

Faculté des sciences

Interpolation spatiale des niveaux de retour de précipitations extrêmes

Comparaison de différentes approches
statistiques

Auteur: **Martin BEN HADDOU**
Promoteur: **Anna KIRILIOUK**
Lecteur: **Christian HAFNER**
Année académique 2025–2026
Master [120] en statistique, orientation générale

Remerciements

Je tiens à remercier toutes les personnes qui, de près ou de loin, m'ont aidé à réaliser ce mémoire dans les meilleures conditions.

Je remercie particulièrement la Professeure Anna Kiriliouk pour sa confiance et son suivi tout au long de l'année. Je remercie également le Professeur Christian Hafner pour le temps accordé à la lecture de ce mémoire.

Abstract

Les précipitations extrêmes constituent un enjeu important pour la gestion des risques hydrologiques, mais leur rareté rend difficile l'estimation de niveaux de retour en des sites non observés. Ce mémoire compare trois approches permettant d'exploiter l'information spatiale afin de prédire des niveaux de retour de précipitations extrêmes. La première consiste à ajuster indépendamment une distribution GEV à chaque station observée, puis à interpoler les niveaux de retour obtenus par pondération inverse de la distance. La deuxième repose sur les mêmes ajustements marginaux, mais remplace la proximité géographique par une pondération fondée sur la dépendance spatiale des extrêmes, estimée à l'aide du F-madogramme. La troisième modélise directement les paramètres de la GEV comme des fonctions de covariables spatiales.

Les trois méthodes sont d'abord évaluées dans deux cadres de simulation, reposant respectivement sur un modèle GEV spatial et sur un modèle max-stable de Brown–Resnick, afin d'étudier leur comportement sous différentes structures marginales et de dépendance. Elles sont ensuite appliquées aux maxima hivernaux de précipitations observés entre 1984 et 2024 dans 13 stations météorologiques situées aux Pays-Bas. Les performances prédictives sont évaluées à l'aide d'une procédure leave-one-station-out pour des périodes de retour de 10, 50 et 100 ans.

Les résultats montrent que les performances dépendent à la fois de la période de retour et de la structure spatiale sous-jacente. Sur les données réelles, l'interpolation par distance et le modèle GEV spatial présentent des performances globalement proches, avec un léger avantage pour le modèle GEV spatial aux périodes de retour les plus élevées. À l'inverse, l'utilisation directe de la dépendance extrême pour construire les poids d'interpolation n'améliore pas la prédiction, malgré l'identification d'une structure de dépendance stable principalement liée à la distance géographique et à l'orientation. Dans l'ensemble, ces résultats montrent qu'un modèle plus complexe ou intégrant une information spatiale plus détaillée n'offre pas nécessairement de meilleures performances prédictives, en particulier lorsque le nombre de stations et d'observations extrêmes reste limité.

Table des matières

Remerciements	I
Abstract	III
1 Introduction et objectifs	1
2 Fondements théoriques	2
2.1 Théorie univariée des valeurs extrêmes	2
2.1.1 Maxima par blocs et résultat asymptotique	2
2.1.2 Distribution généralisée des valeurs extrêmes	3
2.1.3 Inférence statistique, niveau et période de retour	5
2.1.4 Validation du modèle	6
2.1.5 Modèle non stationnaire	7
2.2 Interpolation spatiale	9
2.2.1 Interpolation par pondération inverse de la distance	9
2.2.2 Distance de Haversine	10
2.2.3 Limites de l'interpolation IDW	11
2.3 Dépendance spatiale des extrêmes	12
2.3.1 F-madogramme et dépendance extrême	12
2.3.2 Covariables physiques de dépendance spatiale	13
3 Revue de littérature	14
3.1 Précipitations extrêmes et niveaux de retour	14
3.2 Estimation spatiale des extrêmes pluviométriques	16
3.3 Interpolation spatiale des précipitations extrêmes et des niveaux de retour	17
3.3.1 Méthodes d'interpolation fondées sur la proximité géographique	18
3.3.2 Méthodes géostatistiques et prise en compte de la structure spatiale	19
3.3.3 Dépendance spatiale des extrêmes et interpolation par similarité extrême	20
3.4 Synthèse et positionnement du mémoire	21
4 Données	23
4.1 Origine des données	23
4.2 Disponibilité, sélection et prétraitement des données	23
4.3 Analyse descriptive	24
4.3.1 Répartition des stations météorologiques	24
4.3.2 Distribution des précipitations journalières hivernales	25
4.3.3 Distribution des maxima de précipitations hivernaux	27
5 Méthodologie	29
5.1 Objectif général de l'étude	29
5.2 Notation et validation par retrait de stations	30
5.3 Cadres de simulation	30
5.3.1 Simulation à partir d'un modèle GEV spatial	30
5.3.2 Simulation à partir d'un modèle max-stable	31
5.4 Méthode 1a : ajustements marginaux et interpolation IDW	32
5.5 Méthode 1b : ajustements marginaux et interpolation par dépendance	33
5.6 Méthode 2 : modèle GEV spatial	35
5.7 Critères d'évaluation	35
5.7.1 Évaluation dans les cadres de simulation	36
5.7.2 Évaluation sur les données réelles	36

6	Résultats et discussion des simulations	37
6.1	Résultats sous le cadre GEV spatial	37
6.1.1	Modèle générateur	37
6.1.2	Présentation des résultats	38
6.1.3	Discussion des résultats	39
6.2	Résultats sous le cadre max-stable	40
6.2.1	Modèle générateur	40
6.2.2	Présentation des résultats	40
6.2.3	Discussion des résultats	42
6.3	Comparaison des deux cadres de simulation	42
7	Application aux données réelles	44
7.1	Performances prédictives globales	44
7.2	Répartition spatiale des erreurs	45
7.3	Incertitude sur les niveaux de retour de référence	47
7.4	Structures spatiales sélectionnées	47
7.4.1	Modèle de dépendance de la méthode 1b	48
7.4.2	Structure du modèle de GEV spatiale de la méthode 2	49
7.5	Mise en perspective avec les résultats des simulations	50
8	Conclusion	52
A	Maxima saisonniers de précipitations pour les 13 stations	61
B	Incertitude sur les niveaux de retour de référence utilisés pour l'application aux données réelles	62
C	Modèles sélectionnées pour les méthodes 1b et 2 lors du procédé <i>leave-one-station-out</i> appliqué aux données réelles	63

1 Introduction et objectifs

Les précipitations extrêmes constituent un enjeu important pour la gestion des risques hydrologiques et l'aménagement du territoire. Des épisodes de pluie particulièrement intenses peuvent notamment contribuer à la formation de crues, provoquer des inondations ou mettre sous pression les systèmes de drainage et les infrastructures hydrauliques. L'évaluation de ces risques nécessite donc de pouvoir quantifier l'intensité d'événements rares, mais aussi d'estimer cette intensité dans des zones où les observations météorologiques sont limitées ou inexistantes. Cette dernière difficulté est particulièrement importante dans le cas des précipitations, puisque les relevés disponibles proviennent généralement d'un réseau ponctuel de stations météorologiques alors que le phénomène étudié varie de manière continue dans l'espace.

La modélisation statistique classique s'intéresse le plus souvent au comportement moyen ou central d'un phénomène aléatoire. Or, dans de nombreux domaines appliqués, les conséquences les plus importantes ne sont pas associées aux valeurs ordinaires, mais aux réalisations rares situées dans les queues de distribution. C'est le cas, par exemple, des crues exceptionnelles, des tempêtes, des vagues de chaleur, des pertes financières extrêmes ou encore des fortes précipitations. Dans ces situations, l'objectif n'est pas seulement de décrire les valeurs déjà observées, mais également d'estimer la probabilité d'événements plus rares que ceux contenus dans l'échantillon disponible [1, 2, 3]. C'est dans l'étude et la modélisation de ce type de phénomènes que la théorie des valeurs extrêmes s'inscrit comme un cadre statistique spécifiquement adapté. Elle permet de décrire le comportement des observations situées dans les queues de distribution et fournit des outils permettant d'estimer la probabilité d'événements d'une intensité supérieure – ou inférieure dans le cas de minima – à celles déjà observées. Dans le cadre de ce mémoire, la nécessité de représenter des données extrêmes apparaît dans l'étude des fortes pluies aux Pays-Bas. En effet, la modélisation de celles-ci ne peut pas se limiter à une analyse de la précipitation moyenne : elle doit porter spécifiquement sur la queue supérieure de la distribution des précipitations.

La théorie des valeurs extrêmes est appliquée aux maxima saisonniers de précipitations hivernales. Ceux-ci sont modélisés à l'aide de la distribution généralisée des valeurs extrêmes (GEV), afin d'estimer des niveaux de retour associés à différentes périodes. Le niveau de retour, dans la théorie des valeurs extrêmes, permet de quantifier l'intensité d'événements rares, et constituera la grandeur principale étudiée dans la suite du travail. Cependant, l'estimation directe d'un niveau de retour n'est possible qu'aux stations pour lesquelles des observations sont disponibles. L'objectif est donc également spatial : il s'agit d'estimer ces niveaux de retour en des sites non observés à partir de l'information fournie par les stations voisines.

Trois approches sont comparées dans ce mémoire. La première, appelée méthode 1a, repose sur des ajustements GEV marginaux réalisés séparément aux différentes stations, suivis d'une interpolation des niveaux de retour par pondération inverse de la distance (*Inverse Distance Weighting*, IDW). C'est une méthode de référence simple, fondée uniquement sur l'hypothèse que les stations géographiquement proches sont les plus informatives pour prédire un site cible.

La deuxième approche, appelée méthode 1b, conserve les mêmes ajustements marginaux mais cherche à intégrer une information plus directement liée au comportement des extrêmes. Les poids d'interpolation sont construits à partir d'une mesure de dépendance extrême, le F-madogramme, dont la variation est modélisée en fonction de caractéristiques des paires de stations, notamment leur distance, leur différence d'altitude et leur orientation. Cette méthode permet ainsi d'évaluer si la prise en compte explicite de la dépendance entre les extrêmes améliore l'interpolation fondée uniquement sur la proximité géographique.

Enfin, la méthode 2 repose sur un modèle GEV spatial. Contrairement aux deux méthodes précédentes, elle n'interpole pas directement des niveaux de retour estimés station par station. Les paramètres de la distribution GEV sont modélisés comme des fonctions de covariables spatiales telles que la latitude, la longitude, l'altitude et la distance à la côte. Cette approche permet donc d'étudier si une modélisation spatiale intégrée des distributions d'extrêmes apporte un avantage par rapport aux méthodes reposant sur des ajustements marginaux suivis d'une interpolation.

L'objectif principal du mémoire est ainsi de comparer les performances prédictives de ces trois méthodes et d'évaluer le compromis entre simplicité et complexité de modélisation. La comparaison est réalisée pour des périodes de retour de 10, 50 et 100 ans, afin d'étudier également l'évolution des performances lorsque l'extrapolation dans la queue de distribution devient plus importante. Au-delà des simples erreurs de prédiction, l'étude a aussi pour objectif de déterminer quelles informations supplémentaires peuvent être apportées par les méthodes plus complexes, notamment concernant la structure de dépendance entre les extrêmes et l'influence des covariables géographiques sur les paramètres de la GEV.

Afin d'évaluer les méthodes dans différents contextes, la comparaison est d'abord réalisée sur des données simulées. Deux cadres sont utilisés : un modèle GEV spatial, qui introduit une structure spatiale

dans les distributions marginales, et un modèle max-stable de Brown–Resnick, qui ajoute une dépendance spatiale explicite entre les extrêmes. Les simulations permettent de comparer les prédictions à des niveaux de retour théoriques connus et d’examiner dans quelle mesure les conclusions dépendent du mécanisme générateur.

Les trois méthodes sont ensuite appliquées à des données réelles de précipitations. L’étude porte sur treize stations disposant de séries continues de maxima hivernaux entre 1984 et 2024. Les performances sont évaluées au moyen d’une procédure *leave-one-station-out*, dans laquelle chaque station est successivement retirée avant d’y prédire les niveaux de retour à partir des autres stations. La confrontation des résultats issus des simulations et des données réelles permettra d’identifier les avantages, les limites et les domaines d’intérêt respectifs des trois méthodes d’estimation spatiale étudiées.

2 Fondements théoriques

2.1 Théorie univariée des valeurs extrêmes

Dans ce chapitre, la théorie des valeurs extrêmes sera parcourue et s’intéressera au comportement statistique de maxima. Soit $X_1, \dots, X_n$ une suite d’observations d’une variable aléatoire X , par exemple des cumuls journaliers de pluie enregistrés à une station météorologique. Le maximum de ces n observations est défini par

$$M_n = \max(X_1, \dots, X_n),$$

qui représente l’observation la plus élevée sur une période donnée. Si les données sont regroupées par années, le maximum annuel correspond ainsi à la plus forte valeur observée au cours d’une année. Plus généralement, dans l’approche dite des *maxima par blocs*, la série de données temporelle est divisée en blocs de même longueur, puis seul le maximum de chaque bloc est conservé. Lorsque les blocs sont annuels, on obtient une série de maxima annuels, qui peut ensuite être modélisée par une loi des valeurs extrêmes.

