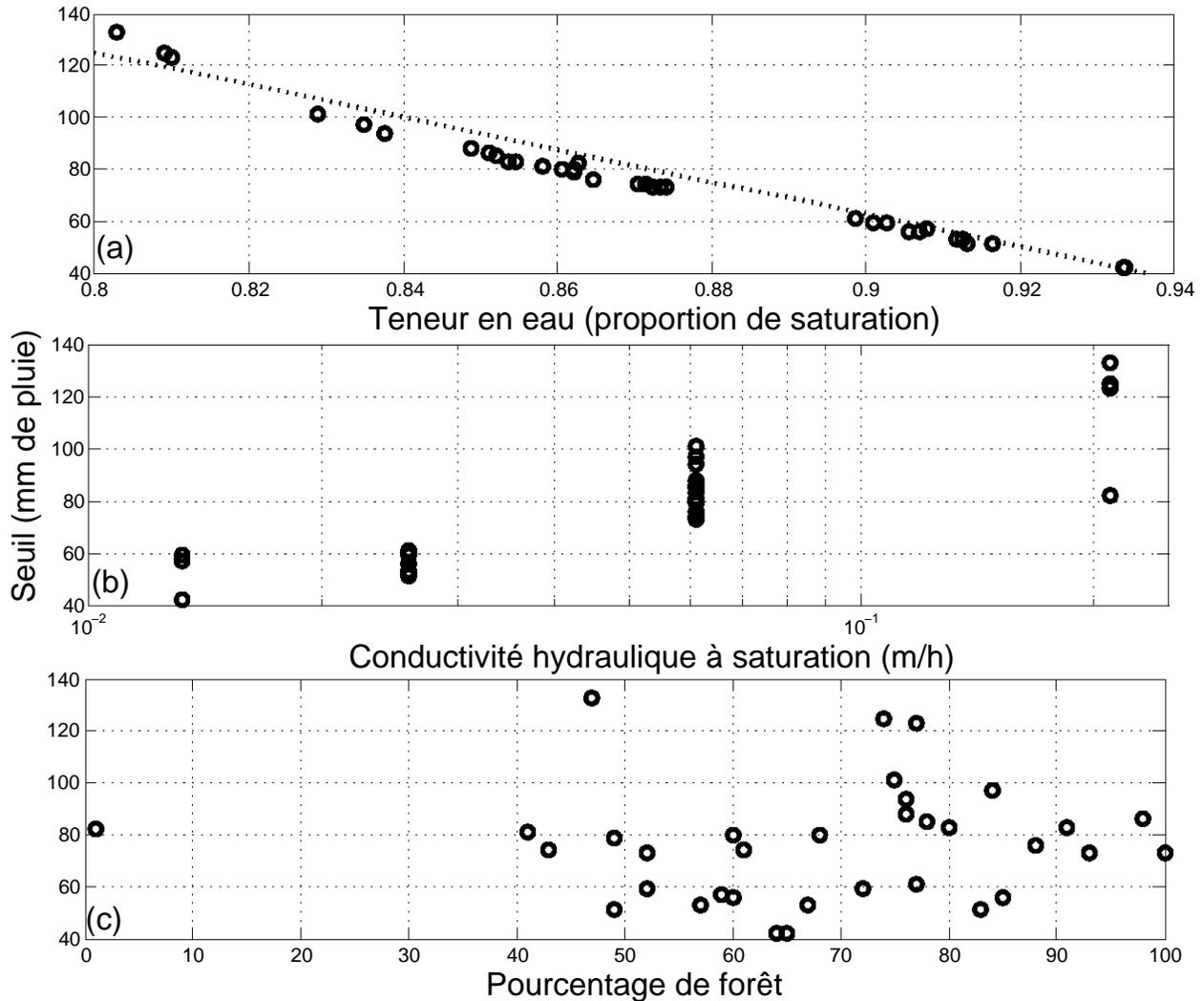


Figure F.2 – Seuil de précipitation en fonction (a) de la teneur en eau du sol au début de l'événement, (b) du type de sol et (c) de l'occupation du sol sur le sous-bassin. Le type de sol est représenté par sa conductivité hydraulique à saturation et l'occupation du sol par la pourcentage de forêt sur le sous-bassin.

