



HAL
open science

Cartographie des événements hydrologiques extrêmes et estimation SCHADEX en sites non jaugés

David Penot

► **To cite this version:**

David Penot. Cartographie des événements hydrologiques extrêmes et estimation SCHADEX en sites non jaugés. Hydrologie. Université de Grenoble, 2014. Français. NNT : 2014GRENU022 . tel-01233267

HAL Id: tel-01233267

<https://theses.hal.science/tel-01233267v1>

Submitted on 24 Nov 2015

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

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

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Océan, Atmosphère et Hydrologie**

Arrêté ministériel : 7 août 2006

Présentée par

David PENOT

Thèse dirigée par **Michel LANG**

et codirigée par **Emmanuel PAQUET**

préparée au sein du **CENTRE DE LYON-VILLEURBANNE D'IRSTEA**

en collaboration avec **EDF-DTG**

dans l'**École Doctorale Terre, Univers, Environnement**

CARTOGRAPHIE DES ÉVÉNEMENTS HYDROLOGIQUES EXTRÊMES ET ESTIMATION SCHADEX EN SITES NON JAUGÉS

Thèse soutenue publiquement le **17 octobre 2014**,
devant le jury composé de :

Mme. Anne-Catherine FAVRE

Université de Grenoble, Présidente

M. Pierre RIBSTEIN

Université Pierre et Marie Curie Paris VI, Rapporteur

M. András BÁRDOSSY

Institut für Wasserbau (Universität Stuttgart), Rapporteur

M. Roger MOUSSA

INRA Montpellier, Examineur

M. Michel LANG

Irstea Lyon, Directeur de thèse

M. Emmanuel PAQUET

EDF-DTG Grenoble, Co-Directeur de thèse



THÈSE

Pour obtenir le grade de
DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Océan, Atmosphère et Hydrologie**

Arrêté ministériel : 7 août 2006

Présentée par

David PENOT

Thèse dirigée par **Michel LANG**
et codirigée par **Emmanuel PAQUET**

préparée au sein du **centre de Lyon-Villeurbanne d'IRSTEA**

en collaboration avec **EDF - DTG**

dans l'**École Doctorale Terre Univers Environnement**

CARTOGRAPHIE DES ÉVÉNEMENTS HYDROLOGIQUES EXTRÊMES ET ESTIMATION SCHADEX EN SITES NON JAUGÉS

Présentée et soutenue publiquement le **17 octobre 2014**,
devant le jury composé de :

Mme. Anne-Catherine FAVRE
M. Pierre RIBSTEIN
M. András BÁRDOSSY
M. Roger MOUSSA
M. Michel LANG
M. Emmanuel PAQUET

Université de Grenoble
Université Pierre et Marie Curie Paris VI
Institut für Wasserbau
INRIA Montpellier
IRSTEA Lyon
EDF-DTG Grenoble

Présidente
Rapporteur
Rapporteur
Examineur
Directeur de thèse
Co-Directeur de thèse

Résumé

Depuis 2006, à EDF, les études de crues extrêmes sont réalisées avec la méthode SCHADEX (Simulation Climato-Hydrologique pour l'Appréciation des Débits EXtrêmes). Elle s'appuie sur un modèle probabiliste MEWP (distribution saisonnière utilisant une classification par type de temps) pour décrire l'aléa pluie et sur une simulation stochastique croisant l'aléa pluie et l'aléa de saturation du bassin.

Les approches par simulation, type SCHADEX, ont montré de bonnes performances pour estimer les distributions de crues extrêmes (projet ANR ExtraFlo, 2013). Cependant, l'utilisation de SCHADEX en l'absence de données (pluie, température, débit) sur le bassin à étudier reste problématique. Cette thèse propose une adaptation de la méthode en site non jaugé en essayant de conserver ses points forts, à savoir :

- une structuration spatiale et probabiliste des précipitations conditionnée par les types de temps.
- un croisement des aléas pluie et saturation du bassin par simulation stochastique.

Ce travail s'est limité au pas de temps journalier afin d'aborder la problématique de régionalisation avec un maximum de données. La démarche s'est alors articulée autour de quatre grands axes :

- proposer une méthode de régionalisation des précipitations journalières extrêmes ponctuelles et construire des cartes de pluies aux temps de retour remarquables. Évaluer l'intérêt d'une classification par type de temps pour la régionalisation des distributions de pluies extrêmes et qualifier l'interpolateur de pluie SPAZM pour l'estimation des pluies extrêmes.
- s'intéresser à la construction de pluies de bassin (ou pluies spatiales) et en particulier à l'impact des choix de construction de cette pluie sur l'estimation des précipitations extrêmes concernant le bassin.
- développer une méthode de simulation stochastique régionale permettant de proposer une distribution de débits journaliers issue d'un croisement des aléas pluies et saturation du bassin.
- étudier le passage de la distribution des débits journaliers à la distribution des débits de pointe.

Les principaux apports de cette thèse sont les suivants :

- la prise en compte des types de temps permet d'améliorer la description des structures spatiales des précipitations extrêmes.
- l'information apportée par les pluies SPAZM se révèle être précieuse pour l'estimation des pluies extrêmes en site non jaugé.
- une étude de sensibilité du calcul de la pluie spatiale en fonction du nombre de postes utilisés (comparaison des pluies SPAZM et Thiessen) donne une indication sur le biais d'estimation.
- le générateur de champs de pluie par bandes tournantes SAMPO permet d'étudier l'abattement sur les précipitations extrêmes et de mettre en place un modèle de correction pour

les quantiles élevés des pluies spatiales SPAZM.

- une nouvelle méthode de simulation stochastique peu paramétrée mais analogue à la méthode SCHADEX (croisement d'un aléa pluie et d'un aléa de saturation du bassin pour produire une distribution des débits journaliers) est proposée pour l'estimation en site non jaugé.
- enfin, un travail préliminaire donne des premiers éléments sur le passage à la distribution des débits de pointe par un générateur d'hydrogrammes s'adaptant à la séquence des débits journaliers simulés.

Tous ces développements et conclusions sont détaillés et justifiés dans le mémoire de thèse.

Mots clés : prédétermination des crues, méthode SCHADEX, simulation stochastique , précipitations extrêmes, contexte non jaugé, types de temps, pluie spatiale, interpolateur de pluie SPAZM, débit de pointe.

Abstract

Since 2006, at EDF, extreme flood estimations are computed with the SCHADEX method (Climatic-hydrological simulation of extreme floods). This method relies on a MEWP probabilistic model (seasonal rainfall distribution using a weather pattern concept) and on a stochastic simulation to cross rainy events hazard and catchment saturation states.

Simulation approaches, as SCHADEX, have shown good performances to estimate extreme flood distributions. However, the use of SCHADEX method without data for a considered catchment (rain, temperature, runoff) remains a main issue. This thesis suggests an adaptation of the method in ungauged context, trying to keep the key points of the SCHADEX method :

- spatial and probabilistic structure of rainfall conditioned by weather patterns.
- a cross of rainfall and catchment saturation hazards by stochastic simulation.

This work is limited to a daily step to address the issue of regionalization with a maximum of data. The approach is then structured around four main points :

- regionalize punctual daily extreme precipitations and construct maps of return period rainfalls. Evaluate the contribution of a weather type classification for the regionalization of extreme rainfall distributions and qualify the SPAZM interpolator for the estimation of extreme rainfall.
- wonder about the construction of an areal rainfall and in particular about the impact of its construction choices on the estimation of extreme precipitations.
- develop a regional stochastic simulation method to estimate a distribution of daily runoffs which crosses rainy events and catchment saturation hazards.
- study the transposition from a daily runoff distribution to a peak flow distribution.

The main contributions of this thesis are :

- taking into account the weather types improves the description of spatial patterns of extreme precipitations.
- information provided by the SPAZM rainfall interpolator proves to be valuable for the estimation of extreme rainfall in ungauged site.
- a sensitivity analysis of the calculation of the areal rainfall based on the number of stations used (comparison SPAZM and Thiessen areal rainfalls) gives an indication of the estimation bias.
- the SAMPO rainfall generator used to study the areal reduction factor of extreme precipitation and implement a correction model for high quantiles of SPAZM areal rainfall.
- a simplified method of stochastic simulation similar to SCHADEX method (cross between a rainfall hazard and a catchment saturation hazard) is developed to produce a distribution of daily flows in ungauged site.
- finally, preliminary work provides a way for the transition to the peak flow distribution using a hydrograph generator adapted to the sequence of daily simulated runoffs.

All these developments and conclusions are detailed and justified in the thesis.

Keywords : flood estimation, SCHADEX method, stochastic simulation, extreme precipitations, ungauged site, weather patterns, areal rainfall, SPAZM rainfall interpolator, peak flow.

Table des matières

I	Introduction	1
I. 1	Pourquoi étudier les débits extrêmes ?	5
I. 2	La méthode SCHADEX pour l'estimation des débits extrêmes	10
I. 3	Estimation des crues extrêmes sur un bassin non jaugé	18
I. 4	La problématique et la méthodologie pour l'estimation des évènements extrêmes en contexte non jaugé	20
II	Socle méthodologique requis	25
II. 1	La classification par « type de temps » EDF	29
II. 1. 1	Contexte	29
II. 1. 2	Une approche bottom-up pour l'identification des types de temps	29
II. 1. 3	Comparaison des classifications EDF2006 et EDF2010	32
II. 2	L'interpolateur de pluie SPAZM	35
II. 2. 1	La zone de construction des pluies SPAZM	35
II. 2. 2	La base pluviométrique d'entrée	35
II. 2. 3	Les ébauches de pluie par type de temps	36
II. 2. 4	Les pluies journalières	37
II. 3	Les modèles probabilistes des pluies	39
II. 3. 1	La loi d'ajustement MEWP	39
II. 3. 2	Les pluies adjacentes	42
I	Les pluies extrêmes	43
III	La régionalisation des pluies extrêmes journalières ponctuelles	45
III. 1	Jeux de données et protocole de calage-validation	49
III. 1. 1	Le jeu de données pour la mise en place des méthodes	50
III. 1. 2	Comment valider les estimations régionales d'extrêmes ?	50
III. 2	Les méthodes de régionalisation des pluies extrêmes	54
III. 2. 1	Méthode 1 : Krigeage des quantiles de pluies	54
III. 2. 2	Méthode 2 : Spatialisation de la distribution MEWP par krigeage des gradex	60
III. 2. 3	Méthode 3 : Estimation des pluies extrêmes locales avec les pluies SPAZM	76
III. 3	Comparaisons des approches de régionalisation	83
III. 3. 1	Évaluation sur le jeu de 770 stations	85
III. 3. 2	Évaluation avec le protocole ExtraFlo	91
III. 3. 3	Les cartes de pluies extrêmes proposées	105
III. 4	Vers une estimation des pluies extrêmes à un pas de temps infra-journalier	109
III. 4. 1	Mise en évidence d'une zone « atypique » pour les précipitations extrêmes journalières	109
III. 4. 2	Vers une distribution MEWP pour l'infra-journalier	112
III. 5	Conclusion	114

IV Les pluies journalières spatiales pour l'estimation d'extrêmes	117
IV. 1 Problématique et objectifs	121
IV. 1. 1 La pluie spatiale avec les polygones de Thiessen	122
IV. 1. 2 La pluie spatiale avec SPAZM	123
IV. 2 Comparaison des deux approches pour l'estimation de pluies spatiales	124
IV. 2. 1 La base de données pour les pluies spatiales	124
IV. 2. 2 Comparaison des cumuls annuels des pluies spatiales	126
IV. 2. 3 Comparaison des estimations des pluies extrêmes	128
IV. 2. 4 Sensibilité de Thiessen aux postes utilisés : Cas de l'Eau d'Olle	132
IV. 3 Utilisation du générateur de champs de pluies SAMPO	135
IV. 3. 1 Présentation du générateur SAMPO	135
IV. 3. 2 Données d'entrée pour le générateur	135
IV. 3. 3 Étude de l'abattement par simulation	138
IV. 4 Conclusion	150
II Les débits extrêmes	151
V Le passage pluie-débit	153
V. 1 La base d'études SCHADEX	157
V. 2 Vers une méthode de simulation stochastique pluie-débit peu paramétrée : la méthode EREBOR	161
V. 2. 1 Présentation du modèle SCS	162
V. 2. 2 Le principe de la simulation avec le SCS	166
V. 3 Régionalisation des paramètres de la méthode EREBOR	173
V. 3. 1 La régionalisation des paramètres	173
V. 3. 2 Validation croisée de la simulation EREBOR	185
V. 4 Automatisation de la simulation SCHADEX	188
V. 4. 1 Sélection des saisons pour les précipitations	188
V. 4. 2 Calage automatique du modèle hydrologique MORDOR	193
V. 4. 3 Évaluation de la simulation SCHADEX automatique	200
V. 5 Passage à la distribution des débits de pointes	202
V. 5. 1 La méthode classique : le coefficient de forme moyen	202
V. 5. 2 La génération d'hydrogrammes pour des scénarios en volume	203
V. 6 Conclusion	210
VI Exemple d'étude de crues extrêmes en non jaugé : l'Eyrieux à Beauvène	213
VI. 1 Présentation du bassin de l'Eyrieux à Beauvène	217
VI. 2 Les modèles probabilistes des pluies	217
VI. 3 Passage pluie-débit en contexte non jaugé : Simulation EREBOR	221
VI. 4 Passage pluie-débit en contexte jaugé : méthode SCHADEX	222
III Conclusion générale	225
Annexe A Le modèle hydrologique MORDOR	235
Bibliographie	244

Table des figures

I.1	Schéma d'un bassin versant topographique.	6
I.2	Exemple d'une distribution des débits de pointe sur le bassin du Tarn à Millau.	8
I.3	Schéma du calcul du temps de base de l'écoulement de l'hydrogramme adimensionné (courbe rouge) pour le bassin de l'Ardèche à Sauze (Garavaglia, 2011b).	12
I.4	Sélection d'épisodes pluvieux pour la méthode SCHADEX (Garavaglia, 2011b).	12
I.5	Schéma conceptuel du modèle MORDOR.	14
I.6	Schématisation de la simulation d'un écoulement (Garavaglia, 2011b).	15
I.7	Procédé de simulation hydrologique de SCHADEX (Garavaglia, 2011b).	15
I.8	Schéma d'un scénario d'écoulement d'une simulation SCHADEX. L'hydrogramme rouge n'est connu que pour les hydrogrammes observés de la sélection (mais pas pour les scénarios SCHADEX).	16
I.9	Exemple de distributions obtenues avec la méthode SCHADEX pour le bassin du Tarn à Millau (2170 km ²).	17
I.10	Schéma du cheminement méthodologique envisagé.	22
II.1	Localisation des 37 radiosondages (points rouges) et des grilles d'analyse utilisées pour la classification EDF2006 (fenêtre bleue) et pour la classification EDF2010 (fenêtre verte).	30
II.2	Méthodologie pour la construction de la classification en types de temps (Garavaglia, 2011b).	31
II.3	Champs géopotentiels moyens (1000hPa-0h) de la classification en types de temps EDF2006. Le cadre montre la fenêtre d'analyse (de 6.2°Ouest à 12.9°Est et de 38.0°Nord à 50.3°Nord) et la flèche indique le flux atmosphérique de basse couche (Mathevet et Garçon, 2010; Garavaglia, 2011b).	33
II.4	Ratio entre la précipitation moyenne de chaque type de temps et la précipitation moyenne globale (tous types de temps confondus). Les zones en bleu sont celles où le type de temps considéré apporte des pluies supérieures à la moyenne globale (Garavaglia, 2011b).	34
II.5	Zone d'étude de SPAZM (Gottardi, 2009).	35
II.6	Illustration des distances euclidiennes et de la distance de franchissement (Gottardi, 2009).	37
II.7	Schéma d'un épisode de pluie utilisé dans SCHADEX (Garavaglia, 2011b).	39
II.8	Loi MEWP sur la station DTG 2834 (Clans Bancairon, Alpes Maritimes). En haut à gauche, l'ajustement sur la saison à risque (de septembre à novembre) avec les huit lois exponentielles des types de temps. En haut à droite, le même ajustement pour la saison complémentaire. En bas, la distribution globale issue de la recombinaison des deux ajustements saisonniers.	40
II.9	Répartition spatiale des saisons à risque (Garavaglia, 2011b).	41
II.10	Boxplots des maxima sur trois mois glissants de l'archive régionale A et B (Garavaglia, 2011b).	41
III.1	Réseau de 770 postes pluviométriques	51

III.2 Réseaux de différentes densités. Les points rouges sont les postes conservés pour la régionalisation, les points bleus sont les stations servant à la validation.	52
III.3 Distribution des altitudes des stations utilisées en validation pour chaque réseau considéré.	53
III.4 Définition de la portée, du palier et de la pépite d'un variogramme.	55
III.5 Semi-variogrammes (en mm^2) des quantiles de pluies de temps de retour 10, 100, 1000 et 10 000 ans calés sur la totalité du réseau.	56
III.6 Cartes des quantiles de pluie de différents temps de retour obtenues par krigeage sur la totalité des 770 postes.	57
III.7 Scores pour différentes densités de réseau pour la pluie millénale.	58
III.8 Graphe quantile-quantile des pluies centennales (en mm) estimées localement et régionalement en Jackknife.	59
III.9 Graphe scatter-plot des pluies centennales (en mm) estimées localement et régionalement en Jackknife. Les droites en pointillé représentent un intervalle de +/- 20%.	59
III.10 Comparaison des paramètres estimés sur les séries observées et sur les séries SPAZM pour la saison à risque.	64
III.11 Représentation des huit classes de types de temps dans l'espace à deux dimensions des deux premières composantes principales.	67
III.12 Vue en plan des deux transects réalisés.	68
III.13 Comportement des gradex et des pluies journalières moyennes par type de temps et globaux pour un transect du Jura.	69
III.14 Comportement des gradex et des pluies journalières moyennes par type de temps et globaux pour un transect des Cévennes.	70
III.15 Semi-variogrammes par types de temps. En vert foncé le semi-variogramme du gradex de la saison (tous types de temps confondus) et en noir le semi-variogramme du gradex global. Les variables explicatives : Pluie annuelle et distance à la Méditerranée.	71
III.16 Exemple de reconstruction d'une loi MEWP.	75
III.17 Carte des ratios entre les estimations de pluies millénales SPAZM en Jackknife et les estimations locales.	77
III.18 Scatter-plot des estimations des pluies (en mm) de temps de retour 10, 100 et 1000 ans avec et sans correction sur les quantiles estimés avec la pluie SPAZM en Jackknife.	79
III.19 Correction des pluies de temps de retour 1000 ans estimées avec SPAZM.	81
III.20 Scatter-plot des précipitations millénales locales et obtenues avec SPAZM en Jackknife (avec ou sans correction). Les écarts-types σ_{KO} sont issus d'un krigeage ordinaire en Jackknife.	82
III.21 Les grandes étapes de chacune des méthodes.	83
III.22 Scatter-plot entre les estimations locales et régionales pour la pluie millénale sur un réseau de validation de 50%.	86
III.23 Scores pour chaque méthode sur la pluie décennale en fonction de la densité de réseau utilisée.	88
III.24 Scores pour chaque méthode sur la pluie centennale en fonction de la densité de réseau utilisée.	88
III.25 Scores pour chaque méthode sur la pluie millénale en fonction de la densité de réseau utilisée.	89
III.26 Scores pour chaque méthode sur la pluie décamillénale en fonction de la densité de réseau utilisée.	89
III.27 Jeu de données pour les méthodes locales-régionales.	91
III.28 Décomposition calage/validation pour les méthodes locales-régionales.	92
III.29 Décomposition pour le test de la robustesse à l'information régionale.	93
III.30 Décomposition pour le test de la robustesse à l'information locale.	94
III.31 Décomposition pour le test de la robustesse générale.	94
III.32 Graphes types de score FF (Garavaglia, 2011b).	96
III.33 Principe du modèle SHYPRE (Aubert, 2012).	98

III.34	Scores de justesse en calage et en validation pour une durée d'information locale de 20 ans.	100
III.35	Scores de justesse en calage et en validation pour une durée d'information locale de 10 ans.	101
III.36	Scores de justesse en calage et en validation pour une durée d'information locale de 5 ans.	102
III.37	Schéma explicatif pour la construction des cartes de pluies journalières de temps de retour T	105
III.38	Cartes des pluies annuelles (PAM) et des quantiles de pluies journalières PJ(T).	106
III.39	Cartes des gradex de chaque type de temps pour la saison à risque et des pluies journalières moyennes de ce type de temps.	108
III.40	Mise en évidence d'une zone atypique (sur le jeu des 770 postes).	110
III.41	Influence du paramètre de forme sur le comportement asymptotique d'une distribution.	110
III.42	Score ExtraFlo selon la zone.	111
III.43	Ajustement de différentes lois sur des échantillons tirés dans une loi exponentielle.	113
IV.1	Schéma illustratif pour la construction d'une pluie spatiale avec les polygones de Thiessen.	122
IV.2	Schéma illustratif pour le calcul d'une pluie spatiale avec la pluie SPAZM.	123
IV.3	Cartes des bassins versants par classe de surfaces.	125
IV.4	Disponibilité des données selon les années.	126
IV.5	Sélection des bassins selon la disponibilité d'information (les points gris représentent les postes disponibles).	126
IV.6	Ratio des volumes annuels SPAZM sur ceux de Thiessen selon l'écart entre l'altitude médiane du bassin et l'altitude du panier Thiessen.	127
IV.7	Ratio de la pluie journalière moyenne SPAZM sur Thiessen en fonction de la fréquence (moyenne sur toute la base de bassins).	127
IV.8	Ratio de la pluie millénale Thiessen corrigée sur celle de SPAZM selon le nombre de postes avec Thiessen.	128
IV.9	Différentes pluies spatiales avec Thiessen sur le bassin de l'Eau d'Olle à La Pernière (169 km ²)	133
IV.10	Exemple de champ d'intermittence généré par SAMPO (en noir : les zones pluvieuses, en blanc : les zones où les pluies sont nulles).	136
IV.11	Exemple de champ de pluie généré par SAMPO.	137
IV.12	Graphes fréquences-fréquences pour les différentes lois des pluies ponctuelles non nulles P_{nm}	138
IV.13	Schéma de la création de champs de pluie SAMPO et SPAZM.	139
IV.14	Exemple de champ journalier obtenu avec SAMPO et SPAZM (les points noirs représentent le réseau sol).	140
IV.15	Courbes des abattements de SAMPO pour les pluies de temps de retour T selon la surface du bassin.	142
IV.16	Comparaison des abattements de SAMPO et SPAZM pour la pluie décennale selon la surface du bassin.	143
IV.17	Lien entre l'abattement sur les quantiles et sur les écarts-types.	145
IV.18	Dispersion des N postes (points rouges) utilisées par SPAZM pour calculer la pluie spatiale avec une densité de postes constante et des bassins de surfaces variables.	145
IV.19	Lien entre l'abattement sur les écarts-types β et l'aire d'effet AOE.	147
IV.20	Estimation du ratio d'abattement SPAZM et SAMPO.	148
IV.21	Organigramme de synthèse de la procédure préconisée pour l'estimation des forts quantiles d'une pluie spatiale SPAZM.	149
V.1	Les débits mensuels normalisés des bassins de la base et les cinq classes de régimes hydrologiques.	158
V.2	Bassins versants de la base.	159
V.3	Distinction des différents types d'écoulement dans un hydrogramme de crue.	164
V.4	Un exemple d'événement simulé pour le bassin de l'Agout à Fraïsse.	167

V.5	Influence des différents paramètres du modèle SCS sur la simulation d'un événement.	168
V.6	Distribution par simulation avec le SCS avec ou sans addition d'un risque nival.	169
V.7	Schéma d'une simulation EREBOR.	171
V.8	Schéma d'un évènement pluie et de l'évènement débit associé.	173
V.9	Distributions des contributions nivales F_{CN} par régime hydrologique.	174
V.10	Boxplot des valeurs de l'hydrogramme unitaire MORDOR sur la base des 45 études par régime hydrologique.	175
V.11	Boxplot des valeurs de l'hydrogramme unitaire par régime hydrologique et selon deux classes de surface.	176
V.12	Hydrogrammes unitaires choisis par classe de régimes hydrologiques.	176
V.13	Distributions de débits obtenues pour différents hydrogrammes unitaires. Cas de la Romanche à Champeau (Nivo-glaciaire, 998 km ²).	177
V.14	Distributions des états de saturation λ via MORDOR pour les bassins pluviaux selon les stocks MORDOR considérés.	178
V.15	Distributions des états de saturation λ par régime hydrologique.	179
V.16	Différentes distributions des états de saturation λ et distributions EREBOR associées. Cas de la Romanche à Champeau (Nivo-glaciaire, 998 km ²).	180
V.17	Illustration de l'optimisation due S_{max} sur le cas de l'Agout à Fraïsse (Pluvial, 45 km ²).	181
V.18	Comparaison des écoulements journaliers de temps de retour T obtenus avec la simulation SCHADEX et avec la simulation EREBOR (S_{max} optimisé).	182
V.19	Sens hydrologique des capacités maximales de rétention obtenues pour la base des 45 études SCHADEX disponibles.	183
V.20	Cartographie des régimes hydrologiques selon les douze classes établies par Sauquet <i>et al.</i> (2008).	184
V.21	Comparaison des écoulements journaliers de temps de retour T obtenus avec la simulation SCHADEX et avec la simulation EREBOR (en validation croisée).	186
V.22	Schéma de l'algorithme de sélection des saisons.	190
V.23	Comparaison des pluies de temps de retour T obtenues avec les saisons expertisées et les saisons automatisées sur la base de 45 études SCHADEX.	192
V.24	Les distributions de débits journaliers obtenues avec la méthode SCHADEX sur le bassin de la Creuse à Fresseline (1238 km ²).	194
V.25	Période de calage et de validation d'un modèle i parmi les 100 modèles du boot-strap.	195
V.26	Mise en évidence de deux groupes de distributions aux comportements asymptotiques distinctes pour la Maulde à Peyrat Le Château (174 km ²).	197
V.27	Espace à 2 dimensions pour identifier les meilleurs modèles selon $KGE_{Q_{mensuel}}$	197
V.28	Espace à 2 dimensions pour identifier le meilleur modèle selon $KGE_{Q_{reg}}$	198
V.29	Exemple de réduction progressive du faisceau de boot-strap pour le Lot à Castelnau (1628 km ²).	199
V.30	Comparaison des quantiles d'écoulements (en mm) entre les distributions SCHADEX de référence et automatisée pour différents temps de retour.	201
V.31	Exemple de simulation pour l'Eau d'Olle à Pernière – hydrogramme du donneur n°1.	205
V.32	Exemple de simulation pour l'Eau d'Olle à Pernière – les 10 hydrogrammes synthétiques sélectionnés.	206
V.33	Hydrogramme synthétique généré pour le scénario de crue l'Eau d'Olle à Pernière.	207
V.34	Reconstitution d'un échantillon de crues historiques – comparaison des approches.	208
VI.1	Localisation du bassin de l'Eyrieux à Beauvène (454 km ²).	217
VI.2	Détermination automatique de la saison.	218
VI.3	Modèle probabiliste des pluies centrales – Ajustement MEWP sur les pluies journalières SPAZM pour l'Eyrieux à Beauvène (454 km ²), période 1953-2005.	219
VI.4	L'Aire d'Effet (AOE) de SPAZM pour le bassin de l'Eyrieux à Beauvène (454 km ²).	220
VI.5	Bassins similaires à l'Eyrieux à Beauvène (454 km ²) (en rouge) – Chassezac à Malarce (484 km ²) et à Sainte Marguerite (414 km ²) (en vert).	221
VI.6	Modèle MORDOR pour l'Eyrieux à Beauvène (454 km ²).	222

VI.7 Distributions des débits moyens journaliers pour l'Eyrieux à Beauvène (454 km ²).	223
VI.8 Synthèse des résultats.	227
VI.9 Illustration de la sous-estimation des plus fortes valeurs pour la pluie millénale (avec la méthode de régionalisation des paramètres MEWP).	229
VI.10 Exemple de l'AOE (aire du disque bleu) pour deux bassins versants.	230
A.1 Schéma conceptuel et équations du modèle MORDOR.	236

Liste des tableaux

II.1	Les huit types de temps de la classification EDF et leurs occurrences pour chacune des deux classifications.	32
II.2	Modèle des pluies adjacentes. Tableau de contingence pour le bassin versant de l'Ar-dèche à Sauze (2260 km ²), période 1959-2008 (Garavaglia, 2011b).	42
III.1	Information sur les différents réseaux utilisés pour la validation	52
III.2	Paramètres des semi-variogrammes des quantiles de pluies de temps de retour 10, 100, 1000 et 10 000 ans pour différentes densités de réseaux.	56
III.3	Coefficient de détermination entre les gradex par type de temps et des variables explicatives potentielles.	66
III.4	Score du krigeage des gradex par type de temps (TT) en Jackknife avec les 770 stations (pour la saison à risque).	73
III.5	Score en Jackknife sur les pluies SPAZM pour différents temps de retour.	76
III.6	Différents ratios entre les pluies SPAZM et les pluies journalières mesurées	76
III.7	Score en Jackknife sur les pluies SPAZM pour différents temps de retour avec une correction sur les quantiles estimés.	78
III.8	Les trois méthodes envisagées et leurs variantes.	83
III.9	Scores de justesse en validation pour les différentes méthodes évaluées dans le projet ExtraFlo.	103
III.10	Score de robustesse à l'information régionale pour les différentes méthodes régionales.	104
IV.1	Répartition des ratios en volumes et en quantiles millénaux selon la classe de surface.	130
IV.2	Données d'entrée pour le générateur SAMPO.	139
IV.3	Coefficient $a(T)$ de la régression linéaire entre $K_{P(T)}$ et $\frac{\beta_{SAMPO}}{\beta_{SPAZM}}$	145
V.1	Les 45 études SCHADEX au pas de temps journalier de la base.	160
V.2	Sensibilité des paramètres MORDOR à la période de calage sur la Durance à La Clapière.	162
V.3	Capacités maximales de rétention médianes par classe de régime hydrologique.	185
VI.1	Modèle probabiliste des pluies adjacentes – Tableau de contingence pour l'Eyrieux à Beauvène (454 km ²), période 1953-2005.	220
VI.2	Les écoulements (en mm) pour différents temps de retour estimés avec la méthode SCHADEX et la méthode EREBOR.	223
A.1	Paramètres du modèle hydrologique MORDOR.	235

Introduction

La prédétermination des débits de crues est un axe de recherche important de l'hydrologie du fait des enjeux socio-économiques. À EDF-DTG, l'estimation des crues extrêmes est réalisée avec la méthode SCHADEX (Paquet *et al.*, 2013). Cette approche permet une évaluation des crues extrêmes pour des sites disposant de suffisamment de données pour son application. Cependant, son utilisation en l'absence de ces données reste difficile et constitue la problématique de cette thèse. Ce chapitre d'introduction présente la notion d'estimation des débits extrêmes et le principe de la méthode SCHADEX. La problématique de l'utilisation de la méthode SCHADEX en site non jaugé et le cheminement méthodologique envisagé pour répondre à ce problème sont également détaillés.



SOMMAIRE DU CHAPITRE

I. 1	Pourquoi étudier les débits extrêmes ?	5
I. 2	La méthode SCHADEx pour l'estimation des débits extrêmes	10
I. 3	Estimation des crues extrêmes sur un bassin non jaugé	18
I. 4	La problématique et la méthodologie pour l'estimation des évènements extrêmes en contexte non jaugé	20

I. 1 Pourquoi étudier les débits extrêmes ?

Selon l'IDD¹, les inondations représentent l'une des catastrophes naturelles les plus fréquentes et coûteuses dans le monde : elles concernent plus de 3 milliards de personnes (parmi lesquelles environ 7 millions de victimes) avec des dommages évalués à plus de 350 milliards d'euros sur la période 1901-2010. La France est directement concernée avec 8% de la population potentiellement concernée (5 millions de personnes, 16 000 communes), 700 millions d'euros de dommages évalués (coût annuel sur la période 1995-2006) et 185 victimes (sur la période 1980-2000). Nous avons tous en tête les dernières crues dévastatrices, pour n'en citer que quelques-unes : les crues du Borne (Grand-Bornand, 1987), de l'Ouvèze (Vaison-la-Romaine, 1992), de la Corse (1993), du Var (1994), de l'Aude (1999), de la Laïta à Quimperlé (2000), des Gardons (2002, 2005), de Nîmes (1988, 2005), du Rhône (2003), de la Moselle (2006), du Gave de Pau (2013), etc.

Les crues induisent, entre autres, un risque pour les ouvrages hydrauliques tels que les barrages. En 1950, il n'y avait que 500 barrages importants² dans le monde. Aujourd'hui, 45 000 sont répertoriés, ce qui correspond à un rythme de 2 nouveaux barrages mis en service chaque jour. Depuis 1800, près de 150 ruptures ont été dénombrées, dont certaines ont fait plus de 1000 morts³. D'après la base de données GRanD⁴, ces ruptures de barrage peuvent avoir plusieurs origines :

- techniques (défaut de fonctionnement des évacuateurs de crue, vices de conception, de construction ou de matériaux, vieillissement des installations),
- naturelles (séismes, crues exceptionnelles, glissements de terrain),
- humaines (insuffisance des études préalables et du contrôle d'exécution, erreurs d'exploitation, de surveillance et d'entretien, malveillance).

Dans environ 30% des cas, la rupture est consécutive à une crue (évacuateur sous-dimensionné, mauvaise gestion du risque de crue). Pour illustrer un accident de ce type, nous pouvons citer par exemple la rupture du barrage Machhu (Inde) le 1^{er} août 1979, qui a fait plus que 2 000 morts.

Ces statistiques justifient complètement les préoccupations de l'homme liées aux crues, notamment à celles les plus rares et intenses, dont l'estimation est primordiale pour le dimensionnement des barrages.

Les séries de débits

Avant de rentrer plus en détail dans les méthodes utilisées pour estimer les débits rares ou extrêmes, nous allons rappeler quelques fondamentaux de l'hydrologie. La plupart du temps, l'unité hydrologique de référence est le bassin versant. Il est défini par un exutoire – une section d'une rivière – et par la surface drainant toute l'eau passant par l'exutoire. Un bassin versant est délimité par des frontières naturelles, appelées lignes de partage des eaux. Ces lignes sont généralement matérialisées par les lignes de crêtes⁵ (figure I.1).

1. International Disaster Database est une base de données disponible au lien suivant : <http://www.emdat.be/database>.

2. Grands barrages dont la hauteur sur fondation dépasse 15 mètres.

3. Quelle que soit la hauteur de l'ouvrage, 144 ruptures de barrage ont été dénombrées dans le monde dont : 71 sans victime, 31 avec moins de 10 morts, 17 avec un nombre de morts compris entre 10 et 99 et 25 avec 100 morts et plus. Les données, fournies par l'Hydrocoop, peuvent être plus au moins fiables en fonction de la politique de diffusion adoptée par chaque pays concerné.

4. Global Reservoir and Dam database.

5. On parle alors de bassin versant topographique. Un bassin versant hydrologique peut aussi être défini. Les lignes de partage des eaux tiennent alors compte des transferts souterrains et du drainage. Il peut alors être différent d'un bassin versant topographique.

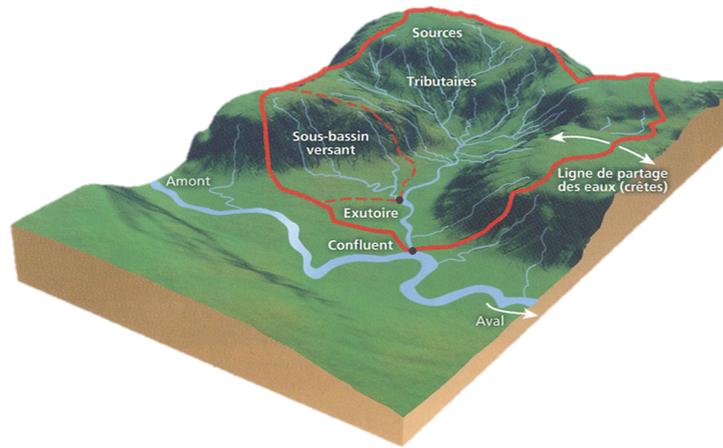


FIGURE I.1 – Schéma d'un bassin versant topographique.

Les écoulements d'un bassin versant sont appréhendés au travers des débits qui transitent au niveau de l'exutoire. La courbe qui montre l'évolution des débits dans le temps est appelée hydrogramme. Les hydrogrammes présentent des variations plus ou moins marquées au cours de l'année en lien avec la climatologie, la géologie du bassin versant, les influences anthropiques. . . Globalement, les hydrogrammes se composent d'une succession de hautes eaux (réponse plus ou moins rapide à la pluviométrie, à la fonte de neige. . .) et de basses eaux, aussi appelées étiages. Nous nous intéressons ici uniquement aux forts débits qui sont observés lors de crues. Dans le glossaire international d'hydrologie (WMO, 1972), la crue est définie comme étant une montée, en général rapide, du niveau d'un cours d'eau jusqu'à un maximum, suivi d'une descente plus lente. Les crues sont donc des phénomènes naturels de hautes eaux résultant de la propagation d'un volume d'eau important dans un réseau hydrographique (rivières, lacs. . .). Les crues doivent être distinguées des inondations qui sous-entendent qu'un cours d'eau (ou lac) sort de son lit mineur (ou de ses berges). Il peut donc y avoir une crue sans pour autant qu'elle occasionne une inondation.

Pour simplifier la suite du travail, quelques notations sur les débits vont être définies. Les débits instantanés, qui correspondent au débit passant au droit d'une section à un instant précis, sont à distinguer des débits moyens sur une période. Les premiers sont notés Q . Dans la pratique, les mesures de hauteur sont réalisées sur des pas de temps plus ou moins fins (pas de temps variables ou fixes en horaire, 10 heures, journalier. . .) et nous ne disposons pas toujours des débits instantanés. Cependant, une mesure au pas de temps horaire (ou plus fin) peut être considérée comme descriptive de l'allure de l'hydrogramme des débits instantanés Q pour les surfaces de bassins versants qui concernent les barrages hydro-électriques français. Les débits moyens journaliers (ou plus généralement, les débits moyens sur une durée d) sont calculés comme étant la moyenne de débits Q sur un jour (ou sur la durée d). Les débits moyens journaliers seront notés QJ (ou Q_d , les débits moyens sur une durée d). Le débit de pointe de crue QX représente le débit instantané Q maximum pendant un événement de crue. Le débit moyen maximum, noté QJX (ou Q_dX), est toujours plus faible que le débit instantané QX . Un ratio pointe-volume, appelé coefficient de forme K_C , peut alors être défini par :

$$K_C = \frac{QX}{Q_dX} \geq 1 \quad (\text{I.1})$$

En hydrologie, le débit moyen (flux en m^3/s) est souvent transformé en une lame d'eau moyenne équivalente⁶ sur le bassin versant. Les écoulements, notés E , s'expriment en milli-

6. Hauteur d'eau moyenne en millimètres correspondant au volume écoulé à l'exutoire sur un temps donné et réparti uniformément sur la surface du bassin.

mètres et sont utilisés pour la réalisation de bilans hydrologiques :

$$E = Q \cdot \frac{\text{Pas de temps}}{\text{Surface}} \quad (\text{I.2})$$

L'estimation des crues extrêmes

Parmi les méthodes couramment utilisées pour le calcul de crues extrêmes, deux grandes familles peuvent être distinguées : les méthodes maximalistes et les méthodes probabilistes. Une approche maximaliste aborde souvent l'estimation des crues extrêmes d'un point de vue physique⁷ et s'appuie sur le concept de la crue maximale probable (PMF – Probable Maximum Flood). La PMF est définie comme la crue résultant du croisement pessimiste de conditions météorologiques et hydrologiques (e.g. forte pluie et sol saturé) raisonnablement possibles sur le bassin versant considéré (CFGB, 1994). Une approche probabiliste est fondée sur des modèles statistiques qui s'appuient sur le concept de variable aléatoire appliqué au débit. En France, les approches probabilistes ont été historiquement privilégiées aux méthodes maximalistes, contrairement aux pays anglo-saxons. La réglementation française demande des estimations en terme probabiliste et nous ne parlerons que de ces méthodes dans ce mémoire. Le choix d'une approche probabiliste amène implicitement à parler de prédétermination des crues. La prédétermination consiste à associer une intensité à une probabilité d'occurrence. Si nous définissons un évènement comme le dépassement d'un débit Q_0 de référence, deux grandeurs peuvent lui être associées :

- la fréquence annuelle F au non-dépassement du débit Q_0 : $F = \text{Prob}(Q < Q_0)$
- le nombre moyen d'années T séparant deux occurrences d'un tel évènement. On parle alors de période de retour T ou de temps de retour T .

La relation entre ces deux variables est :

$$F = 1 - \frac{1}{T} \quad (\text{I.3})$$

Par exemple, une crue millénale est une crue qui revient en moyenne tous les mille ans et a une probabilité de 0.999 de ne pas être dépassée une année donnée. Une fréquence au non-dépassement peut être définie pour toutes les valeurs de débit.

La courbe représentant les débits en fonction de leur probabilité au non-dépassement F est appelée fonction de répartition. Lorsque l'on s'intéresse aux fortes valeurs de débits, la distribution est habituellement tracée dans un espace de Gumbel de façon à zoomer sur la gamme des fortes valeurs (figure I.2). Dans un tel espace, l'abscisse est la variable de Gumbel U telle que :

$$U = -\ln(-\ln(F)) \quad (\text{I.4})$$

Selon le temps de retour, une valeur est dite courante ($T < 10 \text{ ans}$), rare ($10 \text{ ans} \leq T$ et $T < 1000 \text{ ans}$) ou extrême ($T \geq 1000 \text{ ans}$). Les débits rares ou extrêmes sont, par définition, des valeurs très rarement observées et leur estimation demande donc d'extrapoler à partir des observations à disposition. Ceci explique le foisonnement de méthodes probabilistes utilisées pour calculer les débits associés à une période de retour donnée. En s'appuyant sur le travail réalisé par Lang et Lavabre (2007) et par Hingray *et al.* (2009), différentes méthodes peuvent être listées.

7. Les méthodes maximalistes s'appuient souvent sur une approche déterministe et physique mais peuvent également utiliser des outils plus probabilistes afin d'estimer une crue de probabilité de non dépassement de valeur 1.

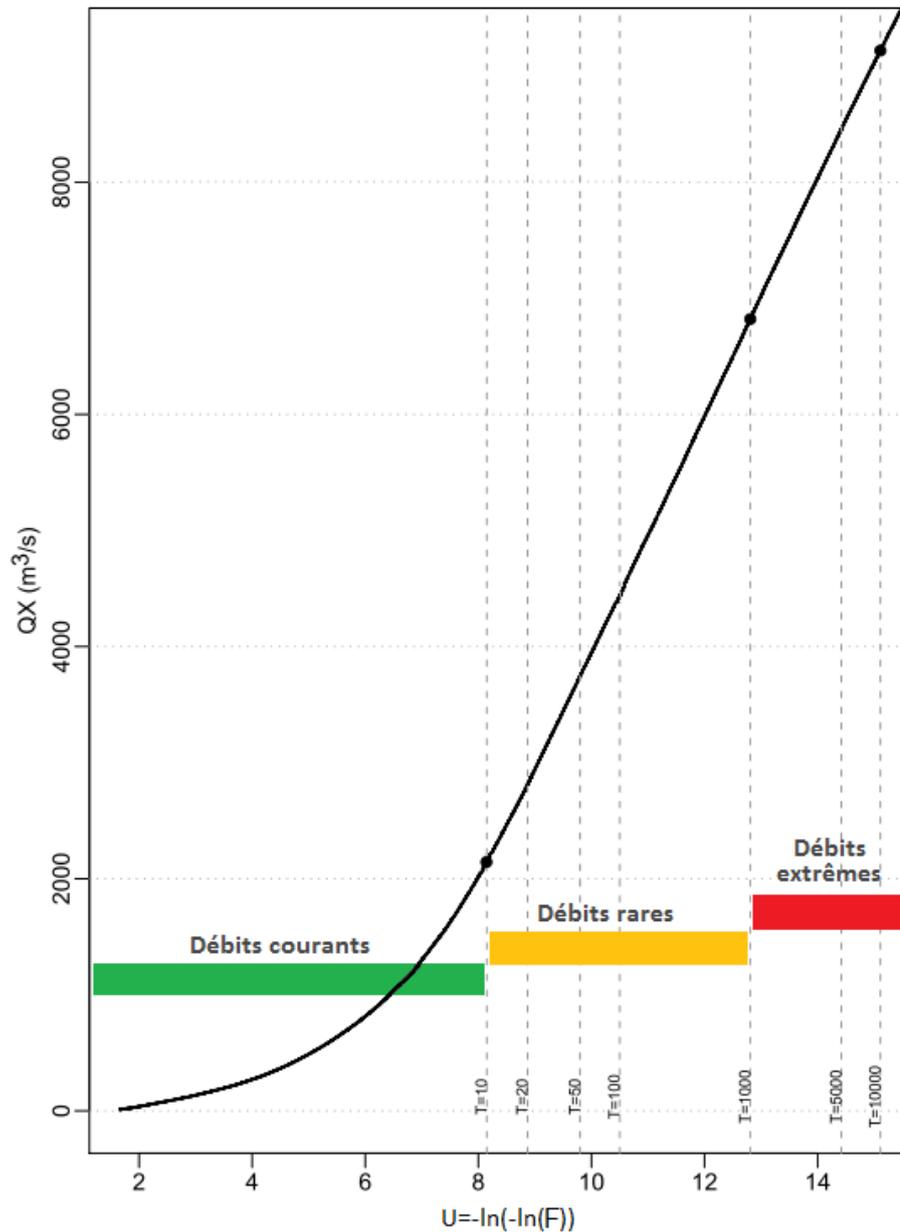


FIGURE I.2 – Exemple d'une distribution des débits de pointe sur le bassin du Tarn à Millau.

L'analyse fréquentielle des débits de crue. Il s'agit d'ajuster une loi de distribution statistique théorique (généralement issue de la théorie des valeurs extrêmes) à la loi de distribution empirique déduite directement des observations de débits au point d'étude considéré.

Les échantillons des débits de crue sur lesquels s'effectue l'analyse fréquentielle peuvent être construits de différentes façons, en sélectionnant soit les maxima par bloc (e.g. les maxima annuels), soit des débits sup-seuil. La sélection par valeurs supérieures à un seuil présente l'intérêt d'une sélection plus rationnelle des valeurs de crue, avec un nombre plus élevé de valeurs et la possibilité de retenir uniquement les plus fortes valeurs. Ainsi certaines crues maximales annuelles de faible ampleur ne sont pas sélectionnées. En contrepartie, il est nécessaire de choisir un seuil d'extraction de façon à ce que les crues retenues soient indépendantes, et d'imposer des contraintes d'espacement entre deux épisodes.

L'analyse fréquentielle peut s'effectuer sur les débits de pointe de crue ou sur les débits moyens maximaux correspondant à différentes durées.

Les méthodes régionales. Ces méthodes s'appuient sur l'information provenant de plusieurs postes afin de consolider l'analyse fréquentielle des débits de crue d'un site. Elles sont utilisées dans les cas où les observations sur le site d'étude sont insuffisantes ou absentes, ou pour réduire l'incertitude d'échantillonnage en site jaugé. Parmi ces approches, peuvent être mentionnées la méthode de l'indice de crue (Dalrymple, 1960), l'analyse régionale fréquentielle par les L-moments (Hosking et Wallis, 1997) et l'application d'un modèle bayésien de combinaison d'informations locale et régionale (Ribatet, 2007).

Les approches historiques. Il s'agit de s'appuyer sur des crues historiques pour compléter les chroniques de débits disponibles. Une analyse fréquentielle sur l'ensemble de l'échantillon (crues observées et crues historiques) est toujours réalisée, mais en faisant en sorte d'adapter la technique d'ajustement afin de prendre en compte l'information supplémentaire sur la période considérée. Des exemples de ces approches sont illustrés par les travaux de Ouarda *et al.* (1998); Naulet *et al.* (2005); Payraastre (2005).

Les approches naturalistes. Deux catégories sont envisageables parmi ces approches :

- L'approche hydrogéomorphologique. Cette méthode géographique appliquée étudie le fonctionnement naturel des cours d'eau en analysant la structure des vallées sur la base des observations et de l'interprétation du terrain naturel (Garry *et al.*, 1996).
- L'approche paléohydrologique. Cette méthode se fonde sur l'analyse des dépôts laissés par les crues (e.g. dans les champs d'inondation, dans les grottes perchées), sur la stratigraphie et sur la datation des dépôts, pour la reconstitution des débits de crue à utiliser dans une approche statistique (Sheffer *et al.*, 2003; Benito et Thorndycraft, 2006; Dezileau *et al.*, 2014).

Les approches basées sur la pluie. Il s'agit de toutes les méthodes qui s'appuient sur l'information pluviométrique pour extrapoler la distribution des crues extrêmes. Les types d'approches suivants peuvent notamment être distingués :

- Méthodes probabilistes simplifiées basées sur des hypothèses hydrologiques, comme la méthode du GRADEX (Guillot et Duband, 1967) et ses développements, parmi lesquels le modèle AGREGEE (Margoum, 1992; Margoum *et al.*, 1994; Lang, 1995).
- Méthodes fondées sur un simulateur d'averses couplé à un modèle pluie-débit, auxquelles appartiennent la méthode SCHADEX (Paquet *et al.*, 2013), le modèle SHYPRE (Cernesson, 1993; Arnaud, 1997) et le modèle SHYREG (Lavabre *et al.*, 2003; Aubert, 2012). Ce dernier est basé sur la régionalisation des paramètres du simulateur de pluies SHYPRE et sur un modèle pluie-débit simplifié.

À EDF, l'appréciation des débits extrêmes est réalisée à l'aide de la méthode SCHADEX. Ce travail de thèse repose sur les développements et principes de cette méthode, c'est pourquoi la méthode SCHADEX va être abordée plus en détail.

I. 2 La méthode SCHADEX pour l'estimation des débits extrêmes

Pour le dimensionnement des organes d'évacuation des crues de ses barrages, EDF a utilisé, depuis la fin des années 1960, la méthode du GRADEX. Celle-ci a été développée à EDF et a fait l'objet d'importantes modifications méthodologiques depuis. La méthode du GRADEX bénéficie d'une importante notoriété nationale et est abordée dans de nombreuses publications : Guillot et Duband (1967); Duband (1973); Guillot (1979, 1981, 1993), pour n'en citer que quelques-unes.

Cette méthode, en s'appuyant sur l'information pluviométrique, a constitué une amélioration substantielle des méthodes de prédétermination des crues basées sur le traitement statistique des seules valeurs de débits observés. Pour autant, malgré ses qualités, il ne s'agit que d'un modèle qui s'appuie nécessairement sur certaines approximations de modélisation :

1. La fonction de répartition des précipitations tombant sur le bassin, au pas de temps retenu comme pertinent, est asymptotiquement à décroissance exponentielle simple.
2. Le déficit d'écoulement, défini comme la différence entre le volume d'eau précipité et le volume d'eau écoulé à l'exutoire du bassin versant pendant un pas de temps, n'augmente plus, en valeur moyenne, au-delà d'une certaine valeur de pluie.
3. Quelle que soit la période de retour, le débit de pointe de crue peut être calculé à partir du volume écoulé (sur le pas de temps) en utilisant un coefficient multiplicatif moyen caractéristique du bassin versant.

Les ingénieurs d'EDF ont dû répondre au besoin de reformuler les risques pluviométriques et hydrologiques de manière plus approfondie. Pour cette raison, un travail exploratoire a été mené à partir de 1999 pour hiérarchiser les problèmes et définir de nouvelles pistes de développement, notamment sur l'ajustement des pluies. Le développement de la méthode SCHADEX (Simulation Climato-Hydrologique pour l'Appréciation des Débits EXtrêmes) a ensuite été mené de 2004 à 2006 à EDF-DTG, ce qui a permis d'aboutir à la première version opérationnelle de la méthode (Paquet *et al.*, 2006, 2013; Garavaglia, 2011b). La méthode SCHADEX apporte deux améliorations substantielles à la méthode du GRADEX :

- un modèle probabiliste utilisé pour la description de l'aléa pluviométrique, basé sur un échantillonnage par types de situations météorologiques (ou types de temps) ;
- l'utilisation d'un modèle hydrologique simulant la réponse du bassin à des précipitations de tous ordres pour approcher la distribution des écoulements rares et exceptionnels.

Les conclusions de la thèse de Garavaglia (2011b) stipulent que la méthode SCHADEX est applicable sur des bassins versants répondant aux conditions suivantes :

- disponibilité de chroniques climatologiques de pluie, température et débit sur une période suffisante pour le procédé de simulation hydrologique semi-continue et la procédure de calage des paramètres du modèle hydrologique pluie-débit (e.g. pour un bassin dont le pas de temps caractéristique est égal ou supérieur à 24 h, un minimum de 15 ans de données de pluie, de température et de débit est nécessaire) ;
- durée caractéristique des situations météorologiques génératrices de fortes précipitations ne dépassant pas les 72 h, i.e. un bassin versant où la corrélation des pluies permet de modéliser l'aléa pluviométrique avec un processus de génération stochastique d'épisodes pluvieux sur trois pas de temps successifs ;
- validité de l'hypothèse selon laquelle l'aléa pluviométrique joue un rôle prépondérant dans la genèse des crues extrêmes ;

- surface ne dépassant pas une certaine limite (i.e. quelques dizaines de milliers de km^2) sous laquelle l'approximation de l'aléa pluviométrique par une pluie globale du bassin versant et les concepts de modélisation hydrologique conceptuelle restent valables.

SCHADEX est une méthode de simulation stochastique semi-continue du fait qu'un événement de débit est issu d'un événement de pluie superposé à une simulation continue avec un modèle pluie-débit qui génère un historique des états de saturation du bassin. Nous allons maintenant aborder les différentes étapes réalisées lors d'une étude de crues extrêmes sur un bassin versant avec la méthode SCHADEX. Une description plus approfondie de la méthode peut être trouvée dans la thèse de Garavaglia (2011b) ou l'article de Paquet *et al.* (2013).

Étape 1 : Collecte de données

La collecte des données et leur critique est une étape importante et parfois délicate. Les données constituent le point de départ de l'étude sur lequel les modèles (hydrologiques ou probabilistes) sont calés. Des incohérences, des biais ou des instationnarités dans les données peuvent mener à des erreurs non négligeables sur les résultats de l'étude. Il est donc primordial d'essayer de collecter le maximum de données possible (spatialement et temporellement) afin de réaliser une critique robuste et fiable des données utilisées.

Étape 2 : Sélection des hydrogrammes de crues

La simulation stochastique SCHADEX permet de simuler des volumes sur un pas de temps caractéristique du bassin versant (e.g. journalier, 12 heures...). Un facteur, appelé coefficient de forme, permet de passer de la distribution des volumes à la distribution des débits de pointe. L'estimation du pas de temps caractéristique et du coefficient de forme est issue de l'analyse d'une population d'hydrogrammes de crues significatives sur le bassin d'intérêt.

L'étude des crues extrêmes est menée sur une durée caractéristique des principaux épisodes de crues observés sur le bassin. Généralement, elle est considérée égale au temps de base moyen des hydrogrammes de crue, c'est-à-dire à la durée qui s'écoule entre la montée significative de débit et le moment où la descente amorce une décroissance plus lente (Le Clerc, 2004). La durée du pas de temps adapté à la dynamique des crues du bassin correspond à une valeur de coefficient de forme moyen proche de 1.4 (notion détaillée dans l'étape 7). Le pas de temps caractéristique du bassin pdt est donc un compromis entre cette durée et le pas de temps des données disponibles (Garavaglia, 2011b).

À partir d'une sélection d'hydrogrammes contrôlée avec un échantillonnage sup-seuil (Lang *et al.*, 1999), un hydrogramme moyen adimensionnalisé (i.e. la moyenne de tous les hydrogrammes de la sélection, synchronisés et normalisés sur la pointe de crue) est calculé. Par convention, le temps de base de l'écoulement de cet hydrogramme est considéré comme égal à la durée entre le décollement du débit (début de l'hydrogramme) et le moment où le débit redescend sous un niveau égal à $1/3$ du débit de pointe (début de la récession). Le pas de temps caractéristique du bassin doit alors être choisi inférieur ou égal au temps de base de l'écoulement (figure I.3).

Étape 3 : Construction d'une pluie spatiale

Une pluie dite « de bassin » ou « spatiale » est ensuite construite à partir de l'information pluviométrique disponible sur le bassin versant. Cette démarche est une conséquence directe de l'utilisation d'un modèle hydrologique global qui s'appuie sur ce type d'entrée pour décrire la pluie. La combinaison de postes pluviométriques est déterminée de façon à fournir une série de pluie spatiale avec des postes représentant au mieux les pluies du bassin (proximité, altitude...) et dont le cumul annuel moyen est cohérent avec le bilan hydrologique du bassin versant.

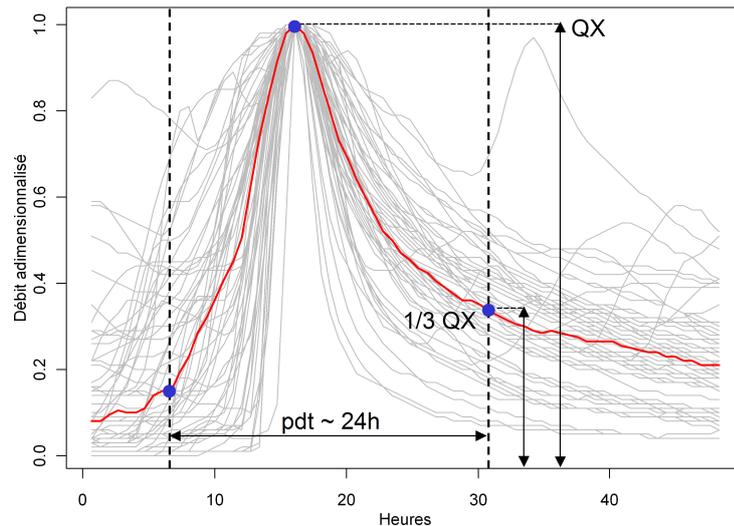


FIGURE I.3 – Schéma du calcul du temps de base de l'écoulement de l'hydrogramme adimensionné (courbe rouge) pour le bassin de l'Ardèche à Sauze (Garavaglia, 2011b).

Étape 4 : Calage des modèles probabilistes des pluies

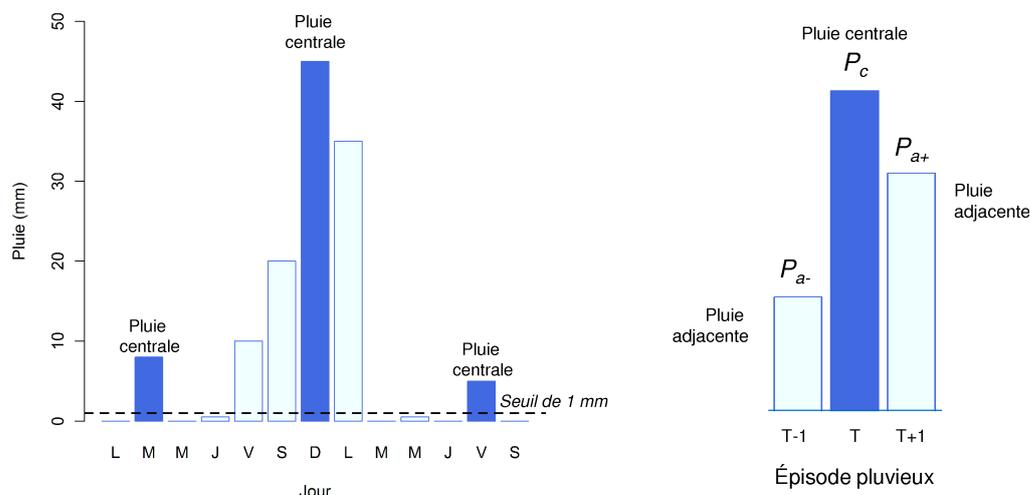
Dans la méthode SCHADEX, les événements de pluies considérés sont des événements triangulaires distribués sur trois pas de temps. Par construction, le maximum est au milieu et il est dénommé pluie centrale (P_C). Les deux pluies de chaque côté d'une pluie centrale sont appelées pluies adjacentes (P_{a-} et P_{a+}). Pour qu'une pluie soit définie comme une pluie centrale, elle doit respecter les trois conditions suivantes :

Condition 1 : excéder 1 mm ($P_C \geq 1\text{mm}$).

Condition 2 : être supérieure à la pluie du pas de temps précédent ($P_C \geq P_{a-}$).

Condition 3 : être supérieure à la pluie du pas de temps suivant ($P_C > P_{a+}$).

La figure I.4 illustre graphiquement ces concepts.



(a) Identification graphique des pluies centrales dans une chronique pluviométrique.

(b) Division d'un épisode pluvieux en pluie centrale et pluies adjacentes.

FIGURE I.4 – Sélection d'épisodes pluvieux pour la méthode SCHADEX (Garavaglia, 2011b).

Cette schématisation a été introduite exclusivement pour des raisons liées au processus de simulation hydrologique de SCHADEX pour la prédétermination des crues extrêmes. Le modèle probabiliste utilisé pour décrire les pluies centrales est une distribution MEWP saisonnière (Multi-Exponential Weather Pattern) mise en place par Garavaglia *et al.* (2010) et qui sera détaillée dans le chapitre suivant. Il est important de noter que dans SCHADEX, les distributions de pluies sont ajustées sur une population de pluies centrales et non pas de maxima annuels. Cependant, les pluies maximales annuelles sont incluses dans les pluies centrales. La distribution des pluies centrales converge asymptotiquement vers celle des pluies maximales annuelles.

Les probabilités des pluies adjacentes P_{a^-} et P_{a^+} sont également décrites par l'intermédiaire de tableaux de contingence. Ces tableaux indiquent les probabilités des ratios P_{a^-}/P_C et P_{a^+}/P_C selon la valeur des pluies centrales P_C . Le fonctionnement de ces tableaux sera abordé plus précisément dans le chapitre suivant.

À partir de la chronique de pluie spatiale au pas de temps de l'étude et des modèles probabilistes ajustés, l'aléa des épisodes pluvieux est totalement décrit.

Étape 5 : Calage du modèle MORDOR

Un modèle hydrologique pluie-débit MORDOR (MOdèle à Réservoirs de Détermination Objective du Ruissellement) est calé sur les observations critiquées (pluie spatiale, température sur le bassin et débits observés à l'exutoire).

Le modèle hydrologique MORDOR (Garçon, 1996; Garçon *et al.*, 1999; Paquet et Garçon, 2000; Garçon, 2000; Andréassian *et al.*, 2006) est utilisé à EDF-DTG depuis plus de 20 ans pour la prévision des débits à des échéances pouvant aller de quelques heures à quelques mois, ainsi que pour des études hydrologiques variées et dont les valeurs doivent être identifiées par calibration sur des données d'observations.

C'est un modèle conceptuel global à réservoirs qui intègre dans son algorithme un module d'accumulation et de fonte de la neige. À partir de données de pluies et de températures journalières, il fait évoluer schématiquement l'état d'un certain nombre de réserves (stock de neige, nappe. . .) qui interagissent et alimentent la rivière à l'exutoire du bassin. Le schéma conceptuel du modèle MORDOR est reporté sur la figure I.5.

Les transferts d'eau entre l'atmosphère, les différents réservoirs et les écoulements superficiels sont réglés par des formulations mathématiques simples (annexe A). Le modèle cherche à représenter la dynamique des processus lents ou rapides du bassin versant (dynamique de nappe, ruissellement, écoulement superficiel, accumulation et fonte de la neige, évapotranspiration, infiltration). Les transferts d'eau entre les différents réservoirs sont régis par une trentaine de paramètres, dépendants du bassin versant étudié.

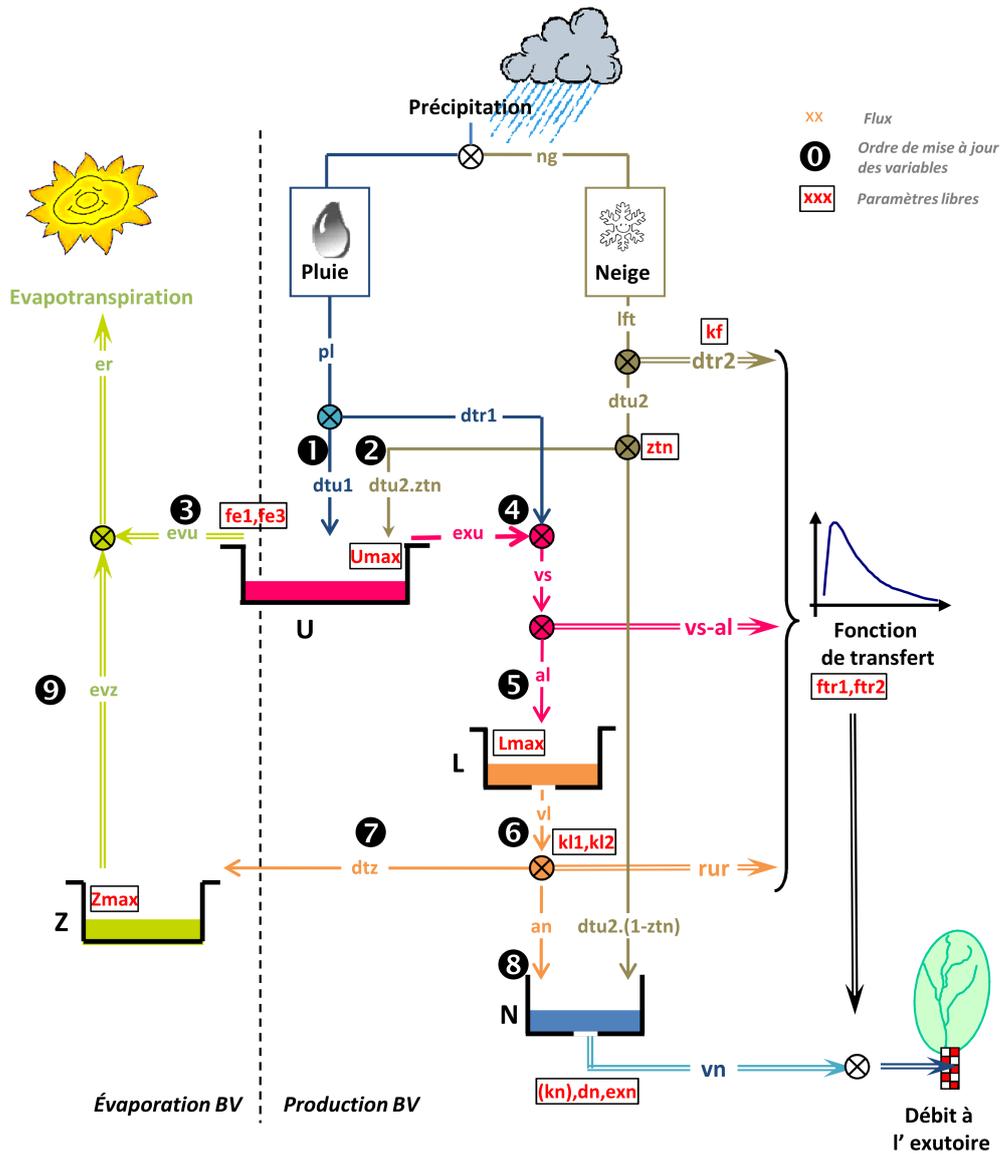


FIGURE I.5 – Schéma conceptuel du modèle MORDOR.

Étape 6 : Simulation stochastique des débits (au pas de temps d'étude)

Cette étape permet de croiser deux aléas à l'aide d'un processus de simulation à la fois hydrologique et stochastique.

1. Une date t pour laquelle une pluie centrale a été observée dans la chronique de pluie spatiale est sélectionnée. La date t appartient à la saison S_i .
2. Une valeur de pluie centrale P_C^* est tirée uniformément entre 1 mm et 1000 mm. La probabilité de P_C^* est donnée par la loi MEWP saisonnière de la saison S_i : $Prob(P_C^* | S_i)$.
3. Les valeurs des ratios $P_{a^-}^*/P_C^*$ et $P_{a^+}^*/P_C^*$ sont tirées uniformément entre 0 et 1. Leurs probabilités conditionnelles respectives sont notées $Prob(P_{a^-}^*/P_C^* | P_C^*)$ et $Prob(P_{a^+}^*/P_C^* | P_{a^-}^* + P_C^*)$ et sont données par les tableaux de contingence des pluies adjacentes.
4. Les trois valeurs de pluies ainsi générées constituent un épisode de pluie $P_{epi}^* = \{P_{a^-}^*; P_C^*; P_{a^+}^*\}$ de probabilité :

$$\begin{aligned}
 Prob(P_{epi}^* | S_i) = & Prob(P_C^* | S_i) \\
 & \cdot Prob(P_{a^-}^*/P_C^* | P_C^*) \\
 & \cdot Prob(P_{a^+}^*/P_C^* | P_{a^-}^* + P_C^*)
 \end{aligned} \tag{I.5}$$

5. L'épisode de pluie P_{epi}^* est substitué à un épisode pluvieux réellement observé P_{epi} pour la date t dans la chronique de pluie spatiale du bassin (figure I.7). La date t est alors caractérisée par un épisode de pluie P_{epi}^* et des conditions hydrologiques du bassin représentées par le modèle MORDOR calé sur l'historique des observations.
6. Le modèle simule alors un événement de débit (au pas de temps de l'étude pdt) associé à cet événement de pluie et aux conditions hydrologiques du bassin à la date t .

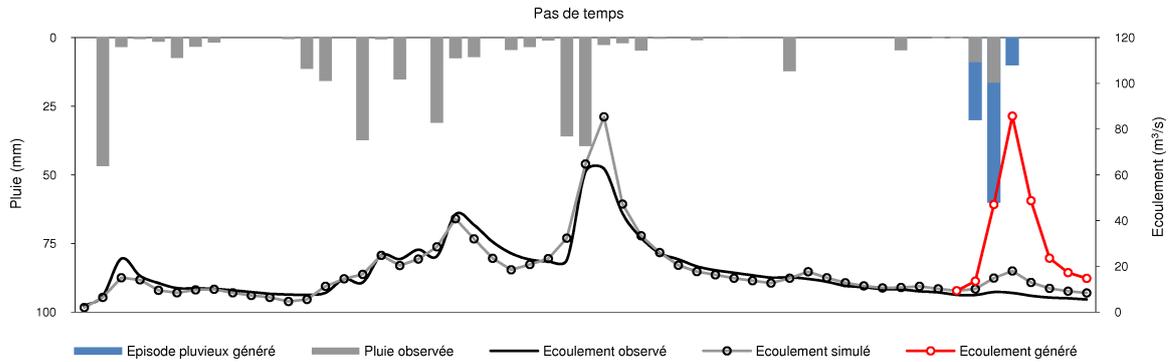


FIGURE I.6 – Schématisation de la simulation d'un écoulement (Garavaglia, 2011b).

Cette opération est répétée un grand nombre de fois (figure I.7) : pour chaque épisode pluvieux observé $P_{epi,i=1,\dots,n}$ dans la chronique, plusieurs centaines de tirages d'événements pluvieux $P_{epi,j=1,\dots,k}^*$ sont effectués et transformés en écoulements $E_{i,j=1,\dots,k}$ avec le modèle. La distribution des débits (au pas de temps d'étude) peut être calculée en triant les $n \cdot k$ débits générés dont la probabilité⁸ est celle des épisodes de pluies (formule (I.5)). Cette distribution est issue d'un croisement exhaustif de l'aléa pluie et l'aléa hydrologique du bassin représenté par le modèle.

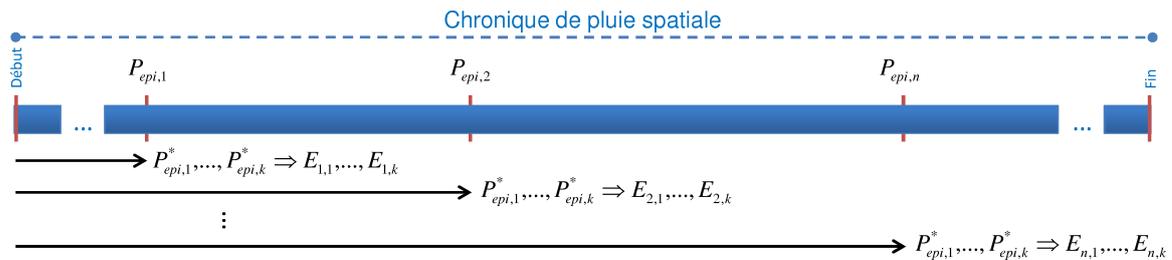


FIGURE I.7 – Procédé de simulation hydrologique de SCHADEX (Garavaglia, 2011b).

Étape 7 : Transformation de la distribution des débits (au pas de temps d'étude) en une distribution des débits de pointe

Lors d'une étude des débits extrêmes sur un bassin versant, une distribution des débits de pointe doit également être estimée. Le coefficient de forme permet de passer de la distribution donnée en sortie de simulation (i.e. la distribution des débits au pas de temps de l'étude pdt) à une distribution des pointes de crues. Actuellement, deux méthodes sont classiquement utilisées pour déterminer ce coefficient de forme.

8. Pour être exact, la probabilité des évènements pluvieux permet de définir une pondération probabiliste pour les évènements d'écoulements.

La méthode du coefficient de forme moyen : le coefficient de forme est une moyenne des ratios du débit de pointe QX et du volume maximum $Q_{pdt}X$ observés sur la sélection d'hydrogrammes (étape 2).

$$K_C = \overline{\left(\frac{QX}{Q_{pdt}X} \right)} \quad (I.6)$$

La distribution des débits de pointe est alors obtenue en multipliant la distribution des débits au pas de temps pdt par cette valeur moyenne de K_C .

La méthode du coefficient de forme variant (Garavaglia, 2011b) : les scénarios d'écoulements générés lors de la simulation SCHADEX (étape 6) ou issus de la sélection d'hydrogrammes observés (étape 2) fournissent un écoulement moyen maximal E_0 et deux écoulements adjacents E_{-1} et E_{+1} sur le pas de temps pdt (figure I.8). Un coefficient de forme volumique peut être défini par :

$$K_V = \frac{3 \cdot E_0}{E_{-1} + E_0 + E_{+1}} \quad (I.7)$$

À l'aide du catalogue d'hydrogrammes sélectionnés (étape 2), les valeurs du coefficient de forme K_C et du coefficient de forme volumique K_V sont estimées pour chaque crue sélectionnée. La dépendance entre K_V et K_C est alors modélisée à travers une relation linéaire :

$$K_C = a(K_V - 1) + 1 \quad (I.8)$$

Ce modèle linéaire $K_V - K_C$ construit sur la sélection d'hydrogrammes, est ensuite utilisé pour estimer un K_C à partir du coefficient K_V de chaque scénario d'écoulement généré par la simulation stochastique de SCHADEX. Le coefficient de forme volumique K_V étant différent pour chaque scénario, la variabilité du coefficient de forme est prise en compte par l'intermédiaire de la variable K_V . La population de débits de pointe alors obtenue permet de construire leur distribution.

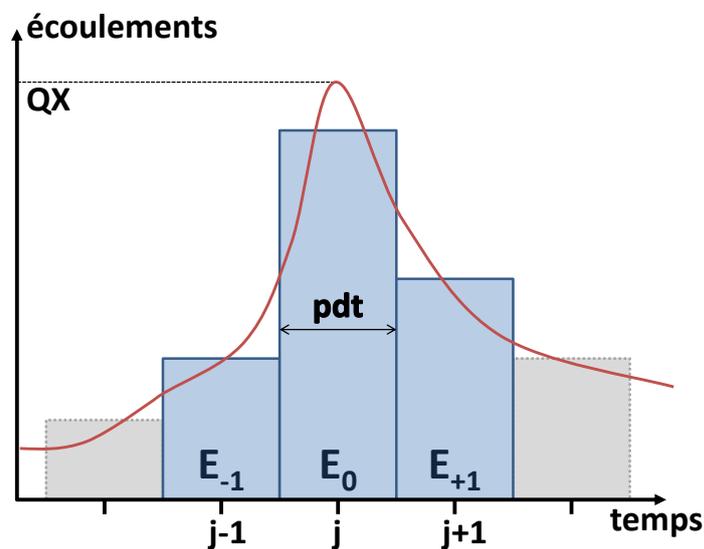


FIGURE I.8 – Schéma d'un scénario d'écoulement d'une simulation SCHADEX. L'hydrogramme rouge n'est connu que pour les hydrogrammes observés de la sélection (mais pas pour les scénarios SCHADEX).

Bilan : la méthode SCHADEX intègre un processus de simulation stochastique semi-continue qui permet de construire une distribution des écoulements en croisant de façon exhaustive l'aléa pluie et l'aléa de saturation du bassin. Le calage du modèle MORDOR, la mise en place des modèles probabilistes pour les pluies, le passage aux débits de pointe utilisent toute l'information disponible (pluie, température, débit, hydrogramme, types de temps) pour obtenir une estimation des débits extrêmes crédible sur le bassin (figure I.9). Cependant, ce fort besoin en données d'entrée pose le problème de son utilisation en contexte non jaugé. Comment caler les modèles probabilistes pour la génération de scénarios de pluies ? Comment estimer une pluie spatiale pour le bassin versant ? Comment passer des pluies aux débits ?

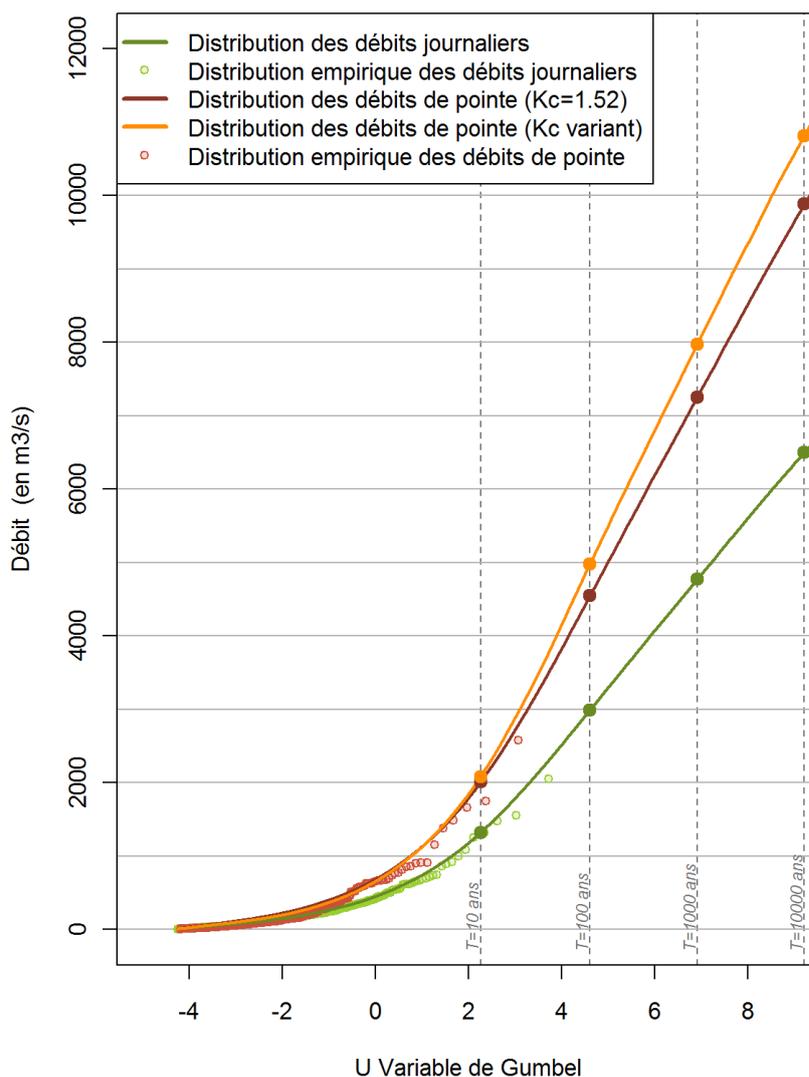


FIGURE I.9 – Exemple de distributions obtenues avec la méthode SCHADEX pour le bassin du Tarn à Millau (2170 km²).

I. 3 Estimation des crues extrêmes sur un bassin non jaugé

Dans la majorité des cas, l'hydrologue n'a pas à disposition toutes les données hydrologiques qu'il souhaiterait pour réaliser son étude. Il est alors compliqué d'ajuster des modèles hydrologiques ou statistiques qui requièrent une disponibilité suffisante des données pour fournir des estimations fiables. Ce problème de manque de données ou de contexte non jaugé peut s'expliquer par plusieurs raisons :

- aucun point de mesure n'existe sur le site (e.g. absence pure et simple de données de pluie ou de débit) ;
- les données existent, mais sont de mauvaise qualité et sont donc inutilisables ;
- les données existent, mais ne sont disponibles que sur des périodes trop courtes qui ne suffisent pas à caler les modèles hydrologiques ou statistiques.

Le problème qui se pose est de savoir comment réaliser une estimation des débits extrêmes dans de telles conditions. Les hydrologues s'appuient en général sur des procédures d'estimation régionales qui permettent de transférer les données hydrologiques disponibles d'autres sites vers le site d'intérêt, où l'estimation hydrologique est requise. Une méthode de régionalisation consiste donc dans un premier temps à définir un groupe de stations considéré comme homogène. Il s'agit de regrouper différents sites jaugés avec un comportement hydrologique ou climatique similaire de façon à décrire le site non jaugé considéré. Il existe différentes techniques pour constituer des regroupements et différents critères pour évaluer l'homogénéité de ces régions (Hosking et Wallis, 1997; Ribatet, 2007).

Une fois qu'une région homogène est choisie pour décrire le site cible considéré, plusieurs approches peuvent être envisagées pour estimer les forts quantiles de débits. Nous allons présenter les grandes familles ainsi que les méthodes les plus représentatives pour chacune d'entre elles.

Les estimations par régressions multiples. Ce sont des formules empiriques généralement de la forme :

$$QX(T) = a \cdot S^b$$

où $QX(T)$ est le débit de pointe (en m^3/s) relatif à la période de retour T , S est la superficie du bassin versant (en km^2), et a et b des coefficients fonction du temps de retour T .

Ces formulations découlent de synthèses régionales permettant d'exprimer a et b selon différentes grandeurs. La valeur du coefficient a est liée au régime pluviométrique et à l'aptitude au ruissellement du bassin. Le coefficient b est un exposant généralement compris entre 0.7 et 0.8. Ce type d'approche est souvent très schématique et imprécis, mais a l'avantage de fournir facilement un premier ordre de grandeur.

Il existe de nombreuses méthodes d'estimation s'appuyant sur des régressions multiples telles que :

- Les courbes enveloppes régionales (Francou et Rodier, 1967) regroupant les maximum observés sur différents bassins, sans homogénéité des temps de retour T .
- La synthèse des débits de période de retour 1000 ans calculés par la méthode du GRADEX (Duband *et al.*, 1988)
- La formule Crupedix (CTGREF *et al.*, 1980-1982) pour le temps de retour $T = 10$ ans.
- La formule Socose (CTGREF *et al.*, 1980-1982) pour les temps de retour de $T = 2$ ans à $T = 50$ ans.

La méthode de l'indice de crue. Cette méthode considère que tous les sites d'une même région homogène ont des distributions de débits identiques à un facteur d'échelle près. Les échantillons de débits de chaque station du voisinage sont normalisés par leur valeur moyenne. Cette valeur moyenne est appelée indice de crue et est propre à chaque station. Le principe de cette approche régionale est de considérer que tous les échantillons normalisés ont une fonction de répartition similaire et représentative de celle du site cible. Une distribution régionale est alors construite à partir des échantillons de crue normalisés aux stations du voisinage homogène défini (Dalrymple, 1960; Hosking et Wallis, 1997). L'estimation de la distribution de débits au site cible est alors obtenue en multipliant la distribution régionale issue des échantillons normalisés par l'indice de crue de ce site. La valeur de l'indice au site cible peut être calculée par des régressions multiples ou des méthodes géostatistiques (krigeage, top-krigeage).

Les modèles pluie-débit simplifiés. Ces méthodes sont généralement dédiées à l'estimation du débit de pointe de crue d'une fréquence donnée. Elles s'appuient sur une hypothèse forte qui est que la probabilité du débit est identique à la probabilité de la pluie (hypothèse rarement vérifiée en hydrologie rurale). La modélisation pluie-débit permet alors de générer un volume écoulé puis une forme d'hydrogramme de crue de projet. Ce passage pluie-débit repose sur la définition d'une fonction de production et d'une fonction de transfert. La fonction de production détermine la part de la pluie qui s'écoule effectivement à l'exutoire du bassin versant pendant la crue. Le transfert hydraulique des volumes jusqu'à l'exutoire est alors assuré par la fonction de transfert. La formulation de la relation pluie-débit est souvent très simplifiée, avec des paramètres d'écoulement qui peuvent tout de même être difficiles à estimer en l'absence d'information hydro-pluviométrique.

Il existe un grand nombre de méthodes de ce type, parmi lesquelles deux méthodes simples et souvent utilisées en ingénierie :

- La méthode rationnelle (Stifel et Mc Pherson, 1964)
- La méthode SCS (Soil-Conservation-Service, 2003)

Les méthodes de simulation s'appuyant sur un aléa pluviométrique et un aléa hydrologique.

Les estimations de débits extrêmes sont réalisées par simulation en réalisant un croisement entre l'aléa pluie et un état initial de saturation du bassin. Un exemple de ce type est la méthode SHYREG. C'est une application régionale de la méthode SHYPRE, qui associe un générateur de pluies horaires et une approche régionale de la modélisation pluie-débit. SHYREG aboutit à une base de données à la résolution du km² avec en tout point les distributions de pluies et de débits sur différentes durées. La fonction de transfert est fixée et les paramètres du modèle hydrologique (modèle GR d'Irstea à deux réservoirs) utilisés sont soit régionalisés soit fixés (Aubert, 2012).

La méthode SCHADEX, dont le principe est détaillé en section I. 2, permet une estimation de la distribution des débits extrêmes en site jaugé. Cependant, son utilisation en contexte non jaugé reste problématique. Il est alors logique de s'intéresser à une approche régionale qui permettrait d'accéder à une estimation des forts quantiles de débit (débits rares ou extrêmes), en l'absence d'information sur un bassin versant, tout en gardant les concepts clés utilisés dans la méthode SCHADEX.

I. 4 La problématique et la méthodologie pour l'estimation des événements extrêmes en contexte non jaugé

La méthode SCHADEX permet d'estimer les crues extrêmes et propose une approche très détaillée des aléas hydrologique et climatologique en intégrant les principales observations : situation météorologique, pluie, température, débit, hydrogrammes. Elle a fait l'objet de nombreuses validations : thèses de Garavaglia (2011b) et de Brigode (2013), d'articles Paquet *et al.* (2006, 2013) et elle a été testée lors de projets d'inter-comparaison comme ExtraFlo⁹ ou Flood-Freq¹⁰. Le projet ANR ExtraFlo a montré que les méthodes de simulation stochastique s'appuyant sur l'aléa pluie (comme SCHADEX ou SHYPRE) sont des méthodes performantes pour l'estimation des débits extrêmes.

Si dans le cas de SHYPRE, le processus a été adapté pour une utilisation régionale avec la méthode SHYREG, ce n'est pas le cas de la méthode SCHADEX. Pour de nombreux bassins, les données nécessaires pour réaliser un calcul SCHADEX complet ne sont pas toutes disponibles et la méthode doit alors être adaptée pour pouvoir fournir un résultat. Différents cas de figure sont possibles et conduisent, à différents niveaux de la méthode, à réaliser des transpositions de données, de paramètres ou de résultats afin de proposer une valeur de crue extrême « raisonnable » et spatialement cohérente pour le dimensionnement des ouvrages. L'enjeu est donc de s'intéresser à cette utilisation non optimale de la méthode SCHADEX en sites non jaugés et de développer des solutions permettant de mieux apprécier le risque de crue pour de tels sites.

La problématique scientifique, qui se pose alors, est celle de l'estimation des extrêmes en sites non jaugés. Les développements sur cette thématique sont depuis longtemps un thème de recherche très actif, notamment sur la spatialisation des phénomènes extrêmes, pour répondre à l'importante demande sociétale et institutionnelle de cartographie des aléas naturels. Le foisonnement scientifique autour de cette problématique est très important, avec des développements statistiques parfois complexes permettant d'intégrer des dérives spatiales (comme une évolution du risque avec l'altitude ou l'éloignement à la Méditerranée).

Enjeux et difficultés de la régionalisation de la méthode SCHADEX

Dans l'état actuel des choses, l'approche SCHADEX utilisée pour réaliser des études de crues extrêmes est applicable sur des bassins jaugés. Sur la colonne de gauche de la figure I.10 sont reportées les grandes étapes d'une application de la méthode SCHADEX qui ont été présentées en section I. 2.

En reprenant les grandes étapes de l'application de la méthode SCHADEX, il apparaît clairement que la régionalisation de la méthode peut intervenir à différents stades. La première difficulté est la régionalisation de l'aléa pluie. En contexte non jaugé, la construction d'une pluie spatiale pertinente et l'estimation des forts quantiles de pluies nécessitent d'être étudiées. Dans un second temps, le calage d'un modèle hydrologique est impossible sans données. Il faut donc s'interroger sur un moyen de régionaliser l'étape de passage pluie-débit. Un troisième et dernier point concerne le passage des débits moyens aux débits de pointe avec la problématique de la régionalisation du coefficient de forme.

Ce travail de thèse a pour but de répondre dans la mesure du possible à toutes ces questions. Ce sera également l'occasion de s'intéresser à des problématiques scientifiques annexes telles

9. Projet français de l'Agence Nationale de Recherche dont l'objectif est de procéder à une inter-comparaison des principales méthodes de prédétermination des pluies et crues extrêmes utilisées en France : <https://extraflo.cemagref.fr/>

10. Projet de recherche pour une étude comparative des différentes méthodes d'analyse fréquentielle des crues à l'échelle européenne : <http://www.cost-floodfreq.eu/home>

que la construction de cartes de pluies extrêmes de temps de retour T , l'évaluation de l'apport d'une classification par types de temps pour la régionalisation des précipitations extrêmes ou encore l'identification de grandeurs conditionnant la réponse hydrologique d'un bassin versant. Le paragraphe suivant présente le cheminement méthodologique et les hypothèses envisagées pour répondre à ces questionnements et aux enjeux industriels sous-jacents.

Hypothèses et cheminement méthodologique envisagé

La méthodologie générale envisagée pour fournir une approche régionale pour les études de crues consiste à s'inspirer du schéma de SCHADEX en s'intéressant dans un premier temps à l'aléa pluviométrique puis à l'aléa hydrologique (colonne de droite de la figure I.10).

Un certain nombre d'hypothèses sous-tendent cette méthodologie :

- **Distribution des pluies** : La distribution utilisée pour les ajustements de pluies est une distribution MEWP (Garavaglia *et al.*, 2010). C'est une loi composite avec des lois marginales par types de temps et par saison. Les détails de cette distribution seront abordés dans le chapitre suivant.
- **Classification par types de temps** : La classification par type de circulation atmosphérique¹¹ (ou type de temps), utilisée par la suite, est celle mise en place et utilisée à EDF-DTG. Elle est déjà intégrée dans de nombreux développements, dont la distribution MEWP et l'interpolateur de pluie SPAZM (Gottardi *et al.*, 2012).
- **Le pas de temps d'étude** : Le travail de régionalisation des extrêmes est réalisé pour un pas de temps journalier afin de borner l'étude régionale à un pas de temps pour lequel les données (mesures de débit et de pluie et études SCHADEX) sont les plus souvent disponibles. Il faut garder à l'esprit que ce pas de temps n'est pas adapté pour des petits bassins réactifs. L'objectif est donc de réaliser une étude exploratoire sur ces bassins pour évaluer les incertitudes induites et pour proposer des pistes de développement futur sur ce sujet.

Sous ces hypothèses, le travail s'est structuré en deux grandes parties.

Une première partie concerne la régionalisation des pluies extrêmes et s'intéresse à une régionalisation de la distribution MEWP des pluies journalières en traitant différents sujets comme :

- l'apport des « types de temps » dans la compréhension de la structure spatiale des pluies extrêmes ;
- l'évaluation de l'interpolateur de pluies journalières SPAZM développé par Gottardi (2009) pour l'estimation des pluies extrêmes ;
- la construction de cartes pour les forts quantiles de pluies ;
- l'influence des choix effectués lors de la construction de pluies spatiales sur l'estimation des extrêmes et l'étude des abattements statistiques réalisés sur ces pluies.

L'objectif final de cette première partie est de traiter la problématique de la description de l'aléa pluie en contexte non jaugé en s'intéressant dans un premier temps aux précipitations ponctuelles journalières, puis dans un second temps aux pluies spatiales.

11. Le terme type de temps est un abus de langage puisque les classifications utilisées sont établies sur des champs de pression. Elles fournissent donc des types de circulations atmosphériques.

Une deuxième partie de la thèse s'intéresse à la description du passage pluie-débit classiquement réalisé grâce au modèle hydrologique MORDOR dans la méthode SCHADEX. L'objectif est alors de comprendre quels critères conditionnent les comportements asymptotiques des distributions d'écoulement obtenues avec SCHADEX et de trouver des grandeurs climatologiques, hydrologiques ou morphologiques du bassin déterminant pour le passage d'une distribution de pluies journalières à une distribution d'écoulements journaliers. **Dans cette seconde partie, un processus de simulation stochastique, utilisant un modèle événementiel simplifié au lieu du modèle MORDOR, est proposé. Il permet d'estimer la distribution de débits journaliers tenant compte de l'aléa pluie et de l'aléa de saturation du bassin. Pour terminer, une approche régionale est envisagée pour l'estimation de la distribution des débits de pointe, à l'aide d'un générateur d'hydrogramme synthétique régional.**

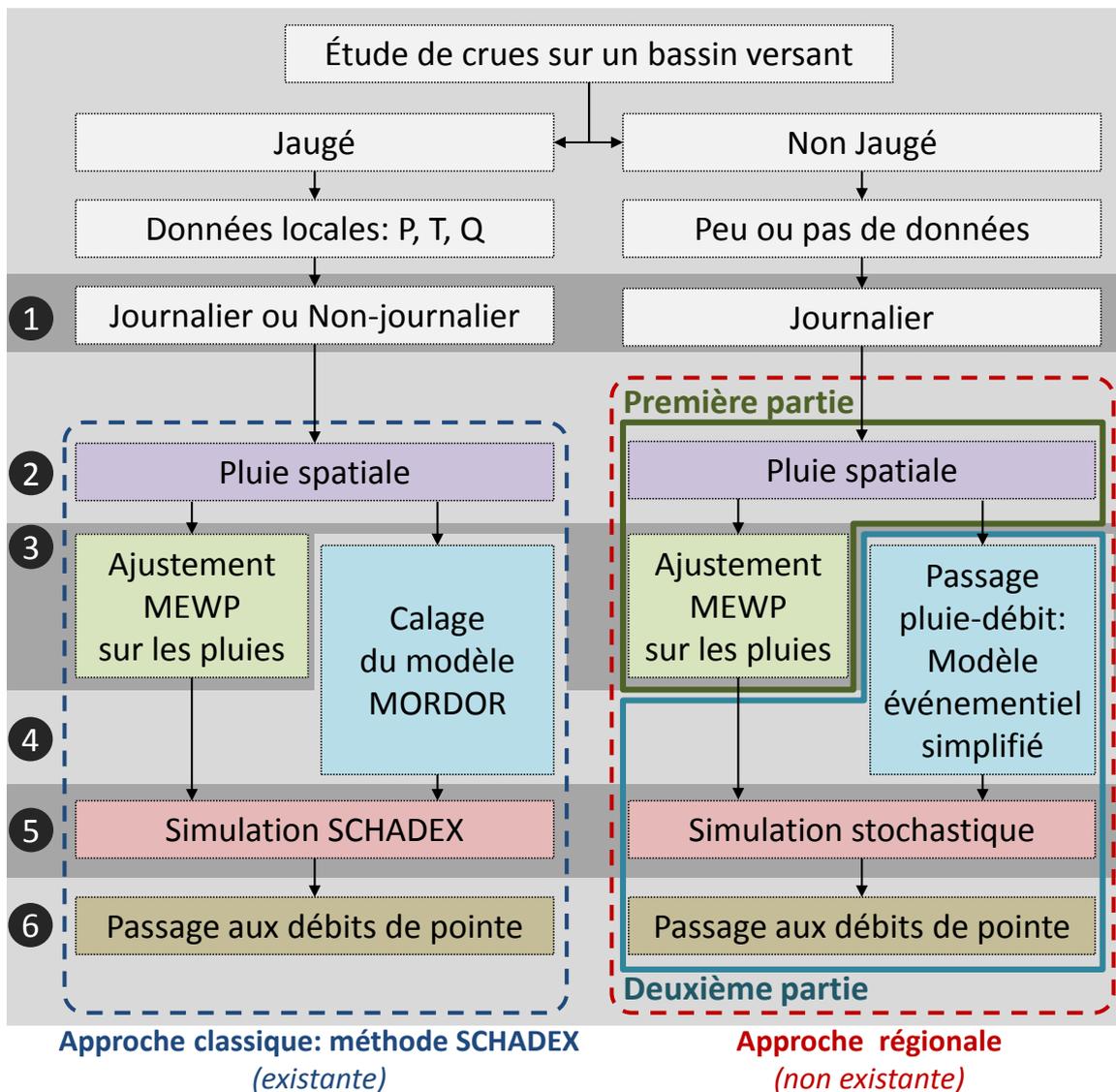


FIGURE I.10 – Schéma du cheminement méthodologique envisagé.

Après un chapitre exposant quelques prérequis nécessaires pour la compréhension des différents résultats, la totalité de ce mémoire se structure autour de ces deux grandes parties :

Première partie : Chapitre III

- Comment estimer les pluies extrêmes ponctuelles ?
- Le concept de type de temps permet-il une meilleure appréhension de la structure spatiale des fortes précipitations ?
- L'interpolateur de pluies journalières SPAZM peut-il être envisagé pour une estimation des précipitations extrêmes ?