Une autre approche classique consiste à étudier les *dépassements de seuil*. Dans cette approche, au lieu de ne retenir qu’un maximum par bloc, on fixe un seuil élevé u et l’on conserve uniquement les observations qui dépassent ce seuil, c’est-à-dire

$$X_i > u, \quad i = 1, \dots, n.$$

Les deux approches répondent donc à une logique différente. L’approche dite *maxima par blocs* est simple à interpréter et bien adaptée lorsque les données extrêmes sont naturellement résumées par périodes, par exemple par années. Elle peut cependant écarter des événements importants qui ne sont pas les maxima de leur bloc. À l’inverse, l’approche par *dépassements de seuil* utilise davantage d’information dans la queue de distribution, mais elle dépend du choix du seuil et peut demander de traiter la dépendance entre dépassements proches dans le temps [4, 5].

Dans ce mémoire, c’est cette première approche de *maxima par blocs* qui sera privilégiée. En effet, cette méthode permet d’éviter une potentielle dépendance temporelle entre des précipitations extrêmes enregistrées plusieurs jours à la suite. On considérera donc l’hypothèse supposant les maxima des différents blocs indépendants [6, 7]. D’un autre côté, l’approche par *dépassements de seuil* soulève plusieurs difficultés. Dans un contexte spatial, le choix du seuil peut varier d’une station à l’autre, selon les localisations spatiales. De plus, les épisodes de fortes pluies peuvent générer plusieurs dépassements temporellement dépendants, ce qui impose des procédures supplémentaires pour traiter cette dépendance [8, 9].

Les sections suivantes introduisent les fondements univariés de la théorie des valeurs extrêmes, en se concentrant sur l’approche des maxima par blocs, avant d’aborder l’estimation, les niveaux de retour et la modélisation non stationnaire des paramètres.

2.1.1 Maxima par blocs et résultat asymptotique

L’approche *par maxima par blocs* repose sur l’étude du comportement limite du maximum d’un échantillon. Soit $(X_1, \dots, X_n)$ une suite de variables aléatoires indépendantes et identiquement distribuées, de fonction de répartition commune F . La fonction de répartition de M_n , le maximum de l’échantillon (défini dans la section 2.1) s’écrit alors

$$\mathbb{P}(M_n \leq z) = \mathbb{P}(X_1 \leq z, \dots, X_n \leq z) = \{F(z)\}^n.$$

Cette relation montre que le maximum dépend surtout du comportement de la distribution F dans ses valeurs les plus élevées. En effet, lorsque $n \rightarrow \infty$, cette probabilité converge vers 0 si $F(z) < 1$, et vers 1 si $F(z) = 1$. Autrement dit,

$$\mathbb{P}(M_n \leq z) \longrightarrow \begin{cases} 0, & \text{si } F(z) < 1, \\ 1, & \text{si } F(z) = 1. \end{cases}$$

Sans transformation, cette suite de maxima conduit donc généralement à une limite dégénérée. Si la distribution possède une borne supérieure, les maxima se rapprochent progressivement de cette borne. Si elle n'en possède pas, les maxima peuvent tendre vers l'infini.

Pour obtenir une limite non dégénérée, il est nécessaire de normaliser les maxima. On considère alors des suites constantes $a_n > 0$ et $b_n \in \mathbb{R}$, et on définit le maximum normalisé par

$$M_n^* = \frac{M_n - b_n}{a_n}.$$

Le rôle de b_n est de recentrer le maximum, tandis que a_n ajuste son échelle lorsque n augmente. Cette normalisation joue, dans l'étude des maxima, un rôle comparable à la normalisation utilisée dans le théorème central limite pour approximer la distribution des moyennes d'échantillons par la distribution normale. L'objectif est donc de rechercher les distributions limites de M_n^* en choisissant a_n et b_n appropriés, plutôt que celles de M_n .

Le théorème de Fisher–Tippett–Gnedenko (ou théorème des types extrêmes) établit les distributions limites existantes : s'il existe des suites de constantes $a_n > 0$ et $b_n \in \mathbb{R}$ telles que, si $n \rightarrow \infty$,

$$\mathbb{P}\left(\frac{M_n - b_n}{a_n} \leq z\right) \longrightarrow G(z)$$

où G est une fonction de répartition non dégénérée, alors G appartient nécessairement à l'une des trois familles suivantes [6, 10, 11] :

$$\text{I : } G(z) = \exp \left\{ -\exp \left[-\left(\frac{z-b}{a} \right) \right] \right\}, \quad -\infty < z < \infty;$$

$$\text{II : } G(z) = \begin{cases} 0, & z < b, \\ \exp \left\{ -\left(\frac{z-b}{a} \right)^{-\alpha} \right\}, & z > b; \end{cases}$$

$$\text{III : } G(z) = \begin{cases} \exp \left\{ -\left[-\left(\frac{z-b}{a} \right) \right]^\alpha \right\}, & z < b, \\ 1, & z \geq b. \end{cases}$$

Ces trois familles I, II et III sont également plus connues sous les noms respectifs de Gumbel, Fréchet et Weibull. Elles ont toutes un paramètre de localisation b et d'échelle $a > 0$. Les distributions de Fréchet et Weibull possèdent également un paramètre de forme $\alpha > 0$. Ces trois types de distributions des valeurs extrêmes constituent les seules limites possibles pour les distributions de M_n^* [3, 6, 10, 11, 12].

2.1.2 Distribution généralisée des valeurs extrêmes

Les trois familles limites introduites précédemment peuvent être réunies dans une seule famille paramétrique, appelée distribution généralisée des valeurs extrêmes, ou distribution GEV. Cette formulation unifiée permet de modéliser les trois comportements possibles des maxima normalisés à l'aide d'un unique paramètre de forme. Une variable aléatoire Z suit une distribution GEV de paramètres de localisation $\mu \in \mathbb{R}$, d'échelle $\sigma > 0$ et de forme $\xi \in \mathbb{R}$, notée

$$Z \sim \text{GEV}(\mu, \sigma, \xi),$$

si sa fonction de répartition est donnée par

$$G(z) = \exp \left\{ -\left[1 + \xi \left(\frac{z-\mu}{\sigma} \right) \right]^{-1/\xi} \right\},$$

définie pour les valeurs de z vérifiant

$$1 + \xi \left(\frac{z - \mu}{\sigma} \right) > 0.$$

Le paramètre μ contrôle la position de la distribution : une augmentation de μ déplace la distribution vers des valeurs plus élevées. Le paramètre σ contrôle sa dispersion : plus σ est grand, plus les maxima présentent une variabilité importante. Enfin, le paramètre ξ détermine le comportement de la queue de distribution. Il est particulièrement important, car il contrôle la probabilité d'observer des valeurs élevées. Il influence donc directement les estimations associées aux événements rares [3, 4, 6, 10, 11, 12]. Selon le signe de ξ , la distribution GEV correspond à l'un des trois types de lois limites :

$$\begin{cases} \xi = 0, & \text{type Gumbel,} \\ \xi > 0, & \text{type Fréchet,} \\ \xi < 0, & \text{type Weibull.} \end{cases}$$

Le cas $\xi = 0$ est obtenu par passage à la limite lorsque $\xi \rightarrow 0$, ce qui donne

$$G(z) = \exp \left\{ -\exp \left[-\left(\frac{z - \mu}{\sigma} \right) \right] \right\}, \quad z \in \mathbb{R}.$$

La distribution appartient alors au domaine de Gumbel. Son support est non borné et sa queue décroît de manière exponentielle.

Lorsque $\xi > 0$, la distribution appartient au domaine de Fréchet. La queue droite est alors lourde, ce qui signifie que des événements très extrêmes conservent une probabilité plus importante que dans le cas de Gumbel.

Lorsque $\xi < 0$, la distribution appartient au domaine de Weibull. Dans ce cas, les maxima ne peuvent pas dépasser une valeur limite finie.

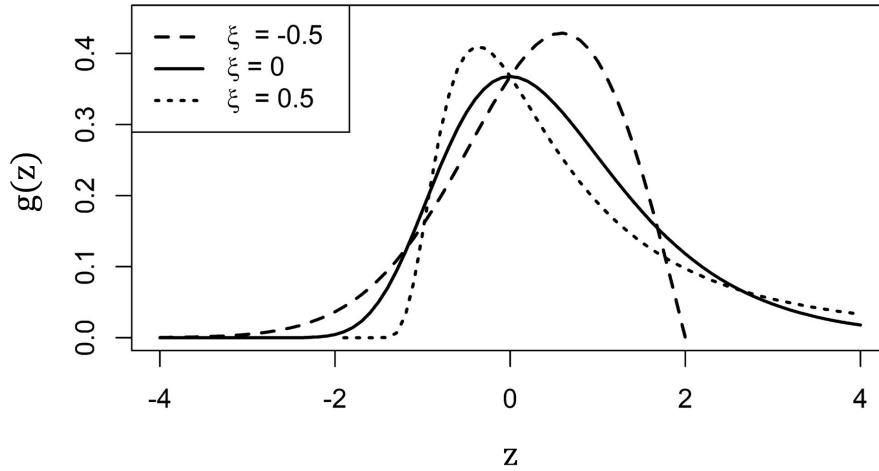


FIGURE 1 – Densités de probabilité de la distribution GEV standard pour différentes valeurs du paramètre de forme ξ [13].

La figure 1 [13] illustre la densité de probabilité $g(z)$ correspondant à la fonction de répartition $G(z)$ de la GEV. Cette densité montre l'effet du paramètre de forme ξ sur la densité d'une GEV standard, c'est-à-dire avec $\mu = 0$ et $\sigma = 1$. Pour $\xi = -0.5$, la densité présente une borne supérieure finie, ce qui correspond au cas Weibull. Pour $\xi = 0$, la distribution correspond au cas Gumbel, avec une queue droite non bornée qui décroît de manière exponentielle. Pour $\xi = 0.5$, la distribution correspond au cas Fréchet : la queue droite est plus lourde, il y a donc une probabilité plus élevée d'observer des valeurs très extrêmes. La distribution admet également une borne inférieure finie : sa fonction de répartition est nulle en dessous de cette borne, puis commence à croître à partir de celle-ci.

Dans le cadre de ce mémoire, la distribution GEV sera utilisée pour modéliser les maxima annuels de précipitations. Pour une station donnée, si $Z_1, \dots, Z_m$ désignent les maxima annuels observés sur m années, on suppose alors que

$$Z_i \sim \text{GEV}(\mu, \sigma, \xi), \quad i = 1, \dots, m.$$

Dans un premier cadre stationnaire, les paramètres μ , σ et ξ sont considérés constants pour une même station. Cette hypothèse sera ensuite généralisée dans le cadre non stationnaire, où certains paramètres pourront dépendre de covariables spatiales.

Le regroupement des trois familles limites au sein d'une même distribution présente un avantage important pour l'inférence statistique. Il n'est pas nécessaire de choisir a priori entre les familles de Gumbel, Fréchet et Weibull : l'estimation du paramètre de forme ξ permet aux données d'indiquer le type de comportement de queue le plus adapté. La GEV fournit ainsi un cadre général pour modéliser les maxima par blocs, et son utilisation pratique repose sur l'estimation de ses paramètres μ , σ et ξ , qui sera abordée dans la section suivante.

2.1.3 Inférence statistique, niveau et période de retour

Après avoir défini la distribution GEV, l'étape suivante consiste à estimer ses paramètres à partir des maxima observés. Dans le cadre stationnaire, pour une station donnée, on suppose que les maxima annuels $Z_1, \dots, Z_m$ suivent une même distribution GEV de paramètres μ , σ et ξ (respectivement paramètres de localisation, d'échelle et de forme). L'objectif de l'inférence statistique est alors d'obtenir des estimations de ces paramètres, notées $\hat{\mu}, \hat{\sigma}, \hat{\xi}$.

L'estimation des paramètres de la GEV est généralement réalisée par maximum de vraisemblance. Cette méthode consiste à rechercher les valeurs de μ , σ et ξ qui rendent les maxima observés les plus probables sous le modèle choisi [3, 6, 10, 11, 12]. En pratique, cette estimation repose sur la maximisation de l'équation de log-vraisemblance qui s'écrit, pour $\xi \neq 0$,

$$\ell(\mu, \sigma, \xi) = -m \log \sigma - \left(1 + \frac{1}{\xi}\right) \sum_{i=1}^m \log \left[1 + \xi \left(\frac{z_i - \mu}{\sigma}\right)\right] - \sum_{i=1}^m \left[1 + \xi \left(\frac{z_i - \mu}{\sigma}\right)\right]^{-1/\xi}.$$

La maximisation est effectuée avec des méthodes numériques car les équations associées ne possèdent généralement pas de solution explicite simple [6]. Lors de celle-ci, il faut respecter les contraintes de la GEV, notamment la positivité du paramètre d'échelle σ et la contrainte de support

$$1 + \xi \left(\frac{z_i - \mu}{\sigma}\right) > 0$$

pour chaque maximum observé z_i .

De manière générale, l'objectif principal de l'ajustement de modèles extrêmes est d'estimer les probabilités d'événements rares ou les risques qui en découlent. Une fois les paramètres estimés, la GEV permet de calculer les niveaux de retour et les périodes de retour, qui sont des quantités directement interprétables pour l'étude du risque. Pour une période de retour T , le niveau de retour z_T est défini comme le niveau dépassé avec une probabilité $1/T$ sur un bloc. Dans le cas de maxima annuels, cela signifie que

$$\mathbb{P}(Z > z_T) = \frac{1}{T},$$

ou

$$G(z_T) = 1 - \frac{1}{T}.$$

Par exemple, un niveau de retour associé à $T = 100$ ans correspond à une valeur dont la probabilité de dépassement par le maximum annuel est égale à $1/100$ sous les hypothèses du modèle. En d'autres termes, en calculant le niveau de retour z_T associé à une période de retour T ans, on calcule l'intensité d'un événement qui n'est dépassée que tous les T ans en moyenne.