Le seuil est visible pour 34 des tronçons étudiés. Trois facteurs peuvent influencer sur la valeur du seuil : les conditions initiales du bassin, le type de sol et l'occupation du sol. La Figure F.2 exprime le seuil en fonction de ces trois facteurs. Les conditions initiales sont données par la teneur en eau du sol, soit le pourcentage de saturation dans le sol, au début de l'événement. Le type de sol peut être représenté numériquement par sa conductivité hydraulique à

saturation. Les plus petites valeurs sont pour les sols argileux et les plus grandes pour les sols sableux. Le tableau F.1 présente la conductivité hydraulique à saturation pour les types de sols présents sur le bassin. Dans son sous-modèle d'écoulement sur la partie terrestre du bassin, HYDROTEL sépare les différentes occupations du sol en deux catégories : les milieux forestiers (conifères, feuillus et friches) et les autres. L'occupation du sol est donc représenté par le pourcentage du sous-bassin qui est forestier. Le pourcentage de zones urbaines aurait aussi pu être considéré puisqu'en milieu urbain, le sol est imperméable dans HYDROTEL. Cependant, les zones urbaines sont négligeables sur le bassin.

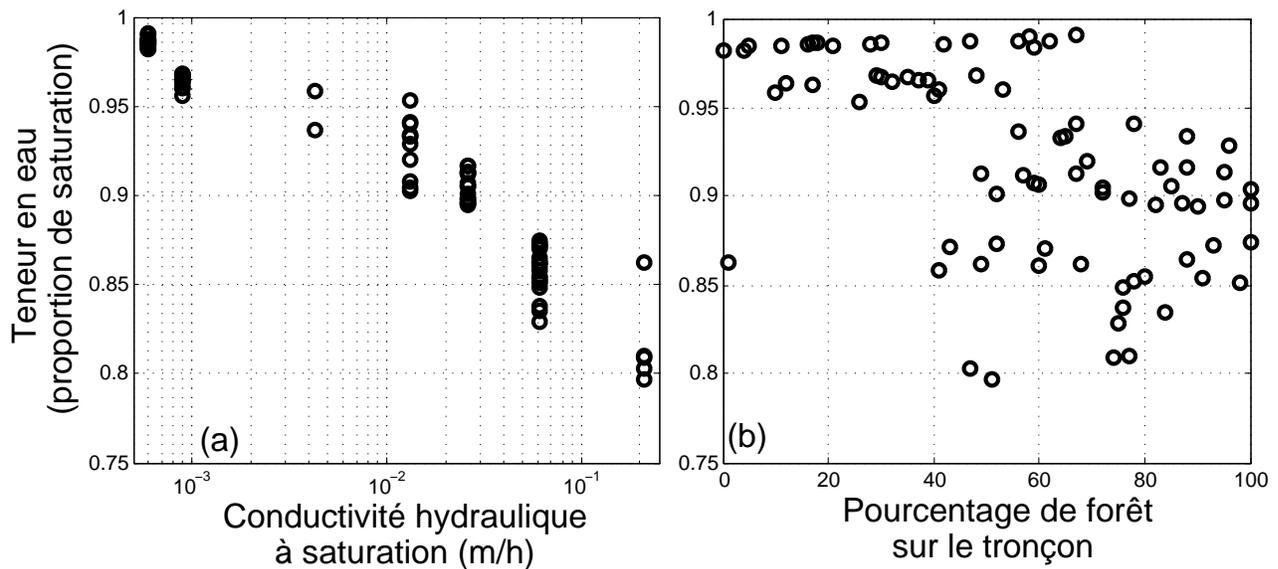


Figure F.3 – Teneur en eau du sol au début de l'événement en fonction (a) du type de sol et (b) de l'occupation du sol sur le sous-bassin. Le type de sol est représenté par sa conductivité hydraulique à saturation et l'occupation du sol par la pourcentage de forêt sur le sous-bassin.