Première partie : Chapitre IV

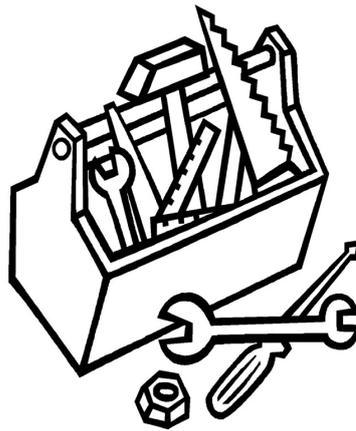
- Comment estimer une pluie spatiale et quels sont les impacts des choix de construction sur l'estimation des fortes précipitations ?

Deuxième partie : Chapitre V

- Comment réaliser un passage pluie-débit et estimer une distribution des débits journaliers prenant en compte l'aléa pluie et l'aléa de saturation du bassin versant ?
- Comment passer d'une distribution de débits journaliers à une distribution de débits de pointe ?

Socle méthodologique requis

Le travail présenté dans ce manuscrit s'appuie en grande partie sur différents développements réalisés à EDF-DTG. Afin de faciliter la compréhension des chapitres suivants, quelques notions vont être présentées dans ce chapitre : la classification par type de temps, l'interpolateur spatial de pluie SPAZM et la distribution probabiliste des pluies MEWP utilisant tous deux cette classification. Une grande partie des résultats présentés par la suite feront référence à des méthodes ou principes présentés dans ce chapitre.



SOMMAIRE DU CHAPITRE

II. 1 La classification par « type de temps » EDF	29
II. 1. 1 Contexte	29
II. 1. 2 Une approche bottom-up pour l'identification des types de temps	29
II. 1. 3 Comparaison des classifications EDF2006 et EDF2010	32
II. 2 L'interpolateur de pluie SPAZM	35
II. 2. 1 La zone de construction des pluies SPAZM	35
II. 2. 2 La base pluviométrique d'entrée	35
II. 2. 3 Les ébauches de pluie par type de temps	36
II. 2. 4 Les pluies journalières	37
II. 3 Les modèles probabilistes des pluies	39
II. 3. 1 La loi d'ajustement MEWP	39
II. 3. 2 Les pluies adjacentes	42

II. 1 La classification par « type de temps » EDF

II. 1. 1 Contexte

La relation entre circulations atmosphériques à large échelle et précipitations a été largement étudiée, surtout en Europe occidentale (Yarnal *et al.*, 2001; Boe et Terray, 2008; Martínez *et al.*, 2008). De nombreux auteurs se sont notamment concentrés sur la région Méditerranéenne. Romero *et al.* (1999); Littmann (2000); Martínez *et al.* (2008) montrent par exemple que l'analyse des situations synoptiques peut apporter d'importantes informations sur la distribution des événements pluviométriques intenses. Il en ressort qu'une classification fondée sur un nombre limité de situations synoptiques pertinentes (ou « types de temps ») est un outil apte à synthétiser l'influence des processus météorologiques sur l'aléa de précipitation.

Cette section décrit brièvement la classification par type de temps utilisée à EDF. Une première version de la classification en type de temps (dénommée par la suite EDF2006) a été établie par Gailhard en 2006 (Paquet *et al.*, 2006). Cette classification en type de temps s'appuie sur une archive de champs géopotentiels fournie par Météo France. Cette archive, calculée sur une grille de 110 points, était initialement interpolée à partir d'un réseau de 37 radiosondages. Pour des questions pratiques (i.e. valeurs des champs géopotentiels plus facilement actualisables) et surtout d'homogénéité (i.e. valeurs des champs géopotentiels plus homogènes sur la période 1948-2010) la classification en type de temps a été recalculée en s'appuyant sur une archive des champs de géopotentiels fournis par le NCEP¹ (National Centers for Environmental Prediction). Ce calcul a abouti à une deuxième version de la classification dénommée EDF2010. Il est important de souligner que les deux classifications, EDF2006 et EDF2010, sont calculées en utilisant la même technique de classification – c'est-à-dire avec la même base de données pluviométrique et la même stratégie de classification – et que la seule différence est dans l'archive des champs de géopotentiels utilisée.

La classification EDF a déjà montré son intérêt sur plusieurs problématiques et en particulier pour les ajustements probabilistes des pluies extrêmes. Une séparation de la population de pluies selon le type de temps de la journée permet d'obtenir des sous-populations moins hétérogènes et respectant mieux l'hypothèse i.i.d. (variables indépendantes et identiquement distribuées) pour l'utilisation de méthodes probabilistes.

II. 1. 2 Une approche bottom-up pour l'identification des types de temps

La classification par type de temps EDF permet de représenter des situations synoptiques journalières sur la zone d'intérêt pour EDF (i.e. fenêtre centrée sur les bassins versants à enjeux pour EDF dans les Pyrénées, le Massif central et les Alpes). L'approche utilisée pour construire cette classification a également été appliquée à la Corse, la Norvège (Lawrence *et al.*, 2013), l'Autriche et la Colombie Britannique (Brigode, 2013). Pour la construction des classifications EDF2006 et EDF2010, une base de données de champs de géopotentiels², ayant les caractéristiques suivantes, a été considérée :

- Analyse centrée sur le sud-est de la France (fenêtre d'analyse pour EDF2006 : grille bleue de la figure II.1 ; fenêtre d'analyse pour EDF2010 : grille verte de la figure II.1).
- Champs de géopotentiels à 700 et 1000 hPa, à 0 h et 24 h. Pour la classification EDF2006, les valeurs des champs sont disponibles depuis le 1^{er} janvier 1953 et sont mesurées

1. <http://www.ncep.noaa.gov/>

2. Un champ de géopotential à une pression de x hPa est le champ indiquant en chaque point l'altitude pour laquelle la pression atmosphérique vaut x hPa.

sur 110 points (points bleus) par une interpolation à partir de 37 radiosondages (points rouges). Pour la classification EDF2010, les valeurs sont mesurées en 56 points (points verts de la fenêtre verte) et sont issues du modèle NCEP.

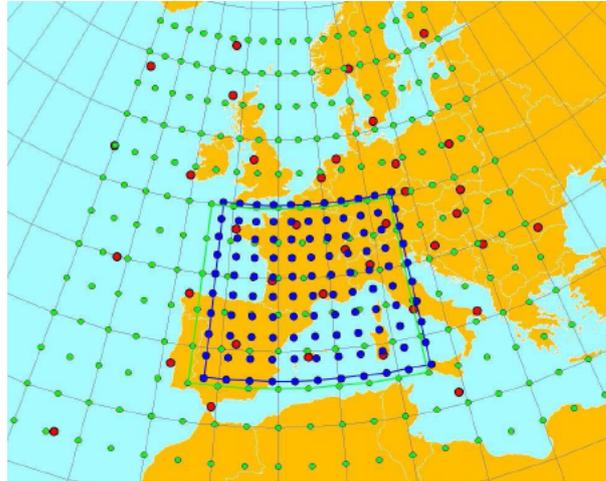


FIGURE II.1 – Localisation des 37 radiosondages (points rouges) et des grilles d'analyse utilisées pour la classification EDF2006 (fenêtre bleue) et pour la classification EDF2010 (fenêtre verte).

De cette façon, chaque jour peut être défini par sa position dans un espace mathématique à 440 dimensions (\mathfrak{R}^{440}). Les 440 dimensions correspondent à la projection des 4 champs géopotentiels calculés sur les 110 points de la grille considérée. Le processus de classification suit une approche « bottom-up ». Les journées pluvieuses sont classifiées en utilisant l'information pluviométrique, puis ces classes sont projetées dans l'espace des champs géopotentiels (\mathfrak{R}^{440}), afin de calculer des centres de gravité par classe. Chaque journée est ensuite projetée dans ce même espace et classée en fonction de sa proximité avec un des centres de gravité identifiés. La méthodologie de la classification peut être résumée en quatre étapes illustrées par le schéma II.2 :

- **Étape 1** : Pour décrire le champ pluviométrique journalier, 54 chroniques de pluies de bassin (combinaisons linéaires de plusieurs postes de mesure) disponibles sur la période 1956-1996 ont été considérées. Seuls les jours pour lesquels ont été observés au moins 5 mm de pluie, en moyenne, ont été retenus (3084 jours). Cet échantillon définit les « jours pluvieux » et représente 21% des jours disponibles. Pour chaque journée pluvieuse, la pluie journalière pour un bassin donné est ensuite normalisée par la précipitation moyenne du jour concerné (calculée sur les 54 chroniques de pluies). Cette normalisation permet de considérer, d'une certaine façon, la « forme » du champ pluviométrique
- **Étape 2** : Une classification ascendante hiérarchique (CAH) est ensuite réalisée sur la population des « formes » du champ pluviométrique des journées pluvieuses qui est définie dans un espace mathématique à 54 dimensions (\mathfrak{R}^{54}). La CAH a montré que le choix de sept classes pluvieuses est raisonnable. Une fois les sept classes définies, les journées restantes (79% de la totalité des jours) peuvent être rangées dans une huitième classe peu pluvieuse.
- **Étape 3** : Les centres de gravité (centroïdes) des huit classes sont calculés dans l'espace de champs géopotentiels (\mathfrak{R}^{440}).
- **Étape 4** : Il est dès lors possible d'attribuer à chaque journée de l'historique la classe, parmi les huit proposées, pour laquelle la journée en question et le centre de gravité de la

classe sont les plus proches. La distance, dans l'espace \mathcal{R}^{440} , entre la journée à classifier et le centre de gravité d'une classe est calculée à l'aide du score de Teweles-Wobus (Teweles et Wobus, 1954). Ce critère a été choisi parce qu'il a été optimisé pour ce type de calcul et parce qu'il a déjà été employé dans la méthode de prévision des précipitations par analogues (Guilbaud, 1997; Obled *et al.*, 2002; Bontron, 2004; Ben Daoud, 2010).

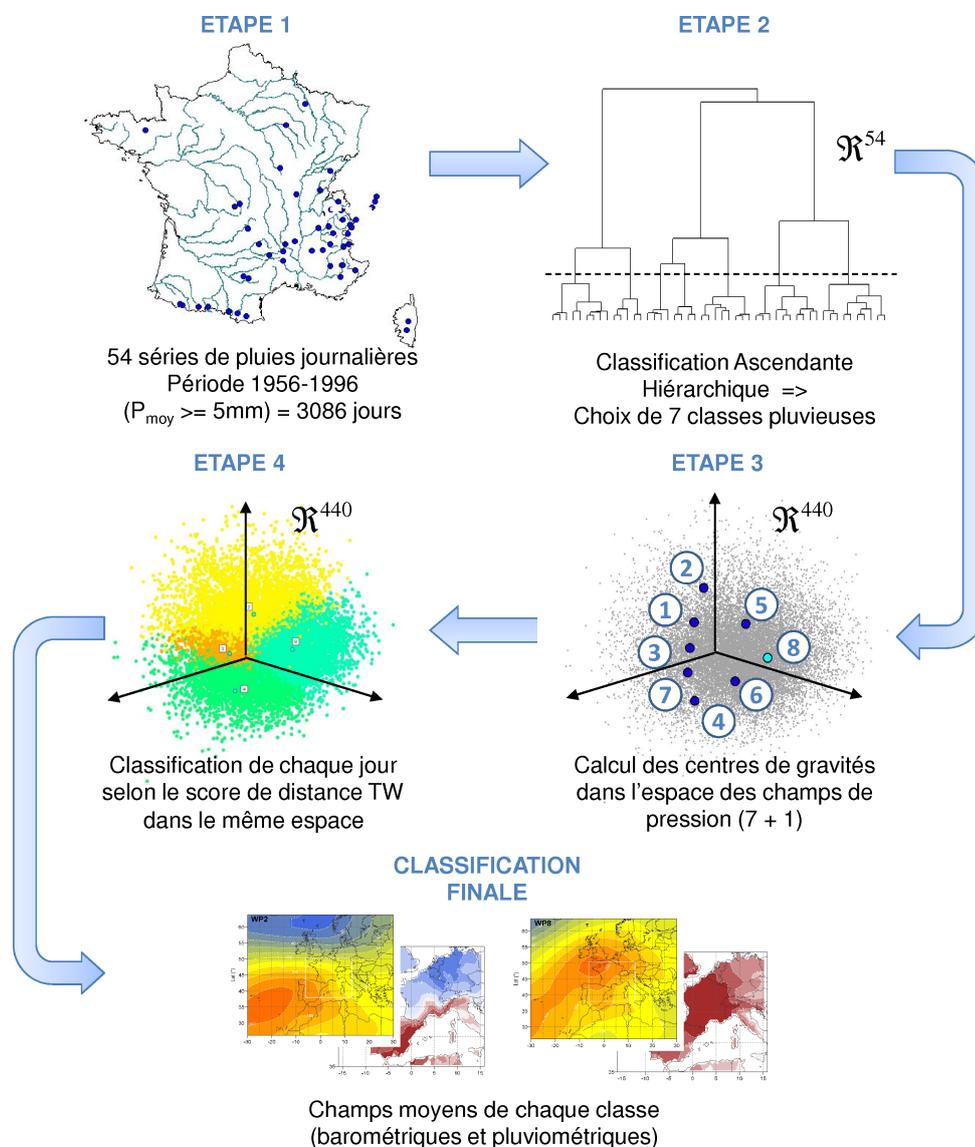


FIGURE II.2 – Méthodologie pour la construction de la classification en types de temps (Garavaglia, 2011b).

Les huit types de temps ainsi obtenus sont illustrés en figure II.3 à travers leurs champs géopotentiels moyens à 1000 hPa à 0h. Pour une raison pédagogique, ces champs barométriques moyens ne sont pas présentés dans un ordre numérique, mais en terme de direction de circulation atmosphérique : 2, 1, 3, 7, 4, 6, 5 et 8 (voir figure II.2 - étape 3). Pour chaque type de temps (sauf pour le type de temps 8), une flèche indique le flux atmosphérique de basse couche induit par la situation synoptique moyenne.

La figure II.4 illustre les champs de précipitation de chaque type de temps sur la partie ouest de l'Europe sous la forme de ratio entre la moyenne de la précipitation des jours du type de temps i et la précipitation moyenne de tous les jours. La base de données de précipitation journalière, définie sur une grille spatiale de résolution $0.5^\circ \times 0.5^\circ$ et sur la période 1953-2005, fournie par le ECA&D (European Climate Assessment and Data - Haylock *et al.* (2008)), a été

utilisée pour calculer ces ratios. Cette figure donne une image régionale du caractère discriminant des huit types de temps.

Les types de temps sont dénommés en relation à la circulation atmosphérique associée (tableau II.1). Le TT2 (Flux atlantique stationnaire), le TT1 (Onde atlantique) et le TT3 (Circulation de Sud-Ouest) correspondent à des circulations océaniques d'Ouest, dont le TT1 est le plus pluvieux par rapport à la région considérée. Le TT7 (Dépression centrale) et le TT4 (Circulation de Sud) correspondent à des circulations favorisant les perturbations méditerranéennes et ils concernent surtout le sud-est de la France où ils amènent des épisodes pluvieux particulièrement intenses. Le TT6 (Retour d'Est) correspond aussi à une circulation méditerranéenne (parfois dénommée « lombarde » dans les Alpes), mais généralement limitée à la frontière franco-italienne et le Roussillon. Le TT5 (Flux de Nord-Est) représente une circulation continentale. Pour terminer, le TT8 (Situation anticyclonique) caractérise des situations de hautes pressions ou, parfois, de « marais barométrique ». Il regroupe les journées globalement peu pluvieuses à l'échelle de la fenêtre d'étude .

TT	Dénomination	Occurrences EDF2006	Occurrences EDF2010
1	Onde Atlantique	7.5%	7.8%
2	Flux atlantique stationnaire	23.6%	21.5%
3	Circulation de Sud-Ouest	7.7%	7.7%
4	Circulation de Sud	17.6%	15.1%
5	Flux de Nord-Est	6.6%	8.5%
6	Retour d'Est	6.1%	6.7%
7	Dépression centrale	3.4%	3.7%
8	Situation anticyclonique	27.5%	29.0%

Tableau II.1 – Les huit types de temps de la classification EDF et leurs occurrences pour chacune des deux classifications.

II. 1. 3 Comparaison des classifications EDF2006 et EDF2010

Dans la suite de ce manuscrit, nous serons amenés à utiliser les deux classifications par types de temps EDF2006 et EDF2010. Tous les travaux récents ont été développés avec la classification EDF2010. Cependant, certains résultats ou développements que nous utiliserons sont plus anciens et se sont appuyés sur la classification EDF2006. Le travail de Garavaglia (2011a) a permis de comparer l'impact de cette montée en version de la classification en types de temps dans la description de l'aléa pluviométrique pour la prédétermination des crues. La conclusion de ce rapport est que les deux classifications sont globalement très proches en terme d'estimation du risque pluviométrique. Ce constat nous permettra de comparer des méthodes utilisant l'une ou l'autre des classifications en sachant que la classification considérée impactera peu l'estimation des pluies extrêmes.

Bilan : la classification par types de temps permet d'associer à chaque journée une information sur la situation météorologique. Elle répartit les précipitations en huit sous-populations a priori plus homogènes du point de vue de la genèse météorologique. Les huit classes définies représentent les grandes situations atmosphériques observables sur la France.

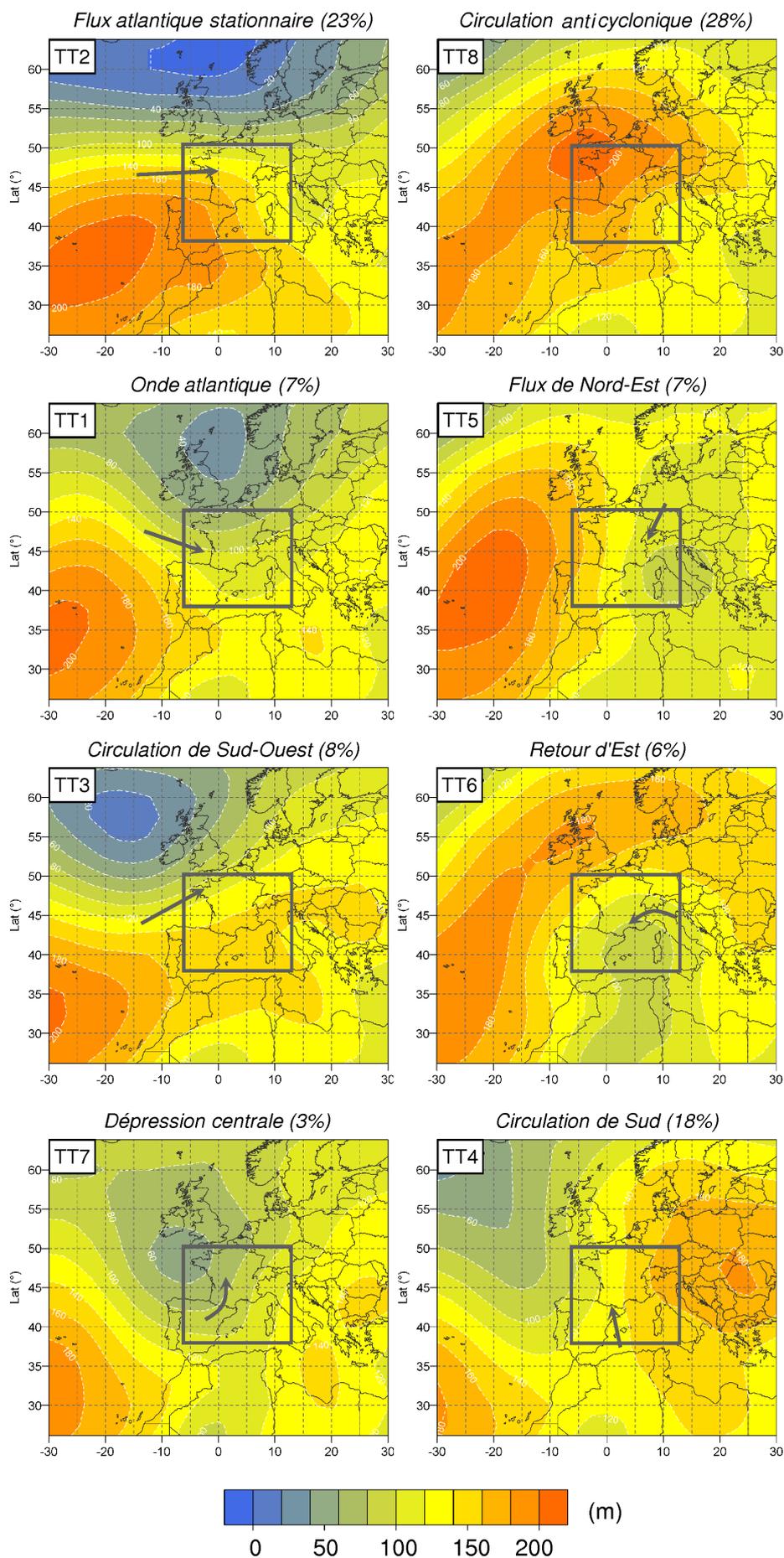


FIGURE II.3 – Champs géopotentiels moyens (1000hPa-0h) de la classification en types de temps EDF2006. Le cadre montre la fenêtre d'analyse (de 6.2°Ouest à 12.9°Est et de 38.0°Nord à 50.3°Nord) et la flèche indique le flux atmosphérique de basse couche (Mathevet et Garçon, 2010; Garavaglia, 2011b).

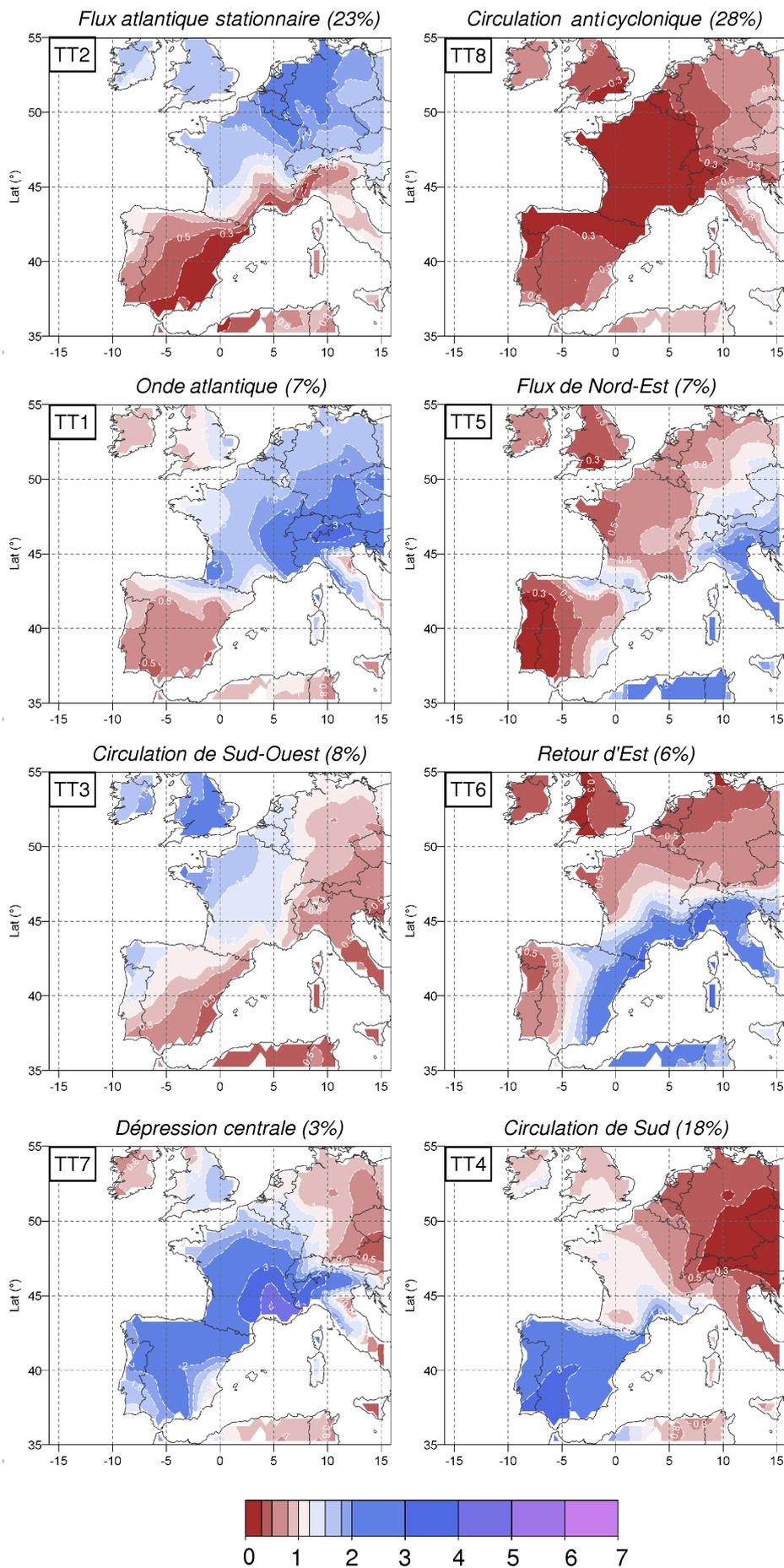


FIGURE II.4 – Ratio entre la précipitation moyenne de chaque type de temps et la précipitation moyenne globale (tous types de temps confondus). Les zones en bleu sont celles où le type de temps considéré apporte des pluies supérieures à la moyenne globale (Garavaglia, 2011b).

II. 2 L'interpolateur de pluie SPAZM

La classification par type de temps a notamment été utilisée afin de régionaliser les précipitations journalières. La méthode SPAZM est un interpolateur de pluies journalières développé par Gottardi (Gottardi, 2009; Gottardi *et al.*, 2012) et qui fournit une estimation spatiale des précipitations en montagne à travers une approche par type de temps. Le modèle d'interpolation utilisée pour les pluies est une évolution de celui de l'approche PRISM (Daly *et al.*, 1994, 2008). Il s'appuie sur l'utilisation d'un modèle local entre les précipitations et le relief par type de temps. Cet interpolateur de pluies journalières sera souvent mentionné dans la suite de ce mémoire et va donc être détaillé dans la section qui suit.

II. 2. 1 La zone de construction des pluies SPAZM

Les pluies SPAZM sont calculées à une résolution de un kilomètre. L'altitude des pixel de 1 km² est issue d'un modèle numérique de terrain (MNT) de même résolution.



FIGURE II.5 – Zone d'étude de SPAZM (Gottardi, 2009).

Les enjeux hydroélectriques d'EDF sont essentiellement concentrés sur les massifs montagneux français dans les Alpes, les Pyrénées et le Massif Central. C'est sur ces massifs que le modèle de pluie SPAZM a été développé. Elle fournit des résultats sur un domaine plus vaste centré sur le sud-est de la France (contour violet sur la figure II.5), afin de limiter les « effets de bord ».

II. 2. 2 La base pluviométrique d'entrée

La construction de la pluie SPAZM s'appuie sur 2201 pluviomètres d'EDF-DTG, Météo France, Météo Suisse, l'ARPA Piémont³, l'INM⁴. Un travail de détection des valeurs aberrantes et des décalages temporels ainsi qu'une détection et correction des ruptures temporelles a déjà été réalisé sur toute la base dans le cadre de la thèse de Gottardi (2009).

3. Agenzia Regionale per la Protezione dell'Ambiente del Piemonte

4. Instituto Nacional de Meteorologia (Météo Espagnole)

II. 2. 3 Les ébauches de pluie par type de temps

Les ébauches sont des cartes des pluies annuelles moyennes par type de temps calculées sur tout le domaine de SPAZM à la maille 1 km². Les huit ébauches, une par type de temps, ont été établies sur la période 1957-1973, période pour laquelle les NPT (Nivo-Pluviomètres Totalisateurs) sont disponibles et fournissent des mesures en altitude.

Les pluies journalières moyennes estimées PE pour un type de temps TT_k donné et un pixel (i, j) donné sont déterminées de la façon suivante :

$$PE_{(TT_k)ij} = a_{(TT_k)ij} \cdot Z_{ij} + b_{(TT_k)ij} \quad (\text{II.1})$$

avec Z_{ij} l'altitude au pixel (i, j) . Les termes $a_{(TT_k)ij}$ et $b_{(TT_k)ij}$ sont déterminés à l'aide d'une sélection judicieuse des précipitations observées PO_{TT_k} se trouvant à proximité du pixel (i, j) . Ces coefficients permettent de définir un gradient altimétrique local par type de temps.

La sélection de ces stations est réalisée par l'intermédiaire d'une cloche de pondération. La cloche de pondération permet de sélectionner les N_s stations les plus proches du pixel $p = (i, j)$ considéré et de leur attribuer un poids W_p . Pour estimer les coefficients $a_{(TT_k)ij}$ et $b_{(TT_k)ij}$ au pixel p , le poids est donné par l'expression :

$$W_p(m) = \lambda \cdot \exp \left[- \left(\frac{d_{3D}(p, m)}{d_0} \right)^\alpha \right] \quad (\text{II.2})$$

où :

- λ ajuste la fourchette des valeurs de pondération ;
- α contraint la forme de la décroissance de la pondération ;
- d_0 détermine la distance à partir de laquelle les stations n'ont plus d'influence au pixel p ;
- $d_{3D}(p, m)$ correspond à la distance de franchissement entre la station m et le pixel p .

La distance de franchissement d_{3D} entre p et m s'exprime de la manière suivante :

$$d_{3D}(p, m) = \sum_{\text{pixels}} \sqrt{(\Delta x^2 + \Delta y^2 + ech_Z \cdot \Delta z^2)} \quad (\text{II.3})$$

avec x et y les coordonnées en Lambert étendue et z l'altitude. La distance de franchissement⁵ est différente d'une distance euclidienne (figure II.6). Le facteur d'échelle ech_Z a été fixé à 20 par Gottardi de façon à prendre en compte la difficulté de franchissement des reliefs. Ainsi cette distance discrimine des stations se trouvant de part et d'autre d'une crête.

5. Il est intéressant de noter que la distance de franchissement n'est pas une distance au sens mathématique du terme car la propriété de l'inégalité triangulaire n'est pas vérifiée. Pour trois points A, B et C, la distance entre A et C peut être supérieure à la somme des distances entre A et B et entre B et C.

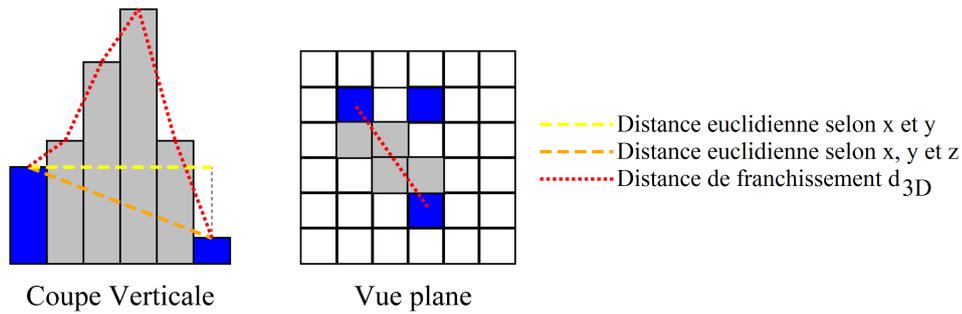


FIGURE II.6 – Illustration des distances euclidiennes et de la distance de franchissement (Gottardi, 2009).

En optimisant de façon à avoir la meilleure interpolation des pluies estimées PE en fonction de l'altitude, les paramètres utilisés pour la construction des ébauches par type de temps sont :

- $N_s = 8$;
- $\lambda = 20$;
- $\alpha = 3$;
- $d_0 = 50\text{km}$.

Remarque : Les ébauches sont construites en utilisant la classification en type de temps EDF2006. Elles n'ont pas été actualisées avec la classification EDF2010, mais les différences seraient minimes.

II. 2. 4 Les pluies journalières

Les précipitations journalières sont calculées sur le même domaine et à la même résolution (1 km^2) que les ébauches. À l'issue du travail de thèse de Gottardi, les pluies journalières étaient disponibles de 1953 à 2005. Un travail complémentaire permet maintenant de disposer de données de pluies journalières de 1948 à 2010. Cependant cette extension de l'interpolation des pluies journalières n'a pas été réalisée sur la même base de données ce qui pose d'éventuels problèmes de raccord dans les séries de pluies journalières sur les différentes périodes. C'est pourquoi nous n'utiliserons que les pluies journalières SPAZM sur la période 1953-2005. Pour chaque journée, l'information suivante est disponible :

- son type de temps à partir de la classification EDF 2006 ;
- l'ébauche associée à ce type de temps ;
- l'ensemble des points de mesure de précipitations journalières (uniquement issu des pluviomètres) avec d'éventuelles lacunes pour le jour considéré.

Pour commencer, l'ébauche PE du type de temps correspondant au jour considéré est déformée afin d'obtenir une première estimation de la pluie journalière \widehat{PJ}_{ij} tombée au pixel (i, j) .

$$\widehat{PJ}_{ij} = \Lambda_{ij} \cdot PE_{ij} \quad (\text{II.4})$$

avec Λ_{ij} le rapport local d'homothétie qui dépend du pixel (i, j) et du jour considéré. Pour travailler de manière locale, les N_p stations les plus proche du pixel (i, j) sont sélectionnées avec la même approche que pour la construction des ébauches. Chacune des N_p stations fournit alors une observations PO_k affectées des poids $W_{i,j}(k)$ (avec $k \in [1, N_p]$). Elles disposent également

des N_p valeurs d'ébauche PE_k associées aux N_p stations. Le rapport homothétique s'exprime alors :

$$\begin{cases} \Lambda_{ij} = \frac{\sum_{k=1}^{N_p} W_{i,j}(k) \cdot PO_k}{\sum_{k=1}^{N_p} W_{i,j}(k) \cdot PE_k} \\ W_{i,j}(k) = \lambda \cdot \exp \left[- \left(\frac{d_{2D}((i,j),k)}{d_0} \right)^\alpha \right] \end{cases} \quad (\text{II.5})$$

Les paramètres déterminés par Gottardi pour l'estimation de pluies journalières sont :

- $N_p = 10$;
- $\lambda = 20$;
- $\alpha = 2$;
- $d_0 = 50 \text{ km}$;
- La distance de sélection est une distance euclidienne d_{2D} (selon les coordonnées x et y) au lieu d'une distance de franchissement comme pour les ébauches.

Un krigeage des résidus ε permet de corriger le champ moyen donné par l'ébauche déformée (en chaque pixel) et d'avoir une estimation exacte au droit de chaque poste. La pluie estimée au pixel (i, j) s'exprime alors par :

$$PJ_{ij}^* = \widehat{PJ}_{ij} + \varepsilon_{ij} = \Lambda_{ij} \cdot PE_{ij} + \varepsilon_{ij} \quad (\text{II.6})$$

À l'issue de ce travail de spatialisation, des précipitations journalières sont disponibles sur tout le domaine d'étude et pour chaque journée de 1953 à 2005.

Bilan : l'interpolateur spatial SPAZM permet une reconstitution des précipitations journalières sur une grande moitié sud de la France et sur la période 1953-2005. SPAZM dispose de plusieurs avantages. Tout d'abord, il peut s'appuyer sur les données journalières du réseau sol avec une forte densité d'information. De plus, il prend en compte un effet orographique local par type de temps (et non un gradient orographique moyen) ce qui lui permet une extrapolation des pluies avec une interaction entre le relief et les flux atmosphériques cohérente. Cet outil de spatialisation des pluies a été développé et validé dans le cadre de la thèse de Gottardi (2009). La cohérence des champs de pluies journaliers et la capacité de SPAZM à représenter les précipitations annuelles pour la réalisation de bilans hydrologiques ont notamment été contrôlées.

II. 3 Les modèles probabilistes des pluies

Dans cette section, nous allons revenir sur la notion d'épisode pluvieux définie lors de la présentation de la méthode SCHADEX dans la section I. 2 (page 10) et en particulier sur les modèles probabilistes développés pour décrire l'aléa des ces épisodes. Les épisodes ont été présentés comme étant composé d'une pluie centrales P_C et de deux pluies adjacentes P_{a-} et P_{a+} (figure II.7). Afin de complètement décrire l'aléa de ce type d'évènements, deux modèles probabilistes ont été mis en place et validés lors de la thèse de Garavaglia (2011b).

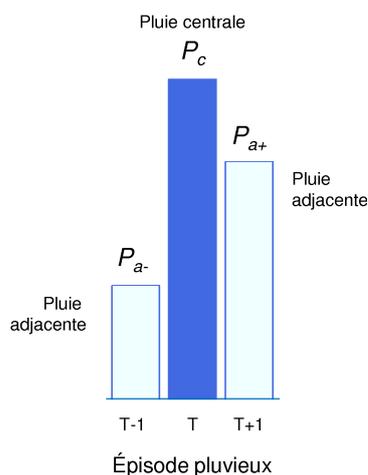


FIGURE II.7 – Schéma d'un épisode de pluie utilisé dans SCHADEX (Garavaglia, 2011b).

II. 3. 1 La loi d'ajustement MEWP

Le modèle probabiliste utilisé pour les pluies centrales est la distribution MEWP qui est détaillée dans l'article de Garavaglia *et al.* (2010). La distribution MEWP ou Multi-Exponential Weather Pattern est une distribution saisonnière. En effet, la population des pluies centrales est répartie en sous-populations selon la saison. Ce sous-échantillonnage par saison est nécessaire afin de considérer des populations de pluies centrales plus homogènes et de mieux représenter les distributions des fortes précipitations (Lang et Desurosne, 1994; Djerboua et Lang, 2007). Pour chaque saison, une loi exponentielle simple est employée comme distribution marginale des pluies centrales au dessus d'un seuil pour chaque type de temps. Les huit lois exponentielles sont alors combinées entre elles en tenant compte de l'occurrence de chaque type de temps afin de revenir à la distribution saisonnière F^i qui s'exprime alors sous la forme suivante :

$$F^i = \sum_{j=1}^8 F_j^i(z_j^i) \cdot p_j^i = \sum_{j=1}^8 \left(1 - \exp\left(-\frac{z_j^i}{\lambda_j^i}\right) \right) \cdot p_j^i \quad (\text{II.7})$$

avec i l'indicateur de la saison considérée et j le type de temps considéré (huit types de temps dans le cas de la classification EDF). Z_j^i est la population des pluies centrales sup-seuil du type de temps j et de la saison i . λ_j^i est le paramètre d'échelle (ou le gradex) de la loi exponentielle et p_j^i le pourcentage d'occurrence relatif de chaque type de temps.

Pour avoir une distribution annuelle, il faut recomposer les ajustements MEWP réalisés sur chaque saison (figure II.8).

Chaque loi exponentielle est ajustée sur un échantillon des pluies sup-seuil du type de temps concerné. Lors d'un ajustement au cas par cas, le seuil peut être ajusté de façon à fournir le

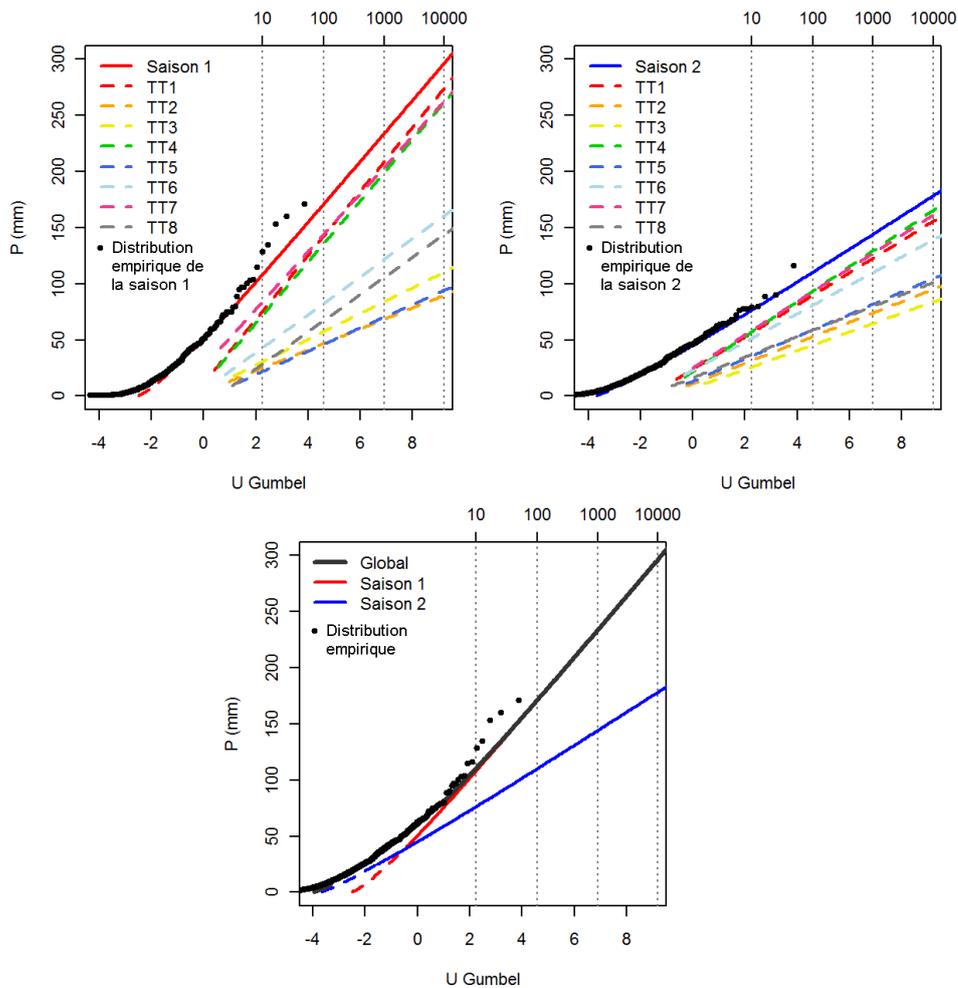


FIGURE II.8 – Loi MEWP sur la station DTG 2834 (Clans Bancairon, Alpes Maritimes). En haut à gauche, l’ajustement sur la saison à risque (de septembre à novembre) avec les huit lois exponentielles des types de temps. En haut à droite, le même ajustement pour la saison complémentaire. En bas, la distribution globale issue de la recombinaison des deux ajustements saisonniers.

meilleur ajustement possible. Pour une approche régionale, le seuil peut être pris égal au quantile 70% de la population considérée. Cela signifie que la loi exponentielle est ajustée sur les 30% des valeurs les plus fortes. La taille de l’échantillon utilisée pour ajuster l’exponentielle peut donc être faible et poser des problèmes de robustesse de l’ajustement. Cependant, ce problème se rencontre pour des sous-populations où le risque de pluies extrêmes est marginale. En dessous de ce seuil, la distribution est la distribution empirique correspondante. Garavaglia (2011b) a montré que cette valeur du seuil permet des ajustements avec un bon compromis justesse et robustesse en s’appuyant sur le large jeu de données présenté en figure II.9. Enfin, dans le cadre de la problématique de régionalisation des pluies extrêmes, des saisons doivent être définies. Garavaglia (2011b) a également cartographié des saisons à risque et a dégagé deux grandes zones (figure II.9) :

- une région A avec une saison à risque d’août à octobre ;
- une région B avec une saison à risque de septembre à novembre.

La saison à risque est la saison pour laquelle la pente asymptotique de la distribution saisonnière – aussi appelé gradex – est la plus forte. Ce découpage montre une certaine cohérence spatiale (figure II.9). Pour arriver à avoir cette nette distribution spatiale, des « compromis » ont été effectués. L’intérêt de cette démarche est d’avoir une répartition qui soit correcte dans l’ensemble. Les stations situées dans la partie ouest et nord-ouest du Massif Central, dans les Vosges, dans

le Morvan et dans la partie est de la Suisse constituent l'archive régionale A. Les stations situées dans les Alpes, Jura, Pyrénées et dans la partie sud-est du Massif central, appartiennent à l'archive régionale B. Il est intéressant de noter une remontée de la région B entre le Morvan et les Vosges, le long de la vallée de la Saône, axe de pénétration des perturbations méditerranéennes. Ce découpage saisonnier a été déterminé sur la base des stations journalières EDF et est donc pertinent pour cette zone et ce pas de temps. Il faudrait généraliser ce travail sur une base plus étendue pour avoir une idée des saisons à risque sur toute la France.

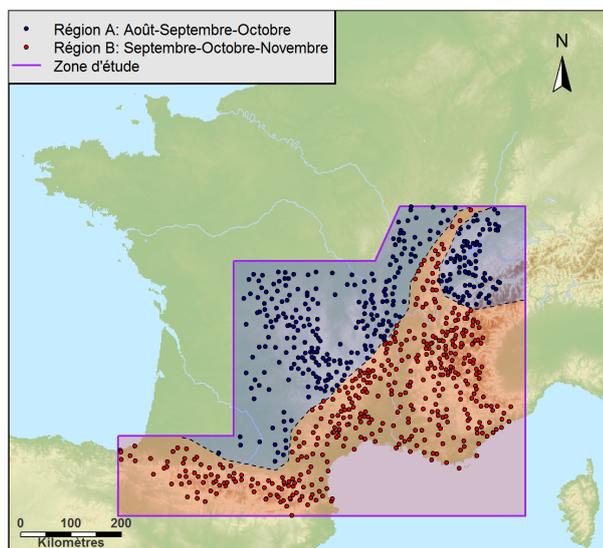


FIGURE II.9 – Répartition spatiale des saisons à risque (Garavaglia, 2011b).

Les figures II.10 (a) et II.10 (b) présentent les boxplots des maxima de pluie sur trois mois glissants respectivement pour l'archive régionale A et B. Comme attendu, les pluies les plus fortes sont observées entre août et octobre (archive régionale A) ou entre septembre et novembre (archive régionale B). Deux saisons seront donc considérées pour les ajustements : une saison de trois mois, dite à risque, et une saison de neuf mois, appelée saison complémentaire.

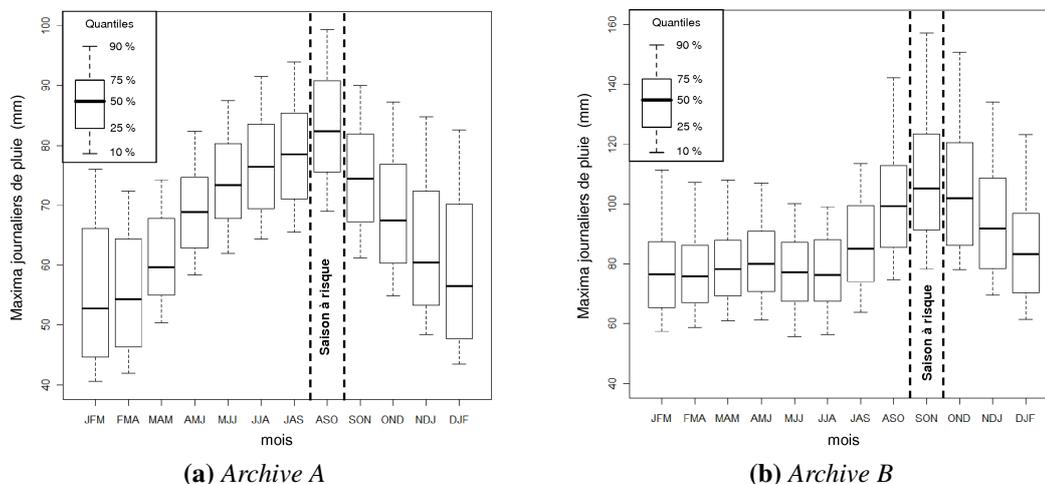


FIGURE II.10 – Boxplots des maxima sur trois mois glissants de l'archive régionale A et B (Garavaglia, 2011b).

II. 3. 2 Les pluies adjacentes

La probabilité des pluies adjacentes est un problème statistique bivarié. Elle est décrite par deux tableaux de contingence qui modélisent les deux probabilités suivantes :

$$\begin{cases} \text{Prob} \left(\frac{P_{a^-}}{P_C} \mid P_C \right) \\ \text{Prob} \left(\frac{P_{a^+}}{P_C} \mid (P_{a^-} + P_C) \right) \end{cases} \quad (\text{II.8})$$

Le principal intérêt des tableaux de contingence est leur simplicité. L'idée est de décrire la probabilité conditionnelle d'une variable aléatoire par une autre variable. Un tableau de contingence permet de décrire les probabilités de P_{a^-}/P_C conditionnées par la valeur de P_C et un autre pour les probabilités de P_{a^+}/P_C conditionnées par la valeur de $P_{a^-} + P_C$.

À titre d'exemple, nous présentons les tableaux de contingence construits pour le bassin de l'Ardèche à Sauze (tableaux II.2). Sur ce bassin, les pluies centrales, issues de la pluie de bassin, sont réparties en quatre classes de façon à ce que chaque classe compte environ deux fois moins d'individus que la précédente. Pour chaque classe de pluie centrale, une probabilité empirique est calculée pour le ratio P_{a^-}/P_C . Le même travail est réalisé pour les ratios P_{a^+}/P_C mais avec cette fois des classes selon la population de $P_{a^-} + P_C$. Les deux tableaux ainsi construits décrivent la probabilité discrète des valeurs P_{a^-}/P_C et P_{a^+}/P_C .

Classe	P_C				Classe	$P_C + P_{a^-}$			
	1	2	3	4		1	2	3	4
Borne Inf. (mm)	1	11	39	76	Borne Inf. (mm)	1	13	49	103
Borne Sup. (mm)	11	39	76	∞	Borne Sup. (mm)	13	49	103	∞
$0 \leq P_{a^-}/P_C < 0.1$	0.4367	0.4219	0.3560	0.2268	$0 \leq P_{a^+}/P_C < 0.1$	0.3965	0.4491	0.4316	0.3579
$0.1 \leq P_{a^-}/P_C < 0.2$	0.1401	0.1641	0.1204	0.1443	$0.1 \leq P_{a^+}/P_C < 0.2$	0.1543	0.1567	0.1789	0.0842
$0.2 \leq P_{a^-}/P_C < 0.3$	0.0789	0.1146	0.1152	0.1546	$0.2 \leq P_{a^+}/P_C < 0.3$	0.1110	0.0992	0.1053	0.0842
$0.3 \leq P_{a^-}/P_C < 0.4$	0.0762	0.0703	0.0576	0.1134	$0.3 \leq P_{a^+}/P_C < 0.4$	0.0677	0.0653	0.0421	0.1158
$0.4 \leq P_{a^-}/P_C < 0.5$	0.0571	0.0286	0.0681	0.0825	$0.4 \leq P_{a^+}/P_C < 0.5$	0.0501	0.0470	0.0579	0.0947
$0.5 \leq P_{a^-}/P_C < 0.6$	0.0449	0.0469	0.0471	0.0928	$0.5 \leq P_{a^+}/P_C < 0.6$	0.0447	0.0366	0.0263	0.0632
$0.6 \leq P_{a^-}/P_C < 0.7$	0.0422	0.0339	0.0681	0.0412	$0.6 \leq P_{a^+}/P_C < 0.7$	0.0460	0.0548	0.0421	0.0526
$0.7 \leq P_{a^-}/P_C < 0.8$	0.0395	0.0443	0.0681	0.0515	$0.7 \leq P_{a^+}/P_C < 0.8$	0.0541	0.0339	0.0368	0.0632
$0.8 \leq P_{a^-}/P_C < 0.9$	0.0272	0.0339	0.0576	0.0515	$0.8 \leq P_{a^+}/P_C < 0.9$	0.0406	0.0157	0.0316	0.0421
$0.9 \leq P_{a^-}/P_C < 1$	0.0571	0.0417	0.0419	0.0412	$0.9 \leq P_{a^+}/P_C < 1$	0.0352	0.0418	0.0474	0.0421

(a) Pluie P_{a^-} du pas de temps précédent (b) Pluie P_{a^+} du pas de temps suivant

Tableau II.2 – Modèle des pluies adjacentes. Tableau de contingence pour le bassin versant de l'Ardèche à Sauze (2260 km²), période 1959-2008 (Garavaglia, 2011b).

Bilan : un épisode pluvieux est finalement caractérisé par :

- la pluie centrale P_C de probabilité $\text{Prob}(P_C)$ donnée par une distribution MEWP ;
- la pluie du pas de temps précédent la pluie centrale P_{a^-} de probabilité $\text{Prob} \left(\frac{P_{a^-}}{P_C} \mid P_C \right)$ donnée par un tableau de contingence ;
- la pluie du pas de temps suivant la pluie centrale P_{a^+} de probabilité $\text{Prob} \left(\frac{P_{a^+}}{P_C} \mid (P_{a^-} + P_C) \right)$ donnée par un tableau de contingence ;
- la probabilité de l'épisode $\text{Prob}_{\text{epi}} = \text{Prob} \left(\frac{P_{a^-}}{P_C} \mid P_C \right) \cdot \text{Prob}(P_C) \cdot \text{Prob} \left(\frac{P_{a^+}}{P_C} \mid (P_{a^-} + P_C) \right)$

Le socle méthodologique ayant été présenté, nous allons pouvoir aborder la première partie de ce travail de thèse : la régionalisation des précipitations extrêmes.

Première partie

Les pluies extrêmes

La régionalisation des pluies extrêmes journalières ponctuelles

Lors d'une application de la méthode SCHADEX, l'objet d'intérêt pour la caractérisation de l'aléa pluviométrique est la pluie de bassin (ou pluie spatiale). Cependant, l'estimation des pluies extrêmes ponctuelles (i.e. en un point précis et non sur la surface d'un bassin versant) – même si elle ne sera pas directement utile pour l'étude des crues extrêmes – comporte plusieurs intérêts. Elle permettra notamment d'évaluer différentes approches pour la régionalisation des pluies extrêmes et de mieux comprendre la structure spatiale des risques de précipitations extrêmes.

Dans ce chapitre, nous allons nous intéresser à l'interpolation spatiale des pluies journalières extrêmes ponctuelle, c'est-à-dire dans notre cas sur une grille de 1 km². Après une présentation du jeu de données utilisé et du principe des différentes méthodes de régionalisation envisagées, nous donnerons les principaux résultats d'une inter-comparaison de ces méthodes et présenterons les cartes de pluies extrêmes produites. Dans un dernier point, nous aborderons certaines évolutions possibles de la distribution de pluie MEWP.



SOMMAIRE DU CHAPITRE

III. 1 Jeux de données et protocole de calage-validation	49
III. 1. 1 Le jeu de données pour la mise en place des méthodes	50
III. 1. 2 Comment valider les estimations régionales d'extrêmes ?	50
III. 2 Les méthodes de régionalisation des pluies extrêmes	54
III. 2. 1 Méthode 1 : Krigeage des quantiles de pluies	54
III. 2. 2 Méthode 2 : Spatialisation de la distribution MEWP par krigeage des gradex . . .	60
III. 2. 3 Méthode 3 : Estimation des pluies extrêmes locales avec les pluies SPAZM . . .	76
III. 3 Comparaisons des approches de régionalisation	83
III. 3. 1 Évaluation sur le jeu de 770 stations	85
III. 3. 2 Évaluation avec le protocole ExtraFlo	91
III. 3. 3 Les cartes de pluies extrêmes proposées	105
III. 4 Vers une estimation des pluies extrêmes à un pas de temps infra-journalier	109
III. 4. 1 Mise en évidence d'une zone « atypique » pour les précipitations extrêmes journalières	109
III. 4. 2 Vers une distribution MEWP pour l'infra-journalier	112
III. 5 Conclusion	114

III. 1 Jeux de données et protocole de calage-validation

De nombreux travaux se sont déjà intéressés à la régionalisation des pluies extrêmes en contexte non jaugé. Szolgay *et al.* (2009) ont réalisé une revue des méthodes couramment envisagées pour cette régionalisation et qui s'appuie sur des outils plus ou moins complexes : géostatistique ou régressions linéaires sur des jeux de variables descriptives appropriées (Kieffer Weiss et Bois, 2000, 2002; Castellarin *et al.*, 2007; Boni *et al.*, 2008; Mailhot *et al.*, 2013), ou utilisation des Max-stable (Blanchet et Lehning, 2010; Gaume *et al.*, 2013), copules (Bardossy et Li, 2008), approches régionales par les L-moments (Hosking et Wallis, 1997) ou encore par approches bayésiennes (Cooley *et al.*, 2007).

Pour simplifier, les méthodes existantes peuvent pour la plupart être réparties au sein de quatre grandes catégories :

1. Des estimations ponctuelles des quantiles extrêmes sont réalisées, à l'aide d'un modèle probabiliste approprié, aux endroits où des observations sont disponibles. Les valeurs des quantiles remarquables (au temps de retour 10, 100 ou 1000 ans) sont spatialisées sur le domaine en utilisant des variables descriptives appropriées (e.g. altitude, distance à la méditerranée...) et des méthodes géostatistiques (krigeage ordinaire...)
2. Une loi de probabilité adaptée est choisie pour le phénomène et le domaine concerné. Ses paramètres sont spatialisés à partir des estimations aux points de mesure, avec des variables descriptives et des méthodes géostatistiques comparables au cas précédent. Par exemple, avec le choix d'une loi GEV (Generalized Extreme Value) pour les pluies, le paramètre de forme est fixé régionalement et les paramètres de position et d'échelle sont expliqués par des variables descriptives (géographique, climatologique...).
3. Un générateur stochastique permet de générer une population quasi exhaustive, statistiquement homologue à une population observée. Les quantiles élevés sont déduits de la population simulée. Les paramètres de la génération sont identifiés grâce aux observations disponibles aux sites jaugés. Les paramètres les plus importants sont alors spatialisés, toujours à l'aide des variables ou des méthodes géostatistiques évoquées pour les deux cas précédents. C'est le principe de la méthode SHYPRE (Arnaud *et al.*, 2007; Cantet et Arnaud, 2014) et de sa version régionale SHYREG (Aubert, 2012; Aubert *et al.*, 2014).
4. Un générateur spatio-temporel permet de générer des champs de pluie par un processus de simulation formalisant les connaissances acquises par l'analyse physique et l'analyse statistique des phénomènes pluviométriques. Il est alors possible de produire des champs de pluies sur tout un domaine d'étude et sur une période choisie (Leblois et Creutin, 2013).

Le but de ce chapitre est de proposer des méthodes pour la régionalisation des pluies extrêmes, s'inspirant du travail déjà réalisé sur ce sujet, mais en essayant d'intégrer différents concepts afin d'améliorer les estimations.

Nous ne nous intéresserons ici qu'aux pluies journalières et nous essaierons notamment de répondre à différentes questions :

- i) Une classification par type de temps permet-il d'améliorer la spatialisation des pluies extrêmes journalières ? La classification permet-elle de mettre en évidence des variables explicatives plus pertinentes pour la régionalisation des pluies extrêmes ?
- ii) Les pluies journalières SPAZM ont fait l'objet d'une validation sur les champs journaliers et sur le cumul annuel des pluies, mais qu'en est-il de leur utilisation pour l'estimation des pluies extrêmes ?

Afin de répondre à ces différentes questions, une base de données a été mise en place, ainsi qu'un protocole de validation qui permettra par la suite de construire et évaluer différentes approches pour construction de cartes de pluies extrêmes.

III. 1. 1 Le jeu de données pour la mise en place des méthodes

Le jeu de données a été construit afin de disposer d'une bonne répartition spatiale et climatique et dispose de séries suffisamment longues pour garantir une bonne robustesse vis-à-vis de l'estimation des extrêmes. Il se compose d'un total de 770 postes pluviométriques au pas de temps journalier issus des 2201 postes de la base de données SPAZM. Il regroupe 689 stations françaises, dont 356 postes de DTG et 333 de Météo France (MFR), 17 postes espagnols de l'INM¹ et 64 stations suisses de Météo Suisse. Les stations espagnoles et suisses permettent de bénéficier d'informations limitant les effets de bords lors de la spatialisation. Les séries de la base sont disponibles sur la période 1953-2005 (ou sur une sous-période).

Les séries ont déjà été critiquées par Gottardi (2009) de façon à détecter et corriger des valeurs ponctuelles aberrantes, des décalages temporels, des ruptures chronologiques. Le problème de ruptures temporelles a été traité par l'étude de stabilité dans le temps des résidus d'un modèle (régression linéaire multiple) reliant les mesures d'une station pluviométrique à celles de stations voisines bien corrélées (Peterson et Easterling, 1994; Gottardi, 2009). Les tests effectués sont basés sur le principe du test d'homogénéité d'Alexandersson (Alexandersson, 1986) et des cumuls de résidus associés aux ellipses de Bois (Bois, 1971; Paquet, 2005).

La détection de tendances dans les séries pluviométriques a été faite avec un test non paramétrique ne nécessitant aucune hypothèse sur la distribution des données (Hamed, 2009). Renard (2006) montre, pour la détection d'une tendance dans des données non auto-corrélées et où la distribution est inconnue, que les tests de Mann-Kendall (Mann, 1945; Kendall, 1975) et de rho Spearman (Lehmann, 1975; Sneyers, 1990) sont parmi les plus performants. Le test utilisé est celui de Mann-Kendall.

De plus, ces séries pluviométriques vérifient différents critères dans le but de pouvoir réaliser des ajustements de pluie fiables et robustes :

- elles possèdent au moins 30 ans de données d'années équivalentes (i.e. sans les données manquantes) ;
- toutes les séries sont au moins disponibles jusqu'en 2005 ;
- les données manquantes représentent moins de 10% des données totales.

Après suppression des doublons Météo France et DTG, il reste finalement 770 stations sur les 2201 critiquées.

Sur la figure III.1, les 770 stations couvrent bien la zone d'étude et il n'y a pas de grandes régions sans poste ce qui pourrait poser des problèmes pour la spatialisation des phénomènes et l'évaluation des méthodes.

III. 1. 2 Comment valider les estimations régionales d'extrêmes ?

La façon dont les méthodes vont être évaluées est très importante. En effet, avec un réseau de pluviomètres très dense, les différentes approches de régionalisation auront toutes des performances assez élevées, car le réseau va contraindre la forme du champ, quelle que soit la méthode utilisée. Réaliser une validation en Jackknife (Tukey, 1958; Quenouille, 1949) ne permettra donc pas de détecter une différence flagrante entre les méthodes puisqu'une seule station est enlevée du réseau à chaque évaluation. **Cette remarque est d'autant plus importante**

1. Instituto Nacional de Meteorologia (Météo Espagnole)

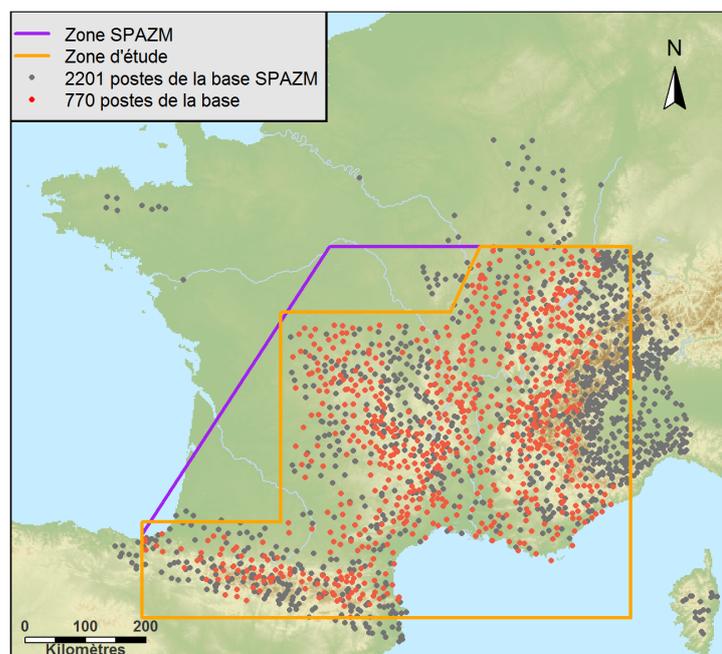


FIGURE III.1 – Réseau de 770 postes pluviométriques

que nous souhaitons évaluer les méthodes en vue d'une utilisation en contexte non jaugé, donc à faible densité de réseau. L'idée est de construire des réseaux de calage et de validation permettant d'évaluer les méthodes dans cette optique.

C'est pourquoi une méthode de validation tenant compte de l'influence de la densité du réseau pluviométrique a été mise en place. Elle consiste à rendre le réseau de moins en moins dense en calage en enlevant les stations les plus proches les unes des autres (figure III.2). Toutes les stations enlevées sont utilisées pour la validation. Une méthode peut alors être comparée par rapport à une autre et évaluée à différentes densités de stations utilisée pour le calage.

Dans notre cas, huit sous-réseaux plus ou moins denses issus du réseau complet constitué des 770 stations de la base de données seront considérés (tableau III.1). Les différents réseaux de validations ont des distributions d'altitude des stations similaires (figure III.3). Nous remarquons cependant que les réseaux utilisés en validation sont peu fournis en stations de plaines (de 0 à 400 mètres). Ceci est dû au fait que c'est dans ces régions que le réseau est le moins dense et que, de par le principe même de construction des réseaux de validation par dédensification, ces stations sont gardées dans les réseaux de calage plus longtemps. Au contraire, la tranche de 400 à 600 mètres est très vite bien représentée dans les réseaux de validation, car elle correspond à des zones disposant d'une bonne densité d'information. Le léger manque de stations pour la validation en plaine n'est pas réellement problématique, car ce n'est pas dans ces zones que se posent les vrais problèmes de spatialisation (la variabilité spatiale étant souvent moins forte, la spatialisation est en générale plus juste). Cette vérification permet de s'assurer que la qualité du jeu de validation reste stable et que l'on a donc des scores comparables d'un réseau de validation à un autre. De plus, les stations utilisées en validation sont réparties sur toute la zone d'étude, ce qui évite de surreprésenter une région dans les scores au risque de fausser l'interprétation de résultats.

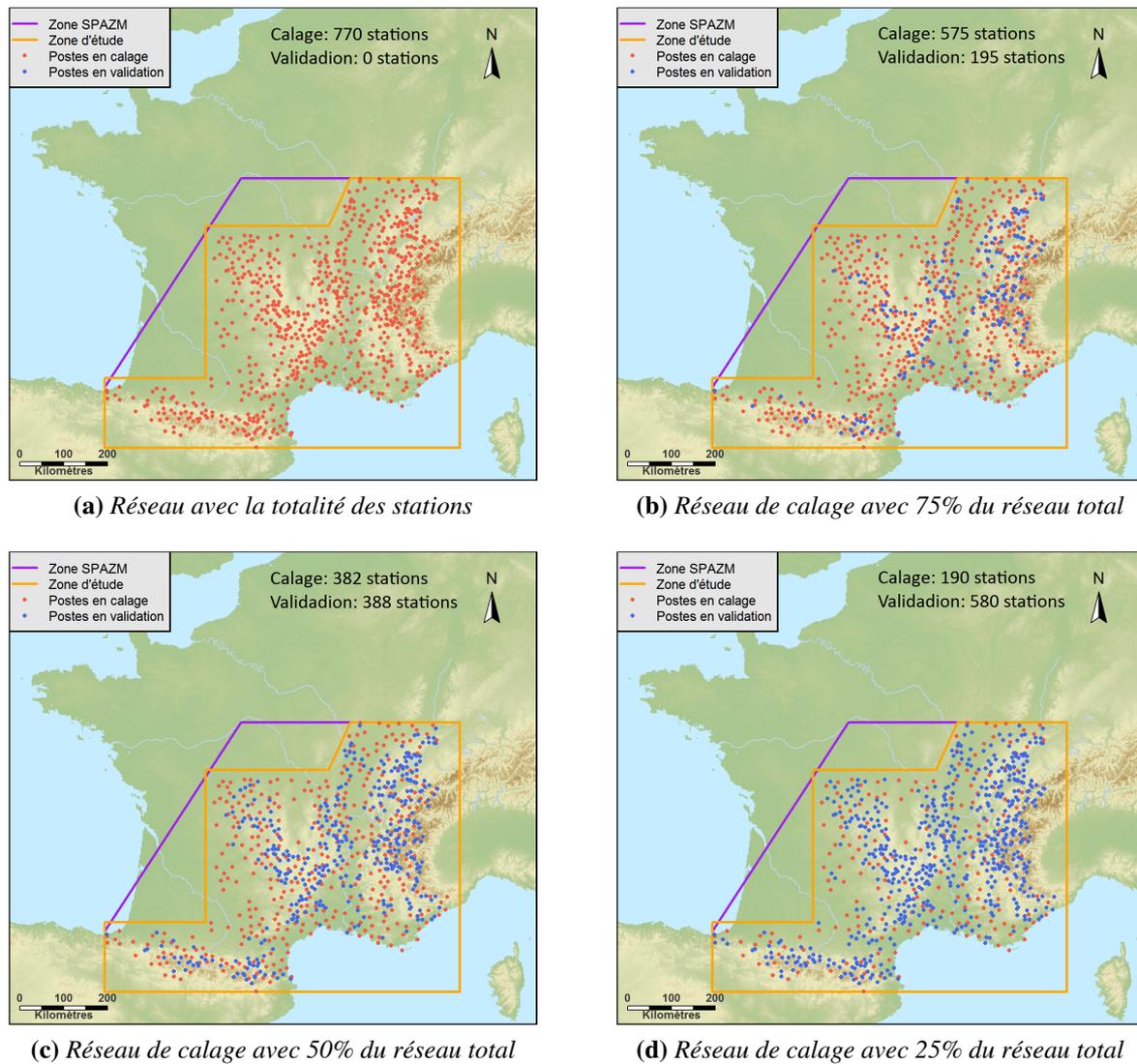


FIGURE III.2 – Réseaux de différentes densités. Les points rouges sont les postes conservés pour la régionalisation, les points bleus sont les stations servant à la validation.

% du réseau en calage	Densité du réseau en calage (km ² /poste)	Nombre de stations en calage	Nombre de stations en validation
90	519	692	78
80	584	614	156
75	623	575	195
60	779	459	311
50	935	382	388
40	1170	304	466
30	1557	228	542
25	1870	190	580

Tableau III.1 – Information sur les différents réseaux utilisés pour la validation

Après la définition de réseaux de calage et de validation, il reste à réfléchir à une métrique pour évaluer les méthodes. Une des difficultés avec l'estimation des valeurs extrêmes est que, par définition même d'une valeur extrême, nous ne disposons d'aucune observation à laquelle confronter nos estimations. Les estimations régionales (obtenues avec les méthodes régionales)

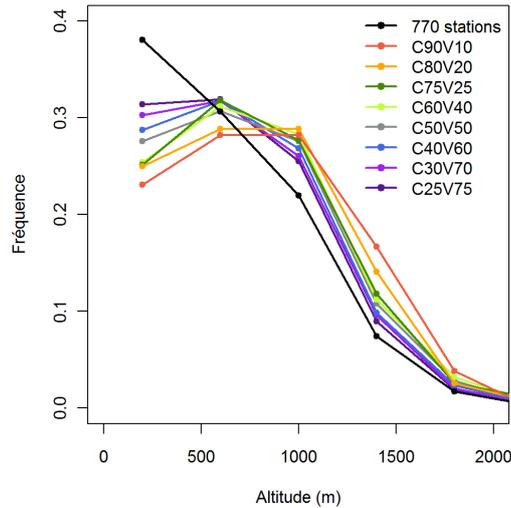


FIGURE III.3 – Distribution des altitudes des stations utilisées en validation pour chaque réseau considéré.

seront donc confrontées aux estimations locales (estimations obtenues par ajustement sur les séries ponctuelles). Cette comparaison est envisageable, car les séries de la base de données ont été choisies de façon à fournir des estimations locales suffisamment robustes pour que nous puissions les considérer comme informatives. Trois différents scores ont alors été choisis pour confronter les valeurs estimées régionalement \hat{X} et les valeurs estimées localement X aux postes utilisés en validation (e.g. X la population d'un quantile donné) :

- 1) Le biais global afin d'évaluer la moyenne des biais locaux sur l'estimation de X :

$$B = 1 - \overline{\left(\frac{\hat{X}}{X}\right)}$$

Une valeur nulle indique qu'en moyenne, la méthode reproduit parfaitement les valeurs X . Au contraire, une valeur positive (réciproquement, négative) met en évidence une sous-estimation (réciproquement, une surestimation) de la moyenne.

- 2) Le critère de NASH (Nash et Sutcliffe, 1970) score la capacité prédictive d'un modèle par rapport à une simple moyenne des observations X :

$$NASH = 1 - \frac{\sum_{i=1}^n (x_i - \hat{x}_i)^2}{\sum_{i=1}^n (x_i - \bar{X})^2}$$

Si $NASH = 1$: modèle idéal, si $NASH \leq 0$: modèle globalement moins bon qu'une simple moyenne.

- 3) L'écart-type des résidus $\sigma(\epsilon)$, avec les résidus s'exprimant de la façon suivante $\epsilon = \hat{X} - X$, afin de juger de la dispersion de l'erreur commise.

III. 2 Les méthodes de régionalisation des pluies extrêmes

À partir de la base de données pluviométriques, différentes approches pour régionaliser les pluies extrêmes ont été mises en place. Trois approches ont été envisagées avec une régionalisation envisagée à différents niveaux du processus :

Méthode 1 : Krigeage des quantiles de pluies. C'est une approche classique qui consiste à estimer les pluies de temps de retour aux postes où l'information est disponible et de construire les cartes de quantiles à partir d'un krigeage de ces valeurs. L'étape de régionalisation est alors envisagée au niveau des quantiles.

Méthode 2 : Spatialisation des paramètres de la distribution MEWP. Avec cette méthode, la régionalisation est réalisée en amont de l'estimation des quantiles. Elle consiste à spatialiser les paramètres de la distribution MEWP de façon à pouvoir reconstruire une distribution en chaque point du domaine d'étude et de pouvoir construire les cartes de quantiles remarquables. L'ajustement MEWP étant une distribution composite par types de temps et par saison, le concept de types de temps peut être utilisé pour la régionalisation des paramètres.

Méthode 3 : Estimation des pluies extrêmes à partir des pluies SPAZM. Cette dernière approche propose d'utiliser les pluies journalières SPAZM comme information. L'étape de régionalisation est donc directement réalisée sur les pluies journalières par l'intermédiaire de l'interpolateur SPAZM. Les ajustements et les forts quantiles sont alors estimés sur les séries régionales de SPAZM.

Le principe de chacune de ces trois méthodes est détaillé dans la suite de façon à expliquer leurs particularités.

III. 2. 1 Méthode 1 : Krigeage des quantiles de pluies

Nous disposons de 770 séries pluviométriques sur la période 1953-2005. Dans un premier temps **une distribution MEWP** est ajustée sur la série de pluie journalière **de 1953 à 2005** à chaque poste en utilisant la **classification par type de temps EDF2010**. La saisonnalité des pluies pour cet ajustement est **la saison à risque de trois mois et sa saison complémentaire de neuf mois** présentée dans la section II. 3. 1. Les pluies au temps de retour souhaité sont alors obtenues au droit de chaque station.

Nous allons étudier la structure spatiale pour les quantiles de pluies puis les kriger afin d'obtenir des cartes de pluies extrêmes pour les temps de retour T ($T=10, 100, 1000$ et $10\ 000$ ans). Le krigeage sera réalisé sur le domaine d'étude, car nous ne possédons que des données sur cette zone et que toute extrapolation en dehors aurait peu de validité.

III. 2. 1.1 Détermination de la structure spatiale des pluies extrêmes

La structure d'un champ d'une variable aléatoire Z peut être décrite par un semi-variogramme. Cette méthode permet d'expliquer la corrélation $\gamma(z_1, z_2)$ entre deux points $z_1(x_1, y_1)$ et $z_2(x_2, y_2)$ uniquement en fonction de leur inter-distance d , la distance euclidienne entre les deux points z_1 et z_2 . Le semi-variogramme de la variable aléatoire Z s'exprime :

$$\gamma(z_1, z_2) = \frac{1}{2} \cdot E \left[(Z(x_1, y_1) - Z(x_2, y_2))^2 \right] \quad (\text{III.1})$$

Sans détailler le principe complet du krigeage (Journel (1977) et Delhomme (1978) pour plus de détails), la seule hypothèse est que les accroissements de la variable aléatoire Z sont stationnaires d'ordre 2, c'est-à-dire qu'ils ne dépendent pas de la position des points dans l'espace

(Ox, Oy) , mais seulement de la distance d entre les points :

$$\begin{cases} E[z_1 - z_2] = 0 \\ \text{var}[z_1 - z_2] = 2\gamma(z_1, z_2) \end{cases} \quad (\text{III.2})$$

Le semi-variogramme s'exprime alors directement en fonction de l'inter-distance d :

$$\gamma_Z(d) = \frac{1}{2} \cdot E \left[(Z(x_1, y_1) - Z(x_2, y_2))^2 \right] \quad (\text{III.3})$$

La construction de tels semi-variogrammes suppose une structure spatiale identique quelle que soit la direction (i.e. isotropique) pour la variable Z .

Avec les stations servant au calage, selon la densité du réseau utilisé, un semi-variogramme empirique est calculé pour les pluies de chaque temps de retour. Les graphes de la figure III.5 montrent que les pluies décennales, centennales, millénales et décamillénales ont des structures similaires. Un semi-variogramme de forme exponentielle est déterminé pour chaque temps de retour (courbes en pointillés bleus). Les variogrammes sont caractérisés par trois paramètres (figure III.4) : la portée, le palier et la pépité. Dans le cas des quantiles de pluie, la portée est de l'ordre d'une centaine de kilomètres et son palier augmente avec le temps de retour.

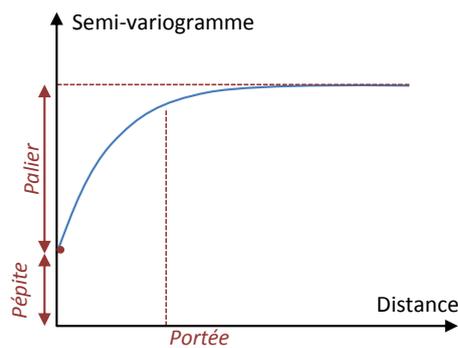


FIGURE III.4 – Définition de la portée, du palier et de la pépité d'un variogramme.

Les variogrammes sont peu dépendants du réseau utilisé pour les construire même si le tableau III.2 montre que plus le réseau est dense, plus la portée du variogramme est faible et que la valeur du palier augmente. Ceci est dû au fait qu'un réseau plus dense capte mieux la variabilité spatiale entre les postes (augmentation de la valeur du palier), mais également une structure spatiale plus complexe et donc a priori plus rapidement décorrélée (portée plus faible). La présence d'un effet pépité (i.e. discontinuité à l'origine du variogramme) est, d'après le travail de thèse d'Emmanuel (2011), due à l'utilisation d'un réseau sol. Les observations pluviométriques obtenues (plus ou moins fiables) à partir du réseau sol ne sont qu'un échantillonnage du champ complet. Dans le cas d'inter-distance très inférieure à l'espacement entre les mesures, la présence d'un effet de pépité traduit cette incapacité à d'écrire la structure à faible distance. Cependant Emmanuel a montré que l'utilisation d'images radar avec une résolution fine (1 km^2) permet de faire disparaître cette pépité.

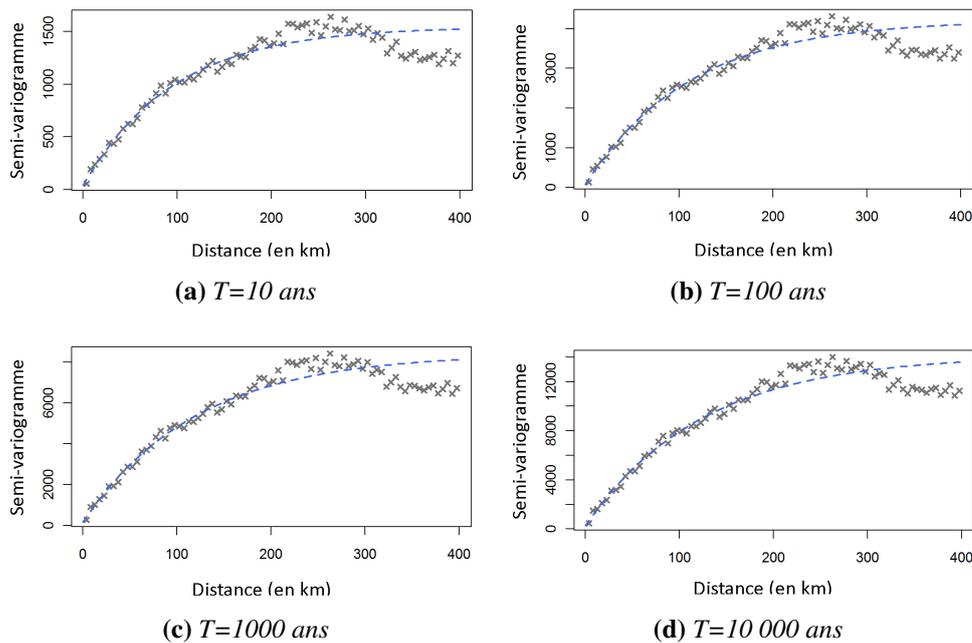


FIGURE III.5 – Semi-variogrammes (en mm^2) des quantiles de pluies de temps de retour 10, 100, 1000 et 10 000 ans calés sur la totalité du réseau.

T (année)	Pépité (mm^2)	Palier (mm^2)	Portée (km)
10	0	1116	132
100	0	3058	147
1000	0	6083	150
10 000	0	10179	149

(a) 25% du réseau en calage

T (année)	Pépité (mm^2)	Palier (mm^2)	Portée (km)
10	69	1211	111
100	132	3419	126
1000	232	6945	133
10 000	380	11788	137

(b) 50% du réseau en calage

T (année)	Pépité (mm^2)	Palier (mm^2)	Portée (km)
10	61	1248	96
100	114	3481	109
1000	210	7002	116
10 000	365	11805	119

(c) 75% du réseau en calage

T (année)	Pépité (mm^2)	Palier (mm^2)	Portée (km)
10	40	1504	96
100	76	4137	111
1000	152	8261	120
10 000	281	13869	124

(d) 100% du réseau en calage

Tableau III.2 – Paramètres des semi-variogrammes des quantiles de pluies de temps de retour 10, 100, 1000 et 10 000 ans pour différentes densités de réseaux.

III. 2. 1.2 Construction des cartes de pluies extrêmes

Pour chaque réseau de calage, un krigeage ordinaire (sans prise en compte de variables explicatives) est réalisé en s'appuyant sur les stations disponibles pour le calage et le variogramme exponentiel déterminé à partir de ces dernières. Des cartes de pluies extrêmes sont produites pour chaque temps de retour à la maille 1 km² dans toute la zone d'étude. Les cartes sont réalisées pour chacun des réseaux considérés.

Les cartes de la figure III.6 ont été construites avec un krigeage ordinaire des quantiles de pluies MEWP déterminés localement en utilisant la totalité des 770 stations. La zone violette correspond à la zone SPAZM et la zone orange à la zone d'étude.

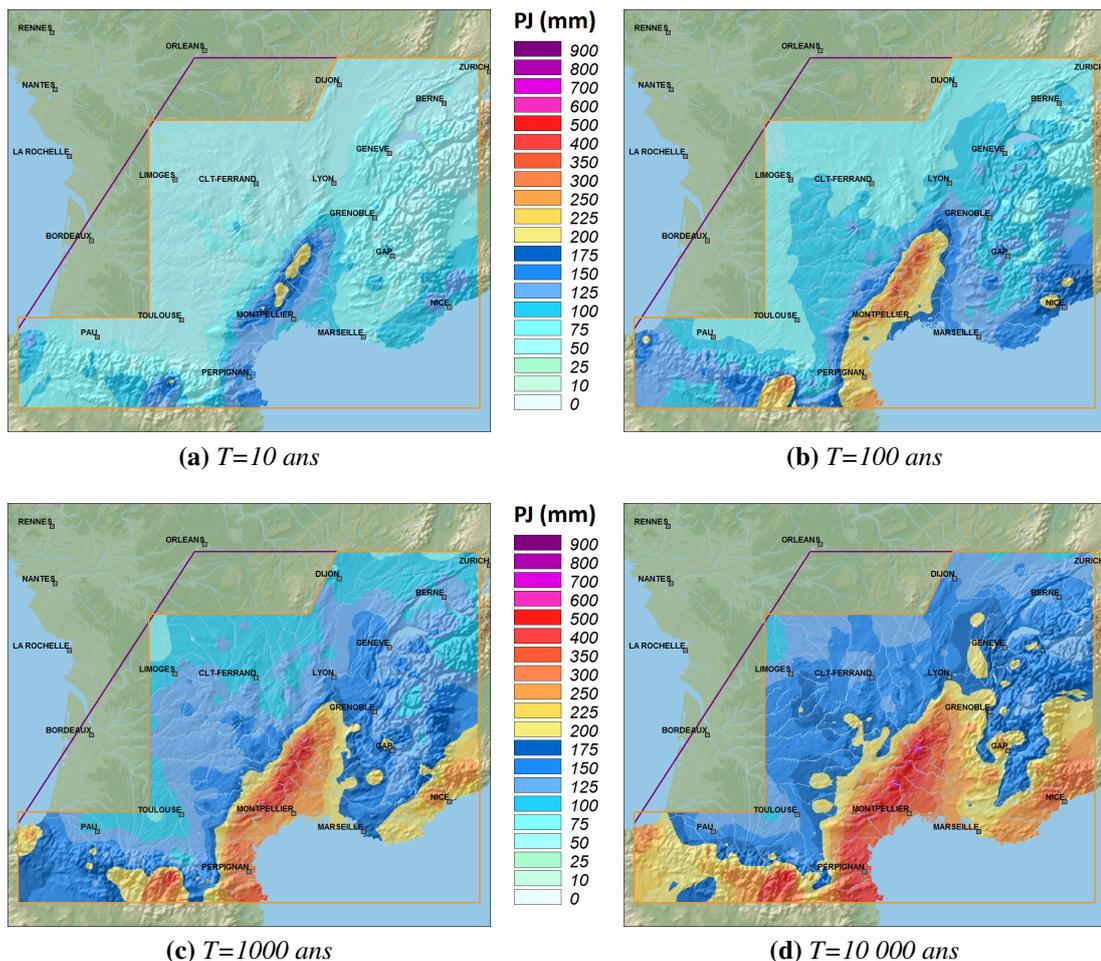


FIGURE III.6 – Cartes des quantiles de pluie de différents temps de retour obtenues par krigeage sur la totalité des 770 postes.

Les cartes obtenues par krigeage des quantiles sont des champs plutôt lisses avec peu d'effets locaux, mais elles mettent en évidence des zones à risque. Les plus forts quantiles de pluies se trouvent sur le pourtour méditerranéen, les Cévennes, le Vivarais, les Pyrénées orientales et les Alpes Maritimes. Nous notons également que les cartes mettent en évidence une remontée des pluies le long de la vallée du Rhône. Les Écrins (4100 m), le mont Ventoux (1912 m) et le Plomb du Cantal (1860 m) sont des reliefs marqués et exposés à des flux méditerranéens pour les deux premiers, et atlantique pour le Plomb du Cantal. Les quantiles de pluies sont plus élevés sur ces reliefs dès le temps de retour 10 ans.

III. 2. 1.3 Évaluation du krigeage des quantiles

Les scores de biais, de NASH et d'écart-type des erreurs entre la population des estimations locales de pluies extrêmes et la population des estimations obtenues par régionalisation sont calculés pour chaque réseau de validation.

Les graphes de la figure III.7 représentent les différents scores obtenus selon la densité de réseau utilisé en calage. Nous constatons que le krigeage ordinaire donne déjà des résultats satisfaisants et ce quel que soit le temps de retour lorsque la totalité du réseau est utilisé en calage. Le biais reste faible et peu dépendant de la densité de réseau utilisé en calage et du temps de retour. Au contraire, les autres scores se dégradent rapidement lorsque la densité du réseau de calage diminue. Les NASH pour les temps de retour plus faibles (en particulier pour la pluie décennale) sont plus faibles. Bien que le NASH semble mettre en évidence une moins bonne spatialisation des faibles temps de retours, l'écart-type des résidus augmente largement avec le temps de retour. Ceci laisse penser que la structure de la pluie décennale étant moins variable que pour des temps de retour plus élevé (figure III.6), le NASH montre que les estimations obtenues expliquent une partie de la variance par rapport à un champ moyen, mais que cet apport est moins important que pour les temps de retour plus élevés où la variabilité est plus forte.

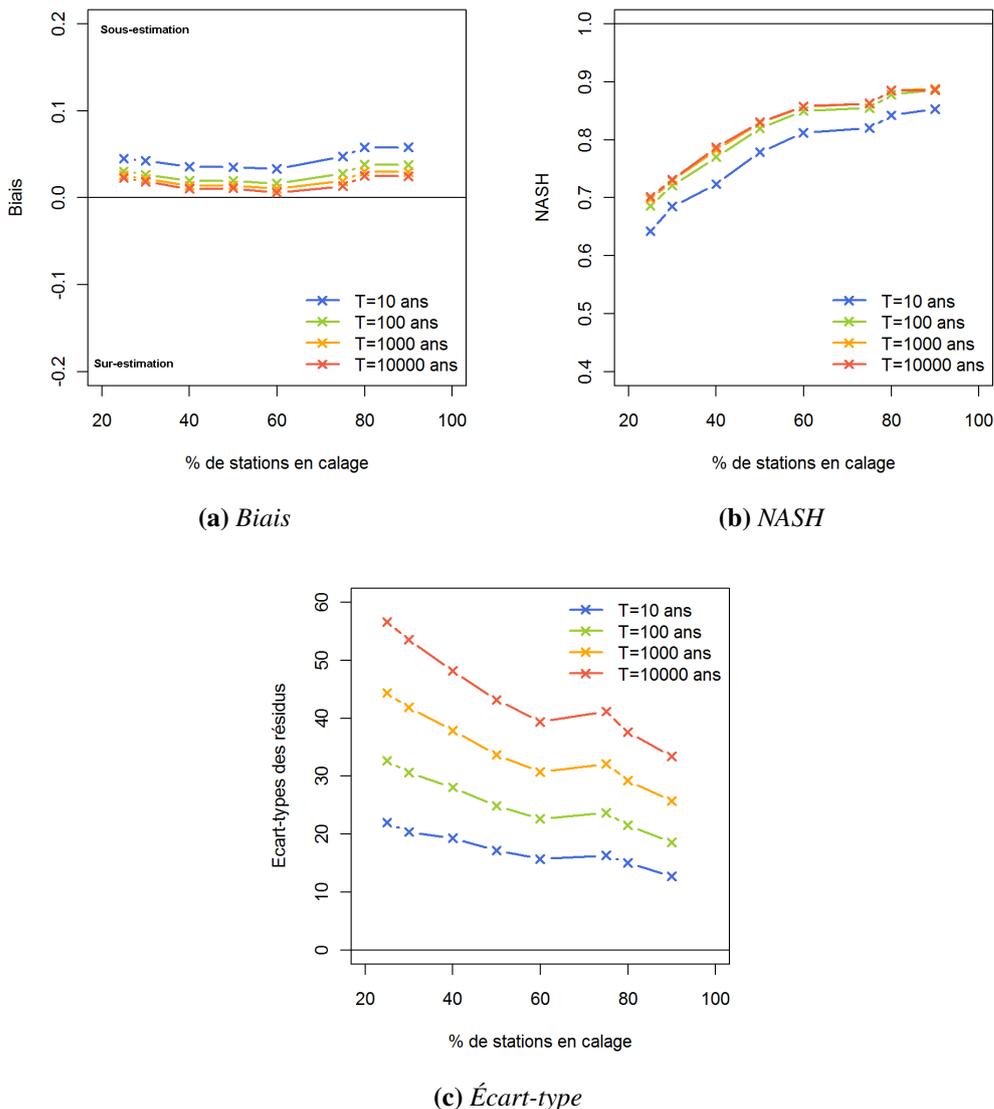


FIGURE III.7 – Scores pour différentes densités de réseau pour la pluie millénale.

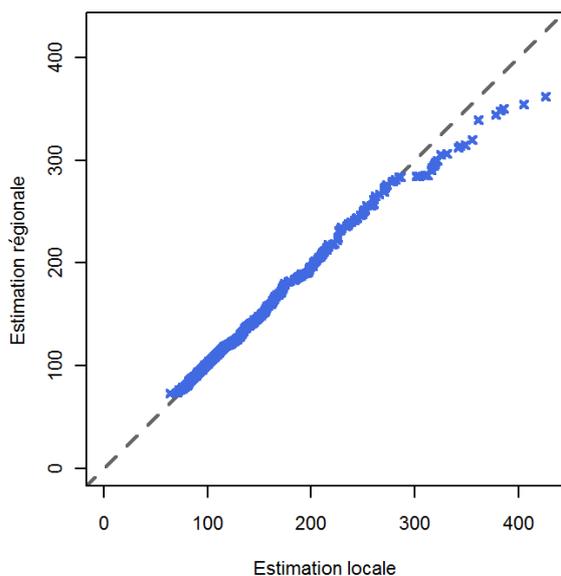


FIGURE III.8 – *Graphique quantile-quantile des pluies centennales (en mm) estimées localement et régionalement en Jackknife.*

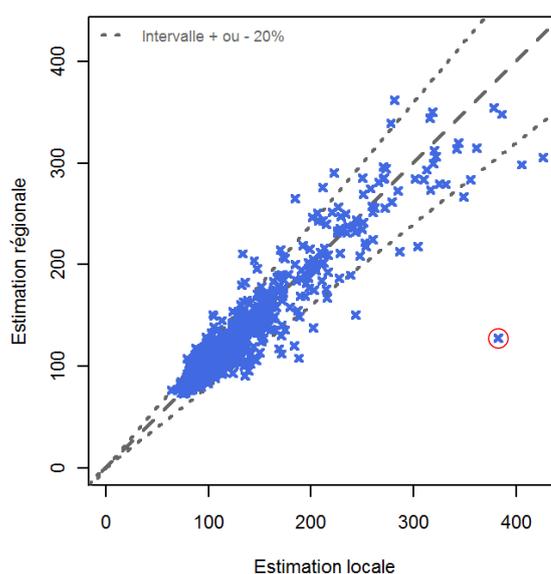


FIGURE III.9 – *Graphique scatter-plot des pluies centennales (en mm) estimées localement et régionalement en Jackknife. Les droites en pointillé représentent un intervalle de +/- 20%.*

Sur les figures III.8 et III.9, la confrontation des estimations locales et régionales en Jackknife² montre que la méthode de krigeage des pluies extrêmes a tendance à sous-estimer les fortes pluies de chaque temps de retour. Ceci est sûrement dû au krigeage qui a tendance à lisser un peu les champs et à ne pas représenter des valeurs plus importantes si les valeurs autour sont plus faibles. Cette sous-estimation se voit sur le scatter-plot avec une tendance à rester sous la première bissectrice pour les valeurs les plus importantes et sur le graphe quantile-quantile où les plus fortes pluies centennales ont une occurrence moins importante en régional que sur les valeurs estimées localement. Nous donnerons la préférence au scatter-plot, car le graphe quantile-quantile réorganise les événements et perd la correspondance entre les estimations locale et régionale d'un point.

Remarque : Sur la figure III.9, pour une station, l'estimation locale du quantile de temps de retour 100 ans est de 380 mm alors que l'estimation régionale est de 109 mm (cercle rouge). Cette station "Pl_ESP_9669" est une station espagnole dont le risque de pluie est nettement plus fort que celui des stations à proximité. Le krigeage ordinaire ne peut reproduire le risque de pluie de cette station, au bord de la zone d'étude, uniquement à l'aide des stations voisines.

L'origine de cette sous-estimation des plus forts quantiles de pluies est liée au processus d'interpolation du krigeage. Sans variable explicative, la variable krigée ne peut être extrapolée à des valeurs plus fortes que les valeurs observées autour du point à estimer. Les champs ont donc tendance à être lissés.

Afin de pallier cette sous-estimation des plus fortes pluies de temps de retour donné, un krigeage avec dérive externe a été envisagé afin de prendre en compte l'impact du relief pour les stations en altitude. Les résultats sont peu différents de ceux obtenus avec le krigeage ordinaire et semblent même se dégrader plus rapidement lors de la dédensification du réseau ce qui laisse penser que le relief n'est pas un facteur explicatif pertinent des quantiles de pluies.

2. Une estimation régionale en un site est obtenue à partir d'un jeu de calage regroupant toutes les stations disponibles sauf celle présente en ce site.

III. 2. 2 Méthode 2 : Spatialisation de la distribution MEWP par krigeage des gradex

Avec le krigeage des quantiles, l'ajustement MEWP a été réalisé en local puis les pluies de temps de retour 10, 100, 1000 et 10 000 ans ont été spatialisées. La seconde approche qui va être détaillée utilise toujours les ajustements MEWP locaux, mais régionalise les paramètres de la distribution afin de pouvoir estimer une distribution MEWP en tout point du domaine.

III. 2. 2.1 La régionalisation des paramètres MEWP

La régionalisation de la loi MEWP demande de spatialiser chacun de ses paramètres. La loi MEWP est ajustée sur un échantillon de pluies centrales (Garavaglia, 2011b). Il faut donc connaître l'occurrence des pluies centrales par rapport à l'ensemble des observations disponibles. Les ajustements étant réalisés sur deux saisons, il faut également connaître les occurrences de chaque saison afin de pouvoir recomposer les deux distributions saisonnières. Pour chaque saison, une loi exponentielle est ajustée sur la population de pluies centrales d'un type de temps seuillées au quantile 70%. Nous devons donc disposer d'une carte des gradex et des seuils par type de temps des lois exponentielles ainsi que de l'occurrence de chaque type de temps afin de recomposer la loi saisonnière. En théorie, la construction d'une carte de distribution MEWP demande donc finalement 51 cartes³ afin de disposer de tous les éléments nécessaires à la recombinaison de la distribution annuelle.