À partir de la fonction de répartition de la GEV, le niveau de retour z_T s'écrit, pour $\xi \neq 0$,

$$z_T = \mu - \frac{\sigma}{\xi} \left[1 - \left\{1 - \log \left(1 - \frac{1}{T}\right)\right\}^{-\xi}\right].$$

Dans le cas limite $\xi = 0$, correspondant au type Gumbel, on obtient

$$z_T = \mu - \sigma \log \left[-\log \left(1 - \frac{1}{T}\right)\right].$$

En pratique, pour obtenir $\hat{z}_T$, les paramètres inconnus μ , σ et ξ sont remplacés par leurs estimations $\hat{\mu}, \hat{\sigma}, \hat{\xi}$.

L'estimation des niveaux de retour peut également être accompagnée d'une quantification de l'incertitude. En effet, le niveau de retour estimé $\hat{z}_T$ est obtenu en remplaçant les paramètres inconnus de la GEV par leurs estimations $\hat{\mu}$, $\hat{\sigma}$ et $\hat{\xi}$. Il dépend donc directement de l'incertitude associée à ces trois estimateurs. Cette incertitude est particulièrement importante pour les grandes périodes de retour, car celles-ci correspondent à une extrapolation dans la queue supérieure de la distribution.

Une manière classique de quantifier cette incertitude consiste à utiliser la matrice d'information de Fisher observée [6, 10]. Si l'on note $\theta = (\mu, \sigma, \xi)$ le vecteur des paramètres de la GEV et $\hat{\theta} = (\hat{\mu}, \hat{\sigma}, \hat{\xi})$ son estimation par maximum de vraisemblance, l'inverse de la matrice d'information de Fisher observée fournit une approximation de la matrice de variance-covariance de $\hat{\theta}$. Cette matrice, notée V , résume l'incertitude associée aux estimations de μ , σ et ξ , ainsi que les dépendances entre ces estimations [6].

La méthode delta permet ensuite de propager cette incertitude vers le niveau de retour. L'idée est de considérer z_T comme une fonction des paramètres de la GEV. En posant

$$q_T = -\log\left(1 - \frac{1}{T}\right),$$

le niveau de retour s'écrit, pour $\xi \neq 0$,

$$z_T = \mu - \frac{\sigma}{\xi} \left(1 - q_T^{-\xi}\right).$$

La variance de $\hat{z}_T$ peut alors être approximée par

$$\text{Var}(\hat{z}_T) \simeq \nabla z_T^\top V \nabla z_T,$$

où ∇z_T désigne le gradient du niveau de retour par rapport aux paramètres μ , σ et ξ , évalué en $\hat{\theta}$. Pour $\xi \neq 0$, ce gradient est donné par

$$\nabla z_T^\top = \left[\frac{\partial z_T}{\partial \mu} \quad \frac{\partial z_T}{\partial \sigma} \quad \frac{\partial z_T}{\partial \xi} \right],$$

également évalué en $(\hat{\mu}, \hat{\sigma}, \hat{\xi})$.

Cette approximation permet d'obtenir une erreur standard pour $\hat{z}_T$, notée $se(\hat{z}_T)$, puis de construire un intervalle de confiance approximatif. Par exemple, un intervalle de confiance à 95 % peut être écrit sous la forme

$$\hat{z}_T \pm 1.96 se(\hat{z}_T).$$

Ces intervalles doivent cependant être interprétés avec prudence. Ils reposent sur une approximation asymptotique et peuvent être peu fiables lorsque le nombre de maxima disponibles est limité. De plus, pour les grandes périodes de retour, l'incertitude augmente fortement, car l'estimation dépend davantage du comportement de la queue de distribution, et donc du paramètre de forme ξ .

Dans le cadre de ce mémoire, les niveaux de retour permettent de transformer les paramètres estimés de la GEV en quantités directement exploitables pour comparer le risque de fortes précipitations entre stations. Ce sont les valeurs de cet indicateur qui seront interpolées dans le but de modéliser l'intensité des fortes pluies sur l'ensemble du territoire étudié, et non uniquement aux emplacements des stations météorologiques.

En plus de tout cela, il faut tout de même préciser que la fiabilité des estimations de niveaux et périodes de retour dépend de la qualité de l'ajustement de la GEV aux maxima observés. Il est donc nécessaire d'évaluer l'adéquation du modèle avant d'interpréter les niveaux de retour estimés, ce qui est discuté dans la section suivante.

2.1.4 Validation du modèle

Une fois les paramètres de la GEV estimés, il est nécessaire d'évaluer si le modèle fournit une représentation cohérente des maxima observés. Cette étape permet d'identifier d'éventuelles déviations importantes entre les données et la distribution ajustée. Plusieurs graphiques diagnostiques sont couramment utilisés dans ce cadre.

Un premier outil est le diagramme de probabilité (PP-plot) [3, 6, 10, 11], qui compare les probabilités empiriques associées aux maxima observés avec les probabilités obtenues à partir du modèle ajusté. Si $z_{(1)} \leq \dots \leq z_{(m)}$ désignent les maxima ordonnés et $\hat{G}$ la fonction de répartition GEV estimée, on compare généralement les points

$$\left(\hat{G}(z_{(i)}), \frac{i}{m+1} \right), \quad i = 1, \dots, m \quad ,$$

où $\hat{G}$ désigne la fonction de répartition de la GEV ajustée, obtenue en remplaçant les paramètres inconnus μ , σ et ξ par leurs estimations $\hat{\mu}$, $\hat{\sigma}$ et $\hat{\xi}$. Si le modèle est adapté, ces points doivent être proches de la droite de référence $y = x$. Ce graphique permet donc de vérifier l'accord global entre la distribution empirique et la distribution ajustée. Cependant, le PP-plot présente une limite importante dans le contexte des valeurs extrêmes. Il met surtout en évidence les écarts dans la partie centrale de la distribution, car les probabilités sont représentées sur une échelle comprise entre 0 et 1. Or, dans l'analyse des extrêmes, l'intérêt principal porte sur les valeurs les plus grandes, c'est-à-dire sur la queue supérieure de la distribution. Ces valeurs sont peu nombreuses et correspondent à des probabilités très proches de 1, ce qui rend les écarts visuellement difficiles à détecter sur ce type de graphique.

Pour cette raison, le diagramme de quantile (QQ-plot) [3, 6, 10, 11] constitue un diagnostic complémentaire plus adapté à l'étude des queues de distribution. Il compare directement les maxima observés ordonnés aux quantiles théoriques obtenus à partir de la GEV ajustée :

$$\left(\hat{G}^{-1} \left(\frac{i}{m+1} \right), z_{(i)} \right), \quad i = 1, \dots, m.$$

Contrairement au PP-plot, ce graphique est exprimé sur l'échelle des observations. Il permet donc de mieux visualiser les écarts pour les valeurs élevées, qui sont précisément les plus importantes, le but final de l'ajustement de modèles de valeurs extrême étant l'extrapolation à des valeurs plus grandes.

Pour finir, un autre diagnostic central est le diagramme de niveau de retour [3, 6]. Ce graphique représente les niveaux de retour estimés en fonction de la période de retour. Les maxima observés sont représentés en fonction de périodes de retour empiriques, calculées à partir de leur rang dans l'échantillon, puis comparés à la courbe des niveaux de retour déduite de la GEV ajustée. Lorsque le modèle est cohérent avec les données, les points empiriques doivent suivre approximativement la courbe estimée, en particulier pour les grandes périodes de retour. Ce graphique permet donc d'évaluer directement la capacité d'extrapolation du modèle vers des événements rares.

2.1.5 Modèle non stationnaire

Les modèles présentés jusqu'ici reposent sur une hypothèse de stationnarité : les paramètres de la GEV sont supposés constants pour l'ensemble des maxima étudiés. Il est possible de distinguer deux types de non-stationnarité : la non-stationnarité temporelle et la non-stationnarité spatiale. Dans le premier cas, pour une station donnée, les maxima $Z_1, \dots, Z_m$ ne sont plus identiquement distribués. Dans le second cas, la distribution des maxima dépend de la localisation spatiale s , où $s_1, \dots, s_d$ désignent les d sites ou stations d'observation, de sorte que les variables $Z_i(s_1), \dots, Z_i(s_d)$ n'ont pas nécessairement la même distribution. Dans le cadre étudié dans ce mémoire, il est nécessaire de considérer ce deuxième type de non-stationnarité. En effet, les maxima annuels de précipitations peuvent dépendre de covariables spatiales propres aux stations météorologiques. Une approche non stationnaire consiste alors à conserver une structure de la loi GEV pour toutes les stations, mais à autoriser un ou plusieurs de ses paramètres à varier en fonction de ces covariables [14, 15, 16, 17, 18].

Soit $Z(s)$ un maximum annuel observé à la station s , associée à un vecteur de covariables $\mathbf{x}(s) = (x_1(s), \dots, x_p(s))$. Dans un modèle GEV non stationnaire, on écrit

$$Z_i(s) \sim \text{GEV}(\mu(s), \sigma(s), \xi(s)), \quad i = 1, \dots, m.$$

où les paramètres peuvent dépendre de $\mathbf{x}(s)$. Une approche possible est alors de paramétriser les paramètres de la GEV via des relations linéaires, par exemple,

$$\mu(s) = \beta_0^{(\mu)} + \beta_1^{(\mu)} x_1(s) + \dots + \beta_p^{(\mu)} x_p(s),$$

et

$$\log(\sigma(s)) = \beta_0^{(\sigma)} + \beta_1^{(\sigma)} x_1(s) + \dots + \beta_p^{(\sigma)} x_p(s).$$

L'utilisation du logarithme pour le paramètre d'échelle permet de garantir que $\sigma_i > 0$, quelle que soit la valeur des covariables [14, 19]. Le paramètre de forme ξ peut également être modélisé en fonction de covariables, mais il est souvent gardé constant en pratique, car son estimation est généralement plus instable que celle des paramètres de localisation et d'échelle [14, 19, 20, 21] :

$$\xi(s) = \xi.$$

Dans les équations qui suivent, le paramètre de forme sera toutefois noté comme dépendant des covariables ($\xi(s)$), afin de conserver une formulation générale du modèle non stationnaire.

Ce type de paramétrisation permet d'intégrer différents types de covariables aux paramètres μ , σ et ξ . Dans certains contextes, il peut s'agir du temps, afin de représenter une tendance dans les extrêmes. Dans le cadre de ce mémoire, l'intérêt porte surtout sur des covariables spatiales, telles que la position géographique ou les caractéristiques physiques des stations. Le modèle ne décrit alors plus seulement les extrêmes observés station par station : il cherche à expliquer une partie de leur variabilité spatiale à partir des informations disponibles sur les stations dont proviennent les relevés pluviométriques.

L'inférence statistique suit les mêmes lignes directrices que dans le cas stationnaire, bien que les paramètres directement estimés ne soient plus seulement μ , σ et ξ , mais également les coefficients des relations reliant ces paramètres aux covariables (coefficients β_j). L'estimation est également réalisée de manière générale par maximum de vraisemblance et par optimisation numérique. Dans le cas où $\xi(s) \neq 0$, la log-vraisemblance peut s'écrire :

$$\ell(\beta) = - \sum_{s=1}^S \sum_{t=1}^m \left\{ \log \sigma(s) + \left(1 + \frac{1}{\xi(s)} \right) \log \left[1 + \xi(s) \left(\frac{z_{s,t} - \mu(s)}{\sigma(s)} \right) \right] + \left[1 + \xi(s) \left(\frac{z_{s,t} - \mu(s)}{\sigma(s)} \right) \right]^{-1/\xi(s)} \right\},$$

sous les contraintes

$$\sigma(s) > 0 \quad \text{et} \quad 1 + \xi(s) \left(\frac{z_{s,t} - \mu(s)}{\sigma(s)} \right) > 0, \quad s = 1, \dots, S \quad \text{et} \quad t = 1, \dots, m.$$

Avec ce genre de paramétrisation, le choix du modèle devient une étape importante. Ajouter des covariables améliore généralement l'ajustement aux données, mais augmente aussi la complexité du modèle et peut conduire à un surajustement. Il faut donc respecter le principe de parcimonie, et choisir le modèle le plus simple qui explique correctement la variabilité observée. La comparaison entre modèles s'appuie sur la vraisemblance, sur des critères d'information comme la déviance, l'AIC, le BIC ou encore le TIC.

Dans ce mémoire, le critère utilisé est le TIC (*Takeuchi Information Criterion*) [22, 23]. Ce critère permet de comparer plusieurs modèles estimés par maximum de vraisemblance en tenant compte à la fois de leur qualité d'ajustement et de leur complexité. Comme l'AIC, il repose sur une pénalisation de la log-vraisemblance maximisée : un modèle est favorisé s'il ajuste bien les données, mais il est pénalisé lorsqu'il devient trop complexe. L'AIC est défini par

$$\text{AIC} = -2\ell(\hat{\theta}) + 2p,$$

où $\ell(\hat{\theta})$ désigne la log-vraisemblance maximisée et p le nombre de paramètres estimés. La pénalité de l'AIC dépend donc uniquement du nombre de paramètres du modèle [24].

Le TIC généralise cette idée en remplaçant le nombre de paramètres p par une mesure de complexité effective du modèle [22, 23]. Il peut s'écrire sous la forme

$$\text{TIC} = -2\ell(\hat{\theta}) + 2\text{tr} \left(\hat{J}^{-1} \hat{K} \right),$$

où $\hat{J}$ représente l'information liée à la courbure de la log-vraisemblance et $\hat{K}$ représente la variabilité des contributions individuelles à la vraisemblance. Le terme $\text{tr} \left(\hat{J}^{-1} \hat{K} \right)$ remplace donc le nombre de paramètres p utilisé dans l'AIC. Lorsque les hypothèses du modèle sont bien respectées, ce terme est proche de p et le TIC se rapproche de l'AIC. En revanche, lorsque certaines hypothèses sont imparfaites, cette pénalité peut s'adapter à la complexité effective du modèle.