La relation est très directe entre le seuil et la teneur en eau au début de l'événement (Figure F.2(a)). Tel qu'attendu, plus le sol est près de la saturation, plus le seuil est petit. Il y a aussi un lien fort entre le seuil et le type de sol (Figure F.2(b)). Plus la conductivité hydraulique est grande (*i.e.* plus le sol est sableux), plus le seuil est élevé. En fait, il y a un lien fort entre la teneur en eau au début de l'événement et le type de sol (Figure F.3(a)). Plus le sol est sableux (*i.e.* plus la conductivité hydraulique du sol est grande), moins le sol est près de la saturation. Cela peut s'expliquer par le fait que dans HYDROTEL, plus la conductivité hydraulique d'un type de sol est grande, plus l'écoulement retardé, soit l'écoulement latéral de la deuxième couche de sol, se produira rapidement, ce qui a pour effet d'assécher plus vite le sol. C'est pourquoi qu'au début de l'événement, les sols argileux sont souvent plus près de la saturation que les sols sableux. Ainsi, pour une hauteur de précipitation donnée,

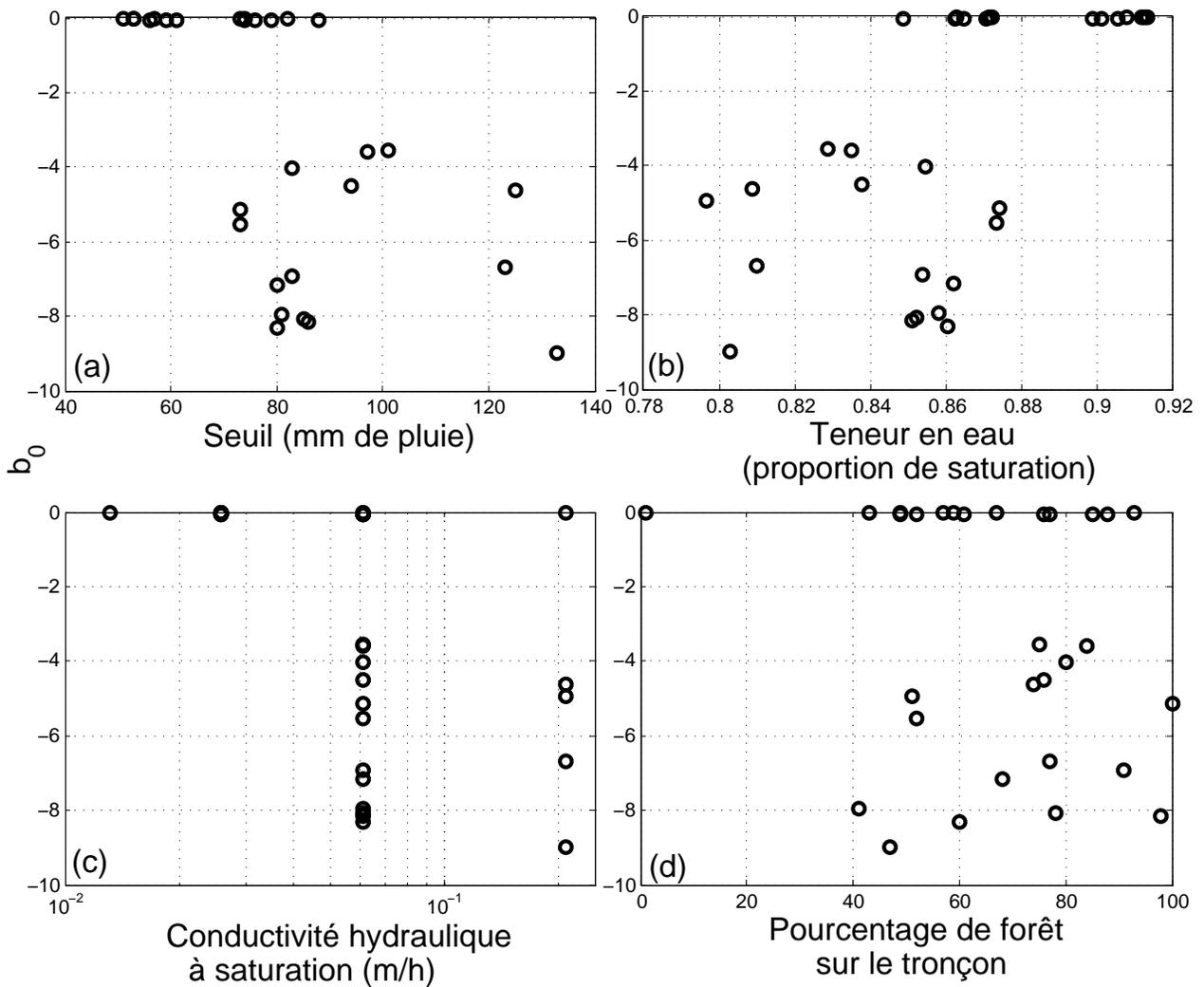


Figure F.4 – Valeur de  $b_0$  de l'équation (F.1) en fonction (a) du seuil de précipitation, (b) de la teneur en eau du sol au début de l'événement, (c) du type de sol et (d) de l'occupation du sol sur le sous-bassin. Le type de sol est représenté par sa conductivité hydraulique à saturation et l'occupation du sol par la pourcentage de forêt sur le sous-bassin.

un UHRH argileux devrait produire plus de ruissellement de surface qu'un UHRH sableux (puisqu'il est plus près de la saturation), mais moins d'écoulement retardé (puisque sa conductivité hydraulique est plus faible). Pour un événement intense, l'écoulement retardé est négligeable par rapport au ruissellement de surface, ce qui explique pourquoi les sols argileux répondent beaucoup plus intensément que les sols sableux.

Il n'y a pas de lien fort entre le seuil et l'occupation du sol (Figure F.2(c)). Il semble y avoir une relation négative entre la teneur en eau et le pourcentage de forêt (Figure F.3(b)), mais cette relation est très faible par rapport à la relation entre la teneur en eau et la conductivité hydraulique à saturation (Figure F.3(a)). Cela corrobore les résultats de Lavigne *et al.*