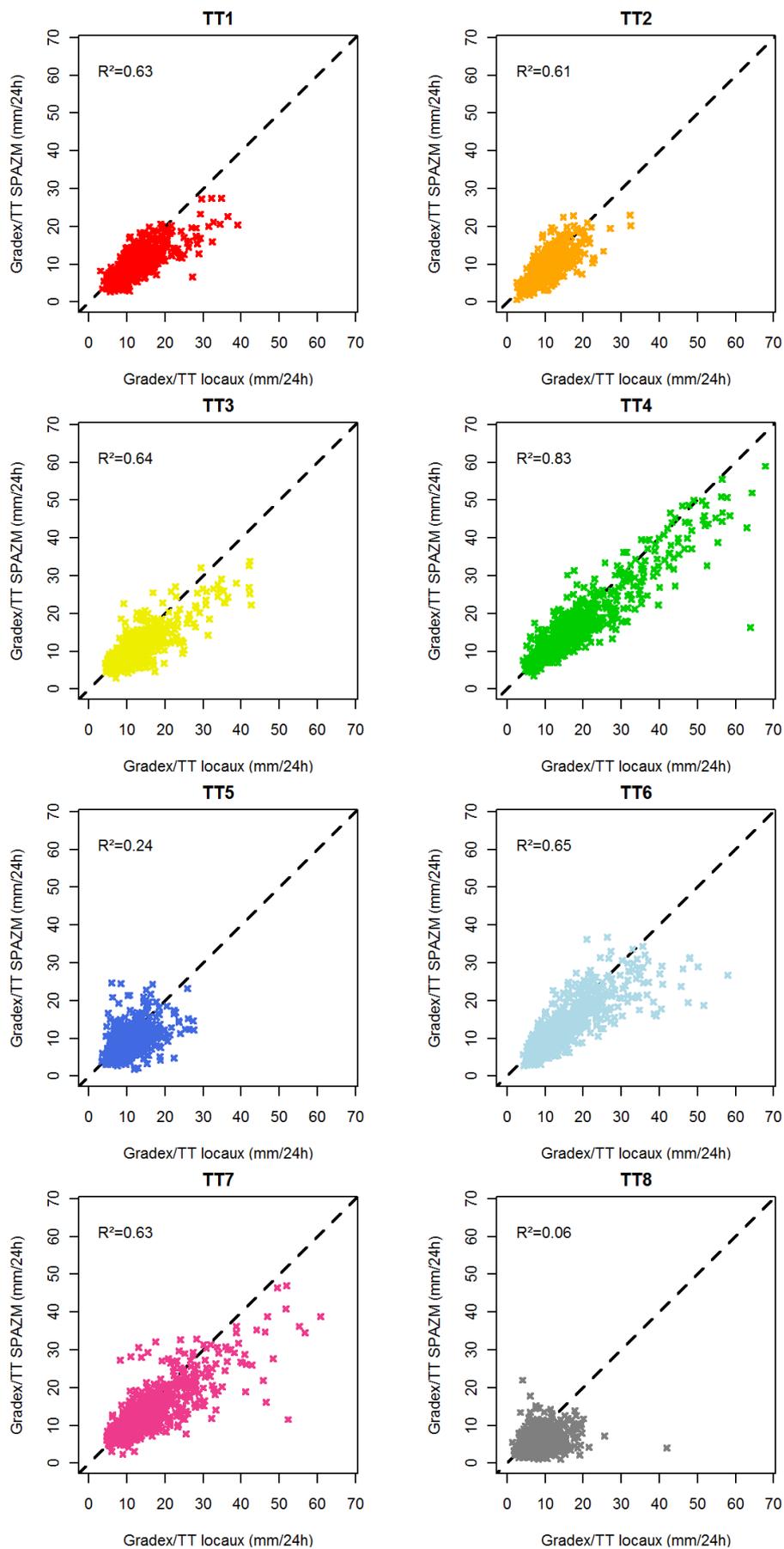
Afin de simplifier la spatialisation de tous ces paramètres, certains d'entre eux peuvent directement être calculés sur la série des pluies journalière SPAZM. Une distribution MEWP est ajustée sur la série de pluies journalières mesurée en un point. Une seconde distribution MEWP est ajustée sur la série des pluies journalières issue de SPAZM en ce même point. SPAZM est un interpolateur exact de pluie journalière, ce qui signifie que la valeur de pluie au pixel d'une station vaut chaque jour la valeur observée à cette station. Pour réaliser une validation rigoureuse avec les données SPAZM, il faudrait refaire l'ensemble de son processus de calage sans l'information à la station à évaluer (construction et déformation des ébauches, krigeage des résidus). Dans notre cas, les ébauches de pluies n'ont pas été recalculées et les étapes de déformation des ébauches et le krigeage de résidus ont été réalisées sans l'information de la station considérée. Nous parlerons alors de pluie SPAZM en Jackknife. Ce choix de recalculer SPAZM en Jackknife en omettant seulement une station plutôt que d'utiliser les réseaux de calage et de validation définis page 50 s'explique de deux façons :

- un intérêt majeur de l'interpolateur SPAZM est d'être capable de digérer une grande quantité d'information, y compris celle de séries avec très peu de données. Recaler la méthode sur le jeu de données avec les 770 stations alors qu'elle utilise normalement plus de 2200 postes serait très pénalisant et peu représentatif de sa capacité à intégrer toute l'information disponible.
- la difficulté de refaire le calage de la totalité du processus SPAZM qui assimile des données de différentes sources pour la construction des ébauches (mesure de pluviomètre, de NPT et NRC⁴ pour la neige). Une reconstruction des ébauches serait donc lourde à mettre en place.

Les pluies SPAZM en Jackknife permettent ainsi d'évaluer l'intérêt de l'information SPAZM dans la régionalisation des quantiles élevés de précipitations en conservant l'avantage majeur de la méthode SPAZM qui permet d'assimiler toute l'information pluviométrique disponible.

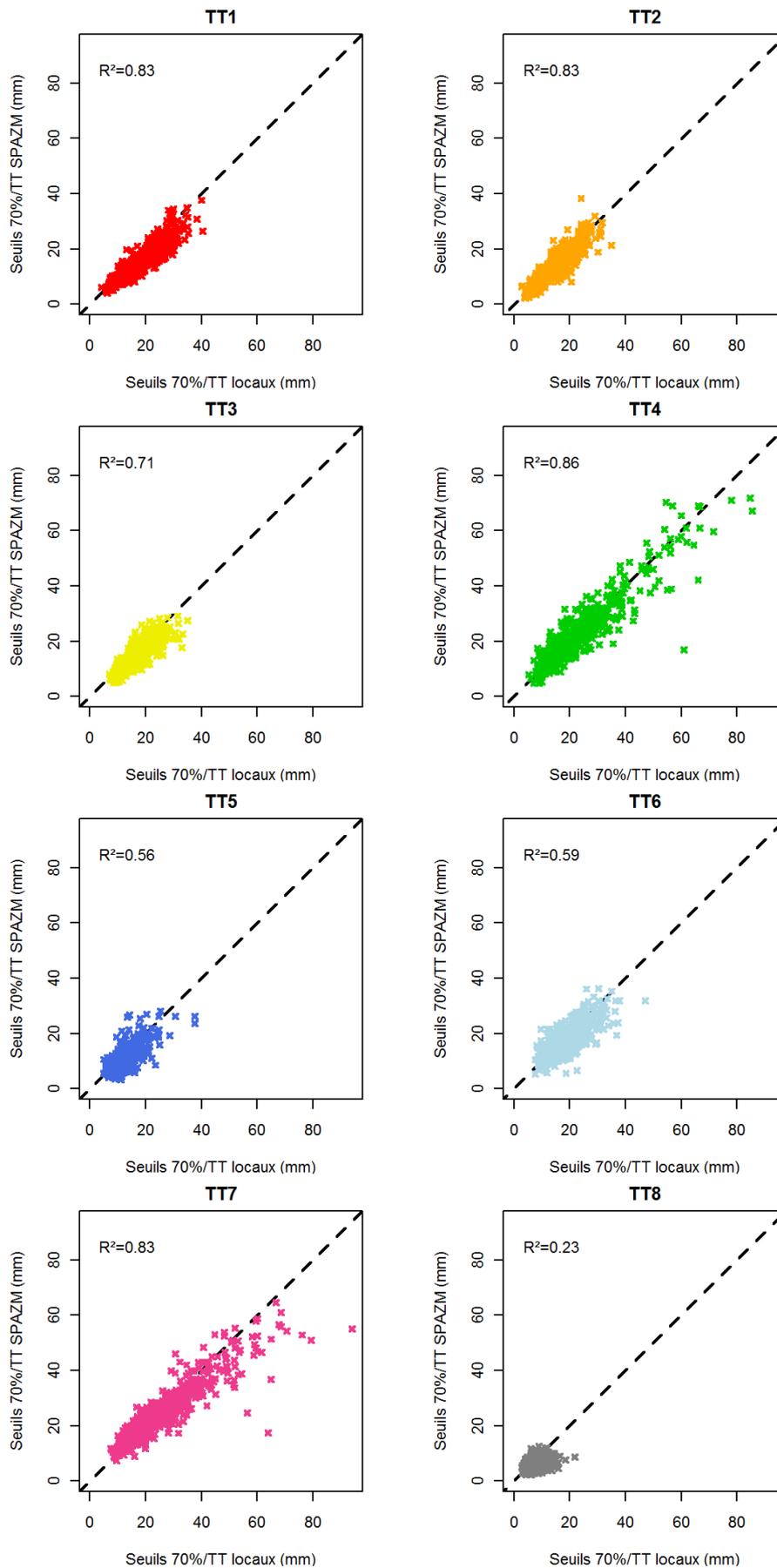
3. 2 saison × 8 type de temps × 3 paramètres + 2 occurrences pour les saisons + 1 occurrence de pluie centrale.

4. Nivo-pluviomètres Totalisateurs et Nivomètres à Rayonnement Cosmique



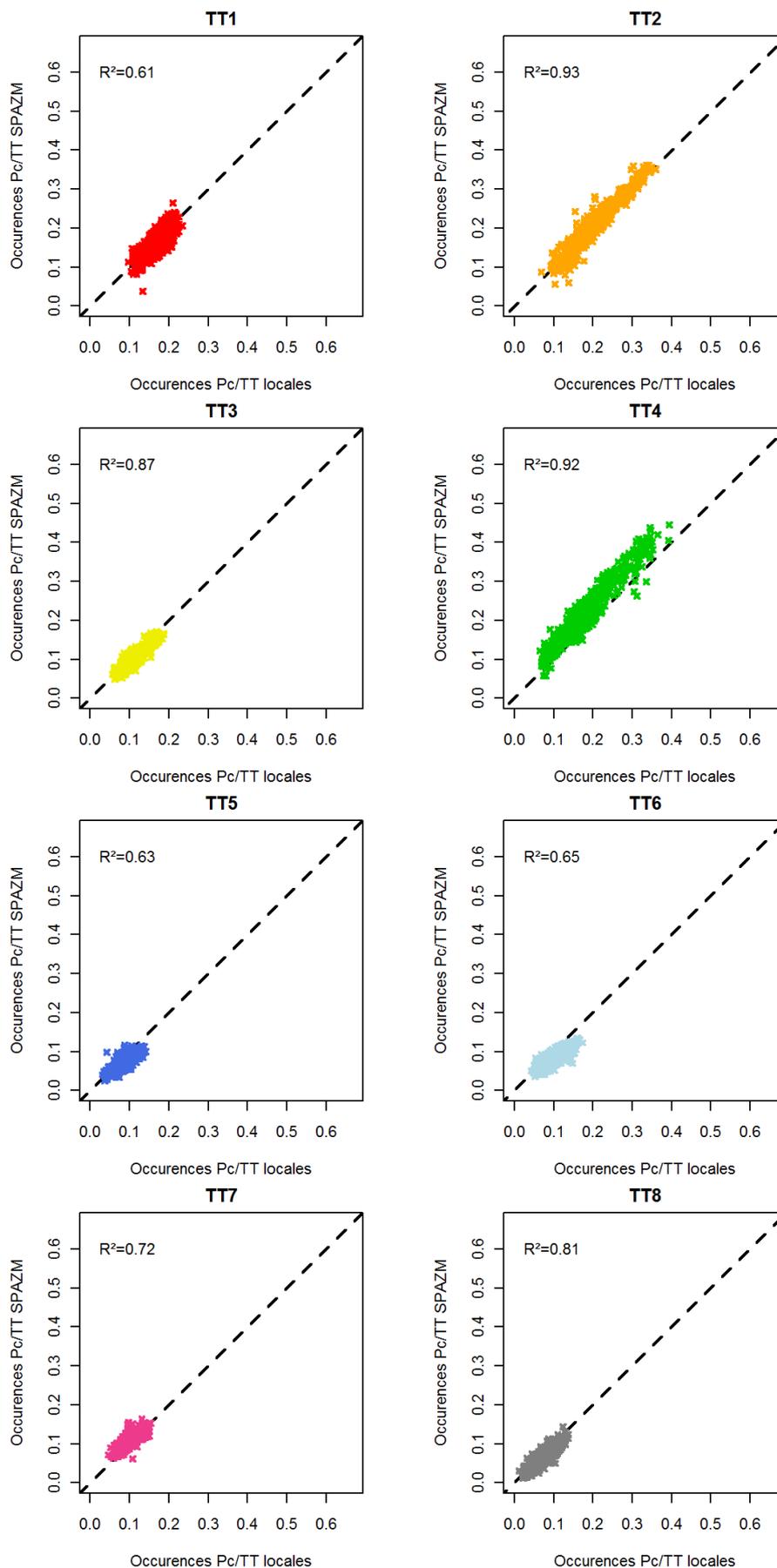
(a) Gradex par type de temps

FIGURE III.10 – Comparaison des paramètres estimés sur les séries observées et sur les séries SPAZM pour la saison à risque.



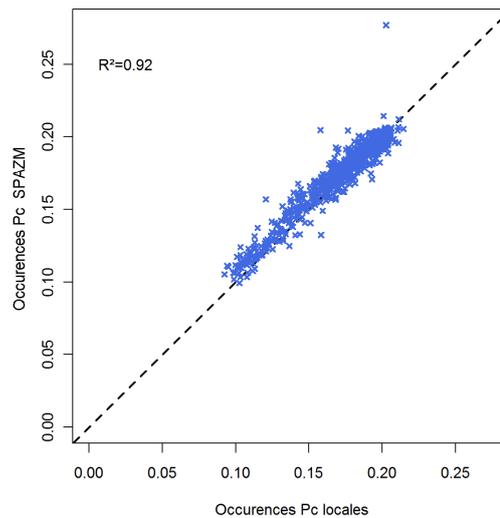
(b) Seuils par type de temps

FIGURE III.10 – Comparaison des paramètres estimés sur les séries observées et sur les séries SPAZM pour la saison à risque.



(c) Occurrences des types de temps

FIGURE III.10 – Comparaison des paramètres estimés sur les séries observées et sur les séries SPAZM pour la saison à risque.



(d) Occurrence des pluies centrales

FIGURE III.10 – Comparaison des paramètres estimés sur les séries observées et sur les séries SPAZM pour la saison à risque.

La figure III.10 permet de comparer les paramètres issus de l’ajustement sur le poste pluviométrique et ceux de l’ajustement sur les pluies SPAZM en Jackknife pour les 770 stations. Les résultats sont présentés pour la saison à risque (de trois mois), mais les conclusions sont les mêmes sur la saison non à risque (de neuf mois).

Dans l’ensemble, la pluie SPAZM permet une bonne estimation des paramètres. Sur la figure III.10, les seuils (sous-figure (b)) et les occurrences (sous-figures (c) et (d)) pourront donc être calculés avec les pluies SPAZM. Par contre, les gradex (sous-figure (a)) semblent être moins bien estimés avec SPAZM, or ce sont eux qui vont influencer le plus la valeur des quantiles extrêmes de pluies. **Les gradex par type de temps et par saison vont donc être spatialisés par une autre méthode de façon à améliorer leur estimation.** Ainsi, pour déterminer la loi MEWP en un pixel donné, la pluie SPAZM servira à déterminer les 16 seuils, les 16 occurrences des types de temps (2 saisons et 8 types de temps) et l’occurrence des pluies centrales. L’occurrence des saisons représente la fréquence de pluies centrales de chacune des deux saisons. Elle est en fait très peu différente de la fréquence de la saison dans l’ensemble des données qui est elle connue puisque nous travaillons toujours avec une saison à risque de trois mois. L’occurrence de la saison à risque est donc 0.25 et celle de la saison complémentaire de 0.75. Enfin, les 16 gradex par type de temps seront issus d’un krigeage.

III. 2. 2.2 Recherche de variables explicatives pour les gradex par types de temps

La façon dont les gradex par type de temps ont été régionalisés va être détaillée dans la partie qui suit.

Recherche de variables explicatives pour les gradex par type de temps

Afin d'améliorer le krigeage des gradex par type de temps GTT_i , des variables explicatives potentiellement pertinentes ont été envisagées :

- la position X en Lambert II étendue ;
- la position Y en Lambert II étendue ;
- le relief avec l'altitude Z ;
- la distance à la Méditerranée $dMed$;
- la distance à l'océan Atlantique $dAtl$;
- la pluie annuelle moyenne PAM ;
- les pluies annuelles par type de temps ETT_i (i.e. les ébauches par type de temps de SPAZM).

Le tableau III.3 donne la valeur du coefficient de détermination entre les gradex par type de temps et les variables explicatives. Les variables les plus explicatives pour chaque type de temps sont communes aux deux saisons (valeurs surlignées en rouge). Deux grands jeux de variables sont définis.

Le premier est un jeu de **variables globales** qui sont des descriptifs classiques comme l'altitude, la distance à la Méditerranée, le cumul de pluie... Avec ces variables, le type de temps 2, responsable des forts cumuls de pluie, est bien corrélé à la pluie annuelle. En effet, le type de temps 2 correspond au type de temps le plus courant et est souvent à l'origine d'importants volumes de précipitation. Les types de temps méditerranéens (types de temps 4, 6 et 7) sont quant à eux plus liés à la distance à la Méditerranée.

Un deuxième jeu de variables, dit de **variables explicatives par type de temps**, vise à tester la corrélation des champs de gradex par type de temps aux cumuls des pluies de ce type de temps. L'idée est d'essayer de dégager des similarités entre les cumuls de pluies et les gradex par types de temps. Nous pourrions nous attendre à ce que le cumul d'un type de temps i explique le gradex de ce type de temps i . Les tableaux III.3 soulignent différentes tendances. Tout d'abord, les types de temps 1, 2 et 3 (qui sont des représentants des influences d'ouest ou de sud-ouest) sont plutôt faiblement corrélés à leur cumul. Les gradex de type de temps 1 et 3 semblent par contre bien corrélés au cumul du type de temps 7 et les gradex de type de temps 2 au cumul du type de temps 1. Le type de temps 5 ne possède pas de structure bien définie sur la zone d'étude ce qui explique la difficulté à trouver des variables explicatives intéressantes pour ce type de temps. De même, le type de temps 8 est un type de temps regroupant les orages d'été locaux et n'a donc pas de véritable structure spatiale (au pas de temps journalier). Enfin, les types de temps 4, 6 et 7 (qui sont des flux méditerranéens) sont assez bien expliqués par leurs cumuls respectifs. Il est également intéressant de remarquer que le cumul de type de temps 4 explique très bien les gradex de ces trois types de temps.

Pour mieux comprendre pourquoi certains gradex sont plus ou moins expliqués par certains cumuls de types de temps, nous pouvons regarder la position relative de chaque classe dans l'espace des champs de pressions (espace \mathbb{R}^{440} dans lequel ont été définies les huit classes). Sur la figure III.11, les barycentres des huit classes sont représentés dans l'espace du premier plan de l'analyse en composantes principales. Les deux composantes correspondent plus ou moins à des flux atmosphériques sud-ouest/nord-est en abscisse et sud-est/nord-ouest en ordonnée.

Nous pouvons constater que les types de temps 1, 3 et 7 – dont les gradex sont corrélés au cumul annuel de pluie du type de temps 7 – sont des classes proches dans le nuage. De même,

		<i>GTT1</i>	<i>GTT2</i>	<i>GTT3</i>	<i>GTT4</i>	<i>GTT5</i>	<i>GTT6</i>	<i>GTT7</i>	<i>GTT8</i>
Variables globales	<i>X</i>	0.10	0.00	0.02	0.03	0.00	0.07	0.04	0.00
	<i>Y</i>	0.02	0.02	0.02	0.24	0.11	0.36	0.03	0.06
	<i>Z</i>	0.00	0.10	0.04	0.00	0.00	0.02	0.01	0.04
	<i>dMed</i>	0.09	0.00	0.12	0.36	0.08	0.37	0.16	0.04
	<i>dAtl</i>	0.09	0.01	0.01	0.03	0.00	0.06	0.03	0.00
	<i>PAM</i>	0.03	0.40	0.01	0.00	0.00	0.01	0.01	0.00
Variables par types de temps	<i>ETT1</i>	0.00	0.26	0.00	0.18	0.02	0.21	0.04	0.01
	<i>ETT2</i>	0.08	0.14	0.07	0.23	0.11	0.21	0.16	0.03
	<i>ETT3</i>	0.00	0.11	0.00	0.02	0.00	0.10	0.00	0.02
	<i>ETT4</i>	0.11	0.02	0.16	0.69	0.06	0.34	0.34	0.01
	<i>ETT5</i>	0.02	0.14	0.07	0.05	0.01	0.00	0.07	0.00
	<i>ETT6</i>	0.14	0.09	0.06	0.28	0.12	0.38	0.12	0.04
	<i>ETT7</i>	0.47	0.15	0.26	0.17	0.06	0.01	0.30	0.00
	<i>ETT8</i>	0.01	0.16	0.04	0.02	0.01	0.01	0.01	0.00

(a) Pour la saison à risque

		<i>GTT1</i>	<i>GTT2</i>	<i>GTT3</i>	<i>GTT4</i>	<i>GTT5</i>	<i>GTT6</i>	<i>GTT7</i>	<i>GTT8</i>
Variables globales	<i>X</i>	0.08	0.00	0.04	0.03	0.00	0.05	0.00	0.01
	<i>Y</i>	0.07	0.03	0.02	0.26	0.17	0.36	0.08	0.23
	<i>Z</i>	0.01	0.13	0.02	0.01	0.00	0.02	0.01	0.01
	<i>dMed</i>	0.15	0.00	0.10	0.34	0.10	0.32	0.19	0.19
	<i>dAtl</i>	0.07	0.00	0.03	0.03	0.00	0.03	0.00	0.01
	<i>PAM</i>	0.10	0.50	0.04	0.01	0.04	0.00	0.04	0.04
Variables par types de temps	<i>ETT1</i>	0.04	0.33	0.00	0.15	0.00	0.15	0.06	0.00
	<i>ETT2</i>	0.03	0.24	0.05	0.23	0.04	0.20	0.14	0.04
	<i>ETT3</i>	0.01	0.15	0.02	0.01	0.00	0.12	0.00	0.00
	<i>ETT4</i>	0.12	0.01	0.19	0.67	0.08	0.29	0.51	0.23
	<i>ETT5</i>	0.00	0.29	0.03	0.03	0.09	0.00	0.03	0.01
	<i>ETT6</i>	0.26	0.07	0.10	0.34	0.26	0.46	0.23	0.24
	<i>ETT7</i>	0.42	0.08	0.25	0.19	0.06	0.00	0.29	0.15
	<i>ETT8</i>	0.05	0.23	0.00	0.00	0.05	0.00	0.00	0.05

(b) Pour la saison complémentaire

Tableau III.3 – Coefficient de détermination entre les gradex par type de temps et des variables explicatives potentielles.

les classes 7, 4 et 6 – dont les gradex sont bien corrélés au cumul annuel du type de temps 4 – sont aussi des classes adjacentes dans le nuage. La classification ayant été construite sur toutes les journées pluvieuses et non sur les journées de fortes précipitations, il est possible que certaines classes de pluies regroupent des pluies à risques similaires. Par exemple, le type de temps 4 est un bon représentant des pluies extrêmes avec une influence méditerranéenne et explique donc bien les gradex des types de temps 4 et 7. Le type de temps 6 (retour d'Est) étant un peu à part, il est bien expliqué par son propre cumul annuel. Les gradex de types de temps océaniques (types de temps 1 et 2) et de sud-ouest (type de temps 3) sont quant à eux plus difficilement corrélés à leur cumul annuel et il semble plus difficile de trouver un explicateur simple pour leur spatialisation excepté pour le type de temps 2 qui est très corrélé à la pluie annuelle moyenne *PAM*. Cette corrélation peut s'expliquer par le fait que le type de temps 2 est un risque de pluies souvent accompagné de fort cumul de précipitation. Le cumul annuel global est alors un meilleur descripteur de ce type de temps que le cumul annuel de type de temps 2 seul (le champs *ETT₂* est moins contrasté que le champ des *PAM*).

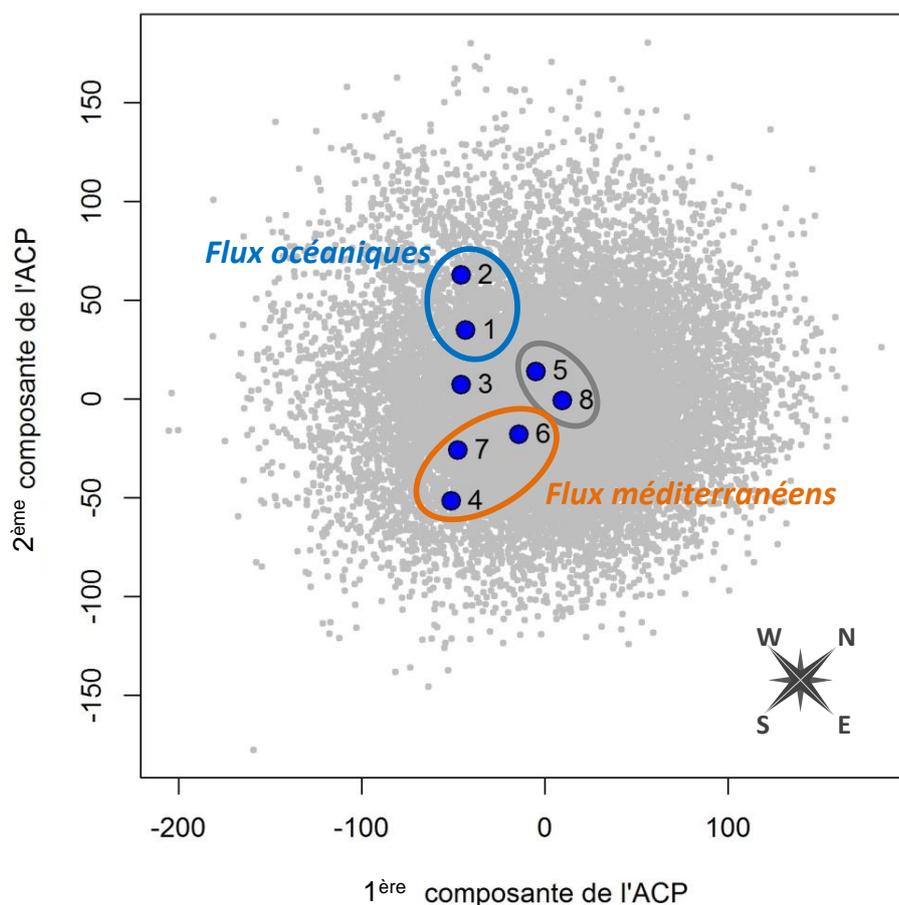


FIGURE III.11 – Représentation des huit classes de types de temps dans l'espace à deux dimensions des deux premières composantes principales.

Influence du relief sur les pluies extrêmes

Dans le tableau III.3, les gradex sont très peu liés à l'altitude des postes. Ceci laisse penser que les pluies extrêmes journalières sont globalement peu expliquées par le relief. Desrosne (1992) et Desrosne *et al.* (1994) se sont déjà intéressés à l'analyse fréquentielle des précipitations extrêmes et à leur liaison avec l'altitude en utilisant entre autres les données du TPG⁵ dans les Pré-Alpes du Nord.

L'étude de transects va permettre de visualiser le comportement des pluies extrêmes avec le relief (figures III.12, III.13 et III.14). Dans ces figures, les points représentent les valeurs des variables (relief ou gradex) évaluées au droit des postes à partir de l'information locale. Les courbes représentent soit le profil d'altitude (surface marron) soit les profils des pluies moyennes⁶ par type de temps ou annuelles.

Sur le transect du Jura (figures III.13), les types de temps océaniques 1 et 2 ont des cumuls annuels de précipitation assez bien corrélés avec le relief (les coefficients de détermination avec le relief sont respectivement de 0.80 et 0.85) alors que l'effet orographique est moins visible sur les cumuls des types de temps 3, 4 et 6. Les cumuls annuels de précipitations par type de temps

5. Transect de Pluviographes pour l'analyse et la modélisation des gradients d'intensités d'altitude, mis en service en 1987 et initié par IRSTEA.

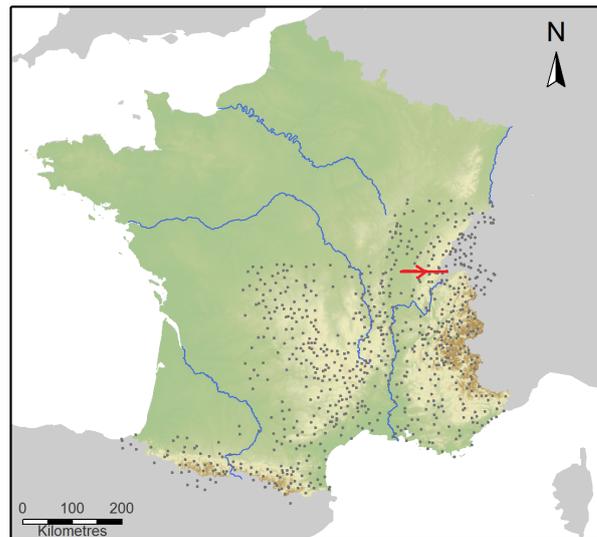
6. Les profils de pluies moyennes par type de temps ou de pluies moyennes annuelles sont issus des champ SPAZM.

permettent donc de prendre en compte une évolution des précipitations avec le relief dépendant de l'exposition des versants.

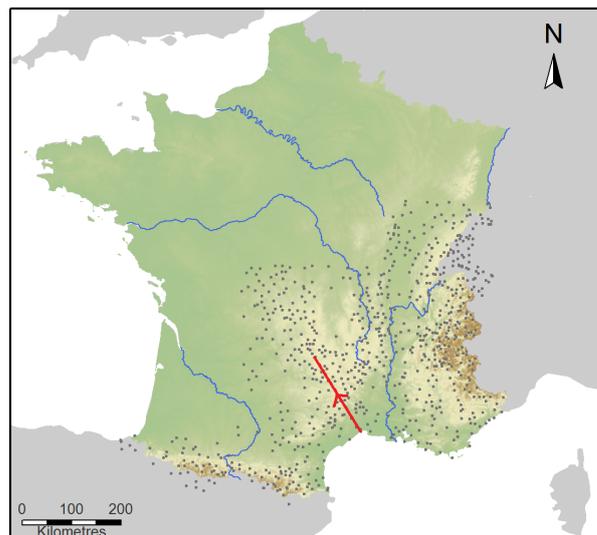
Les gradex de la distribution annuelle des postes présents sur ce transect ne sont pas corrélés au relief et semblent même avoir une tendance à diminuer avec celui-ci. Le gradex le plus fort est observé pour un poste dans le piémont du massif.

Le transect des Cévennes (figures III.14) permet d'étudier le comportement des cumuls annuels et des gradex de pluies pour un relief exposé aux flux de sud. Le cumul de type de temps 4 croît rapidement avec le relief puis chute une fois le premier relief important franchi. Le coefficient de détermination entre le cumul de type de temps 4 et le relief est quasi nul ($R^2 = 0.001$). Les cumuls liés aux flux océaniques ne sont pas significatifs.

Finalement, le gradex des précipitations est en grande partie expliqué par le gradex du type de temps 4 ($R^2 = 0.99$) qui évolue de façon analogue au cumul annuel de ce type de temps.

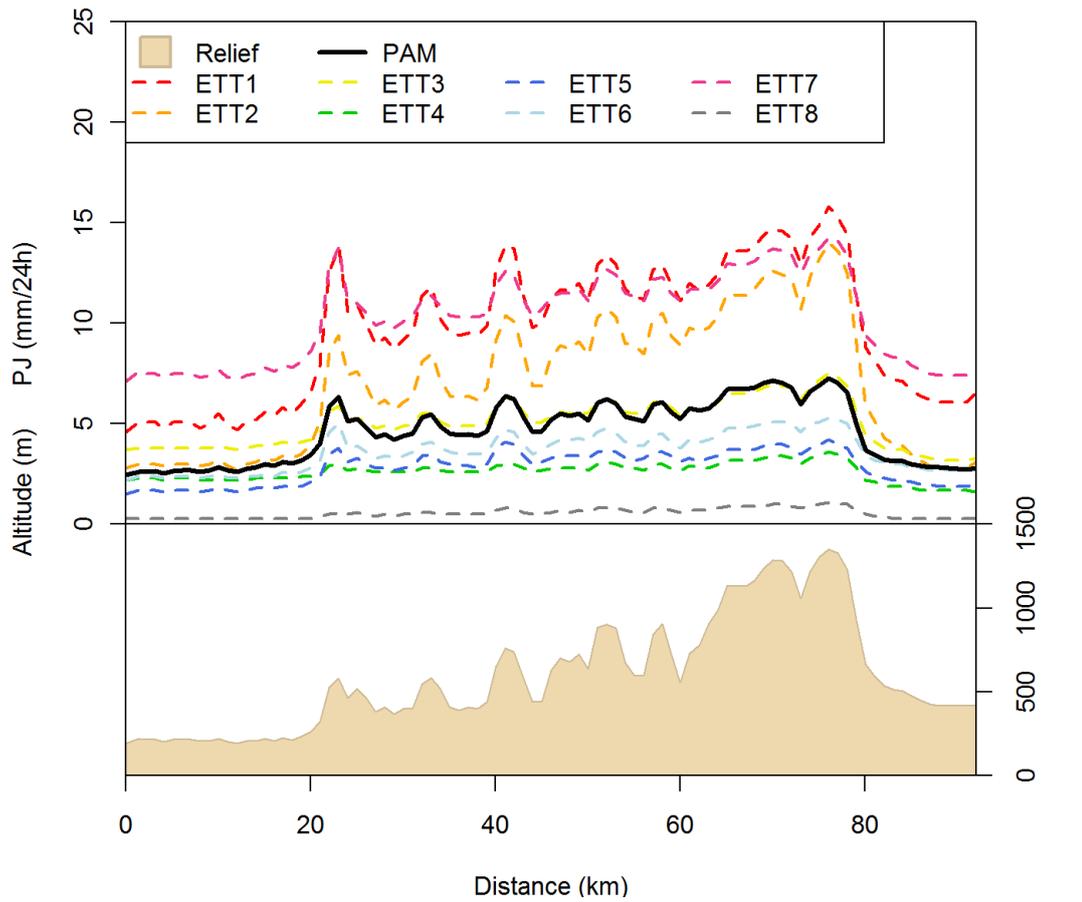


(a) Transect dans le Jura

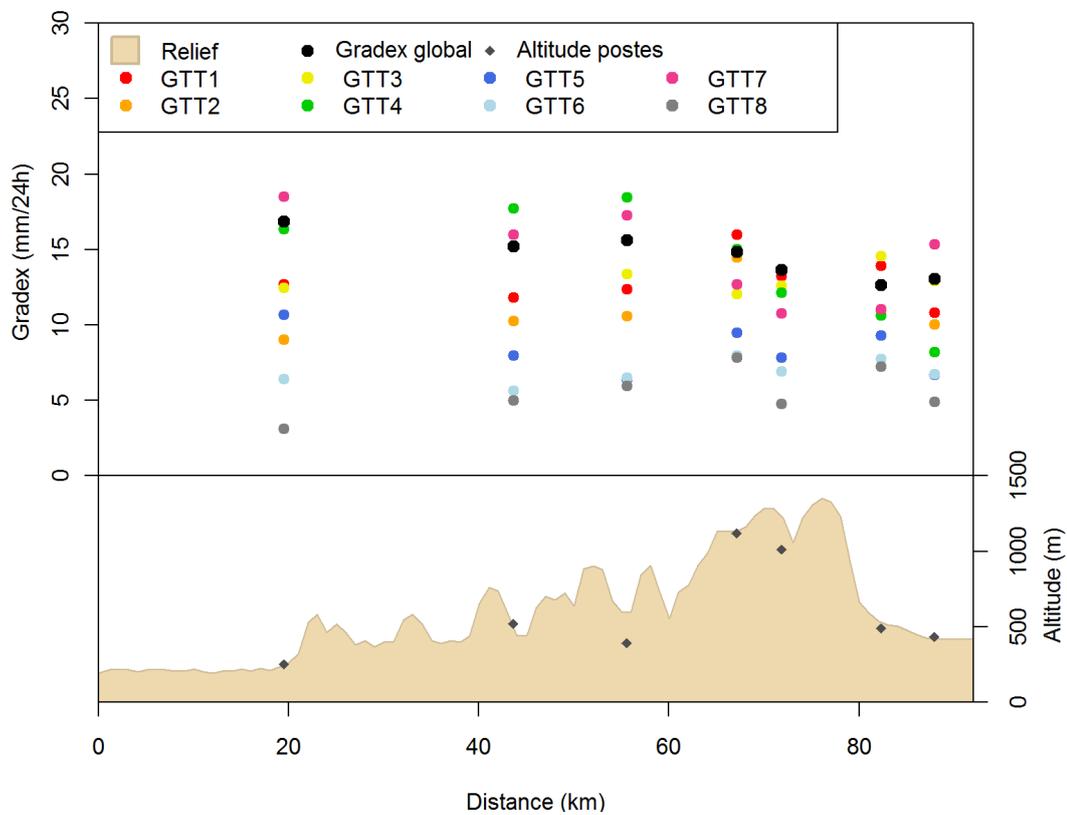


(b) Transect dans les Cévennes

FIGURE III.12 – Vue en plan des deux transects réalisés.

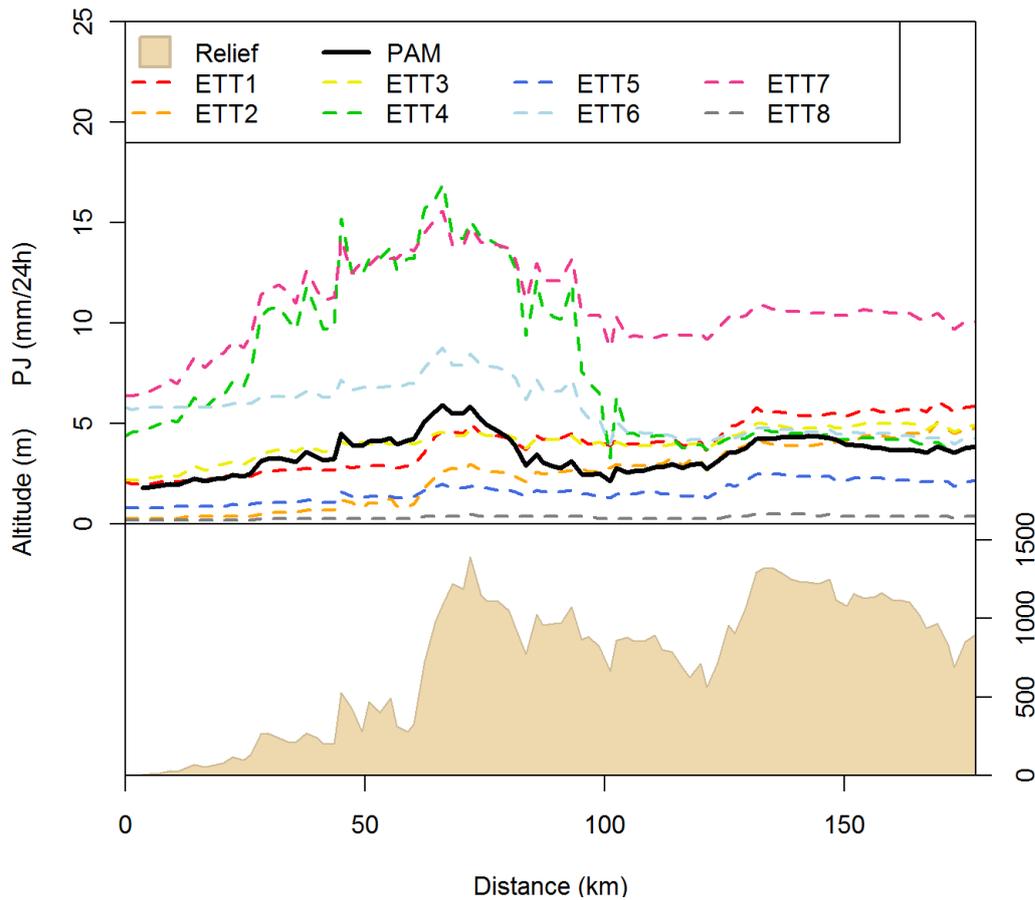


(a) Les pluies journalières moyennes

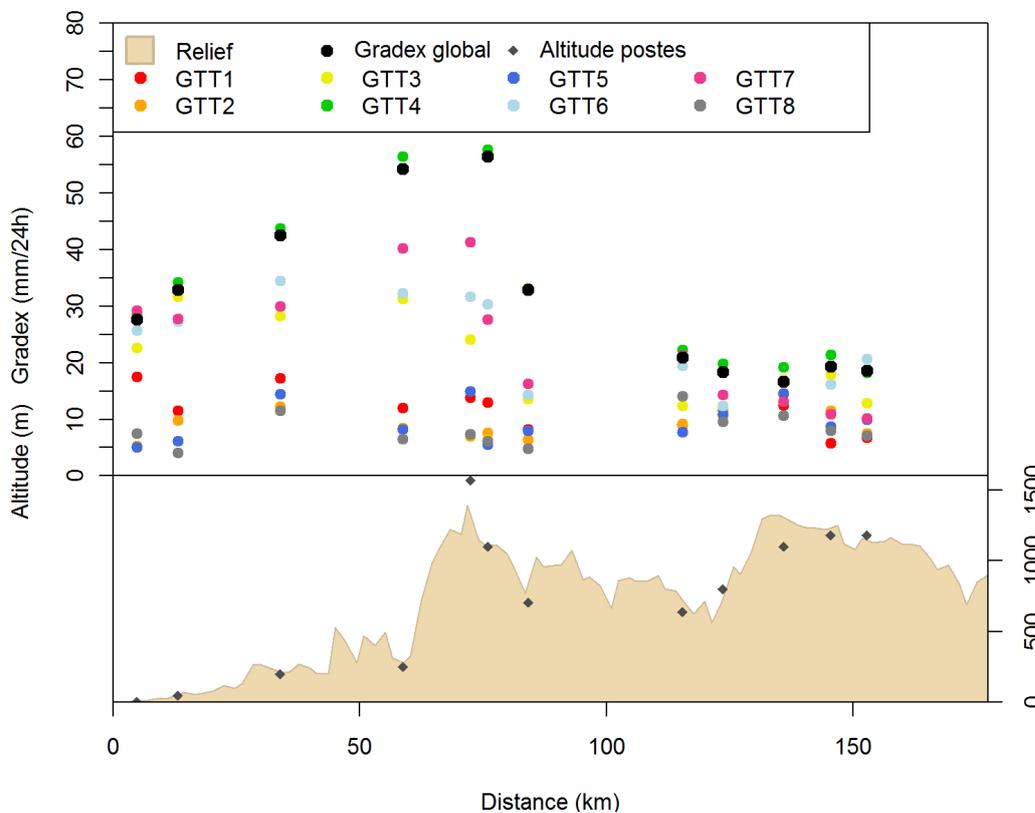


(b) Les gradex

FIGURE III.13 – Comportement des gradex et des pluies journalières moyennes par type de temps et globaux pour un transect du Jura.



(a) Les pluies journalières moyennes



(b) Les gradex

FIGURE III.14 – Comportement des gradex et des pluies journalières moyennes par type de temps et globaux pour un transect des Cévennes.

Bilan : l'altitude locale ne semble donc pas être une variable explicative pertinente pour les pluies extrêmes journalières. C'est souvent sur les premiers reliefs rencontrés et sur leur piémont que le risque de pluie est fort puis il est moins important sur les reliefs franchis par la suite. Il peut y avoir deux raisons à cela : l'événement pluvieux diminue en intensité au fur et à mesure de son avancée, ou bien l'événement pluvieux doit être significatif pour franchir le relief (effet de seuil) ce qui protège la zone située à l'arrière de ce relief⁷. L'exposition du relief semble par contre expliquer les différences des cumuls annuels ou de gradex selon les types de temps. D'ailleurs, Kieffer Weiss et Bois (2000) avaient montré qu'à des pas de temps supérieurs à 3 heures, les fortes précipitations sont plutôt contraintes par des influences d'ouest ou méditerranéennes et non plus par la topographie locale. Ce constat laisse penser que des variables explicatives utilisant les pluies moyennes par type de temps seront plus efficaces pour mettre en évidence une telle typologie des risques de pluies. En effet, elles permettront de donner un poids plus ou moins important à certains types de temps dans certaines régions, suivant que l'on a affaire à des influences méditerranéennes ou autres. Nous avons d'ailleurs déjà pu constater que le cumul annuel d'un type de temps est parfois bien corrélé au gradex d'un autre type de temps. Ces cumuls vont permettre d'apporter un descripteur de l'effet local du relief par type de temps en distinguant ainsi des zones soumises à des flux plutôt océaniques ou méditerranéens.

III. 2. 2.3 Spatialisation des gradex par type de temps

Structure spatiale des gradex par type de temps

Après avoir calculé les gradex de chaque type de temps pour chaque station du réseau utilisé en calage et avoir choisi les variables explicatives, la structure spatiale de chaque gradex est décrite à l'aide d'un semi-variogramme (un pour chaque type de temps de chaque saison).

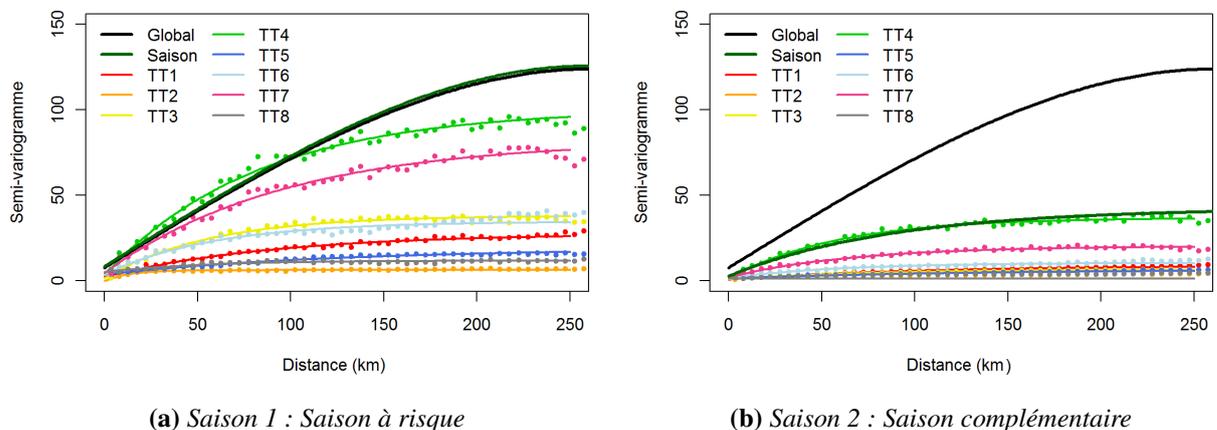


FIGURE III.15 – Semi-variogrammes par types de temps. En vert foncé le semi-variogramme du gradex de la saison (tous types de temps confondus) et en noir le semi-variogramme du gradex global. Les variables explicatives : Pluie annuelle et distance à la Méditerranée.

La détermination de structures bien distinctes justifie l'approche par types de temps et par saison (figure III.15). Les risques sont bien différents selon les types de temps et selon la saison avec une variance plus forte pour la saison à risque et pour les types de temps 4 et 7 qui sont ceux qui impliquent les événements avec les plus fortes précipitations sur le terri-

7. L'épuisement du potentiel précipitant et du potentiel d'ascendance d'une masse d'air explique que le relief est un impact plus ou moins déterminant.

toire français. Les variogrammes des gradex de la distribution annuelle (en noire) et de chaque distribution saisonnière (en vert foncé) sont également représentés à titre indicatif.

Krigeage des gradex par type de temps

À l'aide des variogrammes, seize cartes de gradex, une par type de temps et par saison, sont construites par krigeage. Avec une validation en Jackknife, une évaluation des estimations des différents gradex par type de temps est réalisée et permet de comparer les différentes variables explicatives envisagées pour leur krigeage. Afin de simplifier l'analyse des résultats, seuls les scores pour les gradex par type de temps de la saison à risque seront commentés, mais les résultats sont identiques sur la saison complémentaire.

Le tableau III.4 (a) présente les scores en Jackknife sur le réseau des 770 stations en ne prenant en compte aucune variable explicative pour les gradex par type de temps (krigeage ordinaire). Les gradex des types de temps 1, 4, 6 et 7 sont bien représentés, et ce même sans prise en compte de variables explicatives. Les gradex des types de temps 5 et 8 ont des NASH assez faibles. Cependant, leurs écarts-types des résidus sont plus faibles. Ce constat peut s'expliquer par le fait que les champs de ces deux variables ont des variations plus faibles ce qui impacte le score en NASH mais ne veut pas nécessairement dire que les estimations obtenue par krigeage sont mauvaises. De plus, le type de temps 8 n'est généralement pas le type de temps à risque en journalier et contraint donc peu la forme de la queue de la distribution. Une moins bonne représentation de ce type de temps n'est donc pas trop pénalisante sur le résultat final. De même, les pluies intenses du type de temps 5 sont peu observées dans notre zone d'étude puisqu'elles se rencontrent principalement dans la région des Vosges. Suite au travail sur les variables explicatives réalisé précédemment (page 65), deux jeux de variables ont été envisagés pour la régionalisation des gradex par types de temps.

Le tableau III.4 (b) synthétise les scores obtenus pour les gradex en réalisant un krigeage avec dérive externe tenant compte des variables explicatives (nommées VE dans le tableau) dites globales. La pluie annuelle moyenne améliore significativement la spatialisation du type de temps 2. Les variables globales, telles que la distance à la Méditerranée ou à l'Atlantique, ont plutôt tendance à dégrader les résultats.

Le tableau III.4 (c) est analogue au tableau III.4 (b) mais les variables explicatives considérées sont cette fois les cumulés annuels de pluie par type de temps. Les scores obtenus sont assez similaires à ceux sans variables explicatives. Cependant, les ébauches des types de temps 4 et 6 sont très corrélées à leurs gradex respectifs. Ainsi même si les scores sont très similaires, l'utilisation de ces deux variables explicatives semble crédible pour une régionalisation à faible densité de réseau. Nous notons également que les gradex de type de temps 2 sont moins bien expliqués par leur cumul annuel que par la pluie annuelle moyenne *PAM*.

Le dernier tableau III.4 (d) reprend les variables explicatives validées et considère le cumul annuel du type de temps 4 comme descripteur du gradex de type de temps 7 car nous avons pu constater une bonne corrélation entre ces deux variables (page 65). L'intérêt de variables explicatives peut sembler limité au vu des faibles différences entre les scores des krigeages avec ou sans variables. Cependant, les différentes corrélations et les études de transects laissent penser que ces variables explicatives sont pertinentes et prendront tout leur intérêt lorsque la densité d'information pour la réalisation des krigeages sera plus faible.

TT	Biais	NASH	Écart-type	VE	TT	Biais	NASH	Écart-type	VE
1	0.02	0.81	1.28	–	1	0.02	0.76	1.45	dAtl
2	0.04	0.67	1.78	–	2	0.02	0.78	1.46	PAM
3	0.02	0.69	1.42	–	3	0.02	0.59	1.62	dMed
4	0.03	0.85	2.62	–	4	0.02	0.85	2.62	dMed
5	0.03	0.64	1.54	–	5	0.03	0.64	1.54	–
6	0.02	0.80	1.93	–	6	0.02	0.73	2.26	dMed
7	0.03	0.75	2.17	–	7	0.03	0.63	2.66	dMed
8	0.03	0.56	1.39	–	8	0.03	0.56	1.39	–

(a) Sans variable explicative

TT	Biais	NASH	Écart-type	VE	TT	Biais	NASH	Écart-type	VE
1	0.02	0.81	1.30	ETT1	1	0.02	0.81	1.28	–
2	0.03	0.65	1.84	ETT2	2	0.02	0.78	1.46	PAM
3	0.02	0.67	1.46	ETT3	3	0.02	0.69	1.42	–
4	0.02	0.85	2.62	ETT4	4	0.02	0.85	2.62	ETT4
5	0.02	0.59	1.64	ETT5	5	0.03	0.64	1.54	–
6	0.02	0.80	1.96	ETT6	6	0.02	0.80	1.96	ETT6
7	0.02	0.64	2.62	ETT7	7	0.03	0.75	2.18	ETT4
8	0.03	0.53	1.45	ETT8	8	0.03	0.56	1.39	–

(b) Variables globales

(c) Variables types de temps

(d) Variables choisies

Tableau III.4 – Score du krigeage des gradex par type de temps (TT) en Jackknife avec les 770 stations (pour la saison à risque).

Une approche avec des régressions puis un krigeage des résidus a été testée à la place d'un krigeage avec dérive externe, mais sans donner de nettes améliorations.

Bilan : l'utilisation de variables explicatives améliore le krigeage de certains gradex par type de temps. La pluie annuelle moyenne permet une meilleure estimation pour le type de temps 2. La distance à la Méditerranée ne semble pas apporter beaucoup d'information pour les types de temps méditerranéens (3, 4, 6 et 7) par contre la pluie annuelle moyenne de type de temps 4 semble être un bon descripteur du risque méditerranéen et améliore sensiblement le krigeage du type de temps 7. Les cumuls de pluies par types de temps permettent de mieux décrire la structure spatiale des gradex par types de temps en donnant un impact plus ou moins important au relief selon l'orientation par rapport aux flux. Cependant, aucune variable explicative pertinente n'a été trouvée pour les types de temps 1 et 3. Les cumuls de pluies ou l'altitude ne sont pas toujours de bons descripteurs des champs de gradex.

III. 2. 2.4 Recomposition de la loi MEWP

Pour tout pixel du domaine d'étude, les huit gradex par types de temps de chacune des deux saisons sont fournis par les cartes obtenues par krigeage. Les autres paramètres – seuils à 70% et pourcentages d'occurrence des pluies centrales d'un type de temps et le pourcentage d'occurrence de toutes les pluies centrales – sont calculés sur la série de pluie journalière SPAZM au pixel considéré. La distribution MEWP au pixel est alors recomposée afin d'associer un temps de retour à toutes les valeurs de pluies supérieures au plus grand seuil des huit types de temps.

La reconstruction se déroule en trois étapes (figure III.16) :

- **Étape 1** : La sous-population de pluies centrales de la saison i et du type de temps j est notée (X_j^i) . La loi exponentielle pour la sous-population (X_j^i) est ajustée sur un échantillon sup-seuil de (X_j^i) avec un seuil μ_j^i égal au quantile 70% de (X_j^i) (estimé

avec la pluie journalière SPAZM). La population des pluies centrales sup-seuil du type de temps j et de la saison i est alors notée $Z_j^i = X_j^i - \mu_j^i$.

La fonction de répartition de (Z_j^i) est estimée par la loi exponentielle de paramètre d'échelle (ou gradex) GTT_j^i (obtenue par krigeage).

$$F_j^i(z) = 1 - \exp\left(-\frac{z}{GTT_j^i}\right) \quad (\text{III.4})$$

avec F_j^i , la distribution exponentielle du type de temps j de la saison i .

- **Étape 2 :** Comme les pluies centrales de la saison i sont l'union des pluies centrales de chaque type de temps de la saison, la distribution saisonnière est calculée par un mélange de loi des distributions de chaque type de temps de la saison j . Cette distribution saisonnière s'exprime alors :

$$F^i(z) = \sum_{j=1}^8 F_j^i(z) \cdot p_j^i \quad (\text{III.5})$$

avec F^i , la distribution saisonnière de la saison i et p_j^i le pourcentage d'occurrence relative du type de temps j dans la saison i (estimé avec la pluie journalière SPAZM).

- **Étape 3 :** La distribution annuelle finale est alors issue du mélange des deux distributions saisonnières et s'écrit :

$$F(z) = \sum_{i=1}^2 (F^i(z) \cdot p^i) \quad (\text{III.6})$$

avec F , la distribution annuelle et p^i le pourcentage d'occurrence relative de la saison i (fixé à 0.25 pour la saison à risque de trois mois et à 0.75 pour la saison complémentaire de neuf mois). Cet ajustement annuel sur la population des pluies centrales peut être exploité dans l'espace des maxima annuels en utilisant l'occurrence des pluies centrales p (estimé avec la pluie journalière SPAZM).

Bilan : La régionalisation de la distribution MEWP consiste à spatialiser les 51 paramètres afin de pouvoir recomposer une distribution en chaque point et de déterminer le quantile de pluie de n'importe quel temps de retour (en pratique, au delà de 10 ans). Les pourcentages d'occurrences des types de temps, des saisons et des pluies centrales et les seuils par types de temps sont des variables avec une structure spatiale moins variable spatialement et donc plus facilement régionalisable. De plus, la distribution des pluies est peu sensible à ces paramètres. Une grande partie des paramètres est finalement estimée grâce à la pluie SPAZM (35 paramètres sur les 51). Au contraire, les gradex font l'objet d'une étude plus approfondie, car leur structure spatiale est plus variable et leur impact sur la partie extrapolée de la distribution est important. Des variables explicatives, s'appuyant sur les cumuls de pluie par type de temps, ont notamment permis de prendre en compte un impact des reliefs différents selon les types de risque (méditerranéen ou océanique) et d'améliorer la spatialisation des gradex des différents types de temps.

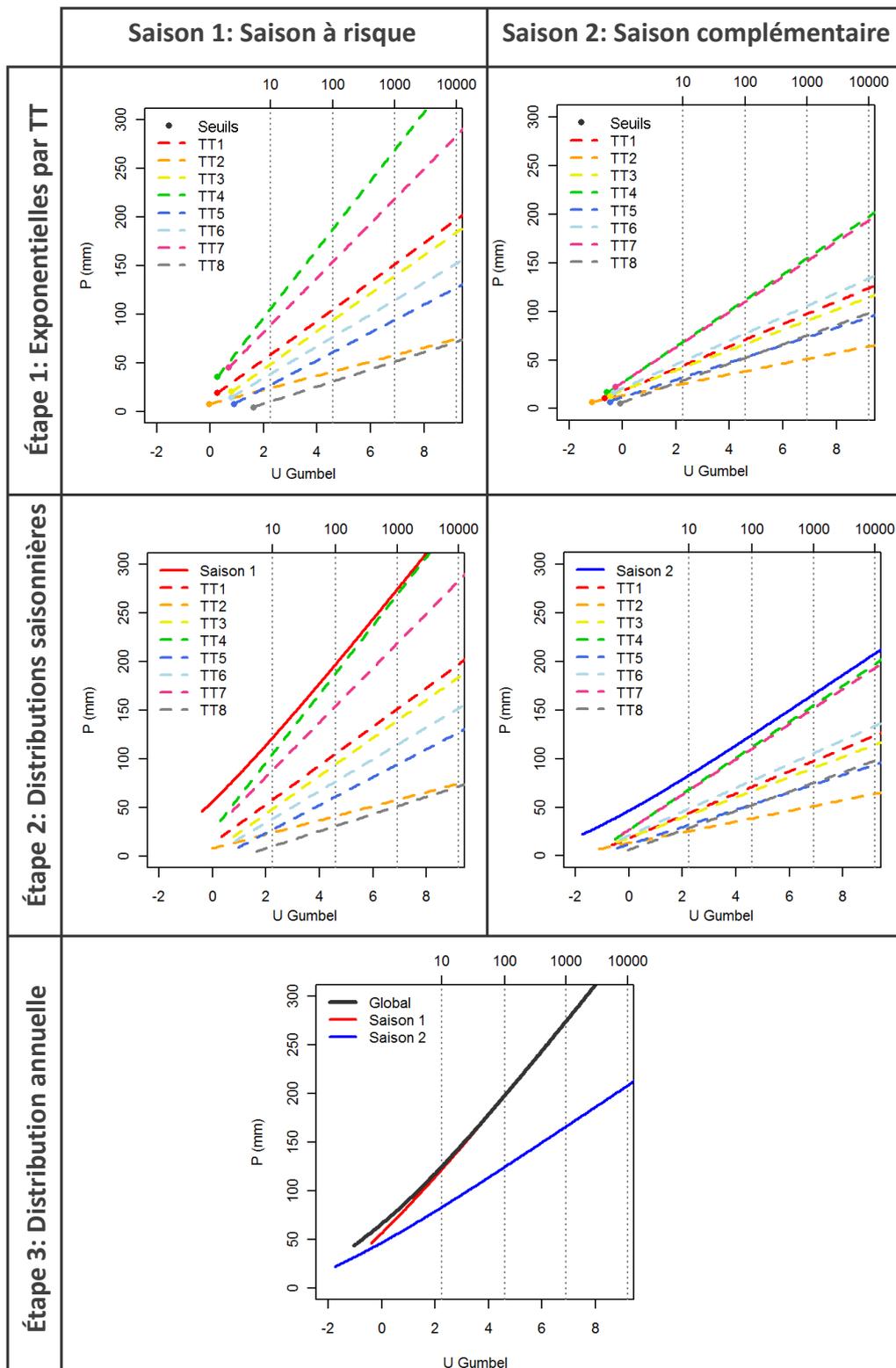


FIGURE III.16 – Exemple de reconstruction d’une loi MEWP.

III. 2. 3 Méthode 3 : Estimation des pluies extrêmes locales avec les pluies SPAZM

La régionalisation des paramètres de la distribution MEWP s'appuie en partie sur les pluies journalières SPAZM pour régionaliser certains paramètres ou comme variables explicatives des gradex par type de temps (avec les ébauches par types de temps). Se pose alors la question de savoir si un ajustement MEWP directement sur les pluies SPAZM fournit une estimation des quantiles élevés de précipitation acceptable. Les pluies journalières SPAZM sont disponibles de 1953 à 2005 en tout pixel de la zone d'étude. L'utilisation de cette information faciliterait donc la construction d'une carte des pluies extrêmes. Il suffirait de réaliser un ajustement MEWP sur les pluies SPAZM à chaque pixel, mais la validité d'une telle utilisation SPAZM n'a jamais été vérifiée puisque il n'a pas été conçu pour le calcul de valeurs extrêmes.

Au droit de chacun des 770 postes, deux séries de pluies journalières sont disponibles. La série des pluies mesurées au pluviomètre et la série de pluie SPAZM. Comme pour la régionalisation des paramètres (section III. 2. 2, page 60), les séries SPAZM utilisées sont issues d'un calage en Jackknife⁸. Une distribution MEWP est ajustée sur chacune des séries de pluies journalières sur la **période 1953-2005** avec la **classification par type de temps EDF2010**. Les saisons considérées sont toujours celles présentées dans la section II. 3. 1. Les estimations de quantiles de pluies de temps de retour 10, 100, 1000 et 10 000 ans entre l'ajustement local et l'ajustement sur les pluies SPAZM peuvent alors être comparées.

Temps de retour	Biais	NASH	Écart-type
2 ans	0.08	0.79	9.86
5 ans	0.08	0.80	12.44
10 ans	0.08	0.80	14.55
100 ans	0.09	0.81	22.29
1000 ans	0.09	0.80	30.75
10 000 ans	0.09	0.80	39.60

Tableau III.5 – Score en Jackknife sur les pluies SPAZM pour différents temps de retour.

Les ajustements sur les pluies SPAZM donnent une estimation cohérente des pluies extrêmes (figure III.18 : colonne de gauche et tableau III.5). Cependant, nous remarquons systématiquement une sous-estimation globale de l'ordre de 10% par rapport aux ajustements locaux. Afin d'améliorer les estimations des forts quantiles de pluies, une majoration des quantiles SPAZM pourrait être envisagée. La mise en place d'un modèle de correction pour cette sous-estimation demande de s'intéresser aux origines de celle-ci. Les séries de précipitations journalières observées et les séries construites avec SPAZM en Jackknife ont été comparées. Les ratios moyens des pluies annuelles et des variances des pluies journalières ainsi que des moyennes et des variances des maxima annuels sont calculés sur les 770 postes et reportés dans le tableau III.6.

	Rapport SPAZM/Observés
Pluie annuelle moyenne	1.01
Variance des pluies journalières	0.93
Moyenne des maxima annuels	0.94
Variance des maxima annuels	0.88

Tableau III.6 – Différents ratios entre les pluies SPAZM et les pluies journalières mesurées

8. Nous rappelons que tout le procédé SPAZM est réalisé en omettant la station à évaluer sauf pour les ébauches par types de temps qui ne sont pas reconstruites.

La méthode SPAZM n'est pas biaisée sur le cumul annuel des précipitations, mais elle sous-estime la variabilité des pluies. Ceci laisse penser que la méthode sous-estime les précipitations les plus importantes. Ce constat peut être vérifié sur les maxima annuels. La pluie SPAZM sous-estime de 6% les pluies maximales annuelles ainsi que leur variabilité.

Ce phénomène est intrinsèque à la méthode qui moyenne 10 postes voisins pour calculer une pluie au pixel. Plus généralement, toute méthode d'interpolation qui moyenne l'information provenant de différents postes a tendance à sous-estimer la variance des pluies. Ce problème est encore plus marqué sur les pluies extrêmes qui sont spatialement plus variables que les pluies plus courantes⁹. Le biais de 10% sur les forts quantiles a été calculé de façon globale sur toute la zone d'étude, mais pourrait être ajusté selon les zones en déterminant une structure spatiale des ratios des quantiles SPAZM sur les quantiles locaux. En regardant la carte des ratios III.17, une sous-estimation sur les plaines peut être notée mais le bilan est plus variable sur les reliefs de montagne avec des sous-estimations ou sur-estimations fortes selon les stations. **La correction à effectuer sur les quantiles élevés SPAZM ne semble donc pas avoir une structure spatiale bien définie et met juste en évidence une plus forte incertitude sur les zones de montagne** (car il y a moins de postes disponibles et une variabilité spatiale des précipitations plus forte).

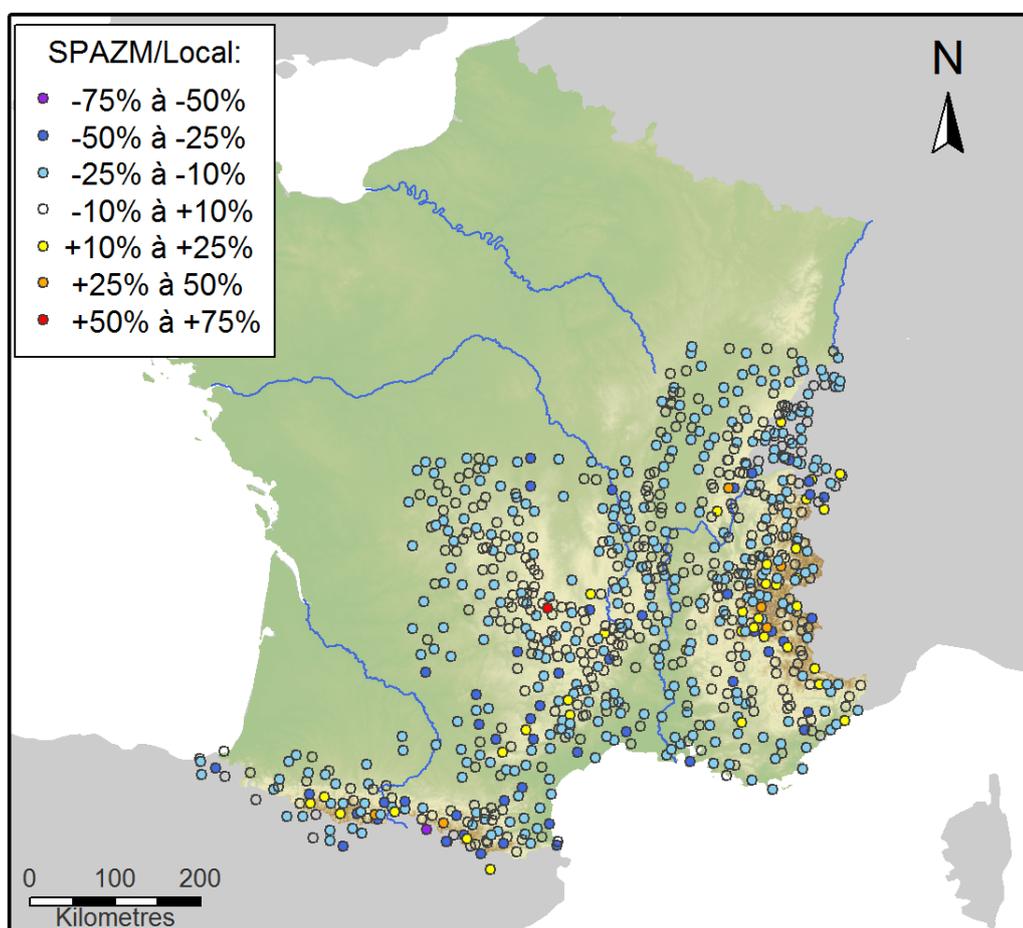


FIGURE III.17 – Carte des ratios entre les estimations de pluies millénales SPAZM en Jackknife et les estimations locales.

9. De la même manière, la fréquence des pluies nulles est sous-estimée par SPAZM.

Application d'une majoration quantile-quantile aux estimations SPAZM

La méthode SPAZM a été optimisée et validée dans le cadre d'analyse hydrologique. Ainsi, les pluies journalières SPAZM permettent une bonne spatialisation des cumuls de pluie, mais induisent une sous-estimation des pluies extrêmes. Dit différemment, les pluies interpolées par SPAZM permettent une bonne estimation du moment d'ordre 1 de la série de pluie, mais pas du moment d'ordre 2 (auquel les valeurs extrêmes sont très liées). Une majoration linéaire des estimations SPAZM à partir d'une correction quantile-quantile dépendant du temps de retour est alors proposée (figure III.18 et tableau III.7). La majoration est de 6% pour le temps de retour annuel, et atteint une valeur proche de 10% pour tous les temps de retour supérieurs à 10 ans. Cette majoration continue vise à avoir une majoration progressive évitant de créer des ruptures dans l'ajustement MEWP obtenu.

Temps de retour	Biais	NASH	Écart-type
2 ans	0.03	0.82	10.07
5 ans	0.01	0.84	12.63
10 ans	0.01	0.85	14.72
100 ans	0.00	0.85	22.47
1000 ans	0.01	0.85	30.91
10 000 ans	0.01	0.85	39.77

Tableau III.7 – Score en Jackknife sur les pluies SPAZM pour différents temps de retour avec une correction sur les quantiles estimés.

Cette correction permet d'améliorer les scores en NASH sans dégrader de façon significative les écarts-types des résidus. La méthode est par construction non biaisée sur les quantiles de temps de retour supérieurs à 10 ans. Un ajustement MEWP sur les pluies SPAZM avec une correction quantile-quantile fournit des estimations acceptables des quantiles élevés de précipitations.

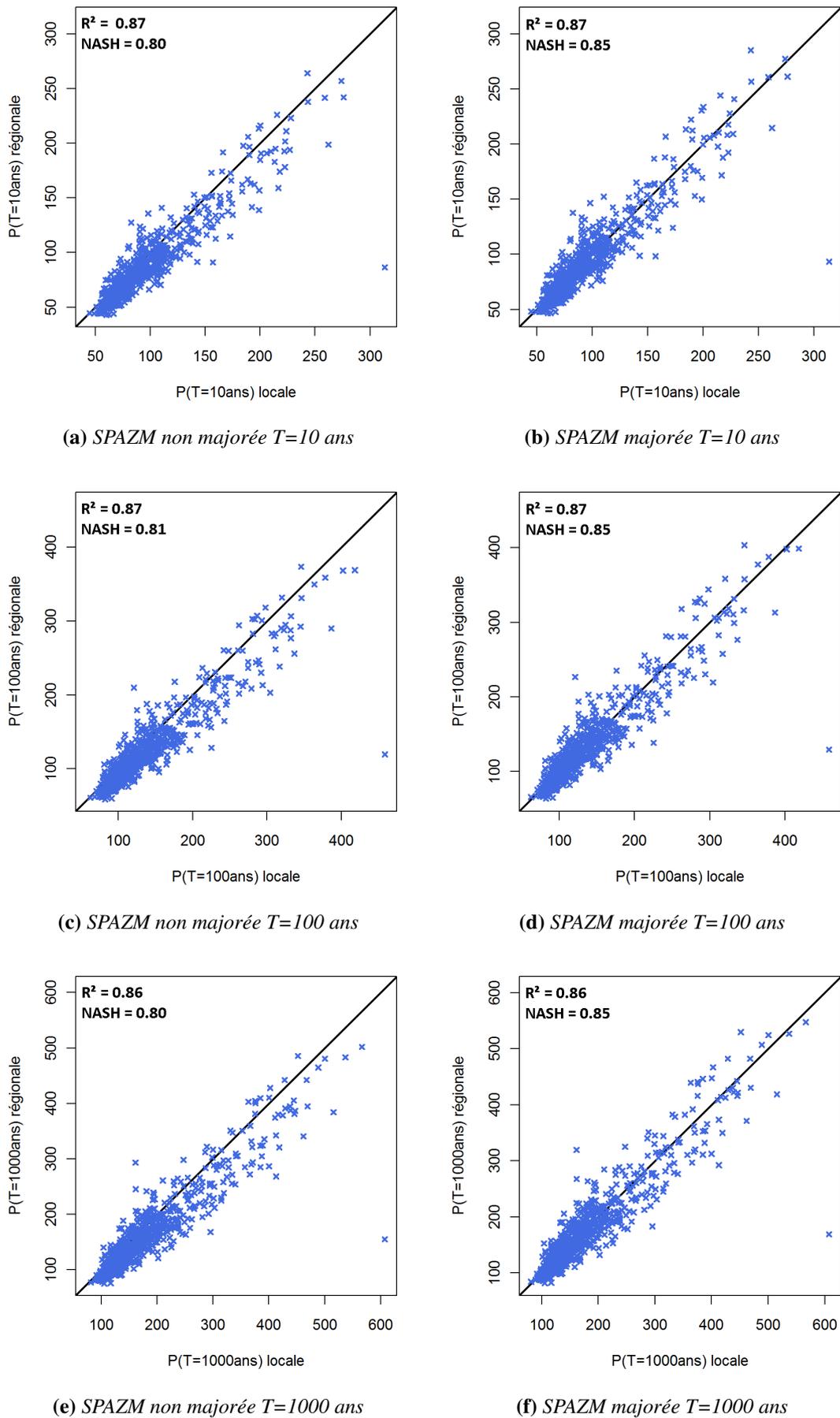


FIGURE III.18 – Scatter-plot des estimations des pluies (en mm) de temps de retour 10, 100 et 1000 ans avec et sans correction sur les quantiles estimés avec la pluie SPAZM en Jackknife.

Comment intégrer la correction dans la base SPAZM ?

La majoration de 10% sur les quantiles SPAZM corrige la sous-estimation moyenne observée avec SPAZM sur les pluies extrêmes. Cependant, SPAZM est un interpolateur exact des champs de pluies journaliers. Au droit d'un poste, la donnée SPAZM est égale à la donnée observée. Les pluies extrêmes estimées avec un ajustement MEWP sur la pluie SPAZM sont donc égales aux pluies extrêmes estimées sur la série observée (à condition d'avoir suffisamment de données observées au poste pour avoir un ajustement robuste) et la majoration de 10% n'a donc pas lieu d'être. L'idée est donc de chercher un moyen de corriger le biais sur les quantiles SPAZM en tout point du domaine en tenant compte du fait que la correction est nulle au droit des postes.

Il a été vu que le procédé de régionalisation de SPAZM permet de construire des séries de pluies journalières qui fournissent une bonne estimation du moment d'ordre 1, mais pas du moment d'ordre 2 qui est souvent sous-estimé. Or, les valeurs extrêmes estimées avec la loi MEWP – dont le comportement asymptotique est exponentiel – sont très liées au moment d'ordre 2 de l'échantillon.

Cette sous-estimation de la variabilité de la pluie SPAZM est due à plusieurs problèmes. Tout d'abord SPAZM est une méthode régionale combinant l'information d'un panier de plusieurs stations. Elle mélange l'information issue d'une zone plus ou moins étendue pour réaliser une estimation ponctuelle. Elle lisse donc la pluie, en particulier les fortes pluies, ce qui est à l'origine de la sous-estimation de la variance des précipitations. Ce phénomène est lié à deux sources principales :

- 1) la densité du réseau autour du point à estimer. Plus on dispose d'informations proches, meilleure sera l'estimation.
- 2) la variabilité des champs de pluies. Même si l'on dispose d'une bonne densité d'information autour du point à estimer, mais que tous les postes se trouvent dans une zone moins exposée que le pixel à estimer (un pixel sur le versant sud est des postes disponibles sur le versant nord), alors SPAZM risque de sous-estimer le risque. En effet, l'extrapolation avec les reliefs ne suffira peut-être pas pour capter le risque au poste.

Pour chaque poste, les écarts-types de la population des pluies centrales supérieures au quantile 70%¹⁰ sont calculés sur la série observée et sur la série SPAZM obtenue en Jackknife. Le choix de population considérée pour le calcul de l'écart-type permet d'être cohérent avec les populations utilisées pour l'ajustement de la distribution MEWP. Les ratios des écarts-types observés σ_{local} sur les écarts-types SPAZM σ_{SPAZM} sont alors comparés aux ratios des quantiles de pluies.

Sur le graphe III.19 (a), nous constatons que les ratios des écarts-types expliquent bien les ratios de pluies millénales. Dans certains cas, la variance de SPAZM est plus forte que la variance observée. Ceci est dû au fait que SPAZM extrapole les pluies avec le relief. Les quantiles élevés de SPAZM $PJ_{SPAZM}(T)$ peuvent alors être corrigés à partir du biais sur les écarts-types pour fournir une nouvelle estimation $PJ^*(T)$:

$$PJ^*(T) = \frac{\sigma_{local}}{\sigma_{SPAZM}} \cdot PJ_{SPAZM}(T) \quad (III.7)$$

Sur la figure III.19 (b), la courbe verte représente la densité de probabilité de ratios de la pluie de temps de retour 1000 ans estimée localement et avec SPAZM en Jackknife. La courbe bleue est la densité des ratios après correction des estimations SPAZM avec la valeur des écarts-types locaux (déterminés sur les séries mesurées aux postes). Le modèle de correction permet à

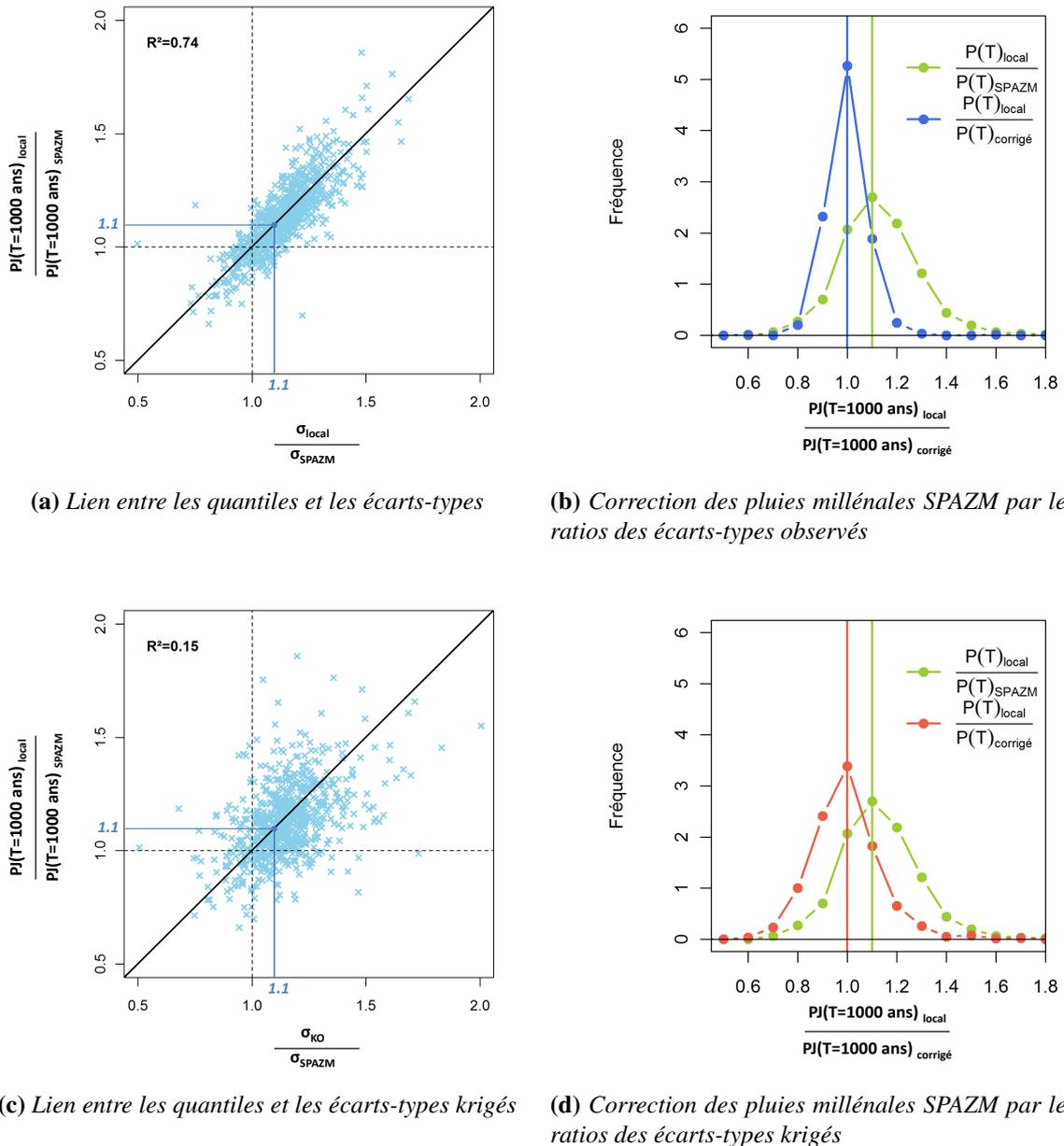


FIGURE III.19 – Correction des pluies de temps de retour 1000 ans estimées avec SPAZM.

la fois de corriger le biais de 10% sur les quantiles mais également de réduire considérablement la variabilité des ratios.

Pour pouvoir corriger les quantiles SPAZM, il faudrait donc estimer le ratio des écarts-types $\sigma_{local}/\sigma_{SPAZM}$. L'écart-type de la pluie SPAZM est connu en tout point, mais pas celui de la pluie observée. Nous proposons alors de construire une carte des écarts-types σ par un krigeage ordinaire. L'intérêt est double. Premièrement, le krigeage permet une estimation non biaisée du moment d'ordre 2 c'est-à-dire sans le biais de 10% sur les forts quantiles. Deuxièmement, le krigeage fournit une interpolation exacte de l'écart-type des pluies au droit des postes et donc une absence de correction sur les quantiles SPAZM à ces postes, comme souhaité.

Les figures III.19 (c) et III.19 (d) montrent effectivement que les écarts-types krigés σ_{KO} expliquent en partie la sous-estimation des quantiles SPAZM ($R^2 = 0.15$). Les écarts-types

10. Le seuillage au quantile 70% s'applique à la population des pluies centrales.

krigés σ_{KO} n'expliquent pas aussi bien les ratios sur les quantiles que les écarts-types locaux. Ils permettent néanmoins de débiaiser les estimations et de réduire légèrement la variabilité des erreurs. Cette correction est alors utilisable en chaque pixel en réalisant une correction quantiles-quantiles avec :

$$PJ^*(T) = \frac{\sigma_{KO}}{\sigma_{SPAQM}} \cdot PJ_{SPAQM}(T) \quad (\text{III.8})$$

Le scatter-plot III.20 montre la capacité à corriger les quantiles SPAQM en tenant compte de l'erreur faite sur la variabilité des pluies. Les points verts représentent les estimations de pluies millénales obtenues avec SPAQM (en Jackknife) sans aucune correction (donc biaisés de 10%). Les points bleus sont les quantiles SPAQM corrigés avec la formule (III.7), c'est-à-dire dans le cas idéal où l'écart-type local σ_{local} est connu en tout point. Dans ce cas, la correction débiaise et réduit nettement la variabilité des écarts entre les estimations régionales et locales. Cependant dans la réalité les valeurs σ_{local} ne sont connues qu'au droit des postes de mesure. Une estimation des écarts-types σ_{KO} par un krigeage ordinaire permet de proposer une estimation des écarts-types et donc une correction des quantiles élevés de SPAQM avec la formule (III.8). Cette correction permet de débiaiser les estimations, mais ne réduit pas de façon claire la variabilité des erreurs. Cependant, ceci montre qu'avec une estimation plus juste de l'écart-type σ en tout point du domaine, les estimations des quantiles élevés de SPAQM pourraient être améliorées de façon non négligeable.

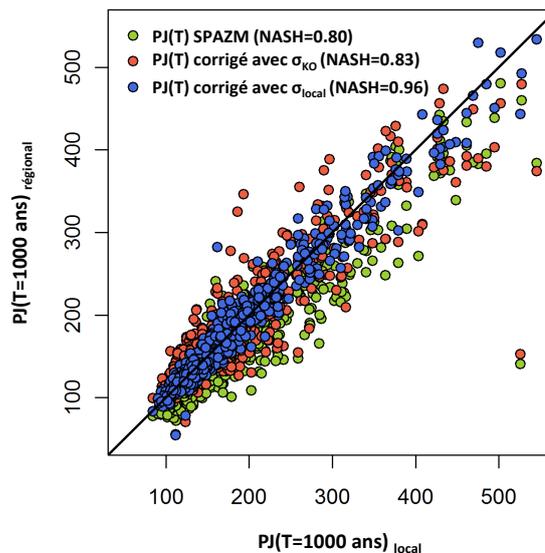


FIGURE III.20 – Scatter-plot des précipitations millénales locales et obtenues avec SPAQM en Jackknife (avec ou sans correction). Les écarts-types σ_{KO} sont issus d'un krigeage ordinaire en Jackknife.

Bilan : un ajustement MEWP directement sur la pluie SPAQM conduit à une sous-estimation des quantiles élevés de pluies. Une majoration de ces quantiles par un biais global de 10% calculé sur la base de données permet de débiaiser la méthode et d'améliorer globalement l'estimation des forts quantiles. Une correction en chaque point du domaine peut être proposée de façon à tenir compte du fait que SPAQM fournit des estimations exactes de forts quantiles au droit des points où des mesures sont disponibles. Elle consiste à s'intéresser à la sous-estimation des écarts-types des pluies par SPAQM afin de proposer un modèle de correction en tout point du domaine permettant de corriger le biais global de 10% sur les quantiles élevés.

III. 3 Comparaisons des approches de régionalisation

Les trois méthodes présentées pour régionaliser les pluies extrêmes journalières vont maintenant être comparées. Chaque approche envisagée a pour objectif de produire des cartes de précipitations journalières aux temps de retour 10, 100, 1 000 et 10 000 ans. La figure III.21 rappelle les grandes étapes de chacune des méthodes en distinguant les étapes utilisant de l'information locale (en rouge) du travail de régionalisation (en vert). Il apparaît que les trois méthodes envisagent le processus de régionalisation plus ou moins en amont de la construction des cartes ce qui permet de prendre en compte des informations différentes pour la régionalisation.

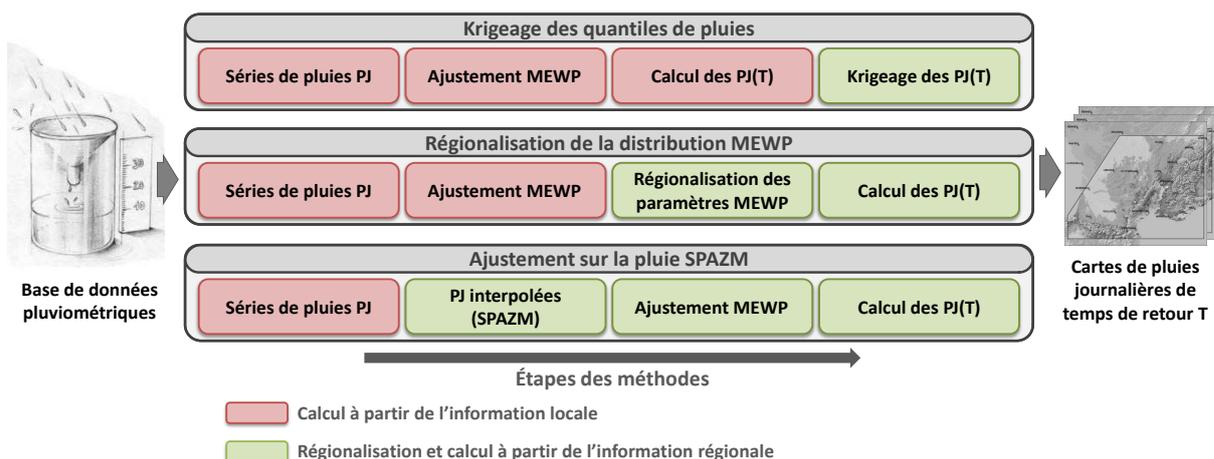


FIGURE III.21 – Les grandes étapes de chacune des méthodes.

Le tableau III.8 synthétise les variantes évaluées pour chaque méthode ainsi que la dénomination utilisée dans la suite pour chacune d'entre elles.

Principe	Variantes pour chaque méthode	Sigles
Krigage de pluies de temps de retour T (<i>section III. 2. 1, page 54</i>)	Krigage ordinaire sans aucune variable explicative	KO P(T)
	Krigage avec dérive externe pour prendre en compte le relief	KED P(T)
Régionalisation des paramètres de la distribution MEWP (<i>section III. 2. 2, page 60</i>)	Les gradex sont spatialisés avec des variables explicatives globales	MEWP Rég. V1
	Les gradex sont spatialisés avec des variables explicatives par types de temps	MEWP Rég. V2
Ajustement sur les pluies SPAZM (<i>section III. 2. 3, page 76</i>)	Sans correction	SPAZM V1
	Avec majoration quantile-quantile de 10%	SPAZM V2

Tableau III.8 – Les trois méthodes envisagées et leurs variantes.

Une première évaluation des ces trois approches a été faite sur le jeu de données de 770 postes. Elle vise notamment à tester les méthodes avec des réseaux de calage de moins en moins denses. Ce travail va permettre de faire ressortir des comportements différents selon les méthodes.

Une seconde évaluation, s'appuyant sur le protocole de validation du projet ANR ExtraFlo (2009-2013), a également été mise en place afin de comparer les méthodes à celles utilisées dans ce projet. L'objet du projet ExtraFlo était de procéder à une inter-comparaison des principales méthodes de prédétermination des pluies et crues extrêmes utilisées en France, de façon à mieux connaître leurs domaines respectifs d'application. Un effort particulier a été réalisé pour la constitution de jeux de données de référence (numérisation de longues séries, jeux régionaux), afin de porter un diagnostic sur les points forts et les lacunes de chaque approche et leur sensibilité à la densité d'information. Ce travail visait à faire évoluer et progresser les méthodes de prédétermination des pluies et crues extrêmes, à étudier leur capacité à prendre en compte les effets attendus du changement climatique, et à mettre à disposition des acteurs concernés par la prévention du risque d'inondation, toute une série d'outils pratiques d'estimation en fonction des caractéristiques hydrologiques des bassins, mais aussi des informations disponibles.

Les membres du projet étaient : Irstea (Lyon et Aix), Météo-France (Toulouse), HydroSciences (Montpellier), EDF-DTG (Grenoble), GéoSciences (Montpellier). Plus d'informations sur ce projet sont disponibles dans les rapports accessibles depuis le site internet <http://extraflo.irstea.fr>.

III. 3. 1 Évaluation sur le jeu de 770 stations

L'objectif de cette évaluation est de tester chaque méthode pour l'estimation de quantiles remarquables de précipitations (pour les temps de retour 10, 100, 1000 et 10 000 ans) et en particulier l'influence de la densité d'information utilisée en calage (avec les réseaux présentés en III. 1. 2, page 50). Les scores qui seront présentés pour la méthode SPAZM sont issus d'une validation en Jackknife (calage sur tout le réseau SPAZM de 2201 postes sauf sur la station cible) et seront donc à considérer différemment de ceux obtenus pour les autres méthodes (calage sur $x\%$ du réseau de 770 postes et validation sur les $100 - x\%$). Le choix de calculer les scores SPAZM de cette façon a été motivé par le fait que :

- 1) SPAZM permet d'assimiler un grand nombre de postes avec des données journalières. En effet, contrairement aux deux autres méthodes qui nécessitent des postes avec des séries de données suffisamment longues pour réaliser un ajustement des pluies, SPAZM peut traiter des postes avec peu de données. Recaler SPAZM sur le réseau de 770 postes, comme les autres méthodes, reviendrait à ne pas exploiter un des avantages de SPAZM qui est de pouvoir utiliser plus d'information pour la régionalisation.
- 2) Une deuxième raison est la difficulté d'avoir à refaire la totalité des étapes de SPAZM : construction des ébauches par types de temps, déformation des ébauches et krigeage des résidus. La reconstruction d'ébauches serait en effet assez lourde à mettre en place.

Dans un premier temps, l'idée était plutôt de voir comment se comportaient les autres méthodes par rapport à une estimation des pluies extrêmes avec SPAZM qui assimile un réseau plus dense et de juger ainsi de leur intérêt.

III. 3. 1.1 Les scores par méthodes

Sur la figure III.22, un scatter-plot des pluies de temps de retour 1000 ans pour les 388 postes utilisés en validation (50% du réseau en calage) est disponible pour chaque méthode. Il apparaît que seules les méthodes exploitant l'information SPAZM ne sous-estiment pas les plus fortes valeurs de pluies millénales. Les cumuls de précipitations annuels par type de temps (utilisés en tant que variables explicatives) ou l'utilisation des pluies journalières SPAZM permettent de prendre en compte des gradients orographiques locaux par type de temps. L'information SPAZM améliore nettement l'estimation des plus fortes valeurs pour un temps de retour donné en permettant une meilleure prise en compte du relief. Nous ne montrons ici que les graphes pour la pluie millénale et pour un réseau de calage à 50% afin d'alléger la présentation des résultats. Les conclusions sont similaires pour les temps de retour 10, 100, 1000 et 10 000 ans mais les performances des méthodes dépendent par contre de la densité du réseau utilisé en calage.

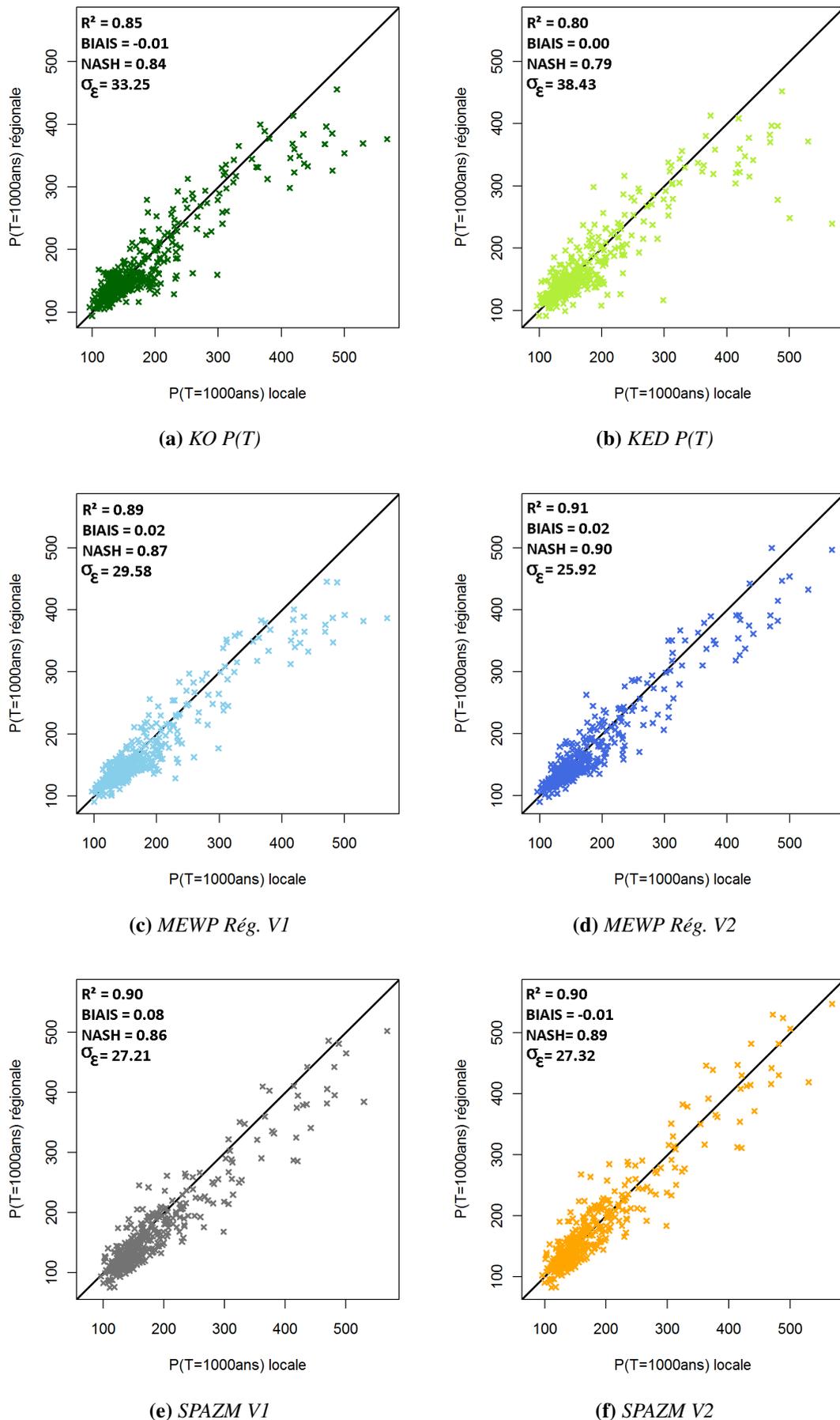


FIGURE III.22 – Scatter-plot entre les estimations locales et régionales pour la pluie millénale sur un réseau de validation de 50%.