Cette différence est importante dans le cadre de ce mémoire. Les stations météorologiques ne peuvent pas être considérées comme totalement indépendantes : deux stations proches peuvent être influencées par les mêmes épisodes de pluie extrême. L'information apportée par ces deux stations peut donc être partiellement redondante. Dans ce contexte, l'AIC peut être trop simple, car il pénalise uniquement le nombre de paramètres sans tenir compte du fait que les observations ne fournissent pas nécessairement une information indépendante. Le TIC est donc plus adapté ici, car il fournit une pénalisation moins naïve de la complexité du modèle lorsque l'hypothèse d'indépendance complète entre stations est discutable.

Enfin, la validation d'un modèle non stationnaire repose sur les mêmes principes que pour le modèle stationnaire, mais elle ne peut pas être appliquée directement aux maxima observés de la même manière [14]. Dans le cas stationnaire, toutes les observations sont supposées suivre une même distribution GEV ajustée, ce qui permet de construire directement des PP-plots, QQ-plots ou diagrammes de niveau de retour à partir de $\hat{G}$. Dans le cas non stationnaire, cette homogénéité n'est plus vérifiée : chaque observation peut être associée à des paramètres différents, calculés à partir de ses covariables.

Il faut donc d'abord ramener les observations sur une échelle commune. Pour une observation z_s associée aux paramètres estimés $\hat{\mu}(s)$, $\hat{\sigma}(s)$ et $\hat{\xi}(s)$, on peut définir une variable standardisée par

$$\tilde{z}_s = \frac{1}{\hat{\xi}(s)} \log \left[1 + \hat{\xi}(s) \left(\frac{z_s - \hat{\mu}(s)}{\hat{\sigma}(s)} \right) \right],$$

dans le cas où $\hat{\xi}(s) \neq 0$. Si le modèle non stationnaire est correctement spécifié, les valeurs transformées $\tilde{z}_s$ doivent alors suivre approximativement une loi de Gumbel standard, de fonction de répartition

$$\mathbb{P}(\tilde{Z} \leq z) = \exp\{-\exp(-z)\}.$$

Les graphiques diagnostiques peuvent ensuite être construits à partir de ces valeurs standardisées. Les PP-plots et QQ-plots ne comparent donc plus directement les maxima bruts à une seule GEV commune, mais les valeurs transformées à la distribution de référence de Gumbel. L'objectif reste le même que dans le cas stationnaire : vérifier si les points suivent approximativement la droite de référence, ce qui indiquerait une bonne adéquation du modèle.

Pour l'une des méthodes utilisées dans ce mémoire, la modélisation spatiale des maxima annuels de précipitations repose sur une GEV dont les paramètres peuvent varier selon des covariables dépendant des stations : la latitude et la longitude des stations, leur altitude ainsi que leur distance à la côte. Cette approche est mise en œuvre à l'aide de la fonction `fitspatgev()` du package `SpatialExtremes`, qui ajuste une loi GEV non-stationnaire en maximisant la fonction de log-vraisemblance présentée ci-dessus [25].

2.2 Interpolation spatiale

L'interpolation spatiale consiste à estimer la valeur inconnue d'un point à partir des points mesurés dans la zone d'étude. Dans le cadre de ce travail, les observations disponibles proviennent d'un ensemble de stations météorologiques localisées aux Pays-Bas. L'objectif est d'estimer, en un site cible s_0 , une valeur $y(s_0)$ à partir des valeurs connues ou estimées aux stations voisines $s_1, \dots, s_n$.

De manière générale, une interpolation spatiale peut s'écrire sous la forme d'une moyenne pondérée [26, 27, 28] :

$$\hat{y}(s_0) = \sum_{i=1}^n w_i(s_0) y(s_i),$$

où $y(s_i)$ représente la valeur connue ou estimée à la station s_i , et où $w_i(s_0)$ désigne le poids attribué à cette station pour prédire la valeur au site cible s_0 . Les poids vérifient généralement les contraintes suivantes :

$$w_i(s_0) \geq 0, \quad \sum_{i=1}^n w_i(s_0) = 1.$$

Le principe général est donc que chaque station contribue à l'estimation en s_0 , mais avec une importance variable selon sa proximité ou selon d'autres critères de similarité. Le poids attribué à chaque point connu est donc fonction de ces critères de similarité. Dans les méthodes les plus simples, cette similarité est uniquement fondée sur la distance géographique entre les stations.

Dans la continuité de la section précédente 2.1, où les niveaux de retour z_T ont été définis à partir de la distribution GEV comme des indicateurs de l'intensité d'événements rares, l'interpolation spatiale est utilisée ici pour estimer ces niveaux de retour en des sites non observés ou temporairement retirés de l'échantillon. L'interpolation intervient donc après l'ajustement marginal des modèles GEV aux stations disponibles. La valeur interpolée peut par exemple être un niveau de retour estimé à partir de ces ajustements marginaux.

2.2.1 Interpolation par pondération inverse de la distance

L'interpolation par pondération inverse de la distance, ou *Inverse Distance Weighting* (IDW), est une méthode d'interpolation reposant sur l'hypothèse que les stations les plus proches du site cible doivent avoir une influence plus importante que les stations plus éloignées [26, 29, 30]. Elle constitue une approche

simple et fréquemment utilisée lorsque l'on souhaite obtenir une estimation spatiale sans spécifier un modèle complet de dépendance spatiale.

Dans le cas de l'IDW, les poids sont construits directement à partir des distances entre le site cible s_0 et les stations disponibles s_i . On définit d'abord un poids non normalisé $W_i(s_0)$ par :

$$W_i(s_0) = \frac{1}{d(s_0, s_i)^p},$$

où $d(s_0, s_i)$ désigne la distance entre le site cible s_0 et la station s_i , et où $p > 0$ est un paramètre de puissance contrôlant la vitesse de décroissance des poids avec la distance. Les poids utilisés dans l'interpolation sont ensuite obtenus par normalisation :

$$w_i(s_0) = \frac{W_i(s_0)}{\sum_{k=1}^n W_k(s_0)} = \frac{d(s_0, s_i)^{-p}}{\sum_{k=1}^n d(s_0, s_k)^{-p}}.$$

Ainsi, plus une station est proche du site cible, plus sa distance $d(s_0, s_i)$ est faible, et plus son poids $w_i(s_0)$ est élevé. À l'inverse, les stations éloignées reçoivent un poids plus faible. Le paramètre p contrôle l'intensité de cet effet : lorsque p est élevé, l'interpolation donne beaucoup plus de poids aux stations les plus proches, et inversement lorsque p est faible. Une valeur de puissance couramment utilisée dans l'interpolation IDW est $p = 2$, ce qui revient à pondérer les stations selon l'inverse du carré de leur distance au site cible. C'est cette puissance qui est retenue dans ce mémoire pour prédire les niveaux de retour z_T .

L'estimateur IDW de $z_T(s_0)$ s'écrit donc :

$$\hat{z}_{T, \text{IDW}}(s_0) = \sum_{i=1}^n \frac{d(s_0, s_i)^{-2}}{\sum_{k=1}^n d(s_0, s_k)^{-2}} z_T(s_i).$$

2.2.2 Distance de Haversine

Dans l'interpolation IDW, les poids attribués aux stations dépendent directement de la distance entre le site cible s_0 et les stations disponibles s_i . Le choix de la métrique utilisée pour calculer cette distance est donc important, car il influence directement la valeur des poids $w_i(s_0)$, et donc l'estimation finale du niveau de retour $\hat{z}_T(s_0)$.

Dans ce mémoire, les stations météorologiques sont localisées par leurs coordonnées géographiques, c'est-à-dire par leur latitude et leur longitude. Ces coordonnées ne sont pas des coordonnées cartésiennes, mais des coordonnées angulaires définies sur la surface de la Terre. Il est donc nécessaire de calculer la distance entre les stations autrement que via une simple distance euclidienne. Pour cette raison, la distance utilisée dans l'interpolation IDW est la distance de Haversine [31, 32], qui permet d'approximer la distance géodésique entre deux points situés à la surface de la Terre.

Soient deux stations s_i et s_j , de coordonnées géographiques respectives (α_i, λ_i) et (α_j, λ_j) , où α désigne la latitude et λ la longitude. Avant d'appliquer la formule, les latitudes et longitudes sont converties de degrés en radians. On calcule ensuite les différences angulaires

$$\Delta\alpha = \alpha_j - \alpha_i, \quad \Delta\lambda = \lambda_j - \lambda_i.$$

La formule de la distance de Haversine entre deux stations s'écrit alors :

$$a = \sin^2\left(\frac{\Delta\alpha}{2}\right) + \cos(\alpha_i) \cos(\alpha_j) \sin^2\left(\frac{\Delta\lambda}{2}\right),$$

puis

$$d_{\text{Hav}}(s_i, s_j) = 2R \arcsin(\sqrt{a}).$$

où R désigne le rayon moyen de la Terre, qui est égal à environ 6371 kilomètres. Dans ce mémoire, la distance $d_{\text{Hav}}(s_i, s_j)$ est donc exprimée en kilomètres.

Dans le cadre de l'interpolation IDW, cette distance de Haversine est utilisée pour calculer la distance entre le site cible s_0 et chaque station disponible s_i . Les poids IDW deviennent alors

$$w_i(s_0) = \frac{d_{\text{Hav}}(s_0, s_i)^{-2}}{\sum_{k=1}^n d_{\text{Hav}}(s_0, s_k)^{-2}}.$$

où $d_{\text{Hav}}(s_0, s_i)$ représente la distance de Haversine entre le site cible et la station s_i . Pour rappel, le facteur de puissance utilisé dans ce travail est $p = 2$.

L'estimation IDW du niveau de retour en s_0 peut donc s'écrire

$$\hat{z}_{T,\text{IDW}}(s_0) = \sum_{i=1}^n \frac{d_{\text{Hav}}(s_0, s_i)^{-2}}{\sum_{k=1}^n d_{\text{Hav}}(s_0, s_k)^{-2}} \hat{z}_T(s_i).$$

où $\hat{z}_T(s_i)$ désigne le niveau de retour estimé à la station s_i à partir de l'ajustement marginal de la GEV. La distance de Haversine intervient donc uniquement dans la construction des poids d'interpolation : elle ne modifie pas les niveaux de retour estimés aux stations, mais détermine l'importance relative accordée à chaque station dans la prédiction au site cible.

2.2.3 Limites de l'interpolation IDW

L'interpolation IDW présente l'avantage d'être simple à mettre en œuvre et facile à interpréter. Cependant, elle repose sur une hypothèse forte : la similarité entre deux stations est supposée dépendre uniquement de leur distance géographique. Autrement dit, deux stations proches sont considérées comme plus informatives l'une pour l'autre que deux stations éloignées. Cette hypothèse peut être raisonnable pour des phénomènes simples ou spatialement réguliers, mais elle devient plus discutable dans des cadres plus complexes, par exemple pour la modélisation de précipitations extrêmes.

Dans le cas étudié dans ce mémoire, l'objectif n'est pas d'interpoler une variable moyenne, mais des niveaux de retour $\hat{z}_T(s)$ obtenus après ajustement marginal d'une GEV aux maxima de précipitations mesurés à chaque stations météorologiques. Ces niveaux de retour résument donc le comportement de la queue supérieure de la distribution. Pour ce type de données, on peut supposer que la proximité géographique ne suffit pas nécessairement à décrire la similarité entre stations. En effet, les phénomènes de pluies dépendent de manière générale de plusieurs facteurs géographiques : l'altitude, la latitude et la longitude, l'occupation du sol, la distance à la côte ou encore d'autres paramètres topographiques [33, 34, 35, 36]. Deux stations proches peuvent donc potentiellement présenter des niveaux de retour très différents si elles sont influencées par des caractéristiques physiques distinctes. À l'inverse, deux stations plus éloignées peuvent avoir des comportements extrêmes plutôt similaires si elles partagent des conditions climatiques ou géographiques comparables.

Cette limite est applicable pour les données de précipitations aux Pays-Bas utilisées dans ce mémoire. Même si le domaine spatial étudié est relativement restreint, les stations ne sont pas nécessairement équivalentes du point de vue des mécanismes générant les fortes pluies [37, 38, 39]. Les stations proches de la côte peuvent par exemple présenter des pluies plus intenses, car elles sont davantage influencées par les dépressions venant de la mer du Nord. De manière similaire, même si les variations d'altitude restent limitées aux Pays-Bas par rapport à des régions montagneuses, elles peuvent tout de même participer à la variabilité spatiale des précipitations extrêmes. Une interpolation fondée uniquement sur la distance de Haversine ne tient pas compte de ces effets.

Une autre limite de l'IDW est qu'elle ne modélise pas directement la dépendance entre les extrêmes. Les poids $w_i(s_0)$ sont déterminés uniquement par la distance $d(s_0, s_i)$, indépendamment du degré de dépendance empirique entre les maxima observés aux stations. Pourtant, dans un contexte de précipitations extrêmes, deux stations peuvent être touchées par les mêmes épisodes pluvieux intenses sur une même période, ce qui crée une dépendance entre leurs maxima. Cette dépendance peut décroître avec la distance, mais elle peut également dépendre de la direction, de la distance à la côte, des différences d'altitude ou d'autres caractéristiques physiques.