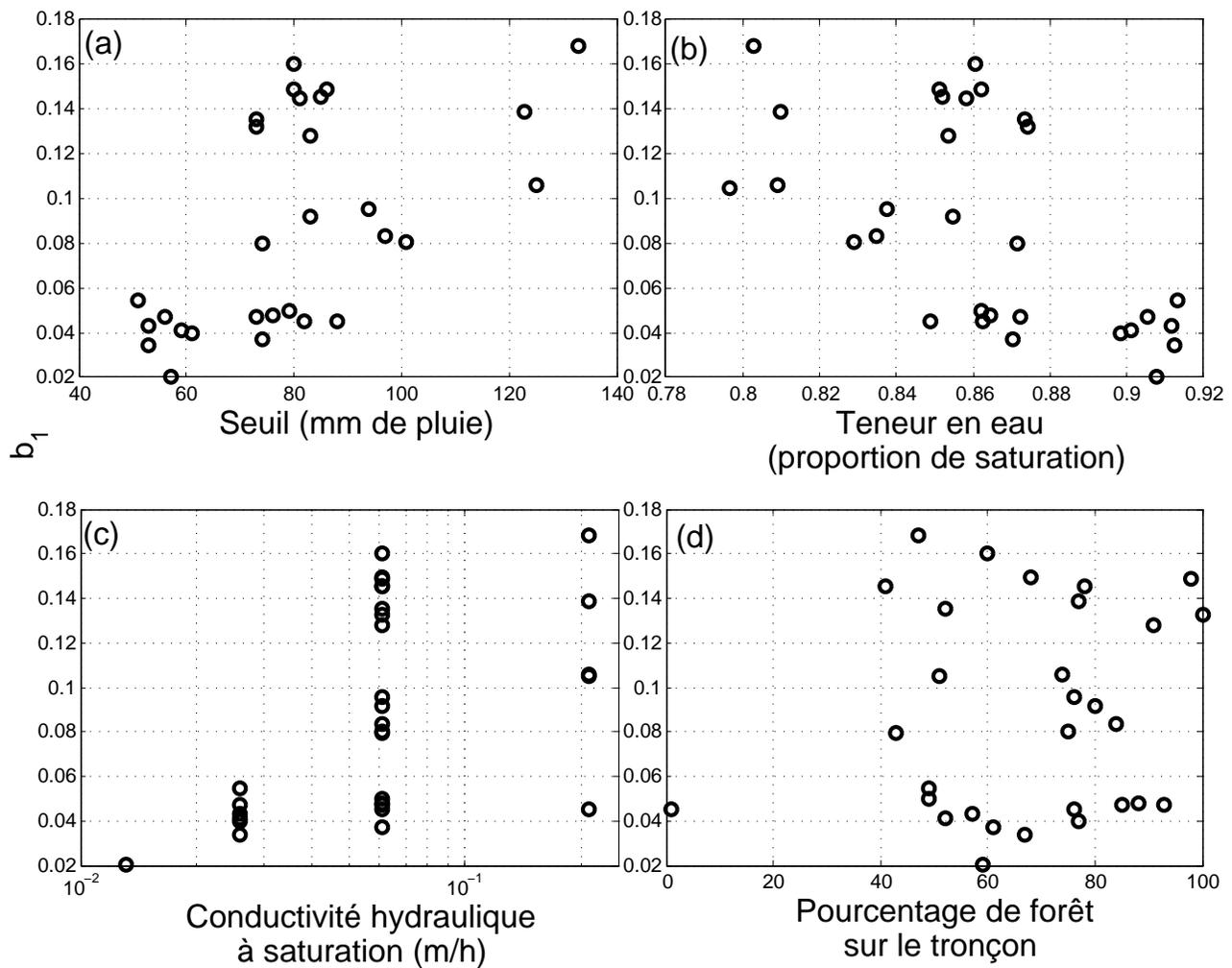


Figure F.5 – Valeur de  $b_1$  de l'équation (F.1) en fonction (a) du seuil de précipitation, (b) de la teneur en eau du sol au début de l'événement, (c) du type de sol et (d) de l'occupation du sol sur le sous-bassin. Le type de sol est représenté par sa conductivité hydraulique à saturation et l'occupation du sol par la pourcentage de forêt sur le sous-bassin.

(2004) qui ont obtenu que l'occupation du sol avait un impact négligeable sur un événement intense.

Il a été décidé d'exprimer le seuil  $S$  par rapport à la teneur en eau  $E$  précédent l'événement. Une simple régression linéaire a montré que pour cet événement et pour les valeurs de paramètres utilisés,  $S$  peut s'exprimer par :