III. 3. 1.2 Les scores par densité de réseau

Les figures III.23, III.24, III.25 et III.26 permettent de visualiser les scores (biais, NASH, écarts-types des résidus) selon la densité de postes utilisés en calage et pour quatre périodes de retour (10, 100, 1000 et 10 000 ans).

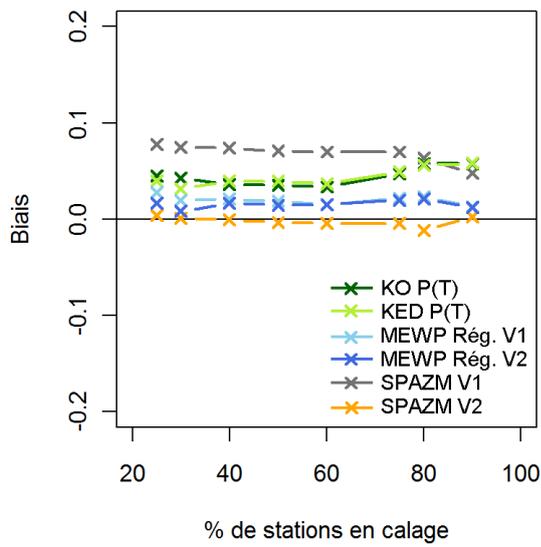
Les krigeages des quantiles (KO P(T) et KED P(T)) sont les deux méthodes qui donnent les moins bonnes estimations. Les critères de NASH sont plus faibles et les écarts-types plus élevés pour tous les temps de retour et toutes les densités de réseaux.

La régionalisation de MEWP par krigeage des gradex donne déjà de meilleurs résultats. Les biais sont très proches de ceux du krigeage des quantiles et sont tout à fait acceptables. Les NASH et les écarts-types sont nettement améliorés par rapport au krigeage de quantiles et cette amélioration est d'autant plus marquée que la densité de réseau utilisée en calage diminue (en particulier pour les temps de retour 10 et 100 ans). Ceci est dû au fait qu'une partie des paramètres de la loi MEWP (seuils, occurrences) sont issus de la pluie SPAZM qui s'appuie sur un réseau dense. Le bas de la distribution MEWP est alors mieux contraint et la pluie de temps de retour 10 ans est alors bien estimée, et ce même à faible densité de réseau. En dessous de 50% du réseau utilisé en calage, les scores décroissent plus rapidement.

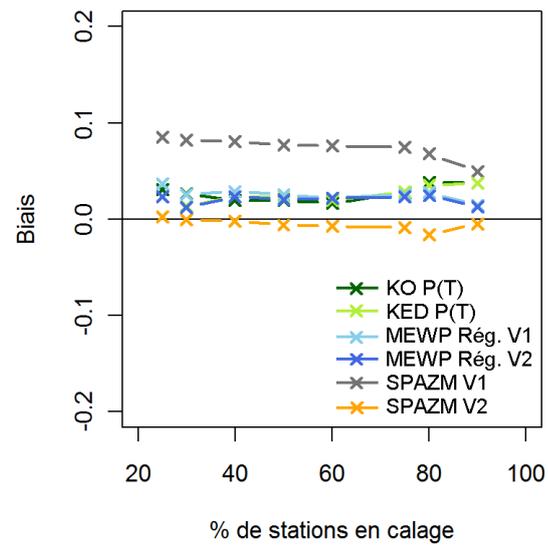
Remarque : Les scores avec 90% des stations en calage chutent légèrement. Ce comportement inattendu s'explique par la faible taille de l'échantillon de validation (moins de 80 stations).

Deux jeux de variables explicatives pour régionaliser les gradex par type de temps ont été envisagés dans cette évaluation. Le premier s'appuie sur la distance à la Méditerranée et la pluie annuelle (MEWP Rég. V1). Le deuxième utilise aussi les pluies annuelles moyennes par type de temps (MEWP Rég. V2). L'étude faite sur l'influence du relief sur les pluies extrêmes (section III. 2. 2, page 67) laisse penser que les pluies moyennes par type de temps permettraient une meilleure explication du gradex. La méthode MEWP Rég. V2 donne en effet de meilleurs scores que la méthode MEWP Rég. V1 en NASH et en écarts-types des résidus.

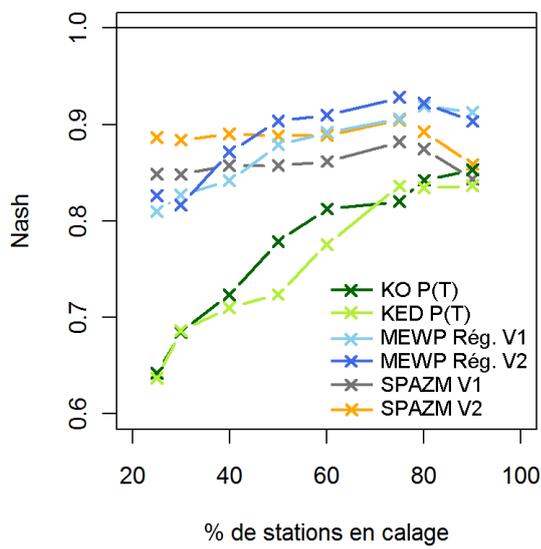
Il reste à comparer les ajustements MEWP sur les séries SPAZM avec ou sans majoration. Nous rappelons que SPAZM a été calé en Jackknife et que les scores obtenus pour cette méthode ne sont donc pas comparables à ceux des autres méthodes. Cependant, ils permettent de juger si les méthodes mises en place permettent d'obtenir un réel progrès par rapport à un ajustement MEWP sur SPAZM. La majoration quantile-quantile appliquée aux estimations des quantiles élevés de SPAZM V1 améliore les performances. La régionalisation de MEWP par le krigeage des gradex (MEWP Rég. V2) et un ajustement corrigé sur SPAZM (SPAZM V2) sont deux méthodes avec de bonnes performances et ce même avec une densité de poste en calage plus faible. Cependant la méthode MEWP Rég. V2 est lourde à mettre en place et ne permet pas un gain significatif par rapport à la méthode SPAZM V2.



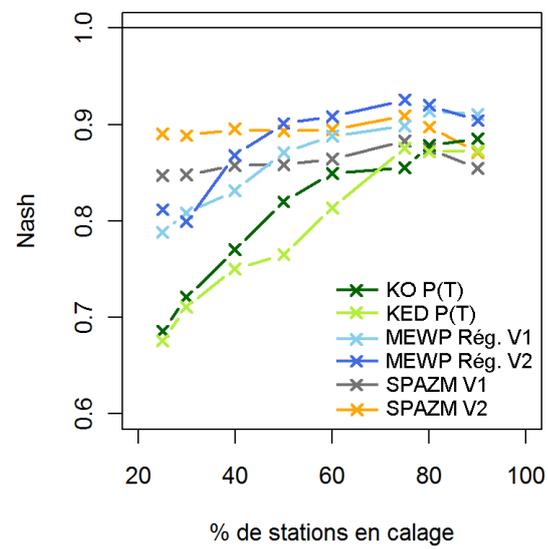
(a) Biases



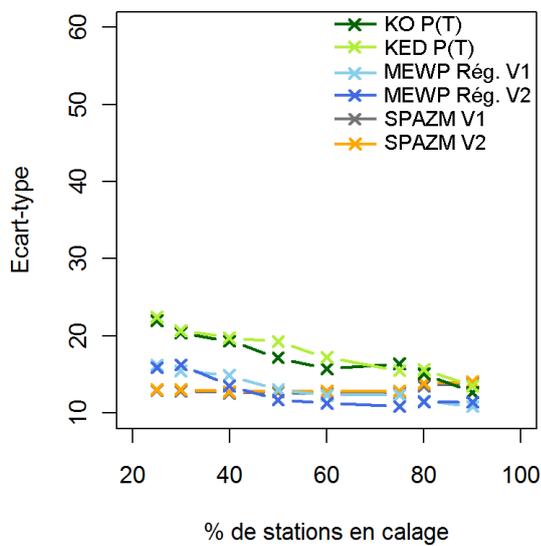
(a) Biases



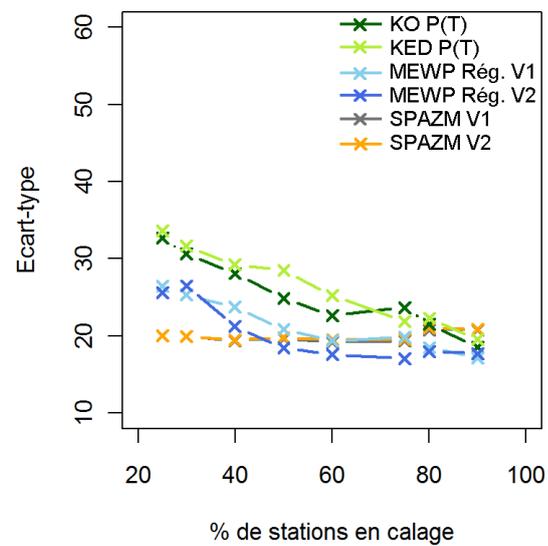
(b) NASH



(b) NASH



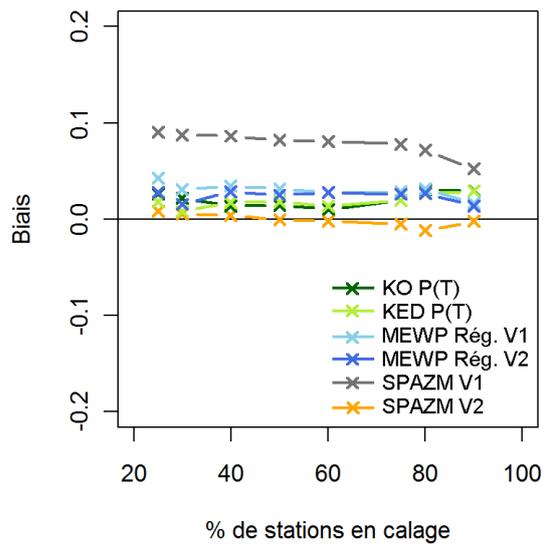
(c) Écart-types des résidus



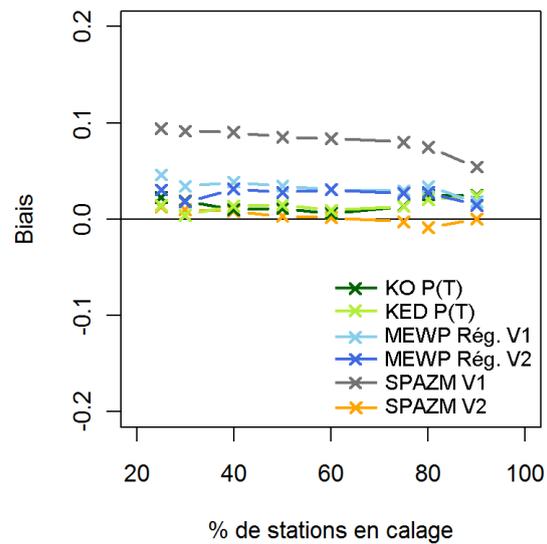
(c) Écart-types des résidus

FIGURE III.23 – Scores pour chaque méthode sur la pluie décennale en fonction de la densité de réseau utilisée.

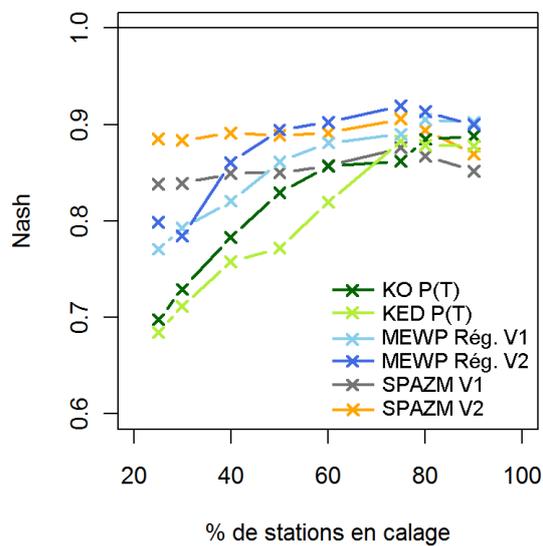
FIGURE III.24 – Scores pour chaque méthode sur la pluie centennale en fonction de la densité de réseau utilisée.



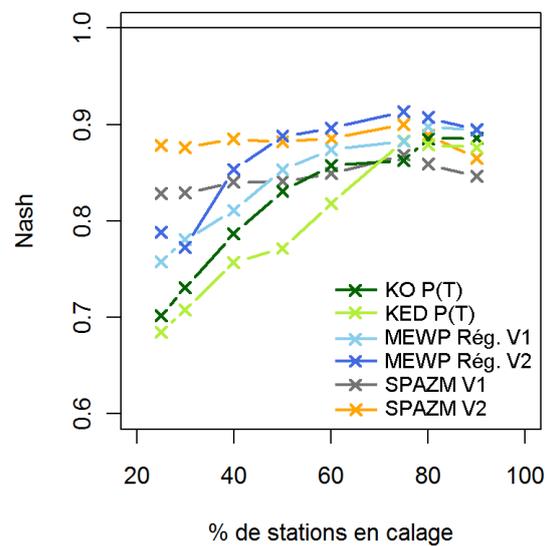
(a) Biases



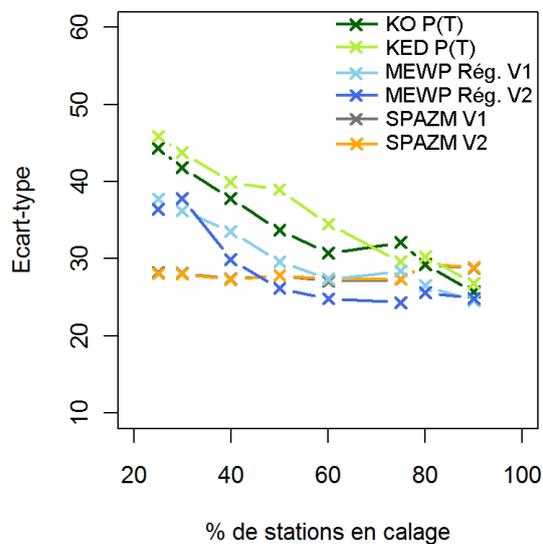
(a) Biases



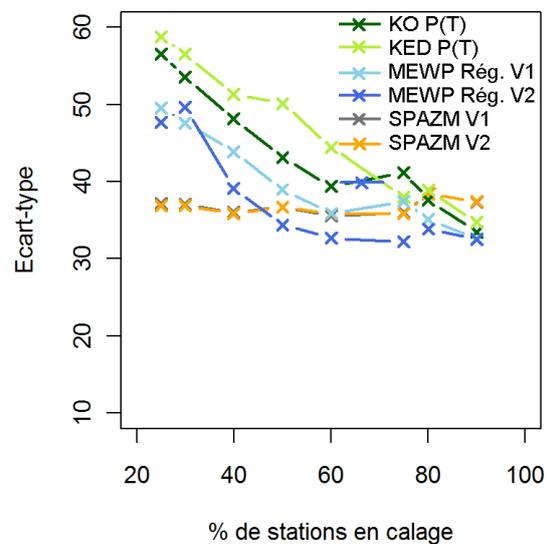
(b) NASH



(b) NASH



(c) Écart-types des résidus



(c) Écart-types des résidus

FIGURE III.25 – Scores pour chaque méthode sur la pluie millénale en fonction de la densité de réseau utilisée.

FIGURE III.26 – Scores pour chaque méthode sur la pluie décennale en fonction de la densité de réseau utilisée.

Bilan : ce travail de comparaison a permis d'évaluer différentes approches pour l'estimation des quantiles remarquables de pluies à différents temps de retour. Il est apparu que l'utilisation d'information issue de SPAZM (ébauches par type de temps ou séries journalières) améliore l'estimation des forts quantiles de précipitation.

La spatialisation de MEWP par krigeage des gradex en utilisant les pluies moyennes par type de temps (MEWP Rég. V2) est la méthode ayant les meilleures performances, mais elle est lourde à mettre en place. Au contraire, un ajustement sur les pluies SPAZM avec majoration (SPAZM V2) a l'avantage d'être facilement utilisable et de fournir des scores presque aussi bons que la méthode précédente. De plus, la méthode SPAZM permet d'utiliser toutes les données de pluies journalières disponibles pour l'interpolation des pluies, contrairement aux autres méthodes qui ont besoin de postes disposant de longues séries de données. Dans une optique d'utilisation en contexte non jaugé, cette capacité d'assimilation de toutes les données est un atout.

Enfin, **l'utilisation de SPAZM avec une correction quantiles-quantiles pour l'estimation des pluies extrêmes semble être un bon compromis entre simplicité de mise en place et qualité des estimations.** De plus, les estimations SPAZM pourraient encore être améliorées en consolidant le modèle de correction proposé dans la section III. 2. 3 (page 80), notamment en travaillant sur la construction d'une carte des écarts-types des pluies. La méthode d'ajustement sur les pluies SPAZM, va maintenant être appliquée avec le protocole de validation du projet ExtraFlo.

III. 3. 2 Évaluation avec le protocole ExtraFlo

Lors de la comparaison des méthodes sur le jeu de 770 stations, un ajustement MEWP majoré sur les pluies SPAZM est apparu être un moyen simple, avec de bonnes performances pour estimer les forts quantiles de pluies et capable d'exploiter plus d'information pour son calage que les deux autres approches. Les jeux de données et protocoles mis à disposition par le projet ExtraFlo vont permettre d'évaluer cette méthode avec celles déjà présentes dans le projet.

III. 3. 2.1 Protocole d'évaluation

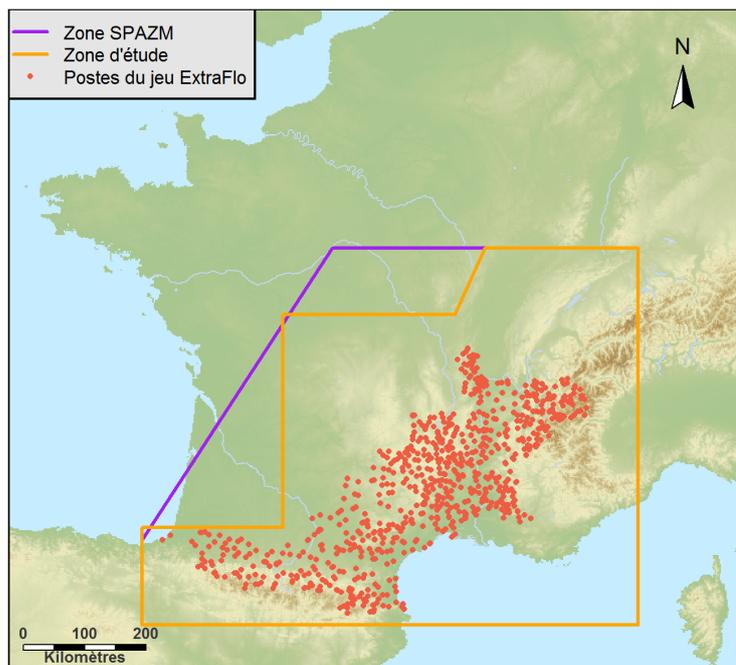


FIGURE III.27 – Jeu de données pour les méthodes locales-régionales.

Le protocole utilisé pour cette évaluation est celui de l'action III.5 (ANR ExtraFlo III.5, 2013) visant à comparer les méthodes purement locales, purement régionales et locales-régionales pour la prédétermination des pluies. Le jeu de données est celui présenté en figure III.27. Il regroupe 822 stations avec 411 stations ayant de 2 à 34 ans de données et les 411 stations restantes ayant plus de 46 ans de données.

Calage des méthodes locales-régionales

Une méthode est dite locale-régionale lorsqu'elle combine à la fois une information régionale et une information locale. Le calage de ces méthodes nécessite donc un calage pour la partie régionale et un autre pour la partie locale. Dans le cadre d'ExtraFlo (ANR ExtraFlo III.5, 2013), un protocole de calage-validation a été défini afin d'évaluer les méthodes pour l'estimation des pluies extrêmes en justesse et en robustesse.

Un modèle est juste lorsqu'il estime la « bonne » probabilité de dépassement des valeurs élevées (concrètement cette qualité ne peut être évaluée que par rapport aux observés).

Un modèle robuste est un modèle qui fournit des estimations similaires lorsqu'une faible perturbation est introduite dans les données. Cette propriété est particulièrement importante dans le cas de l'extrapolation des quantiles élevés afin d'éviter une estimation trop sensible à l'effet d'échantillonnage.

La décomposition calage-validation pour l'évaluation de la justesse sera effectuée de la manière suivante (figure III.28) :

- 50% des stations sont utilisées pour le calage de l'information régionale (vert). Toutes les années disponibles sont utilisées.
- Sur le reste des stations, n années sont choisies aléatoirement (indépendamment pour chaque station) et sont utilisées pour le calage de l'information locale (bleu). À noter que le choix d'un nombre fixe d'années a été fait plutôt qu'un pourcentage des années disponibles. Le poids de la donnée locale pourra être évalué par rapport à l'information régionale en faisant varier n (e.g. $n = 5, 10, 20$ ans)
- La validation sera effectuée sur les années complémentaires (rose) des 50% de stations utilisées en calage local.

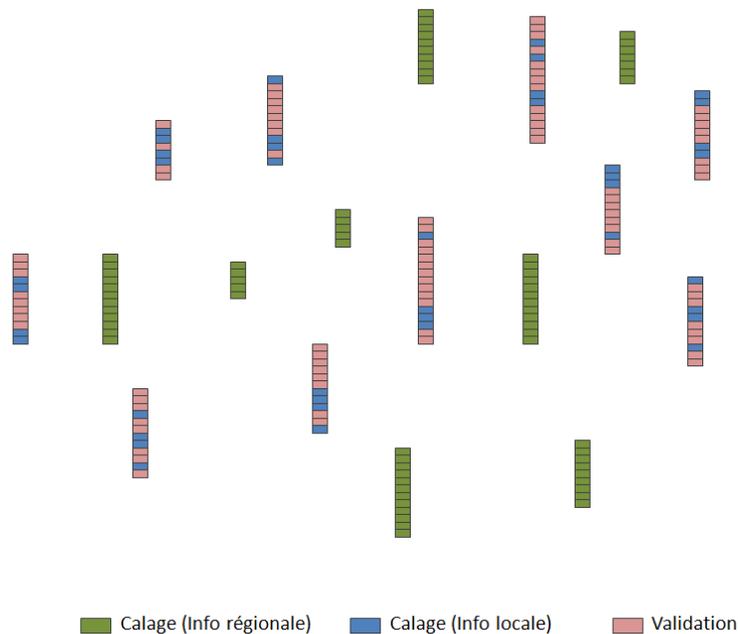


FIGURE III.28 – *Décomposition calage/validation pour les méthodes locales-régionales.*

Le mode de décomposition adopté permet d'inclure dans l'inter-comparaison les méthodes purement locales (uniquement calées sur l'information locale) et les méthodes purement régionales (uniquement calées sur l'information régionale).

Afin d'évaluer la robustesse, trois modes de décomposition supplémentaires ont été envisagés.

La première décomposition vise à évaluer la robustesse à l'information régionale (figure III.29) :

- 33% des stations sont utilisées pour constituer un premier jeu de calage de l'information régionale (R1) (toutes les années disponibles sont utilisées).
- 33% des stations sont utilisées pour constituer un second jeu de calage de l'information régionale (R2) (toutes les années disponibles sont utilisées).
- Sur les 33% de stations restantes, n années sont choisies au hasard et sont utilisées pour le calage de l'information locale (L).
- Finalement, on dispose de deux calages $C1=R1+L$ et $C2=R2+L$. Autrement dit, les deux jeux de calage utilisent des informations régionales différentes, mais avec la même information locale.

La seconde décomposition vise à évaluer la robustesse à l'information locale (figure III.30) :

- 33% des stations sont utilisées pour constituer le jeu de calage de l'information régionale (R) (toutes les années disponibles sont utilisées).
- Sur les 66% de stations restantes, n années sont choisies au hasard et sont utilisées pour un premier jeu de calage de l'information locale (L1). On tire n autres années pour constituer un second jeu de calage de l'information locale (L2).
- Finalement, on dispose de deux calages $C1=R+L1$ et $C2=R+L2$. Autrement dit, les deux jeux de calage utilisent des informations locales différentes, mais avec la même information régionale.

La dernière décomposition vise à évaluer la robustesse générale des méthodes (figure III.31) :

- 33% des stations sont utilisées pour constituer un premier jeu de calage de l'information régionale (R1) (toutes les années disponibles sont utilisées).
- 33% des stations sont utilisées pour constituer un second jeu de calage de l'information régionale (R2) (toutes les années disponibles sont utilisées).
- Sur les 33% de stations restantes, n années sont choisies au hasard et sont utilisées pour un premier jeu de calage de l'information locale (L1). On tire n autres années pour constituer un second jeu de calage de l'information locale (L2).
- Finalement, on dispose de deux calages $C1=R1+L1$ et $C2=R2+L2$. Autrement dit, les deux jeux de calage diffèrent à la fois par les informations locales et régionales utilisées.

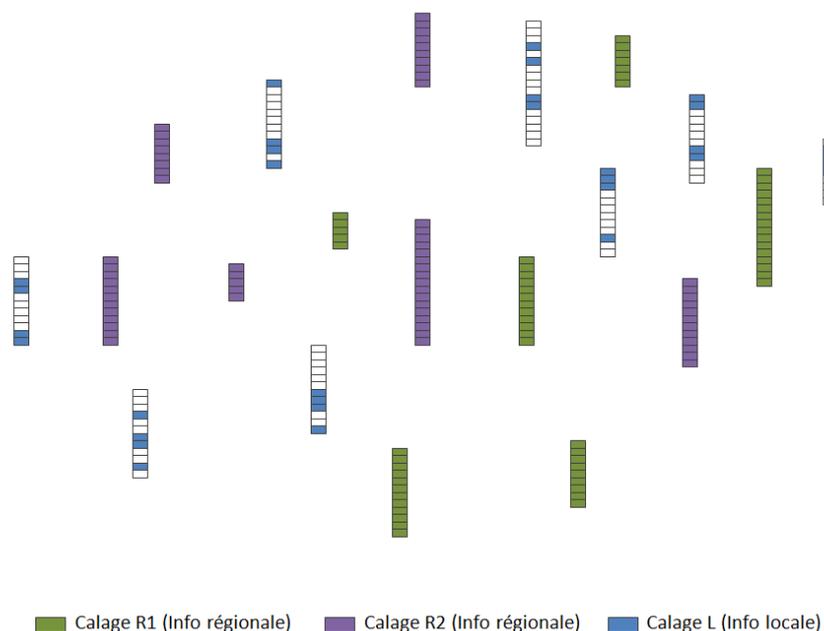


FIGURE III.29 – Décomposition pour le test de la robustesse à l'information régionale.

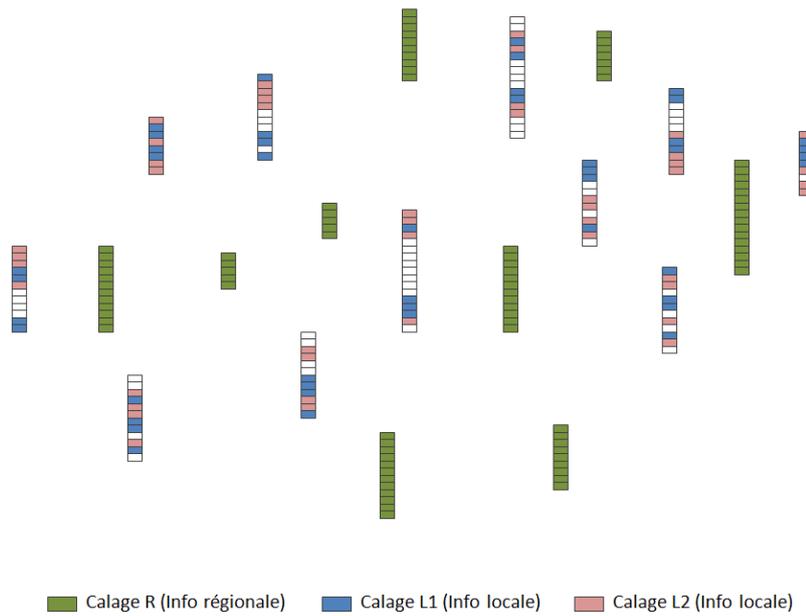


FIGURE III.30 – *Décomposition pour le test de la robustesse à l'information locale.*

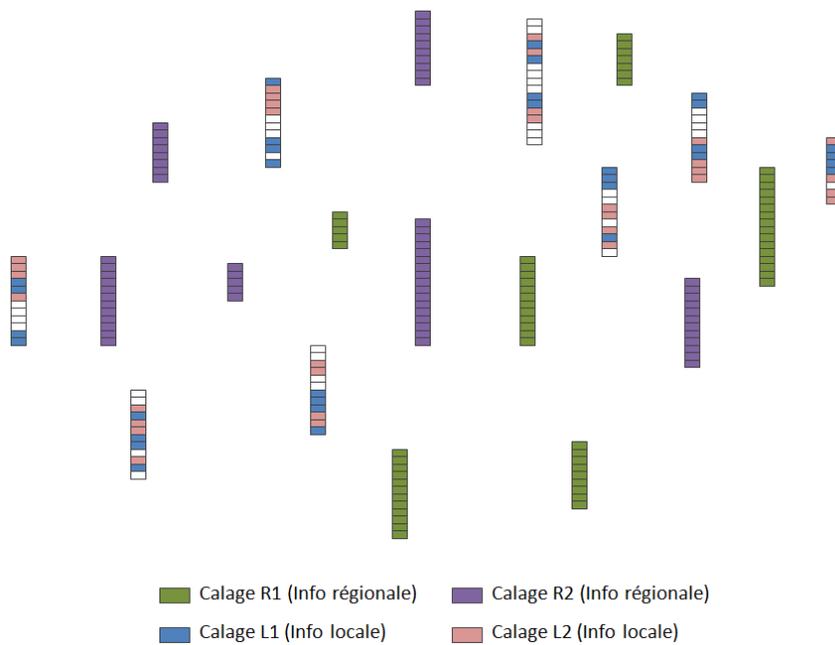


FIGURE III.31 – *Décomposition pour le test de la robustesse générale.*

Scores pour la justesse et la robustesse

Dans le cadre du projet ExtraFlo, des scores ont été mis en place afin d'évaluer et de comparer les méthodes aussi bien en justesse qu'en robustesse (Renard *et al.*, 2013).

Critère de justesse FF : ce critère est utilisé pour estimer la justesse de l'estimation de la probabilité de non-dépassement de la valeur maximale observée d'un échantillon de taille fixée. Il est basé sur une procédure de division d'un échantillon, introduite par Garçon (1995). Chaque échantillon de la station i est divisé en un échantillon de calage **1** de taille $N_{1,i}$ et un échantillon de validation **2** de taille $N_{2,i}$.

Soit $m_{2,i}$ la valeur maximale de l'échantillon de validation **2** de la station i . Soit $F_{1,i}$ la fonction de répartition ajustée sur l'échantillon de calage **1**. Si le modèle est parfait, $F_{1,i}(m_{2,i})$ suit une distribution Beta de paramètres $N_{2,i}$ et 1, i.e. $\mathcal{B}[N_{2,i}, 1]$ (Kumaraswamy, 1980).

$$FF_i = [F_{1,i}(m_{2,i})]^{N_{2,i}} \quad (\text{III.9})$$

Nous calculons alors la différence entre les FF_i et la première bissectrice. Soient (FF'_i) les (FF_i) classés en ordre croissant.

$$FF = 1 - \frac{2}{N} \cdot \sum_{i=1}^N \left| FF'_i - \frac{i}{N+1} \right| \quad (\text{III.10})$$

Le critère FF est compris entre 0 et 1 (1 étant le score optimal).

Critère de justesse NT : ce critère est utilisé pour vérifier si les quantiles calculés sont cohérents avec les observations de l'échantillon de validation. Soit $P_{T,i}$ le quantile estimé associé à la durée de retour T , et $N_{T,i}$ le nombre d'observations de l'échantillon de validation supérieures à $P_{T,i}$. Si l'estimation est juste, $N_{T,i}$ est une réalisation d'une distribution binomiale, le dépassement du quantile $P_{T,i}$ est une épreuve de Bernoulli, de probabilité de succès $Pr(X > P_{T,i}) = \frac{1}{T}$.

$N_{T,i}$ est le nombre de succès parmi $N_{2,i}$ essais, et suit donc une loi binomiale $N_{T,i} \sim \text{Binomiale}(N_{2,i}, \frac{1}{T})$. Soit $N'_{T,i}$ les probabilités de dépassement des $N_{T,i}$ classées en ordre croissant.

$$N_T = 1 - \frac{2}{N} \cdot \sum_{i=1}^N \left| N'_{T,i} - \frac{i}{N+1} \right| \quad (\text{III.11})$$

Le critère N_T est compris entre 0 et 1 (1 étant le score optimal).

Sous cette forme, le score N_T n'est pas adapté à la durée de retour 100 ans : du fait de son caractère discret, il n'est pas possible d'obtenir un score N_T proche de 1, même avec une bonne justesse de l'estimation. En effet, pour une durée de retour T grande devant $N_{2,i}$ les nombres $N_{T,i}$ seront souvent nuls du fait de la rareté de l'événement, la longueur des séries étant limitée à 50 années : un grand nombre de $N'_{T,i}$ seront égaux à $Pr(N_{T,i} > 0) = (1 - \frac{1}{T})^{N_{2,i}}$ et seront donc éloignés de la première bissectrice. Pour éviter cela, les probabilités de dépassement $N'_{T,i}$ ont été modifiées de la façon suivante (Renard *et al.*, 2013) :

$$\begin{cases} f_1 = \text{Prob}(X > N_{T,i}) \\ f_2 = \begin{cases} \text{Prob}(X > N_{T,i} - 1), & \text{si } N_{T,i} \neq 0 \\ 0, & \text{sinon} \end{cases} \\ N'_{T,i} = f_2 + U(0,1) \cdot (f_1 - f_2) \end{cases}$$

où $U(0,1)$ est un nombre tiré dans la loi uniforme entre 0 et 1.

Des courbes du type de celles montrées en figure III.32 permettent de visualiser les différents scores. Nous porterons les fréquences empiriques sur l'axe des abscisses et les $N_{T,i}$ ou

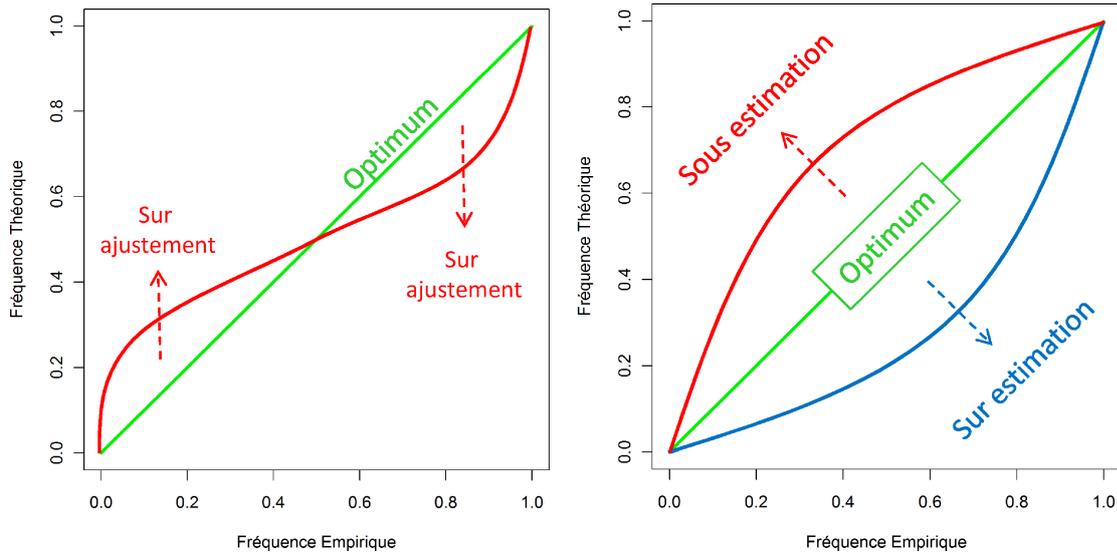


FIGURE III.32 – Graphes types de score FF (Garavaglia, 2011b).

FF_i classés en ordre croissant sur l'axe des ordonnées. Les graphes de $N_{T,i}$ et FF_i permettent de déterminer deux propriétés des modèles (Garavaglia *et al.*, 2010) : si les estimations des quantiles sont surestimées ou sous-estimées (courbes tracées avec les scores évalués sur le jeu de validation), et si les modèles sont sur-paramétrés (courbes tracées avec les scores évalués sur le jeu de calage).

Si la courbe des $N_{T,i}$ est toujours au-dessous de la bissectrice, les estimations des quantiles sont sous-estimées. Il y a trop de cas dans l'échantillon de validation où les valeurs sont supérieures à la pluie décennale estimée avec l'échantillon de calage. Inversement, si la courbe des N_T est toujours au-dessus de la bissectrice, les estimations des quantiles sont surestimées. C'est le contraire avec le critère FF_i . Si la courbe des FF_i est toujours au-dessous de la bissectrice, la durée de retour associée à la valeur maximale est sous-estimée, et donc le quantile que nous calculons pour la durée de retour T sera surestimé. Inversement, si la courbe des FF_i est toujours au-dessus de la bissectrice, les estimations des quantiles sont sous-estimées. L'analyse des courbes des FF_i et des $N_{T,i}$ permet de savoir si les modèles sont sur-paramétrés. En effet, si la courbe est au-dessus de la bissectrice jusqu'à un certain point, et au-dessous de la bissectrice après ce point, le modèle est sur-paramétré. Ses prévisions dépendent trop des données de l'échantillon de calage.

Critère de robustesse SPAN : ce critère est utilisé pour évaluer la stabilité de l'estimation de la durée de retour T , en calculant la différence entre les estimations faites sur deux échantillons différents d'une même station. Il a été proposé par Garavaglia *et al.* (2010). Pour chaque station i et pour chaque durée de retour T , nous calculons un score positif $SPAN_T$, la valeur optimale du score étant 0.

$$SPAN_{T,i} = 2 \cdot \frac{|P_{T,i}(C1) - P_{T,i}(C2)|}{P_{T,i}(C1) + P_{T,i}(C2)} \quad (\text{III.12})$$

La valeur optimum de ce score est 0 et correspond au cas où les deux estimations sont identiques. La valeur $SPAN_{T,i}$ peut être supérieure à 1 si les deux estimations sont très différentes. La formulation pour calculer le score global de chaque durée de retour est :

$$SPAN_T = 1 - \frac{1}{2 \cdot N} \cdot \sum_{i=1}^N SPAN_{T,i} \quad (\text{III.13})$$

III. 3. 2.2 Les méthodes évaluées dans le projet ExtraFlo

Trois types d'approches peuvent être distinguées. Les méthodes purement locales qui s'appuient uniquement sur l'information disponible au site où l'on souhaite réaliser l'estimation des pluies extrêmes. Ce type de méthodes est donc réservé à des sites jaugés. Au contraire, les méthodes purement régionales s'appuient exclusivement sur des données voisines au site d'intérêt. Elles permettent donc une estimation en site non jaugé. Le dernier type d'approche s'appuie à la fois sur de l'information régionale et locale. Les méthodes locales-régionales ont pour objectif de réaliser une estimation locale en un site jaugé mais elles peuvent utiliser de l'information régionale supplémentaire pour consolider leur estimation. Les différentes méthodes comparées dans le projet ExtraFlo vont être détaillées dans les paragraphes suivants (ANR ExtraFlo III.1, 2013; ANR ExtraFlo III.3, 2013; ANR ExtraFlo III.5, 2013).

La méthode SIGEV : La méthode SIGEV est une interpolation spatiale des paramètres d'une distribution GEV (Generalized Extreme Value) locale (Jenkinson, 1955).

Sur un ensemble de N postes – sur lesquels on dispose de l'échantillon des pluies maximales annuelles $(X_j)_{j=1\dots N}$ – une loi GEV est ajustée. Les paramètres μ_j , σ_j et ξ_j sont estimés par la méthode des L-moments. Selon la valeur du paramètre de forme ξ_j ($\xi_j < 0$, $\xi_j = 0$ ou $\xi_j > 0$), le comportement de la distribution est respectivement dans le domaine de Weibull, de Gumbel ou de Fréchet.

$$\begin{cases} F(x_j|\mu_j, \sigma_j, \xi_j) = \exp\left(-\left[1 + \xi_j\left(\frac{x_j - \mu_j}{\sigma_j}\right)\right]^{-\frac{1}{\xi_j}}\right) & , \text{ si } \xi_j \neq 0 \\ F(x_j|\mu_j, \sigma_j, 0) = \exp\left(-\exp\left(-\frac{x_j - \mu_j}{\sigma_j}\right)\right) & , \text{ si } \xi_j = 0 \end{cases} \quad (\text{III.14})$$

Les paramètres de la GEV sont alors interpolés spatialement avec la même technique que les index values dans la méthode de la loi régionale en mode régional. Une estimation des paramètres μ_i , σ_i et ξ_i est alors disponible pour le site cible non jaugé i et la GEV $F(x_i|\mu_i, \sigma_i, \xi_i)$ au site i peut alors être estimée.

La loi GEV régionale (en mode régional ou local-régional) : La variable aléatoire des pluies maximales annuelles au site i est notée X_i et $\hat{F}(x_i)$ est une estimation de la fonction de répartition de X_i . L'idée est d'estimer la fonction de répartition en un site i à partir des N_v stations situées dans un voisinage homogène du site i . La taille de l'échantillon permettant d'estimer $\hat{F}(x_i)$ est ainsi augmentée. La méthode utilisée pour cette estimation est la RFA (Regional Frequency Analysis) de Hosking et Wallis (1997). La mise en œuvre se décline en trois étapes (Neppel *et al.*, 2007) :

1. Définition d'un voisinage homogène associé au site i . Les N_v stations $(X_j)_{j=1\dots N_v}$ de ce voisinage vont servir à l'estimation en i . Chaque échantillon X_j est normalisé par la moyenne des pluies maximales annuelles en j , \bar{X}_j , aussi appelée index value en j . Un échantillon unique des N_v échantillons normalisés $Y_j = \frac{X_j}{\bar{X}_j}$ est alors constitué.
2. Estimation de la loi GEV régionale $\hat{F}(y_i)$ à partir de l'échantillon normalisé régional (Y_j) . Les quantiles $\hat{y}_i(p)$ de probabilité de non-dépassement p sont estimés.
3. Estimation de l'index value \bar{X}_i^* au site i .
4. Retour à une estimation de $\hat{x}_i(p)$ (c'est-à-dire de $\hat{F}(x_i)$) en multipliant $\hat{y}_i(p)$ par la valeur de l'index value en i : $\hat{x}_i(p) = \bar{X}_i^* \cdot \hat{y}_i(p)$. La différence entre la loi GEV régionale en mode purement régional et la loi régionale en mode local-régional réside uniquement dans la

manière d'estimer l'index value \bar{X}_i^* au site i . En mode purement régional, l'index value est interpolé spatialement avec une régression s'appuyant sur des variables explicatives selon la méthode Sliced Inverse Regression (Li, 1991). En mode local-régional, l'index value est estimé à partir de l'information disponible au site d'intérêt i .

Les méthodes SHYPRE et SHYREG : La méthode SHYPRE¹¹ (Cernesson, 1993; Arnaud, 1997; Arnaud *et al.*, 2007) a été conçue pour étudier les distributions de variables hydrologiques (pluies et débits). Elle combine un modèle stochastique pour la pluie horaire avec un modèle pluie-débit (figure III.33). L'extrapolation de la distribution de la pluie vers les grandes durées de retour est obtenue en générant beaucoup d'événements différents sur une grande période de simulation plutôt qu'en ajustant directement une distribution de probabilité théorique sur des valeurs observées.

Le modèle SHYPRE est généralement initialisé avec des données horaires, mais dans cette étude il a été adapté à des observations quotidiennes. Ce générateur de pluies horaires est généralement calé à partir d'information de pluies horaires permettant une analyse des caractéristiques des hyétoigrammes, en vue de leur reconstitution. En l'absence d'information horaire, le générateur peut être calé par une information journalière. Dans ce cas, certains paramètres sont fixés (car peu variables ou peu sensibles) et d'autres sont estimés à partir de variables issues de pluies journalières. Cette version régionalisée (Aubert, 2012) est appelée SHYREG (pour SHYPRE régionalisé) : SHYREG-local si les paramètres journaliers sont déterminés à partir d'une information journalière locale (série pluviométrique) et SHYREG-régional si les paramètres ont été régionalisés.

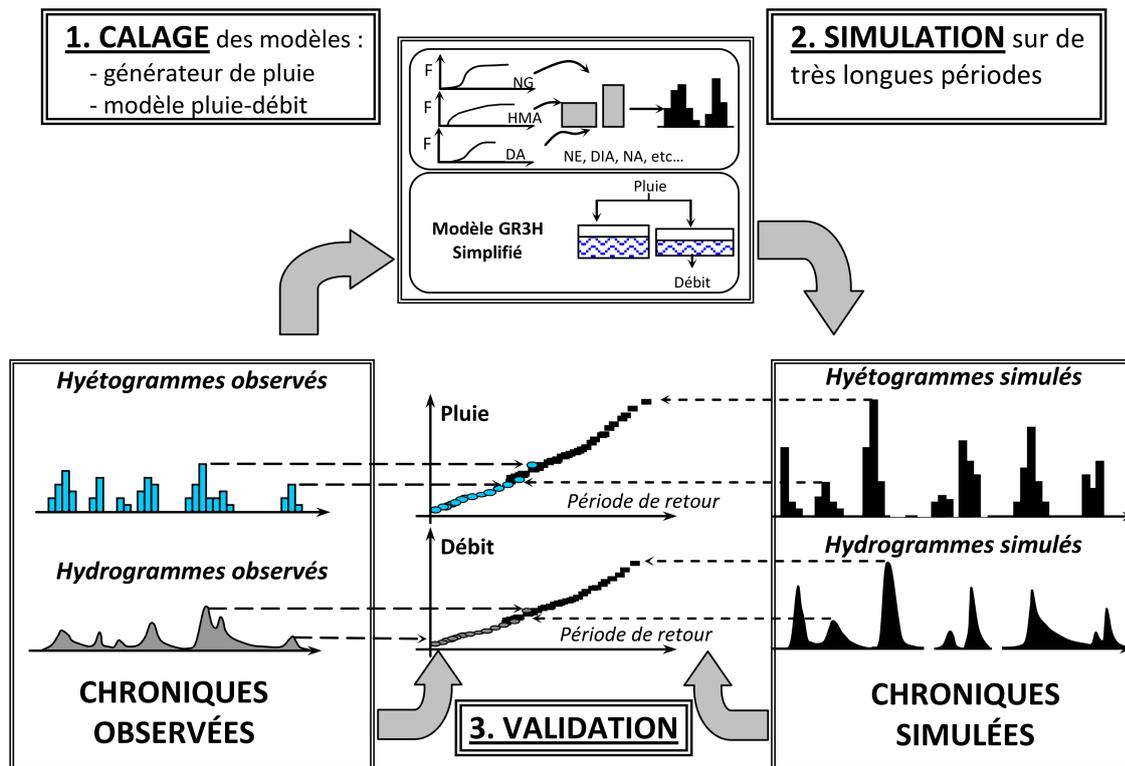


FIGURE III.33 – Principe du modèle SHYPRE (Aubert, 2012).

11. Simulated HYdrographs for flood PRobability Estimation

La loi GP locale : La fonction de répartition au site i est estimée avec un ajustement d'une distribution GP (Generalized Pareto) sur la population de pluies sup-seuil X_i du site i . La population sélectionnée garde environ quatre observations par an. Les paramètres de la GP (Pickands, 1975) sont estimés avec les moments pondérés.

$$\begin{cases} F(x_i|\mu_i, \sigma_i, \xi_i) = 1 - \left[1 + \xi_i \left(\frac{x_i - \mu_i}{\sigma_i} \right) \right]^{-\frac{1}{\xi_i}} & , \text{ si } \xi_i \neq 0 \\ F(x_i|\mu_i, \sigma_i, 0) = 1 - \exp\left(-\frac{x_i - \mu_i}{\sigma_i}\right) & , \text{ si } \xi_i = 0 \end{cases} \quad (\text{III.15})$$

La méthode MEWP locale : Une distribution MEWP est calée localement au site i . Pour chaque type de temps, la loi exponentielle est ajustée par maximum de vraisemblance sur la population de pluie sup-seuil du type de temps. Le seuil est le quantile 70% de la population de ce type de temps.

III. 3. 2.3 Évaluation des distributions MEWP ajustées sur SPAZM

En utilisant les scores en justesse et en robustesse, ainsi que les différentes décompositions calage-validation, nous pouvons réaliser une inter-comparaison des méthodes de spatialisation des précipitations extrêmes.

Ce protocole est appliqué à la méthode d'estimation des pluies extrêmes avec un ajustement MEWP sur les pluies SPAZM. **La méthode SPAZM est recalée de façon à suivre le protocole de validation. Les ébauches ne seront pas changées, mais la déformation de l'ébauche et le krigeage des résidus sont quant à eux réalisés uniquement sur le réseau de calage régional.**

L'information SPAZM à un pixel est alors considérée comme une information locale. Une distribution MEWP est ajustée sur les n années de pluies SPAZM sélectionnées pour l'information locale. **Enfin, l'ajustement MEWP sur les pluies SPAZM est une méthode purement régionale.**

Les conclusions de l'action III.5 du projet Extraflo ont montré que la distribution MEWP locale avait de bonnes performances de justesse et des performances de robustesse comparables aux méthodes avec une composante régionale.

Les graphes III.34, III.35 et III.36 montrent les différents scores de justesse, en calage et en validation. Quatre méthodes sont représentées : un ajustement MEWP sur les données locales (MEWP local), un ajustement MEWP sur la donnée SPAZM sans la correction globale de 10% sur les quantiles (MEWP SPAZM V1), un ajustement MEWP sur la donnée SPAZM avec la correction globale de 10% (MEWP SPAZM V2) et SHYREG en local-régional. Les courbes SHYREG et MEWP locale montrent que ces deux méthodes ont de bonnes performances en justesse et robustesse, ce qui permet de juger MEWP SPAZM vis-à-vis de ces méthodes. Il est cependant important de rappeler que cette comparaison mélange des méthodes purement locales, purement régionales et locales-régionales qui n'ont donc pas les mêmes domaines d'applications (jaugé ou non jaugé). En particulier, la méthode d'ajustement sur SPAZM est une méthode purement régionale qui ne s'appuie sur aucune information locale (hormis l'information locale apportée par les ébauches).

La correction quantile-quantile sur MEWP SPAZM V2 permet de corriger le problème de sous-estimation des quantiles de pluies. Nous constatons également que MEWP SPAZM V2 atteint des performances, en calage ou validation, tout à fait acceptables au vu de celles des méthodes utilisant l'information locale. Les scores en calage pour le NT100 pour des durées d'information locale de 5 ou 10 ans sont meilleurs qu'avec une distribution MEWP locale.

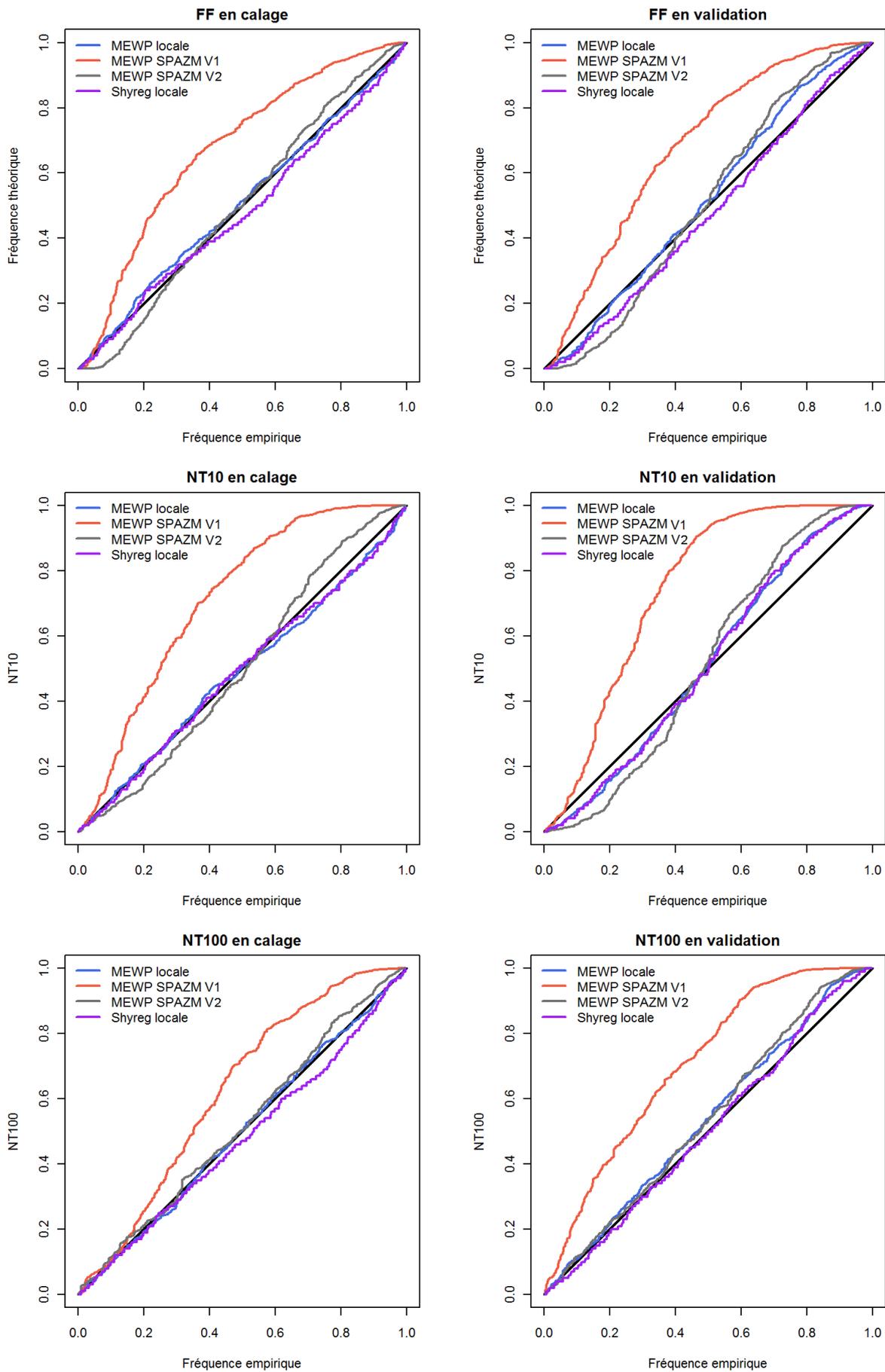


FIGURE III.34 – Scores de justesse en calage et en validation pour une durée d'information locale de 20 ans.

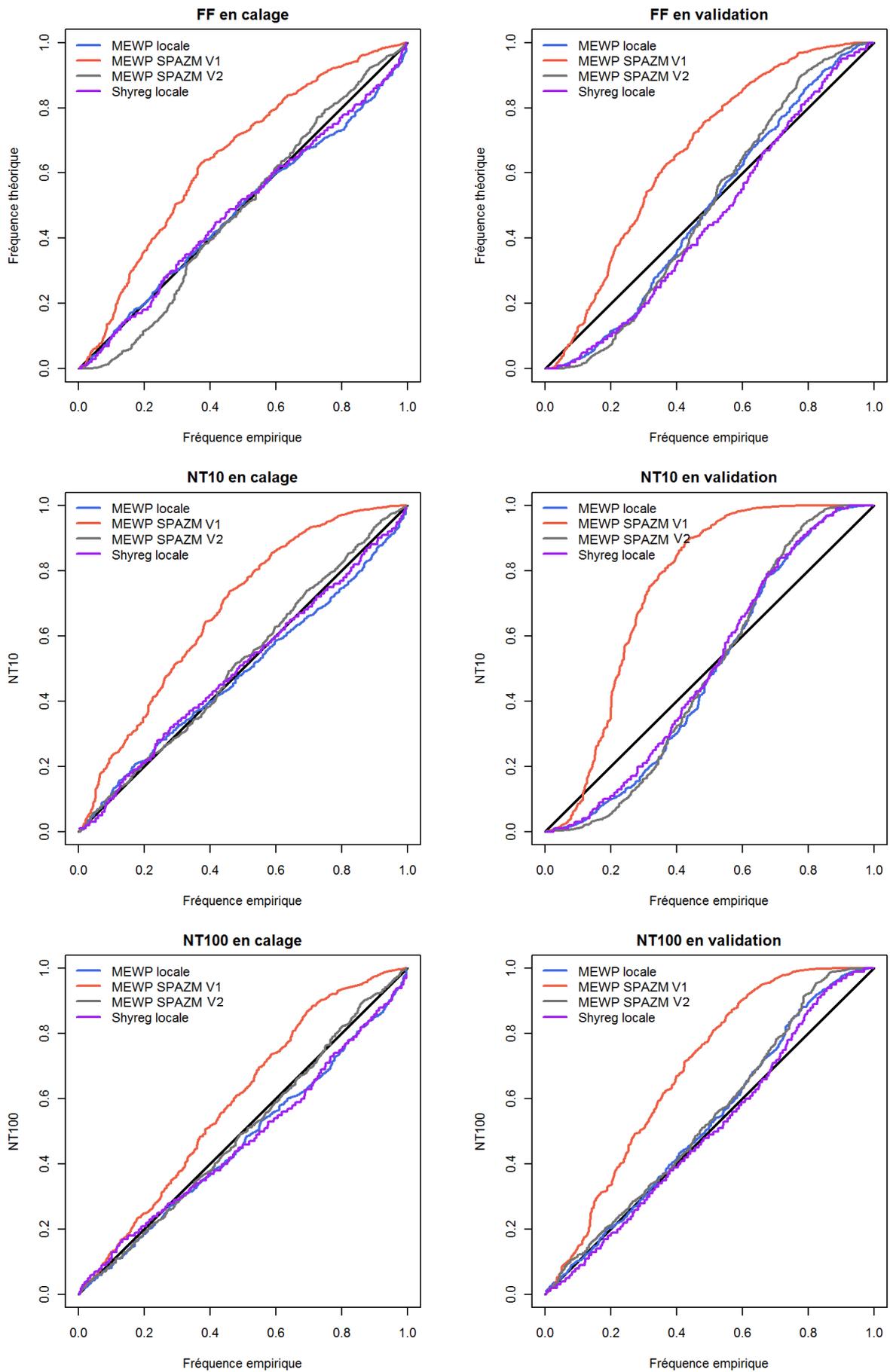


FIGURE III.35 – Scores de justesse en calage et en validation pour une durée d’information locale de 10 ans.

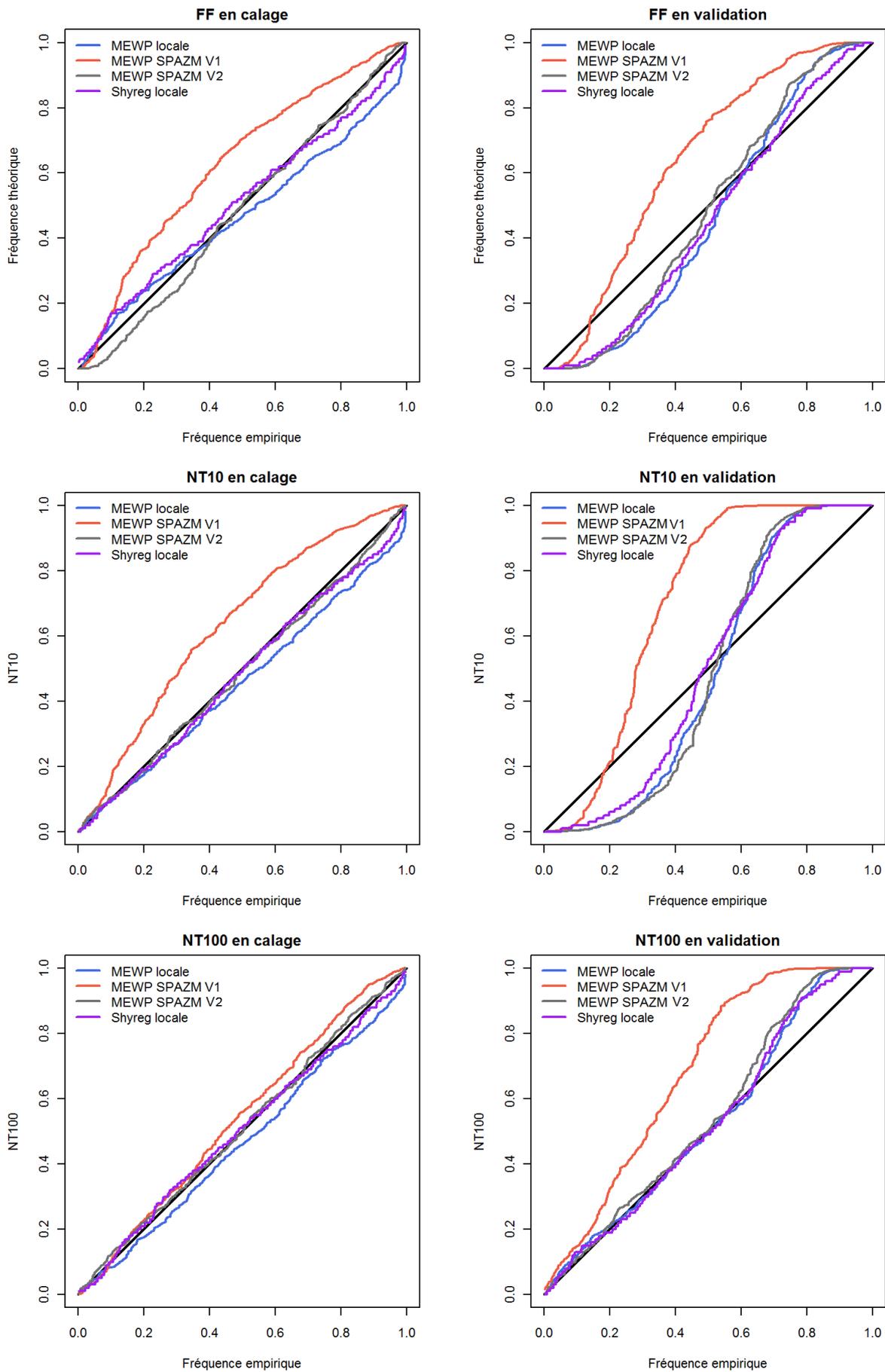


FIGURE III.36 – Scores de justesse en calage et en validation pour une durée d'information locale de 5 ans.

Le tableau III.9 synthétise les scores globaux de justesse en validation pour les différentes méthodes évaluées dans ExtraFlo. Les valeurs optimum des scores sont 1. Les méthodes sont réparties en trois catégories : régionales, locales et locales-régionales. Pour les méthodes régionales, seule l'information régionale est prise en compte pour le calage (aucune information locale). Ainsi, ces méthodes ne sont donc pas déclinées selon la longueur d'information locale utilisée. Cependant, la méthode SPAZM V2 est fournie pour une période de 10 et 20 ans d'information « locale ». Ceci correspond à deux ajustements MEWP estimés réciproquement sur 10 et 20 ans de pluies journalières SPAZM.

La méthode SPAZM V2 construite suivant le protocole ExtraFlo est purement régionale (comme pour une utilisation en contexte non jaugé). Dans la pratique, la méthode SPAZM peut normalement être considérée comme une méthode locale-régionale puisqu'elle utilise l'information régionale mais également l'information locale disponible au site à estimer. Dans ce cas, SPAZM étant un interpolateur exact, les résultats avec la méthode SPAZM V2 seraient proches de ceux obtenus avec la méthode MEWP local.

En comparant MEWP SPAZM V2 aux autres méthodes régionales, nous constatons qu'elle a une meilleure justesse que la méthode SIGEV et est équivalente à la loi GEV régionale en mode régional. Les méthodes utilisant de l'information locale sont plus justes. En comparant MEWP SPAZM V2 aux méthodes utilisant de l'information locale (méthodes locale et locale-régionale), il apparaît que pour des périodes de données locales de 20 ans MEWP local, SHYREG locale-régionale et la GEV en mode local-régional ont de meilleures performances. Cependant lorsque la quantité d'information locale diminue (10 ans), alors les écarts de performance entre les méthodes régionales – MEWP SPAZM V2 et la GEV en mode régional – et les méthodes locales et locales régionales – MEWP local, SHYREG locale-régionale et la GEV en mode local-régional– se réduisent de façon significative.

	Méthodes	NT10	NT100	FF
Régional	GEV en mode régional	0.84	0.95	0.88
	SIGEV	0.72	0.65	0.81
	MEWP SPAZM V2 (10 ans)	0.82	0.92	0.86
	MEWP SPAZM V2 (20 ans)	0.84	0.92	0.88
Local	GP (10 ans)	0.75	0.86	0.81
	GP (20 ans)	0.84	0.89	0.90
	MEWP local (10 ans)	0.83	0.93	0.90
	MEWP local (20 ans)	0.88	0.91	0.96
Local-Régional	GEV en mode local-Régional (10 ans)	0.85	0.92	0.92
	GEV en mode local-Régional (20 ans)	0.89	0.94	0.96
	SHYREG locale-régionale (10 ans)	0.86	0.95	0.89
	SHYREG locale-régionale (20 ans)	0.91	0.97	0.94

Tableau III.9 – Scores de justesse en validation pour les différentes méthodes évaluées dans le projet ExtraFlo.

Le tableau III.10 compare la robustesse à l'information régionale pour les méthodes régionales. Comme pour la justesse, la méthode MEWP SPAZM V2 est similaire à la GEV en mode régional et est plus robuste que la méthode SIGEV.

	Méthodes	SPAN10	SPAN100	SPAN1000
Régional	GEV en mode régional	0.94	0.92	0.89
	SIGEV	0.83	0.82	0.77
	MEWP SPAZM V2 (10 ans)	0.92	0.91	0.90
	MEWP SPAZM V2 (20 ans)	0.92	0.91	0.90

Tableau III.10 – Score de robustesse à l'information régionale pour les différentes méthodes régionales.

Bilan : pour une méthode purement régionale, l'ajustement d'une loi MEWP sur les pluies SPAZM avec une correction quantiles-quantiles de +10% (MEWP SPAZM V2) fournit des résultats tout à fait acceptables aussi bien en justesse qu'en robustesse à l'information régionale. Elle donne des résultats assez similaires à ceux obtenus avec la loi GEV régionale en mode régional. Par ailleurs, du fait du protocole ExtraFlo, la méthode SPAZM n'exploite pas d'information locale (hormis celle assimilée dans les ébauches de pluie). En pratique, l'utilisation de SPAZM permet de disposer d'une méthode locale-régionale capable de traiter l'information locale si disponible et donc de réaliser des ajustements MEWP plus proches de ceux obtenus avec la loi MEWP locale.

III. 3. 3 Les cartes de pluies extrêmes proposées

L'objectif initial était de construire des cartes de pluies journalières pour les temps de retour 10, 100, 1000 et 10 000 ans. Les cartes qui vont être présentées dans la suite ont été construites avec un ajustement MEWP sur les pluies SPAZM pixel par pixel. Ces cartes ne prennent pas en compte la correction quantile-quantile appliquée à SPAZM mais elles illustrent l'allure générale des champs de quantiles remarquables de pluie.

Les étapes de constructions des cartes de pluie de temps de retour T sont détaillées dans le schéma III.37. En chaque pixel p , la série de pluies journalières SPAZM est extraite sur la période 1953-2005. Une distribution MEWP est ajustée sur cette série en prenant en compte les saisons régionales définies en section II. 3. 1 (page 39). Sur cette série, nous calculons l'écart-type σ_{SPAZM} des pluies centrales supérieures au quantile 70%. La valeur $\sigma_{krigé}$ au pixel p , estimée par un krigeage des écarts-types au droit des postes, est également récupérée. La correction quantile-quantile effectuée au pixel p est alors $K(p) = \frac{\sigma_{krigé}(p)}{\sigma_{SPAZM}(p)}$. Les pluies de temps de retour T en p sont estimées en récupérant les quantiles de la distribution MEWP ajustée aux pluies SPAZM $PJ_{SPAZM}(p, T)$ et en y appliquant la correction multiplicative $K(p)$. Une itération sur tous les pixels produit alors des cartes de quantiles SPAZM corrigés $PJ^*(p, T)$.

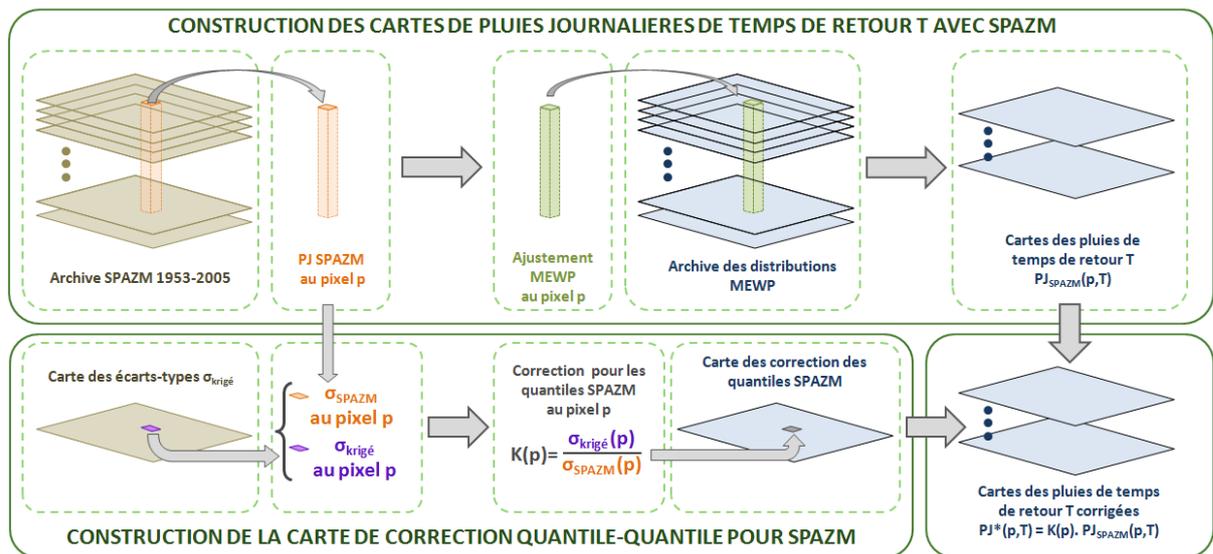
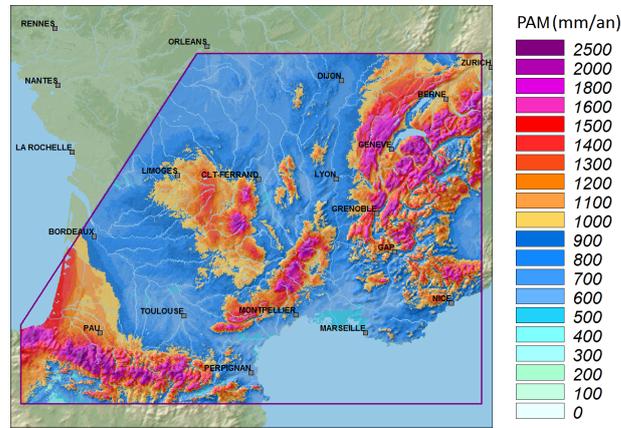
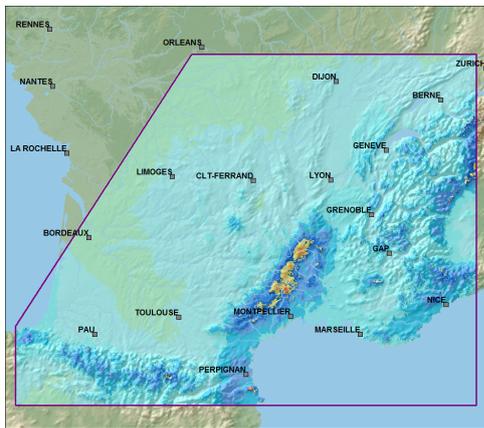


FIGURE III.37 – Schéma explicatif pour la construction des cartes de pluies journalières de temps de retour T .

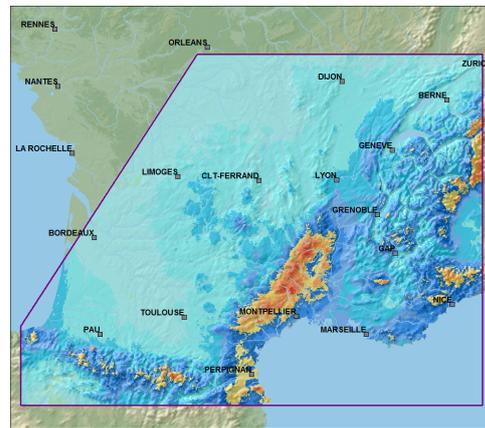
Comme nous disposons d'une distribution MEWP à chaque pixel, les gradex par types de temps sont connus en tout point. Les cartes de la figure III.39 représentent les gradex (à gauche) et les cumuls de pluie annuels (à droite) de chaque type de temps. Elles ont surtout un intérêt qualitatif et illustrent des structures du risque pluviométrique bien différentes d'un type de temps à l'autre. Elles mettent également en évidence le problème déjà abordé d'un effet orographique sur les fortes précipitations dépendant de l'exposition du relief.



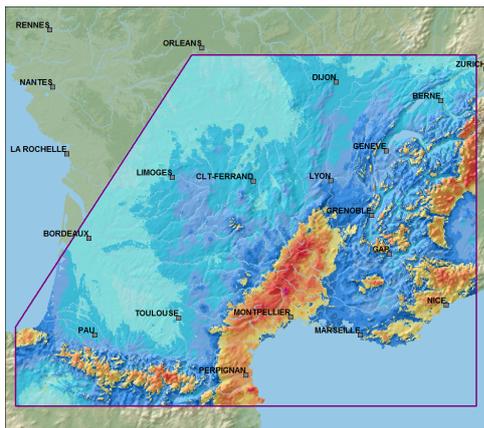
(a) Cumul de pluie annuelle (en mm)



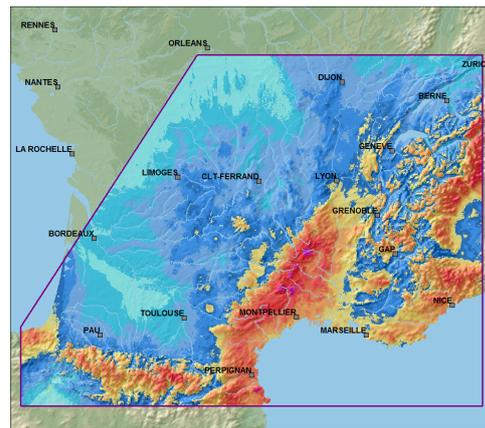
(b) $T=10$ ans



(c) $T=100$ ans



(d) $T=1000$ ans



(e) $T=10\ 000$ ans

FIGURE III.38 – Cartes des pluies annuelles (PAM) et des quantiles de pluies journalières PJ(T).

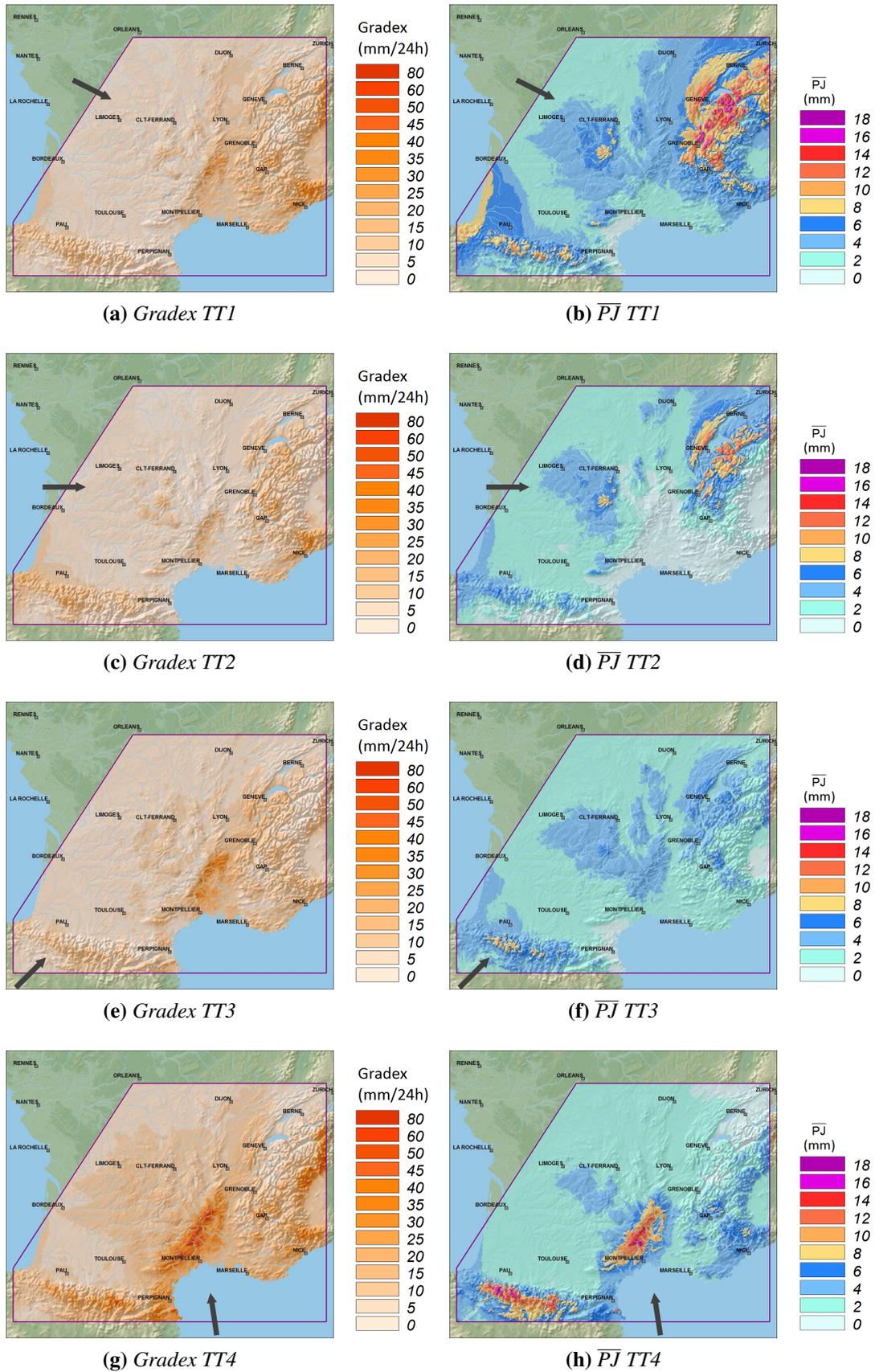


FIGURE III.39 – Cartes des gradex de chaque type de temps pour la saison à risque et des pluies journalières moyennes de ce type de temps.

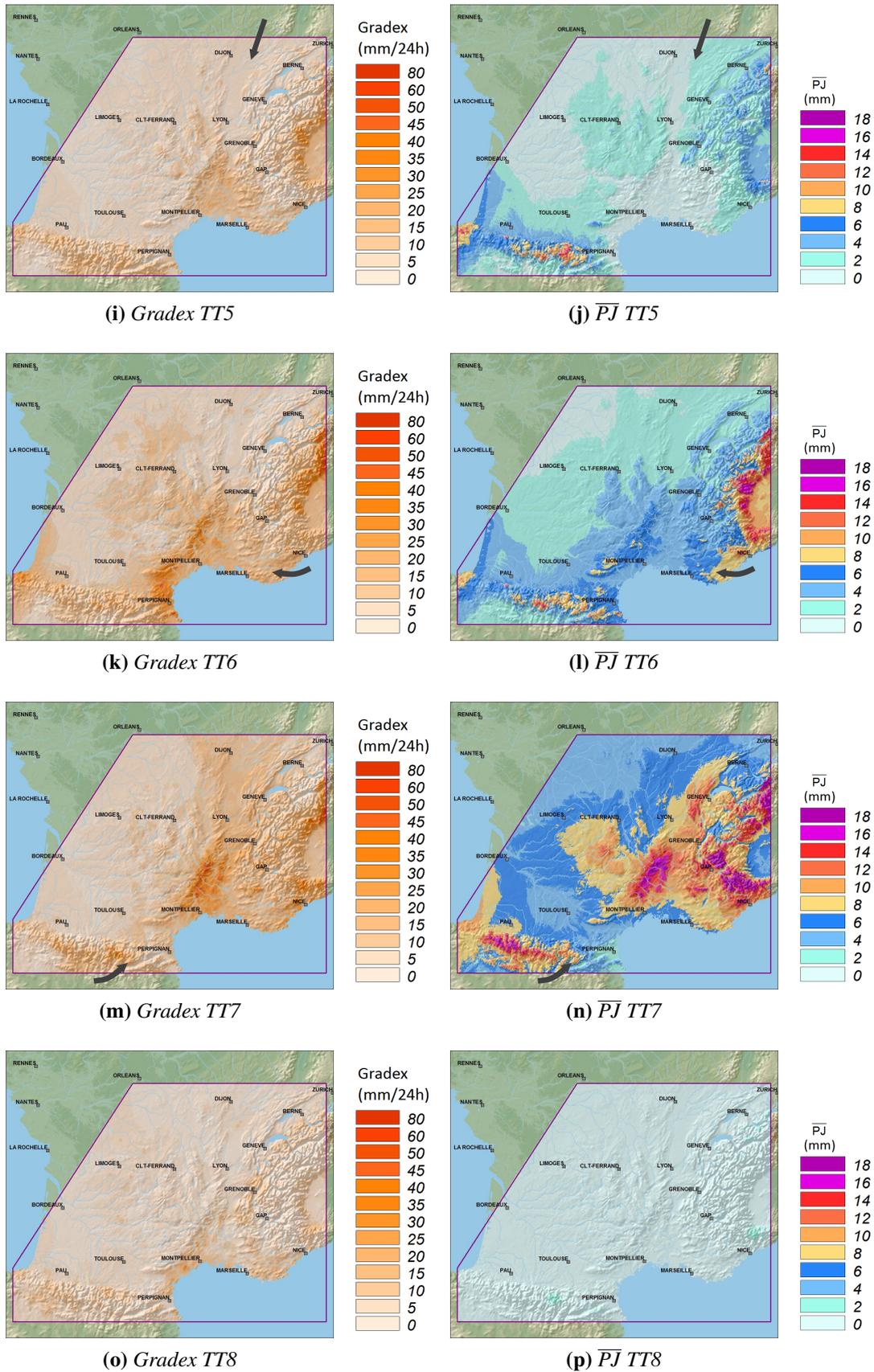


FIGURE III.39 – Cartes des gradex de chaque type de temps pour la saison à risque et des pluies journalières moyennes de ce type de temps.

III. 4 Vers une estimation des pluies extrêmes à un pas de temps infra-journalier

Tout le travail de régionalisation des pluies extrêmes ponctuelles s'est borné au pas de temps journalier et constitue un travail exploratoire pour une généralisation aux pas de temps infra-journaliers. En effet, les processus de précipitations pour des pas de temps plus fins peuvent être différents et nécessiteront donc une adaptation du processus utilisé en journalier pour construire des cartes de quantiles de pluies.

Les conclusions du projet ExtraFlo ont mis en évidence que la distribution MEWP donne de bons résultats en terme de robustesse et de justesse mais qu'elle sous-estimait légèrement les forts quantiles de pluies dans le pourtour méditerranéen (ANR ExtraFlo III.1, 2013; ANR Extraflo V, 2013). Il a notamment été suggéré qu'une raison à cette sous-estimation soit liée à l'hypothèse d'un comportement asymptotique exponentiel pour la distribution des précipitations de certains types de temps.

Le travail réalisé dans ce chapitre a pour but de mettre en évidence l'existence d'une zone dans laquelle les fortes précipitations pourraient avoir un comportement différent de celui envisagé avec la distribution MEWP et de proposer des solutions pour adapter la distribution et améliorer les estimations dans cette zone. **Ces propositions ont pour objectif principal de définir une approche pour traiter la problématique de l'estimation des précipitations extrêmes aux pas de temps infra-journaliers.**

III. 4. 1 Mise en évidence d'une zone « atypique » pour les précipitations extrêmes journalières

Afin de mettre en évidence l'existence d'une zone atypique du point de vue des fortes précipitations, nous envisageons l'indicateur local $\frac{PJX}{PA}$ avec \overline{PJX} la moyenne des maxima annuels et PA la précipitation annuelle moyenne. Cet indicateur permet d'apprécier l'importance des plus fortes pluies par rapport au cumul annuel. Il est calculé pour chacun des 770 postes de la base. La carte obtenue est présentée sur la figure III.40. Une très nette cohérence spatiale peut être notée pour cet indicateur. Le croissant méditerranéen est caractérisé par des valeurs élevées de cet indicateur. Une zone atypique, dénommée zone B, est délimitée par le quantile 10% de cette valeur (qui correspond à une zone où les pluies maximales annuelles représentent plus de 9% du cumul annuel). Cette zone englobe la côte méditerranéenne, les Pyrénées orientales, le front de Cévennes et remonte dans la vallée du Rhône.

La zone atypique B construite montre un comportement sur les fortes pluies différent de ce qui est observé sur le reste du domaine que nous appellerons zone A. Se pose alors la question du comportement asymptotique de la distribution des précipitations dans chacune des zones.

À titre indicatif, une distribution GEV (Generalized Extreme Value) a été calée localement sur chaque poste de la base. La loi GEV est décrite par trois paramètres : un paramètre de position, un paramètre d'échelle et un paramètre de forme. Nous nous intéresserons uniquement à la valeur du paramètre de forme ξ qui décrit l'asymétrie de la densité de probabilité, c'est-à-dire l'allure de la queue de la distribution des pluies (figure III.41). Les cas où $\xi = 0$, $\xi > 0$ et $\xi < 0$ correspondent, respectivement, aux domaines de Gumbel, Fréchet et Weibull (loi bornée).

La distribution du paramètre de forme a été déterminée dans les deux zones (figure III.40 (b)). Les paramètres de forme dans la zone atypique B sont nettement supérieurs à ceux obtenus dans la zone A, ce qui veut dire que le comportement de la GEV est bien plus dans le domaine de Fréchet. En particulier, la valeur médiane du paramètre de forme dans la zone atypique B est de 0.12 (contre 0.05 dans la zone A). Cette valeur médiane rappelle le résultat du travail de

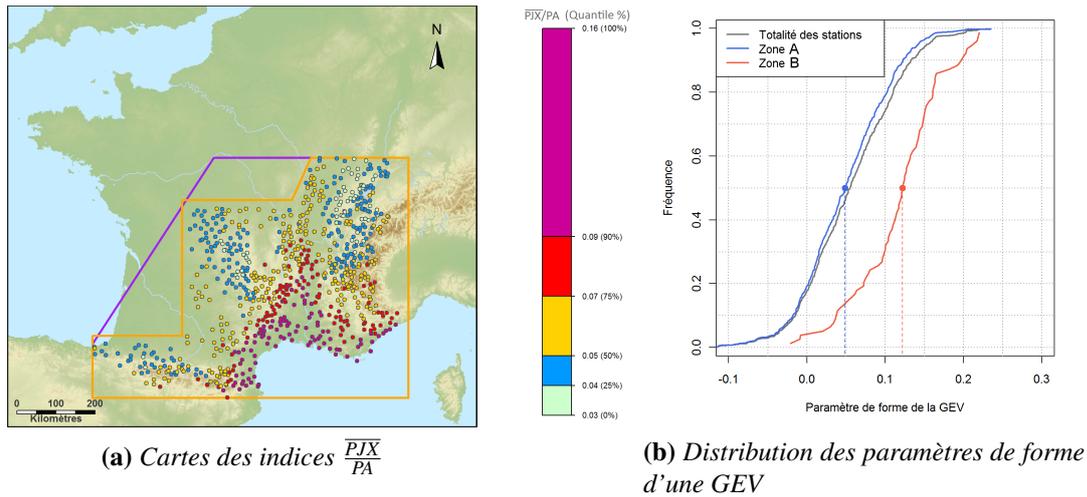


FIGURE III.40 – Mise en évidence d'une zone atypique (sur le jeu des 770 postes).

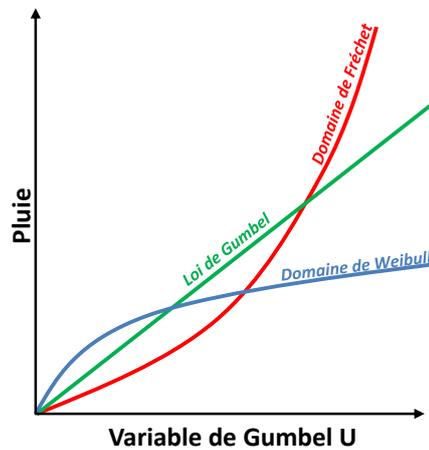


FIGURE III.41 – Influence du paramètre de forme sur le comportement asymptotique d'une distribution.

Neppel *et al.* (2007) qui avaient conclu à une valeur moyenne du paramètre de forme de 0.13 en région méditerranéenne, sur la zone des Cévennes.

L'indicateur local $\frac{\overline{PJX}}{PA}$ a le mérite d'être indépendant de toute méthode d'ajustement et de mettre en évidence une organisation spatiale du comportement des pluies extrêmes pas toujours évidente au regard de la répartition spatiale des paramètres de forme de loi d'ajustement. Par ailleurs, les ajustements d'une loi GEV sur le pourtour méditerranéen semblent plus cohérents avec une distribution annuelle des pluies ayant un comportement asymptotique¹² dans le domaine de Fréchet.

Nous proposons alors de regarder les performances de la distribution MEWP selon la zone considérée. Les scores de justesse FF (page 95) en calage et en validation sont alors calculés sur la zone A (soit sur 90% des postes) et sur la zone atypique B (soit sur 10% des postes) avec un ajustement MEWP local et avec l'ajustement MEWP sur les pluies SPAZM (figure III.42).

Les deux méthodes ont de bonnes performances en justesse sur la zone A. Les scores FF mettent par contre en évidence une sous-estimation des fortes précipitations dans la zone B.

12. Le terme comportement asymptotique n'est pas exact puisque que nous sommes toujours dans la zone de mélange.

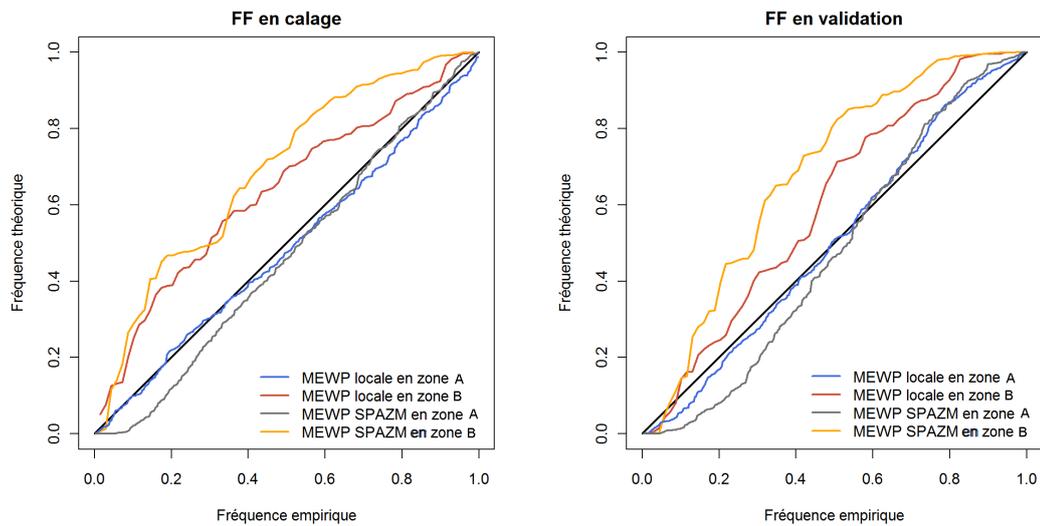


FIGURE III.42 – *Score ExtraFlo selon la zone.*

La méthode régionale MEWP SPAZM montre une sous-estimation plus forte que l’ajustement MEWP local sur cette zone.

Bilan : la mise en évidence d’un comportement des fortes pluies différent sur le pourtour méditerranéen prouve l’intérêt d’adapter la distribution MEWP pour mieux capter cette différence dans un premier temps pour les pluies journalières mais surtout dans la perspective de s’intéresser à l’infra-journalier. En effet, cette zone « atypique » mériterait d’être reconstruite pour des pas de temps plus fins et d’être comparée à celle obtenue en journalier.

III. 4. 2 Vers une distribution MEWP pour l'infra-journalier

La distribution MEWP a de bonnes performances en robustesse et en justesse, mais la différence entre les scores sur la zone A et la zone B prouve que la distribution pourrait être adaptée afin d'améliorer ses performances sur la zone atypique B et dans l'optique de l'utiliser à terme pour des pas de temps infra-journaliers.

Garavaglia (2011b) avait envisagé un modèle MGPWP (Multi Generalized Pareto Weather Pattern) afin de remplacer les lois exponentielles, utilisées comme distributions marginales des pluies centrales par types de temps, par des distributions de Pareto généralisées (GP). Ce changement avait pour but de prendre en compte un comportement asymptotique dans le domaine de Fréchet. Cependant, l'ajustement local de cette distribution MGPWP manque de robustesse. Un ajustement local d'une loi à trois paramètres est en effet très sensible à l'échantillon utilisé en calage.

Afin d'illustrer ce problème, un exercice de style est présenté sur la figure III.43. Nous considérons un point où une distribution des pluies doit être estimée (point bleu). Un panier de stations homogène (points rouges) est constitué (sur un critère de proximité géographique avec le point considéré par exemple). Dans le cadre de cet exercice, nous considérons que la distribution des pluies au point à estimer est une loi exponentielle connue (loi mère). Un échantillon de pluie est alors généré par un tirage dans la loi mère pour le poste bleu. Pour chaque poste rouge du panier, une population de pluie est tirée dans la même loi mère. L'idée est d'estimer un ajustement à partir de ces échantillons de pluie tous issus de la même loi et d'apprécier la qualité du caractère prédictif de la distribution obtenue par rapport à la distribution mère.

Sur la figure (b), la courbe bleue représente la loi exponentielle mère. Les croix représentent les distributions empiriques des populations issues des tirages (les croix bleues pour le point à estimer et les croix rouges pour les points du panier). La courbe orange montre l'ajustement d'une GP sur les données locales (croix bleues). Cet ajustement représente bien l'échantillon de calage en allant chercher les points les plus hauts dans la distribution empirique, mais il ne représente pas du tout la loi exponentielle mère. Le caractère prédictif d'une telle loi est donc peu fiable. Ce constat illustre le problème de calage local d'une loi à trois paramètres évoqué précédemment. La loi verte représente la distribution obtenue par ajustement d'une loi exponentielle sup-seuil sur l'échantillon local avec un seuil fixé au quantile 70% de la population. Cette exponentielle sup-seuil est plus proche de l'ajustement attendu et a un meilleur caractère prédictif que la distribution GP calée localement. C'est le type de loi qui est utilisé dans MEWP pour avoir un ajustement par type de temps (le seuil peut être ajusté). La dernière courbe (courbe rouge) est une loi GP ajustée en tenant compte de l'information régionale disponible. La méthode régionale utilisée est la méthode RFA (Regional Frequency Analysis) par les L-moments de Hosking et Wallis (1997). Même si la loi utilisée est une distribution GP, l'ajustement finalement obtenu est plus proche de l'exponentielle mère que les deux autres. L'information régionale améliore le caractère prédictif et permet un ajustement plus robuste car moins sensible au problème d'échantillonnage. Contrairement à l'ajustement de l'exponentielle sup-seuil, la loi GP régionale a été calée sans a priori sur la loi (exponentielle ou GP) ou sur la valeur du seuil (pris à 70% avec l'exponentielle sup-seuil).

Dans la pratique, la loi mère n'est pas connue. Juger du type de loi à la seule vue d'une distribution empirique locale (qu'en elle est disponible) peut être délicat comme l'a montré l'exemple précédent.