La présence de ces limites motive l'utilisation de poids d'interpolation construits autrement que par la seule distance géographique. Une possibilité consiste à remplacer les poids purement géographiques de l'IDW par des poids fondés sur une mesure de dépendance entre les extrêmes. Dans cette approche, une station reçoit un poids élevé non pas seulement parce qu'elle est proche du site cible, mais parce qu'elle présente un comportement extrême similaire à celui-ci. L'idée générale est donc de construire des poids reflétant une proximité statistique plutôt qu'une simple proximité spatiale.

Dans ce mémoire, une méthode d'interpolation alternative sera construite sur base de la dépendance estimée des comportements extrêmes entre les différentes stations. Cette approche cherche à mieux représenter la structure spatiale des précipitations extrêmes aux Pays-Bas, en tenant compte du fait que la similarité entre deux stations ne se réduit pas à leur distance géographique. Elle constitue ainsi une proposition d'interpolation plus complète que l'IDW : la forme générale de l'interpolation reste une moyenne pondérée des niveaux de retour estimés aux stations voisines, mais la construction des poids repose sur une information plus directement liée au comportement conjoint des extrêmes.

2.3 Dépendance spatiale des extrêmes

La théorie univariée des valeurs extrêmes permet de modéliser le comportement des maxima observés en une station donnée. Dans le contexte de fortes précipitations, cette approche fournit par exemple un modèle pour estimer, station par station, des niveaux de retour associés à différentes périodes de retour. Cependant, une telle modélisation marginale ne décrit pas la manière dont les extrêmes se comportent ensemble dans l'espace. Par exemple, deux stations peuvent présenter des distributions marginales similaires tout en ayant des comportements très différents du point de vue de leur dépendance spatiale.

Cette distinction est importante lorsque l'on s'intéresse à des phénomènes structurés dans l'espace. Les épisodes de précipitations extrêmes ne sont généralement pas limités à une seule station : ils peuvent affecter simultanément plusieurs sites, avec des intensités variables. La proximité géographique entre deux stations peut favoriser une forte dépendance entre leurs maxima, mais elle ne suffit pas toujours à caractériser entièrement cette dépendance. Des facteurs physiques tels que la différence d'altitude ou l'orientation relative entre les stations peuvent également influencer la similarité de leurs comportements extrêmes.

L'objectif de cette sous-section est donc d'introduire les notions nécessaires pour quantifier et interpréter la dépendance spatiale des extrêmes entre deux sites. Le F-madogramme sera présenté comme une mesure empirique de dissimilarité extrême, fondée sur des observations transformées sur une échelle marginale commune. Ensuite, il sera démontré que cette dissimilarité peut être interprétée de manière inverse, afin d'exprimer la force de dépendance entre les extrêmes. Autrement dit, une faible dissimilarité correspond à une forte dépendance entre les extrêmes, tandis qu'une dissimilarité élevée traduit une dépendance plus faible.

Enfin, cette section introduit également plusieurs covariables physiques définies par paire de stations, notamment la distance géographique, la différence d'altitude et l'orientation spatiale, aussi appelée *bearing*. Ces covariables seront utilisées plus tard pour construire une méthode d'interpolation, mais sont ici dans un premier temps présentées comme des grandeurs susceptibles d'expliquer une partie de la structure spatiale de dépendance entre les extrêmes pluviométriques.

2.3.1 F-madogramme et dépendance extrême

Le F-madogramme est une mesure de dissimilarité entre les comportements extrêmes observés en deux sites [40, 41, 42]. Avant de calculer cette dissimilarité, il est nécessaire d'effectuer une transformation marginale des observations. Celle-ci est nécessaire car les maxima bruts de précipitations observés à deux stations différentes peuvent avoir des distributions marginales différentes. Comparer directement leurs valeurs reviendrait alors à mélanger les différences de niveau marginal et la dépendance spatiale réelle entre les stations.

Soient M_i et M_j les maxima observés aux sites s_i et s_j , et soient F_i et F_j les fonctions de répartition marginales associées à ces deux sites. Les variables transformées

$$U_i = F_i(M_i) \quad \text{et} \quad U_j = F_j(M_j)$$

sont alors définies sur une échelle commune. Si les fonctions de répartition marginales sont continues et correctement spécifiées, ces variables suivent une loi uniforme sur $[0, 1]$. Cette transformation permet donc de comparer les deux séries de maxima non plus à partir de leurs niveaux bruts, mais à partir des positions relatives de chaque maximum dans sa propre distribution marginale.

Le F-madogramme entre les sites s_i et s_j , noté $\nu_F(s_i, s_j)$, est défini par

$$\nu_F(s_i, s_j) = \frac{1}{2} \mathbb{E} [|U_i - U_j|] = \frac{1}{2} \mathbb{E} [|F_i(M_i) - F_j(M_j)|].$$

Il correspond donc à la moitié de l'écart absolu moyen entre les deux variables transformées sur l'échelle uniforme. Plus les maxima des deux sites ont tendance à occuper simultanément des positions similaires dans leurs distributions marginales respectives, plus cet écart est faible. À l'inverse, si les extrêmes observés aux deux sites se comportent de manière plus indépendante, les valeurs transformées sont plus éloignées et le F-madogramme augmente.

En pratique, lorsque les fonctions marginales F_i et F_j ne sont pas connues, elles peuvent être remplacées par des estimations empiriques ou par des fonctions de répartition ajustées. Un estimateur empirique du F-madogramme peut alors s'écrire, pour un échantillon de maxima annuels $(M_i^{(t)}, M_j^{(t)})$ relevés aux

stations i et j pendant n ans :

$$\hat{\nu}_F(s_i, s_j) = \frac{1}{2n} \sum_{t=1}^n \left| \hat{F}_i(M_i^{(t)}) - \hat{F}_j(M_j^{(t)}) \right|.$$

Le F-madogramme est donc interprété comme une mesure de dissimilarité extrême. Dans le cas d'une dépendance parfaite entre les deux sites, les variables transformées sont identiques et l'on obtient

$$\nu_F(s_i, s_j) = 0.$$

À l'inverse, dans le cas d'indépendance entre les extrêmes, le F-madogramme atteint sa valeur maximale [42, 43] par calcul direct

$$\nu_F(s_i, s_j) = \frac{1}{6}.$$

Ainsi, une valeur faible de $\nu_F(s_i, s_j)$ indique une forte similarité entre les comportements extrêmes des deux stations, tandis qu'une valeur élevée indique une dépendance plus faible.

Cette dissimilarité peut ensuite être interprétée de manière inverse, afin d'exprimer la force de dépendance entre les extrêmes. Par exemple, la quantité $1 - 6\nu_F(s_i, s_j)$ fournit une transformation facilement interprétable du F-madogramme : elle vaut 1 en cas de dépendance parfaite et 0 en cas d'indépendance. Cette transformation permet de passer d'une lecture en termes de dissimilarité à une interprétation intuitive en termes de dépendance extrême.

Le F-madogramme ne se limite pas à fournir une mesure de dissimilarité pouvant être interprétée, par inversion, comme une mesure de dépendance. Il est également directement relié au coefficient extrême, qui constitue une mesure classique de la dépendance en théorie des valeurs extrêmes [40, 41, 43]. Ce lien est donné par

$$\theta_{ij} = \frac{1 + 2\nu_F(s_i, s_j)}{1 - 2\nu_F(s_i, s_j)}.$$

Cette relation montre que le F-madogramme et le coefficient extrême contiennent la même information de dépendance, exprimée sur deux échelles différentes. Lorsque $\nu_F(s_i, s_j) = 0$, on obtient $\theta_{ij} = 1$, ce qui correspond à une dépendance parfaite. Lorsque $\nu_F(s_i, s_j) = \frac{1}{6}$, on obtient $\theta_{ij} = 2$, ce qui correspond à l'indépendance. Dans la suite, le F-madogramme sera donc retenu comme mesure empirique de dissimilarité extrême entre paires de stations.

Cette relation entre le F-madogramme et le coefficient s'inscrit dans un cadre théorique plus large : celui des modèles max-stables, qui constituent les modèles limites pour décrire la dépendance spatiale entre maxima [40]. Dans un champ max-stable, la structure de dépendance entre deux sites est entièrement caractérisée par le coefficient d'extrêmes θ_{ij} , lequel quantifie la probabilité que des événements extrêmes se produisent simultanément aux deux stations.

2.3.2 Covariables physiques de dépendance spatiale

Le F-madogramme permet de quantifier la dissimilarité extrême entre deux stations à partir des observations. Cependant, cette dissimilarité ne doit pas être considérée indépendamment des caractéristiques physiques des paires de stations lors de l'établissement d'un modèle d'interpolation par dépendance spatiale. Pour deux stations données, certaines caractéristiques géographiques peuvent aussi contribuer à expliquer pourquoi les comportements extrêmes sont plus ou moins proches. Dans ce mémoire, trois covariables physiques sont considérées pour décrire cette relation géographique entre paires de stations : la distance géographique, la différence d'altitude et l'orientation spatiale du segment reliant les deux sites.

Pour la distance géographique entre deux stations, on considère la distance de Haversine d_{Hav} (section 2.2.2). Celle-ci est calculée à partir des coordonnées géographiques des stations s_i et s_j , respectivement (α_i, λ_i) et (α_j, λ_j) . Dans un contexte de dépendance spatiale des extrêmes, elle est une covariable naturelle : deux sites proches ont plus de chances d'être affectés par un même événement pluvieux intense que deux sites éloignés. Cette relation n'est toutefois pas nécessairement suffisante (section 2.2.3), et il convient d'inclure d'autres covariables pour construire un modèle plus adapté à la réalité.

La deuxième covariable est la différence absolue d'altitude entre les deux stations. Si h_i et h_j désignent les altitudes respectives des stations s_i et s_j , on définit

$$\Delta_{\text{alt},ij} = |h_i - h_j|.$$

Deux stations situées à des altitudes différentes peuvent être soumises à des régimes pluviométriques d'intensités distinctes, même lorsqu'elles sont relativement proches. Dans le cas des Pays-Bas, les différences

d'altitude sont limitées par rapport à des régions montagneuses, mais cette covariable peut néanmoins jouer un rôle dans la dépendance géographique, notamment entre zones littorales, basses terres et zones légèrement plus élevées.

La troisième covariable est l'orientation spatiale du segment reliant deux stations, aussi appelée *bearing*. Elle permet de tenir compte d'une éventuelle anisotropie de la dépendance spatiale. En effet, deux stations séparées par une même distance peuvent présenter des niveaux de dépendance différents selon leur orientation relative. Dans le contexte de fortes pluies, cette covariable peut par exemple être liée aux mouvements des masses d'air ou à l'orientation du segment reliant les stations par rapport à la mer. Dans un de ses rapports, le KNMI souligne que la considération de l'anisotropie pourrait améliorer l'interpolation des données de pluies [44], et d'autres auteurs considèrent dans leur travail la potentielle anisotropie de la dépendance entre les stations de relevés de précipitation [45, 46].

Soit φ_{ij} l'angle décrivant l'orientation du segment reliant les stations s_i et s_j . Cet angle peut être calculé à partir des coordonnées géographiques des deux stations [47], et converti en radians. Dans ce mémoire, la direction du segment n'est pas considéré comme importante : la paire (s_i, s_j) représente la même relation spatiale que la paire (s_j, s_i) . Les angles φ_{ij} et $\varphi_{ij} + \pi$ doivent être considérés comme équivalents.

Pour représenter cette orientation sans discontinuité entre 0 et 2π , on utilise les composantes trigonométriques :

$$\cos(2\varphi_{ij}) \quad \text{et} \quad \sin(2\varphi_{ij}).$$

Le facteur 2 permet de rendre équivalentes deux directions opposées, puisque φ_{ij} et $\varphi_{ij} + \pi$ donnent les mêmes valeurs de $\cos(2\varphi_{ij})$ et $\sin(2\varphi_{ij})$. Les deux quantités $\cos(2\varphi_{ij})$ et $\sin(2\varphi_{ij})$ fournissent donc une représentation continue et non orientée du bearing entre deux stations, utilisable comme covariables pour décrire une possible variation de la dépendance spatiale en fonction de leur orientation relative. Dans ce mémoire, le bearing est calculé aux moyens de la fonction `bearing()` du package R `geosphere` [48].

Ces trois covariables sont donc complémentaires. La distance décrit l'éloignement horizontal entre les stations, la différence d'altitude représente leur contraste topographique, tandis que l'orientation permet de tenir compte d'une éventuelle variation directionnelle de la dépendance. Leur combinaison fournit une description parcimonieuse de trois dimensions distinctes de la configuration spatiale d'une paire de stations : la proximité, la similarité du relief et l'alignement spatial.

D'autres covariables auraient également pu être envisagées. La distance à la côte constitue notamment une variable pertinente dans le contexte étudié, puisque la proximité maritime peut influencer les régimes de précipitations et leur structure spatiale [49]. Son exclusion dans la caractérisation de la dépendance spatiale ne signifie donc pas qu'elle est considérée comme dépourvue d'effet. Il a toutefois été choisi de ne pas l'intégrer au modèle de dépendance principal. Contrairement à la distance, à la différence d'altitude et à l'orientation, sa traduction en une caractéristique propre à une paire de stations est moins univoque : elle pourrait être représentée par la différence de distance à la côte entre les deux sites, par leur distance moyenne à la côte, ou par ces deux informations simultanément. De plus, une partie de son influence sur le niveau des précipitations sera déjà prise en compte dans la modélisation des marges, avant leur transformation sur une échelle commune. Le modèle de dépendance est donc volontairement centré sur les trois caractéristiques décrivant directement la géométrie et le contraste topographique des paires de stations. La distance à la côte reste néanmoins une extension possible du modèle et son absence se traduit par un choix de spécification plutôt que comme la considération de son absence d'influence.