$$S = 596,9 \text{ mm} \times (1 - E) \tag{F.2}$$

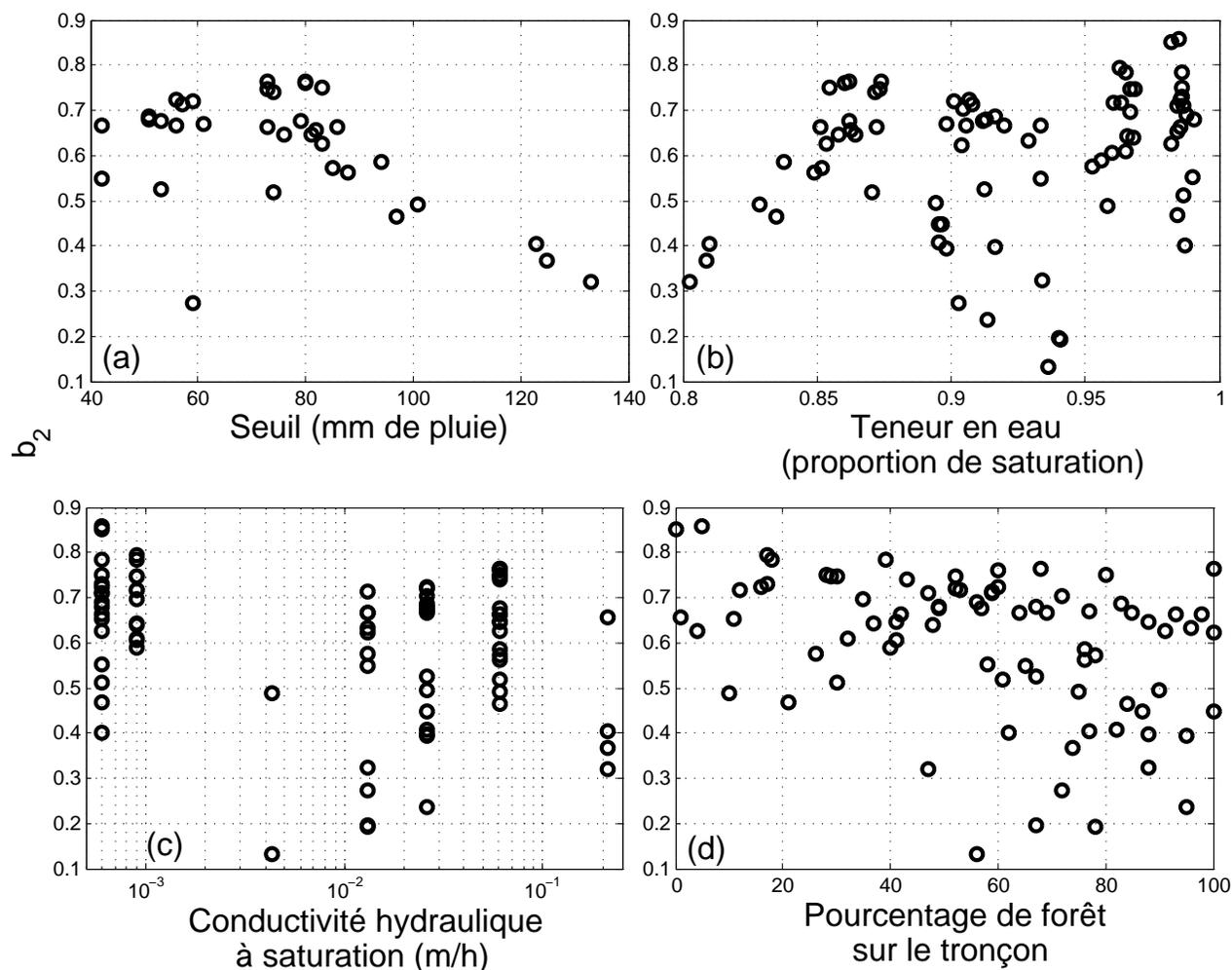


Figure F.6 – Valeur de  $b_2$  de l'équation (F.1) en fonction (a) du seuil de précipitation, (b) de la teneur en eau du sol au début de l'événement, (c) du type de sol et (d) de l'occupation du sol sur le sous-bassin. Le type de sol est représenté par sa conductivité hydraulique à saturation et l'occupation du sol par la pourcentage de forêt sur le sous-bassin.

où  $S$  est en millimètre.

Puisque la corrélation est forte entre la conductivité hydraulique à saturation du sol et la teneur en eau au début de l'événement, le seuil aurait aussi pu s'exprimer en fonction du type de sol. Une fois  $S$  estimé pour chaque tronçon, les valeurs choisies des paramètres  $b_0$ ,  $b_1$  et  $b_2$  de l'équation (F.1) sont celles qui minimisent la somme du carré des erreurs pour chacun des 81 tronçons étudiés. Les Figures F.4, F.5 et F.6 expriment les valeurs estimées des paramètres en fonction des variables explicatives. Pour les Figures F.4 et F.5, seuls les 34 tronçons où le seuil est visible sont présentés. Dans l'ensemble, il y a une faible dépendance entre les valeurs estimées des paramètres et les variables. Par