Deux pistes d'amélioration du modèle MEWP peuvent être suggérées afin de répondre au problème de sous estimation dans la zone B :

- Une approche régionale semble être un bon moyen d'améliorer la robustesse des estimations mais également de supprimer le choix a priori d'un seuil pour le calage des

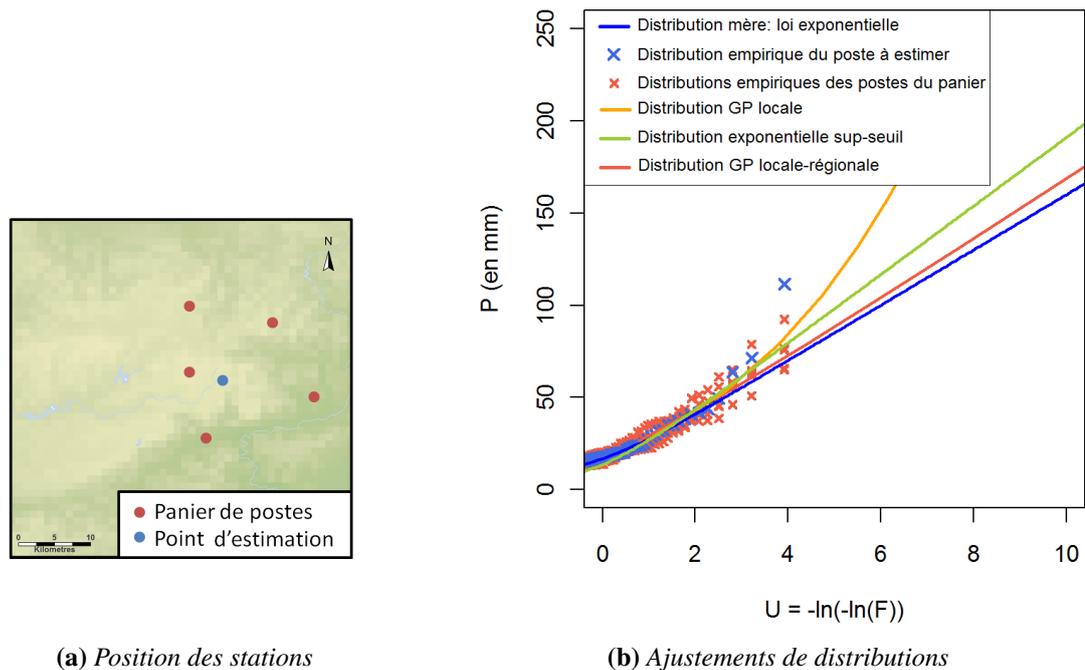


FIGURE III.43 – Ajustement de différentes lois sur des échantillons tirés dans une loi exponentielle.

exponentielles par type de temps. L'utilisation d'une approche régionale est indispensable si l'utilisation de loi à trois paramètres est envisagée pour régler le problème de sous-estimation de la loi MEWP en zone B.

En effet, l'utilisation d'une approche RFA avec les L-moments permettrait d'intégrer une information locale et régionale. Cette évolution du modèle probabiliste MEWP permettrait ainsi de prendre en compte un paramètre de forme régional robuste pour les lois marginales par type de temps (e.g. loi GP). Ce point a fait l'objet d'un stage (Douchez, 2013) qui a produit des cartes de paramètres de forme par types de temps qui montrent une tendance à observer des paramètres de formes positifs pour certains types de temps dans le pourtour méditerranéen.

- Le problème observé dans la zone B peut également venir de la classification des types temps et non pas du choix de la loi marginale. En effet, la classification par type de temps EDF a été construite sur une fenêtre centrée sur la France (section II. 1, page 29) et sans utiliser de postes pluviométriques en bord de mer. Il est donc possible que les huit types de temps identifiés ne soit pas assez représentatifs de la climatologie du pourtour méditerranéen. Une solution pour améliorer les estimations MEWP dans la zone B serait alors de redéfinir une classification pour cette zone.

Un projet de recherche avec le LTHE (Laboratoire d'étude des Transferts en Hydrologie et Environnement) devrait voir le jour pour traiter cette problématique. L'objectif est de définir une nouvelle formulation de modèle MEWP, de la valider pour l'estimation des pluies extrêmes journalières, puis de s'intéresser à une utilisation sur des pas de temps infra-journalier et enfin d'étudier la cohérence des estimations entre les différents pas de temps.

III. 5 Conclusion

L'étude des pluies journalières extrêmes locales avait pour objectif de répondre à plusieurs questions. La comparaison de différentes approches a montré l'intérêt de travailler avec une information par type de temps. En effet, la classification EDF en huit classes permet de tenir compte de sous-populations plus homogènes, mais également de s'intéresser à des processus climatiques différents ayant leur propre structure spatiale. Un krigeage ordinaire des quantiles de pluies se révèle être une méthode peu robuste dès que la densité de postes pluviométriques est faible et tend à sous-estimer les plus forts quantiles de chaque temps de retour.

Une idée simple et souvent envisagée pour pallier cette sous-estimation consiste à prendre en compte l'effet orographique. Cependant, **bien qu'il soit clair que le relief ait un impact certain sur les précipitations extrêmes, intégrer son influence sur les champs de quantiles est loin d'être évident.** Le lien entre le relief et les pluies extrêmes est souvent très complexe et dépend en grande partie de la région considérée et des influences climatiques en jeu. **La décomposition en types de temps permet de prendre en compte cette influence du relief selon les processus climatiques prédominants dans une région.** Ainsi, des comportements très différents – influences méditerranéennes, influences atlantiques et retours d'Est – ont pu être dégagés et ont mis en évidence des effets orographiques plus ou moins marqués selon l'exposition aux flux de chaque type de temps. La chaîne des Cévennes constitue un obstacle majeur aux flux méditerranéens ce qui explique que son relief soit un excellent descripteur du risque du type de temps 4 et 7 dans cette région. De même, les retours d'Est (TT6) concernent la région du Languedoc Roussillon, les Pyrénées orientales et le versant des Alpes italiennes. Afin de prendre en compte l'influence plus ou moins forte du relief selon les sites, l'information fournie par la pluie SPAZM s'est révélée particulièrement utile.

L'interpolateur de pluies journalières SPAZM fait la synthèse d'un grand nombre de données. Il s'appuie sur un réseau pluviométrique très dense et réalise une interpolation des pluies tenant compte d'un gradient altimétrique local des pluies par type de temps. La prise en compte de ce gradient altimétrique local par type de temps, plutôt qu'un gradient moyen global, corrige systématiquement le problème de sous-estimation sur les fortes valeurs des pluies d'un temps de retour T ($T \geq 10$ ans). Ceci peut être réalisé soit par l'utilisation de variables explicatives (cumul annuel des pluies par type de temps) pour la régionalisation des gradex des types de temps, soit par l'utilisation directe des séries de pluies journalières SPAZM. Dans tous les cas, **l'information SPAZM améliore nettement les performances des méthodes de régionalisation des quantiles de pluies extrêmes et en particulier lorsque la densité du réseau sol est faible.**

Ce travail d'évaluation autour de la spatialisation des extrêmes a également été l'occasion de s'intéresser à la validation de l'utilisation des séries de pluies journalières SPAZM pour estimer les précipitations extrêmes. Le travail de Gottardi (2009) avait permis de valider cet interpolateur pour la construction de champs de pluies journaliers et pour le calcul de volumes annuels, mais il n'avait jamais fait l'objet d'une étude précise pour les valeurs extrêmes. Les résultats obtenus ont montré que, comme beaucoup de méthodes d'interpolation, la méthode SPAZM provoque un effet de « lissage » et a tendance à sous-estimer les pluies extrêmes en moyenne de 10% sur les pluies de temps de retour supérieur à 10 ans. Une correction moyenne de cette sous-estimation permet d'atteindre des estimations de pluies extrêmes tout à fait acceptables. La validation de la méthode en contexte purement régional, avec les données du projet Extra-Flo, conclut sur de bonnes performances de la méthode, à égalité avec une loi GEV en mode régional.

Un travail approfondi, visant à construire une correction plus efficace qu'une majoration moyenne de 10% sur les quantiles extrêmes, a mis en évidence que l'interpolation SPAZM

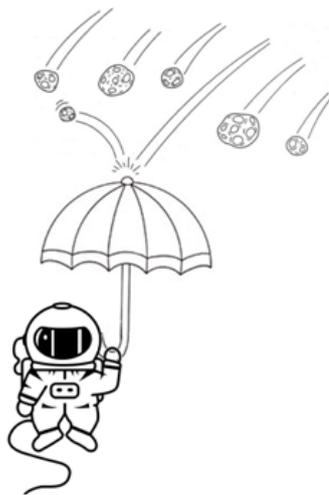
fournit une estimation des pluies journalières satisfaisante concernant le moment d'ordre 1 (moyenne), mais sous-estime souvent leur variance. Or l'écart-type de la population des pluies est directement lié au gradex des pluies. Un krigeage des écarts-types des pluies est alors réalisé en parallèle de façon à fournir un estimateur non biaisé du moment d'ordre 2. Cette carte des écarts-types permet une correction de SPAZM en tout point du domaine et fournit une correction nulle au droit des postes où les estimations SPAZM sont exactes. Une amélioration éventuelle, du krigeage des écarts-types permettrait d'améliorer significativement les estimations de quantiles. Finalement, une archive de distributions MEWP au km² sur toute la zone SPAZM a été construite. Des cartes de quantiles aux temps de retour 10, 100 et 1 000 et 10 000 ans en ont été extraites.

L'ensemble des résultats obtenus sont valables pour les précipitations journalières. Un travail de recherche similaire mais pour les pas de temps infra-journaliers devrait débiter. Cette perspective offrirait notamment l'occasion d'envisager une approche régionale pour prendre en compte l'utilisation d'une loi de Pareto généralisée par type de temps dans la composition des lois par types de temps et d'étudier la cohérence entre les estimations de forts quantiles aux différents pas de temps.

Les pluies journalières spatiales pour l'estimation d'extrêmes

La partie précédente a permis de construire des cartes de distributions de pluies journalières sur toute la zone d'étude au pixel du km^2 . Ces distributions permettent de décrire le risque de pluie en un point donné. Cependant, la méthode SCHADEX vise à estimer les quantiles de débits extrêmes à l'exutoire d'un bassin versant. Pour cela, l'objet utilisé pour décrire l'aléa pluviométrique est la pluie spatiale.

La pluie spatiale est la lame d'eau totale collectée par le bassin à chaque pas de temps et répartie uniformément sur celui-ci. Dans un cas parfait où des mesures de pluies journalières seraient disponibles en tous les pixels du bassin, nous pourrions déterminer une série de pluie spatiale journalière pour ce bassin. Chaque jour la valeur de la pluie spatiale serait donnée par la moyenne des valeurs observées en tous les pixels du bassin versant. En réalité, les pluies ne sont souvent accessibles qu'au travers d'un réseau de mesures ponctuelles. Dans cette partie, nous allons étudier l'impact des choix réalisés lors de la construction de pluies spatiales sur l'estimation des pluies extrêmes du bassin. Nous pouvons notamment nous interroger sur la sensibilité des méthodes aux postes disponibles et à la façon de combiner ces informations.



SOMMAIRE DU CHAPITRE

IV. 1 Problématique et objectifs	121
IV. 1. 1 La pluie spatiale avec les polygones de Thiessen	122
IV. 1. 2 La pluie spatiale avec SPAZM	123
IV. 2 Comparaison des deux approches pour l'estimation de pluies spatiales	124
IV. 2. 1 La base de données pour les pluies spatiales	124
IV. 2. 2 Comparaison des cumuls annuels des pluies spatiales	126
IV. 2. 3 Comparaison des estimations des pluies extrêmes	128
IV. 2. 4 Sensibilité de Thiessen aux postes utilisés : Cas de l'Eau d'Olle	132
IV. 3 Utilisation du générateur de champs de pluies SAMPO	135
IV. 3. 1 Présentation du générateur SAMPO	135
IV. 3. 2 Données d'entrée pour le générateur	135
IV. 3. 3 Étude de l'abattement par simulation	138
IV. 4 Conclusion	150

IV. 1 Problématique et objectifs

De nombreuses méthodes ont été proposées pour calculer la pluie spatiale d'un bassin. Nous pouvons regrouper les méthodes en quatre classes principales, sachant que chacune présente des variantes :

1. **Moyenne arithmétique.** C'est la méthode la plus simple. Il s'agit de calculer la moyenne arithmétique des différentes stations, en considérant l'information apportée par chaque pluviomètre comme le résultat d'un sondage classique, i.e. tirage d'une observation indépendante des autres dans une large population. Cependant, cette hypothèse pose un problème d'échantillonnage spatial.
2. **Moyenne pondérée.** Cette méthode représente une évolution de la précédente. Dans cette approche, chaque pluviomètre est pondéré par un coefficient, tenant compte de la densité des postes et de leur répartition sur le bassin. Pour le calcul des coefficients de pondération, il est possible d'utiliser, par exemple, la méthode des polygones de Thiessen.
3. **Méthode des isohyètes.** Il s'agit de déterminer des zones d'égale pluviométrie à partir du réseau de mesure. La pluie de bassin est donc calculée en sommant la pluie moyenne tombée entre deux isohyètes¹ pondérée par la superficie comprise entre ces deux courbes de niveau.
4. **Méthodes dites « géostatistiques ».** Dans cette approche, le bassin est discrétisé selon une grille spatiale. En chaque point de cette grille, la hauteur de précipitation est estimée, par exemple en interpolant spatialement par krigeage les mesures ponctuelles des postes les plus proches au point considéré. Les différentes valeurs sont ensuite agrégées pour former la pluie spatiale du bassin. La méthode SPAZM fait partie de cette catégorie.
5. **Méthodes de simulations stochastiques.** Ces méthodes considèrent la pluie comme une fonction aléatoire dans l'espace et dans le temps. Une simulation aléatoire dans l'espace consiste à générer, à l'aide de techniques numériques, des réalisations de la fonction aléatoire dans un domaine, en respectant les propriétés statistiques de la fonction aléatoire considérée, généralement inférées à partir des observations disponibles. Un modèle numérique de simulation peut répondre au besoin de connaître la valeur de la fonction en tout point d'un domaine (en général, sur une grille régulière). Il faut noter que ce type d'approche diffère des méthodes d'interpolation par le fait que les valeurs simulées ne sont pas les meilleures estimations de la grandeur étudiée, mais elles sont des réalisations de la fonction aléatoire qui reproduisent le mieux les fluctuations du phénomène (Journel et Huijbregts, 1978).

Il est important de spécifier que la « véritable » pluie spatiale est inconnue. Quelle que soit la méthode utilisée pour construire la pluie spatiale d'un bassin, il est légitime de s'interroger sur la cohérence de cette pluie du point de vue du volume et de l'estimation des extrêmes qui seront les deux critères importants pour juger de leur pertinence pour réaliser une étude SCHADDEX.

Concernant l'estimation des extrêmes, la pluie spatiale doit correctement intégrer l'abatement spatial sur les précipitations. Lorsque l'on s'intéresse à la pluie de temps de retour T sur le bassin versant, il est important de faire la distinction entre la pluie de temps de retour T estimée à partir de la pluie spatiale représentative du bassin et les pluies de temps de retour T estimées en chaque pixel du bassin (i.e. les cartes construites dans le chapitre précédent, page 105). En effet, la probabilité pour avoir une pluie de temps de retour T sur le bassin est différente de la

1. Ligne reliant des points d'égales quantités de précipitations tombées en une période déterminée

probabilité d'avoir une pluie de temps de retour T à chaque pixel du fait de la conjonction des aléas et de la dépendance spatiale des pluies en chaque pixel.

Considérons un bassin versant de surface S . Supposons que chaque pixel reçoive une pluie centennale. Le temps de retour de la pluie spatiale correspondante à cette situation est bien supérieur à 100 ans, car il est très peu probable qu'une pluie centennale tombe sur tous les pixels du bassin en même temps. Vue du bassin, une pluie centennale sur la surface S peut être obtenue par une conjonction de pluie de temps de retour inférieurs à 100 ans sur les pixels du bassin. Ce phénomène, appelé abattement spatial des précipitations, est d'autant plus marqué que le bassin est grand. L'abattement exprime la différence entre un risque de pluie ponctuelle et spatiale.

Un deuxième point à évaluer est la sensibilité des pluies spatiales aux mesures disponibles pour les construire. En effet, les pluies spatiales représentatives d'un bassin versant sont la plupart du temps construites à partir d'observations pluviométriques limitées en quelques points du bassin. Se pose alors la question de connaître la sensibilité des estimations de pluies extrêmes à cette disponibilité des données. Nous allons tenter de répondre à ces questions dans ce chapitre.

Deux méthodes ont été envisagées pour l'estimation des pluies spatiales : une pluie spatiale construite avec les polygones de Thiessen et une pluie spatiale construite avec la pluie SPAZM. Le principe de chacune d'entre elles est détaillé dans les paragraphes suivants.

IV. 1. 1 La pluie spatiale avec les polygones de Thiessen

La méthode du polygone de Thiessen (Thiessen, 1911) est la plus couramment utilisée du fait de sa facilité de mise en place. Cette méthode permet de réaliser une combinaison pondérée de N postes à proximité du bassin versant (de surface S). Chaque pixel du bassin est rattaché au pluviomètre le plus proche. Chaque pluviomètre i (i de 1 à N) dispose ainsi d'une zone d'influence d'aire S_i . La valeur $\lambda_i = \frac{S_i}{S}$ représente le facteur de pondération de la valeur locale. Les différentes zones d'influence sont déterminées par un simple découpage géométrique (la frontière entre deux postes est la médiatrice). La précipitation spatiale pour le bassin se calcule alors en effectuant la somme des précipitations P_i de chaque station, multipliées par leur facteur de pondération λ_i . La pluie spatiale au jour t sur le bassin s'écrit :

$$\begin{cases} PS_{Thiessen}(t) = \sum_{i=1}^N \lambda_i(t) \cdot P_i(t) \\ \sum_{i=1}^N \lambda_i(t) = 1 \end{cases} \quad (IV.1)$$

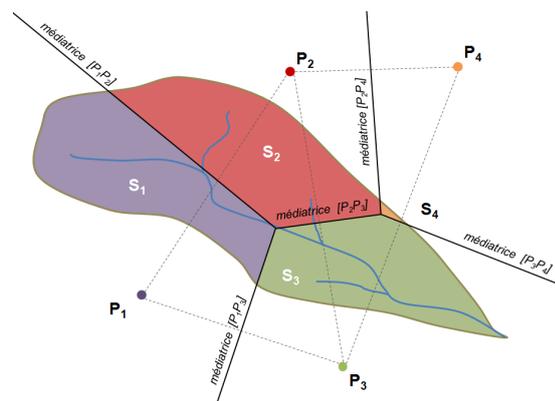


FIGURE IV.1 – Schéma illustratif pour la construction d'une pluie spatiale avec les polygones de Thiessen.

Normalement, les pondérations $\lambda_i(t)$ dépendent de la disponibilité des postes le jour t et devraient donc être recalculées selon cette disponibilité. Par souci de simplification, les poids λ_i sont calculés en considérant tous les postes (avec ou sans données). Ces poids λ_i sont indépen-

dants du temps. Les valeurs $\lambda_i(t)$ sont alors calculées en omettant les poids λ_i des postes sans données disponibles le jour t et faisant en sorte que la somme des $\lambda_i(t)$ vaille 1.

$$\begin{cases} \lambda_i(t) = \frac{1_{\mathbf{R}^+}(P_i(t)) \cdot \lambda_i}{\sum_{i=1}^N 1_{\mathbf{R}^+}(P_i(t)) \cdot \lambda_i} \\ 1_{\mathbf{R}^+}(x) = \begin{cases} 1 & \text{si } x \in \mathbf{R}^+ \\ 0 & \text{sinon} \end{cases} \end{cases} \quad (\text{IV.2})$$

La méthode des polygones de Thiessen peut facilement être mise en place, mais la pondération effectuée sur les différents postes est uniquement dépendante de leur disposition et non du relief ou d'autres grandeurs pouvant influencer la pluviométrie du bassin.

IV. 1. 2 La pluie spatiale avec SPAZM

Une deuxième approche pour estimer une pluie spatiale, plutôt que d'utiliser une combinaison de postes, est d'utiliser les champs de pluies journalières interpolées avec SPAZM. La pluie spatiale au jour t calculée avec SPAZM est une moyenne sur tous les pixels du bassin des valeurs de pluies journalières SPAZM pour ce jour t . Une série de pluie spatiale journalière est obtenue sur une période pour laquelle la donnée SPAZM est disponible. Contrairement à la pluie de Thiessen qui est une moyenne pondérée de N postes, la pluie SPAZM est une moyenne des n_p pixels du bassin.

$$PS_{SPA\text{ZM}}(t) = \frac{1}{n_p} \sum_{p=1}^{n_p} P_{SPA\text{ZM}}(p, t) \quad (\text{IV.3})$$

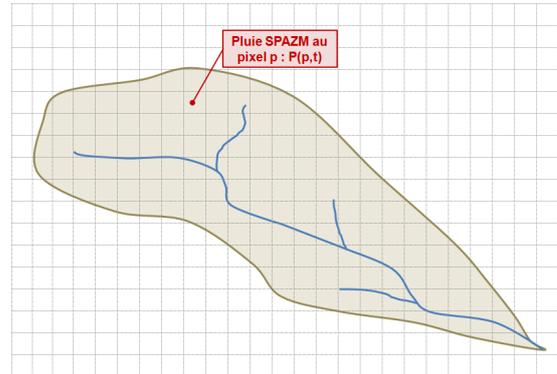


FIGURE IV.2 – Schéma illustratif pour le calcul d'une pluie spatiale avec la pluie SPAZM.

La méthode SPAZM a l'avantage de fournir une estimation des champs de pluies journaliers. Le calcul de la pluie spatiale intègre donc l'abattement par agrégation de l'information jour par jour sur les pixels du bassins. De plus SPAZM tient compte de l'information du relief. Cependant, comme toutes les méthodes d'interpolation spatiale s'appuyant sur une combinaison d'informations, un effet de lissage des champs est présent et induit une sous-estimation de la variabilité des pluies.

IV. 2 Comparaison des deux approches pour l'estimation de pluies spatiales

L'objectif du travail qui va suivre est de comparer les deux approches pour l'estimation de pluies spatiales (polygones de Thiessen et SPAZM) et de voir les éventuelles différences. Nous pouvons notamment nous poser plusieurs questions comme l'impact du nombre de postes utilisés avec Thiessen ? L'influence du schéma de pondération utilisé ? Quelle est la conséquence d'une correction en volume sur les pluies extrêmes ?

IV. 2. 1 La base de données pour les pluies spatiales

Nous disposons d'une base de 1956 bassins versants² d'une surface de 1 à 15 000 km^2 . Les bassins sont issus d'un regroupement de la base EDF et de la base IRSTEA. La base de bassins IRSTEA provient d'un algorithme d'extraction mis en place dans la thèse de Sauquet (2000). Les bassins versants sont construits en s'appuyant sur un modèle numérique de terrain (MNT) et le réseau hydrographique de la base BD Carthage. Pour un exutoire, le bassin versant associé peut être défini en remontant les lignes de crêtes avec le relief du MNT. Tous ces bassins sont ensuite répartis selon six classes de surfaces (figure IV.3) :

- de 0 à 100 km^2 (666 bassins) ;
- de 100 à 500 km^2 (889 bassins) ;
- de 500 à 1500 km^2 (262 bassins) ;
- de 1500 à 2500 km^2 (57 bassins) ;
- de 2500 à 5000 km^2 (45 bassins) ;
- de 5000 km^2 et plus (37 bassins).

Une pluie spatiale SPAZM et une pluie spatiale Thiessen vont être calculées pour chacun de ces bassins. Le réseau de pluviomètres utilisé est le réseau qui fonde l'interpolateur SPAZM (figure III.1, page 51) afin de disposer d'une information commune pour les deux méthodes. La période 1961-1981 est la période où le pourcentage de données disponibles est le plus élevé (figure IV.4) avec au moins 75% des données disponibles. Ceci permettra également d'éviter au maximum la présence de valeurs de pluies de Thiessen issues d'une réattribution des poids des stations où les données manquent. En comparant les pluies spatiales sur cette période, la disponibilité de données assure des ajustements MEWP robustes.

Les pluies spatiales Thiessen ont été construites en utilisant les postes présents dans un cercle de rayon égal à la longueur du rectangle circonscrit au bassin et centré sur le centroïde du bassin. Seuls les bassins avec au moins un poste dans cette zone ont été conservés (soit 1553 bassins sur la base des 1956 disponibles). Ce choix vise à garantir que nous disposons d'un minimum d'information à proximité du bassin pour construire une pluie de Thiessen acceptable.

2. La base est composée de bassins versants et de bassins versants intermédiaires.

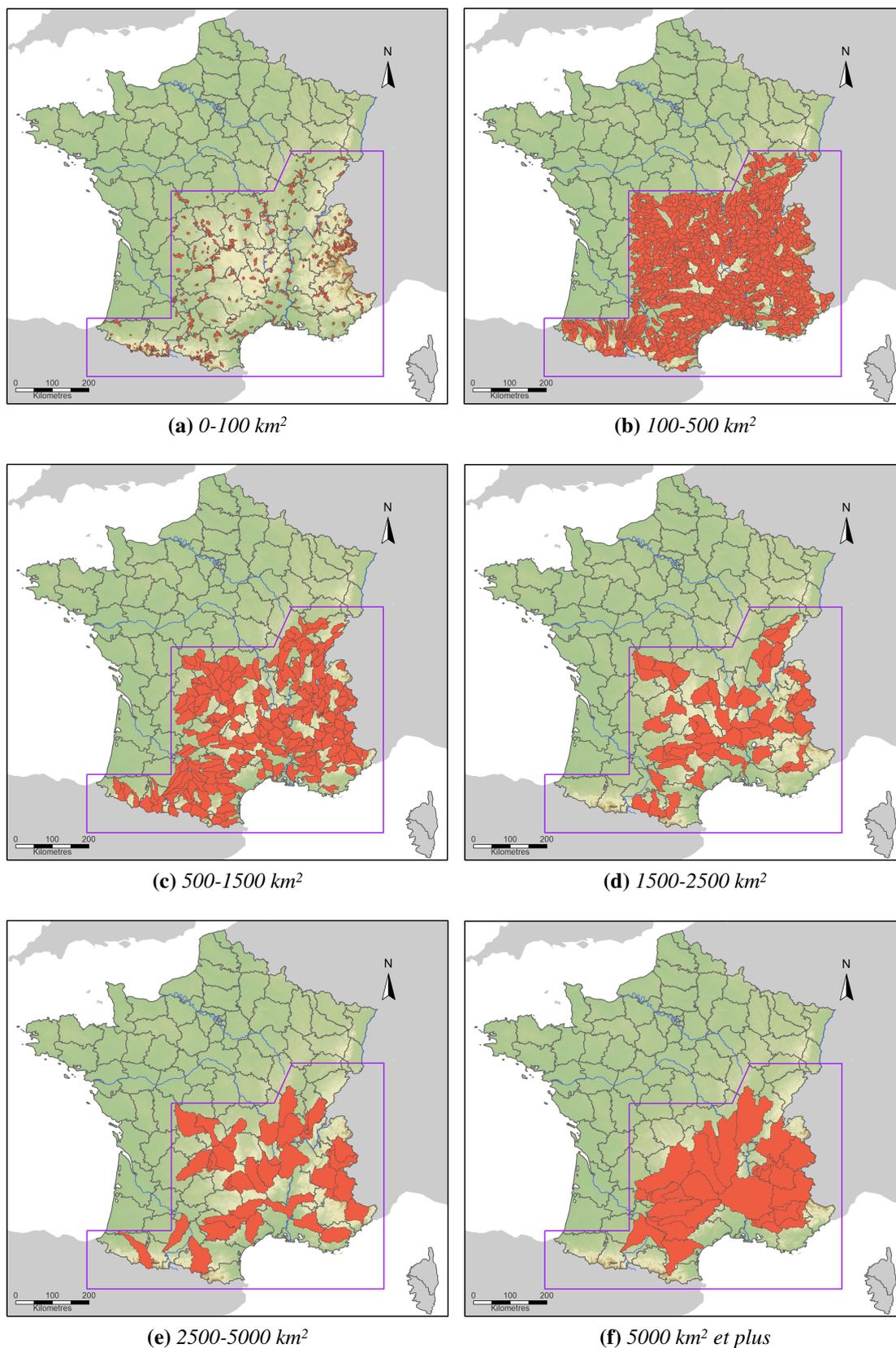


FIGURE IV.3 – Cartes des bassins versants par classe de surfaces.

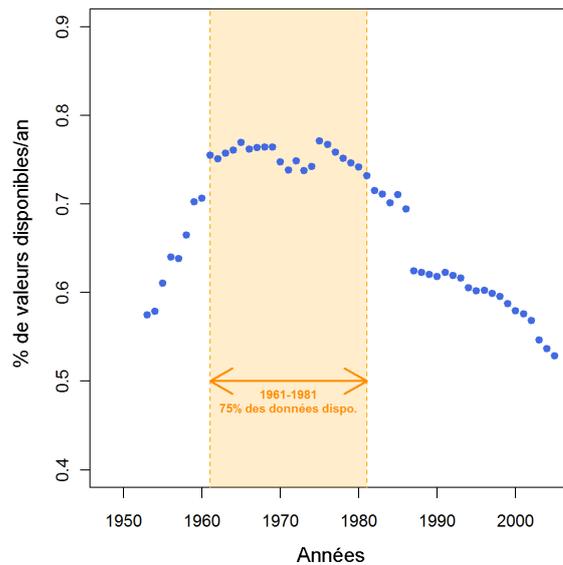


FIGURE IV.4 – Disponibilité des données selon les années.

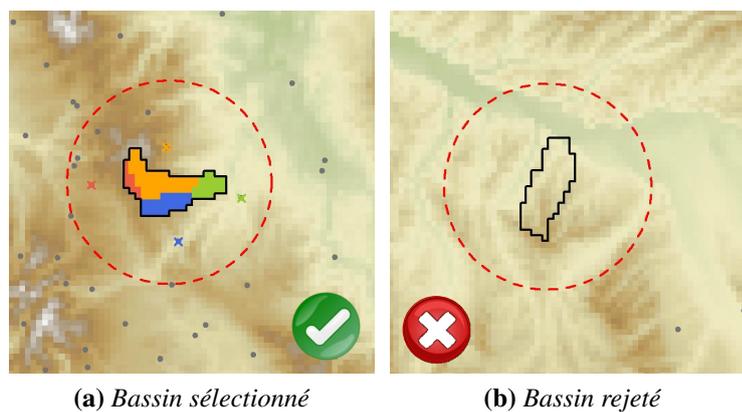


FIGURE IV.5 – Sélection des bassins selon la disponibilité d'information (les points gris représentent les postes disponibles).

IV. 2. 2 Comparaison des cumuls annuels des pluies spatiales

Avant de comparer les estimations de précipitations extrêmes avec les pluies spatiales de Thiessen ou SPAZM, nous commençons par nous intéresser au cumul annuel. Les précipitations annuelles moyennes des pluies SPAZM sont parfois différentes de celles des pluies Thiessen. La pluie de Thiessen est donc corrigée par un coefficient multiplicatif afin de la ramener au volume de la pluie SPAZM qui permet normalement un meilleur bilan que Thiessen (Gottardi, 2009). Cette correction a pour but de permettre une comparaison des estimations de pluies extrêmes à volume global précipité égal.

Sur la figure IV.6, nous constatons que la différence sur les cumuls annuels des pluies est en partie expliquée par le fait que l'altitude de la pluie Thiessen (moyenne des altitudes des postes utilisés pondérées par les poids de Thiessen) ne représente pas l'altitude médiane du bassin. Ces différences sont d'autant plus fortes que la variabilité du relief est forte et que la densité de réseau est faible.

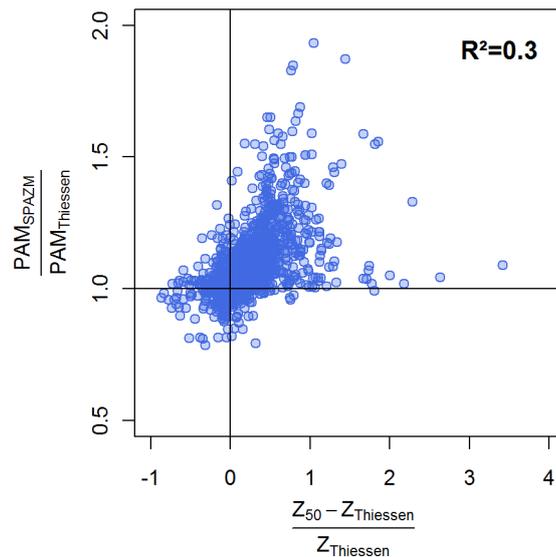


FIGURE IV.6 – Ratio des volumes annuels SPAZM sur ceux de Thiessen selon l'écart entre l'altitude médiane du bassin et l'altitude du panier Thiessen.

La correction de Thiessen en volume vise à ramener la pluie spatiale Thiessen au volume de la pluie SPAZM afin de ne pas subir l'influence d'une sous-estimation du volume de pluie sur l'estimation des extrêmes. Nous espérons notamment corriger le problème de représentativité de l'altitude du bassin avec Thiessen, car le bilan SPAZM tient compte de l'effet du relief. Cependant, la correction appliquée est la même pour tous les jours d'une série. Ainsi de fortes pluies seront corrigées de la même manière que des pluies fréquentes.

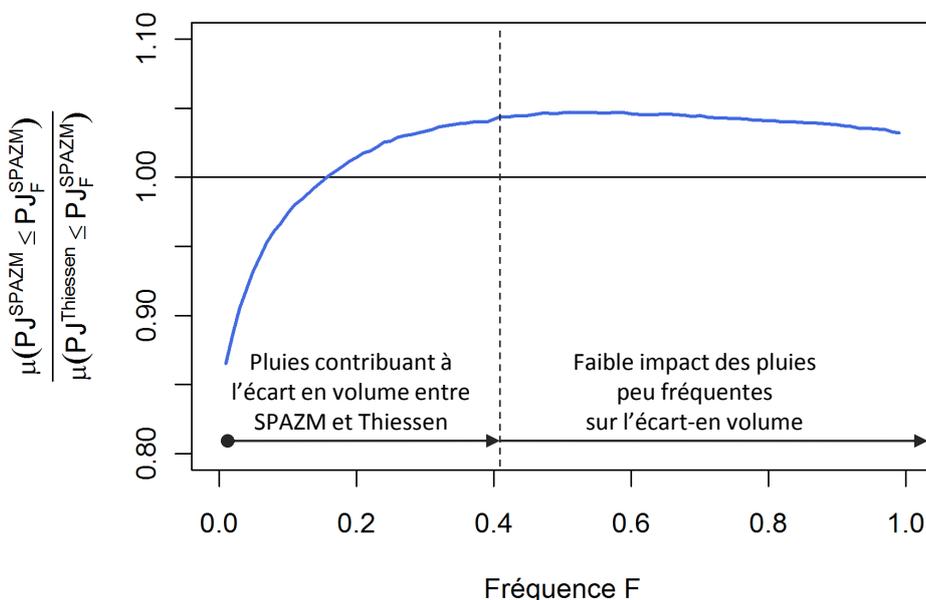


FIGURE IV.7 – Ratio de la pluie journalière moyenne SPAZM sur Thiessen en fonction de la fréquence (moyenne sur toute la base de bassins).

La figure IV.7 permet d'illustrer le problème d'une correction globale du volume. L'axe des abscisses est la fréquence en non-dépassement F . Le quantile empirique SPAZM PJ_F^{SPAZM}

associé à la fréquence F est alors calculé. Nous déterminons alors la pluie journalière moyenne des pluies non nulles et inférieures au quantile $PJ_F^{SPA\text{ZM}}$ avec la pluie SPAZM et Thiessen. Le ratio de la pluie moyenne SPAZM sur la pluie moyenne Thiessen pour la fréquence F est calculé pour tous les bassins de la base. Nous constatons que seules les pluies suffisamment fréquentes ($F < 0.4$) contribuent aux différences de volume entre les deux pluies spatiales. Les pluies plus rares ont souvent un poids moins fort sur le cumul annuel (sauf dans des zones où il pleut très peu et où la majeure partie du volume est dû à ces fortes pluies³ i.e. dans la zone méditerranéenne). La correction en volume utilisée pour corriger la pluie spatiale Thiessen vise à rattraper une mauvaise estimation du moment d'ordre 1 sur la pluie, mais rien ne prouve que cette correction corrige aussi une mauvaise estimation des fortes valeurs qui sont elles plutôt liées au moment d'ordre 2. Nous pouvons alors nous demander quel sera l'impact de cette correction des volumes sur l'estimation des pluies extrêmes.

IV. 2. 3 Comparaison des estimations des pluies extrêmes

Pour chaque bassin versant, les séries de pluies journalières spatiales SPAZM et Thiessen sont disponibles de 1961 à 1981. La pluie de Thiessen a été corrigée de façon à avoir le même cumul annuel que la pluie SPAZM. Les ajustements MEWP sur les deux pluies spatiales fournissent des estimations du risque de pluie sur chaque bassin versant. Les ajustements ont été réalisés avec une unique saison de 12 mois et un seuil de 70% pour l'ajustement des exponentielles sup-seuil par type de temps.

Afin de comparer les ajustements obtenus sur les deux pluies spatiales, nous nous intéresserons aux quantiles obtenus avec les deux distributions MEWP :

$$K_{P(T)} = \frac{PJ(T)_{\text{Thiessen}}}{PJ(T)_{\text{SPA\text{ZM}}}} \quad (\text{IV.4})$$

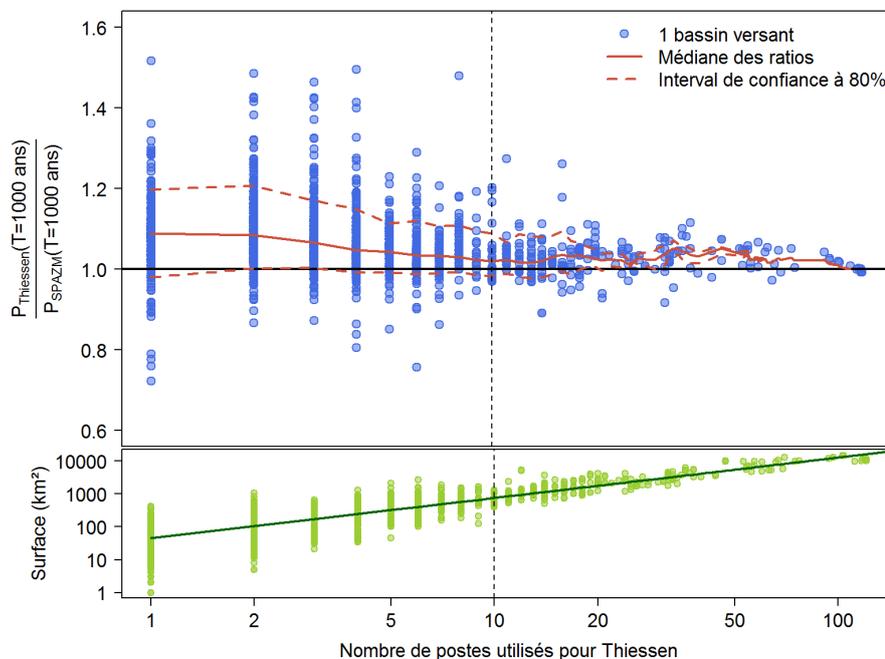


FIGURE IV.8 – Ratio de la pluie millénaire Thiessen corrigée sur celle de SPAZM selon le nombre de postes avec Thiessen.

3. Ce sont les zones identifiées par l'indicateur \overline{PJX}/PAM présenté en section III. 4, page 109.

Sur la figure IV.8, l'axe des abscisses représente le nombre de postes utilisés pour le calcul de la pluie de Thiessen. Les points verts représentent la surface des bassins et la droite verte la médiane de ces surfaces (pour un nombre de postes donné). Nous voyons que le nombre de postes utilisé dans Thiessen est très lié à la taille du bassin versant. Les points bleus correspondent à un bassin versant et ont pour ordonnée le ratio de la pluie millénale Thiessen corrigée par la pluie millénale SPAZM $K_{P(T=1000\text{ans})}$. Les courbes rouges représentent la courbe médiane (en trait plein) et les bornes de l'intervalle 80% (en pointillés). La même analyse pourrait être faite pour des temps de retour 10, 100 et 10 000 ans avec des conclusions similaires.

Globalement, les estimations Thiessen sont supérieures à celles avec la pluie SPAZM, mais l'analyse du graphe peut être décomposée en deux parties :

- ✓ **Les pluies Thiessen avec plus de 10 postes.** Lorsque Thiessen utilise plus de 10 postes, soit pour des surfaces de bassins supérieures à 700 km², les estimations de pluies millénales sont assez similaires et les ratios $K_{P(T=1000\text{ans})}$ sont très peu variables. Il est donc raisonnable de penser que les pluies spatiales ainsi construites fournissent une estimation proche de la pluie spatiale réelle.
- ✗ **Les pluies Thiessen avec moins de 10 postes.** Dans le cas où Thiessen est construit avec moins de 10 postes, les ratios $K_{P(T=1000\text{ans})}$ sont très variables. Globalement, les estimations SPAZM sont plus faibles que les estimations Thiessen corrigées. Pour les pluies Thiessen à 1 ou 2 postes, SPAZM donne une valeur plus faible en moyenne de 10%, et pour 80% des bassins un ratio $K_{P(T=1000\text{ans})}$ compris entre 1 et 1.2. Cette valeur de 10% obtenue avec un poste correspond à la sous-estimation de SPAZM sur les estimations de pluies extrêmes ponctuelles de la partie précédente. La forte variabilité des $K_{P(T=1000\text{ans})}$ peut avoir plusieurs raisons :
 - Sur la figure IV.8, la population de bassins sur lesquels Thiessen ne dispose que d'un seul poste regroupe des bassins de surfaces très variables (de 1 à près de 1000 km²). Cette forte variabilité dans la gamme de surface se retrouve également sur le ratio des quantiles de pluies millénales. Le **dispersion du réseau** peut alors être le critère majeur des écarts entre les méthodes dans de tels cas.
 - Avec peu de postes ou pour des surfaces de bassin plus petites, le risque de mal décrire le risque de pluie est d'autant plus important. En particulier, les petits bassins sont souvent des bassins de montagne avec une forte variabilité des champs de pluies. Il est alors courant de constater que les pluies spatiales sont uniquement construites avec des postes en fond de vallée ou se trouvant dans une vallée adjacente soumise à une climatologie différente. Dans ce cas, les écarts entre méthodes sont plutôt à expliquer par le **manque de données descriptives de la pluie perçue par le bassin**. Pour de tels cas SPAZM a l'avantage de tenter une extrapolation avec le relief et qui, au vu de la partie précédente, est capable de réaliser une estimation des extrêmes cohérente.
 - Une dernière raison à ces écarts peut être la façon dont sont calculées les pluies spatiales. En effet, pour de grands bassins, l'abattement induit par le mode de calcul de SPAZM et de Thiessen est proche de la réalité (suffisamment de postes, moyenne sur une grande surface). Pour des bassins plus petits, le comportement entre SPAZM et Thiessen peut être très différent. Le cas extrême étant un bassin de quelques km² avec un poste à proximité. La pluie Thiessen sera la pluie de ce poste alors que la pluie SPAZM s'appuiera sur un panier d'une dizaine de postes plus ou moins proches du bassin. La pluie de Thiessen est alors équivalente à une pluie ponctuelle et la pluie SPAZM correspond à une pluie spatiale sur une surface bien supérieure à celle du bassin. Ce problème peut également se retrouver lorsqu'un des pluviomètres

utilisés par Thiessen porte une pondération bien supérieure aux autres. Le problème est alors de savoir quelle méthode génère le meilleur **abattement pour la pluie spatiale**.

Afin de creuser plus en avant les résultats du graphe IV.8, nous proposons de détailler les résultats précédents en catégorisant les bassins selon les corrections en volume appliquées à Thiessen et les résultats sur les quantiles de pluies millénaux. Pour chaque classe de surface, les bassins ont été répartis dans 9 catégories selon la correction en volume appliquée à la pluie Thiessen et selon la valeur de $K_P(T = 1000 \text{ ans})$ (ratio entre la pluie millénaire de Thiessen corrigée et de SPAZM). Les tableaux IV.1 donnent le pourcentage de bassin de chaque classe de surface appartenant à chacune des catégories.

		$PA_{Thiessen}/PA_{SPAZM}$		
		> 1.1	[0.9, 1.1]	< 0.9
$K_{P(T=1000 \text{ ans})}$	> 1.1	0%	22%	11%
	[0.9, 1.1]	0%	41%	23%
	< 0.9	0%	1%	2%

(a) De 100 km² à 500 km²

		$PA_{Thiessen}/PA_{SPAZM}$		
		> 1.1	[0.9, 1.1]	< 0.9
$K_{P(T=1000 \text{ ans})}$	> 1.1	1%	22%	10%
	[0.9, 1.1]	2%	54%	12%
	< 0.9	0%	0%	1%

(b) De 500 km² à 1500 km²

		$PA_{Thiessen}/PA_{SPAZM}$		
		> 1.1	[0.9, 1.1]	< 0.9
$K_{P(T=1000 \text{ ans})}$	> 1.1	0%	10%	6%
	[0.9, 1.1]	0%	60%	22%
	< 0.9	0%	0%	1%

(c) De 1500 km² à 2500 km²

		$PA_{Thiessen}/PA_{SPAZM}$		
		> 1.1	[0.9, 1.1]	< 0.9
$K_{P(T=1000 \text{ ans})}$	> 1.1	0%	7%	5%
	[0.9, 1.1]	0%	67%	21%
	< 0.9	0%	0%	0%

(d) De 2500 km² à 5000 km²

		$PA_{Thiessen}/PA_{SPAZM}$		
		> 1.1	[0.9, 1.1]	< 0.9
$K_{P(T=1000 \text{ ans})}$	> 1.1	0%	0%	0%
	[0.9, 1.1]	0%	74%	26%
	< 0.9	0%	0%	0%

(e) Plus 5000 km²

Tableau IV.1 – Répartition des ratios en volumes et en quantiles millénaux selon la classe de surface.

Les volumes ou les quantiles sont considérés comme comparables si le ratio est compris entre 0.9 et 1.1. Plusieurs constats peuvent être réalisés :

- Toutes classes de surface confondues, les pluies SPAZM et Thiessen sont cohérentes dans 80% des cas du point de vue de l'estimation de la pluie millénale.
- **SPAZM ne trouve jamais des volumes plus faibles que Thiessen** ce qui est plutôt rassurant puisque la méthode SPAZM a été calibrée pour respecter au mieux les volumes en tenant compte du relief⁴.
- Quelle que soit la classe de surface, dans au moins 65% des cas, **SPAZM et Thiessen sont cohérents en volume**. De plus, plus les bassins sont grands, plus SPAZM et Thiessen sont en accord (75% des cas pour les plus grands bassins).
 - Dans plus de 2/3 des cas, les pluies SPAZM et Thiessen sont aussi cohérentes en quantiles. Dans moins de 1/3 des cas, SPAZM donne des quantiles plus faibles que Thiessen (lissage de SPAZM ou une pluie Thiessen avec une variance plus forte du fait du faible nombre de postes utilisés pour les petits bassins).
 - À volume égal, SPAZM ne donne jamais de quantiles plus forts que Thiessen sauf pour quelques bassins. Ce sont, dans la plupart des cas, des bassins d'altitude pour lesquels SPAZM extrapole alors que Thiessen utilise des postes de basse altitude.
- **SPAZM trouve un volume plus fort que Thiessen** dans environ 35% des bassins (moins pour les bassins de grande surface). La pluie Thiessen est alors corrigée pour être ramenée au volume SPAZM.
 - Dans 2/3 des cas, la correction en volume permet une estimation des quantiles avec Thiessen similaire à celle de SPAZM.
 - Dans 1/3 des cas SPAZM, donne des quantiles plus faibles qu'avec Thiessen. La correction peut conduire à une augmentation trop forte des quantiles Thiessen. Les raisons à ces différences peuvent être que SPAZM génère un abattement trop fort (lissage des champs) ou que Thiessen n'en génère pas assez (e.g. utilisation d'un seul poste) ou bien que la correction du cumul annuel de Thiessen n'est pas adaptée aux pluies extrêmes.
 - Dans de très rares cas, les quantiles et volumes SPAZM sont supérieurs à ceux de Thiessen. Les bassins où sont effectués ce genre de constat sont souvent des bassins où les postes utilisés pour Thiessen se trouvent en fond de vallée et ne représentent pas le risque sur le bassin. SPAZM – qui tient compte de l'effet du relief – fournit alors une estimation plus crédible aussi bien en volume qu'en pluies extrêmes. Avec une pluie Thiessen aussi peu représentative du bassin la correction en volume conduit à des estimations de pluies extrêmes souvent très fortes.
- Dès que les bassins sont suffisamment grands, SPAZM et Thiessen sont de plus en plus similaires en quantiles de pluies et les corrections en volumes appliquées semblent justifiées (observations faites précédemment sur la figure IV.8).

Afin d'illustrer les différents problèmes abordés dans la comparaison de la pluie SPAZM et Thiessen sur la base de bassins, nous allons étudier un cas représentatif.

4. La pondération et les altitudes des postes de Thiessen entraînent souvent une sous-estimation de l'altitude médiane du bassin.

IV. 2. 4 Sensibilité de Thiessen aux postes utilisés : Cas de l'Eau d'Olle

Le bassin de l'Eau d'Olle à La Pernière (169 km²), de par sa situation sur l'envers du massif de Belledonne (Alpes) et son orientation, est très arrosé (figure IV.9 (a)). Il est exposé à tous les régimes et reçoit notamment de fortes précipitations en automne en présence d'un flux de sud-ouest chaud à l'avant d'une perturbation atlantique. Le cumul de pluie annuel est plus élevé sur la partie sud-ouest du bassin qui est plus exposée à ces flux et de façon moins importante sur les reliefs à l'est du bassin (figure IV.9 (b)). La hauteur de ses reliefs favorise également les pluies orageuses d'été qui peuvent être très intenses. La carte IV.9 (c) représente les quantiles de pluie de temps de retour 1000 ans. Les risques de pluies les plus forts se trouvent sur le sud-ouest du bassin.

Trois pluies de Thiessen différentes (figure IV.9 (d), (e), (f)) ont été calculées sur ce bassin ainsi que la pluie spatiale SPAZM. Les zones de couleurs représentent la zone assignée au poste de même couleur par la méthode des polygones de Thiessen. Les distributions empiriques de la pluie SPAZM, Thiessen et Thiessen avec la correction en volume pour retrouver le volume SPAZM sont tracées sur les figures IV.9 (g), (h) et (i) pour chaque cas envisagé.

Dans le cas 1 et 2, deux postes ont été choisis pour construire la pluie Thiessen. Dans le cas 1, les postes *P1* et *P3* ont été sélectionnés pour décrire au mieux, le volume sur le bassin et le risque de pluie. La pluie Thiessen obtenue sous-estime le volume de 7% par rapport à SPAZM qui donne une pluie annuelle moyenne de 1720 mm. Les pluies spatiales Thiessen et SPAZM sont donc plutôt cohérentes en volume. Du point de vue des distributions empiriques, les pluies spatiales sont en accord.

Dans le cas 2, nous considérons que le poste *P1* n'est pas disponible. Nous utilisons à la place le poste *P2* situé à 6 km au sud, dans la même vallée. Ce poste a un risque de pluie un peu plus faible et un volume annuel très inférieur au poste *P1*. La pluie Thiessen obtenue sous-estime alors le volume de 44% par rapport à SPAZM alors que le biais sur le haut de la distribution empirique est plutôt de l'ordre de 20% entre la pluie SPAZM et Thiessen non corrigé en volume. Corriger le volume de Thiessen avec un facteur multiplicatif de 1.44 sur toutes les pluies conduit alors à une distribution des pluies bien au-dessus de celle de SPAZM.

Dans le cas 3, les trois postes *P1*, *P2* et *P3* participent pour construire la pluie Thiessen. La pluie Thiessen ne sous-estime alors plus que de 24% le volume et cette fois la correction en volume de Thiessen amène une cohérence avec SPAZM à la fois sur l'estimation du cumul annuel et sur la distribution des pluies.

Finalement, plusieurs choses sont à retenir de cet exemple. **Il est clair qu'une correction globale en volume sur une série de pluie, sans tenir compte de leur intensité, n'est absolument pas adaptée à une estimation des extrêmes.** La correction corrige certes le problème en volume, mais sa répercussion sur les fortes précipitations est très incertaine. Deux postes peuvent très bien observer des précipitations annuelles différentes tout en ayant des quantiles élevés de pluies comparables (et réciproquement).

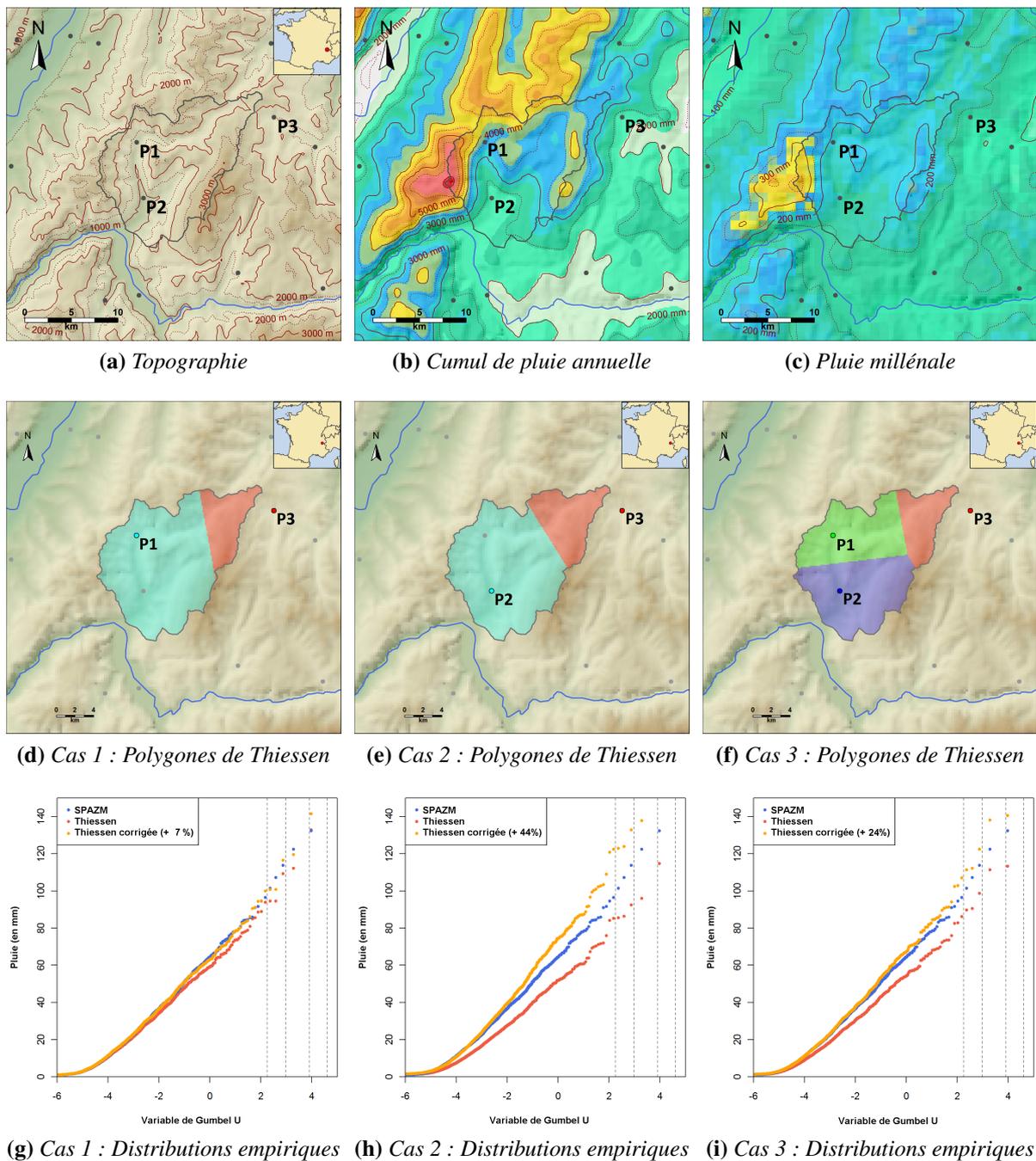


FIGURE IV.9 – Différentes pluies spatiales avec Thiessen sur le bassin de l'Eau d'Olle à La Périère (169 km²)

De plus, le cas de l'Eau d'Olle montre bien que, sur des petits bassins⁵, le choix des postes est très important pour réaliser une pluie Thiessen. Le cas 3 utilise trois postes et fournit une pluie de Thiessen moins crédible que le cas 1 avec deux postes choisis intelligemment. Une pluie de Thiessen est très sensible aux postes disponibles et choisis. Il faut être vigilant et s'attacher à construire une pluie représentative de la structure spatiale des précipitations annuelles et du risque de pluies. **Si aucune donnée plus descriptive du bassin n'est disponible, mieux vaut se fier à une pluie SPAZM qui assure un minimum de représentativité du point de vue de la structure spatiale et en particulier du relief.**

5. Petit bassin dans le sens où la surface du bassin est trop petite pour disposer d'au moins dix postes pour calculer la pluie spatiale Thiessen (nombre de postes à partir duquel SPAZM et Thiessen semblent cohérents, figure IV.8).

Bilan : globalement, dans 80% des cas, les pluies spatiales SPAZM et Thiessen corrigées en volume estiment des quantiles élevés de pluies analogues. Les conclusions sont cependant assez contrastées selon la classe de surface considérée.

Les pluies spatiales de Thiessen et SPAZM fournissent des estimations cohérentes pour des bassins versants de surfaces suffisamment grandes (supérieure à 700 km²). Si une majoration de la pluie Thiessen est nécessaire pour corriger le cumul annuel, elle améliore la plupart du temps la cohérence entre les estimations des quantiles extrêmes.

Au contraire, sur des petits bassins, la construction d'une pluie de Thiessen est beaucoup plus sensible au choix des stations du fait de la dispersion des postes autour du bassin et d'une éventuelle variabilité spatiale des pluies dans la zone (e.g. zone montagneuse). La correction en volume essaie alors de corriger un souci de représentativité de la pluie de Thiessen due à un manque d'information. La correction semble alors moins adaptée aux fortes précipitations et engendre souvent des estimations de pluies extrêmes avec Thiessen beaucoup plus fortes qu'avec SPAZM.

Finalement, la construction d'une pluie spatiale avec la méthode des polygones de Thiessen est très dépendante des données disponibles et des postes choisis. Ce problème est en partie dû au fait que l'attribution des pondérations pour la combinaison de postes est purement géométrique. Elles sont donc très dépendantes de la densité du réseau sol et de la variabilité spatiale des fortes pluies qui accentuent le risque de mal représenter la pluviométrie du bassin. L'utilisation de SPAZM est donc à privilégier pour calculer une pluie spatiale. En effet, la pluie SPAZM permet de prendre en compte des effets orographiques par types de temps et donc d'extrapoler les pluies plus intelligemment selon la variabilité des champs. De plus, elle permet de mieux respecter le cumul de pluie annuel. Cependant la méthode a tendance à sous-estimer les fortes pluies du fait de son processus d'interpolation. Ce phénomène de lissage peut générer un abattement des pluies plus fort que celui qui serait observé sur des champs réels surtout sur des bassins de petites surfaces.

La suite du travail va donc consister à évaluer l'abattement réalisé par SPAZM vis-à-vis de l'abattement réel qui devrait être observé. Nous nous attendons à ce que les résultats soient plus prévisibles que ce qui pourrait être observé avec des pluies de Thiessen qui intègrent une grande part de subjectivité dans leur construction.

IV. 3 Utilisation du générateur de champs de pluies SAMPO

Jusqu'à maintenant, nous nous sommes contentés de comparer les deux méthodes d'estimation de pluies spatiales, SPAZM et Thiessen. De cette comparaison sont ressortis certains avantages et inconvénients des deux méthodes. Thiessen s'est révélé être très sensible aux postes (que ce soit au choix des postes disponibles ou tout simplement à la disponibilité même de poste descriptif pour le bassin). La pluie SPAZM est apparue être une méthode de calcul de la pluie spatiale plus robuste cependant une question reste en suspens : Qu'en est-il de la justesse des pluies spatiales estimées avec SPAZM ? Sans connaissance des champs de pluie réels, il est difficile de juger de la justesse de SPAZM. Un simulateur de pluie permet de créer un environnement virtuel dans lequel ils sont connus. **L'objectif est donc maintenant d'estimer si les pluies estimées par SPAZM sont représentatives des pluies spatiales réelles en utilisant un simulateur de champs de pluie.**

IV. 3. 1 Présentation du générateur SAMPO

L'incertitude sur l'estimation des pluies spatiales est généralement due à une mauvaise description du champ qui n'est connu qu'au travers d'informations ponctuelles (Moulin *et al.*, 2009; Severino et Alpuim, 2005). La simulation de champs de pluie est une méthode géostatistique qui permet de simuler plusieurs champs plausibles d'après les mesures ponctuelles aux différents pluviomètres (Vischel *et al.*, 2009). La méthode de simulation utilisée dans cette partie est un générateur de pluie par bandes tournantes appelée SAMPO⁶.

Le générateur de pluie SAMPO (Domingues Ramos, 2002; Lepioufle, 2009; Renard *et al.*, 2011; Leblois et Creutin, 2013) permet de simuler des champs statistiquement cohérents avec les statistiques de pluies données en entrée du modèle. Des champs de pluies peuvent également être générés sur une période donnée avec un conditionnement par des mesures du réseau sol. Des pluies spatiales peuvent être calculées à partir des champs simulés et comparées à celles obtenues avec d'autres approches.

Le générateur SAMPO – dans son utilisation optimale – permet de prendre en compte des advections ou des cohérences temporelles d'une simulation à une autre. Dans l'étude qui va suivre, nous n'utiliserons pas ces modules. Le générateur SAMPO ne servira qu'à simuler des champs de pluies journaliers indépendants les uns des autres (éventuellement conditionnés par les données du jour) et respectant les caractéristiques statistiques de la pluie fournies en entrée.

IV. 3. 2 Données d'entrée pour le générateur

Pour pouvoir effectuer des simulations de champs de pluie, plusieurs informations sur les pluies sont nécessaires. Chacune des entrées de SAMPO va être détaillée et expliquée. Les entrées sont déterminées sur toute la population des pluies ou sur les pluies d'un type de temps donné selon que l'on souhaite, ou pas, prendre en compte une structure des pluies différente d'un type de temps à l'autre.

6. Simulation of Advected Mesoscale Precipitations and their Occurrence.

Variogrammes de l'intermittence

L'intermittence des pluies correspond au fait qu'il y ait des zones de pluie et de non-pluie, créant ainsi des discontinuités dans le champ de pluie. C'est une valeur au pixel qui est égale à 0 s'il n'a pas plu et égale à 1 s'il a plu. La structure spatiale de l'intermittence est décrite par un variogramme exponentiel ajusté sur les observations disponibles sur la zone de simulation. Un exemple de champ de d'intermittence est donné en figure IV.10.

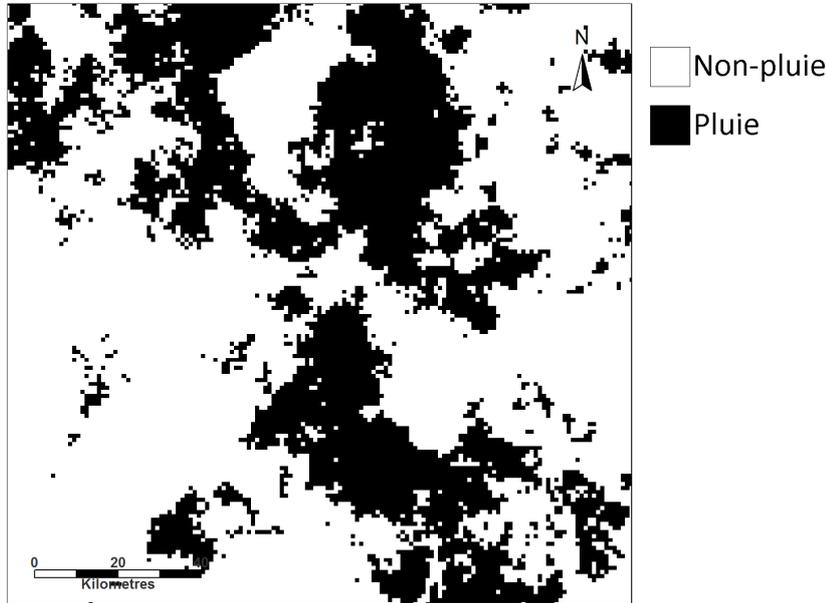


FIGURE IV.10 – Exemple de champ d'intermittence généré par SAMPO (en noir : les zones pluvieuses, en blanc : les zones où les pluies sont nulles).

Loi ponctuelle des pluies non nulles

Une fois la distinction faite entre les zones avec pluie et les zones sans pluie, il faut s'intéresser à la description des pluies non nulles. Tout d'abord, il faut choisir une loi représentative des pluies à un pixel donné. La loi ponctuelle des pluies non nulles est supposée être une gaussienne inverse dans SAMPO :

$$F(x|\mu, \lambda) = \Phi \left(\sqrt{\frac{\lambda}{x}} \cdot \left(\frac{x}{\mu} - 1 \right) \right) + \exp \left(\frac{2\lambda}{\mu} \right) \cdot \Phi \left(-\sqrt{\frac{\lambda}{x}} \cdot \left(\frac{x}{\mu} + 1 \right) \right) \quad (\text{IV.5})$$

où $\mu > 0$ est son espérance et $\lambda > 0$ est un paramètre de forme tel que $\lambda = \frac{\mu^3}{\sigma^2}$. La fonction Φ est la fonction de répartition de la loi normale centrée réduite $\mathcal{N}(0, 1)$.

Ses paramètres peuvent être déterminés à partir de la moyenne et de l'écart-type des pluies non nulles observées sur la zone de simulation. Le générateur SAMPO n'utilise pas des valeurs de moyenne et d'écart-type ponctuelles pour la loi ponctuelle des pluies non nulles (stationnarité spatiale). Une gaussienne-inverse unique est définie pour toute la zone de simulation et ses paramètres sont estimés avec la moyenne et l'écart-type moyen sur cette zone.

Variogrammes des pluies non nulles

Dans un second temps, la structure spatiale des pluies non nulles est décrite par un variogramme exponentiel estimé avec les postes disponibles. Les champs de pluies non nulles sont alors simulés selon le principe de bandes tournantes et le champ obtenu est recoupé avec celui de l'intermittence. Un exemple de champ de pluie est donné sur la figure IV.11.

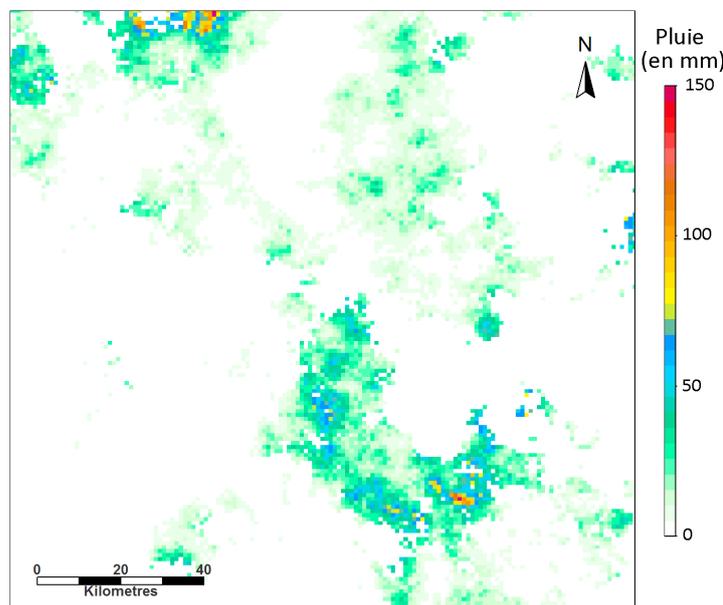


FIGURE IV.11 – Exemple de champ de pluie généré par SAMPO.

Conditionnement

Le générateur permet également de simuler des champs de pluie conditionnés avec le réseau d'observation. Les champs simulés sont alors contraints pour respecter les valeurs de pluies journalières observées par les postes au sol. Pour distinguer les deux types de champs, nous parlerons de champs libres pour les champs simulés sans conditionnement et de champs conditionnés pour les champs avec un conditionnement.

Le conditionnement des champs d'un simulateur est un problème complexe. Les champs conditionnés avec SAMPO sont calculés par déformation des champs libres simulés pour qu'ils respectent les valeurs imposées aux points de mesure.

Problèmes de concordance entre les lois ponctuelles des pluies non nulles avant et après conditionnement : Le générateur SAMPO utilise une gaussienne inverse comme distribution théorique pour les pluies non nulles. Le graphe IV.12 permet de comparer la fréquence empirique des pluies non nulles à la fréquence donnée par la gaussienne inverse de paramètres $\mu = \overline{P_{nn}}$ et $\lambda = \frac{\mu^3}{\sigma^2}$ (avec σ l'écart-type des pluies non nulles P_{nn}). La courbe bleue montre que la gaussienne inverse choisie pour la simulation ne représente pas parfaitement la distribution empirique des pluies non nulles observées dans la zone de simulation. La courbe rouge représente les pluies non nulles des champs libres. La fréquence empirique de ces pluies simulées respecte bien la loi théorique imposée. Nous pourrions nous attendre à ce que la prise en compte de données de conditionnement fournisse des pluies non nulles dont la distribution serait une distribution intermédiaire entre la gaussienne-inverse théorique imposée lors de la génération des champs libres et de la distribution empirique imposée par les postes utilisés pour le conditionnement. Cependant, nous observons que les pluies non nulles simulées avec un conditionnement

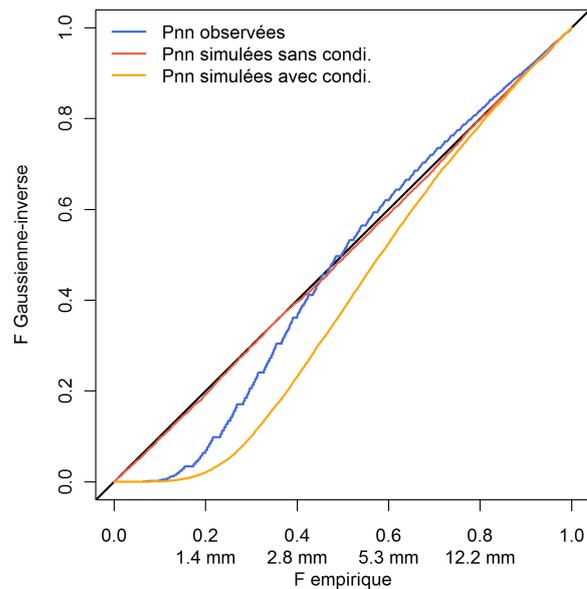


FIGURE IV.12 – Graphes fréquences-fréquences pour les différentes lois des pluies ponctuelles non nulles P_m .

sont encore plus éloignées de la gaussienne inverse que les pluies non nulles observées. Une explication plausible de cet effet est lié à l'anamorphose permettant de passer du champ gaussien au champ final dans le processus de simulation de SAMPO. En effet, la fonction d'anamorphose permet de passer d'un espace dans lequel les pluies non nulles suivent la gaussienne inverse donnée en entrée à un espace gaussien. Cependant, si les données de conditionnement ne suivent pas exactement la loi parentale (i.e. gaussienne inverse théorique), alors l'anamorphose n'est plus adaptée et le conditionnement du champ induit une mauvaise représentation de la loi des pluies. Dans la suite, l'utilisation de SAMPO se limitera donc à la génération de champs libres.

IV. 3. 3 Étude de l'abattement par simulation

Le but de cette partie est d'utiliser SAMPO pour simuler des champs de pluies et de considérer ces champs comme une représentation plausible de la variabilité spatiale de ces champs sur un domaine donné. Le simulateur SAMPO permet de produire un grand nombre de champs libres stationnaires respectant les propriétés statistiques des pluies données en entrée. Nous considérons des champs libres afin de nous affranchir d'éventuels problèmes dus au conditionnement. Les champs ne comportent donc aucun effet orographique et sont tous indépendants entre eux.

Construction de champs de pluie journaliers

Le générateur SAMPO est un outil pour simuler des champs de pluies statistiquement réalistes. Nous allons donc l'utiliser pour construire une base de champs de pluies journaliers de référence qui seront considérés comme représentatifs de la réalité (figure IV.13). Comme nous avons mis en évidence un problème de mauvaise représentation de la distribution des pluies non nulles lors de son utilisation avec un réseau sol pour le conditionner, SAMPO sera utilisé en génération libre (i.e. sans conditionnement par les données observées). La zone de simulation choisie est une zone de 150 kilomètres par 150 kilomètres. Les différents paramètres d'entrée sont calculés à partir d'une zone du sud des Cévennes de même taille. Les caractéristiques des

pluies déterminées sont reportées dans le tableau IV.2. La classification par type de temps n'est pas utilisée c'est-à-dire qu'une unique population de pluie journalière est considérée et non huit sous-populations avec chacune ses propres paramètres d'entrée.

L'intermittence des pluies		
Pourcentage de pluie nulle :	70%	
Variogramme exponentiel :	portée= 29 km	palier = 1
Les pluies non nulles		
Moments des P_{nn} :	$\overline{P_{nn}} = 9.41 \text{ mm}$	$\sigma(P_{nn}) = 14.64 \text{ mm}$
Variogramme exponentiel :	portée= 15.2 km	palier = 214 mm

Tableau IV.2 – Données d'entrée pour le générateur SAMPO.

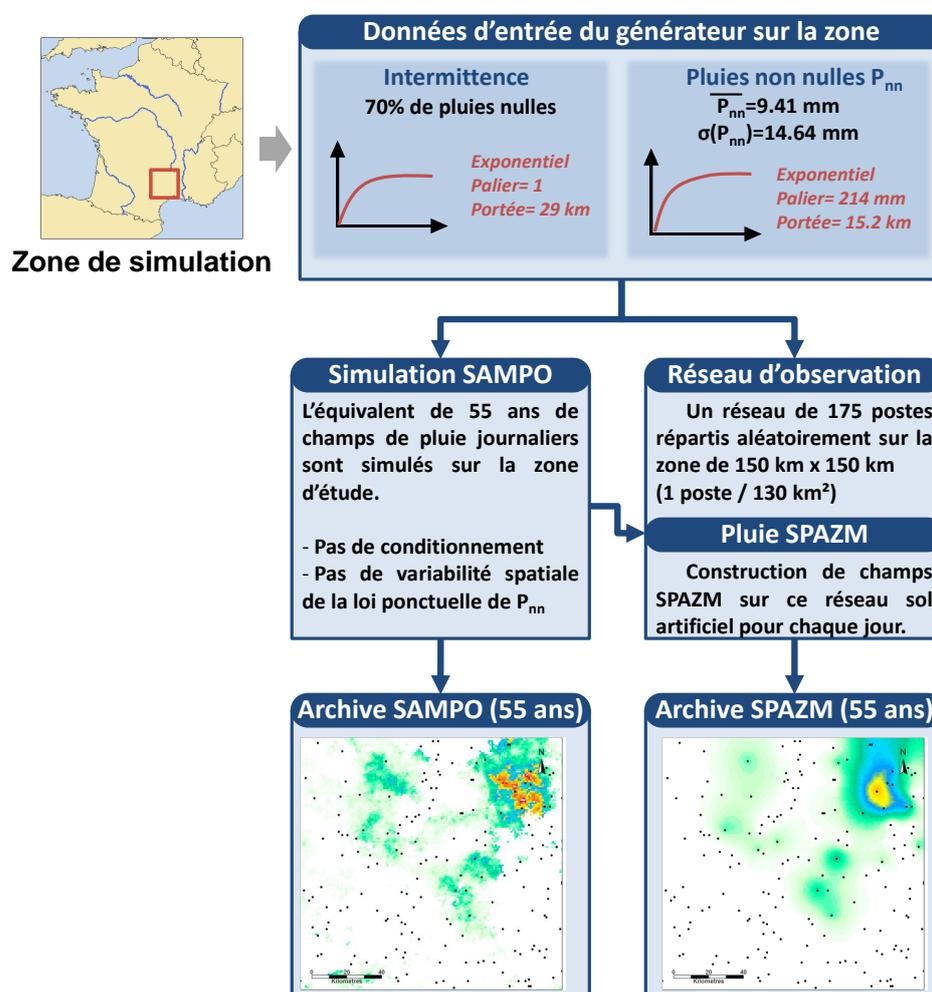


FIGURE IV.13 – Schéma de la création de champs de pluie SAMPO et SPAZM.

Le générateur SAMPO simule ensuite environ 20 000 champs stationnaires de pluies journalières en adéquation avec les caractéristiques statistiques et structurales imposées en entrée. Les champs sont dits stationnaires, car ils sont générés sans conditionnement et avec des paramètres de pluies uniformes sur toute la zone. Il n'a donc aucun effet orographique ou autre instationnarité spatiale. Chaque champ généré est indépendant des autres champs simulés ce qui implique que la pluie journalière obtenue en un point n'a aucune cohérence temporelle d'un

jour à l'autre, mais la cohérence fréquentielle, qui nous intéresse, est respectée. Finalement, une archive de champs de pluie journaliers sur 55 ans est constituée, avec chaque jour, un champ de pluie au km² sur la zone de simulation (figure IV.14 (a)).

Un réseau artificiel de 175 pluviomètres est tiré uniformément sur la zone de simulation. Le nombre de pluviomètres a été choisi de façon à représenter la densité moyenne d'instrumentation disponible dans la réalité qui est d'environ un poste tous les 130 km². Le réseau généré (points noirs sur la figure IV.14) est plus ou moins dense à des endroits de la zone de simulation. La pluie observée à chacun de ces postes est la série journalière de 55 ans extraite des simulations SAMPO. Nous disposons alors d'un réseau sol de 175 pluviomètres avec 55 ans de données. Par cet exercice de style, un réseau pluviométrique proche de celui à disposition en pratique a été construit, mais en connaissant les champs de pluie réels sous-jacents.

En s'appuyant sur le réseau artificiel de pluviomètres, des champs de pluies journaliers peuvent être estimés en utilisant la méthode SPAZM. Une archive de 55 ans de champs journaliers SPAZM au km² est construite et constitue une estimation des champs SAMPO par la méthode SPAZM. Comme les champs de pluie SAMPO sont stationnaires et qu'un unique type de temps a été considéré, SPAZM se résume à un interpolateur géostatistique proche d'un krigage journalier s'appuyant sur le réseau. Sur la figure IV.14, le champ (a) journalier SAMPO est juxtaposé au champ (b) estimé avec SPAZM sur le réseau sol (points noirs) pour le même jour. Le réseau de pluviomètres n'est qu'un échantillonnage du champ de pluie (i.e. du champ SAMPO). Il peut donc manquer une grande partie de l'évènement. De plus, le processus d'interpolation lisse la structure du champ. Les pluies SPAZM ont donc tendance à sous-estimer les pluies extrêmes localement (sauf au droit des postes). Cependant, la pluie spatiale est calculée par une moyenne des pluies sur tous les pixels du domaine. Ainsi, selon la surface considérée, l'impact peut être plus ou moins fort sur les pluies extrêmes.

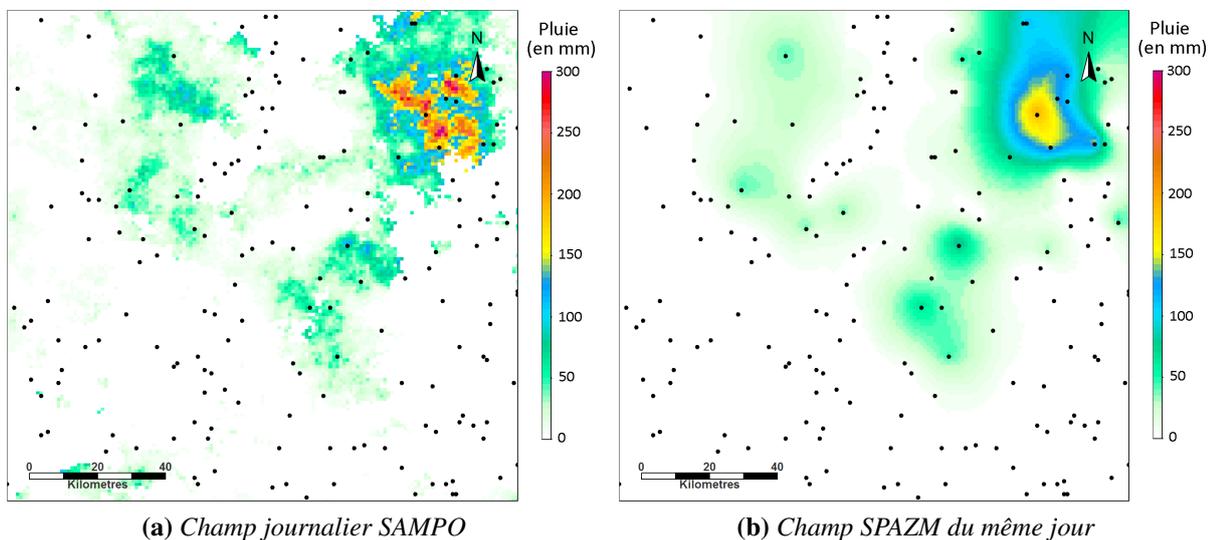


FIGURE IV.14 – Exemple de champ journalier obtenu avec SAMPO et SPAZM (les points noirs représentent le réseau sol).

Afin d'évaluer les pluies spatiales, nous allons nous intéresser à l'abattement sur les pluies extrêmes de temps de retour T (Galea *et al.*, 1982). L'abattement $\alpha(T)$ est une valeur variant entre 0 et 1 et qui décrit une diminution du risque de pluie (par rapport à l'agrégation des quantiles ponctuels) plus ou moins forte du fait de l'agrégation de la pluie sur une surface. L'abattement vaut 1 pour un bassin de 1 km² (i.e. assimilable à la valeur ponctuelle) ce qui signifie que la pluie spatiale et la pluie ponctuelle portent le même risque de pluie. L'abattement est d'autant plus fort que $\alpha(T)$ diminue.

$$\begin{aligned}\alpha(T) &= \frac{\text{Pluie spatiale } PS \text{ de temps de retour } T}{\text{Moyenne des pluies de temps de retour } T \text{ aux pixels } p \text{ du bassin}} \\ &= \frac{PS(T)}{\overline{(P_p(T))}_{p \in \text{Bassin}}}\end{aligned}\tag{IV.6}$$

Un abattement est alors défini sur la pluie spatiale SAMPO – qui représente l'abattement réel attendu – et sur la pluie spatiale SPAZM :

$$\begin{cases} \alpha_{SPAZM}(T) = \frac{PS_{SPAZM}(T)}{\overline{(P_p(T))}_{p \in \text{Bassin}}} \\ \alpha_{SAMPO}(T) = \frac{PS_{SAMPO}(T)}{\overline{(P_p(T))}_{p \in \text{Bassin}}}\end{cases}\tag{IV.7}$$

Des bassins de formes et de surfaces variables sont alors placés aléatoirement sur la zone. Pour chaque bassin, des pluies spatiales journalières SAMPO et SPAZM sont calculées en effectuant la moyenne journalière des valeurs de pluie des pixels du bassin sur leurs champs respectifs. Nous disposons ainsi d'une population de bassins variés avec pour chacun, la pluie spatiale SAMPO (pluie spatiale de référence) et la pluie spatiale SPAZM. Les abattements de chacune des pluies peuvent alors être déterminés avec les formules (IV.7).

La figure IV.15 montre l'évolution de l'abattement de SAMPO $\alpha_{SAMPO}(T)$ pour les différents temps de retour T en fonction de la surface des bassins versants⁷. Les pluies de temps de retour T sont estimées empiriquement. Avec 55 années de données, l'estimation de pluie décennale est suffisamment robuste. À noter que les résultats sur le temps de retour 20 ans sont plus incertains. Il est couramment admis qu'il faut environ une série de longueur $5T$ pour estimer de façon empirique la pluie de temps de retour T (Stifel et Mc Pherson, 1964). Une moyenne des pluies de temps de retour à chaque pixel du bassin conduit à surestimer le risque par rapport à l'estimation de la pluie de temps de retour sur la pluie spatiale. L'abattement tend à être très fort pour de grands bassins. Sa valeur semble par contre quasi-indépendante du temps de retour considéré.

7. Les bassins considérés ne dépassent pas une surface de 1000 km² car pour cette valeur, un comportement asymptotique est déjà observé sur la courbe d'abattement et pour des raisons de lourdeurs de calcul.

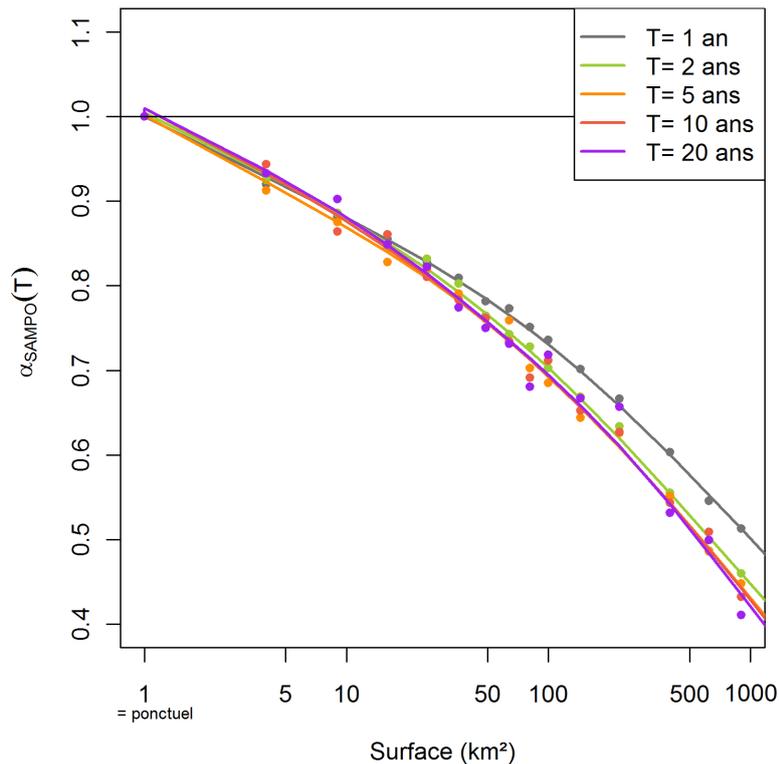


FIGURE IV.15 – Courbes des abattements de SAMPO pour les pluies de temps de retour T selon la surface du bassin.

La figure IV.16 permet de regarder le rapport entre l'abattement réalisé par SAMPO et celui réalisé par SPAZM pour le temps de retour 10 ans. Plus simplement, regarder le ratio $\frac{\alpha_{SAMPO}(T)}{\alpha_{SPAZM}(T)}$ revient à s'intéresser à $K_{P(T)} = \frac{PS_{SAMPO}(T)}{PS_{SPAZM}(T)}$. C'est donc un graphe équivalent au graphe IV.8 (page 128) qui compare les quantiles de pluie de SPAZM et de Thiessen. SAMPO et SPAZM sont globalement cohérents pour des surfaces de bassin supérieures à 700 km² (droite en pointillés) qui était également la limite de surface à partir de laquelle SPAZM et Thiessen fournissaient des résultats similaires. Pour des surfaces plus faibles, SPAZM sous-estime fortement les quantiles et la sous-estimation est très variable pour une surface de bassin donnée. Les pluies spatiales calculées avec SPAZM peuvent générer des abattements sur les pluies extrêmes 50% plus forts que ceux de SAMPO. La sous-estimation médiane sur une surface de 1 km² est de 50% alors que celle observée lors de l'étude des pluies ponctuelles extrêmes était plutôt de l'ordre de 10%. Cette forte différence peut s'expliquer par le fait que la méthode SPAZM sur les champs stationnaires SAMPO s'appuie sur une ébauche uniforme sur toute la zone de simulation perdant ainsi un part d'information apportée par l'ébauche SPAZM (données en altitude...). De plus, tout les champs sont générés avec des paramètres moyens de toute la population des pluies non nulles. Des simulations avec des caractéristiques des pluies par types de temps permettraient des populations de champs de pluies plus diversifiées. Cependant, le comportement observé sur les champs générés reste pertinent.

Lors de l'étude des estimations de pluies extrêmes ponctuelles avec SPAZM, il avait été montré que la sous-estimation des quantiles était en grande partie due à une mauvaise estimation de l'écart-type de la population des pluies. Nous avons alors corrigé les estimations sur les pluies extrêmes locales en nous appuyant sur les erreurs d'estimation des écarts-types des

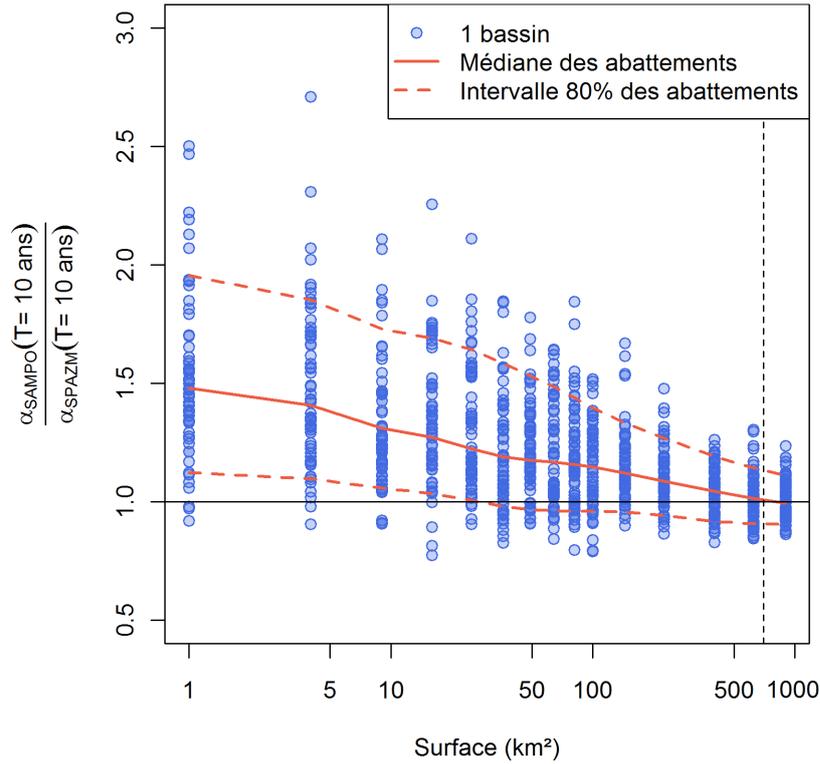


FIGURE IV.16 – Comparaison des abattements de SAMPO et SPAZM pour la pluie décennale selon la surface du bassin.

pluies. Il avait été suggéré que la densité du réseau et la variabilité du champ pouvaient entraîner un lissage plus ou moins fort sur les valeurs extrêmes de pluie avec SPAZM. Les champs stationnaires SAMPO vont permettre d'étudier l'impact de la dispersion des postes. L'objectif est notamment de réussir à estimer le ratio $K_P(T) = \frac{\alpha_{SAMPO}(T)}{\alpha_{SPAZM}(T)}$ de façon à pouvoir corriger les quantiles SPAZM à partir de la sous-estimation sur les écarts-types des pluies spatiales.

Par analogie avec les abattements sur les quantiles $\alpha(T)$, un abattement β sur les écarts-types des pluies peut alors être défini :

$$\begin{aligned} \beta &= \frac{\text{Ecart-type de la pluie spatiale PS}}{\text{Moyenne pondérée des écarts-types aux postes contribuant à la pluie spatiale SPAZM}} \\ &= \frac{\sigma(PS)}{\sum_{i=1}^N \omega_i \cdot \sigma(P_i)} \end{aligned} \quad (\text{IV.8})$$

Les poids $\omega = (\omega_i)_{i \in \{1, N\}}$ sont des poids représentant l'importance des postes comme source d'information pour la construction d'une pluie spatiale SPAZM sur le bassin. Pour chaque pixel (x, y) du bassin, les poids $\lambda_{i \in \{1, 10\}}(x, y)$ des 10 postes les plus proches⁸ du pixel (x, y) sont déterminés. Les poids $\lambda_i(x, y)$ sont calculés avec une fonction de pondération similaire à la

8. Les 10 postes les plus proches ne sont pas forcément les mêmes pour chaque pixel du bassin.

cloche de pondération SPAZM :

$$\left\{ \begin{array}{l} \lambda_i(x,y) = 20 \cdot e^{-\left(\frac{d_i}{d_0}\right)^2} \\ d_i, \text{ distance euclidienne entre le poste } i \text{ et le pixel } (x,y) \\ d_0 = \min(d_i) \\ \text{Les } \lambda_i \text{ sont normalisés de façon à avoir : } \sum_{i=1}^{10} \lambda_i(x,y) = 1 \end{array} \right. \quad (\text{IV.9})$$

La seule différence avec la cloche de pondération SPAZM est la distance d_0 utilisée. En effet, la cloche SPAZM utilise une distance $d_0 = \max(d_i)$. Cette différence a pour conséquence de donner un poids λ_i plus fort pour les postes les plus proches du pixel (x,y) que ce que fournit la cloche de pondération SPAZM. La raison de ce choix vient du fait que SPAZM utilise certes une cloche de pondération en $\lambda_i(x,y) = 20 \cdot e^{-\left(\frac{d_i}{d_0}\right)^2}$ avec $d_0 = \max(d_i)$ mais les champs de pluies subissent ensuite un krigeage des résidus. En fin de compte, la valeur estimée en un pixel est à la fois issue de la déformation de l'ébauche $\Lambda_{(x,y)}$ (calculée avec la cloche de pondération) et du krigeage des résidus $\varepsilon_{(x,y)}$ (fonction du variogramme utilisé).

$$PJ_{(x,y)}^* = \Lambda_{(x,y)} \cdot EB_{(x,y)} + \varepsilon_{(x,y)} \quad (\text{IV.10})$$

La valeur estimée est un mélange de deux « couches de pondérations ». Le krigeage des résidus fait porter un peu plus d'information aux postes les plus proches ce qui justifie le choix d'une pondération en $d_0 = \min(d_i)$ pour décrire la contribution de chaque poste pour la pluie spatiale SPAZM.

En chaque pixel (x,y) , les dix postes les plus proches ont un poids $\lambda_i(x,y)$ et les autres ont des poids nuls. La contribution à la pluie spatiale ω_i d'un poste i est la moyenne de ses contributions pour chaque pixel (x,y) du bassin.

$$\left\{ \begin{array}{l} \omega_i = \overline{(\lambda_i(x,y))}_{(x,y) \in \text{Bassin}} \\ \sum_{i=1}^N \omega_i = 1 \end{array} \right. \quad (\text{IV.11})$$

Pour le calcul de l'abattement β , le choix de pondération n'est pas primordial, puisque la valeur au dénominateur de β a pour but principal d'adimensionnaliser la valeur de β et de la rendre comparable d'un bassin à l'autre, mais elle trouvera son intérêt plus tard.

Pour chaque bassin, une valeur de β est calculée pour chaque bassin avec les pluies spatiales SAMPO et SPAZM.

$$\left\{ \begin{array}{l} \beta_{SPAZM} = \frac{\sigma(PS_{SPAZM})}{\sum_{i=1}^N \omega_i \cdot \sigma(P_i)} \\ \beta_{SAMPO} = \frac{\sigma(PS_{SAMPO})}{\sum_{i=1}^N \omega_i \cdot \sigma(P_i)} \end{array} \right. \quad (\text{IV.12})$$

Nous pouvons alors étudier les différences entre les abattements sur les écarts-types réalisés par SAMPO et par SPAZM en calculant le critère $\frac{\beta_{SAMPO}}{\beta_{SPAZM}}$.

Sur la figure IV.17, nous pouvons voir que la valeur $K_{P(T)}$ est très corrélée à $\frac{\beta_{SAMPO}}{\beta_{SPAZM}}$ et qu'un lien linéaire peut être envisagé :

$$K_{P(T)} - 1 = a(T) \cdot \left(\frac{\beta_{SAMPO}}{\beta_{SPAZM}} - 1 \right) \quad (\text{IV.13})$$

avec $a(T)$ qui augmente avec le temps de retour.

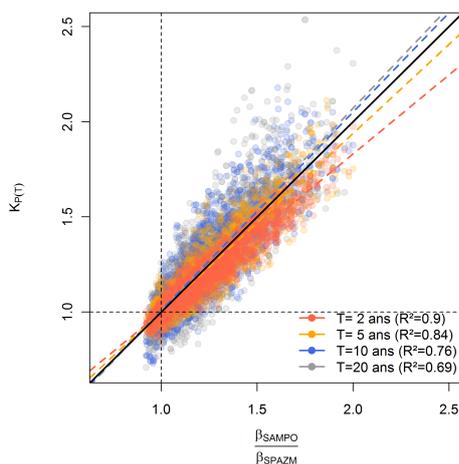
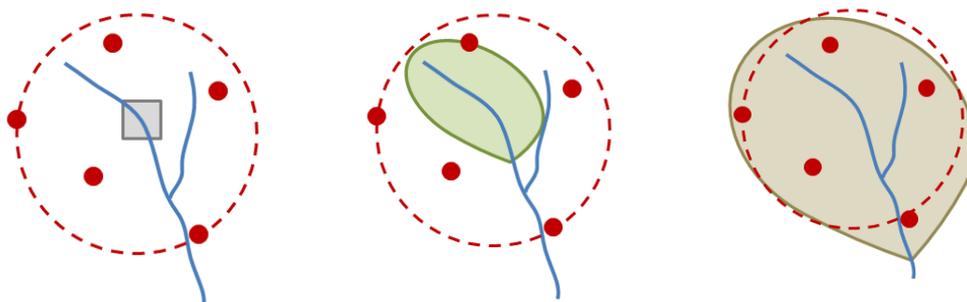


FIGURE IV.17 – Lien entre l'abattement sur les quantiles et sur les écarts-types.

T (ans)	$a(T)$
2	0.83
5	0.93
10	1.04
20	1.06

Tableau IV.3 – Coefficient $a(T)$ de la régression linéaire entre $K_{P(T)}$ et $\frac{\beta_{SAMPO}}{\beta_{SPAZM}}$.

Une piste pour corriger les quantiles de la pluie spatiale SPAZM est donc d'estimer $K_{P(T)}$ en s'appuyant sur les ratios $\frac{\beta_{SAMPO}}{\beta_{SPAZM}}$. L'abattement réalisé sur les écarts-types est fortement lié au processus de lissage, que ce soit de par le processus d'interpolation de SPAZM ou de par le principe même d'agrégation des pluies sur les pixels du bassin pour construire la pluie spatiale. Dans les deux cas, c'est le fait de moyenner de l'information de plusieurs sources qui explique cet abattement. **La différence entre l'abattement SAMPO et l'abattement SPAZM vient donc en grande partie du fait que SPAZM réalise une pluie spatiale à partir d'information plus dispersée autour du bassin.**



(a) Dispersion des postes SPAZM autour d'un pixel. (b) Dispersion des postes SPAZM autour d'un petit bassin. (c) Dispersion des postes SPAZM autour d'un grand bassin.