Le rôle méthodologique précis de ces covariables sera présenté plus tard dans ce travail, lorsque la méthode d'interpolation par dépendance sera décrite.

3 Revue de littérature

3.1 Précipitations extrêmes et niveaux de retour

Les précipitations extrêmes font l'objet d'étude central en hydrologie, car elles sont directement associées aux risques d'inondation et à d'autres catastrophes. Contrairement aux précipitations moyennes, l'importance des pluies extrêmes ne tient pas vraiment à leur fréquence, mais surtout à leur intensité et à leurs conséquences potentielles. Plusieurs auteurs [50, 51, 52, 53] rappellent ainsi que l'augmentation observée de la fréquence des événements pluviométriques intenses nécessite l'amélioration et le développement des techniques d'estimation des pluies rares, en particulier pour l'aménagement du territoire, le dimensionnement des ouvrages hydrauliques et la gestion des risques naturels. Ces études partagent

toutes une même logique : élaborer des modèles probabilistes à partir des observations météorologiques locales pour améliorer l'évaluation du risque [51, 52, 54].

Les travaux consacrés à la modélisation des précipitations extrêmes s'inscrivent largement dans le cadre de la théorie des valeurs extrêmes, qui permet de se concentrer spécifiquement sur le comportement de la queue supérieure de la distribution [50, 53]. Parmi les approches disponibles, la distribution GEV, adaptée à la modélisation des maxima annuels ou saisonniers, est particulièrement répandue dans la littérature et a été appliquée à des contextes climatiques variés [52, 53, 54, 55].

Dans les travaux cités ci-dessus, l'objectif principal n'est pas d'estimer les paramètres des distributions en tant que tels, mais d'estimer des niveaux de retour z_T associés à une période de retour T . Les périodes de retour étudiées varient selon les objectifs. Wu [52] estime des niveaux associés à des périodes de 2 à 100 ans afin d'évaluer les pluies rares à court et long termes dans une région japonaise récemment touchée par des événements intenses. Ahuja et al. [54] considèrent des périodes allant jusqu'à 500 ans dans un contexte himalayen, ce qui montre une volonté de quantifier des risques extrêmes à très long termes. Gentilucci et al. [51] se concentrent sur le niveau de retour à une période de 100 ans, car il s'agit d'une référence usuelle pour les infrastructures et l'aménagement du territoire. Lazoglou et al. [53], quant à eux, utilisent les niveaux de retour pour comparer les méthodes basées sur la théorie des valeurs extrêmes et classer les stations méditerranéennes selon leur niveau de risque. Ces différentes études montrent que le niveau de retour joue un rôle bien spécifique : c'est une grandeur de synthèse permettant de comparer des lieux, des méthodes ou des périodes d'observation.

Malgré son usage répandu, l'estimation marginale des niveaux de retour présente plusieurs limites. La première concerne la longueur de la série utilisée. Gentilucci et al. [51] montrent que les niveaux de retour à 100 ans estimés par GEV varient fortement selon que l'on utilise une série longue de 70 ans ou une série plus récente de 30 ans. Dans leur étude, la série courte 1991–2020 conduit à des niveaux de retour plus élevés que la série 1951–2020, et l'ajout d'un événement extrême récent augmente encore les estimations. Ces propos sont également appuyés par les travaux de Marra et al. [56] et Papalexiou & Koutsoyiannis [57], qui confirment que ce type d'approche génère de larges incertitudes lors de l'ajustement des paramètres sur des échantillons de faible taille. De plus, la variabilité interne du système climatique induit une énorme variabilité dans les événements de précipitations extrêmes. L'estimation du comportement de queue est en particulier caractérisée par une incertitude élevée [56, 57]. Pour des estimations de ce type, une longue série peut mélanger plusieurs régimes climatiques et donc lisser les extrêmes récents, tandis qu'une série plus courte peut être plus représentative du régime actuel, mais plus instable statistiquement [51].

La seconde limite concerne l'incertitude croissante lorsque la période de retour augmente. Wu [52] montre que la GEV peut donner un ajustement acceptable pour des périodes de retour T modérées, mais les intervalles de confiance associés aux niveaux de retour s'élargissent avec T . C'est un comportement logique : plus on extrapole loin au-delà des observations disponibles, plus l'estimation dépend fortement de la forme de la queue et donc du paramètre de forme ξ , souvent difficile à estimer avec précision. Toute la difficulté se retrouve alors dans l'estimation la plus précise possible de ce paramètre de forme, souvent peu fiable en raison de la faible longueur des séries pluviométriques. Une petite erreur dans l'estimation du paramètre de forme peut entraîner une erreur bien plus importante dans l'estimation des niveaux de retour [58, 59].

Enfin, le choix même de la méthode d'estimation des paramètres du modèle peut impacter fortement les estimations. Lazoglou et al. [53] montrent que les niveaux de retour peuvent changer selon que les paramètres sont estimés par maximum de vraisemblance, L-moments ou méthode bayésienne. Le problème est également un problème méthodologique : deux méthodes statistiquement défendables peuvent produire des niveaux de retour différents. Cette différence entre les méthodes d'estimation est surtout marquée pour l'estimation du paramètre de forme des distributions GEV. Dans leur comparaison entre l'estimation des paramètres par maximum de vraisemblance et L-moments, Wehner et al. [60] concluent que les L-moments donnent des estimations moins variables pour les petits échantillons, lorsque les MLE peuvent être moins biaisées, mais parfois instables ou non convergents. Dans tous les cas, le choix de la méthode « optimale » doit dépendre des données disponibles et du contexte. L'étude menée par Hossain et al. [61] nuance les deux études précédentes : si la méthode d'estimation influence effectivement l'estimation des paramètres, cet effet reste faible au regard de l'impact beaucoup plus marqué de la taille de l'échantillon.

De manière générale, les modèles marginaux simples supposent souvent que chaque station peut être traitée indépendamment et que la distribution des extrêmes reste stable dans le temps. Or, toutes les conclusions des travaux cités dans cette section se rejoignent sur un point : les méthodes non stationnaires sont nécessaires pour améliorer la stabilité des estimations. Que ce soit dans un contexte de climat changeant ou face à des séries d'observations trop courtes, le recours à des modèles tenant compte de

la non-stationnarité des données apparaît comme une piste essentielle pour améliorer l'estimation des niveaux de retour. Milly et al. [62] remettent directement en cause l'usage de la stationnarité comme hypothèse par défaut en gestion des ressources en eau, au motif que le changement climatique modifie les régimes hydroclimatiques et donc les probabilités associées aux extrêmes. Katz et al. [50] soulignent également que les modèles d'extrêmes gagnent à intégrer des covariables physiques, des tendances ou des effets régionaux, plutôt que de se limiter à des ajustements station par station. Cette critique prépare naturellement le terrain pour considérer des approches spatiales, capables d'emprunter de l'information entre stations et de tenir compte de facteurs géographiques ou climatiques. En effet, l'estimation marginale station par station est nécessaire, mais souvent insuffisante pour prédire en des sites non observés.

3.2 Estimation spatiale des extrêmes pluviométriques

Cette section a pour but d'explorer les différentes méthodes mises en place dans la littérature pour construire des modèles exploitant les données de plusieurs stations de relevé de précipitations simultanément.

Une méthode d'abord utilisée dans certains travaux est l'analyse fréquentielle régionale, couramment utilisée en hydrologie pour estimer les pluies ou débits rares. Dans cette approche, les stations sont regroupées en régions supposées homogènes, puis les données sont normalisées avant d'ajuster une distribution régionale commune. Cette logique permet d'augmenter la quantité d'information disponible pour l'estimation de quantiles extrêmes, comme dans l'étude de Fowler et Kilsby [63], qui applique une analyse fréquentielle régionale aux maxima de pluie au Royaume-Uni. Dans leur travail, Szolgay et al. [26] concluent que l'analyse fréquentielle régionale peut être considérée comme une méthode appropriée en vue d'établir une référence pour la cartographie des précipitations, et permet de pallier le problème du manque de données. Par contre, cette méthode repose sur une hypothèse forte : les régions considérées sont homogènes et clairement identifiables. Or, cette hypothèse peut être difficile à vérifier lorsque les stations considérées présentent des caractéristiques topographiques variables, qui peuvent créer des régimes pluviométriques extrêmes différents. Deidda et al. [64] soulignent dans leur travail que les approches par régions homogènes peuvent mal représenter les caractéristiques locales des précipitations extrêmes, en particulier pour des terrains complexes. Le Gall et al. [65] confirme ces propos en rappelant que, bien que l'identification de régions climatiquement cohérentes peut améliorer l'inférence des niveaux de retour, cette étape reste délicate et conditionne fortement la qualité des estimations obtenues. Ces limites motivent le développement de modèles spatiaux plus continus, dans lesquels les paramètres des distributions de valeurs extrêmes varient directement avec des covariables géographiques ou climatologiques.

Deux familles de modèles spatiaux ont été identifiées dans la littérature : la première, plus classique, consiste à conserver le cadre GEV tout en rendant ses paramètres dépendants de la localisation [19, 66, 67, 68]. La deuxième est reposée sur les modèles hiérarchiques bayésiens spatiaux [69, 70, 71]. Dans cette analyse, c'est la première famille de méthode qui sera présentée et discutée, la seconde étant plus complexe et éloignée du sujet de ce mémoire. Au lieu d'ajuster indépendamment une distribution à chaque station, on suppose que les paramètres $\mu(s)$, $\sigma(s)$ et éventuellement $\xi(s)$ varient selon des covariables mesurées en chaque site (cf. section 2.1.5). Cette stratégie permet de produire une distribution d'extrêmes continue en tout point du domaine, à condition d'avoir les covariables spatiales à disposition aux endroits où l'on veut estimer le modèle. Cette méthode se distingue donc d'une simple interpolation a posteriori des niveaux de retour ou des paramètres de distributions marginales : la structure spatiale est intégrée directement dans le modèle probabiliste des extrêmes.

Les travaux de Van de Vyver [66], Fischer et al. [67], Hammami et al. [68] et Boluwade [19] illustrent différentes manières de mettre en œuvre cette méthodologie. Leur point commun est de considérer que la variabilité spatiale des extrêmes pluviométriques peut être résumée, au moins partiellement, par des covariables physiques ou géographiques. Van de Vyver [66] relie les paramètres GEV à des covariables comme l'altitude, la distance à la mer ou la pluie annuelle moyenne, et montre que la pluie annuelle moyenne et l'altitude expliquent mieux la variation spatiale des extrêmes que les coordonnées géographiques seules. Fischer et al. [67] utilisent également la longitude, la latitude et l'altitude, mais les introduisent au moyen de polynômes de Legendre afin de décrire des variations spatiales lisses. Hammami et al. [68] élargissent cette logique en comparant des modèles GLM et des réseaux de neurones pour représenter la relation entre paramètres GEV et covariables spatiales, notamment les coordonnées, l'altitude et également une covariable représentant les totaux moyens de précipitations saisonnières. Boluwade [19] adopte une approche plus orientée vers la dépendance spatiale, mais conserve une logique comparable pour les distributions : les paramètres GEV sont exprimés à partir de coordonnées géographiques, de la topographie et d'une

covariable temporelle.

Ces articles montrent aussi que tous les paramètres GEV ne jouent pas le même rôle dans la modélisation spatiale. La localisation $\mu(s)$ est généralement le paramètre pour lequel le signal spatial est le plus net. Van de Vyver [66] met en évidence une forte variabilité régionale du paramètre de localisation, une variabilité plus faible du paramètre d'échelle, et ne retient pas de variation spatiale significative du paramètre de forme. Fischer et al. [67] obtiennent également une structure spatiale principalement portée par le paramètre de localisation, tout en montrant que permettre au paramètre de forme de varier selon la saison et l'espace peut améliorer la performance prédictive. Hammami et al. [68] modélisent les trois paramètres, mais leurs résultats rappellent que cette flexibilité doit être contrôlée, notamment dans les régions où le nombre de stations est faible. En effet, le fait de laisser varier davantage de paramètres augmente la souplesse du modèle, mais accroît aussi le risque d'instabilité, en particulier pour ξ , qui est souvent difficile à estimer précisément.

Il est aussi possible de comparer ces travaux à propos de la forme utilisée pour relier les covariables aux paramètres GEV. Les modèles de régression linéaire, comme ceux de Van de Vyver [66], ont l'avantage d'être interprétables : ils permettent d'identifier et de quantifier directement le rôle de covariables comme l'altitude ou la pluie moyenne annuelle. À défaut d'inclure une covariable de pluie moyenne annuelle, Boluwade [19] introduit quant à lui une covariable temporelle afin d'évaluer l'existence d'une tendance dans les extrêmes, et teste des termes d'interactions et de second ordre dans ses modèles de régression linéaire. Les modèles de Fischer et al. [67] restent paramétriques, mais utilisent des bases de fonctions pour obtenir une variation spatiale et saisonnière plus souple. À l'inverse, Hammami et al. [68] montrent que les réseaux de neurones peuvent mieux capturer des relations non linéaires entre covariables et paramètres GEV, mais au prix d'un réglage plus délicat de la complexité du modèle. Leur comparaison entre GLM et ANN indique que le choix du modèle dépend fortement de la densité du réseau de stations, de la taille de l'échantillon et des corrélations entre covariables. Cela met en évidence le fait qu'une méthode plus flexible n'est pas automatiquement préférable si les données ne suffisent pas à l'estimer correctement.