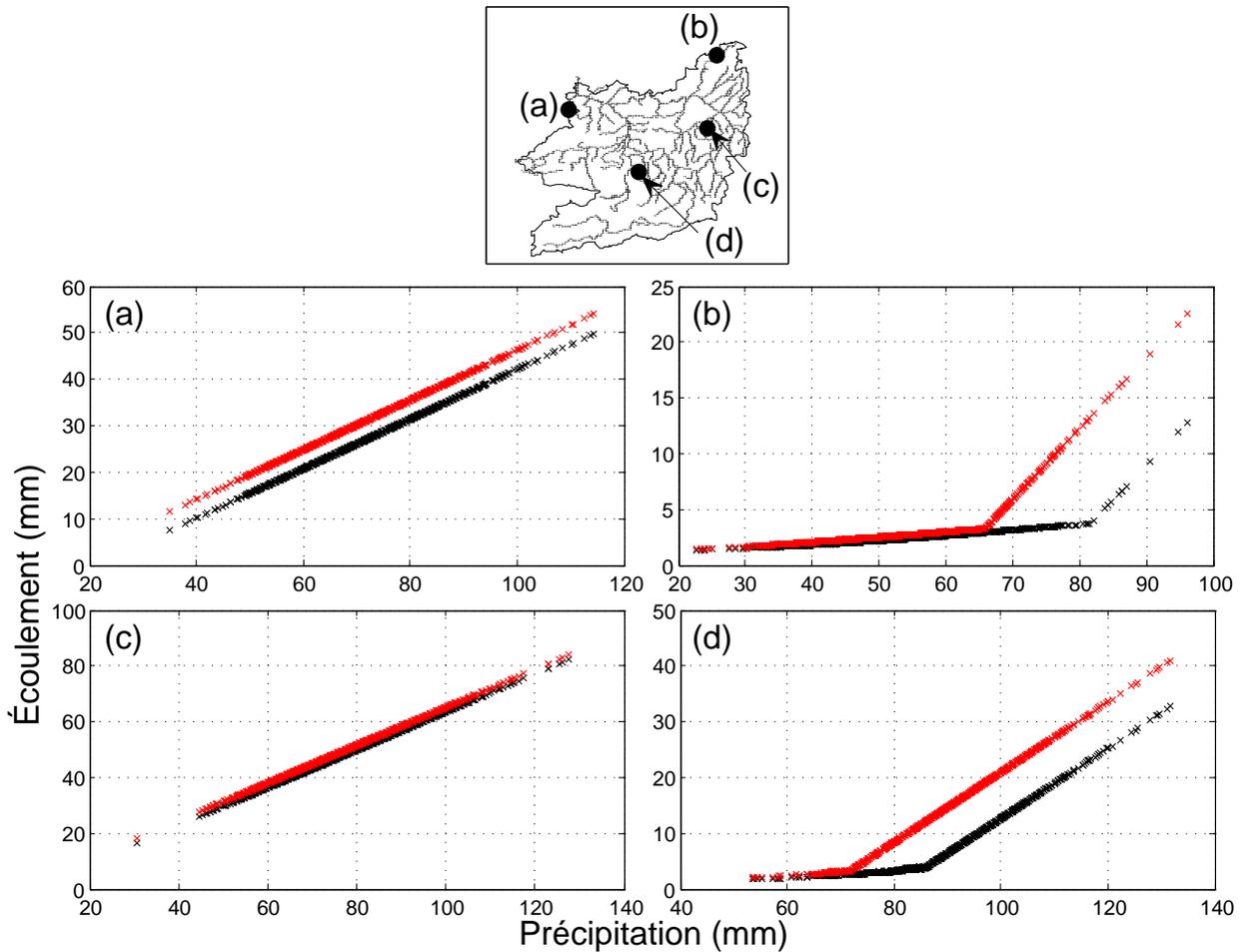


Figure F.7 – Écoulement journalier simulé à partir des calages à 4,4 (noir) et à 26,4 km (rouge) en fonction de la pluie du 11 juin 2002 pour les 1000 séries désagrégées à 4,4 km sur les tronçons (a) 2, (b) 15, (c) 43 et (d) 98.

contre,  $b_1$  (respectivement  $b_0$ ) est corrélé positivement (négativement) avec le seuil et négativement (positivement) avec la teneur en eau au début de l'événement. Cela est causé par la corrélation négative entre  $b_0$  et  $b_1$ . Pour  $b_2$  (Figure F.6), il y a de faibles tendances, mais pas de relation forte avec les variables explicatives. Cela suggère que les valeurs de paramètres de l'équation (F.1) ne sont pas ou peu dépendantes des propriétés du sous-bassin.

En somme, la relation entre l'écoulement et la précipitation varie uniquement en fonction du seuil, qui peut être exprimé en fonction de la teneur en eau au début de l'événement ou du type de sol. Les mêmes analyses ont été faites à partir de la pluie désagrégée à 4,4 km, mais à partir des paramètres d'HYDROTEL estimés à 26,4 km ( $Z_3$  moins profond ; Tableau 9.5). Les résultats sont identiques, sauf que les seuils sont plus petits, puisque les sols sont plus

près de la saturation lorsqu'ils sont moins profonds. La Figure F.7 illustre la différence entre les deux calages de la relation entre l'écoulement et la pluie du 11 juin 2002 pour quatre sous-bassins.