FIGURE IV.18 – Dispersion des N postes (points rouges) utilisées par SPAZM pour calculer la pluie spatiale avec une densité de postes constante et des bassins de surfaces variables.

La figure IV.18 schématise le problème de dispersion du panier SPAZM. La figure (a) montre un exemple de panier stations (points rouges) utilisé par SPAZM pour calculer la pluie à un pixel (carré gris). Le cercle rouge permet de visualiser la surface sur laquelle s'étalent les postes du panier SPAZM. La pluie au pixel SPAZM est alors une pluie calculée à partir d'information provenant d'une zone bien plus large. C'est ce phénomène qui explique le lissage des champs SPAZM. La pluie SPAZM à chaque pixel est en réalité une « pseudo-pluie spatiale » sur la zone rouge dont la taille varie selon la densité d'information à disposition. Le problème est le même avec des bassins de petite surface. Avec un grand bassin, la zone de sélection de SPAZM et le bassin tendent à avoir une surface équivalente. Les écarts entre les abattements SAMPO et SPAZM viennent alors du fait que plus le bassin est petit, plus la pluie SPAZM calcule une pluie spatiale sur une surface plus grande que celle du bassin. L'abattement avec SPAZM est donc plus important qu'avec SAMPO.

Pour la suite, deux surfaces descriptives de cet étalement de l'information vont être définies (formules IV.14). La surface S_{BV} correspond à la surface du bassin, c'est-à-dire à la surface réelle sur laquelle la pluie spatiale doit être calculée. Nous définissons également une aire d'effet pour SPAZM (ou AOE – Area Of Effect) qui correspond à la surface sur laquelle SPAZM s'appuie pour estimer la pluie spatiale du bassin. La distance D_{SPAZM} est la moyenne pondérée des distances des postes au centre du bassin. La pondération utilisée est celle définie par la formule IV.9 introduite précédemment. Nous considérons simplement que SPAZM fournit une pluie spatiale sur un disque de rayon D_{SPAZM} .

$$\left\{ \begin{array}{l} D_{SPAZM} = \sum_{i=1}^N \omega_i \cdot \left\| \begin{pmatrix} x_i \\ y_i \end{pmatrix}, \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} \right\| \\ \text{avec } (x_i, y_i) \text{ les coordonnées du poste } i \text{ de poids } \omega_i, \\ (x_0, y_0) \text{ les coordonnées du centroïde du bassin de surface } S_{BV}, \\ \text{et la norme } \|\cdot\| \text{ la distance euclidienne selon } (x, y). \\ AOE = \pi \cdot D_{SPAZM}^2 \end{array} \right. \quad (IV.14)$$

Les deux aires AOE définissent une métrique pour évaluer les abattements des pluies spatiales. En effet, plutôt que de tracer des courbes d'abattement en fonction de la surface, elles sont tracées en fonction des AOE définies pour voir si les abattements SPAZM et SAMPO sont comparables une fois considérés sur des surfaces équivalentes. La figure IV.19 représente en vert l'abattement β_{SAMPO} en fonction de l' AOE SAMPO (i.e. la surface réelle du bassin) et en bleu l'abattement β_{SPAZM} en fonction de l' AOE SPAZM.

L'abattement β évoluent linéairement avec le logarithme de l' AOE . La courbe d'abattement obtenues est similaire à celle obtenue en figure IV.16 mais cette fois les abattements SPAZM et SAMPO sont comparables. Cependant, pour les faibles valeurs d' AOE de SPAZM, les abattements sur les écarts-types β_{SPAZM} sont inférieurs à 1. Ceci peut être dû au choix réalisé pour les pondérations ω_i qui impacte directement le calcul de D_{SPAZM} donc de l' AOE .

Nous introduisons deux constantes C_{SPAZM} et C_{SAMPO} :

$$\begin{aligned} C_{SPAZM} &= \frac{1 - \beta_{SPAZM}}{\ln(AOE_{SPAZM})} \\ C_{SAMPO} &= \frac{1 - \beta_{SAMPO}}{\ln(AOE_{SAMPO})} = \frac{1 - \beta_{SAMPO}}{\ln(S_{BV})} \end{aligned} \quad (IV.15)$$

Comme :

$$C_{SPAZM} \approx C_{SAMPO}$$

On peut finalement écrire :

$$\frac{1 - \beta_{SAMPO}}{\ln(S_{BV})} \approx \frac{1 - \beta_{SPAZM}}{\ln(AOE_{SPAZM})} \quad (IV.16)$$

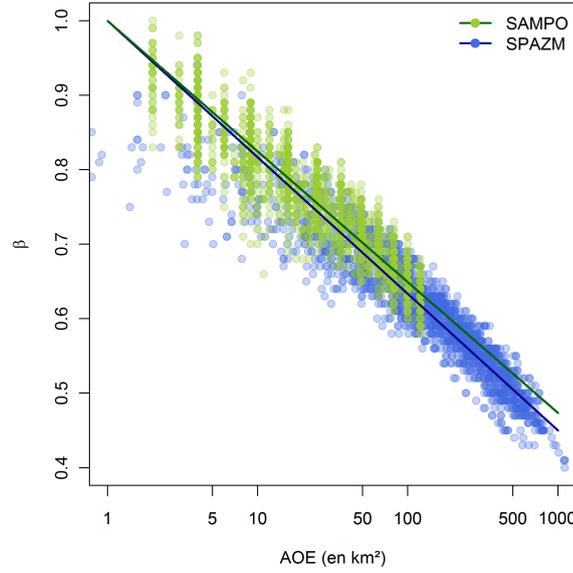


FIGURE IV.19 – Lien entre l'abattement sur les écarts-types β et l'aire d'effet AOE .

La relation (IV.16) permet d'exprimer l'abattement sur les écarts-types en fonction du ratio des AOE .

$$\beta_{SAMPO}^* = 1 - \frac{\ln(S_{BV})}{\ln(AOE_{SPAZM})} \cdot (1 - \beta_{SPAZM}) \quad (IV.17)$$

Lorsque SPAZM et le bassin ont des AOE équivalentes (i.e. SPAZM utilise un panier représentatif du bassin) alors les abattements β sur les écarts-types sont similaires. Si l'aire d'effet de SPAZM est plus petite que l'aire du bassin versant alors SPAZM possède une information concentrée à l'intérieur du bassin. L'abattement β_{SPAZM} est alors égal à l'abattement β_{SAMPO} . Dit autrement, si SPAZM possède suffisamment d'information dans le bassin, il estime correctement l'abattement de la pluie spatiale.

$$\begin{cases} \text{Si } AOE_{SPAZM} > S_{BV} & , \text{ alors } \beta_{SPAZM} < \beta_{SAMPO} \\ \text{Si } AOE_{SPAZM} \simeq S_{BV} & , \text{ alors } \beta_{SPAZM} \simeq \beta_{SAMPO} \\ \text{Si } AOE_{SPAZM} < S_{BV} & , \text{ alors } \beta_{SPAZM} = \beta_{SAMPO} \text{ (Imposé)} \end{cases} \quad (IV.18)$$

Le ratio $\frac{\beta_{SAMPO}}{\beta_{SPAZM}}$ peut alors être estimé uniquement en fonction de l' AOE de SPAZM et de la surface du bassin et de la valeur de l'abattement SPAZM β_{SPAZM} qui sont tous trois calculables dans la pratique à partir des données connues (équation (IV.17) et figure IV.20). Les ratios ainsi estimés ont un coefficient de détermination à 0.92 avec les ratios calculés via SAMPO (figure IV.20 (a)).

Une fois le ratio $\frac{\beta_{SAMPO}}{\beta_{SPAZM}}$ connu, nous pouvons estimer une valeur de $K_{P(T)}$ et ainsi corriger SPAZM pour revenir aux quantiles estimés par SAMPO (figure IV.20 (b)). Cette correction à partir des dispersions des postes, du bassin et de l'abattement sur l'écart-type β_{SPAZM} de SPAZM est utilisable sur des cas réels, sans champs observés à disposition. Elle permet de corriger l'abattement réalisé par SPAZM, souvent trop fort du fait de la dispersion du panier de postes utilisé. Le seul bémol de ce travail est qu'il ne corrige SPAZM que du point de vue de la dispersion des postes, mais pas de la variabilité du champ de pluie à risque. En effet, avec une dispersion donnée des postes, un bassin serait mieux représenté si les champs de pluies sont peu variables. Une solution pour prendre compte cet effet serait d'ajouter une composante liée à la valeur du champ de pluie de temps de retour T à la distance euclidienne utilisée pour le calcul de

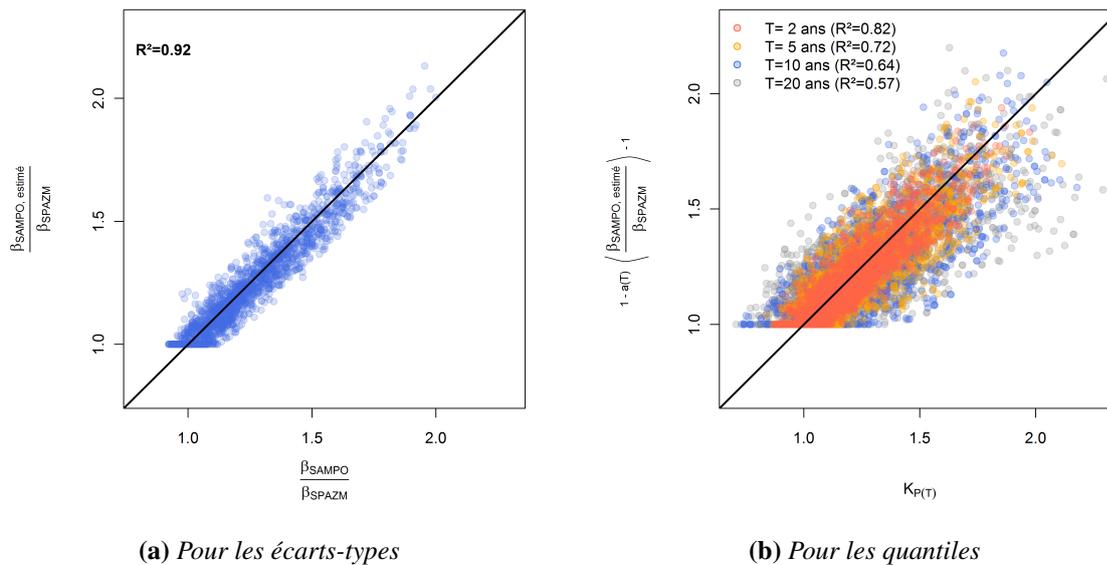


FIGURE IV.20 – Estimation du ratio d'abattement SPAZM et SAMPO.

la dispersion. Les cartes de pluies décennales ou centennales trouvées dans la partie précédente pourraient être utilisées afin de calculer une distance selon les variables $(x, y, P(T))$. Avec une disposition de postes identique vis-à-vis du bassin, des dispersions plus ou moins fortes seraient trouvées selon que les postes représentent bien ou pas les quantiles élevés aux pixels du bassin.

Les champs stationnaires SAMPO utilisés pour ce travail ne tiennent pas compte du relief. Nous ne pouvons donc pas mettre en place cette prise en compte de la variabilité des champs de pluies extrêmes. Pour pouvoir le faire, il faudrait refaire des simulations SAMPO tenant compte de la variabilité des champs de pluie au sein de la zone de simulation. Les champs pourraient être conditionnés par des valeurs observées pour forcer le simulateur à représenter une certaine structure du champ de pluie. Cependant, nous avons mis en évidence quelques limitations concernant le conditionnement de SAMPO qui ne respecte pas la loi de pluies non nulles donnée en entrée ou la distribution des pluies de conditionnement. Une autre approche serait de simuler des champs libres, mais au lieu de fournir une loi de pluies non nulles globale sur toute la zone d'étude, elle pourrait être fixée à chaque pixel en utilisant une carte de moyenne et d'écart-type sur la zone de simulation. Des champs libres avec des instationnarités pourraient être générés. Cependant, l'abattement sur des champs non stationnaires a de fortes chances d'être similaire, à une pondération près, à ceux observés sur les champs stationnaires. Le calcul de l'AOE, même sur des cas réels, est donc un moyen de vérifier qu'une pluie SPAZM dispose de suffisamment d'information pour réaliser une pluie spatiale cohérente.

L'organigramme IV.21 résume la procédure préconisée pour estimer des pluies spatiales extrêmes avec la méthode SPAZM.

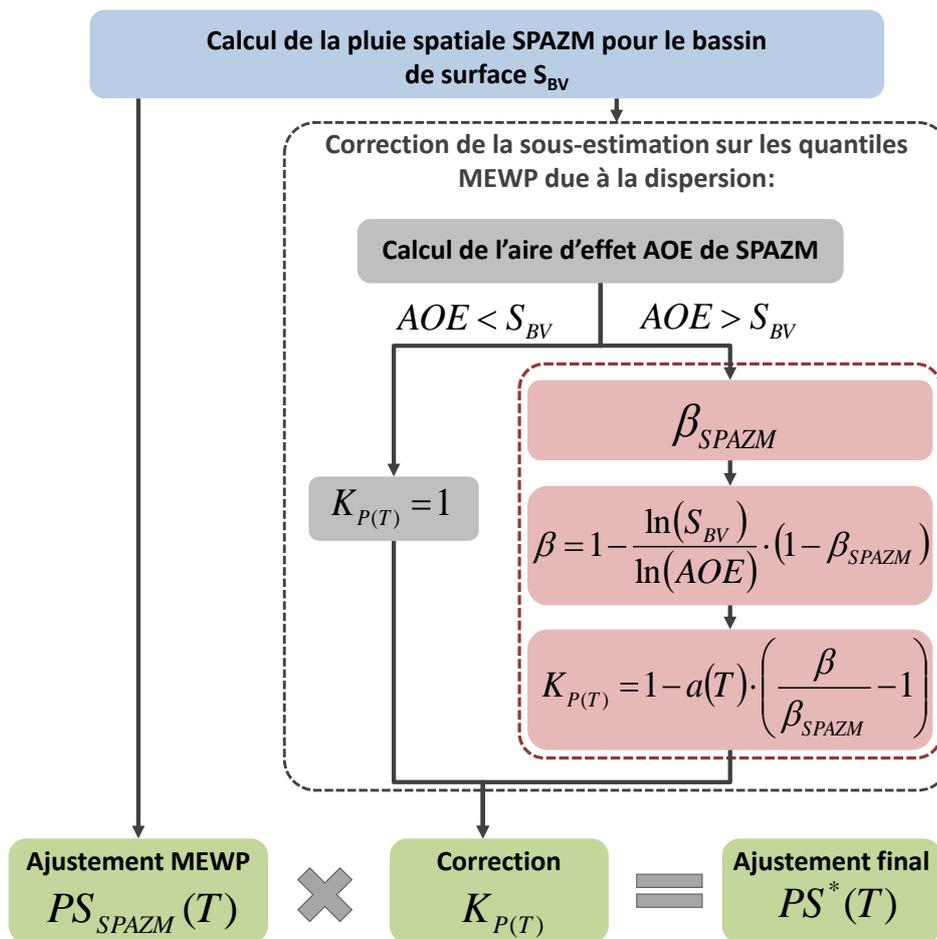


FIGURE IV.21 – Organigramme de synthèse de la procédure préconisée pour l'estimation des forts quantiles d'une pluie spatiale SPAZM.

IV. 4 Conclusion

L'objectif de cette partie était d'étudier la construction de pluie spatiale pour les bassins et d'évaluer la sensibilité des choix réalisés sur l'estimation des pluies extrêmes journalières des bassins considérés. Une comparaison des pluies spatiales construites avec les polygones de Thiessen et avec les pluies SPAZM sur une large base de bassins répartis sur toute la zone de SPAZM et avec une grande variété de surface a permis d'étudier les différences entre les deux approches.

Il est apparu que les pluies spatiales SPAZM et Thiessen sont cohérentes entre elles pour des bassins de surface supérieure à 700 km². **Dit autrement, pour des bassins de surface suffisamment grande, la densité du réseau sol est suffisante pour permettre une estimation de la pluie spatiale pertinente, quelle que soit la méthode.** Une correction du cumul annuel moyen de la pluie Thiessen pour le ramener à la valeur de référence donnée par la pluie SPAZM est, si elle existe, justifiée. Dans plus de 80% des cas concernant des bassins suffisamment grands, SPAZM et Thiessen corrigé fournissent alors des estimations de pluies extrêmes cohérentes.

Pour des bassins versants de surface inférieure à 700 km², conclure sur la méthode la plus performante pour représenter la pluie de bassin la plus proche de la réalité est moins trivial. En effet, SPAZM et Thiessen ont tendance à fournir des précipitations annuelles moyennes et des estimations de pluies extrêmes très variables. Thiessen a tendance à sous-estimer les cumuls de précipitation par rapport à SPAZM en particulier pour des bassins de montagnes pour lesquels il est plus difficile de trouver des postes en altitude représentatifs de la pluviométrie du bassin. L'extrapolation avec le relief de la méthode SPAZM est alors une référence plus fiable pour réaliser une correction en volume sur la pluie spatiale Thiessen. Après correction des volumes entre SPAZM et Thiessen, les pluies spatiales SPAZM estiment souvent des quantiles élevés de pluie plus faibles que ceux des pluies Thiessen. Cependant, il est difficile de savoir laquelle des deux approches est la plus proche de la pluie de bassin attendue. La méthode SPAZM, même si elle conduit, de par son processus d'interpolation, à sous-estimer les pluies extrêmes sur les petits bassins, présente l'avantage par rapport aux polygones de Thiessen de mieux prendre en compte les zones de relief, en général sous-instrumentées. **De plus, les pluies Thiessen sont très sensibles aux choix des postes. La correction en volumes sur Thiessen permet de rattraper la sous-estimation des pluies annuelles moyennes mais fournit des valeurs de pluies extrêmes souvent incohérentes. De ce point de vue, les estimations SPAZM sont de meilleure qualité.**

Une évaluation de l'abattement des pluies SPAZM, en utilisant le générateur de pluie SAMPO, a mis en évidence que la dispersion des postes utilisés pour construire les champs de pluies SPAZM est un facteur explicatif important de la sous-estimation des pluies extrêmes avec SPAZM. L'indice *AOE* (Area of Effect) permet d'évaluer une surface équivalente pour la pluie spatiale SPAZM qui peut être comparée celle du bassin. Un lien fort a été mis en évidence entre le ratio de l'aire *AOE* sur la surface du bassin et la sous-estimation des quantiles de pluies. Cet indice permet d'évaluer la confiance à accorder à la pluie SPAZM et de proposer une correction. Une évaluation sur des bassins réels constitue une étape nécessaire pour effectuer une validation de cette approche de correction sur les quantiles. Cependant la validation restera difficile à mettre en place sans pluie spatiale « réelle » pour servir de référence.

Deuxième partie

Les débits extrêmes

Le passage pluie-débit

À ce stade, nous avons abordé les différents problèmes entourant les pluies journalières extrêmes. Nous avons en particulier étudié la construction de pluie de bassin dans l'optique d'étudier les précipitations extrêmes. Une des raisons de ce travail était de permettre une estimation des pluies extrêmes en site non jaugé. En effet, l'aléa pluie est une variable clé pour l'explication de l'aléa débit. Cependant, le passage de la distribution des pluies à celle des débits n'est pas un problème simple. Avec les hypothèses de la méthode GRADEX, le passage pluie-débit aboutissait à un parallélisme entre les deux distributions. Cependant, une telle réponse du bassin a été remise en cause par plusieurs travaux. Un moyen de pallier à ce défaut est de prendre en compte l'aléa du bassin, c'est-à-dire la façon dont la réponse du bassin est influencée par son état (saturation, débit de base, neige...). C'est ce croisement d'aléa que réalise la méthode SCHADEX. Dans cette partie, nous allons nous intéresser à la façon de passer d'une distribution de pluies journalières à une distribution de débits – d'abord journaliers puis de pointe – pour un bassin non jaugé.

Une méthode de simulation stochastique reprenant les concepts SCHADEX, mais adaptée au contexte non jaugé, a été mise en place et va être présentée. Après avoir détaillé le principe de la méthode envisagée, une première tentative de régionalisation de ses paramètres va être présentée en s'appuyant sur une base d'études SCHADEX. Suite aux difficultés rencontrées pour régionaliser les paramètres sur la seule base des études SCHADEX disponibles, une automatisation de la méthode SCHADEX va être abordée afin de constituer une large base d'études SCHADEX pour la régionalisation. Pour terminer, nous nous intéresserons à l'étape finale, permettant de passer d'une distribution de débits moyens journaliers à une distribution de débit de pointe avec laquelle nous disposerons finalement d'une estimation du risque de crue sur le bassin.



SOMMAIRE DU CHAPITRE

V. 1 La base d'études SCHADEX	157
V. 2 Vers une méthode de simulation stochastique pluie-débit peu paramétrée : la méthode EREBOR	161
V. 2. 1 Présentation du modèle SCS	162
V. 2. 2 Le principe de la simulation avec le SCS	166
V. 3 Régionalisation des paramètres de la méthode EREBOR	173
V. 3. 1 La régionalisation des paramètres	173
V. 3. 2 Validation croisée de la simulation EREBOR	185
V. 4 Automatisation de la simulation SCHADEX	188
V. 4. 1 Sélection des saisons pour les précipitations	188
V. 4. 2 Calage automatique du modèle hydrologique MORDOR	193
V. 4. 3 Évaluation de la simulation SCHADEX automatique	200
V. 5 Passage à la distribution des débits de pointes	202
V. 5. 1 La méthode classique : le coefficient de forme moyen	202
V. 5. 2 La génération d'hydrogrammes pour des scénarios en volume	203
V. 6 Conclusion	210

V.1 La base d'études SCHADEX

La méthode SCHADEX est utilisée à EDF-DTG depuis 2006 pour réaliser les études de crues extrêmes. Près de 80 études ont été menées à ce jour sur différents bassins, en France ou à l'étranger (Norvège (Lawrence *et al.*, 2013), Espagne, Autriche, Colombie Britannique (Brigode, 2013)). Afin de régionaliser la méthode SCHADEX, un travail de collecte des différentes études a été réalisé de façon à construire une base de référence pour l'étude du passage pluie-débit. Seules les études réalisées au pas de temps journalier ont été conservées (tableau V.1). Pour chaque étude, les différents résultats obtenus lors de leur réalisation ont été récupérés :

- Les données pluies, températures et débits journaliers ayant fait l'objet d'une critique ;
- Le choix des saisons pour le risque pluie ainsi que les modèles probabilistes de pluie (loi MEWP, pluie adjacente) obtenus ;
- Le modèle MORDOR calé pour cette étude ;
- La simulation SCHADEX réalisée à partir de ces différents modèles et données et en particulier la distribution des écoulements journaliers ;
- Le coefficient de forme K_C moyen déterminé pour le passage à la pointe.

Tous les résultats d'études ont été expertisés à chaque étape ce qui permet de bénéficier d'une référence fiable pour la distribution des débits, mais également pour le modèle MORDOR et sa modélisation de la réponse hydrologique du bassin. Il est important de noter qu'élargir la base d'étude serait assez difficile du fait de cette part importante d'expertise apportée aux études SCHADEX. En effet, chaque étape de la méthode est validée de façon à représenter au mieux l'idée que l'on se fait de la réalité sur le bassin. Le modèle MORDOR est calé avec des critères objectifs évaluant le modèle du point de vue de sa capacité à simuler les crues, mais l'expertise reste prépondérante pour le choix du modèle hydrologique. Élargir la base d'étude demanderait d'être capable d'approcher cette expertise par des procédures automatiques pour obtenir des quantiles de débits comparables à ceux obtenus avec la méthode SCHADEX.

Les études ont été classées selon cinq classes de régime (figure V.2 (a)). Ce choix de cinq classes est un compromis entre la description de tendances très différentes dans le comportement des débits mensuels au cours de l'année et le nombre d'individus par classe. Les classes de régimes hydrologiques ont été construites par une ACP (Analyse des Composantes Principales) sur les douze débits moyens mensuels (figure V.1). Cette classification a été comparée à celle de Sauquet *et al.* (2008) qui permet de retrouver des classes équivalentes.

Régime pluvial : Ce régime hydrologique est caractérisé par une alternance d'une période de hautes eaux de novembre à avril et d'une période de basses eaux durant l'été. Ce type de régime est retrouvé sur des bassins versants où les précipitations liquides (i.e. la pluie et pas la neige) sont responsables des écoulements observés sur le bassin (figure V.1 (a)).

Régime cévenol : Ce type de régime est un cas particulier du régime pluvial propre à la France. Il concerne la région des Cévennes – à laquelle il doit son nom – et le piémont cévenol. Le contraste est encore plus marqué que pour le régime pluvial avec de très forts débits au début de l'automne (figure V.1 (b)).

Régime nival : Le régime nival est un modèle de régime hydrologique simple, caractérisé par une seule alternance annuelle de hautes et de basses eaux. Il se retrouve dans les bassins versants principalement alimentés par des précipitations sous forme de neige, c'est-à-dire dans les zones montagneuses dans le cas de la France. Le pic de hautes eaux est principalement dû à l'onde de fonte de la neige qui commence au début du printemps et atteint un maximum en mai-juin (figure V.1 (d)).

Régime nivo-pluvial : Le régime nivo-pluvial est un régime hydrologique mixte partageant des traits du régime nival et du régime pluvial. Il se caractérise par deux pics de débit bien marqués : un premier au printemps qui est le plus prononcé et qui est lié à la fonte des neiges, puis un second en automne, lié aux précipitations (figure V.1 (c)).

Régime nivo-glaciaire : Ce régime est assez similaire au régime nival, mais il est caractérisé par un maximum plutôt localisé entre les mois de juin et juillet. Le pic arrive plus tard que pour un bassin nival du fait de la fonte des glaces qui suit la fonte des neiges (figure V.1 (e)).

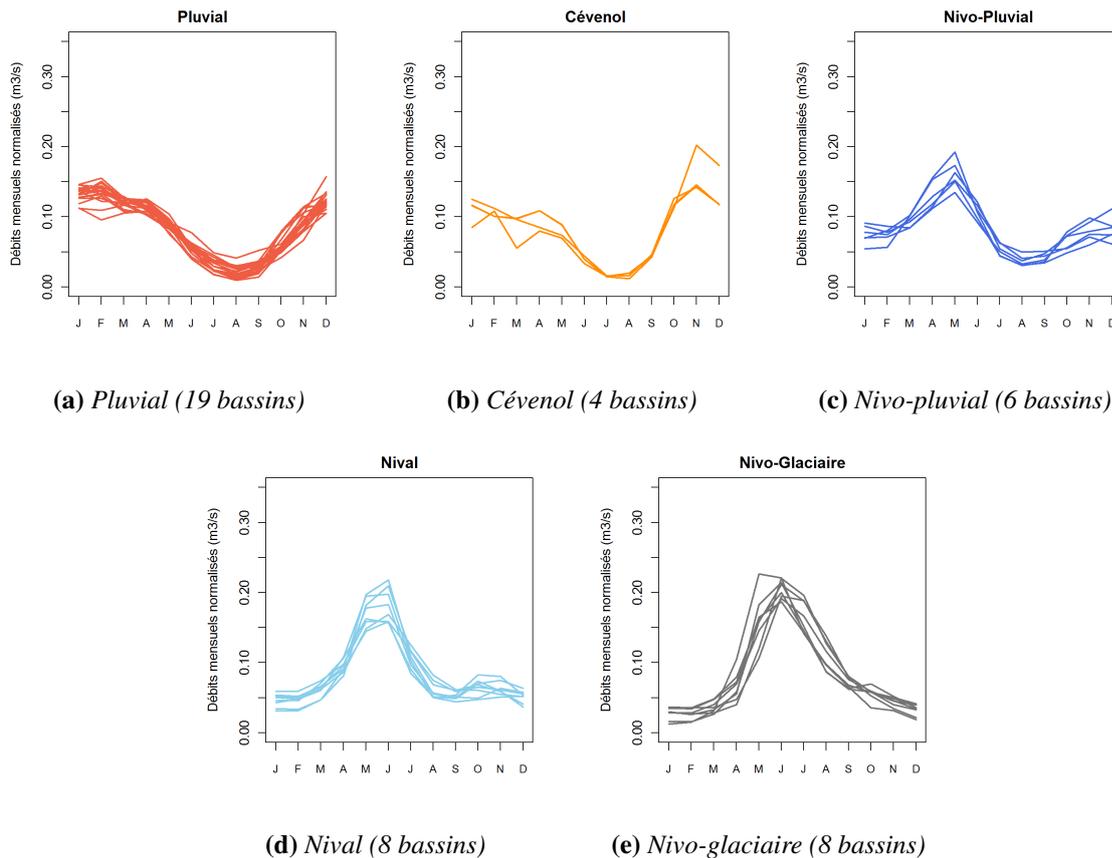
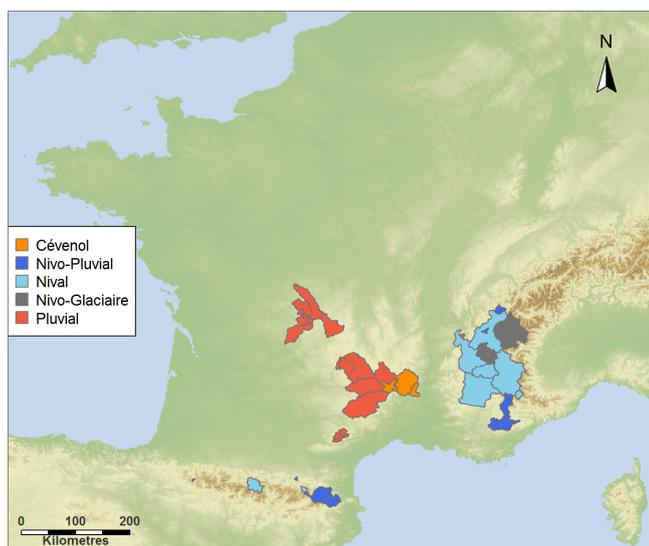
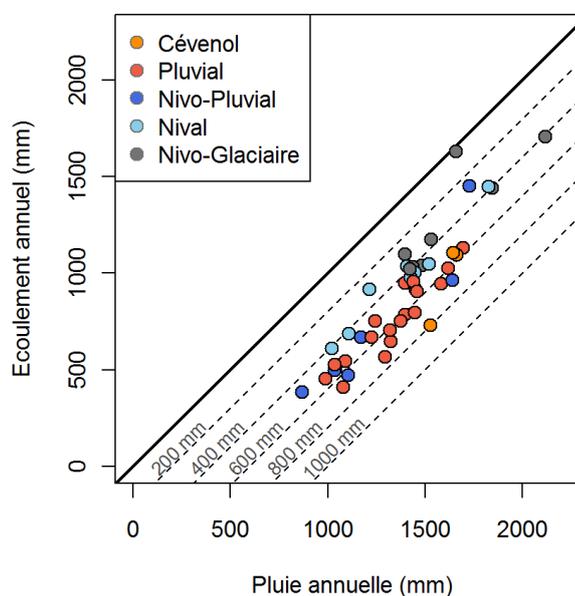


FIGURE V.1 – Les débits mensuels normalisés des bassins de la base et les cinq classes de régimes hydrologiques.

Sur la figure V.1 et la carte V.2 (a), sont représentés les différents bassins selon leur régime. La classe de bassins cévennols ne comporte que quatre individus ce qui peut poser un problème de robustesse pour disposer de conclusions sur cette classe.



(a) Régimes hydrologiques



(b) Bilans hydrologiques

FIGURE V.2 – Bassins versants de la base.

La base d'étude compte donc finalement 45 études SCHADEX. Cette base a fait l'objet d'une expertise humaine et d'une critique approfondie ce qui garantit une certaine fiabilité des études. Cependant, le choix de travailler sur ce type d'études réduit considérablement le nombre de bassins à disposition pour le travail de régionalisation. De plus, l'expertise humaine peut elle même être à l'origine d'inerties (par exemple, du fait de l'auto-apprentissage et de l'évolution des pratiques entre les études les plus anciennes et les plus récentes).

À partir de cette base, nous allons essayer de trouver des processus pouvant conditionner la réponse hydrologique d'un bassin à la pluie qu'il reçoit. L'objectif final est ici de réussir à estimer la distribution des écoulements journaliers à partir de celle des pluies. Les distributions de débits des études SCHADEX serviront de distributions de référence.

ID	Rivière	Exutoire	Superficie (en km ²)	Régime hydrologique	Pourcentage de la surface en glacier
1	Agout	Fraïsse	45	Pluvial	0
2	Agout	La Raviège	371	Pluvial	0
3	Allier	Poutes	1015	Pluvial	0
4	Arc	Saussaz	947	Nivo-Glaciaire	16
5	Ardèche	Sauze	2267	Cévenol	0
6	Ardèche	Vallon	1966	Cévenol	0
7	Arly	Moulin Ravier	222	Nivo-Pluvial	0
8	Arn	Taillade	82	Pluvial	0
9	Bromme	Brommat	106	Pluvial	0
10	Chassezac	Malarce	484	Cévenol	0
11	Chassezac	S ^{te} Marguerite	414	Cévenol	0
12	Creuse	Fresseline	1238	Pluvial	0
13	Dordogne	Bort-Les-Orgues	1012	Pluvial	0
14	Drac	Sautet	990	Nival	2
15	Durance	Escale	6764	Nival	1
16	Durance	Serre Ponçon	3582	Nival	2
17	Eau d'Olle	Grand Maison	48	Nivo-Glaciaire	6
18	Eau d'Olle	La Pernière	169	Nivo-Glaciaire	4
19	Ferrand	Mizoen	81	Nivo-Glaciaire	10
20	Garonne	Plan d'Arem	577	Nival	0
21	Gelon	Rochette	64	Nivo-Pluvial	0
22	Isère	Cevin	1777	Nivo-Glaciaire	8
23	Isère	Grenoble	5704	Nival	5
24	Isère	S ^t Gervais	9983	Nival	5
25	Isère	Tignes	170	Nivo-Glaciaire	15
26	Lot	Castelnau	1628	Pluvial	0
27	Maulde	Peyrat Le Château	174	Pluvial	0
28	Maulde	Vassivière	76	Pluvial	0
29	Morge	Voiron	46	Pluvial	0
30	Oriège	Campauleil	89	Nival	0
31	Romanche	Chambon Amont	228	Nivo-Glaciaire	20
32	Romanche	Champeau	998	Nivo-Glaciaire	16
33	Tarn	Pinet	2624	Pluvial	0
34	Taurion	Roche Talamie	654	Pluvial	0
35	Tech	Reynes	477	Nivo-Pluvial	0
36	Tet	Vinca	945	Nivo-Pluvial	0
37	Tinee	Pont de Belloire	170	Nival	0
38	Touyre	Montferrier	29	Nivo-Pluvial	0
39	Truyère	Grandval	1790	Pluvial	0
40	Truyère	Sarrans	2388	Pluvial	0
41	Vèbre	Laouzas	150	Pluvial	0
42	Verdon	S ^t Croix	1625	Nivo-Pluvial	0
43	Vezière	Saillant	964	Pluvial	0
44	Vienne	Bussy	380	Pluvial	0
45	Vienne	Peyrelevade	61	Pluvial	0

Tableau V.1 – Les 45 études SCHADEx au pas de temps journalier de la base.

V. 2 Vers une méthode de simulation stochastique pluie-débit peu paramétrée : la méthode EREBOR

Le but de cette partie est d'expliquer la mise en place d'une méthode simplifiée pour réaliser le passage de la distribution de pluie à la distribution des écoulements. Dans la méthode SCHADEX, le couplage de l'aléa pluie et de l'aléa du bassin est réalisé via le modèle hydrologique MORDOR. La finalité attendue est une méthode reprenant les grands concepts de la simulation SCHADEX, mais en ayant le moins de paramètres possible afin de rendre sa régionalisation plus aisée.

En supposant que la distribution des pluies est connue, le point clé du passage pluie-débit est la régionalisation du modèle hydrologique MORDOR. Ce modèle s'appuie sur une formulation qui a fait ses preuves pour la prévision et l'analyse hydrologique. Cependant, le modèle complet comporte 20 paramètres calés. De plus, des jeux de paramètres différents sont parfois équifinaux et deux bassins similaires n'ont pas nécessairement des paramètres proches.

Un test intéressant est de considérer un bassin versant et de caler le modèle sur des périodes différentes afin de voir si les jeux de paramètres obtenus sont cohérents entre eux. S'ils sont trop différents cela veut dire que, du point de vue des paramètres, le bassin n'est pas similaire à lui même sur des périodes différentes et qu'ils ne sont donc pas consistants pour décrire régionalement le comportement du bassin. Durant la thèse de Kuentz (2013), un calage du modèle MORDOR sur trois périodes distinctes a été réalisé sur la Durance à la Clapière (tableau V.2). Nous constatons, entre autres, des différences assez marquées sur les paramètres représentant les tailles maximales des réservoirs MORDOR (U_{max} , L_{max} , Z_{max}) et sur les paramètres régissant la vidange du stock intermédiaire L (k_{l1} et k_{l2}). Ces paramètres impactent fortement la réponse du modèle aux fortes pluies donc la façon dont il représente les crues (annexe A). Dans une optique de régionalisation, utiliser un modèle trop paramétré peut donc être problématique. Deux possibilités ont été envisagées. Le modèle MORDOR peut être conservé en essayant de fixer a priori certains de ses paramètres, mais contraindre a priori des paramètres n'est pas évident (Laboureau, 2013). Une seconde option est de substituer un autre modèle, plus simple, à MORDOR.

Après un travail exploratoire, le choix de remplacer MORDOR par un modèle plus simple a été privilégié. L'idée est alors de remplacer le modèle MORDOR par le modèle SCS (Soil Conservation Service), qui est très peu paramétré, afin d'obtenir des paramètres moins nombreux et plus facilement régionalisables pour un bassin versant.

Paramètres	1960-2009	1980-1994	1994-2009
fe1	0.03	0.04	0.03
fe3	-4.90	-4.90	-4.90
kl1	0.17	0.43	0.11
kl2	0.33	0.93	0.33
dn	800.00	800.00	800.00
exn	9.00	9.00	9.00
ftr1	0.55	0.70	0.55
ftr2	4.45	5.85	3.70
Umax	25.00	200.00	200.00
Lmax	185.00	100.00	40.00
Zmax	200.00	15.00	35.00
ltm	0.41	0.15	0.59
prft	99.00	99.00	99.00
efp	1.98	2.00	1.90
eft	-0.29	-0.99	0.16
dbt	0.00	0.00	0.00
kf	2.98	3.49	2.26
kfs	0.20	0.20	0.20
es	0.50	0.50	0.50
ztn	0.00	0.00	0.00

Tableau V.2 – Sensibilité des paramètres MORDOR à la période de calage sur la Durance à La Clapière.

V. 2. 1 Présentation du modèle SCS

Les origines de la méthode du SCS

La méthode du « Soil Conservation Service - Curve Number » (SCS-CN) de l'USDA¹ trouve son origine dans la modélisation pluie-débit via un hydrogramme unitaire (Beven, 2003). Elle se décompose en deux temps. Une première partie, appelée fonction de production, qui vise à prédire la quantité de pluie qui va réellement contribuer aux écoulements (i.e. la pluie efficace). Puis une seconde partie, correspondant à la fonction de transfert, qui s'appuie sur une approche par hydrogramme unitaire et qui permet de déterminer l'hydrogramme résultant de cette pluie efficace. La méthode du SCS-CN (ou plus simplement SCS) s'est appuyée sur des analyses empiriques, menées par l'USDA, sur les écoulements de petits bassins. Mockus (1949) a proposé de représenter ces données par une formule du type :

$$\frac{Q}{P - I_a} = 1 - 10^{-b(P - I_a)} \quad (\text{V.1})$$

ou :

$$\frac{Q}{P - I_a} = 1 - \exp^{-B(P - I_a)} \quad (\text{V.2})$$

avec Q l'écoulement en volume, P le volume de précipitation, I_a les pertes initiales de la pluie dans le sol, et b et B des coefficients. Mockus (1949) a suggéré que le coefficient b était lié aux pluies antérieures, au type de sol, à la saison et à la durée de l'épisode.

Plus tard, Mishra et Singh (1999) ont montré que cette formule pouvait être déterminée à partir :

1. d'une formulation du bilan hydrique : $Q = P_e - L$ avec le volume de pertes L et les pluies efficaces $P_e = P - I_a$;

1. United States Department of Agriculture

2. de l'hypothèse d'un refus proportionnel pour les pertes initiales par rapport aux pluies efficaces : $\frac{dL}{dP_e} \propto -L$ ou $\frac{dL}{dP_e} = -B \cdot L$ avec B la constante de proportionnalité.
3. et de l'hypothèse que la contrainte $B \cdot P_e < 1$ soit vérifiée.

L'hypothèse de proportionnalité permet alors d'écrire :

$$\begin{aligned}
 \frac{dL}{dP_e} = -B \cdot L &\Rightarrow \frac{dL}{L} = -B \cdot dP_e \\
 &\Rightarrow \ln(L) = -B \cdot P_e + C \quad (\text{par intégration}) \\
 &\Rightarrow L = P_e \cdot e^{-B \cdot P_e} \quad (C = \ln(P_e) \text{ car } L \xrightarrow{P_e \rightarrow 0} P_e) \\
 &\Rightarrow \frac{Q}{P_e} = 1 - e^{-B \cdot P_e} \quad (\text{en utilisant l'équation } Q = P_e - L) \\
 &\Rightarrow \frac{Q}{P - I_a} = 1 - e^{-B(P - I_a)}
 \end{aligned} \tag{V.3}$$

Ceci correspond à l'équation (V.2). Un développement en séries fractionnaires d'Euler de la fonction exponentielle sur le membre de droite de l'égalité donne l'équation (V.4), équivalente à la formulation standard du modèle SCS. Cette approximation est raisonnable dans le cas de l'hypothèse $B \cdot P_e < 1$:

$$\frac{Q}{P - I_a} = \frac{B(P - I_a)}{1 + B(P - I_a)} \tag{V.4}$$

ou, en posant $S_{max} = \frac{1}{B}$ avec S_{max} appelé capacité maximale de rétention :

$$\frac{Q}{P - I_a} = \frac{P - I_a}{S_{max} + P - I_a} \tag{V.5}$$

Le modèle SCS permet de calculer le volume de pluie participant à l'écoulement direct d'un événement. En regardant la figure V.3, un hydrogramme peut être distingué en trois grands types d'écoulements. Ce découpage est schématique, mais témoigne de processus bien différenciables du bassin versant.

L'écoulement par dépassement de la capacité d'infiltration du sol est considéré comme pertinent pour expliquer la réponse hydrologique des bassins en climats semi-arides ou lors de conditions de fortes intensités pluviométriques. Les écoulements de surface sont la contribution directe des précipitations tombant sur ces surfaces saturées et sont en général le principal descripteur des fortes crues.

Une partie des précipitations infiltrées chemine quasi horizontalement dans les couches supérieures du sol pour réapparaître à l'air libre, à la rencontre d'un chenal d'écoulement. Cette eau qui peut contribuer rapidement au gonflement de la crue est désignée sous le terme d'écoulement de sub-surface. L'importance de la fraction du débit total qui emprunte la voie sub-superficielle dépend essentiellement de la structure du sol.

Le dernier type d'écoulement, en provenance de la nappe phréatique, est appelé écoulement de base ou écoulement souterrain. À cause des faibles vitesses de l'eau dans le sous-sol, l'écoulement de base n'intervient que pour une faible part dans l'écoulement de crue. De plus, il n'est pas toujours relié au même événement pluvieux que l'écoulement de surface et provient généralement des pluies antécédentes. L'écoulement de base assure, le plus souvent, le débit des rivières en l'absence de précipitations et soutient les débits d'étiage (l'écoulement souterrain des régions karstiques fait exception à cette règle).

Le modèle SCS s'intéresse uniquement à la représentation des écoulements de surface qui sont le processus prédominant en cas de crue. Il permet de calculer la fraction de pluie participant à l'écoulement rapide en tenant compte d'une rétention plus ou moins forte de la part du bassin pour cet événement.

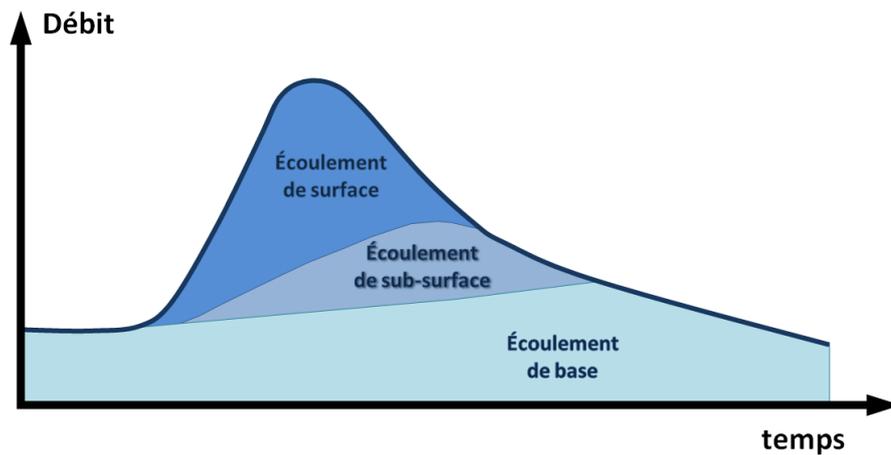


FIGURE V.3 – Distinction des différents types d'écoulement dans un hydrogramme de crue.

Une hypothèse supplémentaire souvent utilisée dans le SCS est que $I_a = \lambda \cdot S_{max}$, avec λ compris entre 0 et 1 (la valeur classiquement utilisée est $\lambda = 0.2$). Sous cette hypothèse, l'écoulement peut alors être déterminé à partir de la formule suivante :

$$Q = \frac{(P - \lambda \cdot S_{max})^2}{P + (1 - \lambda) \cdot S_{max}} \quad (V.6)$$

La valeur de λ peut varier selon l'état du bassin. Une valeur de $\lambda = 0$ indique un bassin très humide où le volume de pluie participant à l'écoulement sera fort. Au contraire $\lambda = 1$ caractérise un bassin très sec donc avec des pluies efficaces plus faibles. Le cas particulier de $S_{max} = 0\text{mm}$ correspond à un bassin sans capacité de rétention pour lequel tout le volume de pluie se retrouve systématiquement en écoulement.

Un hydrogramme unitaire est ensuite appliqué au volume obtenu avec la formule (V.6) afin de représenter le transfert au sein du bassin et de fournir un hydrogramme d'écoulements.

Les conditions de saturations antérieures et les surfaces de contribution.

Le passé de la méthode SCS en fait une méthode purement empirique. Ceci constitue son principal point fort, mais également une limitation du fait qu'elle ne permet pas une interprétation directe en terme de processus. La capacité maximale de rétention S_{max} est une variable qui intègre le type de sol, leur saturation ou toute autres processus pouvant avoir un impact sur la réponse hydrologique du bassin. Le fait que toutes les pertes possibles soient regroupées dans une seule variable fait du modèle SCS un modèle global. Cependant, le travail de Mockus (1949) suggère que la valeur de S_{max} est liée au volume d'infiltration ou stock disponible - selon lequel est le plus faible. Deux études plus récentes ont tenté d'approfondir cette interprétation. Steenhuis *et al.* (1995) ont interprété la méthode comme étant équivalente à supposer une aire de contribution variable pour la génération des écoulements sur laquelle l'événement pluvieux est suffisant pour compenser la valeurs des pertes initiales d'un événement. Le pourcentage du bassin contribuant à la génération des écoulements pour une valeur de pluie donnée est alors directement relié à la valeur des pertes initiales I_a (ou au paramètre λ). Ils ont montré une assez bonne cohérence avec les données observées pour plusieurs bassins australiens et américains (ayant des surfaces de 0 à 70 km²) avec des sols perméables. L'ajustement de la valeur S_{max} pour chaque bassin a permis d'obtenir des valeurs dans une gamme de 80 à 400 mm.

Les limitations de la méthode

Il est bon de rappeler ce qui est calculé avec l'approche SCS. Le modèle fournit un écoulement direct pour un événement de pluie donné, après une certaine rétention initiale précédant le début de cet écoulement. Cet écoulement est routé via un hydrogramme unitaire (ou une autre méthode de routage) afin de prédire l'hydrogramme. Le modèle est donc sujet à tous les problèmes et limitations associés à la distinction du ruissellement rapide dans l'hydrogramme de débit total.

L'approche SCS pour la génération d'écoulement a fait l'objet d'un grand nombre de critiques (Hjelmfelt *et al.*, 1982; Bales et Betson, 1982). Des travaux complémentaires sont nécessaires pour clarifier sous quelles conditions la méthode fournit de bonnes prévisions. Mishra et Singh (1999) ont montré qu'une formule généralisée de la méthode (formule (V.5) avec un paramètre supplémentaire) donne de meilleurs résultats que la formulation originale (avec $\lambda = 0.2$ et $a = 0$) ce qui est logique du fait des deux paramètres supplémentaires pour le calage :

$$\frac{Q}{P - I_a} = \frac{P - I_a}{S_{max} + a \cdot (P - I_a)} \quad (V.7)$$

Hjelmfelt *et al.* (1982) suggèrent que le paramètre S_{max} , plutôt que d'être considéré comme une variable caractéristique de l'état des sols, soit considéré comme une variable stochastique. Leurs analyses des maxima annuels de ruissellement rapide sur deux petits bassins dans l'Iowa laissent penser que la capacité de stockage S_{max} calculée pour différents événements est approximativement distribuée selon une loi log-normale avec un coefficient de variation de l'ordre de 20%. Cependant, ils n'ont trouvé aucune corrélation marquée entre les S_{max} et les conditions antérieures pour des événements.

Bilan : malgré ses limitations, le modèle SCS continue d'être largement utilisé. Ceci est en grande partie dû au fait que le paramètre S_{max} peut être estimé en utilisant les tables de « curve numbers » fournies par l'USDA. Les curve numbers sont des valeurs tenant compte des conditions hydrologiques moyennes ou du type de sol du bassin. **Dans le cadre d'une régionalisation, la méthode SCS fournit une estimation simple avec une quantité limitée de paramètres.** Elle ne modélise que le processus de ruissellement rapide du bassin. Or, dans le cas d'une étude des débits extrêmes, le ruissellement rapide est le type d'écoulement prépondérant. Nous restons conscient que le modèle SCS est très simpliste et n'a pas pour but d'être utilisé en simulation continue, mais seulement en événementiel.

V. 2. 2 Le principe de la simulation avec le SCS

Comme nous l'avons déjà signalé, l'objectif est de remplacer MORDOR par un modèle SCS dans le processus de simulation SCHADEX. À ce stade, nous supposons qu'une distribution sur la pluie spatiale est déjà estimée pour le bassin et seule la problématique du passage pluie-débit nous intéresse.

1. Une distribution MEWP sur les pluies centrales P_C et les tableaux de contingence pour les pluies adjacentes P_{a-} et P_{a+} (modèles probabilistes pour les pluies adjacentes) sont estimés sur la série de pluie spatiale.
2. À partir des tableaux de contingence et de la distributions des pluies centrales, un épisode de pluie (pluie centrale P_C^* et pluies adjacentes P_{a-}^* et P_{a+}^*) est généré puis probabilisé à l'aide de la distribution MEWP et des tableaux de contingence. La probabilité de l'épisode $Proba_{epi}^*$ est le produit des probabilités de chacune des trois pluies. Cette génération de scénarios de pluies permet de prendre en compte l'aléa pluviométrique.
3. Le scénario pluie de probabilité $Proba_{epi}^*$ est alors transformé en écoulement avec le modèle SCS et un hydrogramme unitaire. L'écoulement obtenu dépend du scénario de pluie généré, mais également de l'état du bassin λ et de la capacité maximale de rétention du bassin S_{max} .

Simulation d'un scénario de débit avec le modèle SCS

Les événements pluvieux sur trois pas de temps sont un concept utilisé dans la simulation SCHADEX et qui est conservé dans la simulation simplifiée. Cependant, le SCS ne se prête pas au routage des écoulements d'un pas de temps à l'autre. Afin de ne pas se poser le problème du routage, les écoulements seront calculés sur des durées croissantes (USDA, 1994). Au pas de temps t , le modèle SCS incrémental permet de calculer le cumul d'écoulement $\sum_{k=1}^t E_k^{SCS}$ entre le pas de temps 1 et t en fonction du cumul de pluie $\sum_{k=1}^t P_k$ de 1 à t :

$$\sum_{k=1}^t E_k^{SCS} = \frac{(\sum_{k=1}^t P_k - \lambda \cdot S_{max})^2}{\sum_{k=1}^t P_k + (1 - \lambda) \cdot S_{max}} \quad (V.8)$$

Par différence entre deux pas de temps consécutifs, les volumes d'écoulements E_t^{SCS} peuvent être calculés pour $t \in \{1; 2; 3\}$. En utilisant un hydrogramme unitaire HU , l'hydrogramme des écoulements E_t à l'exutoire du bassin peut être estimé. En considérant un hydrogramme unitaire sur trois pas de temps, les écoulements E_t s'expriment :

$$\begin{cases} E_1 = HU_1 \cdot E_1^{SCS} \\ E_2 = HU_2 \cdot E_1^{SCS} + HU_1 \cdot E_2^{SCS} \\ E_3 = HU_3 \cdot E_1^{SCS} + HU_2 \cdot E_2^{SCS} + HU_1 \cdot E_3^{SCS} \\ E_4 = HU_3 \cdot E_2^{SCS} + HU_2 \cdot E_3^{SCS} \\ \dots \end{cases} \quad (V.9)$$

Un exemple d'événement simulé avec ce modèle SCS incrémental est présenté figure V.4. Le graphe représente le hyétogramme des pluies (barres bleues), les écoulements observés (courbe bleue) ou simulés avec le modèle MORDOR calé pour ce bassin (courbe verte) d'un jour J . La courbe rouge représente l'hydrogramme de crue obtenue avec le modèle SCS incrémental puis passage dans l'hydrogramme unitaire. Le paramètre S_{max} est déterminé à partir de la valeur maximale des stocks MORDOR $S_{max} = U_{max} + Z_{max} + L_{max} = 230 \text{ mm}$ qui représentent les stocks superficiels de MORDOR et qui régissent l'écoulement rapide. L'état du bassin λ est

quant à lui déterminé à l'aide du remplissage de ces stocks au début de l'événement dans MORDOR : $\lambda = 1 - \frac{U(J-1)+Z(J-1)+L(J-1)}{U_{max}+Z_{max}+L_{max}} = 28\%$. La fonction de transfert choisie est celle obtenue dans MORDOR soit $HU = \{0.46; 0.52; 0.02\}$. L'écoulement simulé avec le SCS atteint un pic similaire à celui de l'hydrogramme observé. Le processus de décrue est mal représenté puisque le SCS ne prend en compte aucun processus lent pour modéliser ce comportement. Les débits autour de la pointe, qui nous intéressent, sont plutôt bien représentés dans ce cas.

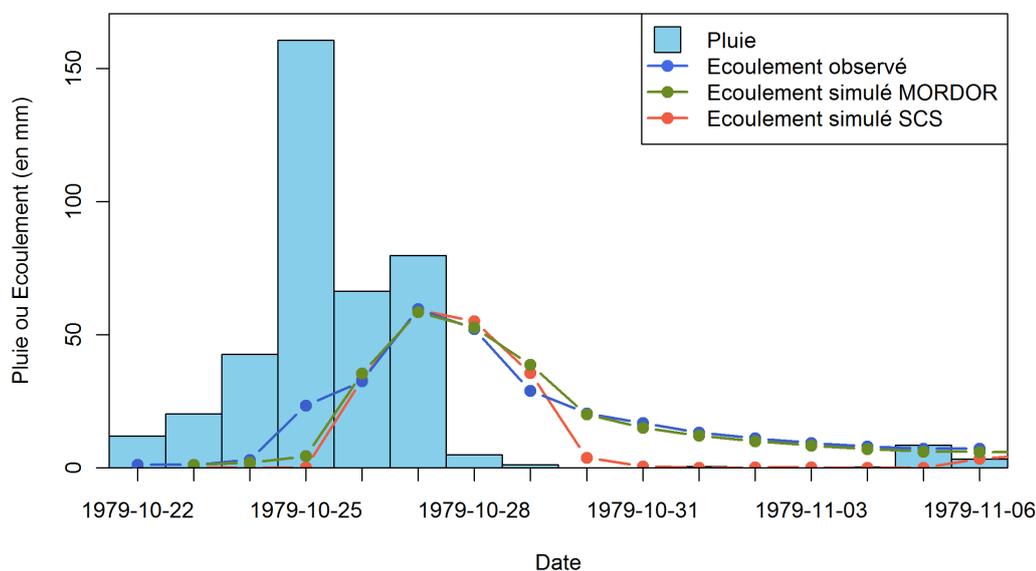


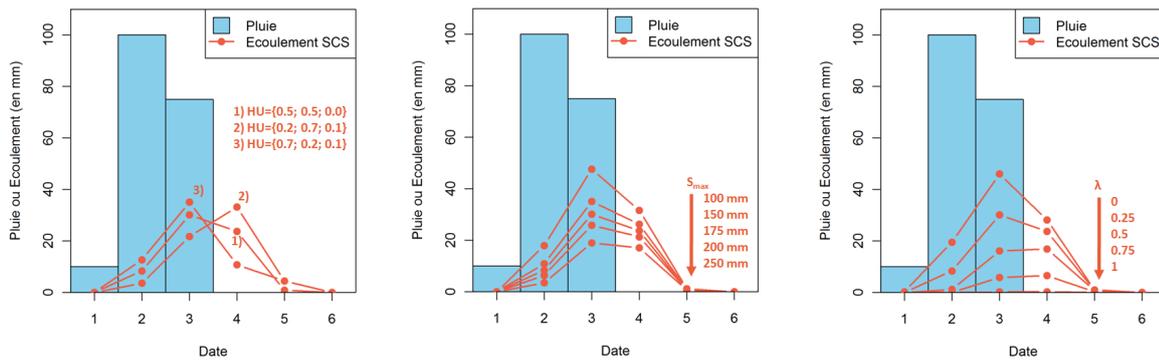
FIGURE V.4 – Un exemple d'événement simulé pour le bassin de l'Agout à Fraïsse.

Influence des différents paramètres du modèle SCS sur la simulation d'un scénario

Les seuls paramètres du modèle SCS sont :

- pour la fonction de production du SCS : la capacité maximale de rétention S_{max} et un état de saturation du bassin λ ,
- pour la fonction de transfert : un hydrogramme unitaire sur trois pas de temps.

Pour un scénario de pluie fixé, la figure V.5 trace les hydrogrammes d'écoulements obtenus selon les jeux de paramètres utilisés. Le jeu de paramètres de référence utilisé est $S_{max} = 175 \text{ mm}$, $\lambda = 0.25$ et l'hydrogramme unitaire $HU = \{0.5; 0.5; 0.0\}$. La forme de l'hydrogramme unitaire HU impacte directement la forme de l'hydrogramme des écoulements. Elle décrit en particulier l'étalement de ce dernier et la position de la pointe (figure V.5 (a)). La capacité maximale de rétention représente la capacité du bassin à absorber de façon plus ou moins importante le volume des précipitations (figure V.5 (b)). Plus la valeur de S_{max} est forte, plus le bassin aura un effet « tampon » important c'est-à-dire plus la pointe de l'hydrogramme sera faible (diminution du volume de la crue). Le choix de l'hydrogramme unitaire HU et des capacités maximales de rétention S_{max} est fait pour le bassin versant et est donc invariant d'un scénario à un autre. Les deux valeurs permettent de décrire le comportement général du bassin. L'état ponctuel du bassin avant un scénario de pluie est quant à lui décrit par le paramètre λ . La valeur de λ varie à chaque scénario de pluie afin de décrire la variabilité des conditions de saturation du bassin au moment de l'événement. Plus sa valeur est élevée, plus la réponse hydrologique du bassin sera faible (figure V.5 (c)). L'état de saturation λ du bassin peut être décrit par une distribution. Ainsi, une valeur donnée de λ est plus ou moins probable selon le bassin.



(a) Influence de l'hydrogramme unitaire. (b) Influence de la capacité maximale de rétention S_{max} . (c) Influence de l'état du bassin λ .

FIGURE V.5 – Influence des différents paramètres du modèle SCS sur la simulation d'un événement.

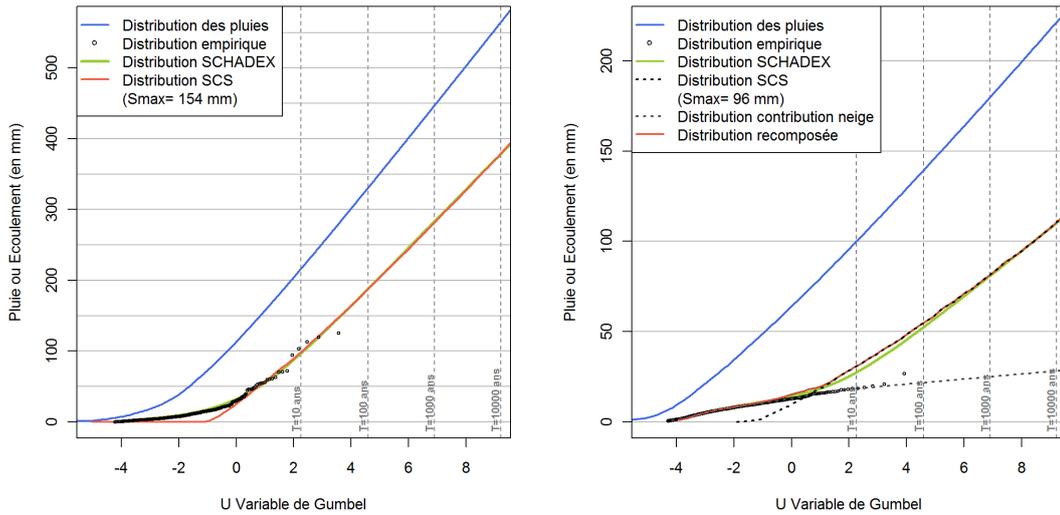
Simulation d'une distribution de débit avec une simulation stochastique et le SCS

Comme il l'a été signalé un peu plus haut, le modèle SCS permet de générer une représentation du ruissellement rapide. Sur des bassins de régime pluvial, le comportement asymptotique de la distribution est essentiellement porté par celui du risque pluie. Pour des bassins de montagne avec une composante nivale forte, ce comportement n'est retrouvé que pour des temps de retour plus élevés. Une partie intermédiaire de la distribution (entre des temps de retour de 10 ans à 100 ans selon les bassins) correspond à des écoulements à risques issus à la fois d'un processus pluviométrique, mais également d'un processus nival lié à l'onde de fonte au printemps (figure V.6). En effet, de fortes crues peuvent survenir lorsque des pluies plus ou moins fortes s'ajoutent à la fonte des neiges au printemps. Les crues des Pyrénées, en 2013, sont une illustration parfaite de ce risque où un réchauffement, suivi de précipitations, a provoqué une forte montée des eaux. Ce mélange de deux processus ne peut donc être simulé uniquement avec le modèle SCS. C'est pourquoi, une distribution de la contribution nivale, visant à représenter ce processus nival, peut être prise en compte. La distribution finale sera une loi composite regroupant la distribution d'écoulements rapides proposée par le SCS et la distribution des contributions nivales.

La simulation stochastique simplifiée mis en place sera dénommée EREBOR (Écoulements Régionaux Extrêmes sur un Bassin à Observations Réduites). Le processus complet d'une simulation EREBOR est présenté sur le schéma V.7. Les données d'entrées sont distinguées en deux catégories, pour décrire l'aléa pluie et pour décrire le bassin.

La description de l'aléa pluie est réalisée par l'intermédiaire des deux modèles probabilistes déjà présentés : une distribution MEWP ajustée sur les pluies centrales P_C et les tableaux de contingences pour les pluies adjacentes P_{a-} et P_{a+} . Ces deux modèles permettent une probabilisation des épisodes pluvieux (voir II. 3, page 39).

La description du comportement hydrologique du bassin s'appuie quant à elle sur la capacité maximale de rétention S_{max} qui traduit la capacité du bassin à retenir une partie des précipitations et sur un hydrogramme unitaire HU sur trois pas de temps journaliers servant de fonction de transfert. Ce sont des paramètres descriptifs du bassin. Ils sont donc fixes pour tout événement pluvieux. Au contraire, l'état initial du bassin λ est considéré comme une variable aléatoire et est amené à varier d'un événement à un autre. Chaque état du bassin est représenté par sa probabilité (un coefficient λ nul – un bassin saturé – est plus ou moins probable selon le bassin considéré).



(a) Purement pluvial – Agout à Fraïsse (45 km²) (b) Composante nivale – Romanche à Champeau (998 km²)

FIGURE V.6 – Distribution par simulation avec le SCS avec ou sans addition d'un risque nivale.

Une fois les différentes entrées définies pour le bassin étudié (capacité de rétention maximale S_{max} , fonction de répartition des états de saturations λ , hydrogramme unitaire HU sur trois pas de temps et éventuellement la distribution du risque nivale), la simulation stochastique à proprement parler peut être abordée.

1. Un événement de pluie est généré puis probabilisé avec les modèles probabilistes des pluies. Une valeur de pluie centrale P_C^* est tirée uniformément entre 0 et 1000 mm puis des valeurs de pluies adjacentes P_{a-}^* et P_{a+}^* sont tirées sachant cette pluie centrale. La probabilité $Proba_{epi}^*$ de cet épisode P_{epi}^* est donnée par la formule (V.10).

$$\begin{aligned}
 Proba_{epi}^* = & \quad Proba\left(\frac{P_{a-}^*}{P_C^*} | P_C^*\right) \\
 & \cdot Proba\left(P_C^*\right) \\
 & \cdot Proba\left(\frac{P_{a+}^*}{P_C^*} | (P_{a-}^* + P_C^*)\right)
 \end{aligned} \tag{V.10}$$

2. En parallèle, un état initial du bassin λ^* pour cet épisode P_{epi} est tiré aléatoirement dans la fonction de répartition de ce paramètre².
3. Le modèle SCS permet alors de transformer le volume de précipitation en volume d'écoulement.

$$\begin{aligned}
 E_1^{SCS} &= \frac{(P_{a-}^* - \lambda^* \cdot S_{max})^2}{P_{a-}^* + (1 - \lambda^*) \cdot S_{max}} \\
 E_1^{SCS} + E_2^{SCS} &= \frac{(P_{a-}^* + P_C^* - \lambda^* \cdot S_{max})^2}{P_{a-}^* + P_C^* + (1 - \lambda^*) \cdot S_{max}} \\
 E_1^{SCS} + E_2^{SCS} + E_3^{SCS} &= \frac{(P_{a-}^* + P_C^* + P_{a+}^* - \lambda^* \cdot S_{max})^2}{P_{a-}^* + P_C^* + P_{a+}^* + (1 - \lambda^*) \cdot S_{max}}
 \end{aligned} \tag{V.11}$$

2. La fonction de répartition sera déterminée empiriquement.

Les valeurs de E_1^{SCS} , E_2^{SCS} et E_3^{SCS} sont calculées par différences entre deux pas de temps consécutifs. Le scénario de volume d'écoulement obtenu correspond alors au volume d'écoulement se retrouvant dans la rivière.

4. Ce scénario est routé par l'hydrogramme unitaire choisi pour le bassin afin de déterminer les écoulements à l'exutoire du bassin pour cet épisode. Le scénario d'écoulement $\{E_1^{SCS}; E_2^{SCS}; E_3^{SCS}\}$ sur trois pas de temps est alors transformé en un scénario d'écoulement sur quatre pas de temps $E_{epi} = \{E_1; E_2; E_3; E_4\}$. Le quatrième pas de temps est calculé, car la fonction de transfert décale les écoulements par rapport aux pluies. Le quatrième pas de temps assure l'observation d'un écoulement maximum puis d'un début de décrue.

$$\begin{aligned} E_1 &= E_1^{SCS} \cdot HU_1 \\ E_2 &= E_1^{SCS} \cdot HU_2 + E_2^{SCS} \cdot HU_1 \\ E_3 &= E_1^{SCS} \cdot HU_3 + E_2^{SCS} \cdot HU_2 + E_3^{SCS} \cdot HU_1 \\ E_4 &= E_2^{SCS} \cdot HU_3 + E_3^{SCS} \cdot HU_2 \end{aligned} \quad (V.12)$$

Au final, le modèle SCS fournit un scénario d'écoulement E_{epi} associé à l'épisode de pluie P_{epi}^* et à l'état de bassin λ^* .

5. La simulation stochastique consiste à répéter cette production de scénario n fois, de façon à parcourir tout l'aléa pluviométrique et tous les états de saturation du bassin. En pratique, $n = 100\,000$ générations sont suffisantes pour avoir un croisement exhaustif des deux aléas. Nous disposons alors d'une population d'écoulements maximum $(\max(E_{epi}))_n$. Les probabilités des épisodes de précipitations $Proba_{epi}^*$ fournissent une pondération probabiliste pour les écoulements maximum $\max(E_{epi})$. La probabilité d'un état de saturation est quant à elle directement intégrée dans la population des écoulements puisque chaque λ^* a été tiré dans sa fonction de répartition. La fonction de répartition des écoulements maximum F_{SCS} est déterminée en triant par ordre croissant la population $(\max(E_{epi}))_n$ et en intégrant la fonction de densité des probabilités normalisées $\frac{Proba_{epi}^*}{\sum_{epi=1}^n Proba_{epi}^*}$.
6. Si le bassin a un risque de crues printanières liées à la fusion nivale, une fonction de répartition F_{CN} des écoulements de fonte peut être considérée³.
7. La distribution finale F est alors une distribution composite tenant compte à la fois du risque dû à des contributions nivales et pluviométriques.

$$F(E) = F_{SCS}(E) \cdot p_{SCS} + F_{CN}(E) \cdot p_{CN} \quad (V.13)$$

avec p_{SCS} (respectivement p_{CN}) le pourcentage d'occurrence des écoulements dus aux pluies (respectivement dus à la fonte de neige). Nous disposons alors de la distribution des écoulements au pas de temps d'étude. Les deux pourcentages d'occurrence sont déterminés selon la période considérée pour les contributions nivales (*par exemple pour trois mois/ans*, $p_{SCS} = 0.8$ et $p_{CN} = 0.2$). Pour revenir à une distribution des débits moyens (sur le pas de temps d'étude) Q_{pdt} , il suffit de convertir les écoulements en débits avec :

$$Q_{pdt} = E \cdot S_{BV} \cdot \frac{1}{3.6 \cdot pdt}$$

avec pdt le pas de temps d'étude (en heure) et S_{BV} la surface du bassin versant (en km²).

3. La fonction de répartition sera déterminée empiriquement.

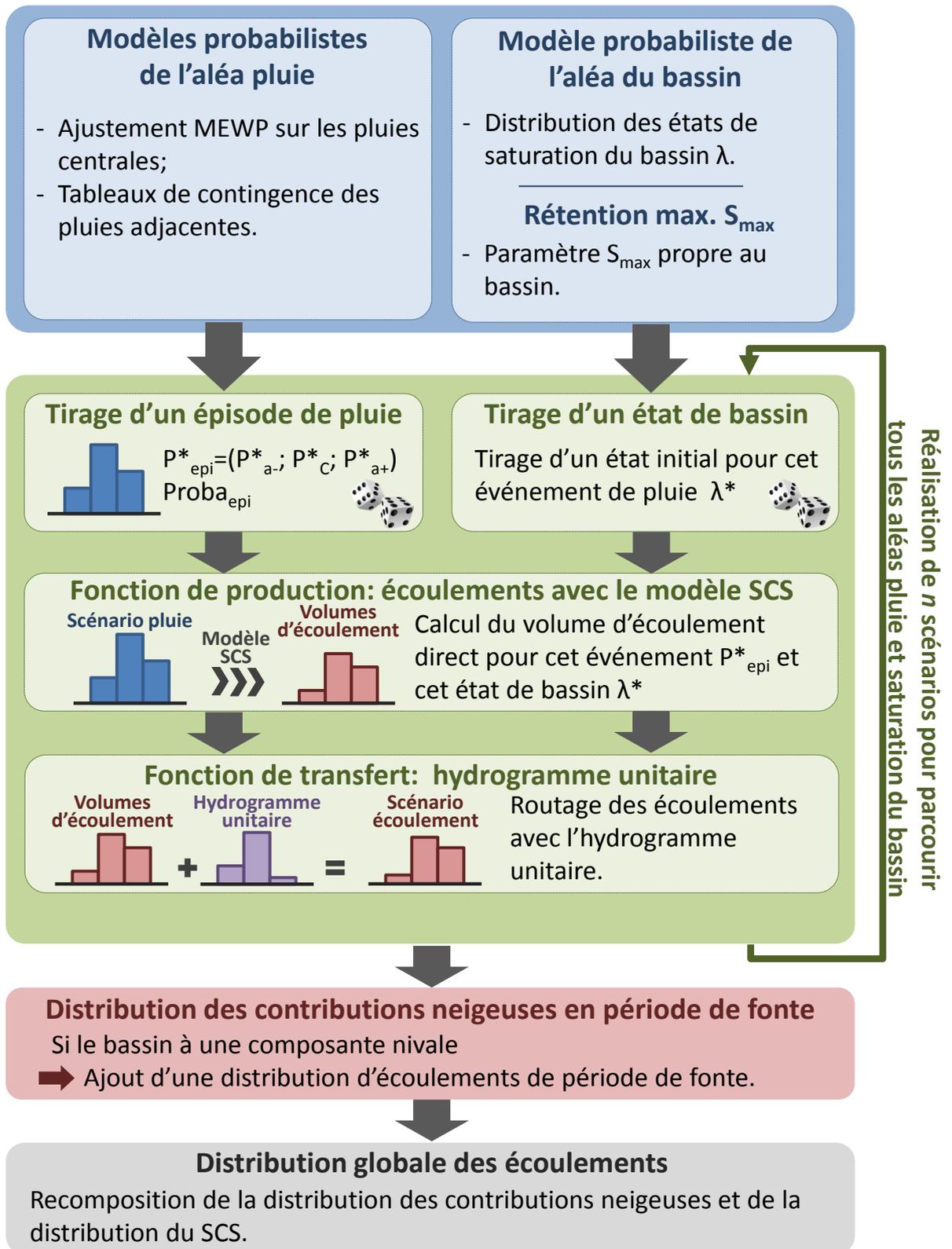


FIGURE V.7 – Schéma d'une simulation EREBOR.

Bilan : la simulation EREBOR permet de fournir une distribution des débits moyens sur le pas de temps d'étude. Dans notre cas, son utilisation s'est limité à un pas de temps journalier même si dans l'absolu la méthode est applicable à un pas de temps quelconque. **La simulation permet de reprendre des concepts utilisés dans la méthode SCHADEX. Elle croise l'aléa pluie et l'aléa saturation du bassin de façon à mieux représenter le risque en débit.** Le principe d'épisodes pluvieux triangulaire de SCHADEX est également réutilisé afin de représenter le risque pluie. L'intérêt majeur, dans l'optique d'une régionalisation du passage pluie-débit, est l'utilisation du modèle SCS en tant que modèle événementiel. **Le nombre de paramètres nécessaires pour lancer une simulation est bien inférieur à celui de la méthode SCHADEX qui utilise le modèle MORDOR.** Cependant, il reste à régionaliser les différents paramètres et à évaluer la capacité de ce modèle à retrouver des distributions similaires à celles obtenues avec la méthode SCHADEX.

V. 3 Régionalisation des paramètres de la méthode EREBOR

La distribution des écoulements obtenue avec la simulation EREBOR dépend essentiellement de trois paramètres (quatre si le bassin a une composante nivale significative). L'objectif étant d'utiliser cette méthode en contexte non jaugé, les paramètres doivent pouvoir être estimés pour un bassin sans disposer d'information locale de débit. Nous allons donc maintenant régionaliser les paramètres suivants :

- l'hydrogramme unitaire du bassin HU ;
- la distribution des différents états λ du bassin ;
- la capacité maximale de rétention S_{max} du bassin ;
- dans le cas d'un bassin nivale, la distribution des contributions nivales du printemps.

Ces paramètres vont être estimés pour chaque étude SCHADEX de la base en utilisant les informations disponibles (données de débit, modèle MORDOR...) puis une régionalisation de ces estimations sera proposée.

V. 3. 1 La régionalisation des paramètres

V. 3. 1.1 La distribution des contributions nivales

La distribution des contributions neigeuses vise à prendre en compte un risque de crue dû à des précipitations sur un débit printanier déjà fort du fait de l'onde de fonte. Elle concerne donc des bassins dont le régime hydrologique est fortement marqué par une composante nivale. À partir des séries journalières de pluies et débits, tous les épisodes de précipitations, composés d'une pluie centrale et des deux pluies adjacentes, sont identifiés. L'évènement en débit consécutif à cet épisode pluvieux est détecté grâce à un algorithme mis en place à EDF-DTG (figure V.8). Une population d'épisodes de pluie sur trois pas de temps et la population des épisodes de débit associés sont alors disponibles. Sur la figure V.8, les écoulements bleus représentent un volume d'écoulements antécédent à l'épisode et les volumes rouges sont les volumes d'écoulement rapide.

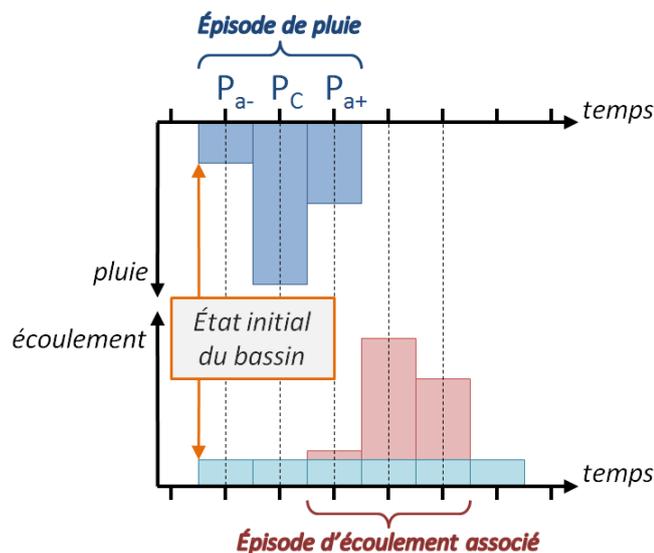


FIGURE V.8 – Schéma d'un évènement pluie et de l'évènement débit associé.

Le SCS s'intéresse à la représentation des écoulements directs qui sont, en général, prépondérants dans le processus des crues. Cependant, il est possible que l'écoulement antécédent la crue ne soit pas négligeable devant son processus rapide. L'idée de l'intégration d'une distribution des contributions nivales F_{CN} est de prendre en compte ce processus pour considérer les forts débits du printemps sur les bassins nivaux. Seuls les écoulements survenant sur les mois de mai-juin-juillet, qui correspondent au mois où la contribution nivale est la plus forte, sont conservés. Les distributions de ces écoulements pour chaque bassin de la base d'études (courbes en pointillés) sont tracées sur les graphes de la figure V.9. Les courbes en trait plein représentent la distribution médiane de chaque régime hydrologique. Nous constatons que les bassins de type cévenol ou pluvial ont des risques de contributions nivales au printemps très faibles. Au contraire, des bassins nivaux ou nivo-glaciaires ont des contributions beaucoup plus fortes au printemps. La classe nivo-pluviale semble être moins marquée avec des distributions, selon les bassins, soit plutôt nivales soit plutôt pluviales. Ceci laisse penser que cette classe n'est pas très crédible et qu'elle mériterait d'être séparée en deux sous-classes. Les bassins du Tech à Reynes, de la Tet à Vinça et du Verdon à Sainte-Croix semblent être plus proches d'un régime pluvial alors que les bassins de l'Arly à Moulin Ravier, le Gelon à Rochette et la Touyre à Montferrier sont plus proches d'un régime nival au vu des distributions des débits de printemps.

Finalement, une distribution médiane par classe de régime apparaît être une première estimation raisonnable des distributions de contributions nivales au printemps (classe des bassins nivo-pluviaux mise à part). Le choix de la saison de fonte sur les mois de mai-juin-juillet est une hypothèse simplificatrice. À terme, il pourrait être judicieux de définir une saison de fonte selon le régime hydrologique et de redéfinir des classes de régime plus adaptées.

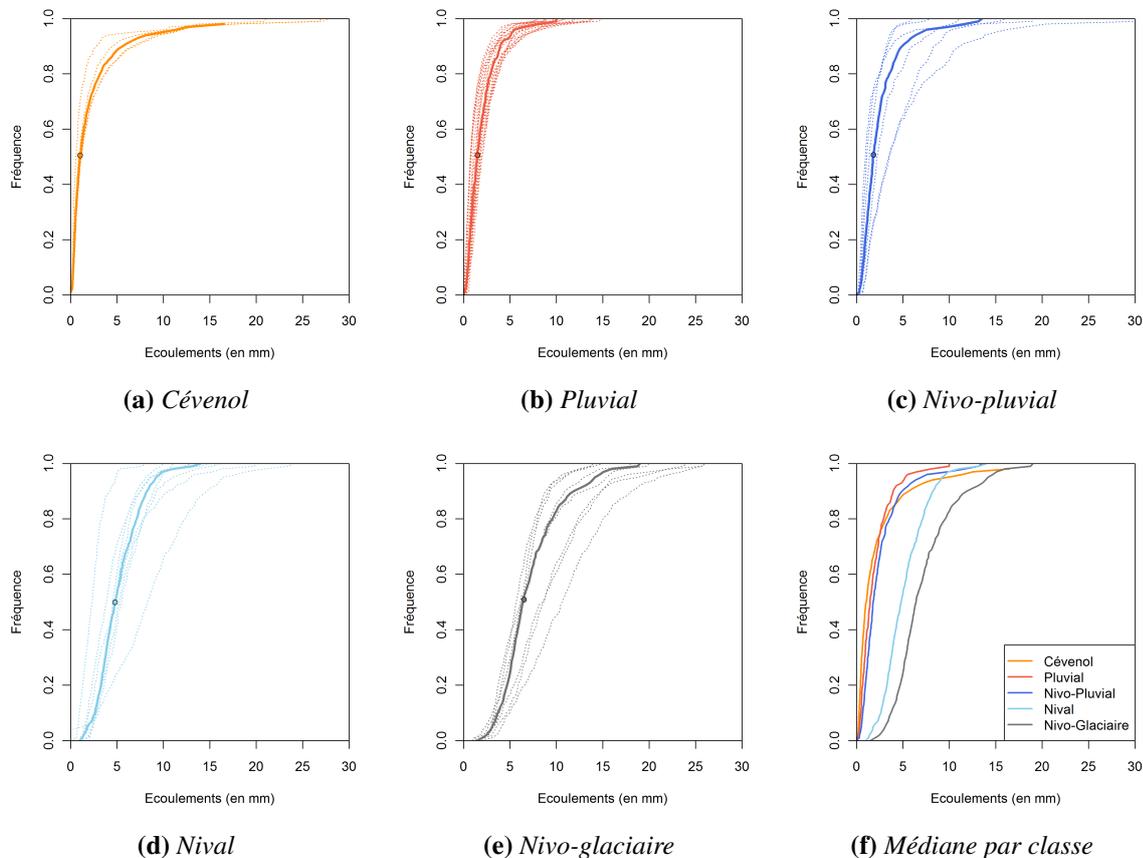


FIGURE V.9 – Distributions des contributions nivales F_{CN} par régime hydrologique.

V. 3. 1.2 L'hydrogramme unitaire

La fonction de transfert choisie est un hydrogramme unitaire sur trois pas de temps, ici journaliers. De nombreux travaux se sont intéressés à la régionalisation de l'hydrogramme unitaire. La forme de l'hydrogramme dépend en partie de la taille du bassin et de sa géomorphologie. Une estimation régionale de l'hydrogramme unitaire s'appuyant sur ces grandeurs pourrait donc être étudiée, par exemple en utilisant des hydrogrammes unitaires géomorphologiques. Cependant, l'hypothèse effectuée est de considérer les fonctions de transferts du modèle MORDOR comme une représentation des hydrogrammes unitaires. Ce choix permet une première régionalisation approximative des hydrogrammes unitaires qui pourra être revue ultérieurement. Dans un premier temps, les hydrogrammes unitaires obtenus avec MORDOR sont donc collectés pour toutes les études de la base. La figure V.10 montre, pour chaque pas de temps, un boxplot des valeurs de l'hydrogramme unitaire HU selon les différents régimes hydrologiques. Les bassins nivo-pluviaux ne figurent pas sur le graphe, car ils ont été répartis entre le pluvial et le nival selon la remarque faite dans le paragraphe précédent. Une nette différence sur la forme médiane des hydrogrammes unitaires peut être constatée selon le régime. Les régimes pluvial et cévenol correspondent plutôt à des fonctions de transfert contrastées, du type $HU = \{20\%; 70\%; 10\%\}$, alors que les régimes nival et nivo-glaciaire sont plutôt caractérisés par un hydrogramme plus plat, du type $HU = \{40\%; 50\%; 10\%\}$. Hydrologiquement parlant, ceci signifie qu'un bassin versant pluvial (ou cévenol) a, en général, tendance à répondre de façon plus impulsive à une précipitation et qu'un bassin nival qui a plutôt tendance à absorber le volume de précipitation et à l'étaler dans le temps.

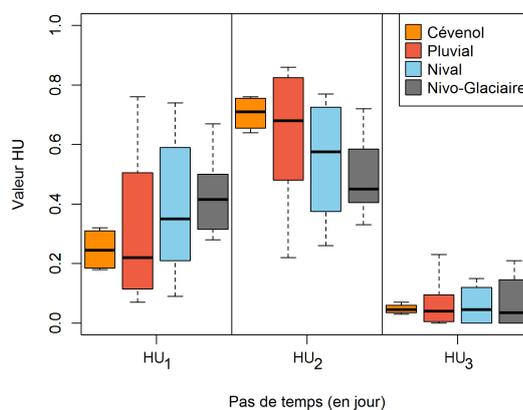


FIGURE V.10 – Boxplot des valeurs de l'hydrogramme unitaire MORDOR sur la base des 45 études par régime hydrologique.

Le même type de graphe peut être fait en séparant cette fois les bassins en deux classes de surface (nous nous limitons à deux classes pour conserver suffisamment d'individus dans chaque sous-population). La figure V.11 (b) montre un comportement similaire à ce qui a déjà pu être observé. Cependant, nous notons une différence de comportement entre le nival et le nivo-glaciaire nettement plus marquée. Pour des bassins de surface plus faible, les fonctions de transfert déterminées par MORDOR sont très variables pour le pluvial. Cette différence de comportement selon la surface est en grande partie due au fait que sur des petits bassins, le pas de temps journalier est moins adapté pour la réalisation des études. Faute de données infra-journalières, les études SCHADEX ont été réalisées à ce pas de temps et la fonction de transfert de MORDOR inscrit tout le processus d'étalement dans un seul pas de temps.

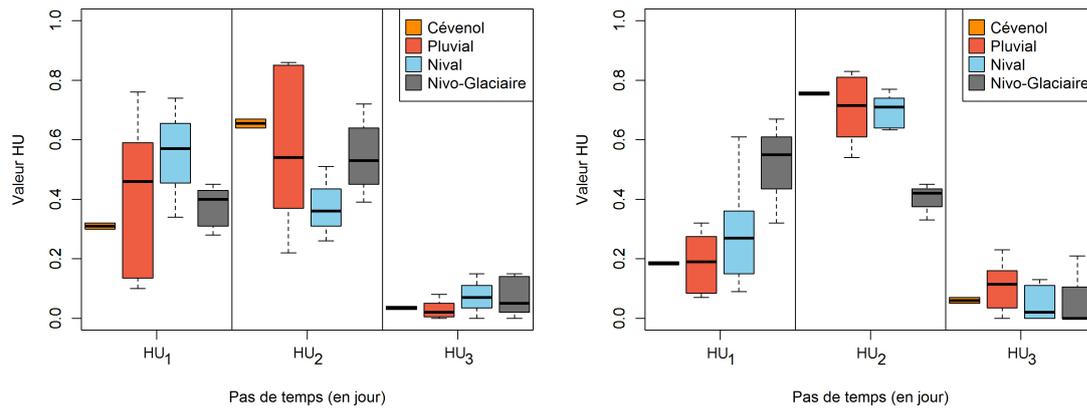
(a) Surface de moins de 700 km² (25 bassins/45)(b) Surface de plus de 700 km² (20 bassins/45)

FIGURE V.11 – Boxplot des valeurs de l’hydrogramme unitaire par régime hydrologique et selon deux classes de surface.

Dans un premier temps, une forme d’hydrogramme est donc fixée par régime hydrologique (figure V.12).

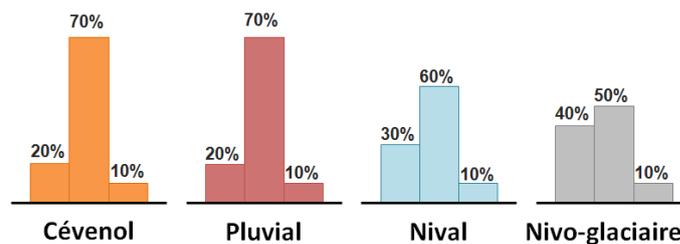


FIGURE V.12 – Hydrogrammes unitaires choisis par classe de régimes hydrologiques.

Influence sur la distribution des écoulements

Considérons comme fixes tous les paramètres d’une simulation EREBOR pour un bassin donné et intéressons nous uniquement à faire varier l’hydrogramme unitaire. L’influence de l’hydrogramme unitaire sur la réponse à un scénario de pluie donnée a été illustrée figure V.5 (a) (page 168). Son impact sur la distribution des écoulements journaliers se fait sentir sur l’asymptote (figure V.13) et en particulier sur le rendement asymptotique⁴.

Le compromis d’un hydrogramme unitaire moyen par régime hydrologique permet donc de générer un rendement asymptotique moyen selon la classe de régime. Des bassins avec des processus nivaux disposent d’un hydrogramme unitaire induisant un étalement des crues plus importants, et donc un rendement asymptotique plus faible, que pour des bassins avec des processus principalement pluviaux. Cette régionalisation de l’hydrogramme unitaire reste un choix schématique qui pourra faire l’objet d’un travail complémentaire pour affiner son estimation.

4. Le rendement asymptotique correspond au ratio du gradex des débits et du gradex des pluies. Par exemple dans le cas de la méthode du gradex, une hypothèse est de considérer le rendement asymptotique égal à 1 c’est-à-dire que les distributions pluie et débit ont des asymptotes parallèles.

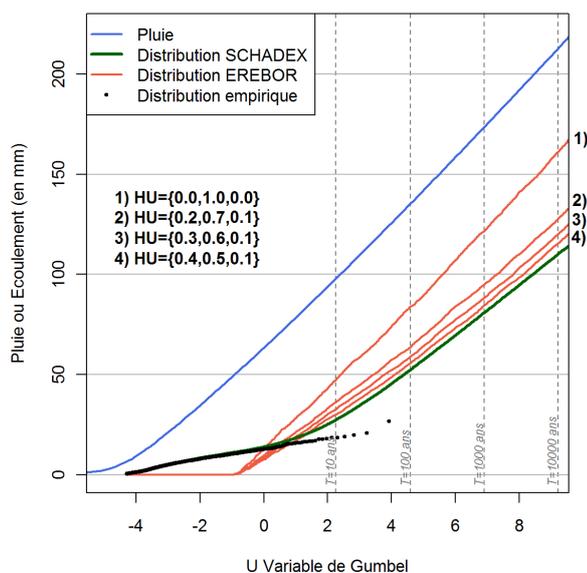


FIGURE V.13 – Distributions de débits obtenues pour différents hydrogrammes unitaires. Cas de la Romanche à Champeau (Nivo-glaciaire, 998 km²).

V. 3. 1.3 La distribution des états de saturation

Pour décrire l'état de saturation d'un bassin non jaugé, une distribution du paramètre λ doit être estimée. Afin d'avoir une idée de la distribution des stocks du bassin, les valeurs des stocks U , Z et L de MORDOR sont récupérées avant chaque épisode de pluie. Ces trois réservoirs contribuent à l'interception des pluies et donc aux écoulements rapides du bassin (section I. 2, 13). Le réservoir N représente les écoulements de base avec une dynamique plus lente et ne nous intéresse donc aucunement ici. Nous disposons ainsi des états initiaux des stocks au début d'un épisode (figure V.8) et de la capacité maximale de chaque stock (U_{max} , Z_{max} et L_{max}). Pour chaque épisode, un pourcentage de remplissage des réservoirs MORDOR est estimé et est assimilé à l'état de saturation du bassin λ :

$$\lambda = 1 - \frac{U + Z + L}{U_{max} + Z_{max} + L_{max}} \quad (\text{V.14})$$

Cependant, l'estimation de λ faite en considérant ces trois réservoirs MORDOR fournit une distribution pour laquelle il est impossible d'observer une saturation générale ($\lambda = 0$) du bassin (figure V.14 (a)). Ceci peut s'expliquer par la formulation même du modèle MORDOR et en particulier les lois de vidange et de remplissage du réservoir L (annexe A). Les apports de L sont inversement proportionnels à son taux de remplissage ce qui induit que le réservoir L peut se remplir puis se vider rapidement et la valeur de L n'est accessible qu'en fin de pas de temps. Cette formulation justifie le palier observé sur la distribution des états du bassin puisque L n'est presque jamais plein en fin de pas de temps⁵. Afin de disposer d'un indicateur pertinent de la réaction du bassin, seuls les deux réservoirs U et Z sont utilisés comme descripteurs de l'allure de la distribution de λ (figure V.14 (b)). Les distributions obtenues sont assez similaires si ce n'est que le palier n'est plus présent.

$$\lambda = 1 - \frac{U + Z}{U_{max} + Z_{max}} \quad (\text{V.15})$$

5. Une valeur de L au milieu du pas de temps réglerait certainement ce problème.

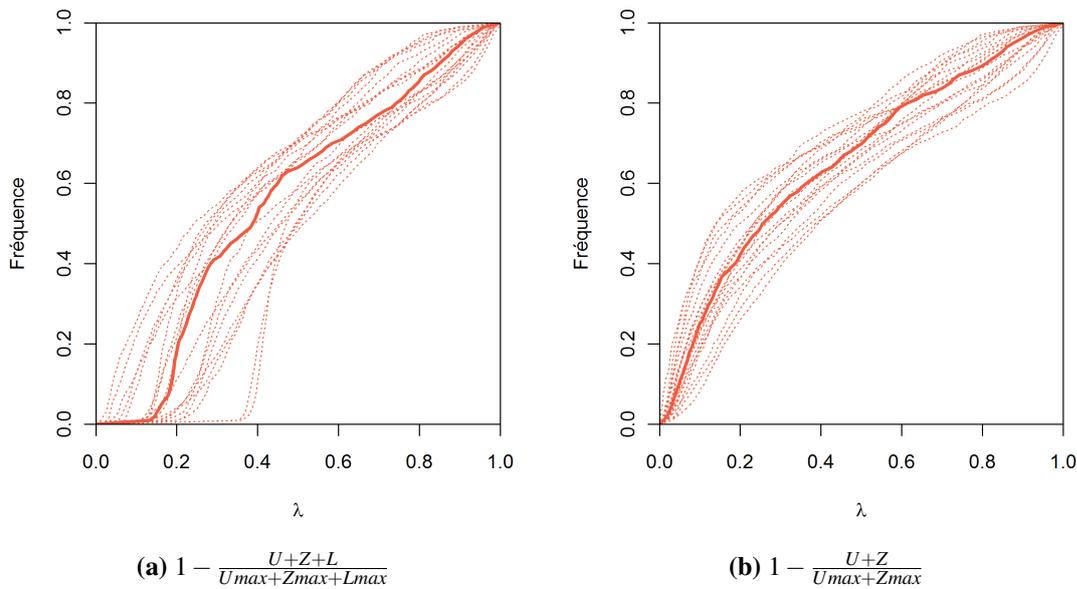


FIGURE V.14 – Distributions des états de saturation λ via MORDOR pour les bassins pluviaux selon les stocks MORDOR considérés.

Les graphes de la figure V.15 montrent les distributions des états de saturation pour les bassins de la base selon leur régime hydrologique (courbes en pointillés). Les distributions médianes pour chaque régime sont tracées en trait plein. Les bassins pluviaux et cévenols ont une distribution du paramètre λ bien définie et peu variable d'un bassin à l'autre. De même, les bassins nivaux et nivo-glaciaires ont également des distributions au comportement assez semblable. Nous notons cependant trois bassins nivo-glaciaires pour lesquels la distribution des λ est plus proche des bassins pluviaux que nivaux (l'Arc à Saussaz, le Ferrand à Mizoën et la Romanche à l'amont du Chambon). Concernant les bassins nivo-pluviaux, la remarque est la même que celle faite pour les distributions de contribution nivale. Les bassins du Tech à Reynes, de la Tet à Vinça et du Verdon à Sainte-Croix semblent être plus proches d'un régime pluvial alors que les bassins de l'Arly à Moulin Ravier, le Gelon à Rochette et la Touyre à Montferrier sont plus proches d'un régime nival au vu des distributions des états de saturation du bassin. Un travail de redéfinition des classes de régimes semble donc nécessaire, par exemple en réalisant une classification sur la base des distributions des états de saturation λ .

Notons que les distributions médianes de chaque régime mettent en évidence le fait que les bassins nivaux ont une probabilité beaucoup plus faible que les bassins pluviaux d'être très sec (i.e. λ proche de 1). Une distribution médiane par régime hydrologique de l'état des stocks (classe des bassins nivo-pluviaux mise à part) est un moyen simple, mais pertinent pour mettre en avant des différences dans le comportement des bassins.