Ces modèles spatiaux GEV ont donc un intérêt direct pour l'estimation des niveaux de retour en sites peu observés ou non observés. Ils permettent de remplacer une estimation strictement locale par une estimation en partie mutualisée, où les observations de plusieurs stations contribuent à l'identification d'une surface spatiale des paramètres. Ils permettent d'introduire des facteurs physiques interprétables dans les modèles, comme l'altitude, la proximité au littoral, la pluie moyenne annuelle ou d'autres covariables spatiales ou climatiques. Leur limite principale est que leur performance dépend fortement du choix des covariables et de la forme imposée aux relations spatiales. Une spécification trop simple peut trop lisser certaines variations locales importantes, lorsqu'une spécification trop flexible peut au contraire devenir instable si le réseau de stations est trop peu dense.

À côté de ce type de modèle intégrant directement des covariables spatiales dans la spécification du modèle, une autre stratégie consiste à estimer les niveaux de retour localement puis à les interpoler vers des sites non observés. Cette approche, plus simple et plus opérationnelle, fait l'objet de la sous-section suivante.

3.3 Interpolation spatiale des précipitations extrêmes et des niveaux de retour

L'interpolation spatiale constitue une seconde manière d'exploiter l'information spatiale disponible entre les stations. Contrairement aux modèles spatiaux intégrés présentés précédemment, elle ne cherche pas à modéliser directement les paramètres de la loi GEV sur l'ensemble du domaine. Elle repose plutôt sur une logique en deux étapes : les distributions marginales sont d'abord ajustées séparément aux stations observées, puis les quantités estimées, comme les niveaux de retour z_T , sont transférées vers des sites non observés par pondération spatiale.

Cette approche présente un intérêt pratique important. Elle permet de conserver la simplicité et l'interprétabilité des ajustements marginaux, tout en utilisant l'information issue des stations voisines pour prédire les niveaux de retour en des points où aucune série temporelle n'est disponible. Toutefois, elle soulève une question centrale : comment définir les poids attribués aux stations observées ? La réponse la plus simple consiste à pondérer les stations selon leur proximité géographique, comme dans l>IDW. Des approches plus élaborées, comme le krigeage, cherchent plutôt à représenter une structure spatiale globale à partir d'un variogramme. Enfin, dans le cas spécifique des extrêmes, il est possible de dépasser la seule notion de proximité géographique en construisant des poids à partir d'une mesure de similarité extrême entre stations.

Les sous-sections suivantes présentent donc progressivement ces différentes stratégies. Les méthodes fondées sur la distance géographique sont d’abord introduites comme référence simple. Les méthodes géostatistiques sont ensuite discutées comme une généralisation tenant compte de la structure spatiale des observations. Enfin, les approches fondées sur la dépendance extrême sont examinées, car elles correspondent plus directement à l’objectif du mémoire : prédire des niveaux de retour pluviométriques en tenant compte non seulement de la proximité entre stations, mais aussi de la similarité de leur comportement dans les événements rares.

3.3.1 Méthodes d’interpolation fondées sur la proximité géographique

L’interpolation spatiale constitue une alternative plus simple et interprétable que les modèles spatiaux intégrés présentés précédemment. Plutôt que de modéliser directement les paramètres de la distribution des extrêmes comme des fonctions de covariables spatiales, elle consiste à estimer d’abord les niveaux de retour aux stations observées, puis à transférer cette information vers un site non observé. Cette approche prolonge donc les deux sections précédentes : la théorie des valeurs extrêmes définit la quantité à prédire et discute la méthode d’ajustement optimale adaptée à des fortes pluies, tandis que l’estimation spatiale des extrêmes souligne l’intérêt d’exploiter l’information issue de plusieurs stations. L’interpolation réunit ces deux concepts de manière plus directe, en supposant généralement – dans les cas les plus simplistes – que les stations les plus proches du site cible sont aussi les plus pertinentes pour estimer son comportement extrême.

L’interpolation spatiale repose sur une idée simple : une quantité pluviométrique inconnue en un site cible peut être estimée comme une moyenne pondérée des valeurs observées ou estimées aux stations disponibles. Comme le rappellent Ly et al. [72], ces méthodes sont très utilisées en hydrologie parce que les modèles hydrologiques nécessitent souvent des champs continus de précipitation alors que les mesures proviennent de réseaux ponctuels de stations de relevés pluviométriques. Tabios et Salas [73] inscrivent cette idée dans un cadre général où les différentes méthodes d’interpolation se distinguent essentiellement par la manière dont les poids sont calculés. Cette approche est importante par rapport au sujet de ce mémoire, car l’IDW et les poids fondés sur une mesure de dépendance peuvent être vus comme des variantes d’une même logique : produire une estimation pondérée à partir des stations disponibles.

Parmi les méthodes les plus simples, l’Inverse Distance Weighting (IDW) est la plus utilisée. Elle attribue les poids en fonction de l’inverse de la distance entre le site cible et les stations observées, de sorte que les stations proches reçoivent davantage d’importance que les stations éloignées (cf. section 2.2.1). Cette méthode est attractive parce qu’elle est simple, rapide, déterministe et facile à interpréter. Ly et al. [72] la présentent ainsi comme une méthode couramment utilisée, notamment lorsque l’objectif est de produire rapidement une surface pluviométrique à partir d’un réseau de stations. Tabios et Salas [73] montrent également que, même si l’IDW est moins performant que des méthodes plus élaborées comme le krigeage ou l’interpolation optimale dans leur étude, il fournit des résultats jugés plutôt satisfaisants par rapport à des méthodes plus rudimentaires comme les polygones de Thiessen ou certaines interpolations polynomiales.

Cette simplicité constitue toutefois aussi la principale limite de l’IDW (cf. section 2.2.3). Plusieurs études montrent que cette hypothèse est trop restrictive. Ly et al. [72] soulignent que le choix de la fonction de pondération, notamment la puissance appliquée à la distance, reste en partie subjectif et que la méthode ne fournit pas directement de mesure d’incertitude. Caloiero et al. [74] confirment cette limite dans une comparaison de méthodes appliquées aux précipitations mensuelles en Nouvelle-Zélande : l’IDW donne les coefficients les plus faibles parmi les méthodes testées et n’est jamais sélectionné comme meilleur interpolateur selon les critères de validation considérés. De manière similaire, Das [75] observe que l’IDW est moins performant que le krigeage ordinaire pour l’estimation des facteurs de croissance et de l’indice pluviométrique utilisés dans l’estimation de pluies extrêmes en sites non observés.

Les études comparatives montrent néanmoins que l’IDW ne doit pas être écarté trop rapidement. Sa performance dépend fortement du type de variable interpolée, de l’échelle temporelle, de la densité du réseau et de la structure spatiale du phénomène. Ly et al. [72] rappellent qu’aucune méthode d’interpolation ne domine toutes les autres pour toutes les situations, et que la méthode optimale dépend surtout du contexte. Dans certains cas, notamment lorsque le réseau de stations est dense ou lorsque la variabilité spatiale est relativement régulière, une méthode simple comme l’IDW peut donc constituer une référence utile. C’est aussi ce que suggèrent indirectement Tabios et Salas [73] : même si les méthodes statistiques sont globalement meilleures dans leur application, l’IDW reste compétitif par rapport à d’autres méthodes déterministes simples. Ces conclusions soulignent un constat largement partagé dans la littérature : il n’existe pas de méthode universelle, et le choix doit être effectué parmi plusieurs approches en

fonction du contexte.

L'application aux Pays-Bas étudiée par Soenario et Sluiter [44] illustre bien le compromis qui doit être fait entre simplicité et réalisme. Dans leur comparaison de méthodes d'interpolation de pluie journalière, l>IDW produit des structures spatiales globalement cohérentes avec les images radar, mais génère aussi des motifs circulaires irréalistes autour de certaines stations, souvent qualifiés de « bulls-eye patterns ». Les auteurs testent une version lissée de l>IDW afin de réduire cet effet, mais celle-ci produit une surface trop lisse et perd une partie de la variabilité spatiale visible dans les images radar.

Une difficulté supplémentaire apparaît lorsque l'on n'interpole plus des précipitations observées, mais des niveaux de retour ou des paramètres de lois d'extrêmes. Dans ce cas, les valeurs disponibles aux stations ne sont pas des mesures directes : ce sont déjà des estimations statistiques, obtenues après ajustement d'un modèle d'extrêmes. Blanchet et Lehning [76] discutent cette situation en comparant deux approches pour cartographier des niveaux de retour de hauteur de neige : interpoler les paramètres GEV estimés station par station, ou ajuster directement un modèle GEV spatial lisse à l'ensemble des stations. Leur étude montre que l'interpolation des paramètres GEV, notamment par IDW ou krigeage, peut fournir une première cartographie des niveaux de retour, mais que sa performance se dégrade dans les zones où le réseau de stations manque de densité. Le modèle GEV spatial lisse obtient alors de meilleurs résultats sur les stations de validation.

3.3.2 Méthodes géostatistiques et prise en compte de la structure spatiale

Les méthodes géostatistiques prolongent les approches simples comme l>IDW en remplaçant la pondération directement fondée sur la distance par une pondération estimée à partir de la structure spatiale des données. Dans l>IDW, la distance intervient directement dans le calcul des poids. Dans le krigeage, elle intervient à travers un semi-variogramme, qui décrit la manière dont la dissimilarité entre observations augmente avec leur séparation spatiale. Cette différence est importante : le krigeage ne suppose pas seulement que les stations proches sont plus informatives, il cherche à quantifier cette relation à partir des données disponibles. Ly et al. [72] présentent ainsi le krigeage comme une famille de méthodes capables de tenir compte de l'autocorrélation spatiale et de fournir une mesure d'incertitude de prédiction. Goovaerts [77] souligne également que l'intérêt de la géostatistique est d'exploiter la corrélation spatiale entre observations voisines, plutôt que de se limiter à une règle purement géométrique.

Les comparaisons empiriques montrent généralement que le krigeage améliore les interpolations fondées uniquement sur la distance, mais que cette amélioration dépend fortement du contexte. Dans l'étude de Goovaerts [77], menée sur des précipitations mensuelles et annuelles au Portugal, les méthodes géostatistiques produisent des erreurs plus faibles que les méthodes ignorant la dépendance spatiale, comme les polygones de Thiessen ou l>IDW. Soenario et Sluiter [44] obtiennent également de meilleures performances pour le krigeage ordinaire que pour l>IDW dans l'interpolation des pluies journalières aux Pays-Bas, tout en montrant que les surfaces obtenues peuvent rester globalement proches. Dans leur comparaison, la principale faiblesse de l>IDW réside surtout dans la présence d'artefacts circulaires irréalistes autour de certaines stations, comme expliqué dans la section 3.3.1.

Cette nuance apparaît aussi dans les travaux portant sur les précipitations extrêmes. Szolgay et al. [26] montrent que l>IDW et le krigeage peuvent produire des cartes relativement similaires, alors que la méthode du nearest neighbour génère des structures spatiales trop simples. Ils soulignent surtout que, dans l'étude des extrêmes, la question n'est pas seulement de choisir une méthode d'interpolation, mais aussi de décider quelle quantité doit être interpolée : les observations, les paramètres de la loi d'extrêmes ou les quantiles finaux. Das et al. [78] reprennent directement cette question en comparant l'interpolation des L-moments, des paramètres de la GEV et des quantiles de période de retour. Leur conclusion est que l'interpolation directe des quantiles est plus simple à appliquer et à interpréter, mais qu'elle peut être moins robuste pour les périodes de retour élevées, car ces quantiles dépendent fortement du paramètre de forme. Dans ce mémoire, c'est précisément une interpolation des niveaux de retour z_T qui est effectuée. Les limites associées à ce choix devront donc être prises en compte dans la discussion des résultats.

Une autre force des méthodes géostatistiques est leur capacité à intégrer des covariables spatiales ou physiques. Dans les variantes comme le krigeage avec dérive externe, le co-krigeage ou le krigeage des résidus, l'interpolation ne dépend plus seulement de la structure spatiale estimée entre stations, mais aussi de variables disponibles sur l'ensemble du domaine. Goovaerts [77] montre ainsi que l'intégration de l'altitude peut améliorer la prédiction des précipitations, tout en soulignant que la régression seule n'est pas suffisante si les résidus restent spatialement corrélés. Ly et al. [72] vont dans le même sens : dans leur revue, les méthodes géostatistiques multivariées sont souvent améliorées par l'utilisation de covariables comme l'altitude, les données radar ou d'autres informations auxiliaires, surtout pour des pas de temps

courts ou des réseaux de stations peu denses. Toutefois, l'utilité des covariables n'est pas automatique. Das et al. [78] montrent que le krigeage avec dérive externe est préférable lorsqu'une covariable significative est identifiée, tandis que le rapport du KNMI indique que l'altitude apporte peu d'information supplémentaire aux Pays-Bas, probablement en raison du relief limité [44]. Des variables comme la direction du vent ou la distance à la mer peuvent alors être plus pertinentes.