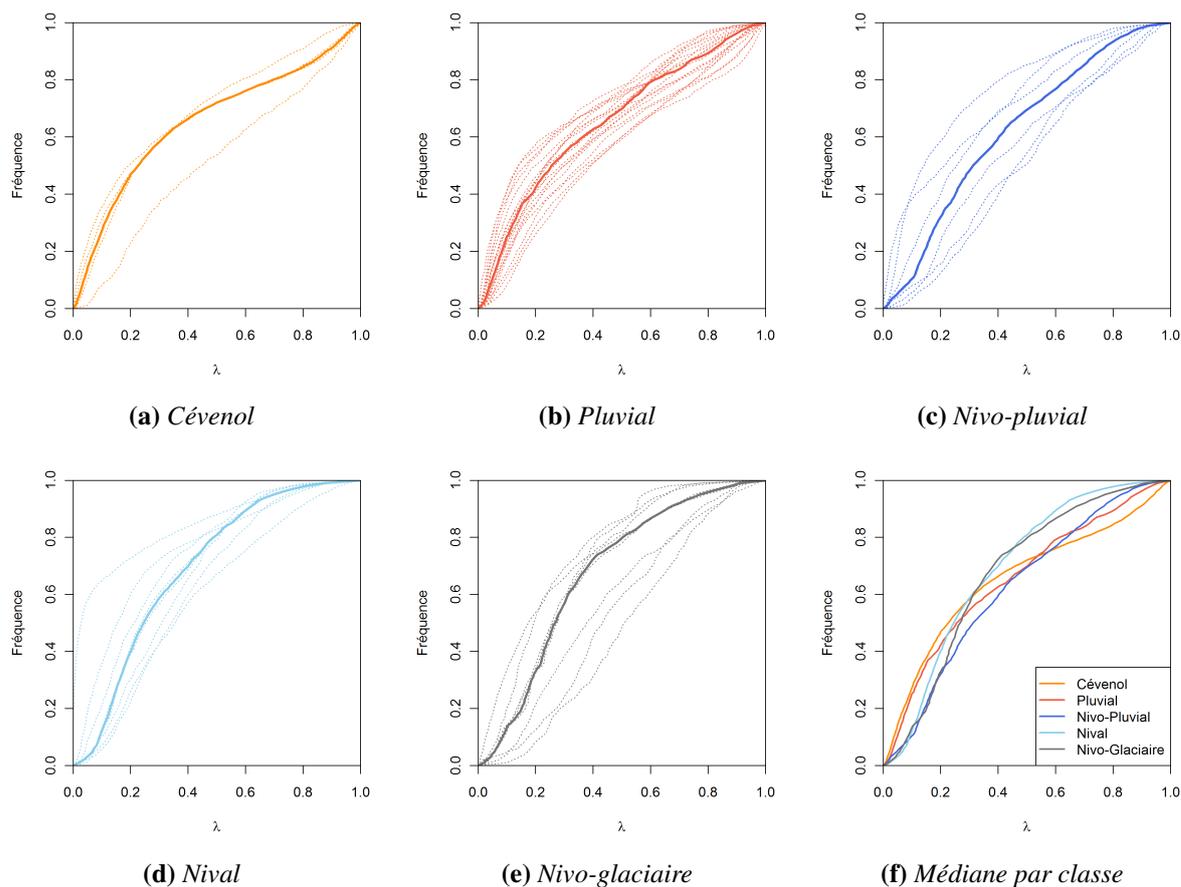
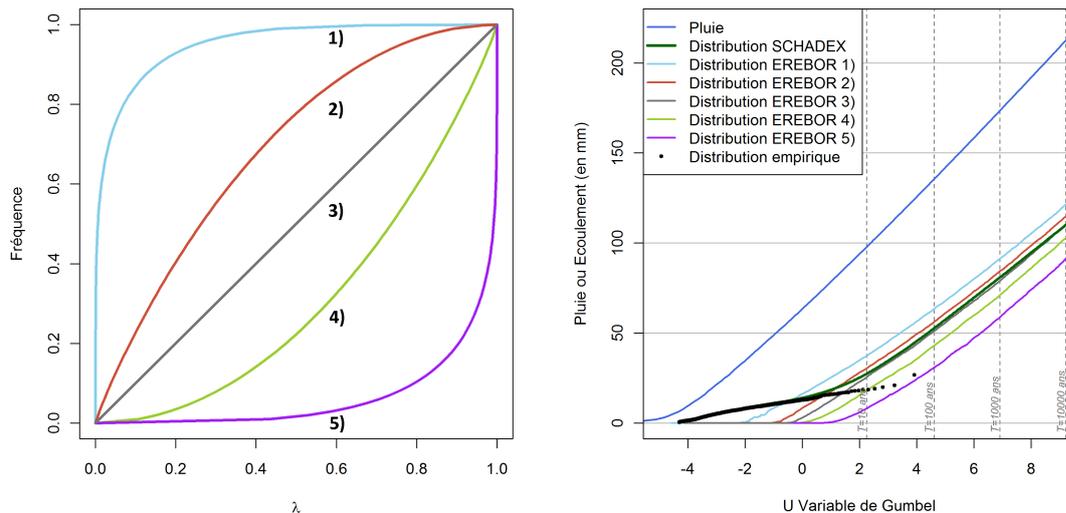


FIGURE V.15 – *Distributions des états de saturation λ par régime hydrologique.*

Influence sur la distribution des écoulements

La figure V.16 permet d'évaluer la sensibilité de la distribution EREBOR à la distribution des états de saturation choisie. Le graphe (a) représente cinq formes de fonctions de répartition potentielles de λ . La courbe 3 correspond à une répartition uniforme des états de saturation du bassin. Les courbes 1 et 2 fournissent une distribution des λ où la probabilité d'avoir de faibles valeurs (i.e. un bassin plutôt saturé) est plus probable. Au contraire, les courbes 4 et 5 ont des probabilités plus fortes pour des λ élevés (i.e. bassin fréquemment très sec). Le graphe (b) expose les différentes distributions d'écoulement obtenues pour les différents cas (tous les autres paramètres étant fixés). Le rendement asymptotique des différentes distributions est peu influencé par le choix de la fonction de répartition des λ , mais il influence la position du point de décollement (i.e. point de la distribution à partir duquel la distribution des débits suit la distribution des pluies).

(a) Distributions des états de saturations λ fixées

(b) Distributions d'écoulements obtenues

FIGURE V.16 – Différentes distributions des états de saturation λ et distributions EREBOR associées. Cas de la Romanche à Champeau (Nivo-glaciaire, 998 km²)

V. 3. 1.4 Les capacités maximales de rétention

Le quatrième et dernier paramètre, S_{max} , vise à décrire la capacité maximale du bassin à absorber le volume de précipitations. L'objectif final est de réussir à reproduire le comportement des distributions obtenues avec la simulation SCHADEX. Chacun des autres paramètres a été contraint par le régime hydrologique du bassin. Les compromis faits pour fixer les valeurs de chaque paramètre ont pour objectif principal de « dépolluer » le paramètre S_{max} d'une grande partie des processus bien distincts sur les bassins :

- la distribution des contributions neige permet pour le nival d'être un minimum représentatif sur le bas de la distribution, sans quoi le paramètre S_{max} aurait tendance à avoir des valeurs très grandes pour expliquer un décollage tardif de la distribution de débits des bassins nivaux ;
- l'hydrogramme unitaire explique les mécanismes de transfert et l'étalement des volumes de ruissellement jusqu'à l'exutoire des bassins ;
- la distribution des états de saturation permet d'ajuster les points de décollage de la distribution d'un régime à un autre.

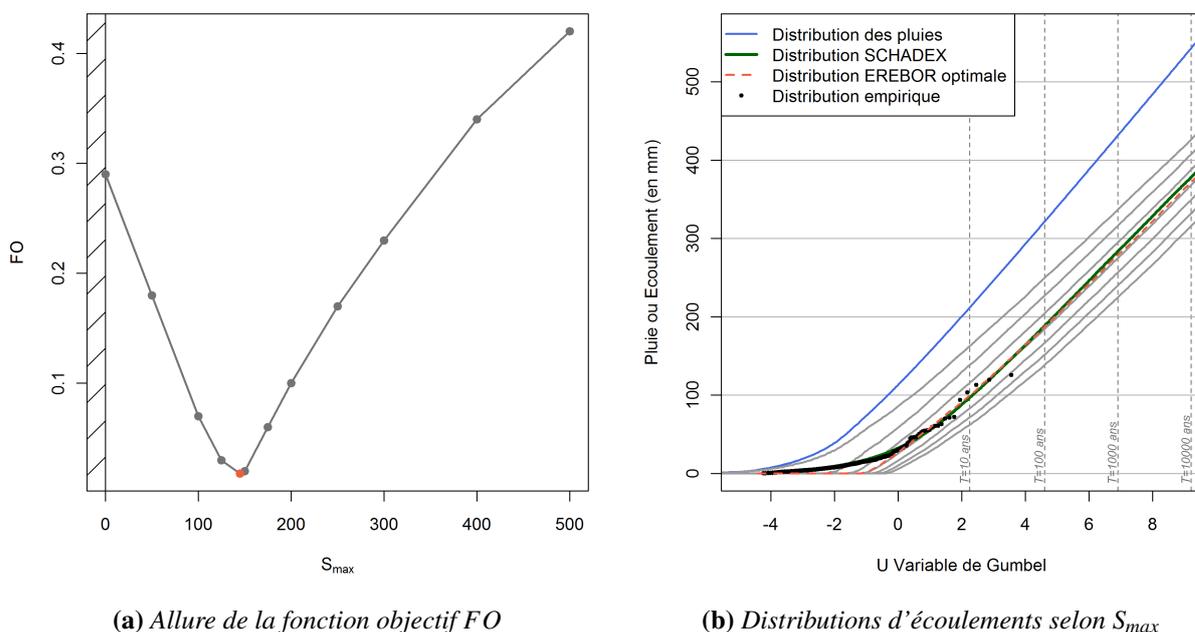
Finalement, en ayant procédé ainsi, nous pouvons espérer avoir un paramètre S_{max} plus représentatif des capacités d'interception superficielles du bassin et donc avec une cohérence accrue entre bassins.

Le paramètre S_{max} doit être défini sur chaque bassin de la base afin de représenter au mieux les quantiles de débits estimés avec SCHADEX. Le paramètre est optimisé pour chaque bassin de façon à minimiser la fonction objectif FO (formule (V.16)). La fonction FO fournit une moyenne des écarts relatifs entre les quantiles d'écoulement E avec la simulation EREBOR et la simulation SCHADEX pour des temps de retour T . Les temps de retour choisis pour la fonction objectif permettent d'évaluer la capacité de la distribution EREBOR à reproduire correctement l'asymptote de la distribution de SCHADEX (à partir du temps de retour 10 ans). Tous les paramètres ayant été fixés, excepté S_{max} , la fonction objectif ne dépend que de ce dernier. La valeur optimale de S_{max} est donc la valeur pour laquelle FO est minimum (figure V.17).

$$FO(S_{max}) = \frac{1}{5} \cdot \sum_T \frac{|E_{EREBOR}(T, S_{max}) - E_{SCHADEX}(T)|}{E_{SCHADEX}(T)} \tag{V.16}$$

avec $T \in \{10, 100, 1000, 5000, 10000\}$

La fonction FO admet toujours un minimum unique en un S_{max} bien déterminé (figure V.17 (a)). Il est donc aisé de définir pour chaque étude de la base la valeur du paramètre S_{max} permettant de reproduire au mieux la distribution obtenue avec la méthode SCHADEX. Sur la figure V.17 (b), les capacités maximales de rétention sont directement liées au décalage entre la distribution pluie et la distribution des écoulements c'est-à-dire un déficit moyen observé à partir d'un certain niveau de précipitation. Cette valeur de déficit moyen dépend du paramètre S_{max} , mais également de la distribution des états de saturation choisie.



(a) Allure de la fonction objectif FO

(b) Distributions d'écoulements selon S_{max}

FIGURE V.17 – Illustration de l'optimisation due S_{max} sur le cas de l'Agout à Fraïsse (Pluvial, 45 km²).

Quelques remarques sur les capacités maximales de rétentions obtenues

Une première vérification consiste à s'assurer qu'une simulation EREBOR est apte à produire une distribution des débits journaliers similaire à la distribution SCHADEX réalisée sur le bassin. Les paramètres de simulation EREBOR sont tous fixés selon le régime du bassin sauf la capacité maximale de rétention S_{max} qui est optimisée à l'aide de la fonction objectif FO . Nous disposons à présent de 45 capacités maximales de rétention S_{max} (une pour chaque étude SCHADEX). Les graphes de la figure V.18 permettent de comparer les écoulements journaliers à un temps de retour T issus de la distribution de l'étude SCHADEX à ceux issus de la distribution EREBOR. Le paramètre S_{max} optimal du point de vue de la fonction objectif FO fournit une estimation des forts quantiles d'écoulements similaires à ceux de la simulation SCHADEX.

Dans un second temps, nous pouvons également nous interroger sur l'information contenue dans le paramètre S_{max} . Nous pouvons par exemple s'assurer de son indépendance au risque pluie. Dans la figure V.19 (a), les valeurs de S_{max} ne sont absolument pas corrélées aux gradex des pluies. Par contre, les capacités maximales de rétention portent un sens hydrologique. Les réservoirs MORDOR U , Z et L – qui représentent la capacité de stockage du bassin pour les

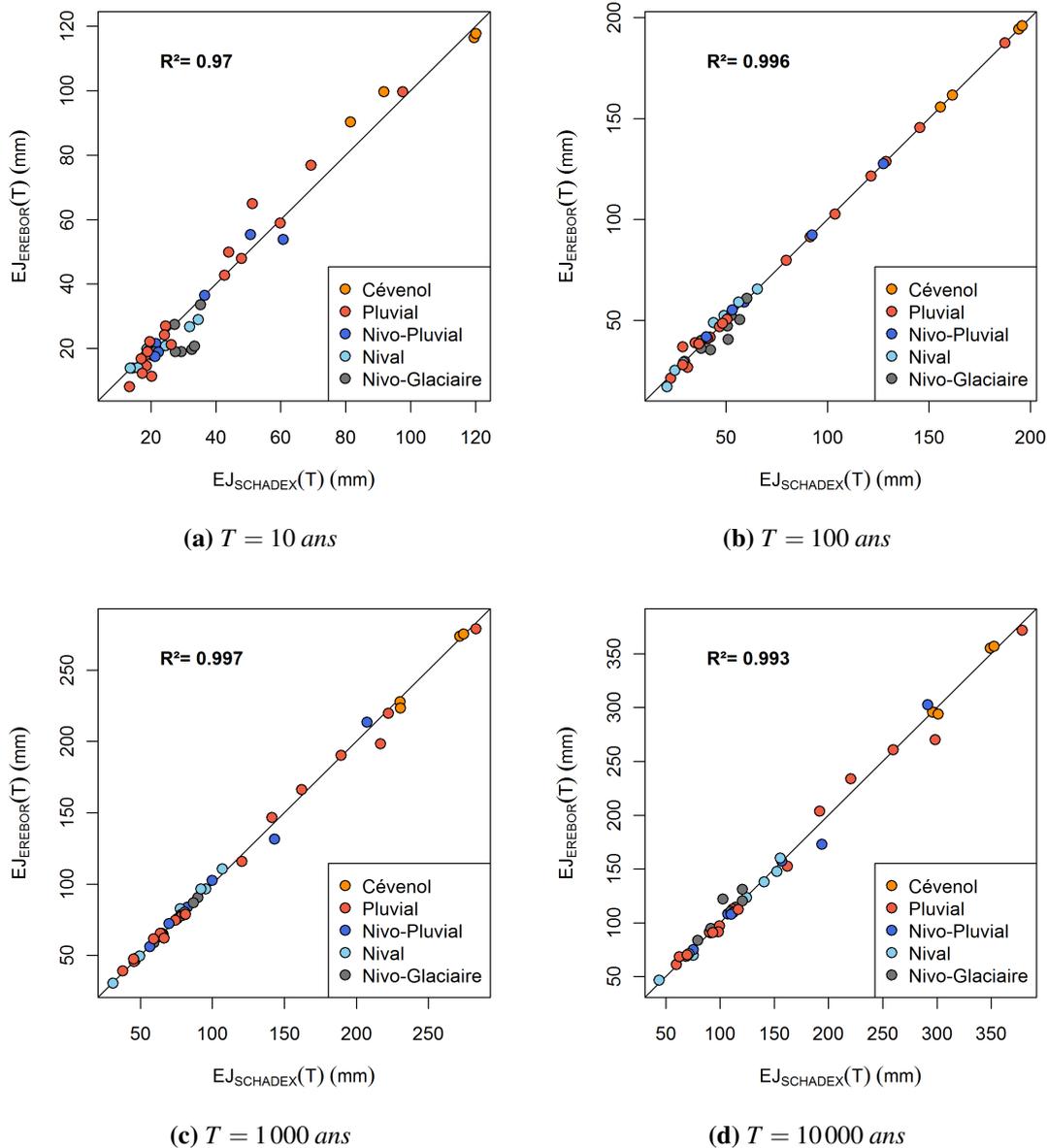


FIGURE V.18 – Comparaison des écoulements journaliers de temps de retour T obtenus avec la simulation SCHADEX et avec la simulation EREBOR (S_{max} optimisé).

processus de surface ou sub-surface⁶ – ont un lien fort avec les valeurs déterminées pour S_{max} (figure V.19 (b)). Deux bassins nivo-glaciaires avec des comportements singuliers sont toutefois notés. L'Arc à Saussaz, avec un paramètre S_{max} de 229 mm, est un bassin nivo-glaciaire avec un rendement asymptotique plutôt faible même pour un bassin glaciaire. Le paramètre S_{max} optimal est ainsi très élevé pour rendre compte de ce comportement particulier. Le bassin de l'Eau d'Olle à Pernière a également un comportement un peu à part avec une valeur de S_{max} faible (129 mm). Sans prendre en compte les bassins nivo-glaciaires, le coefficient de détermination passe de 0.33 à 0.51.

6. La distribution des états de saturation du bassin était déterminée uniquement avec les réservoirs U et Z . Ce choix était en grande partie lié à la dynamique de vidange de L et la façon dont elle est gérée dans le modèle MORDOR. Concernant la capacité de rétention totale du bassin, la capacité des trois réservoirs est prise en compte.

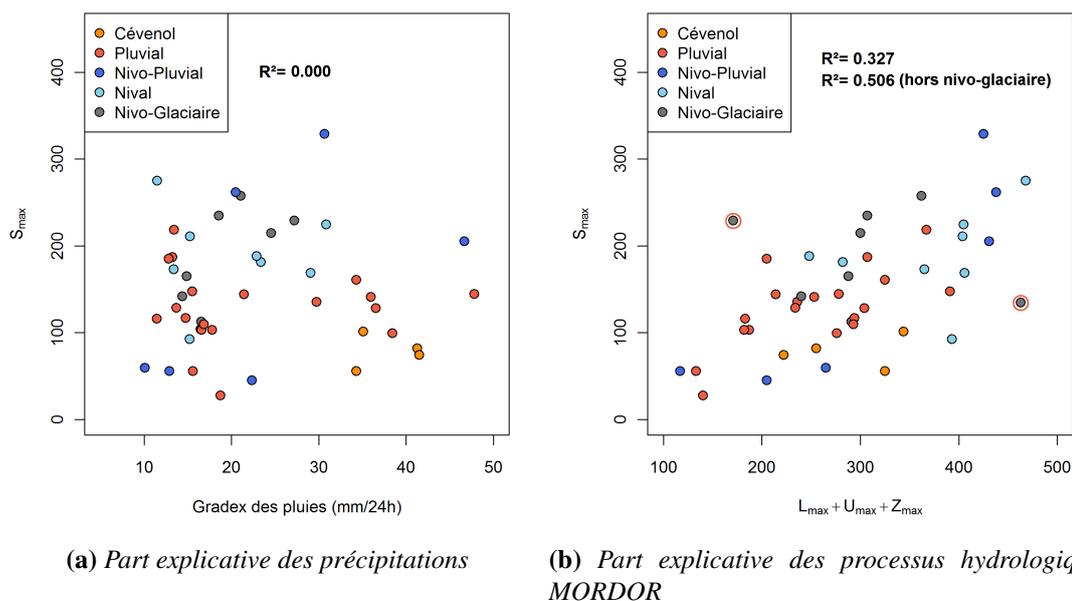


FIGURE V.19 – Sens hydrologique des capacités maximales de rétention obtenues pour la base des 45 études SCHADEX disponibles.

Bilan : les hydrogrammes unitaires HU , la distribution des états de saturation λ et la distribution des contributions nivales ont été fixés selon le régime du bassin versant. Cette régionalisation repose sur le fait que des travaux de régionalisation des régimes hydrologiques ont déjà été réalisés. Sauquet (Sauquet *et al.*, 2008; Sauquet, 2000) a notamment réalisé une cartographie des écoulements mensuels sur toute la France ainsi qu'une cartographie des régimes hydrologiques (figure V.20). Un régime hydrologique peut être déterminé pour tout bassin versant. La forme de l'hydrogramme unitaire et les distributions de contribution nivale et des états de saturation sont alors contraint à partir de ce régime. Finalement, le seul paramètre à déterminer pour chaque bassin est la valeur des pertes initiales maximum S_{max} . Ce travail constitue une première tentative de régionalisation, mais une redéfinition des régimes par rapport aux paramètres EREBOR (et non plus par rapport au débit moyen mensuel) est nécessaire pour améliorer la représentation des différents comportements hydrologiques des bassins.

Nous avons indiqué qu'il existait une valeur optimale de S_{max} permettant de reproduire le comportement asymptotique de la distribution des débits journaliers obtenue avec la méthode SCHADEX. Cette valeur est indépendante du comportement asymptotique de la distribution des pluies journalières ce qui est plus sain pour un modèle de passage pluie-débit visant uniquement à reproduire le comportement du bassin et non le risque pluie. De plus, les capacités maximales de rétention ont un lien fort avec la capacité maximum des stocks MORDOR qui sont eux un descripteur de la réponse hydrologique du bassin versant. **La valeur S_{max} semble donc contenir une information cohérente avec celle portée par MORDOR, même si certains des bassins avec une composante glaciaire semblent plus difficiles à modéliser avec une simulation EREBOR du fait de leur comportement particulier.**

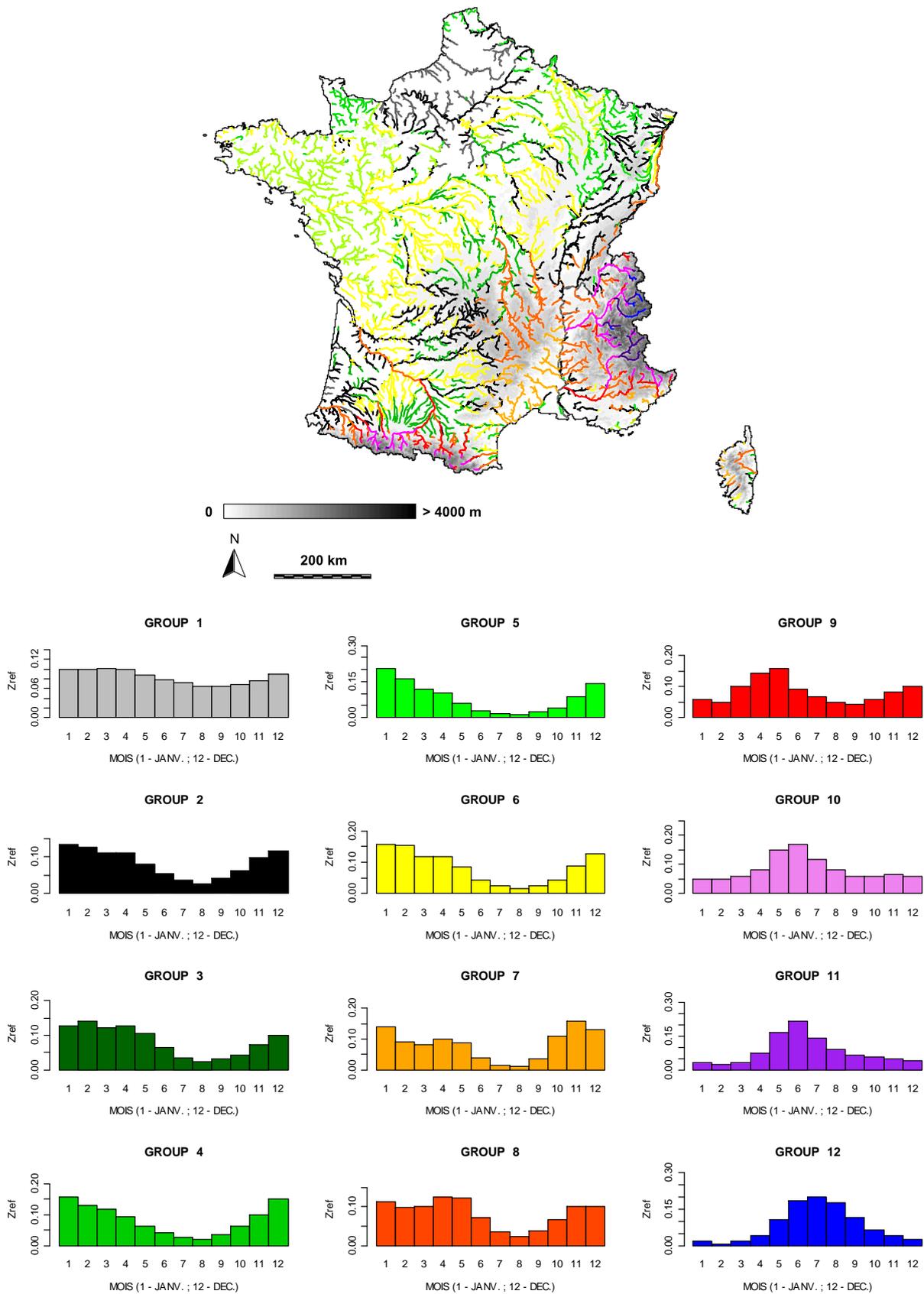


FIGURE V.20 – Cartographie des régimes hydrologiques selon les douze classes établies par Sauquet et al. (2008).

V. 3. 2 Validation croisée de la simulation EREBOR

Une validation croisée de la méthode EREBOR a été réalisée sur la base des 45 études SCHADDEX. Pour chaque bassin, les paramètres sont fixés selon son régime de la manière suivante :

- si le bassin est nival ou nivo-glaciaire, la distribution des contributions nivales est la distribution médiane des bassins nivaux ou nivo-glaciaires établie sur les 44 études restantes ;
- l'hydrogramme unitaire est fixé par le régime hydrologique du bassin ;
- la distribution des états de saturation est fixée à la distribution médiane des états de saturations des bassins de même régime parmi les 44 études restantes ;
- enfin, la valeur de S_{max} est la médiane des S_{max} des bassins de même régime parmi les 44 études restantes.

Faute de disposer d'une meilleure régionalisation pour le S_{max} pour le moment, il est fixé à une valeur selon son régime. La capacité maximale de rétention n'est à priori pas lié au régime, mais plutôt à des descripteurs géomorphologiques. Ce choix, certes imparfait, fournit une estimation régionale de ce paramètre difficilement améliorable sur la seule base des 45 études. Le tableau V.3 témoigne d'une tendance dans la valeur des S_{max} selon le régime. Plus un bassin est pluvial, plus la valeur de S_{max} est faible et plus le décollement de la distribution est situé à des temps de retour faibles.

La figure V.21 montre les écoulements extrêmes estimés avec cette validation croisée avec le régime comme seule information sur le bassin. Les résultats sont acceptables avec de bonnes corrélations entre les quantiles SCHADDEX et EREBOR. En comparant les résultats à ceux de la figure V.18, il apparaît clairement que les performances sont moins bonnes et que la valeur de S_{max} par régime est très schématique. De plus, les quantiles d'écoulements des bassins nivo-pluviaux sont moins bien estimés avec la moitié de la classe qui est systématiquement sous-estimée et l'autre moitié qui est systématiquement sur-estimée. Cette classe, déjà signalée comme peut crédible, devra être redéfinie de façon à mieux représenter les processus hydrologiques des bassins en questions.

Régime hydrologique	S_{max} médian
Cévenol	80 mm
Pluvial	130 mm
Nival	180 mm
Nivo-glaciaire	190 mm

Tableau V.3 – Capacités maximales de rétention médianes par classe de régime hydrologique.

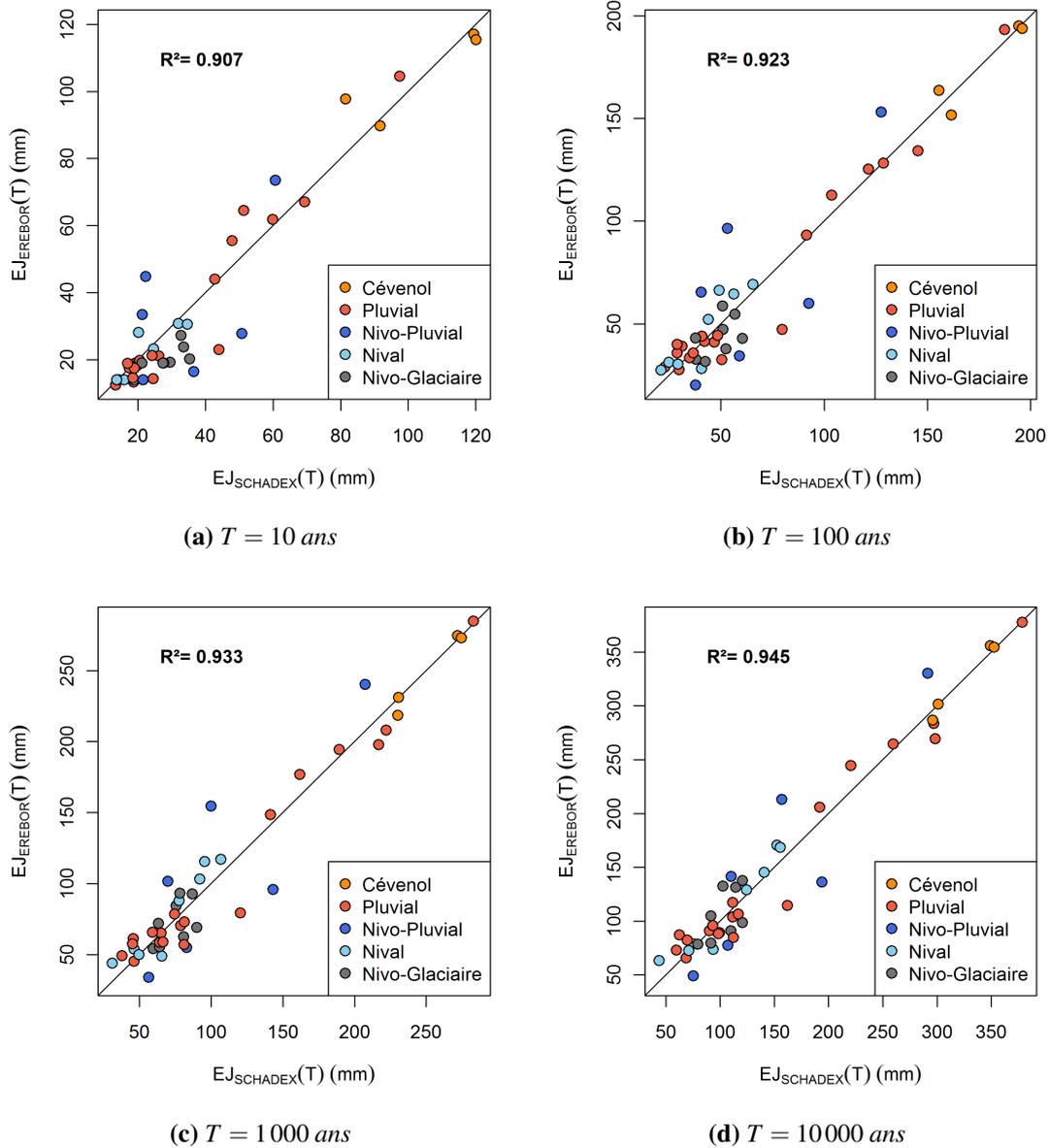


FIGURE V.21 – Comparaison des écoulements journaliers de temps de retour T obtenus avec la simulation SCHADEX et avec la simulation EREBOR (en validation croisée).

Bilan : une première tentative de régionalisation des paramètres de la méthode EREBOR a été mise en place. Elle est pour le moment exclusivement basée sur le régime hydrologique de bassin. Nous avons notamment mis en évidence l'intérêt de chacun des paramètres pour la construction de la distribution des écoulements journaliers. Les valeurs des paramètres ont des comportements assez différents selon le régime hydrologique du bassin. Cependant, pour certains bassins, le régime hydrologique établi sur la dynamique des débits mensuels au cours de l'année ne suffit pas à décrire la dynamique pour les crues. Par exemple, un bassin peut avoir une composante nivale significative et donc visible sur son régime hydrologique, mais les processus hydrologiques intervenant dans la génération des crues peuvent quand à eux être expliqués majoritairement par les précipitations (e.g. Eau d'Olle à Pernière). Plutôt qu'une classification par régimes hydrologiques classiques, des classes de « régimes de crues » pourraient être envisagées afin de mieux décrire le comportement des crues sur des bassins de ce type. Nous pourrions également travailler en allant chercher non plus une classe de régime, mais les n bassins analogues en terme de régime sur lesquels une étude SCHADEX est disponible.

Enfin, un travail complémentaire pour agrandir la base d'étude SCHADEX à disposition est nécessaire. Avec une base de plus grande taille, tout le travail de régionalisation pourrait être refait de façon à définir des régimes plus pertinents pour le risque de crue. La régionalisation des hydrogrammes unitaires, des distributions de contributions nivales et des états de saturation pourra ainsi être affinée. Une régionalisation des capacités maximales de rétention par bassins voisins en terme de géomorphologie et de régime de crue sera également envisageable. En effet, une médiane du paramètre S_{\max} par régime hydrologique fournit une première estimation des forts quantiles de débit cohérente. Il est donc fortement probable qu'une amélioration de la répartition des bassins par régime permette une meilleure estimation régionale du S_{\max} .

Bien que le travail de régionalisation puisse clairement faire l'objet d'améliorations, tout le travail d'exploration a permis de fixer un processus de simulation simplifié permettant de prendre en compte les différents comportements des bassins. Ce modèle est capable de reproduire le comportement asymptotique de SCHADEX en intégrant différents paramètres possédant chacun leur sens hydrologique pour le bassin. Des travaux futurs ne devraient donc pas trop remettre en cause la structure de la méthode, mais plutôt s'intéresser à une amélioration de la régionalisation des différents paramètres par un élargissement de la base des études SCHADEX. Se pose alors le problème de pouvoir faire tourner une simulation SCHADEX sur une grande base de bassins versants et donc d'être capable d'automatiser son application.

V. 4 Automatisation de la simulation SCHADEX

La méthode SCHADEX repose sur des modèles hydrologiques et probabilistes dont le calage sur les données du site d'étude est toujours validé par une expertise humaine. Cette expertise a pour objectif d'assurer un calage des différents modèles en accord avec l'idée que l'on se fait des processus hydrologiques et climatiques sur les bassins versants. Elle assure également une cohérence au sein de l'étude et entre les études similaires (e.g. cohérence régionale entre des bassins voisins ou imbriqués). Cette exigence sur les résultats de simulation est en grande partie due aux forts enjeux industriels qui en découlent pour le dimensionnement des ouvrages.

Vouloir disposer d'une base élargie d'étude SCHADEX nécessite d'aller vers une automatisation de la méthode car une expertise aussi poussée n'est pas possible sur plusieurs centaines d'études. Il faut donc réfléchir à des moyens d'objectiver les choix réalisés en étude et en particulier :

- le choix d'une saisonnalité des précipitations extrêmes ;
- l'ajustement des modèles probabilistes des pluies ;
- le calage du modèle hydrologique MORDOR.

La partie qui va suivre va donc détailler ces différents points. Nous disposons des données pluie, débit et température sur toutes les études SCHADEX réalisées.

V. 4. 1 Sélection des saisons pour les précipitations

Une des étapes primordiales d'une étude SCHADEX est la construction d'une pluie spatiale pour le bassin versant. Nous nous sommes déjà intéressé à la construction de pluies spatiales et à l'estimation des précipitations extrêmes dans le chapitre IV. Cependant, les modèles probabilistes utilisés pour la définition de l'aléa pluie (ajustement MEWP ou tableaux de contingence des pluies adjacentes) reposent sur une pluie spatiale la plus cohérente possible, mais également sur une saisonnalité pertinente de la population des précipitations. En effet, même avec une pluie spatiale idéale (c'est-à-dire une estimation de la pluie spatiale du bassin qui correspond à la pluie spatiale de la réalité), une mauvaise saisonnalité des précipitations peut mélanger différents phénomènes et conduire à une sous-estimation du risque pluie (Allamano *et al.*, 2011). Il est donc important de réfléchir à une méthode de sélection objective et pertinente des saisons pour les précipitations extrêmes.

Algorithme de détection automatique des saisons

L'algorithme présenté figure V.22 explique la méthode proposée pour déterminer les saisons de façon automatique et objective. À l'aide d'une chronique de précipitations (au pas de temps journalier ou autre), les maxima annuels PX_{mois} sont sélectionnés puis la moyenne inter-annuelle⁷ pour chaque mois est réalisée pour obtenir une pluie maximum mensuelle moyenne \overline{PX}_{mois} . L'idée est alors d'essayer de regrouper des mois de risque similaire dans une même saison.

Le regroupement des mois en saison doit respecter certaines règles visant à écarter des problèmes d'échantillonnage ou de sens hydrologique :

- une saison comporte au minimum deux mois. Une saison d'un mois pose des problèmes de taille d'échantillon pour réaliser l'ajustement.
- la taille de chaque saison n'est pas forcément la même. Selon la climatologie des bassins versants étudiés, des saisons peuvent mettre en évidence un risque très marqué sur deux mois comme sur quatre.

7. La valeur médiane a également été testée et fournit des résultats similaires.

- une saison est constituée de mois consécutifs.

Sans a priori sur la durée des saisons, deux périodes sont considérées. Une période automnale (période 2) d'août à décembre et le reste de l'année (période 1) sont définies afin de chercher une saison à risque potentielle dans chacune d'entre elles. Cette séparation vise à faciliter la recherche des saisons en distinguant deux périodes pouvant avoir de fortes précipitations : précipitations automnales et précipitations printanières ou hivernales. Dans chacune des périodes, une fenêtre glissante de deux, trois puis quatre mois maximise la moyenne des \overline{PX}_{mois} (saisons verte et orange). Trois cas (cas 1 à 3) de saisonnalité sont donc définis avec, pour chacun, une saison (de deux à quatre mois) sur chaque période. Les saisons intermédiaires sont alors fixées de façon à faire la jonction entre les deux premières saisons définies (saisons violette et bleue). Si une des saisons intermédiaires ne comporte qu'un seul mois, alors ce mois est intégré à une des saisons adjacentes à ce mois (celle dont la moyenne des \overline{PX}_{mois} est le plus proche du \overline{PX}_{mois} du mois en question). Au final, trois cas de saisonnalité sont envisagés (avec des saisons à risque sur 2, 3 ou 4 mois et avec trois ou quatre saisons chacun).

Le travail consiste à identifier quelle saison représente le mieux les risques pluviométriques. En particulier, savoir quel cas cible le mieux la saison à risque qui portera le comportement asymptotique de la distribution de pluies. Une distribution MEWP est ajustée pour chacun des cas et fournit les distributions saisonnières de chaque saison. Par ailleurs, les douze distributions empiriques de chaque mois sont également connues.

Nous allons alors évaluer si pour un mois i , la distribution saisonnière estimée pour la saison j décrit correctement la distribution empirique du mois i . Un critère KGE ou Kling-Gupta efficiency (Gupta *et al.*, 2009) est utilisé pour comparer les fréquences en non-dépassement empirique F_i et celle de la distribution saisonnière F^j (formule (V.17)). La valeur du KGE croit avec les performances et atteint un maximum de 1. Pour chacun des trois cas, les douze KGE mensuels sont calculés.

$$\left\{ \begin{array}{l} KGE(F_i, F^j) = 1 - \sqrt{(r-1)^2 + (\alpha-1)^2 + (\beta-1)^2} \\ r = Corr(F_i, F^j) \\ \alpha = \frac{\sigma(F_i)}{\sigma(F^j)} \\ \beta = \frac{\overline{F}_i}{\overline{F}^j} \end{array} \right. \quad (V.17)$$

Ensuite, pour chaque cas, deux KGE – un sur la période 1 et un sur la période 2 – sont calculés avec une somme pondérée des KGE mensuels. Les poids donnés à chaque mois sont fonction du risque du mois. Plus un mois est à risque, plus son poids est fort. Pour chaque période, le cas ayant le plus fort KGE est choisi. Les cas sélectionnés (i.e. la saison choisie pour la saison à risque) pour les périodes 1 et 2 ne sont pas nécessairement les mêmes.

$$\left\{ \begin{array}{l} KGE_{période\ 1} = \sum_{i=1}^7 \frac{\overline{PX}_i}{\sum_{k=1}^{12} \overline{PX}_k} \cdot KGE_i \\ KGE_{période\ 2} = \sum_{i=8}^{12} \frac{\overline{PX}_i}{\sum_{k=1}^{12} \overline{PX}_k} \cdot KGE_i \end{array} \right. \quad (V.18)$$

Les saisons intermédiaires entre les deux saisons sélectionnées pour la période 1 et 2 sont alors redéfinies.

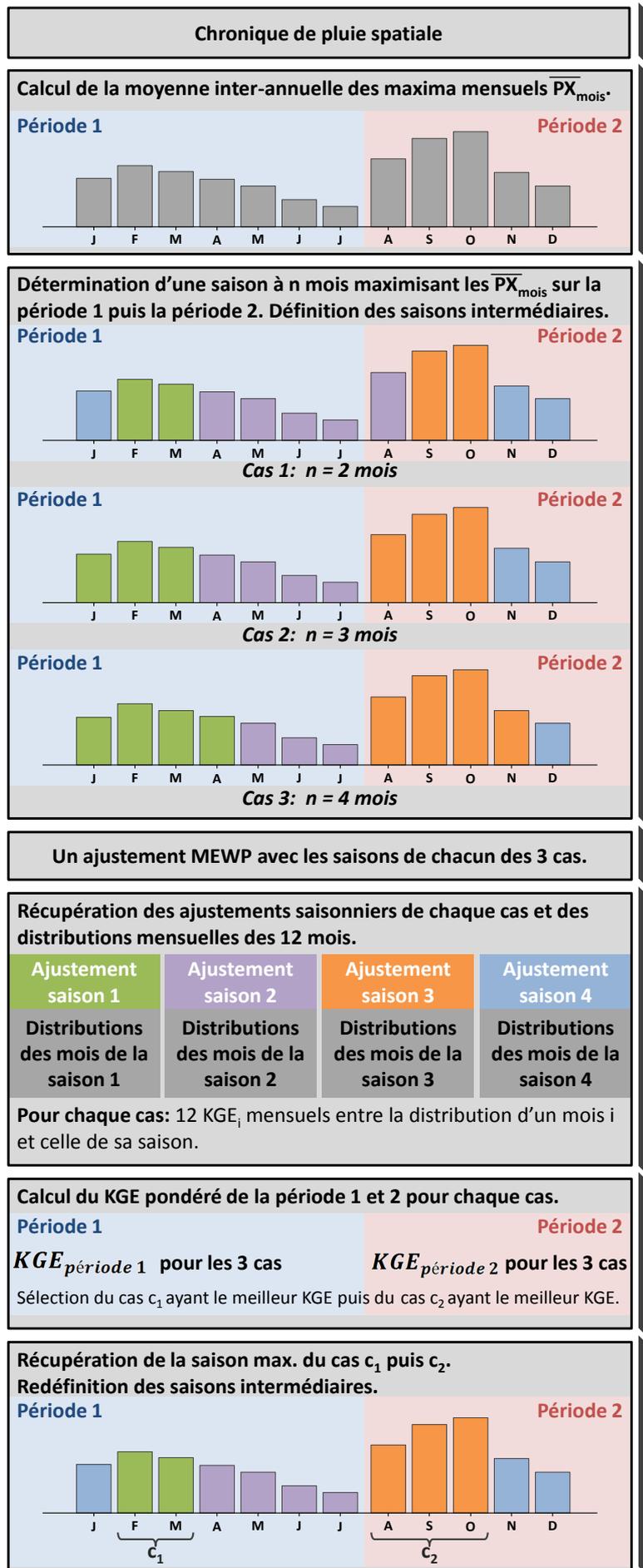


FIGURE V.22 – Schéma de l’algorithme de sélection des saisons.

Évaluation de l'algorithme de choix des saisons

Pour chacune des 45 études SCHADEX, les pluies spatiales, les saisons choisies, ainsi que les ajustements MEWP associés ont été collectés. Les quantiles remarquables de pluies des ajustements MEWP peuvent être très sensibles au choix des saisons et en particulier à celui de la saison à risque.

Une comparaison des saisons fournies par l'algorithme et celles déterminées lors des études d'ingénierie est alors réalisée. Le score de recouvrement f des saisons à risque est défini :

$$f = \frac{\text{nombre de mois communs des saisons à risque}}{\text{nombre de mois de la saison à risque automatique}} \quad (\text{V.19})$$

Le score f indique le pourcentage de mois commun entre la saison à risque automatique et celle de référence pour un bassin. Sa valeur optimale vaut 1 et correspond au cas où tous les mois de la saison à risque automatique appartiennent également à la saison à risque de référence. Il atteint une valeur nulle pour deux saisons disjointes. La moyenne de f sur les 45 études est de 0.86, ce qui signifie que 86% des mois des saisons à risque automatiques sont communs à ceux des saisons expertisées.

La figure V.23 permet de comparer les pluies de temps de retour 10, 100, 1000 et 10 000 ans obtenues avec l'ajustement expertisé et avec l'algorithme de détection automatique des saisons. L'algorithme permet une détection des saisons pertinente et en particulier d'identifier une saison à risque fournissant un comportement asymptotique de la distribution proche de celui obtenu en opérationnel.

La saisonnalité des pluies déterminée par l'algorithme fournit donc un découpage de la population des précipitations pertinent et capable d'appréhender le risque pluie. Les saisons ainsi déterminées sont utilisées pour l'ajustement d'une loi MEWP sur les pluies centrales.

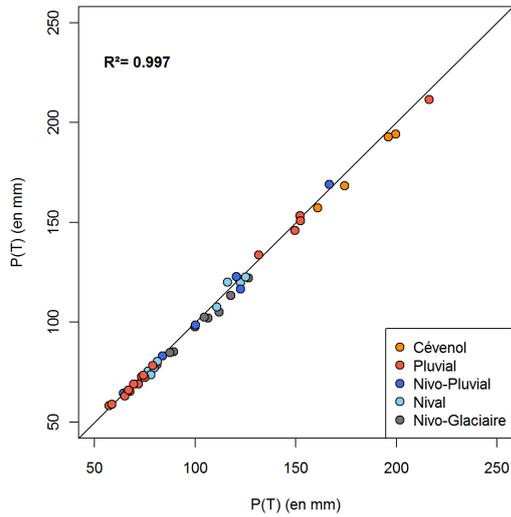
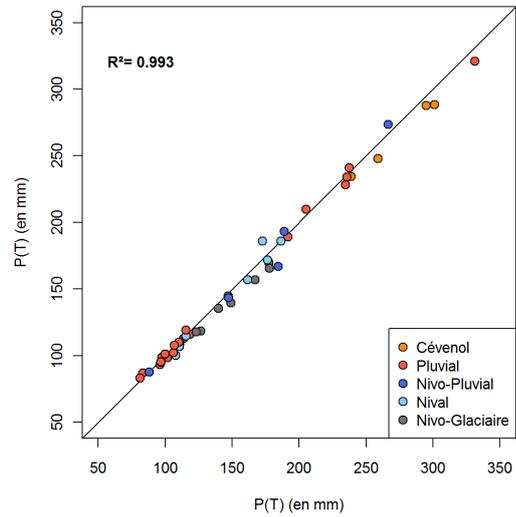
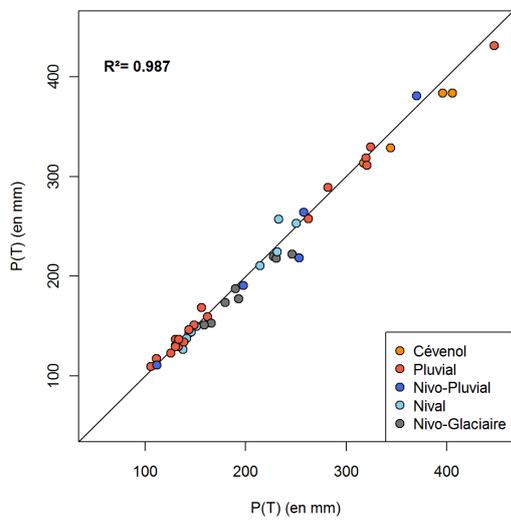
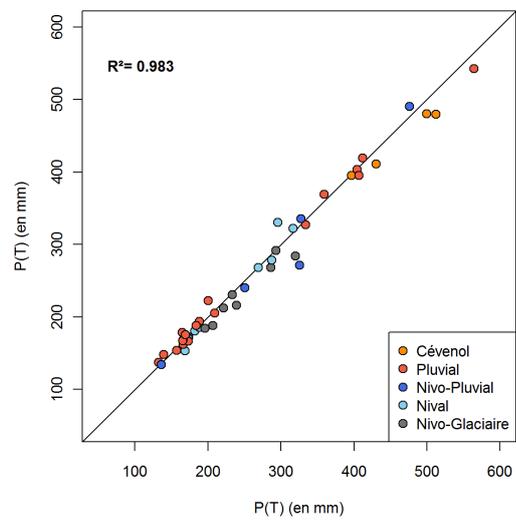
(a) $T = 10 \text{ ans}$ (b) $T = 100 \text{ ans}$ (c) $T = 1000 \text{ ans}$ (d) $T = 10000 \text{ ans}$

FIGURE V.23 – Comparaison des pluies de temps de retour T obtenues avec les saisons expertisées et les saisons automatisées sur la base de 45 études SCHADEX.

V. 4.2 Calage automatique du modèle hydrologique MORDOR

En parallèle de cette étude sur la saisonnalité des pluies, un modèle hydrologique MORDOR est calé sur les données débit, pluie et température du bassin versant. Au sein d'une simulation SCHADEX, le modèle hydrologique MORDOR comporte deux principaux intérêts :

- il permet de représenter l'aléa de saturation du bassin considéré ;
- il transforme plusieurs milliers d'événements de pluie synthétiques en épisodes de crues.

Le modèle hydrologique doit donc être calé de façon à répondre à ces deux objectifs, c'est-à-dire de posséder à la fois de bonnes performances sur le comportement hydrologique moyen du bassin (régime hydrologique) et sur les fortes crues observées sur le bassin. Le modèle hydrologique MORDOR utilisé comporte ou non un modèle pour la gestion du stock de neige⁸ suivant que les processus nivaux sont négligeables ou pas sur le bassin versant. Les paramètres sont calés à l'aide d'un algorithme d'optimisation génétique en utilisant une fonction objective FO_{MORDOR} (formule (V.20)). La fonction objective utilisée est un critère composite construit pour obtenir des modèles polyvalents, aussi bien capables de reproduire les régimes de débits que les distributions mensuelles de débits consécutifs à un épisode de pluie (voir page 173). Ce critère se décompose en trois parties afin d'intégrer une composante temporelle, saisonnière et fréquentielle (Lamb, 1999; Garavaglia, 2011b) :

$$FO_{MORDOR} = \underbrace{KGE(Q, \hat{Q})}_{\text{Temporel}} + \underbrace{KGE(Q_{reg}, \hat{Q}_{reg})}_{\text{Saisonnier}} + \underbrace{KGE(Q_{P_c, mensuel}, \hat{Q}_{P_c, mensuel})}_{\text{Fréquentiel}} \quad (\text{V.20})$$

avec Q les débits observés et \hat{Q} les débits simulés,

avec Q_{reg} (et \hat{Q}_{reg}) le régime interannuel des débits observés (et simulés),

avec $Q_{P_c, mensuel}$ (et $\hat{Q}_{P_c, mensuel}$) l'échantillon mensuel des débits consécutifs à un épisode de pluie sur les débits observés (et simulés),

et $KGE(Q_{P_c, mensuel}, \hat{Q}_{P_c, mensuel})$ la somme pondérée des KGE mensuels pondérés par le débit moyen de chaque mois, pour les écoulements suivant une pluie centrale.

La démarche qui va être présentée a vocation à analyser la sensibilité des résultats de simulation SCHADEX au calage du modèle MORDOR et de réfléchir à une façon d'objectiver le choix d'un modèle pour la réalisation de la simulation. Pour cela, nous pouvons nous inspirer du travail réalisé par Brigode (Brigode, 2013; Brigode *et al.*, 2014). Les modèles probabilistes des pluies sont considérés comme acquis et nous n'envisageons donc pas de test de sensibilité en rapport à ceux-ci. Concernant la sensibilité aux modèles hydrologiques, Brigode a montré que la considération de plusieurs périodes de calage pour le modèle hydrologique, et donc de plusieurs jeux de paramètres de ce modèle, génère une importante dispersion des estimations de débits extrêmes. Des tentatives de définition d'indicateurs pour caractériser les différentes périodes de calage et expliquer la dispersion ont été plutôt infructueuses. Néanmoins, la présence ou non de fortes crues dans les différentes sous-périodes de calage semble jouer un rôle prépondérant dans la dispersion obtenue en terme d'estimation des débits extrêmes. Enfin, la variabilité de l'aléa de saturation du bassin versant à la période utilisée lors de la simulation semi-continue de SCHADEX a également été envisagée dans le travail de Brigode, qui a conclu que son impact sur l'estimation des débits extrêmes était non significatif. Finalement, le point crucial qui a été mis en évidence est le rôle prépondérant des horsains (i.e. les épisodes difficilement comparables aux autres épisodes observés de par leur intensité) au sein d'une période de calage des modèles hydrologiques. Les estimations de crues extrêmes réalisées sur un bassin versant avec la présence de horsains semblent alors être totalement dépendantes de la prise en compte ou non de ces épisodes dans la période de calage du modèle. Dans une étude, la prise

8. Le nombre de paramètres du modèle passe ainsi de 12 à 23 avec l'ajout du module neige de MORDOR.

en compte ou non de ces événements est jugée par le chargé d'étude selon le crédit qu'il donne à ce horsain.

Dans le cadre de la démarche d'automatisation de la méthode SCHADEX, nous nous retrouvons confronté au problème du choix d'un modèle hydrologique MORDOR fiable pour réaliser une estimation des extrêmes. Une étude par boot-strap indépendants permet de caler le modèle pluie-débit MORDOR sur 100 sous-périodes différentes et de fournir 100 jeux de paramètres. La procédure de boot-strap est souvent utilisée en hydrologie pour l'analyse des incertitudes et la caractérisation de la variabilité. Plus récemment, elle a été envisagée pour caler des modèles pluie-débit, notamment au travers des travaux de Ebtehaj *et al.* (2010) et Selle et Hannah (2010) qui ont montré qu'elle permettait une meilleure identification des paramètres de modèle.

Les 45 études de références de la base sont des études expertisées qui utilisent la période la plus longue possible pour caler le modèle pluie-débit MORDOR (période sur laquelle on dispose de données validées après la critique). Le boot-strap sans remise sur des sous-périodes de cette période de calage fournit un foisonnement de jeux de paramètres pouvant être considéré comme représentatif de tous les modèles envisageables pour le bassin considéré. La question qui se pose est alors de savoir quel modèle choisir pour une estimation des débits extrêmes et avec quels indicateurs.

Pour chacune des études, 100 périodes de n années de calage sont tirées aléatoirement avec un tirage sans remise parmi les années de disponibilités des données. Le modèle MORDOR est alors calé selon la fonction objectif (V.20) pour chacune des périodes afin de disposer de 100 jeux de paramètres (i.e. 100 modèles) possibles par bassin. Des simulations SCHADEX sont alors réalisées avec chaque modèle, et produisent 100 distributions de débits (figure V.24). La période de la simulation SCHADEX (période sur laquelle est réalisée la simulation semi-continue) n'a pas d'influence notable sur les résultats de simulation et est donc définie par la période de disponibilité de la pluie et température pour toutes les études.

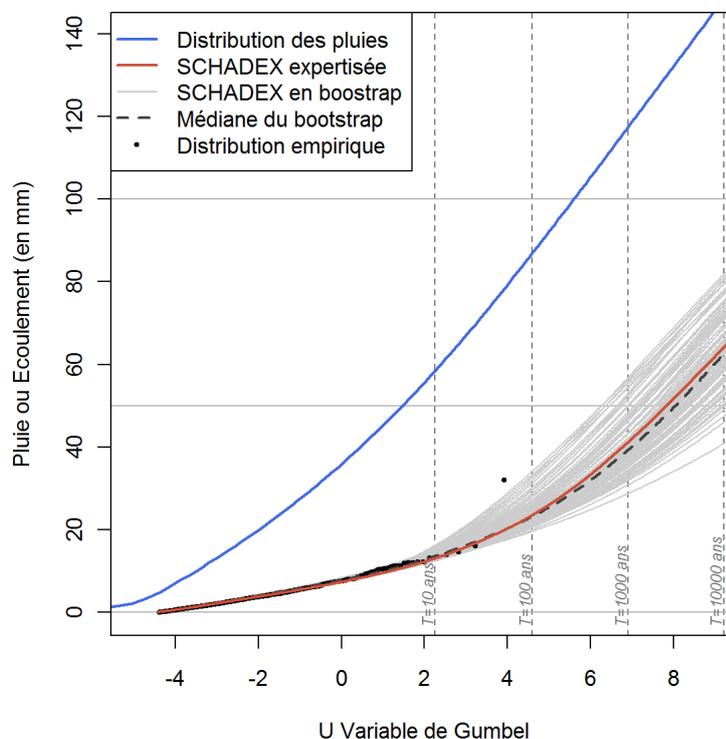


FIGURE V.24 – Les distributions de débits journaliers obtenues avec la méthode SCHADEX sur le bassin de la Creuse à Fresseline (1238 km²).

Le faisceau de courbes grises représente les 100 distributions SCHADEX associées aux 100 modèles du boot-strap. Lorsque le faisceau est très resserré, le modèle est bien défini. Une forte dispersion du faisceau met en évidence un manque de robustesse dans l'identification d'un modèle pour le bassin. Le choix réalisé pour le modèle est alors un facteur clé du comportement en extrapolation des distributions de débit. Tout le problème réside dans le choix du modèle permettant d'avoir un comportement asymptotique de la distribution des débits cohérent avec celui des études SCHADEX expertisées de référence.

Critères de sélection du modèle hydrologique pour la simulation SCHADEX

Afin de simplifier la mise en place du principe de sélection du modèle MORDOR, différents termes vont être distingués. La distribution SCHADEX de référence se rapportera à la distribution des débits obtenue par la méthode SCHADEX expertisée. Les distributions de débits obtenues par simulation SCHADEX avec les 100 modèles MORDOR sur les différentes périodes de calage du boot-strap seront appelées distributions « boot-strap » pour simplifier. Enfin, le modèle choisi pour réaliser une simulation SCHADEX pertinente du point de vue de la distribution de référence sera dénommé modèle « candidat ».

Le modèle utilisé pour la distribution de référence est souvent un modèle hydrologique représentant au mieux les fortes crues observées sur le bassin tout en ayant une performance globale correcte. De plus, les événements les plus forts font souvent l'objet d'une attention particulière et la distribution finale des débits doit prendre en compte ces forts événements de façon à ne pas estimer une probabilité irréaliste d'une crue observée sur le bassin versant. La méthode proposée pour déterminer un modèle candidat est donc de s'appuyer sur cette démarche et d'essayer de reproduire ces règles de décisions de façon automatique.

La sélection du modèle candidat s'organise alors par écrémage progressif des 100 modèles du boot-strap et des distributions boot-strap associées. Les différentes étapes sont détaillées ci-dessous.

étape 1 : Récupération des 100 modèles MORDOR boot-strap, des scores en calage et en validation et des 100 distributions boot-strap.

Pour chaque modèle, les différents scores sont estimés en calage et en validation. Le schéma V.25 représente les différentes périodes sur lesquelles les scores sont calculés. Les scores calculés sur la période où le modèle a été calé sont des scores en calage. Les scores en calage décrivent les performances du modèle à s'ajuster sur les débits de la période de calage. Par analogie, des scores en validation, calculés sur une période n'ayant pas servi au calage du modèle, évaluent la capacité prédictive du modèle c'est-à-dire la performance avec laquelle il représente des débits sur lesquels il n'a pas été calibré.

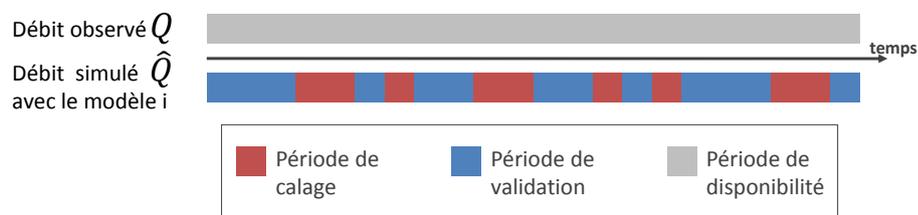


FIGURE V.25 – Période de calage et de validation d'un modèle i parmi les 100 modèles du boot-strap.

Les scores sont calculés pour deux critères :

- Les scores KGE sur la représentativité du régime interannuel des débits en calage ($KGE_{Q_{reg}}^C$) et en validation ($KGE_{Q_{reg}}^V$).
- Les scores KGE sur les distributions mensuelles pondérées de débits consécutifs à un épisode de pluie en calage ($KGE_{Q_{mensuel}}^C$) et en validation ($KGE_{Q_{mensuel}}^V$).

étape 2 : Recherche de faisceaux distincts de distributions.

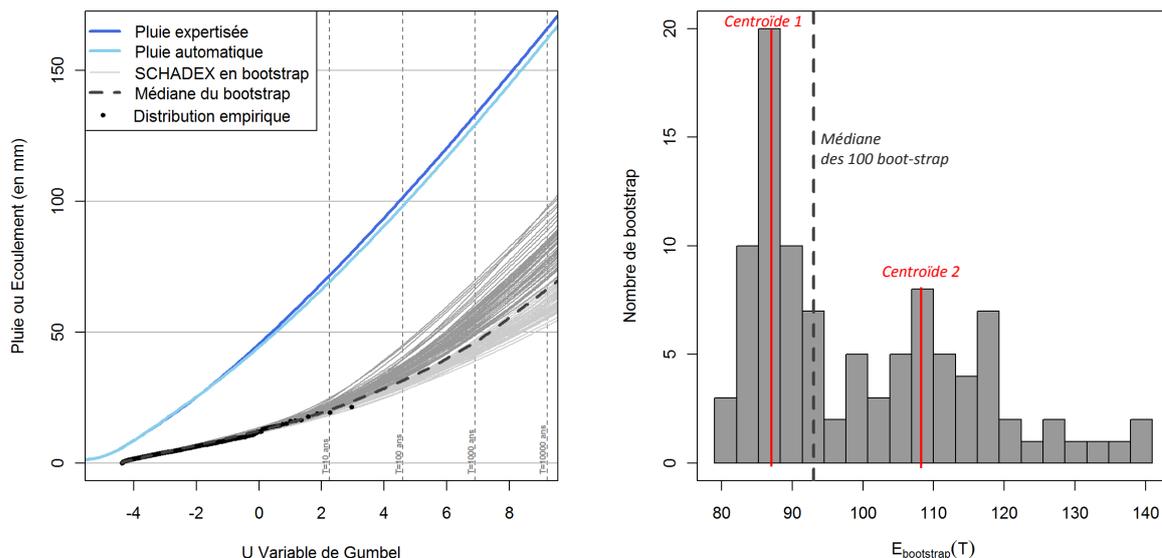
Il est possible que les réalisations du boot-strap construisent un faisceau bimodal. Ce doublement est particulièrement visible sur la densité de probabilité des quantiles de débits qui comporte alors deux maxima (figure V.26 (b)). Deux raisons principales peuvent expliquer la présence de deux sous-populations au comportement distinct. La présence d'événements, type horsain, peut créer deux groupes de distributions selon que ces valeurs sont présentes ou non dans la période de calage. Deuxièmement, des ruptures dans les séries hydro-climatiques peuvent être à l'origine de ce doublement, chacun des faisceaux peut alors être représentatif de risque du bassin avant et après la rupture. Dans le souci de reproduire les choix faits en ingénierie, nous voulons privilégier un modèle candidat prenant en compte les éventuels horsains afin de ne pas sous-estimer la probabilité des plus forts événements observés sur le bassin. C'est pourquoi une première étape consiste à identifier la présence éventuelle de deux sous-populations de distributions boot-strap et de ne sélectionner que le faisceau constitué du risque le plus fort. La densité de probabilité des écoulements millénaux du boot-strap sert à identifier la présence ou non de deux maxima (figure V.26 (b)). Les valeurs d'écoulement des deux maxima sont alors récupérées s'ils existent. Les deux valeurs constituent le centroïde de chaque classe de faisceau. Chaque distribution est alors associée à un des deux faisceaux en fonction de la valeur de son écoulement millénal et de sa proximité aux centroïdes des deux classes (figure V.26 (a)). Les modèles constituant le faisceau avec des risques les moins forts sont écartés, car ils représentent une modélisation du bassin ne s'appuyant pas sur les fortes crues observées (distributions en gris clairs). Tous les autres modèles sont conservés comme modèles candidats potentiels (distributions en gris foncés). Si le boot-strap génère un faisceau de distribution unique alors tous modèles sont conservés.

étape 3 : Recherche des 10 meilleurs modèles selon les critères $KGE_{Q_{mensuel}}^C$ et $KGE_{Q_{mensuel}}^V$.

Pour les modèles gardés, les scores $KGE_{Q_{mensuel}}$ sur les distributions mensuelles sont récupérés en calage et validation. Chaque modèle est alors placé dans l'espace à deux dimensions de ces scores (figure V.27). Sur cette figure, les petits points représentent les modèles écartés par l'étape 2 et les gros points ceux conservés. Il est possible que les modèles écartés soient des modèles dont les performances en calage et en validation sont très satisfaisantes et même meilleures que celles des modèles conservés. Le choix réalisé dans l'étape 2 discrimine les modèles représentant mal les quelques fortes valeurs observées, et ce même si tout le reste de la distribution est bien représenté (donc avec des KGE forts). Finalement, les distributions conservées sont plutôt des distributions capables de bien reproduire le risque des plus fortes valeurs, au détriment des valeurs courantes, ce qui est très pénalisant du point de vue du critère $KGE_{Q_{mensuel}}$. L'utilisation d'un critère fréquence-fréquence FF (page 95) comparant la probabilité empirique de la plus forte valeur d'un échantillon à celle donnée par la distribution du boot-strap serait sûrement un critère plus pertinent, bien que moins robuste, pour évaluer la capacité des distributions à bien prendre en compte les fortes crues observées sur le bassin.

Parmi tous les modèles restants, seuls ceux qui sont les plus performants à la fois en calage et en validation sont recherchés c'est-à-dire les modèles situés sur le haut du nuage⁹. Les

9. Cette sélection des points se rapproche de la recherche d'un front de Pareto sur deux critères.



(a) Distributions boot-strap.

(b) Histogramme des débits millénaux des distributions boot-strap.

FIGURE V.26 – Mise en évidence de deux groupes de distributions aux comportements asymptotiques distinctes pour la Maulde à Peyrat Le Château (174 km²).

10 modèles minimisant le critère C_1 (formule (V.21)) sont conservés (représentés par des points bleus sur la figure V.27). Le critère C_1 représente une distance euclidienne à l'optimum (1; 1). Cette sélection garantit un choix de modèle représentant le mieux possible les distributions de débits de chaque mois.

$$C_1 = \sqrt{\left(1 - KGE_{Q_{\text{mensuel}}}^C\right)^2 + \left(1 - KGE_{Q_{\text{mensuel}}}^V\right)^2} \quad (\text{V.21})$$

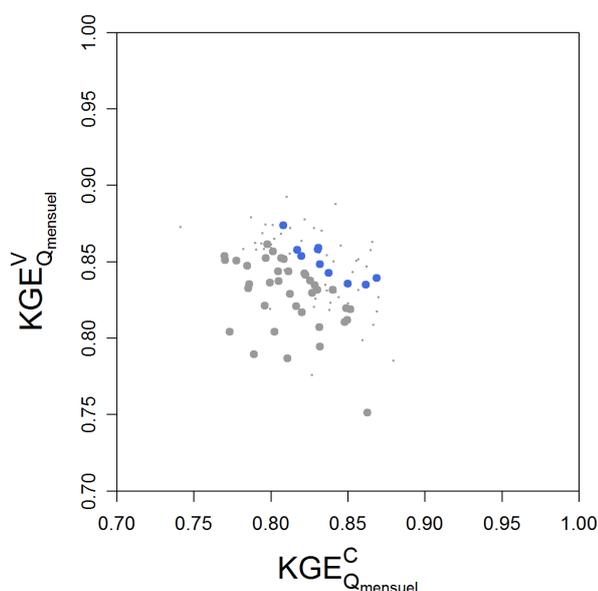


FIGURE V.27 – Espace à 2 dimensions pour identifier les meilleurs modèles selon $KGE_{Q_{\text{mensuel}}}$.

étape 4 : Sélection du modèle candidat parmi les 10 modèles restants qui représente le mieux le régime des débits.

La dernière étape consiste à choisir le modèle candidat parmi les 10 modèles restant à la fin de l'étape 3. Le choix du modèle est fait avec le même principe que dans l'étape 3, mais en s'intéressant cette fois à la représentation du régime avec les scores $KGE_{Q_{reg}}^C$ et $KGE_{Q_{reg}}^V$. Les 10 modèles ont été sélectionnés de façon à être performants du point de vue de leur comportement pour les extrêmes. Nous notons cependant que leur performance sur la représentativité du régime interannuel est assez variable (figure V.28). Le modèle représentant le mieux le régime (point vert) est identifié parmi les 10 modèles restants en minimisant le critère C_2 :

$$C_2 = \sqrt{\left(1 - KGE_{Q_{reg}}^C\right)^2 + \left(1 - KGE_{Q_{reg}}^V\right)^2} \quad (V.22)$$

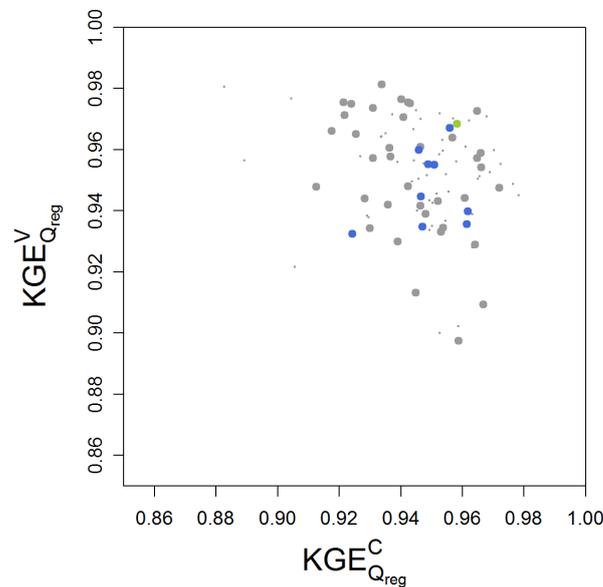


FIGURE V.28 – Espace à 2 dimensions pour identifier le meilleur modèle selon $KGE_{Q_{reg}}$.

Bilan : ce principe d'écrémages successifs vise à reproduire le choix de modèle MORDOR de référence qui sont le résultat de choix et de compromis réalisés par les équipes d'études (figure V.29). Une première sélection consiste à identifier les modèles qui ont la capacité d'assigner une probabilité pertinente aux plus fortes valeurs observées sur le bassin considéré et respectant aux mieux les distributions mensuelles des écoulements. Sous ces conditions, le modèle candidat doit également être capable de représenter convenablement le régime des débits. Des pistes d'améliorations sont envisageables. L'utilisation d'un critère FF pour évaluer la capacité à associer une probabilité cohérente aux plus forts débits observés serait par exemple à tester pour le choix des modèles dans l'étape 2.

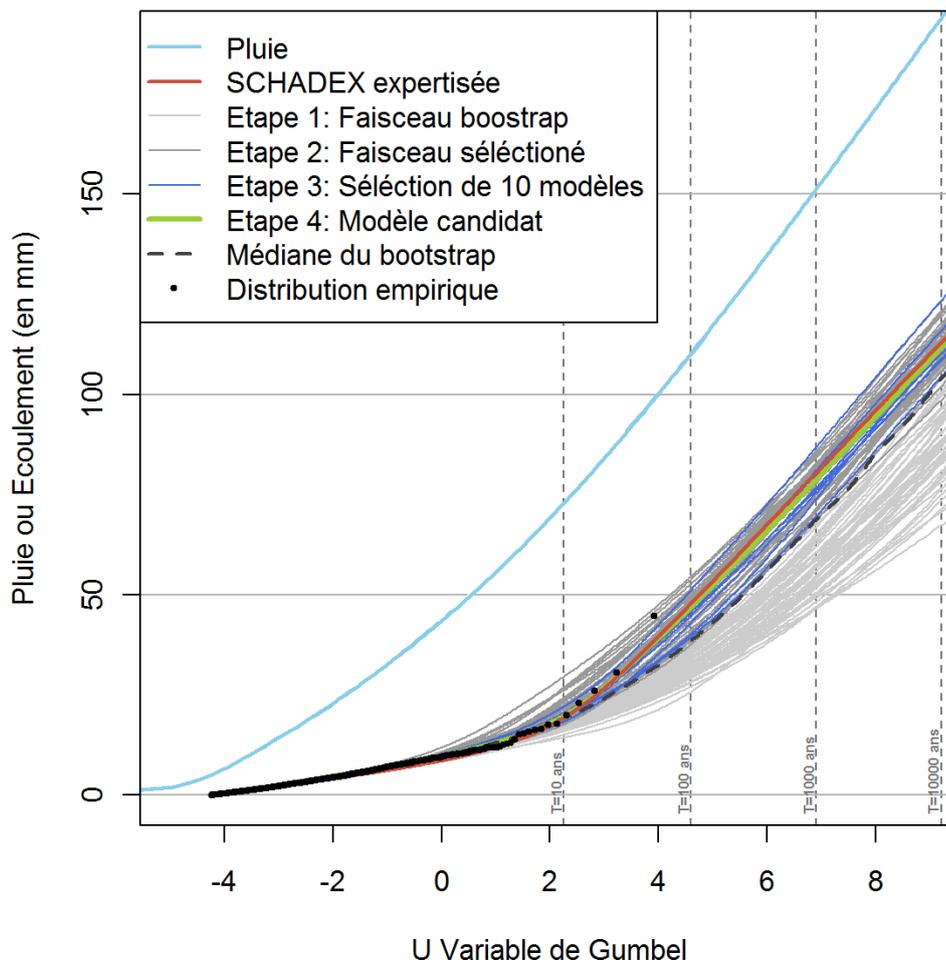


FIGURE V.29 – Exemple de réduction progressive du faisceau de boot-strap pour le Lot à Castelnau (1628 km²).

V. 4. 3 Évaluation de la simulation SCHADEX automatique

Dans cette partie, une comparaison entre la simulation SCHADEX de référence et la simulation SCHADEX automatisée est réalisée sur la base des 45 études de référence. Pour chaque bassin, la saisonnalité des précipitations est déterminée de façon automatique avec l'algorithme présenté dans la section V. 4. 1. Un ajustement MEWP des pluies centrales avec ces saisons et les tableaux de contingence pour les pluies adjacentes sont alors construits. Puis un modèle hydrologique MORDOR est choisi avec la procédure de sélection développée plus haut (dans la section V. 4. 2). La simulation SCHADEX s'appuyant sur les modèles probabilistes de pluie et le modèle hydrologique MORDOR obtenu fournit alors une distribution SCHADEX automatisée.

Sur la figure V.30 (a), un exemple illustre les différentes distributions de pluie et de débit sur le bassin de la Maulde à Peyrat Le Château (174 km²). La distribution des pluies journalières de l'étude SCHADEX de référence est tracée en bleu foncé. La distribution des précipitations obtenue avec la même série de pluie, mais avec une détection automatique des saisons, est représentée par la courbe bleue ciel. Les autres courbes sont des distributions d'écoulements journaliers. La distribution rouge correspond à la distribution SCHADEX de référence. Les courbes en trait plein gris sont les 100 distributions SCHADEX avec le boot-strap. La courbe noire en pointillés est la distribution médiane du boot-strap, et la courbe verte, celle de la distribution SCHADEX automatisée (obtenue avec le modèle candidat choisi).

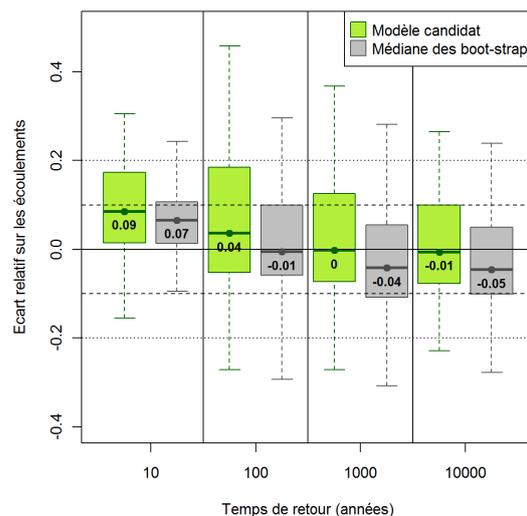
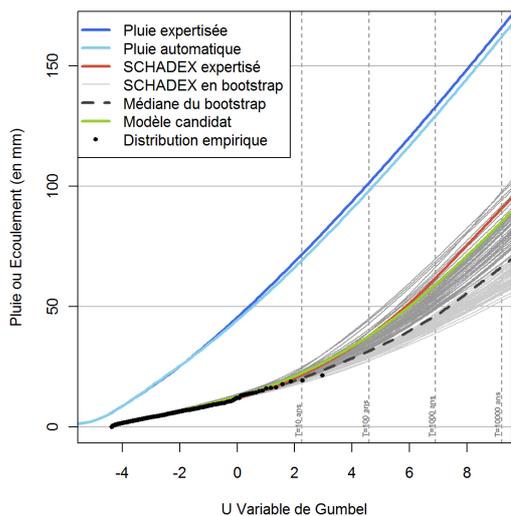
La figure V.30 (b) compare les écoulements de temps de retour 10, 100, 1000 et 10 000 ans de la distribution de référence $E_{\text{référence}}(T)$ à ceux de la distribution automatisée ou de la distribution médiane des boot-strap $\hat{E}(T)$. L'axe des ordonnées correspond aux écarts relatifs :

$$\text{Écart relatif} = \frac{\hat{E}(T) - E_{\text{référence}}(T)}{E_{\text{référence}}(T)}$$

L'utilisation de la médiane des distributions du boot-strap comme distribution candidate pour représenter la distribution de référence conduit à une sous-estimation des débits de temps de retour élevés (1000 et 10 000 ans). Cette sous-estimation est en grande partie due aux cas où deux faisceaux de comportements différents sont présents (figure V.30 (a)). La sélection d'un modèle candidat permet de corriger ce biais sur les écarts relatifs pour les débits exceptionnels. Cependant, cette automatisation du choix du modèle conduit à surestimer de façon quasi systématique les débits décennaux de références avec une surestimation moyenne sur les écarts relatifs de 9%. Les scatters-plots (c), (d), (e) et (f) montrent que les estimations avec la distribution SCHADEX automatique sont assez proches de celles de la simulation de référence.

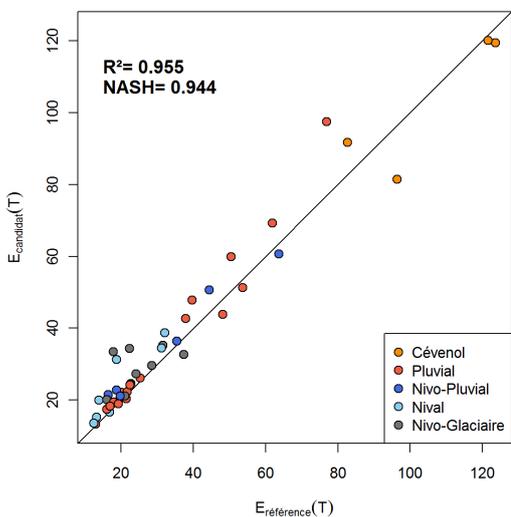
Bilan : Nous rappelons ici que l'objectif de ce travail d'automatisation de la base SCHADEX est de fournir une base étendue pouvant servir à la régionalisation des paramètres d'un modèle régional de passage pluie-débit (en particulier du modèle EREBOR). Ce sera aussi un excellent moyen de fournir rapidement des estimations SCHADEX sur tous les bassins disposant de données et de permettre des comparaisons inter-méthodes avec la méthode SCHADEX sur des bassins n'ayant pas fait l'objet d'étude détaillée.

Le processus d'automatisation proposé pour la méthode SCHADEX permet de fournir une estimation de la distribution des débits cohérente avec celle des études SCHADEX expertisées. Le choix d'une saisonnalité pour les ajustements sur les précipitations et d'un modèle MORDOR ont notamment été objectivés de façon à pouvoir capter au mieux la saison à risque des fortes précipitations et de disposer d'une modélisation pluie-débit capable de reproduire au mieux les crues.

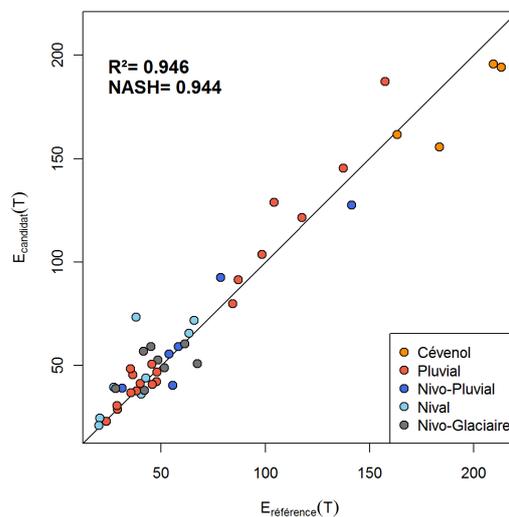


(a) Exemple de la Maulde à Peyrat Le Château (174 km²)

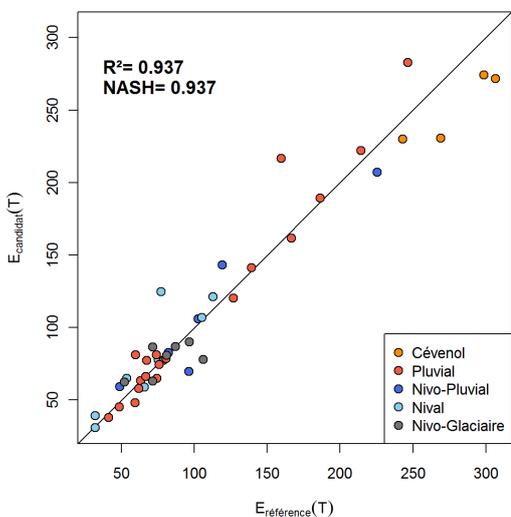
(b) Écarts relatifs sur les quantiles d'écoulements



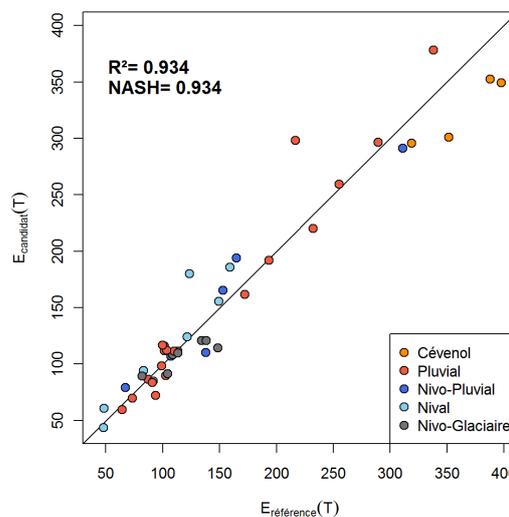
(c) Écoulements décennaux



(d) Écoulements centennaux



(e) Écoulements millénaux



(f) Écoulements décamillénaux

FIGURE V.30 – Comparaison des quantiles d'écoulements (en mm) entre les distributions SCHADEX de référence et automatisée pour différents temps de retour.

V. 5 Passage à la distribution des débits de pointes

Dans la prédétermination des crues et dans ses applications (pour le dimensionnement d'ouvrage), la notion de crue de projet est souvent définie. Elle est caractérisée par un hydrogramme avec un débit maximal de pointe QX de temps de retour donné (e.g. la crue maximale de pointe millénaire). Jusqu'à maintenant, nous nous sommes intéressés aux distributions d'écoulements ou de débits moyens sur le pas de temps d'étude du bassin versant (dans notre cas le pas de temps journalier). Cependant, la distribution des débits de pointe QX (i.e. le débit instantané maximal d'une crue) intéresse également l'hydrologue.

La distribution des débits de pointe permet de définir le débit instantané maximum de temps de retour T noté $QX(T)$. Le plus souvent, les valeurs $QX(T)$ sont utilisées avec un hydrogramme de référence normalisé de façon à avoir une valeur de pointe de 1. Une crue de projet pour le temps de retour T est alors estimée par le produit de cet hydrogramme normalisé et de la valeur du débit de pointe de temps de retour T donné par la distribution.

V. 5. 1 La méthode classique : le coefficient de forme moyen

Dans la méthode SCHADEX, comme auparavant dans la méthode GRADEX, le passage des distributions de volumes à celle des débits de pointe de crues est réalisé par l'application d'un coefficient multiplicatif constant, le coefficient de forme K_C .

La troisième hypothèse de la méthode GRADEX est reprise. Elle affirme que le coefficient de forme peut être considéré comme une variable aléatoire ayant une certaine variabilité et dont la valeur moyenne, calculée sous un échantillon bien choisi de crues, ne dépend pas de l'ampleur de la crue et reste un bon indicateur du rapport entre le débit de pointe et le débit moyen pour des temps de retour élevés.

En pratique, l'évaluation du coefficient de forme est effectuée en deux étapes.

Étape 1 : La collection d'hydrogrammes de crue est sélectionnée manuellement à partir de la visualisation des chroniques de débit horaire, du dépouillement de limnigrammes, ou de relevés réalisés lors des crues sur un aménagement (cote du plan d'eau, position des vannes, etc.). Plusieurs critères qualitatifs sur les hydrogrammes contribuent à rendre cette sélection la plus homogène possible :

- le débit maximal atteint doit être significatif ;
- la dynamique de crue doit être la plus « pure » possible (« type hydrogramme unitaire ») ;
- le débit de base (au début de l'hydrogramme, avant la montée) doit être faible devant le débit de pointe (de 20 à 30% au maximum) ;
- les crues dont les hydrogrammes sont très étalés, influencés par la neige (e.g. abatement nival important qui « casse » la montée en débit, récession soutenue par la fonte), ou trop complexes (e.g. succession rapprochée d'événements) doivent être éliminées.

Pour synthétiser, tout ceci vise à ne retenir que des crues dont la forme pourrait être transposée de manière crédible sur un événement extrême.

Étape 2 : Une fois l'échantillon d'hydrogrammes sélectionné, le coefficient de forme $K_C^i = \frac{QX^i}{QJX^i}$ de chaque crue est calculé. Ces coefficients de forme fournissent alors une estimation du coefficient de forme moyen $K_C = \overline{K_C^i}$.

Un hydrogramme moyen, calculé sur la sélection d'hydrogrammes préalablement normalisés, permet de disposer d'une forme de crue moyenne et unique pour la crue de projet.

Tout en restant une solution applicable d'un point de vue opérationnel, la méthode d'estimation du coefficient de forme présentée dans la section précédente n'est pas totalement satisfaisante. Il faut étudier en détail l'ensemble de la chronique de débit pour sélectionner les hydrogrammes les plus aptes à représenter le coefficient de forme du bassin. La valeur obtenue par la moyenne des coefficients de forme peut donc varier significativement avec un impact multiplicatif direct sur l'estimation du débit de pointe selon la sélection effectuée. De plus, une utilisation en contexte non jaugé est impossible car aucun hydrogramme n'est alors disponible.

Avec cette méthode, la dynamique des crues est représentée par un coefficient de forme et un hydrogramme de crue de projet uniques. Différents travaux ont porté sur la prise en compte d'une variabilité des dynamiques de crues et du coefficient de forme (Le Clerc, 2004; Favre *et al.*, 2004; Kuentz, 2009; Garavaglia, 2011b; Penot, 2011). Le modèle du coefficient de forme variable en est un exemple (section I. 2, étape 7)

V. 5. 2 La génération d'hydrogrammes pour des scénarios en volume

La méthode de génération d'hydrogrammes de crue présentée dans cette section a été développée dans le cadre d'une étude de crues extrêmes d'EDF sur la Basse-Isère (Paquet, 2014). Nous allons détailler le principe méthodologique permettant d'associer un hydrogramme horaire à un scénario en volume et sa validation.

V. 5. 2.1 Principe méthodologique

Que ce soit avec la méthode SCHADEX ou la méthode EREBOR, les scénarios de débits en sortie de simulation sont des épisodes constitués de séquences de débits moyens au pas de temps de l'étude. Pour simplifier les explications qui vont suivre, nous allons considérer un pas de temps journalier même si dans l'absolu tout le processus peut être déroulé pour un pas de temps quelconque. Le problème qui se pose est de réussir à associer un hydrogramme horaire à un scénario d'écoulements journaliers.

L'archive d'hydrogrammes

Une vaste base de données d'hydrogrammes à l'échelle de la France a été constituée. Elle rassemble près de 4700 hydrogrammes de crues issus de 120 stations, à l'exutoire de bassins de surface comprise entre 10 et 2240 km², réparties sur l'ensemble du territoire français. Pour un hydrogramme observé, 24 variantes sont calculées avec 24 heures différentes de référence pour le calcul des débits journaliers. En pratique, cela correspond à déplacer la fenêtre de calcul des débits moyens journaliers au sein d'une journée. Ainsi, ce sont au total plus de 113 000 formes différentes qui pourront être proposées, tous régimes hydrologiques et toutes surfaces de bassin confondus.

L'idée est alors de se servir de cette archive d'hydrogrammes observés, pour proposer un hydrogramme représentatif d'un scénario en volumes journaliers donné.

Le critère de proximité

Trouver les hydrogrammes observés de l'archive les plus pertinents avec le scénario en volume (pour lequel l'hydrogramme n'est pas connu) nécessite de définir un critère de proximité entre les séquences de débits journaliers d'une crue. Pour cela, des indicateurs adimensionnels (formule (V.23)) sont calculés pour la crue simulée et pour l'ensemble des hydrogrammes de la base.

$$\begin{cases} jx \text{ est le jour où le débit journalier est maximal.} \\ r_{-2} = \frac{QJ(jx-2)}{QJ(jx)} & r_{-1} = \frac{QJ(jx-1)}{QJ(jx)} \\ r_{+1} = \frac{QJ(jx+1)}{QJ(jx)} & r_{+2} = \frac{QJ(jx+2)}{QJ(jx)} \end{cases} \quad (\text{V.23})$$

Ces indicateurs sont précalculés pour tous les hydrogrammes de l'archive, soit environ 113 000 valeurs du quadruplet $\{r_{-2}, r_{-1}, r_{+1}, r_{+2}\}$.

Nous pouvons considérer une séquence cible de débits journaliers pour laquelle un hydrogramme doit être généré. La séquence cible de débits journaliers a une dynamique qualifiée par $\{r_{-2}^c, r_{-1}^c, r_{+1}^c, r_{+2}^c\}$. Chaque hydrogramme candidat, qualifié par $\{r_{-2}, r_{-1}, r_{+1}, r_{+2}\}$, est alors caractérisé par la distance pondérée d entre les quadruplets de la séquence cible et de l'hydrogramme candidat :

$$d = \sqrt{\omega_{-2} (r_{-2}^c - r_{-2})^2 + \omega_{-1} (r_{-1}^c - r_{-1})^2 + \omega_{+1} (r_{+1}^c - r_{+1})^2 + \omega_{+2} (r_{+2}^c - r_{+2})^2} \quad (\text{V.24})$$

avec $\omega_{-2} = 0.1$, $\omega_{-1} = 1$, $\omega_{+1} = 0.9$ et $\omega_{+2} = 0.05$. Ces pondérations accentuent très fortement l'analogie sur la dynamique des deux jours encadrant la pointe. Elles ont été optimisées pour représenter au mieux les hydrogrammes des plus grosses crues observées sur les témoins.

Cette distance permet de classer toute la base d'hydrogrammes selon la proximité de la séquence de débits sur les 5 jours de la crue par rapport à la séquence cible. En pratique, seuls les hydrogrammes candidats provenant de bassins de même régime hydrologique que le bassin cible sont recherchés. Quatre catégories ont été retenues : Nival, Pluvio-nival, Pluvial et Méditerranéen. La base d'hydrogrammes candidats est ainsi classés selon leur proximité avec la séquence de volumes journaliers cible.

Les hydrogrammes synthétiques simulés

Pour chacun des candidats disponibles, un hydrogramme de synthèse est élaboré à partir :

- d'une courbe spline reliant toutes les valeurs du débit journalier affecté à la mi-journée sur la séquence cible (courbe pleine bleue et courbe interrompue rouge de la figure V.31)
- de l'hydrogramme d'un donneur issu de la base (courbes pointillées noire et continue rouge de la figure V.31)

Il faut ensuite raccorder le plus continûment possible la forme proposée par l'hydrogramme donneur autour de la pointe aux débits préalables et postérieurs à la crue, en raccordant la spline issue des débits journaliers et l'hydrogramme donneur aux moments où ils sont les plus proches, entre $jx - 2$ et $jx - 1$ avant la pointe, et entre $jx + 1$ et $jx + 2$ après la pointe.

La figure V.31 présente l'exemple du premier donneur possible pour fournir un hydrogramme à la crue modélisée (un des scénarios SCHADDEX) à la Pernière. La partie rouge est issue de l'hydrogramme donneur, les parties bleues sont issues de la séquence de débits journaliers. L'hydrogramme synthétique issu de ce donneur est alors intégré au pas de temps journalier (histogramme bleu) et comparé aux débits journaliers cibles (histogramme hachuré) sur l'ensemble de la séquence simulée. Une erreur quadratique moyenne (RMSE) est calculée entre les deux séries de débits journaliers. Cet hydrogramme sera retenu s'il respecte les deux conditions suivantes :

- i) l'erreur relative sur les volumes $\frac{RMSE}{QJ}$ est inférieure à 15% ;
- ii) une autre modalité de cet hydrogramme (même hydrogramme mais décalé dans le temps) n'est pas déjà sélectionnée parmi les donneurs.

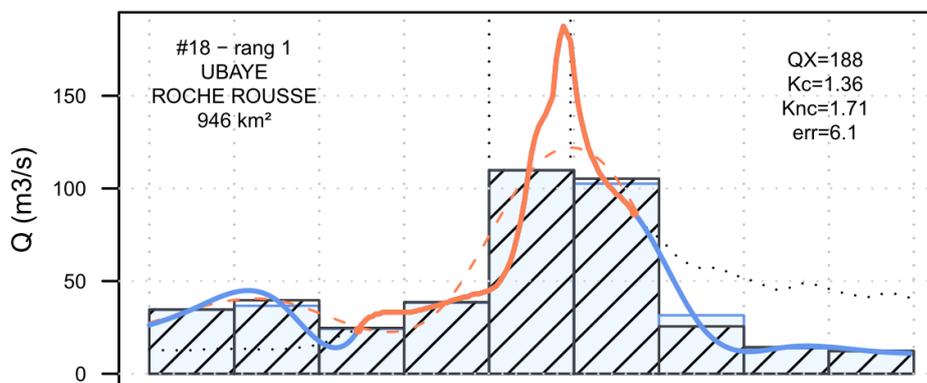


FIGURE V.31 – Exemple de simulation pour l'Eau d'Olle à Pernière – hydrogramme du donneur n°1.

À noter qu'un hydrogramme observé sur le bassin cible peut être candidat, au même titre que les autres hydrogrammes des autres bassins. L'information locale, si elle est pertinente pour l'épisode en question, est donc valorisée dans ce procédé.

L'ensemble de l'archive d'hydrogrammes candidats classés est parcourue, jusqu'à constituer un échantillon de 10 hydrogrammes synthétiques vérifiant les conditions désirées. Ces dix hydrogrammes synthétiques sont issus de 10 hydrogrammes observés de donneurs mis à l'échelle et raccordés grâce à la séquence cible de débits journaliers. La figure V.32 présente les 10 hydrogrammes synthétiques sélectionnés pour représenter une crue forte de l'Eau d'Olle à Pernière (169 km²) issue d'une simulation SCHADEX (débit journalier maximum de 110 m³/s, soit une crue de temps de retour d'environ 50 ans).

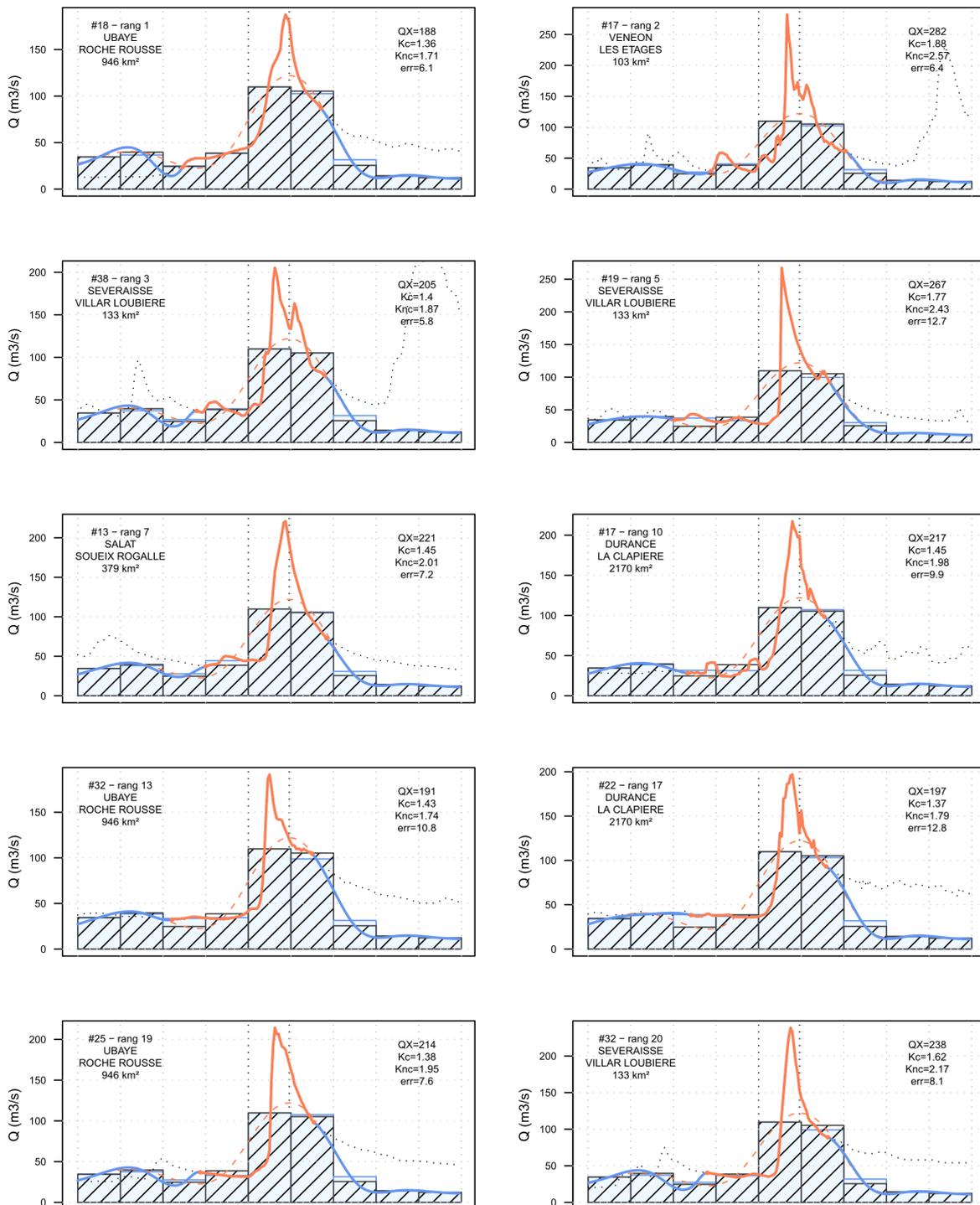


FIGURE V.32 – Exemple de simulation pour l'Eau d'Olle à Pernière – les 10 hydrogrammes synthétiques sélectionnés.

L'étape suivante va permettre de résumer ces 10 formes en prenant, à chaque heure, la médiane des 10 valeurs hydrogrammes synthétiques. Un lissage très léger est effectué pour gommer les instabilités, inévitables dans un tel processus.

L'hydrogramme synthétique finalement proposé à l'Eau d'Olle à Pernière est présenté dans la figure V.33. Le coefficient de forme de cet hydrogramme synthétique vaut 1.37, valeur dans ce cas inférieure au coefficient de forme moyen de l'étude locale ($K_C = 1.57$).

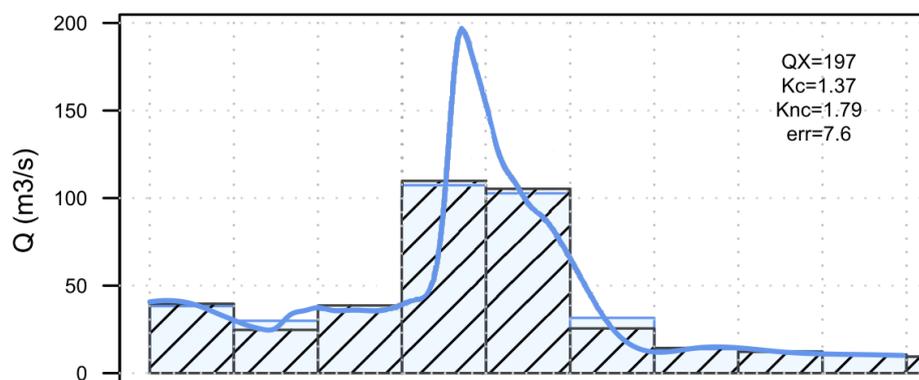


FIGURE V.33 – Hydrogramme synthétique généré pour le scénario de crue l'Eau d'Olle à Pernière.

Cette méthode réalise donc un passage au débit de pointe qui s'affranchit d'une hypothèse importante jusqu'ici imposée dans la méthode SCHADEX : celle d'un coefficient de forme constant affecté aux crues simulées, identifié autant que possible sur une collection d'hydrogrammes locaux. Le procédé exposé propose un scénario hydrologique compatible avec la séquence de débits journaliers simulés, et basé sur des formes de crue issues d'une base de données régionale (au sens large, tous les bassins de régime compatible peuvent être candidats).

Un autre avantage est que l'heure de la pointe n'est pas une hypothèse du procédé, elle découle des formes de crue des donneurs, pour lesquelles toutes les modalités (i.e. les 24 heures possibles d'occurrence de la pointe dans la journée) sont évaluées.

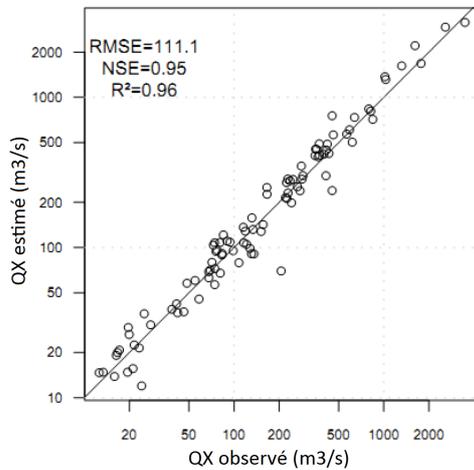
V. 5. 2.2 Validation du procédé

Cette méthode a été qualifiée statistiquement par rapport à la méthode classique du coefficient de forme moyen. Au sein de la base de données, 96 stations disposant d'une bonne collection d'hydrogrammes ont été identifiées. Pour ces 96 stations, le débit de pointe de la crue maximale de la station a été reconstitué de deux manières :

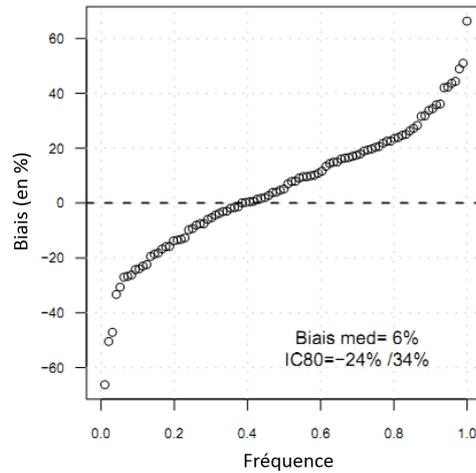
Méthode classique : le coefficient de forme moyen est identifié sur la collection d'hydrogrammes de la station. Le débit de pointe est calculé en multipliant le débit journalier maximum par ce coefficient de forme. C'est la méthode qui est utilisée dans une application standard de la méthode SCHADEX pour passer des débits moyens journaliers aux débits de pointe.

Hydrogramme synthétique : la méthode exposée au paragraphe précédent permet de sélectionner 10 donneurs dont les hydrogrammes s'adaptent au mieux à la séquence de débit journaliers observés. À noter que pour ce protocole d'évaluation, la sélection d'hydrogrammes de la station en question n'est pas autorisée (évaluation purement régionale donc). L'hydrogramme synthétique est alors bâti avec la médiane des 10 donneurs, puis mis à l'échelle grâce au débit journalier maximum du site cible, ce qui permet d'aboutir au débit de pointe.

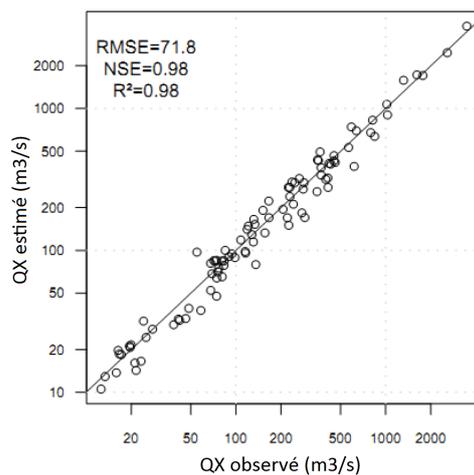
Nous allons donc comparer les statistiques sur la valeur estimée du débit de pointe de la plus grosse crue observée sur les 96 stations. Elles sont présentées dans la figure V.34, au moyen de deux graphiques : le scatter-plot des débits de pointe estimés $QX_{\text{estimé}}$ et observés $QX_{\text{observé}}$ (graphes (a) et (c)) et la distribution des biais $\frac{QX_{\text{estimé}} - QX_{\text{observé}}}{QX_{\text{observé}}}$ (graphes (b) et (d)).



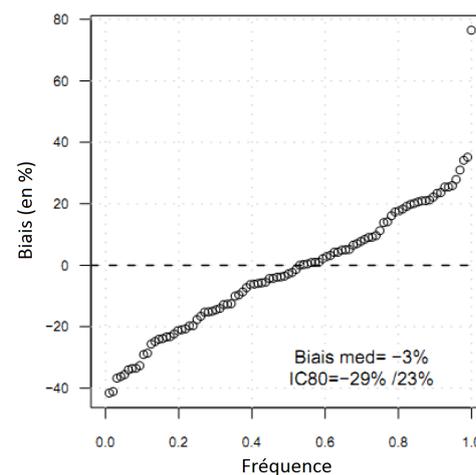
(a) *Méthode classique* – Scatter-plot



(b) *Méthode classique* – Distribution des biais



(c) *Hydrogrammes synthétiques* – Scatter-plot



(d) *Hydrogrammes synthétiques* – Distribution des biais

FIGURE V.34 – Reconstitution d'un échantillon de crues historiques – comparaison des approches.

Sur toutes les statistiques présentées dans ces graphes, l'approche par construction d'un hydrogramme synthétique a des performances légèrement meilleures que l'approche classique utilisant le coefficient de forme moyen pour la reconstitution du débit de pointe des crues maximales observées. De plus, l'approche par hydrogramme synthétique propose une forme de crue adaptée à la dynamique des débits journaliers, alors que pour certains événements, l'hydrogramme moyen local peut s'avérer incompatible avec la séquence des débits journaliers simulés. Le principale avantage de cette approche par hydrogramme synthétique réside en grande partie dans l'utilisation d'une archive régionale d'hydrogrammes. La méthode est applicable indifféremment aux bassins où l'on dispose ou non d'hydrogrammes locaux, alors que le choix d'un hydrogramme de référence pour un bassin non jaugé est une hypothèse lourde.

Bilan : dans un contexte non jaugé, le générateur d'hydrogrammes s'appuie sur une archive régionale d'hydrogrammes pour associer un hydrogramme à une séquence de volumes journaliers. Cet hydrogramme synthétique fournit alors une estimation du débit de pointe prenant en compte la dynamique des débits moyens journaliers. Une validation de cette approche sur un jeu de 96 stations de débits a montré sa capacité à générer des populations d'hydrogrammes pertinentes et plus justes que celle obtenues avec un coefficient de forme moyen.

Ce générateur, utilisable pour les scénarios d'écoulements de la méthode SCHADDEX, est également envisageable sur les scénarios d'écoulements journaliers d'une simulation EREBOR. **Un hydrogramme synthétique pourrait être généré afin de disposer d'une estimation de débit de pointe pour chaque scénario et donc , au final, d'une distribution des débits de pointe. Cependant son utilisation sur les scénarios EREBOR devra faire l'objet d'une évaluation complémentaire.**

V. 6 Conclusion

Le calage des différents modèles statistiques et hydrologiques impliqués dans la méthode SCHADEX repose sur une quantité d'observations pluie, débit et température importante. C'est pourquoi une utilisation en contexte non jaugé nécessite la mise en place d'une méthode plus simple mais qui reprend les principes clés de la méthode SCHADEX de façon à reproduire des distributions de débits issues des mêmes hypothèses que celles faites lors des études industrielles. C'est pour répondre à cet enjeu que la méthode de simulation EREBOR a été développée.

La simulation EREBOR est une méthode de simulation stochastique reprenant la plupart des composants fondamentaux de la méthode SCHADEX :

- la génération d'épisodes de pluie pour parcourir l'aléa pluie,
- l'utilisation d'un modèle hydrologique pour simuler les écoulements en tenant compte d'un état variable du bassin.

La représentation de l'aléa pluie est faite au travers de la génération d'épisodes pluvieux sur trois pas de temps journaliers. Un ajustement MEWP est réalisé sur les pluies centrales et des tableaux de contingence sont déterminés pour les pluies adjacentes. Ces modèles probabilistes sont utilisés pour probabiliser les scénarios de précipitation générés.

Une fois la génération de scénarios pluvieux réalisée, le passage pluie-débit doit être traité. Dans la méthode SCHADEX, le modèle hydrologique MORDOR permet de transformer l'évènement pluvieux en scénario d'écoulements en tenant compte d'un état de saturation du bassin. Cependant l'utilisation de MORDOR pour une approche régionale n'est pas aisée. Un modèle SCS, plus simple et peu paramétré a donc été considéré pour réaliser ce passage pluie-débit. Ce modèle s'appuie sur trois paramètres (quatre dans le cas d'un bassin nival) permettant de transformer un scénario de pluie en écoulement en tenant compte de l'aléa de saturation du bassin. Les différents paramètres sont :

- la capacité maximale de rétention et la distribution des états de saturation du bassin qui gèrent la fonction de production. Les états de saturations sont considérés comme une variable aléatoire de façon à tenir compte des aléas de saturation du bassin ;
- l'hydrogramme unitaire pour transférer et étaler les écoulements jusqu'à l'exutoire du bassin ;
- une distribution des contributions nivales (si la composante nivale du bassin considéré est significative).

Une simulation EREBOR fournit une distribution de débits journaliers croisant l'aléa pluie et l'aléa de saturation du bassin.

Une première tentative de régionalisation de ces paramètres à partir des régimes hydrologiques a été envisagée. Il est apparu que contraindre les différents paramètres selon le régime hydrologique du bassin produit une première estimation de la distribution des débits moyens journaliers cohérente par rapport à celle de la méthode SCHADEX. Cependant, la régionalisation des paramètres pourra clairement faire l'objet d'un travail complémentaire pour améliorer leurs estimations. La base d'études SCHADEX pourra notamment être complétée de façon à disposer de suffisamment de bassins versants pour construire des classes de régimes plus pertinentes et plus robustes (e.g. sans le problème observé pour les bassins nivo-pluviaux). La classification des bassins pourra notamment être envisagée selon d'autres critères que les débits mensuels normalisés comme par exemple des régimes de crues. Sur cette base élargie, la capacité maximale de rétention S_{max} et l'hydrogramme unitaire pourront également être régionalisés de manière plus pertinente que sur la seule information du régime hydrologique. Malgré tout, le travail préliminaire effectué dans ce chapitre a permis de

fixer une première version d'une approche de simulation stochastique simplifiée capable d'estimer une distribution des débits journaliers pour un bassin versant non jaugé.

Dans l'optique de disposer d'études SCHADEX sur une large base de bassins versants, nous nous sommes intéressés dans la seconde partie de ce chapitre à une automatisation de la méthode SCHADEX. En effet, l'expertise dont fait l'objet les études industrielles n'est pas envisageable sur plusieurs centaines d'études. Le choix d'une saisonnalité pour les précipitations a été objectivé de manière à capter au mieux la saison à risque sur la bassin versant et à ne pas sous-estimer les quantiles de pluie. Un protocole a également été mis en place afin de choisir un modèle MORDOR le plus pertinent possible pour réaliser une simulation SCHADEX. La capacité du modèle à associer une probabilité raisonnable aux plus fortes crues observées sur le bassin versant a été un critère important pour sélectionner un modèle apte à reproduire le comportement des distributions des études SCHADEX expertisées. Cette automatisation de la méthode SCHADEX fournit une estimation des débits extrêmes qui permettra de construire une large base de référence d'études SCHADEX pour approfondir la régionalisation des paramètres de la méthode EREBOR, mais également d'être capable de comparer la méthode SCHADEX à d'autres méthodes sur un large jeu de bassins (dans l'esprit du projet ExtraFlo). Ce travail sera également l'occasion de valider les choix d'automatisation réalisés sur un grand nombre de bassins.

Une fois la distribution de débits moyens journaliers estimée, il faut passer à la distribution des débits de pointe. La méthode classiquement utilisée repose sur l'estimation du coefficient de forme moyen. Son estimation à partir d'une sélection d'hydrogrammes est très sensible, dépend exclusivement des données locales disponibles et ne permet pas de représenter la variabilité des dynamiques de crues observées dans la réalité. **Un générateur d'hydrogrammes synthétiques permet de construire un hydrogramme bien adapté aux scénarios de débits moyens journaliers générés en cherchant des situations analogues dans une archive régionale d'hydrogrammes.** Cette approche régionale dispose de performances légèrement meilleures que la méthode classique utilisée en contexte jaugé. Ce générateur permet alors de disposer d'un débit de pointe pour chaque scénario d'écoulement et donc de construire une distribution du débits de pointe, objectif final d'une étude SCHADEX. Son utilisation sur les scénarios d'écoulements journaliers EREBOR devra tout de même l'objet d'une validation.

Chapitre

VI

Exemple d'étude de crues extrêmes en non jaugé : l'Eyrieux à Beauvène

Dans cette partie, nous allons revenir sur un cas plus concret en appliquant les méthodes abordées tout au long de ce manuscrit sur le bassin de l'Eyrieux à Beauvène.



SOMMAIRE DU CHAPITRE

VI. 1 Présentation du bassin de l'Eyrieux à Beauvène	217
VI. 2 Les modèles probabilistes des pluies	217
VI. 3 Passage pluie-débit en contexte non jaugé : Simulation EREBOR	221
VI. 4 Passage pluie-débit en contexte jaugé : méthode SCHADEX	222

VI.1 Présentation du bassin de l'Eyrieux à Beauvène

L'Eyrieux à Beauvène, situé en Ardèche, est un bassin versant de 454 km² avec une altitude médiane de 912 m et un point culminant à 1735 m (figure VI.1 (a)). Sur ce bassin, les fortes précipitations sont essentiellement dues à des perturbations méditerranéennes qui touchent les reliefs au sud du bassin (figures VI.1 (b) et (c)).

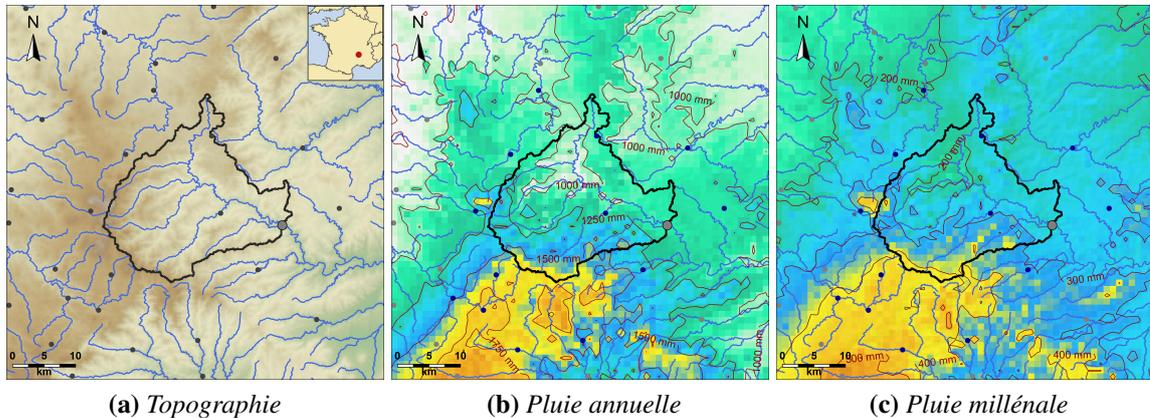


FIGURE VI.1 – Localisation du bassin de l'Eyrieux à Beauvène (454 km²).

Une série de pluies spatiales journalières SPAZM est disponible de 1953 à 2005. La pluie annuelle moyenne est de 1170 mm. Une station de débit fournit les débits journaliers à Beauvène de 1953 à 2005.

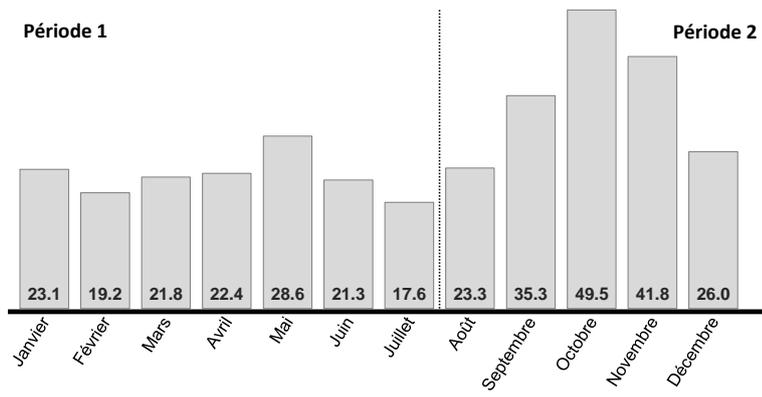
Les modèles probabilistes des pluies seront ajustés sur la pluie SPAZM. La série de débits sera tout d'abord considérée comme inexistante et la méthode EREBOR sera appliquée au bassin (*contexte non jaugé*). Une étude SCHADDEX sera ensuite réalisée sur le bassin de l'Eyrieux en prenant en compte la série de débit de Beauvène (*contexte jaugé*).

VI.2 Les modèles probabilistes des pluies

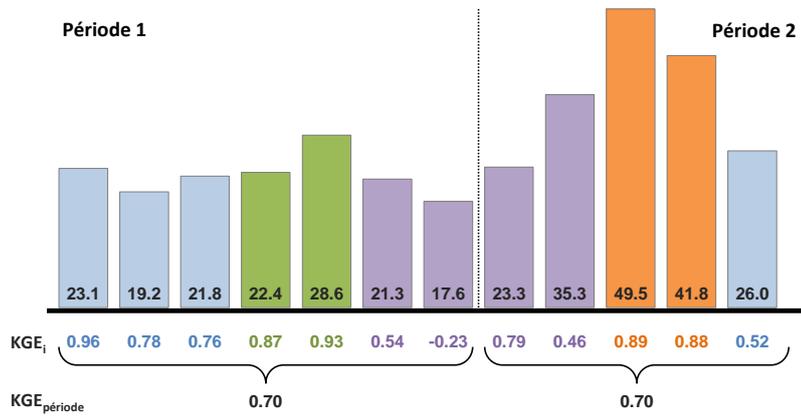
La saisonnalité des pluies journalières

Un modèle probabiliste pour les pluies centrales et pour les pluies adjacentes va être ajusté à partir de la pluie spatiale journalière SPAZM du bassin. Afin de ne pas sous-estimer les quantiles de pluies, la saisonnalité des pluies est déterminée avec l'algorithme présenté dans la section V.4.

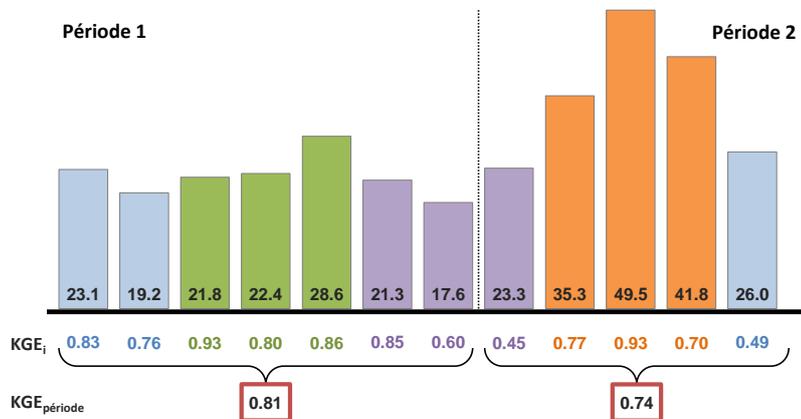
Les pluies mensuelles maximales moyennes sont calculées pour chaque mois (figure VI.2 (a)). Trois cas de saisons sont ensuite envisagés avec une saison à risque de deux, trois puis quatre mois sur chaque période (saison verte et orange, figure VI.2 (b), (c) et (d)) et une distribution MEWP est ajustée sur chacun d'entre eux. Les critères KGE_i mensuels entre la distribution empirique du mois i et la distribution saisonnière correspondante sont ensuite estimés. Pour chaque période et chaque cas, une moyenne pondérée des KGE_i mensuels est calculée. La saison à risque choisie pour chaque période est celle correspondant au cas ayant le meilleur KGE . Ici, le découpage en saison choisi pour la période 1 et 2 est celui du cas 2.



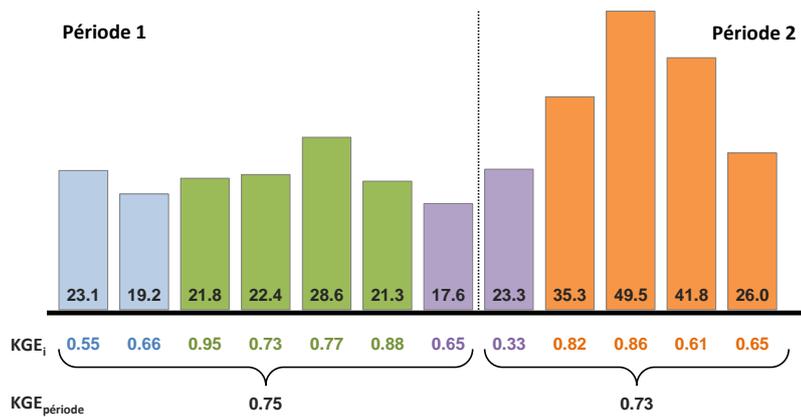
(a) Pluie mensuelle maximale moyenne



(b) Cas 1



(c) Cas 2



(d) Cas 3

FIGURE VI.2 – Détermination automatique de la saison.

Modèles probabilistes pour les pluies journalières

Un ajustement MEWP est réalisé sur les pluies centrales de la pluies SPAZM en considérant la saison déterminée (figure VI.2 (c)). La figure VI.3 montre les quatre ajustements saisonniers et l'ajustement annuel avec un gradex $27.7\text{mm}/24\text{h}$ issu de la recomposition. La saison à risque est la saison de septembre à novembre.

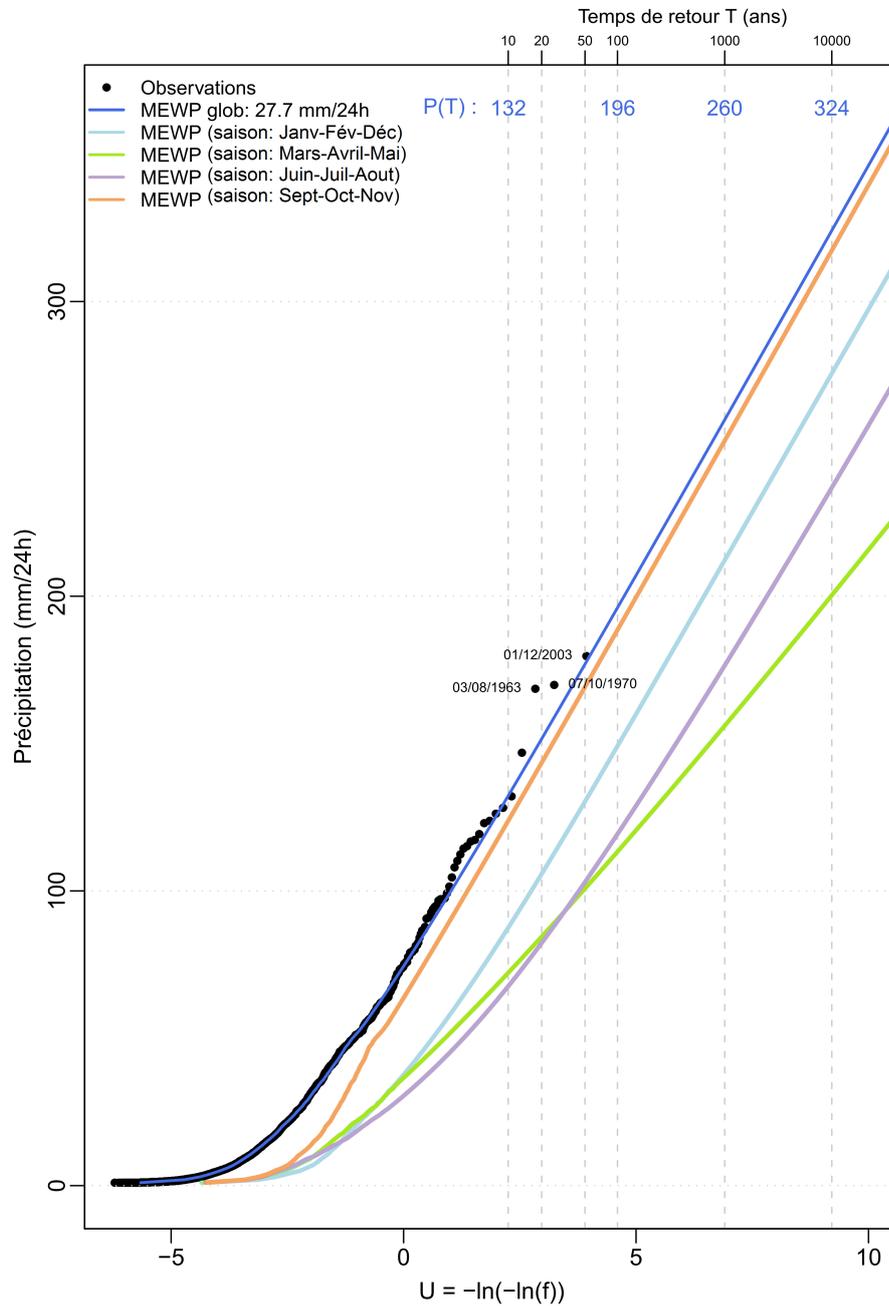


FIGURE VI.3 – Modèle probabiliste des pluies centrales – Ajustement MEWP sur les pluies journalières SPAZM pour l'Eyrieux à Beauvène (454 km²), période 1953-2005.

Les deux tableaux de contingence servant de modèle probabiliste aux pluies adjacentes sont également estimés à l'aide de la pluie spatiale SPAZM (tableaux VI.1).

P_C					$P_C + P_{a-}$				
Classe	1	2	3	4	Classe	1	2	3	4
Borne Inf. (mm)	1	6	19	42	Borne Inf. (mm)	1	8	22	51
Borne Sup. (mm)	6	19	42	∞	Borne Sup. (mm)	8	22	51	∞
$0 \leq P_{a-}/P_C < 0.1$	0.4293	0.4634	0.4408	0.3502	$0 \leq P_{a+}/P_C < 0.1$	0.4144	0.4131	0.4114	0.4326
$0.1 \leq P_{a-}/P_C < 0.2$	0.1317	0.1548	0.1601	0.2396	$0.1 \leq P_{a+}/P_C < 0.2$	0.1561	0.1831	0.1977	0.1814
$0.2 \leq P_{a-}/P_C < 0.3$	0.0983	0.0785	0.0998	0.0829	$0.2 \leq P_{a+}/P_C < 0.3$	0.0956	0.1056	0.1068	0.0977
$0.3 \leq P_{a-}/P_C < 0.4$	0.0611	0.0858	0.0742	0.0829	$0.3 \leq P_{a+}/P_C < 0.4$	0.0665	0.0798	0.0886	0.0558
$0.4 \leq P_{a-}/P_C < 0.5$	0.0566	0.0492	0.0510	0.0323	$0.4 \leq P_{a+}/P_C < 0.5$	0.0587	0.0505	0.0591	0.0465
$0.5 \leq P_{a-}/P_C < 0.6$	0.0533	0.0262	0.0510	0.0461	$0.5 \leq P_{a+}/P_C < 0.6$	0.0436	0.0528	0.0295	0.0744
$0.6 \leq P_{a-}/P_C < 0.7$	0.0360	0.0377	0.0348	0.0507	$0.6 \leq P_{a+}/P_C < 0.7$	0.0393	0.0387	0.0364	0.0279
$0.7 \leq P_{a-}/P_C < 0.8$	0.0450	0.0262	0.0302	0.0230	$0.7 \leq P_{a+}/P_C < 0.8$	0.0417	0.0246	0.0227	0.0372
$0.8 \leq P_{a-}/P_C < 0.9$	0.0443	0.0324	0.0209	0.0415	$0.8 \leq P_{a+}/P_C < 0.9$	0.0369	0.0164	0.0273	0.0279
$0.9 \leq P_{a-}/P_C < 1$	0.0443	0.0460	0.0371	0.0507	$0.9 \leq P_{a+}/P_C < 1$	0.0472	0.0352	0.0205	0.0186

(a) Pluie P_{a-} du pas de temps précédent (b) Pluie P_{a+} du pas de temps suivant

Tableau VI.1 – Modèle probabiliste des pluies adjacentes – Tableau de contingence pour l'Eyrieux à Beauvène (454 km²), période 1953-2005.

Correction de l'ajustement MEWP

Nous avons vu que la pluie spatiale estimée avec la méthode SPAZM pouvait sous-estimer les fortes pluies du fait de la dispersion des postes et une correction a été proposée. Comme présenté dans la section IV. 3, l'AOE (Area of Effect) de SPAZM est calculée selon la disposition des postes utilisés par SPAZM pour calculer la pluie de bassin. La valeur d'AOE trouvée est de 496 km² pour une surface de bassin de 454 km² (cercle bleu, figure VI.4). L'AOE étant faiblement plus forte que la surface du bassin, l'erreur d'abattement sur les fort quantiles réalisée par SPAZM à de forte chance d'être relativement faible. Il a été proposé d'évaluer cette erreur en considérant un abattement sur les écarts-types β . L'abattement de SPAZM β_{SPAZM} sur l'écart-type vaut 0.87. Une valeur de l'abattement réel β^* est alors estimé en considérant l'abattement réalisé par SPAZM et l'AOE de SPAZM :

$$\beta^* = 1 - \frac{\ln(S_{BV})}{\ln(AOE_{SPAZM})} \cdot (1 - \beta_{SPAZM}) = 1 - \frac{\ln(454)}{\ln(496)} \cdot (1 - 0.87) = 0.87 \quad (VI.1)$$

L'abattement réalisé par SPAZM étant égal à celui censé être observé, la correction à effectuer sur les quantiles de l'ajustement MEWP sur la pluie SPAZM est nulle pour ce bassin.

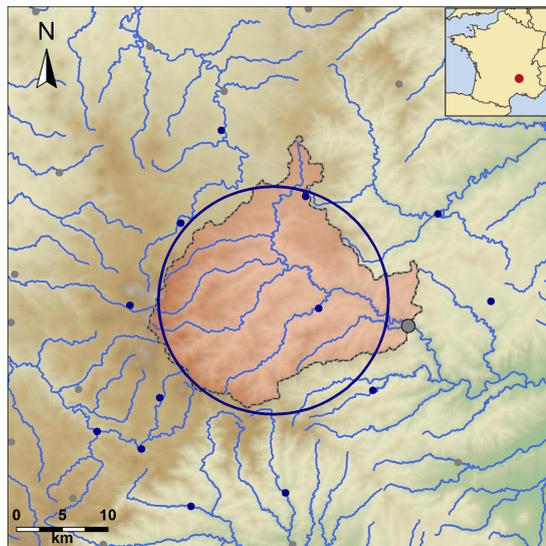


FIGURE VI.4 – L'Aire d'Effet (AOE) de SPAZM pour le bassin de l'Eyrieux à Beauvène (454 km²).

VI. 3 Passage pluie-débit en contexte non jaugé : Simulation EREBOR

Maintenant que tous les modèles probabilistes pour décrire l'aléa des précipitations ont été estimés, le passage à la distribution des débits avec la méthode EREBOR peut être envisagé. Pour cela les différents paramètres de la simulations doivent être déterminés. D'après la régionalisation des régimes de Sauquet *et al.*, le bassin de l'Eyrieux, dispose d'un régime hydrologique plutôt cévenol. Les paramètres utilisés pour la simulation EREBOR sont alors fixés :

- un hydrogramme unitaire $HU = \{20\%; 70\%; 10\%\}$;
- une fonction de répartition des états de saturation égale à la médiane de celle des distribution des bassins cévenols de la base d'étude SCHADEX ;
- aucune distribution des contributions neigeuses sur le bassin ;
- une capacité maximale de rétention S_{max} égale à celles des bassins cévenols du Chassezac à Malarce (484 km²) et à Sainte Marguerite (414 km²) qui sont proches du bassin de l'Eyrieux, de surface équivalente et avec une régime hydrologique assez similaire (figure VI.5). La capacité maximale de rétention $S_{max}(\text{Beauvène}) = 78\text{ mm}$ est prise égale à la moyenne de celle de ces deux bassins : $S_{max}(\text{Malarce}) = 82\text{ mm}$ $S_{max}(\text{St}^{\text{e}} \text{ Margerite}) = 74\text{ mm}$.

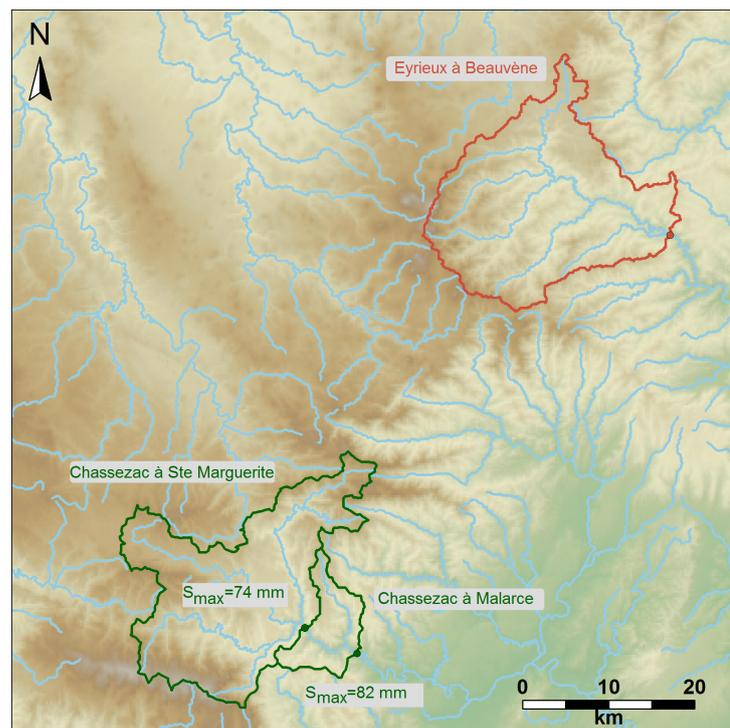
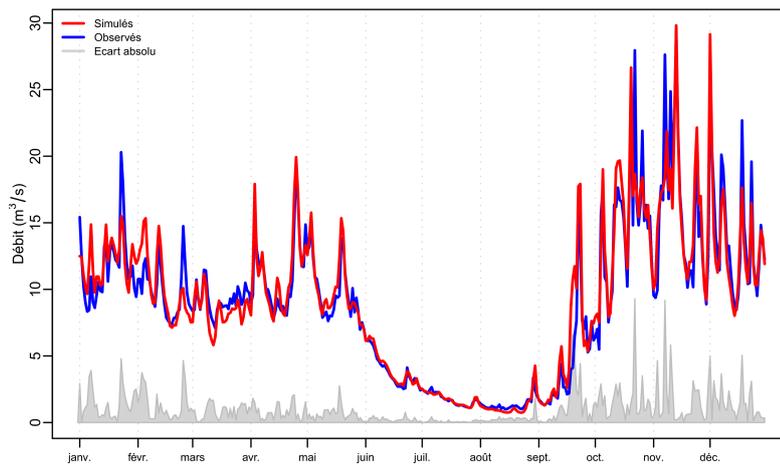


FIGURE VI.5 – Bassins similaires à l'Eyrieux à Beauvène (454 km²) (en rouge) – Chassezac à Malarce (484 km²) et à Sainte Marguerite (414 km²) (en vert).

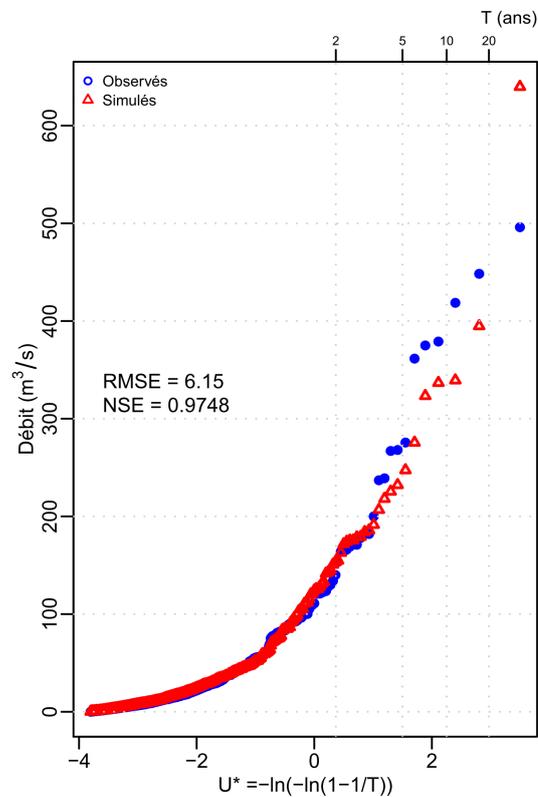
La simulation EREBOR permet alors d'obtenir une distribution des débits moyens journaliers (courbe rouge, figure VI.5). Afin de contrôler la pertinence de la distribution proposée, la distribution empirique estimée à partir de la série de débits de Beauvène est ajoutée. La simulation EREBOR est cohérente avec les fortes observations.

VI. 4 Passage pluie-débit en contexte jaugé : méthode SCHADEX

Afin de comparer la distribution de débits EREBOR à celle d'une étude SCHADEX, un modèle MORDOR a été calé avec les pluies et températures SPAZM et la série de débit à Beauvène sur la période 1972-2005. Le critère KGE entre les débits observés et simulés vaut 0.90 et modélise bien le régime de la rivière avec un KGE de 0.96 (figure VI.6 (a)). La distribution des débits journaliers classés simulés est cohérente avec celle des débits observés avec un KGE de 0.90.



(a) Régime des débits



(b) Débits classés

FIGURE VI.6 – Modèle MORDOR pour l'Eyrieux à Beauvène (454 km²).

Une simulation SCHADEX est ensuite réalisée avec les modèles probabilistes des pluies construits auparavant et le modèle MORDOR du bassin et permet d'obtenir une distribution de débits journaliers (courbe verte, figure VI.5).

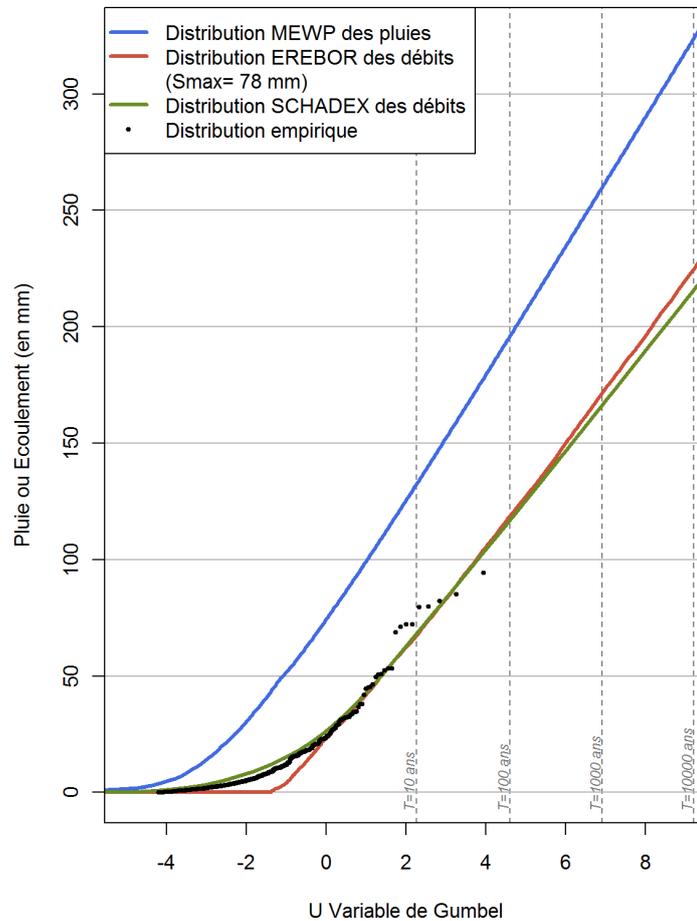


FIGURE VI.7 – Distributions des débits moyens journaliers pour l'Eyrieux à Beauvène (454 km²).

Bilan : la distribution de débits journaliers SCHADEX obtenue en utilisant toute l'information disponible sur le bassin et celle obtenue avec EREBOR sans aucune information de débit sont cohérentes entre elles et respectent la distribution empirique.

Temps de retour	SCHADEX	EREBOR	Écart relatif
10 ans	67	68	1.4%
100 ans	119	117	1.6%
1000 ans	172	166	3.4%
10 000 ans	225	216	4.0%

Tableau VI.2 – Les écoulements (en mm) pour différents temps de retour estimés avec la méthode SCHADEX et la méthode EREBOR.

Troisième partie

Conclusion générale

En reprenant le cheminement méthodologique présenté au début de ce manuscrit (figure VI.8), nous proposons de conclure en faisant un bilan sur les réponses apportées (partie droite du schéma) aux problématiques initiales (partie gauche) ainsi que sur les perspectives futures.

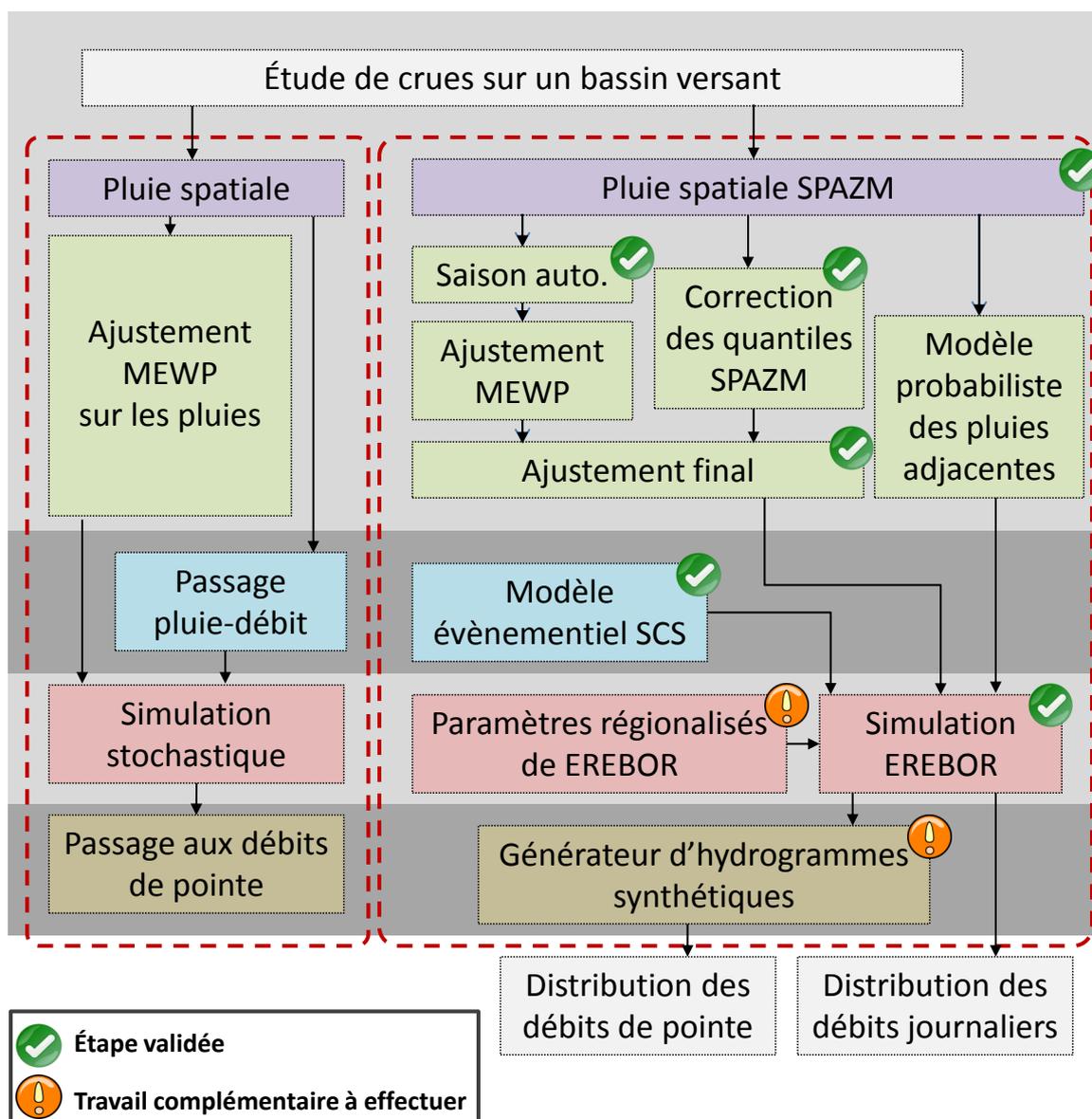


FIGURE VI.8 – Synthèse des résultats.

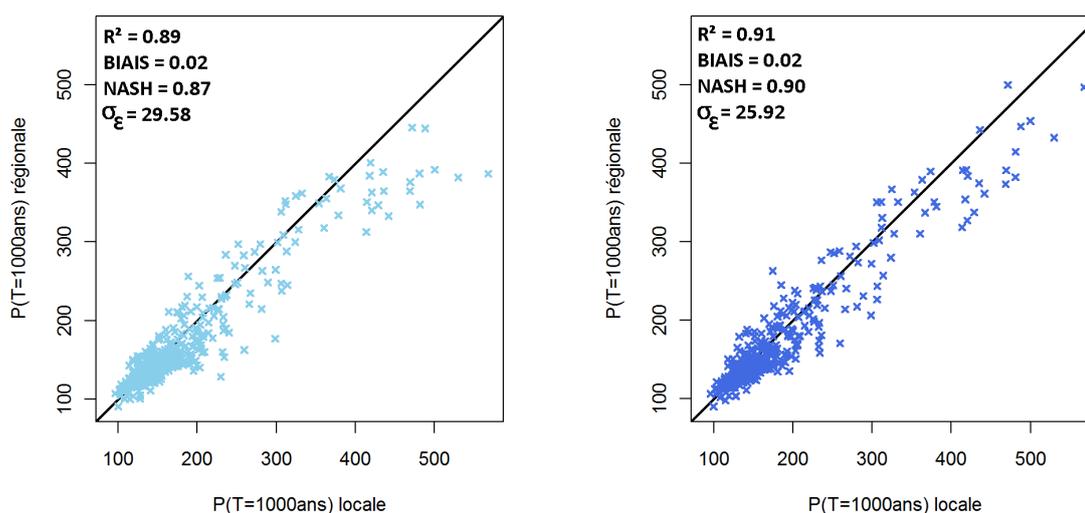
Les pluies extrêmes ponctuelles en contexte non jaugé

Dans un premier temps, nous nous sommes intéressés à une régionalisation des pluies journalières extrêmes. Les méthodes classiques de régionalisation utilisant un krigeage des pluies de temps de retour T ont montré leurs limites. En effet, elles sont peu robustes et leur performance chute rapidement dès que la densité de pluviomètres disponibles diminue. Elles ont notamment une forte tendance à sous-estimer les plus fortes valeurs. Un moyen pour remédier à ce problème est de prendre en compte des variables explicatives permettant une meilleure description de la structure du champ des pluies de temps de retour T . Le relief est la première variable qui peut venir à l'esprit. Cependant, bien que cette variable soit pertinente pour expliquer les cumuls annuels de précipitations, son utilisation pour décrire la structure spatiale des pluies extrêmes est complexe. **L'identification de variables pertinentes est délicate car elles sont souvent peu représentatives du risque de pluie final, qui mélange différents processus selon les régions et le type de situation météorologique. La classification des pluies par type de temps est alors un moyen de résoudre ce problème.** En s'intéressant non plus au comportement des quantiles de pluies extrêmes issus d'un ajustement global, mais au risque de pluie au sein de chaque type de temps, il est possible de faire ressortir un risque de pluie bien différent entre les types de temps et selon les régions. Nous avons ainsi pu mettre en évidence un lien fort entre les flux climatologiques de chaque type de temps et les régions de fort aléa pluviométrique concernées par ces flux. Les flux méditerranéens touchent logiquement le pourtour méditerranéen et sont responsables de fortes précipitations sur ce territoire. Les flux atlantiques concernent quant à eux plutôt les reliefs exposés aux flux d'ouest comme le plomb du Cantal, le puy de Sancy et les Pyrénées occidentales, mais avec une intensité moindre que pour les flux méditerranéens. Cette décomposition du risque en sous-populations ayant chacune leur propre structure facilite l'identification de variables explicatives pour la régionalisation. Ainsi, les cumuls annuels de pluie par type de temps s'avèrent pertinents pour représenter les structures spatiales du risque de pluie (notamment des gradex) par type de temps.

Un autre apport significatif pour améliorer la régionalisation des pluies extrêmes est l'utilisation des réanalyses de pluies SPAZM. Cet interpolateur de pluies utilise le maximum de données journalières disponibles et s'appuie sur des structures par type de temps afin de construire des champs de pluies sur toute la partie sud de la France. Un autre avantage non négligeable de SPAZM est l'utilisation d'un gradient altimétrique local par type de temps pour extrapoler le comportement des précipitations avec le relief. Au vu du constat fait sur la structure spatiale des pluies extrêmes, ces trois particularités font de SPAZM un outil précieux pour la spatialisation de pluies extrêmes : **SPAZM exprime un lien entre l'altitude et les précipitations qui dépend à la fois de la région et du type de temps.** L'utilisation des séries journalières de pluie SPAZM pour réaliser un ajustement MEWP des pluies, ou l'utilisation des cumuls annuels par type de temps de SPAZM comme variables explicatives, permettent d'améliorer l'estimation des quantiles extrêmes de pluies. La figure VI.9 illustre le gain de performance obtenu en intégrant l'information sur les cumuls annuels par type de temps SPAZM.

Après évaluation, un ajustement MEWP sur les pluies journalières SPAZM a donc été privilégié pour estimer la distribution des pluies extrêmes journalières. Les performances des estimations régionales ainsi obtenues sont tout à fait acceptables moyennant une majoration moyenne de 10% des quantiles de pluies de temps de retour supérieur à 10 ans. Nous avons alors construit des cartes de pluies aux temps de retour 10, 100, 1000 et 10 000 ans à une résolution de 1 km². Ces cartes pourront constituer une aide à la décision pour appréhender la structure du risque de pluie sur un bassin versant, mais également un support pour des travaux futurs.

Plusieurs pistes d'améliorations sont d'ailleurs envisageables. La pluie SPAZM décrit une structure cohérente du risque de pluie, mais sous-estime souvent les forts cumuls journaliers du fait de son processus d'interpolation. Un modèle de correction des quantiles de pluies journalières SPAZM, s'appuyant sur une carte des écarts-types des pluies construite par krigeage,



(a) Sans les cumuls annuels par type de temps SPAZM (b) Avec les cumuls annuels par type de temps SPAZM

FIGURE VI.9 – Illustration de la sous-estimation des plus fortes valeurs pour la pluie millénaire (avec la méthode de régionalisation des paramètres MEWP).

propose un facteur correctif sur les quantiles en fonction du ratio de l'écart-type des pluies SPAZM et de l'écart-type krigé. Améliorer ce krigeage des écarts-types serait un moyen simple pour améliorer les estimations des pluies de temps de retour avec SPAZM.

La totalité de ce travail a porté sur les précipitations journalières. Un travail exploratoire pour une estimation des précipitations extrêmes aux pas de temps infra-journaliers a également été mené. Il a mis en évidence l'intérêt de définir une loi composite régionale par type de temps avec des lois de Pareto généralisées comme lois marginales et de redéfinir une classification par type de temps pour le pourtour méditerranéen. Ces pistes de développement devraient être abordées dans le cadre d'un projet de recherche avec le LTHE¹ afin de généraliser le travail de cartographie des pluies extrêmes effectué en journalier à des pas de temps infra-journaliers et de vérifier la cohérence entre les résultats. L'utilisation de données RADAR, avec les réanalyses Météo France, pourrait notamment permettre d'étudier les risques orageux locaux qui ne sont pas toujours perçus par un réseau de pluviomètres.

La construction de pluies spatiales pour l'estimation des pluies extrêmes d'un bassin versant

La construction de pluies spatiales représentatives pour un bassin versant est un problème délicat, car la pluie spatiale réelle n'est pas connue. **Une étude comparative entre une pluie spatiale SPAZM et celle évaluée avec les polygones de Thiessen a été menée. Elle a permis de conclure que la correction de la pluie Thiessen par un facteur multiplicatif, afin de ramener son cumul annuel à celui de SPAZM, est dans 80% des cas justifiée et ne dégrade pas l'estimation des extrêmes.** Cependant, pour beaucoup de petits bassins, avec peu d'information à proximité, la correction appliquée au volume de la pluie Thiessen n'est pas adaptée à l'estimation des extrêmes. Ceci est en partie expliqué par le fait que pour des petits bassins, le risque de construire une pluie spatiale qui ne soit pas représentative des pluies du bassin est plus important (aucun poste disponible en altitude...). La correction des volumes, parfois importante, génère alors des estimations de quantiles de pluies incohérentes.

1. Laboratoire d'étude des Transferts en Hydrologie et Environnement.

Le diagnostic sur les différences constatées entre les deux pluies spatiales est assez dépendant de la gamme de surface. Pour des bassins de plus de 700 km², la densité du réseau de pluviomètres permet de construire des pluies spatiales cohérentes entre les deux méthodes. Avec un panier de 10 postes, les pluies SPAZM et Thiessen fournissent des estimations des pluies extrêmes comparables. Pour des bassins plus petits et moins bien instrumentés, les écarts entre les estimations SPAZM et Thiessen sont alors très variables et montrent en moyenne une sous-estimation de SPAZM par rapport à Thiessen. **Les pluies Thiessen doivent être considérées avec précautions pour des petits bassins car elles sont très incertaines et dépendantes des postes utilisés. Cette incertitude s'ajoute à celle sur la correction du cumul annuel. L'utilisation d'une pluie SPAZM sera donc préférable dans ce cas puisqu'elle aura un caractère plus prévisible et une pluie moyenne annuelle validée. SPAZM, en l'absence d'observation en altitude, tient compte d'un effet du relief et extrapole au mieux avec l'information disponible.**

Une étude comparative de l'abattement généré par SPAZM et de celui estimé par le simulateur de champs de pluies SAMPO a mis en évidence que SPAZM engendre un abattement souvent trop fort par rapport à la réalité du fait de la dispersion de l'information utilisée pour son interpolation. Comme pour les pluies ponctuelles, la pluie SPAZM sous-estime souvent la variabilité des pluies. Le critère AOE (Area of Effect) a été défini afin d'évaluer une surface de référence pour la pluie spatiale SPAZM. Cette surface est représentative de l'information pluviométrique utilisée pour la construction de la pluie spatiale SPAZM (figure VI.10). La conclusion de ce travail est que **la sous-estimation des forts quantiles de pluies avec SPAZM est très corrélée à la dispersion des postes autour du bassin. La prise en compte de l'aire AOE et de la surface du bassin est un moyen efficace de corriger la sous-estimation des quantiles extrêmes de la pluie SPAZM.**

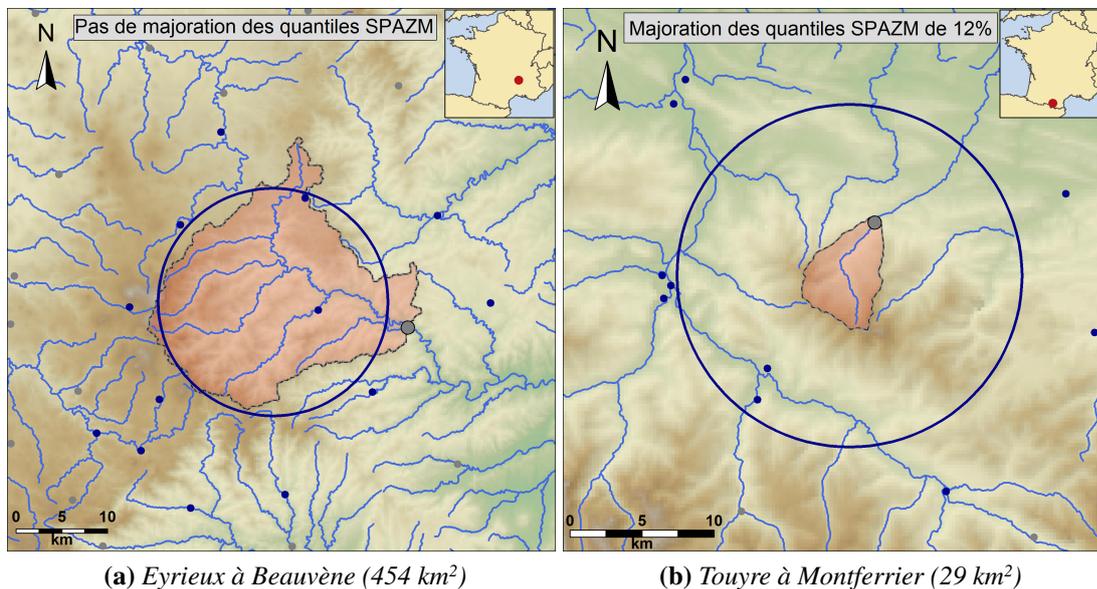


FIGURE VI.10 – Exemple de l'AOE (aire du disque bleu) pour deux bassins versants.

À l'avenir, l'étude de l'abattement spatial des précipitations pourra bénéficier de la mise à disposition de chroniques de lame d'eau radar sur plusieurs dizaines d'années par Météo France. Ce jeu de données permettra d'étudier en détail le champ spatial des précipitations même si les estimations en zone d'altitude resteront incertaines.

Il est ainsi possible de fournir des séries de pluies spatiales journalières fiables pour un bassin versant non jaugé. Cependant, la réalisation d'un ajustement MEWP sur ces séries est

très sensible à la saisonnalité des fortes précipitations choisie pour réaliser cet ajustement. Un algorithme a permis d'automatiser le processus de sélection des saisons. Une évaluation, sur la base des ajustements de référence des études SCHADEX disponibles au pas de temps journalier, a validé la capacité de ce processus à identifier un découpage saisonnier ne sous-estimant pas la distribution des pluies extrêmes.

Un passage pluie-débit régional

À cette étape de la procédure, une série journalière de pluie spatiale ainsi qu'un modèle probabiliste pour les pluies centrales fiables peuvent être utilisés pour décrire l'aléa pluie. À partir de la série de pluies, il est également possible de construire un modèle probabiliste pour les pluies adjacentes aux pluies centrales. L'aléa pluviométrique est donc complètement décrit.

Pour passer à la distribution des débits, nous nous sommes intéressés aux processus hydrologiques du bassin pouvant conditionner le passage d'une distribution des pluies journalières à une distribution des débits moyens journaliers. Dans les études SCHADEX, le passage des pluies aux débits est réalisé par un modèle hydrologique MORDOR. Une régionalisation de MORDOR étant un problème trop complexe, un modèle pluie-débit SCS a été utilisé. Ce modèle événementiel simple et couramment utilisé est adapté à la simulation d'épisodes d'écoulements indépendants. Il permet de transformer un scénario de pluie en un scénario d'écoulements journaliers à l'aide d'un nombre de paramètres restreint et en grande partie contraints en fonction du régime hydrologique du bassin. Un état de saturation du bassin peut être défini pour chaque épisode, l'épisode d'écoulement ainsi généré est issu d'un croisement entre un épisode de pluie et un état du bassin. **Un processus de simulation, dénommé EREBOR (Écoulements Régionaux Extrêmes sur un Bassin à Observations Réduites), permet alors de générer une distribution de débits en réalisant une simulation stochastique analogue à celle de SCHADEX, mais en utilisant le modèle SCS.** Le paramètre clé de ce modèle est la valeur de la capacité maximale de rétention choisie pour le bassin. Ce paramètre contient une information cohérente avec celle des stocks pour le ruissellement rapide du modèle MORDOR et fonde donc son sens hydrologique.

Faute de suffisamment de données, ce paramètre n'a pu être régionalisé de façon à pouvoir appliquer la méthode sur n'importe quel bassin non jaugé. La base d'études utilisée pour la régionalisation est constituée des études SCHADEX ayant fait l'objet d'une expertise. Cette base permet de disposer de données contrôlées mais réduit considérablement le nombre de bassins pour la régionalisation. Un travail préliminaire pour agrandir la base d'études SCHADEX a donc été engagé, ce qui implique d'automatiser la méthode SCHADEX. En effet, une étude de crues extrêmes réalisée avec la méthode SCHADEX est le résultat de mois de travail garantissant des choix les plus fiables et robustes possibles en cohérence avec la connaissance qu'acquière les chargés d'études sur le bassin concerné. Reproduire les choix réalisés, en particulier sur le choix d'un modèle MORDOR avec un comportement adapté pour l'étude de crues, a demandé de définir des critères de sélection appropriés. Une première proposition a été faite pour automatiser le processus de simulation SCHADEX, mais une validation doit encore être pratiquée sur une base élargie de bassins avec des critères objectifs pour évaluer la sensibilité des résultats à chaque étape. Une étude SCHADEX automatisée ne remplacera jamais une expertise humaine, mais devrait produire des résultats fiables, objectifs et homogènes pour un grand nombre de bassins. Ce sera également un moyen de produire rapidement une première estimation pouvant servir de point de départ à une expertise approfondie.

Les travaux futurs devront permettre d'exploiter le protocole d'automatisation de la procédure SCHADEX afin de disposer d'une large base de distributions de débits SCHADEX sur tout le territoire. Cette base d'études SCHADEX sera indispensable pour régionaliser convenablement les paramètres de la méthode régionale EREBOR et la valider à grande échelle.

La distribution du débit de pointe

La distribution du débit de pointe est généralement la dernière étape d'une étude de crues extrêmes. Que ce soit avec la méthode SCHADEX ou avec la méthode EREBOR, la sortie est une distribution des débits moyens journaliers constituée d'un très grand nombre de crues simulées, c'est-à-dire de séquences de débits journaliers. **Une méthode de génération d'hydrogrammes synthétiques a été développée à EDF-DTG. Elle permet d'associer un hydrogramme horaire à un scénario d'écoulement journalier (scénarios disponibles en sortie de simulation SCHADEX et EREBOR).** Une archive régionale d'hydrogrammes horaires a été constituée à partir d'observations provenant de différents sites en France. Une recherche de candidats sur la base d'une analogie des séquences de débits journaliers permet d'identifier des formes d'hydrogrammes cohérentes avec les volumes et de dynamiques variées. La forme médiane de ces hydrogrammes candidats fournit alors un hydrogramme synthétique pour chaque scénario de débits journaliers généré. Le débit de pointe associé à chaque scénario permet alors de construire une distribution des débits de pointe. Cette méthode purement régionale a été validée sur une centaine de stations de débit et affiche des performances équivalentes voire meilleures par rapport à la méthode du coefficient de forme moyen utilisée actuellement (qui est une approche locale donc inutilisable sur des bassins non jaugés). L'utilisation de cette méthode sur les scénarios d'écoulements EREBOR devra cependant faire l'objet d'une évaluation.

Bilan général

Les développements proposés dans cette thèse constituent une approche pour l'étude des événements hydrologiques extrêmes en contexte non jaugé. Comme dans la méthode SCHADEX utilisée sur les bassins jaugés, il a été proposé de s'appuyer sur la description de l'aléa des précipitations et sur un modèle pluie-débit afin d'estimer la distribution des débits journaliers extrêmes. Avec cette approche régionale, une estimation des pluies et débits (journaliers et de pointe) extrêmes est réalisable en contexte non jaugé.

Plusieurs axes de recherche sont nécessaires pour terminer et approfondir le travail réalisé :

- la construction d'une large base d'études SCHADEX sera utilisée pour améliorer la régionalisation des paramètres de la méthode EREBOR ;
- la totalité des étapes de l'approche régionale (des pluies aux débits de pointe) pourront être évaluées sur cette large base de bassins versants au pas de temps journalier ;
- un travail similaire mais pour des pas de temps infra-journaliers devra également être mené afin de fournir des cartes de pluies extrêmes et des estimations de distributions de débits à des pas de temps plus fins.

Annexes

Le modèle hydrologique MORDOR

Paramètre	Description	Valeur
Paramètres caractéristiques du bassin versant		
fp1 à fp4	Coefficient de la fonction arc-tangente pour le calcul de la fraction liquide des précipitations	Calculé
fnts0	Fraction du bassin en glace ou enneigée en permanence	Calculé
Paramètre relatif aux précipitations		
cp	Coefficient correctif appliqué à la précipitation mesurée	Calculé
Paramètres relatifs à la température		
ptjp1	Poids de la température du jour J+1 dans l'éclatement des précipitations	Calé
pnfr	Pondération donnée à la neige la plus froide (neige du jour ou stock) pour le calcul de la température du stock	Calé
prft	Pondération de l'influence de la température du jour sur la température de fonte	Calé
dbt	Bonus de température « de beau temps » pour le calcul de la fonte	Calé
ltm	Pondération des températures antérieures pour le calcul de la température de fonte superficielle	Calé
efp	Décalage de la température de référence avant éclatement des précipitations	Calé
eft	Décalage de la température de référence avant calcul de la lame de fusion	Calé
Paramètres relatifs à la fusion du manteau neigeux		
kfs	Coefficient de fusion sous le manteau	Calé
es	Exposant réglant l'évolution de la surface enneigée lors de la fusion	Calé
ztn	Portion de la lame de fonte qui ne ruisselle pas directement interceptée par le stock superficiel	Calé
kf	Coefficient brut de fusion superficielle	Calé
Paramètres relatifs à l'évapotranspiration		
fe1	Premier paramètre de l'évapotranspiration potentielle	Calé
fe2	Second paramètre de l'évapotranspiration potentielle	2
fe3	Troisième paramètre de l'évapotranspiration potentielle	Calé
Paramètres relatifs aux échanges souterrains (nappes)		
kl1	Coefficient de vidange du stock intermédiaire vers le stock évaporant ou la rivière	Calé
kl2	Coefficient de vidange du stock intermédiaire vers le stock évaporant ou le stock profond	Calé
dn	Coefficient de vidange du stock profond	Calé
exn	exposant de la loi de vidange du stock profond	Calé
Paramètres relatifs à la fonction de transfert du bassin		
ftr1	Premier paramètre de la fonction de transfert du bassin	Calé
ftr2	Second paramètre de la fonction de transfert du bassin	Calé
Paramètres relatifs aux niveaux de remplissage des stocks		
Umax	Valeur maximale du stock superficiel	Calé
Lmax	Valeur maximale du stock intermédiaire	Calé
Zmax	Valeur maximale du stock évaporant	Calé

Tableau A.1 – Paramètres du modèle hydrologique MORDOR.

Bibliographie

- H. ALEXANDERSSON : A homogeneity test applied to precipitation data. *Journal of Climatology*, 6:661–675, 1986.
- P. ALLAMANO, F. LAIO et P. CLAPS : Effects of disregarding seasonality on the distribution of hydrological extremes. *HESS*, 15 (10):3207–3215, 2011.
- V. ANDRÉASSIAN, S. BERGSTROM, N. CHAHINIAN, Q. DUAN, Y.M. GUSEV, I. LITTLEWOOD, T. MATHEVET, C. MICHEL, A. MONTANARI, G. MORETTI, R. MOUSSA, O.N. NASONOVA, K.M. O’CONNOR, E. PAQUET, C. PERRIN, A. ROUSSEAU, J. SCHAAKE, T. WAGENER et Z. XIE : Catalogue of the models used in mopex 2004/2005. *IAHS Publication*, 307,:41–93, 2006.
- ANR EXTRAFL0 III.1 : Tâche III : Inter-comparaison des méthodes probabilistes – rapport III.1 «Comparaison des méthodes locales pour l’estimation des pluies extrêmes». 2013.
- ANR EXTRAFL0 III.3 : Tâche III : Inter-comparaison des méthodes probabilistes – rapport III.3 «Comparaison des méthodes purement régionales pour l’estimation des pluies extrêmes». 2013.
- ANR EXTRAFL0 III.5 : Tâche III : Inter-comparaison des méthodes probabilistes – rapport III.5 «Comparaison des méthodes locales-régionales pour l’estimation des pluies extrêmes». 2013.
- ANR EXTRAFL0 V : Tâche V : Diagnostic sur les méthodes de prédétermination – Rapport V «Domaines d’application des méthodes de prédétermination des pluies et crues extrêmes». 2013.
- P. ARNAUD : *Modèle de prédétermination de crues basé sur la simulation*. Thèse de doctorat, Université Montpellier II - Science et Technique du Languedoc, 1997.
- P. ARNAUD, J.A. FINE et J. LAVABRE : An hourly rainfall generation model applicable to all types of climate. *Atmospheric Research*, 85:230–242, 2007.
- Y. AUBERT : *Estimation des valeurs extrêmes de débit par la méthode Shyreg : réflexions sur l’équifinalité dans la modélisation de la transformation pluie en débit*. Thèse de doctorat, Paris 6, 2012.
- Y. AUBERT, P. ARNAUD, P. RIBSTEIN et J. FINE : La méthode SHYREG débit—Application sur 1605 bassins versants en France métropolitaine. *Hydrological Sciences Journal*, 59(5):993–1005, 2014.
- J. BALES et R. BETSON : *The curve number as an hydrologic index*. Water Resouce Publications, 1982.
- A. BARDOSSY et J. LI : Geostatistical interpolation using copulas. *Water Ressources Recherche*, 44: W07412, 2008.
- A. BEN DAOUD : *Améliorations et développements d’une méthode de prévision probabiliste des pluies par analogie. Application à la prévision hydrologique sur les grands bassins fluviaux de la Saône et de la Seine*. Thèse de doctorat, Université de Grenoble, 2010.

- G. BENITO et V. R. THORNDYCRAFT : Palaeoflood hydrology : insight into rare events and extreme flood discharges. *La Houille Blanche*, 5:91–96, 2006.
- K. J. BEVEN : *Rainfall-Runoff modelling The Primer*. British Library, 2003.
- J. BLANCHET et M. LEHNING : Mapping snow depth return levels : smooth spatial modeling versus station interpolation. *Hydrology and Earth System Sciences Discussions*, 7(4):6129–6177, 2010.
- J. BOE et L. TERRAY : A weather-type approach to analyzing winter precipitation in france : Twentieth-century trends and the role of anthropogenic forcing. *Journal of Climate.*, 21(13):3118–3133, 2008.
- P. BOIS : *Une méthode de contrôle de séries chronologiques utilisées en climatologie et en hydrologie*. Publication du Laboratoire de Mécanique des Fluides, 1971.
- G. BONI, A. PARODI et F. SICCARDI : A new parsimonious methodology of mapping the spatial variability of annual maximum rainfall in mountainous environments. *Journal of Hydrometeorology*, 9:492 – 506, 2008.
- G. BONTRON : *Prévision quantitative des précipitations : adaptation probabiliste par recherche d'analogues - utilisation des réanalyses NCEP/NCAR et application aux précipitations du sud-est de la France*. Thèse de doctorat, INPG, Janvier 2004.
- P. BRIGODE : *Changement climatique et risque hydrologique : évaluation de la méthode SCHADEX en contexte non-stationnaire*. Thèse de doctorat, Université Pierre et Marie Curie, 2013.
- P. BRIGODE, P. BERNARDARA, E. PAQUET, J. GAILHARD, F. GARAVAGLIA, R. MERZ, Z. MIĆOVIĆ, D. LAWRENCE et P. RIBSTEIN : Sensitivity analysis of schadex extreme flood estimates to observed hydro-meteorological variability. *Water Resources Research*, 50:353–370, 2014.
- P. CANTET et P. ARNAUD : Extreme rainfall analysis by a stochastic model : impact of the copula choice on the sub-daily rainfall generation. *Stochastic Environmental Research and Risk Assessment*, 28:1479–1492, 2014.
- A. CASTELLARIN, R. VOGEL et N. MATALAS : Multivariate probabilistic regional envelopes of extreme floods. *Journal of Hydrology*, 336(3-4):376 – 390, 2007.
- F. CERNESSON : *Modèle simple de prédétermination des crues de fréquences courantes à rares sur petits bassins versants méditerranéens*. Thèse de doctorat, Montpellier II, 1993.
- CFGB : Design flood determination by the gradex method. *Bulletin du Comité Français des Grands Barrages-FRCOLD News 96*, 18th congress CIGB-ICOLD n2, nov:108, 1994.
- D. COOLEY, D. NYCHKA et P. NAVEAU : Bayesian spatial modeling of extreme precipitation return levels. *Journal of the American Statistical Association*, 102(479):824–840, 2007.
- CTGREF, SRAE et S.H. DIAME : Synthèse nationale sur les crues des petits bassins versants. fascicule 2 : la méthode socose ; fascicule 3 : la méthode crupedix. Rapport technique, 1980-1982.
- T. DALRYMPLE : Flood frequency analysis. *USGS, Water-Supply Papers:1543A*, 1960.
- C. DALY, M. HALBLEIB, J. I. SMITH, W. P. GIBSON, M. K. DOGGETT, G. H. TAYLOR, J. CURTIS et P. P. PASTERIS : Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous united states. *International Journal of Climatology*, 28(15):2031–2064, 2008.
- C. DALY, R. NEILSON et D. PHILLIPS : A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *Journal of Applied Meteorology*, 33:140–158, 1994.
- J.P. DELHOMME : Application de la théorie des variables régionalisées dans les sciences de l'eau. *Bulletin du B.R.G.M (deuxième série)*, Section III (4):341–375, 1978.

- I. DESUROSNE : *Gradients d'intensité de pluie en zone de relief : expérimentations et premières modélisations des données d'un réseau transalpin, le TPG*. Thèse de doctorat, Université Louis Pasteur, septembre 1992.
- I. DESUROSNE, G. OBERLIN et G. BLANCHET : Towards a better knowledge of high basin rainfall fields at small time and space steps : analysis of rainfall data from a unidirectional northern pre-alps network. *IAHS Publications-Series of Proceedings and Reports-Intern Assoc Hydrological Sciences.*, 221:323–342, 1994.
- L. DEZILEAU, B. TERRIER, J.F. BERGER, P. BLANCHEMANCHE, A. LATAPIE, R. FREYDIER, L. BREMOND, A. PAQUIER, M. LANG et J.L. DELGADO : A multivariate approach applied to historical slackwater flood deposits of the Gardon river, {SE} France. *Geomorphology*, 214(0):56 – 68, 2014.
- A. DJERBOUA et M. LANG : Scale parameter of maximal rainfall distribution : comparison of three sampling techniques. *Revue des sciences de l'eau*, 20:111–125, 2007.
- M. H. DOMINGUES RAMOS : *Analyse de la pluviométrie sous des systèmes nuageux convectifs*. Thèse de doctorat, UJF - Grenoble, 2002.
- V. DOUCHEZ : *Modèle probabiliste de pluies extrêmes : Prise en compte d'un caractère hyper-exponentiel par type de temps.*, 2013.
- D. DUBAND : Evaluation des risques de fortes précipitations sur les régions alpines françaises selon la saison - utilisation du gradex. *AISH Red Book*, 106:383–395, 1973.
- D. DUBAND, C. MICHEL, H. GARROS et J. ASTIER : Evaluation des crues extrêmes et de la crue de projet par la méthode gradex. *San Francisco, CIGB, 16^e congrès*, 1988.
- M. EBTEHAJ, H. MORADKHANI et H. GUPTA : Improving robustness of hydrologic parameter estimation by the use of moving block bootstrap resampling. *Water Resources Research*, 46 (7):W07515, 2010.
- I. EMMANUEL : *Évaluation de l'apport de la mesure de pluie par radar météorologique pour la modélisation pluie - débit de petits bassins versants*. Thèse de doctorat, Ifstar, 2011.
- A.C. FAVRE, S. EL ADLOUNI, L. PERREAULT, N. THIÉMONGE et B. BOBÉE : Multivariate hydrological frequency analysis using copulas. *Water Resources Research*, 40, 2004.
- J. FRANCOU et J.A. RODIER : Essai de classification des crues maximales observées dans le monde. *Cahier de l'Orstom, série hydrologie*, IV (3), 1967.
- G. GALEA, Michel C. et G. OBERLIN : Pluies de bassin - abattement sur une surface des averses de 1 h à 24 h, 1982.
- F. GARAVAGLIA : Classification en type de temps d'EDF : comparaison entre les versions 2006 et 2010 de la classification. Rapport technique, 2011a.
- F. GARAVAGLIA : *Méthode SCHADEX de prédétermination des crues extrêmes. Méthodologie, applications, études de sensibilité*. Thèse de doctorat, Université de Grenoble, 2011b.
- F. GARAVAGLIA, J. GAILHARD, E. PAQUET, M. LANG, R. GARÇON et P. BERNARDARA : Introducing a rainfall compound distribution model based on weather patterns sub-sampling. *Hydrology and Earth System Sciences*, 14(6):951–964, 2010.
- R. GARÇON : Statistical and bayesian methods in hydrological sciences. *In A joint UNESCO International Conference in honour of Jacques Bernier*, Septembre 11-13 1995. Communication orale.
- R. GARÇON : Prévision opérationnelle des apports de la Durance à Serre-Ponçon à l'aide du modèle MORDOR. *La Houille Blanche*, 5:71 – 76, 1996.

- R. GARÇON : Modélisation des crues : retour d'expérience sur le gradex, 2000.
- R. GARÇON, C. CARRÉ et P. LYAUDET : Exemple de prévision et de simulation opérationnelle des débits d'étiage pour les besoins d'edf. *La Houille Blanche*, -(6):37–42, 1999.
- G. GARRY, M. MASSON et J.-L. BALLAIS : Cartographie des zones inondables : approche hydro-géomorphologique. Rapport technique, Paris La Défense, ministère de l'Équipement, ministère de l'Environnement, Les Editions Ville et Territoires, 1996.
- J. GAUME, N. ECKERT, G. CHAMBON, M. NAAIM et L. BEL : Mapping extreme snowfalls in the french alps using max-stable processes. *Water Resources Research*, 49:1 – 20, 2013.
- F. GOTTARDI : *Estimation statistique et réanalyse des précipitations en montagne. Utilisation d'ébauches par type de temps et assimilation de données d'enneigement. Application aux grands massifs montagneux français*. Thèse de doctorat, INPG, 2009.
- F. GOTTARDI, C. OBLED, J. GAILHARD et E. PAQUET : Statistical reanalysis of precipitation fields based on ground network data and weather patterns : Application over french mountains. *Journal of Hydrology*, 432–433:154 – 167, 2012.
- S. GUILBAUD : *Prévision quantitative des précipitations journalières par une méthode statistico-dynamique des recherche d'analogues - Application à des bassins du pourtour méditerranéen*. Thèse de doctorat, ENSHMG - INPG, 1997.
- D. GUILLOT : Letter to the editor on "the gradex method". *Journal des Sciences Hydrologiques*, 26 (3):334–336, 1981.
- D. GUILLOT : The arguments of the gradex method : a logical support to assess extreme floods. *AISH Red Book*, -(213):287–298, 1993.
- D. GUILLOT et D. DUBAND : La méthode du gradex pour le calcul de la probabilité des crues à partir des pluies. In *Colloque international sur les crues et leur évaluation, Leningrad, 15 au 18 août, IASH*, 1967.
- P. GUILLOT : Debits et pluies extreme, 1979.
- H. V. GUPTA, H. KLING, K. K. YILMAZ et G. F. MARTINEZ : Decomposition of the mean squared error and nse performance criteria : Implications for improving hydrological modelling. *Journal of Hydrology*, 377(1-2):80 – 91, 2009.
- K.H. HAMED : Exact distribution of the Mann-Kendall trend test statistic for persistent data. *Journal of Hydrology*, 365(1-2):86 – 94, 2009.
- M. R. HAYLOCK, N. HOFSTRA, A. M. G. Klein TANK, E. J. KLOK, P. D. JONES et M. NEW : A european daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006. *Journal of Geophysical Research - Atmospheres*, 113(20):20119, 2008.
- B. HINGRAY, P. PICOUET et A. MUSY : *Hydrologie 2 : Une science pour l'ingénieur*. Presses polytechniques et universitaires romandes, 2009.
- A. T. HJELMFELT, L. A. KRAMER et R. BURWELL : *Curve numbers as randomness variables*. Water Resouce Publications, 1982.
- J. R. M. HOSKING et J. R. WALLIS : *Regional Frequency Analysis : An approach based on L-moments*. Cambridge University Press, 1997.
- A. F. JENKINSON : The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quarterly Journal of the Royal Meteorological Society*, 81(348):158–171, 1955.

- A.G. JOURNAL : *Géostatistique minière*, volume Tome I et II. Centre de Géostatistique, Paris, 1977.
- A.G. JOURNAL et C.J. HUIJBREGTS : Mining geostatistics. *Academic Press, London*, page 600, 1978.
- M. G. KENDALL : Rank correction methods. *London*, 1975.
- A. KIEFFER WEISS et P. BOIS : Topographic effect on statistical characteristics of heavy rainfall and mapping in the french alps. *Journal of Applied Meteorology*, 40:720–740, 2000.
- A. KIEFFER WEISS et P. BOIS : A comparison of methods for mapping statistical characteristic of heavy rainfall in the french alps : the use of daily information. *Hydrological Sciences Journal*, 47 (5):739–752, 2002.
- A. KUENTZ : Calcul du coefficient de forme des crues à l'aide du modèle QdF-convergent., 2009.
- A. KUENTZ : *Un siècle de variabilité hydro-climatique sur le bassin de la Durance. Recherches historiques et reconstitutions*. Thèse de doctorat, Institut des Sciences et Industries du Vivant et de l'Environnement (AgroPARisTech), 2013.
- P. KUMARASWANY : A generalized probability density functions for double-bounded random processes. *Journal of Hydrology*, 46:79–88, 1980.
- C. LABOUREAU : Transposition des paramètres d'un modèle hydrologique entre plusieurs bassins voisins, 2013.
- R. LAMB : Calibration of a conceptual rainfall-runoff model for flood frequency estimation by continuous simulation. *Water Resour. Res.*, 35:–, 1999.
- M. LANG : "Confidence interval of quantile estimates of the AGREGE model". *CEMAGREF*, pages 181 – 189, 1995. Groupe AMHY Séminaire annuel, septembre 1995.
- M. LANG et I. DESUROSNE : Esquisse des risques de crues a l'échelle euro-méditerranéenne : les premiers resultants du programme FRIEND-AMHY exploitant les modèles AGREGE et TPG. *23emes Journées de l'hydrauliques, Congrès SHF Crues et Inondations.*, 1:123 – 130, 1994.
- M. LANG et J. LAVABRE : *Estimation de la crue centennale pour les plans de prévention des risques d'inondations*. Update Sciences & technologies, 2007.
- M. LANG, T. B. M. J. OUARDA et B. BOBÉE : Towards operational guidelines for over-threshold modeling. *Journal of Hydrology*, 225(3-4):103 – 117, 1999.
- J. LAVABRE, C. FOUCHIER, N. FOLTON et Y. GREGORIS : SHYREG : une méthode pour l'estimation régionale des débits de crue. application aux régions méditerranéennes françaises. *Ingénierie EAT*, n°spécial Risque naturel et aménagement du territoire:97–111, 2003.
- D. LAWRENCE, E. PAQUET, J. GAILHARD et A.K. FLEIG : Stochastic semi-continuous simulation for extreme flood estimation in catchments with combined rainfall–snowmelt flood regimes. *Natural Hazards and Earth System Sciences*, 14:1283–1298, 2013.
- S. LE CLERC : *Revisiter la notion de scénario hydrologique de référence pour la caractérisation des inondations*. Thèse de doctorat, Université Joseph Fourier de Grenoble, 2004.
- E. LEBLOIS et J.D. CREUTIN : Space-time simulation of intermittent rainfall with prescribed advection field : Adaptation of the turning band method. *Water Resources Research*, 49(6):3375–3387, 2013.
- E.L. LEHMANN : Nonparametrics, statistical methods based on ranks. *California*, 1975.
- J. M. LEPIOUFLE : *Modélisation spatio-temporelle d'un champ de pluie : application aux pluies journalières du bassin de la Loire*. Thèse de doctorat, CEMAGREF - Institut Polytechnique de Grenoble, 2009.

- K.C. LI : Sliced inverse regression for dimension reduction. *Journal of the American Statistical Association*, 86:316–327, 1991.
- T. LITTMANN : An empirical classification of weather types in the mediterranean basin and their interrelation with rainfall. *Theoretical and Applied Climatology*, 66:161–171, 2000.
- A. MAILHOT, S. LACHANCE-CLOUTIER, G. TALBOT et A.C. FAVRE : Regional estimates of intense rainfall based on the peak-over-threshold approach. *Journal of Hydrology*, 476:188–199, 2013.
- H.B. MANN : Nonparametric tests against trend. *Econometrica*, 13:245–259, 1945.
- M. MARGOUM : *Estimation des crues rares et extrêmes : le modèle AGREGEE. Conception et premières validations*. Thèse de doctorat, Ecole des Mines de Paris, 1992.
- M. MARGOUM, G. OBERLIN, M. LANG et R. WEINGARTNER : Estimation des crues rares et extrêmes : principes du modèle AGREGEE. *Hydrologie Continentale*, 9 (1):83–98, 1994.
- C. MARTÍNEZ, J. CAMPINS, A. JANSÜ et A. GENOVÉS : Heavy rain events in the western mediterranean : an atmospheric pattern classification. *Advances in Science and Research*, 2:61–64, 2008.
- T. MATHEVET et R. GARÇON : Tall tales from the hydrological crypt : are models monsters ? *Hydrological sciences journal*, 55(6):857 – 871, 2010.
- S. MISHRA et V. P. SINGH : Another look at scs-cn method. *Journal of Hydrological Engineering*, ASCE 4:257–264, 1999.
- V. MOCKUS : Estimation of total (and peak rate of) surface runoff for individual storms. *US Department of Agriculture*, 1949.
- L. MOULIN, E. GAUME et C. OBLED : Uncertainties on mean areal precipitation : assessment and impact on streamflow simulations. *Hydrology and Earth System Sciences*, 13(8):99–114, 2009.
- J.E. NASH et J.V. SUTCLIFFE : River flow forecasting thought. part i. a conceptual models discussion of principles. *Journal of Hydrology*, 10:282–290, 1970.
- R. NAULET, M. LANG, T.B. OUARDA, D. COEUR, B. BOBÉE et D. RECKING, A. and Moussay : Flood frequency analysis on the Ardèche river using French documentary sources from the two last centuries. *Journal of Hydrology*, 313:58–78, 2005.
- L. NEPPEL, P. ARNAUD et J. LAVABRE : Connaissance régionale des pluies extrêmes. : Comparaison de deux approches appliquées en milieu méditerranéen. *Comptes Rendus Geosciences*, 339(13):820–830, 2007.
- C. OBLED, G. BONTRON et R. GARÇON : Quantitative precipitation forecasts : a statistical adaptation of model outputs through an analogues sorting approach. *Atmospheric Research*, 63(3-4):303 – 324, 2002.
- T.B.M.J. OUARDA, P.F. RASMUSSEN, B. BOBÉE et J. BERNIER : Utilisation de l'information historique en analyse hydrologique fréquentielle. *Revue des sciences de l'eau*, 11 (1):41–49, 1998.
- E. PAQUET : Critique et homogénéisation de séries climatologiques - Outil KRITIKOR. Rapport technique, Notice méthodologique pour EDF-DTG, 2005.
- E. PAQUET : Evaluation des crues extrêmes sur la Basse-Isère - Méthodologie et scénarios hydrologiques millénaires retenus, 2014.
- E. PAQUET, J. GAILHARD et R. GARÇON : Evolution de la méthode du gradex : approche par type de temps et modélisation hydrologique. *La Houille Blanche*, -(5):80–90, 2006.

- E. PAQUET, F. GARAVAGLIA, R GARÇON et J. GAILHARD : The schadex method : a semi-continuous rainfall-runoff simulation for extreme flood estimation. *Journal of Hydrology*, 2013.
- E. PAQUET et R. GARÇON : Modèle mordor : notice méthodologique. Rapport technique, Electricité de France, 2000.
- O. PAYRASTRE : *Faisabilité et utilité du recueil de données historiques pour l'étude des crues extrêmes de petits cours d'eau. Etude du cas de quatre bassins versants affluents de l'Aude*. Thèse de doctorat, Ecole Nationale des Ponts et Chaussées, 2005.
- D. PENOT : Génération de scénarios hydrologiques pour les crues de projet calculées par la méthode SCHADEX, 2011.
- T.C. PETERSON et D.R. EASTERLING : Creation of homogeneous composite climatological reference series. *International Journal of Climatology*, 14(6):671–679, Juillet 1994.
- J. PICKANDS : Statistical inference using extreme order statistics. *Annals of Statistics*, 3:119–131, 1975.
- M. H. QUENOUILLE : Problems in plane sampling. *The Annals of Mathematical Statistics*, 20:355–375, 1949.
- B. RENARD : *Détection et prise en compte d'éventuels impacts du changement climatique sur les extrêmes hydrologiques en France*. Thèse de doctorat, INPG, 2006.
- B. RENARD, D. KAVETSKI, E. LEBLOIS, M. THYER, G. KUCZERA et S. FRANKS : Toward a reliable decomposition of predictive uncertainty in hydrological modeling : Characterizing rainfall errors using conditional simulation. *Water Resources Research*, 47, 2011.
- B. RENARD, K. KOCHANEK, M. LANG, F. GARAVAGLIA, E. PAQUET, L. NEPPEL, K. NAJIB, J. CARREAU, P. ARNAUD, Y. AUBERT, F. BORCHI, J.-M. SOUBEYROUX, S. JOURDAIN, J.-M. VEYSSEIRE, E. SAUQUET, T. CIPRIANI et A. AUFRAY : Data-based comparison of frequency analysis methods : a general framework. *Water Resources Research*, 2013.
- M. RIBATET : *Consolidation de l'information hydrologique disponible localement et régionalement pour l'estimation probabiliste du régime des crues*. Thèse de doctorat, INPG / INRS (Québec, Canada), 2007.
- R. ROMERO, G. SUMNER, C. RAMIS et A. GENOVÉS : A classification of the atmospheric circulation patterns producing significant daily rainfall in the spanish mediterranean area. *International Journal of Climatology*, 19(7):765–785, 1999.
- E. SAUQUET : *Cartographie des écoulements annuels et mensuels d'un grand bassin versant structurée par la typologie du réseau hydrographique*. Thèse de doctorat, Ecole doctorale Terre, Univers et Environnement (Grenoble), 2000.
- E. SAUQUET, L. GOTTSCHALK et I. KRASOVSKAIA : Estimating mean monthly runoff at ungauged locations : an application to france. *Hydrology Research*, 39(5 - 6):403 – 423, 2008.
- G. SELLE et M. HANNAH : A bootstrap approach to assess parameter uncertainty in simple catchment models. *Environmental Modelling & Software*, 25 (8):919–926, 2010.
- E. SEVERINO et T. ALPUIM : Spatiotemporal models in the estimation of area precipitation. *Environmetrics*, 16:773–802, 2005.
- N. A. SHEFFER, Y. ENZEL, G. BENITO, T. GRODEK, N. POART, M. LANG, R. NAULET et D. COEUR : Paleofloods and historical floods of the Ardèche river, France. *Water Resour. Res.*, 39(12):1376–, décembre 2003.

- R. SNEYERS : On the statistical analysis of series of observations. *Word Meteorological Organisation*, page page Technical note 143, 1990.
- SOIL-CONSERVATION-SERVICE : *Curve Number (SCS-CN) methodology*. Kluwer academic publishers, 2003.
- T. S. STEENHUIS, M. WINCHELL, J. ROSSING, J. A. ZOLLWEG et M. F. WALTER : SCS runoff equation revisited for variable source runoff area. *Journal of Irrigation and Drainage Engineering*, 121:234–238, 1995.
- W.J. STIFEL et M.B. MC PHERSON : *Handbook of applied hydrology*. McGraw-Hill book Company, 1964.
- J. SZOLGAY, J. PARAJKA, S. KOHNOVÁ et K. HLAVCOVÁ : Comparison of mapping approaches of design annual maximum daily precipitation. *Atmospheric Research*, 92(3):289 – 307, 2009. 7th International Workshop on Precipitation in Urban Areas, 7th International Workshop on Precipitation in Urban Areas.
- J. TEWELES et H. WOBUS : Verifications of prognostics charts. *Bulletin of the American Meteorological Society*, 35 (10):455–463, 1954.
- A.H. THIESSEN : Precipitation averages for large areas. *Monthly Weather Review*, 39(7):1082–1084, 1911.
- J.W. TUKEY : Bias and confidence in not quite large samples. *The Annals of Mathematical Statistics*, 1958.
- USDA : *Engineering and Design – Flood-runoff analysis*. 1994.
- T. VISCHÉL, T. LEBEL, S. MASSUEL et B. CAPPELAERE : Conditionnal simulation schemes of rain fields and their application to the rainfall-runoff modeling studies in the Sahel. *Journal of Hydrology*, 375 (1-2):273–286, 2009.
- WMO : Distribution of precipitation in mountainous areas. Rapport technique, World Meteorological Organization, 1972.
- B. YARNAL, A. C. COMRIE, B. FRAKES et D. P. BROWN : Developments and prospects in synoptic climatology. *International Journal of Climatology*, 21(15):1923–1950, 2001.