Les méthodes géostatistiques présentent aussi un avantage pratique par rapport à l>IDW : elles fournissent une incertitude de prédiction. Le rapport du KNMI illustre cette propriété avec les cartes de variance du krigeage, dont les valeurs augmentent notamment aux frontières du domaine, lorsque les sites à prédire sont moins bien entourés par les observations [44]. Ceresetti et al. [79] associent également une incertitude aux niveaux de retour interpolés par krigeage et montrent que celle-ci peut rester élevée dans le cas des pluies extrêmes. Cette limite rappelle que la performance finale dépend autant du modèle d'extrêmes choisi que de l'interpolateur utilisé. Dans leur comparaison, remplacer le modèle de maxima annuels par une approche par dépassement de seuil améliore davantage les niveaux de retour estimés que remplacer le krigeage par un réseau de neurones [79].

Ainsi, les méthodes géostatistiques occupent une position intermédiaire entre l'interpolation par distance et les modèles spatiaux d'extrêmes plus intégrés. Elles dépassent l>IDW en estimant une structure spatiale à partir des données et en permettant, dans certaines variantes, d'intégrer des covariables physiques. Elles restent cependant limitées par la qualité du semi-variogramme, la densité du réseau, la pertinence des covariables et la stabilité des quantités extrêmes interpolées. Le krigeage modélise donc une dépendance spatiale globale, souvent fondée sur l'ensemble de la variabilité observée, mais il ne caractérise pas nécessairement la dépendance spécifique dans la queue de distribution. C'est pour cette raison que la sous-section suivante s'intéresse à une autre stratégie : construire les poids d'interpolation à partir de mesures de dépendance propres aux extrêmes.

3.3.3 Dépendance spatiale des extrêmes et interpolation par similarité extrême

Les méthodes géostatistiques présentées précédemment modélisent une structure spatiale globale, généralement à travers une covariance ou un semi-variogramme. Cette logique est pertinente pour représenter la variabilité spatiale moyenne d'un champ pluviométrique, mais elle n'a pas été construite pour être adaptée aux extrêmes. En effet, deux stations peuvent présenter des précipitations ordinaires relativement corrélées sans nécessairement connaître simultanément des événements extrêmes. Cette différence est importante dans la littérature sur les extrêmes spatiaux : Davison et al. [80] rappellent que les événements rares sont par nature peu nombreux, ce qui rend l'inférence incertaine, et que cette difficulté est amplifiée dans un cadre spatial, où l'on cherche à extrapoler vers un espace de grande dimension. De manière similaire, Boluwade [19] souligne que les modèles géostatistiques classiques sont insuffisants pour représenter les caractéristiques spatiales propres aux extrêmes, car ils décrivent surtout des niveaux moyens du processus plutôt que la dépendance dans la queue de distribution.

Cette limite motive l'utilisation de cadres spécifiquement construits pour les extrêmes spatiaux. Les processus max-stables constituent l'une des extensions naturelles de la théorie univariée des valeurs extrêmes pour l'adapter au cadre spatial (cf. section 2.3.1). Ils permettent de séparer deux aspects du problème : d'une part, les distributions marginales des maxima, généralement modélisées par des lois GEV ; d'autre part, la structure de dépendance entre les extrêmes observés en différents sites. Davison et al. [80] comparent plusieurs familles de modèles d'extrêmes spatiaux, notamment les modèles à variables latentes, les copules et les processus max-stables. Leur synthèse montre que les modèles à variables latentes peuvent bien ajuster les distributions marginales, mais qu'ils représentent moins bien les distributions conjointes des extrêmes. Les copules adaptées aux extrêmes et les processus max-stables apparaissent donc plus appropriés lorsque l'objectif est de représenter la dépendance extrême elle-même.

La littérature appliquée confirme cette séparation entre modélisation marginale et dépendance extrême. Dans son étude, Van de Vyver [66] modélise les paramètres de la GEV en Belgique comme des fonctions de covariables géographiques ou climatologiques, tout en corrigeant l'incertitude des estimations pour tenir compte de la dépendance spatiale entre stations. Cette approche vise principalement à produire des cartes de niveaux de retour, tout en montrant que l'hypothèse d'indépendance spatiale entre maxima est difficilement acceptable. Boluwade [19] va plus loin en combinant une GEV spatiale, dont les paramètres dépendent notamment de la latitude, de la longitude, de la topographie et du temps, avec un modèle max-stable, qui caractérise la dépendance entre maxima annuels à partir d'un F-madogramme. Yang et al. [81] adoptent une logique similaire dans leur étude des précipitations extrêmes en Allemagne : les marges sont décrites par une GEV spatiale, tandis que la dépendance spatiale est représentée par un

modèle max-stable.

Pour résumer cette dépendance, la littérature utilise souvent le coefficient extrême, défini dans la section 2.3.1. Cette mesure ne décrit pas toute la loi jointe des extrêmes, mais elle fournit un indice interprétable de la propension de deux stations à connaître simultanément des événements rares. Davison et al. [80] présentent ce coefficient comme une mesure synthétique de dépendance extrême, et Yang et al. [81] l'utilisent pour comparer la dépendance spatiale des extrêmes simulés par différents modèles climatiques.

Le F-madogramme, directement lié au coefficient extrême 2.3.1, fournit alors un outil opérationnel pour estimer cette dépendance à partir des données. Vannitsem et Naveau [55] utilisent cette idée pour étudier les dépendances entre maxima de précipitations en Belgique. Leur analyse montre que la dépendance entre extrêmes varie selon la distance entre stations, mais aussi selon la saison et la durée d'accumulation considérée (heures, jours, mois). Les extrêmes hivernaux apparaissent plus spatialement cohérents que les extrêmes estivaux, et l'augmentation de la durée d'accumulation renforce la dépendance spatiale. Ce résultat est important : il montre que la proximité géographique ne suffit pas à caractériser la similarité entre stations, puisque deux stations séparées par une même distance peuvent avoir une dépendance extrême différente selon le régime météorologique ou la durée étudiée.

Ces résultats fournissent un argument direct pour construire des poids d'interpolation à partir d'une notion de similarité extrême. Dans l'IDW, les poids reposent uniquement sur la distance au site cible. Dans une approche fondée sur le F-madogramme, l'idée est différente : une station doit recevoir un poids élevé si elle présente un comportement extrême similaire à celui du site cible, c'est-à-dire si la dissimilarité extrême estimée est faible. Cette logique est cohérente avec les travaux de Vannitsem et Naveau [55], qui montrent que la dépendance extrême ne décroît pas uniquement de manière mécanique avec la distance, et avec ceux de Davison et al. [80], qui insistent sur la nécessité de modéliser explicitement la dépendance jointe des extrêmes. Elle reste tout de même plus simple qu'un modèle max-stable complet : il ne s'agit pas ici d'estimer toute la loi spatiale des extrêmes, mais d'exploiter une mesure bivariable de similarité pour améliorer une interpolation de niveaux de retour.

Cette approche comporte tout de même quelques limites. Le F-madogramme est une mesure empirique, donc sensible à la longueur des séries, à la qualité de l'ajustement marginal et au petit nombre d'événements extrêmes réellement informatifs. Vannitsem et Naveau [55] rappellent également que l'estimation repose sur des hypothèses fortes, notamment l'homogénéité et la stationnarité temporelle du champ des maxima. Yang et al. [81] montrent en outre que la dépendance spatiale et temporelle des extrêmes peut évoluer avec le climat, avec des changements différenciés selon la saison. Une mesure de dépendance estimée sur une certaine période ne doit donc pas être interprétée comme une propriété fixe du système pluviométrique.

Enfin, l'utilisation du F-madogramme pour construire des poids d'interpolation pose une difficulté pratique : si le site cible est non observé, sa dépendance empirique avec les stations voisines n'est pas directement disponible. Une des stratégies possible – testée dans ce mémoire – est alors de relier la dissimilarité extrême à des covariables décrivant les paires de sites, comme la distance géographique, la différence d'altitude, la distance à la côte ou l'orientation relative. Cette stratégie se rapproche des approches spatiales avec covariables utilisées par Van de Vyver [66], Boluwade [19] et Yang et al. [81], mais elle les mobilise dans un objectif plus restreint : prédire une mesure de similarité extrême afin de construire des poids d'interpolation. Cette approche occupe donc une position intermédiaire entre l'IDW et les modèles spatiaux d'extrêmes complets : elle conserve la simplicité d'une interpolation pondérée, tout en remplaçant la seule proximité géographique par une proximité fondée sur le comportement des extrêmes.

3.4 Synthèse et positionnement du mémoire

La revue de littérature met en évidence trois constats principaux. Le premier est que la théorie des valeurs extrêmes fournit un cadre adapté pour résumer le comportement des précipitations rares à travers les niveaux de retour z_T . Dans le cas des maxima annuels, la loi GEV constitue le modèle de référence pour relier une période de retour T à une intensité de précipitation attendue. Toutefois, l'estimation marginale station par station reste fragile, en particulier lorsque les séries disponibles sont courtes ou lorsque la période de retour considérée devient élevée. Cette fragilité est principalement liée à l'incertitude sur le paramètre de forme ξ , dont la valeur est fort variable en fonction des techniques d'ajustement et de la taille de l'échantillon.

Le deuxième constat est que l'information spatiale ne peut pas être ignorée. Les stations météorologiques voisines ne peuvent pas être considérées comme des sources d'information indépendantes : elles

sont soumises à des mécanismes climatiques, géographiques et topographiques communs. Les modèles spatiaux de valeurs extrêmes répondent directement à cette difficulté en faisant varier les paramètres de la loi GEV selon des covariables spatiales, telles que la latitude, la longitude, l'altitude ou la distance à la côte. Cette approche est statistiquement plus intégrée qu'une estimation station par station, puisqu'elle permet de mutualiser l'information entre les sites. En revanche, elle suppose de spécifier correctement la relation entre les paramètres de la GEV et les covariables disponibles. Une spécification trop simple peut manquer des variations locales importantes, tandis qu'une spécification trop complexe peut devenir instable lorsque le réseau de stations est limité.

Le troisième constat concerne les méthodes d'interpolation. Les approches fondées uniquement sur la distance géographique, comme l'IDW, sont simples, interprétables et faciles à appliquer, mais elles reposent sur une hypothèse forte : deux stations proches dans l'espace seraient nécessairement plus informatives l'une pour l'autre que deux stations éloignées. Cette hypothèse est souvent raisonnable pour l'analyse de précipitations moyennes, mais elle devient plus discutable pour les extrêmes. Deux sites proches peuvent présenter des comportements extrêmes différents si ils possèdent des caractéristiques géographiques différentes. Inversement, deux sites plus éloignés peuvent être soumis à des mécanismes pluviométriques similaires et donc présenter une dépendance extrême plus forte que ne le suggère leur seule distance.

C'est pour palier cette limitation de l'interpolation géographique que s'inscrit l'intérêt du F-madogramme. Contrairement à une pondération fondée uniquement sur la proximité spatiale, le F-madogramme fournit une mesure empirique de la similarité du comportement extrême entre deux stations. Son lien avec le coefficient extrême permet de l'interpréter comme un indicateur de dépendance spécifique aux événements rares. Cependant, cette mesure n'est directement observable qu'entre stations pour lesquelles des séries temporelles sont disponibles. Pour prédire un niveau de retour en un site non observé, une des solutions peut être d'estimer la dépendance extrême attendue entre ce site cible et les stations voisines à partir de covariables décrivant les paires de sites, comme la distance géographique, la différence d'altitude, la distance à la côte ou l'orientation relative.

Ce mémoire a pour objectif d'effectuer une comparaison de performances entre différents modèles construits à partir de la littérature. Trois approches sont étudiées afin d'évaluer différentes manières de transférer l'information spatiale pour l'estimation des niveaux de retour associés à des pluies extrêmes aux Pays-Bas. La première approche combine un ajustement marginal de la loi GEV à chaque station avec une interpolation IDW des niveaux de retour estimés. Elle sert de méthode de référence simple, fondée uniquement sur la distance. La deuxième approche conserve l'ajustement marginal des GEV, mais remplace les poids géographiques par des poids fondés sur une similarité extrême estimée à partir du F-madogramme et de covariables spatiales. Elle occupe une position intermédiaire : elle garde la simplicité d'une interpolation pondérée, tout en introduisant une information directement liée à la dépendance des extrêmes. La troisième approche repose sur un modèle GEV spatial, dans lequel les paramètres de la distribution sont modélisés directement en fonction de covariables géographiques ou physiques.

L'objectif du mémoire n'est donc pas de construire un modèle spatial complet de la loi jointe des extrêmes pluviométriques, comme le feraient certains modèles max-stables ou modèles à variables latentes. L'objectif est le suivant : déterminer si une mesure dépendance extrême peut améliorer l'interpolation des niveaux de retour par rapport à une simple interpolation IDW, et comparer cette stratégie à une modélisation GEV spatiale intégrée. Dans le contexte des Pays-Bas, cette comparaison permet d'évaluer si la proximité géographique suffit à prédire les niveaux de retour ou si une information plus spécifique au comportement extrême apporte un gain réel.

Le mémoire exploite donc un assemblage de trois familles de méthodes : l'estimation marginale des valeurs extrêmes, l'interpolation spatiale des quantités estimées et la modélisation de la dépendance spatiale des extrêmes. L'objectif est d'examiner empiriquement si une interpolation par similarité extrême peut constituer une alternative pertinente entre l'IDW, simple mais peu spécifique aux extrêmes, et les modèles GEV spatiaux, plus cohérents statistiquement mais plus sensibles aux choix de spécification.

4 Données

4.1 Origine des données

Les données météorologiques et spatiales utilisées dans ce mémoire proviennent du site du [Koninklijk Nederlands Meteorologisch Instituut \(KNMI\)](#). Les données ont été retenues principalement pour des raisons d’accessibilité, le site de l’Institut météorologique des Pays-Bas mettant à disposition des séries homogènes, de bonne qualité et directement utilisables. Dans un premier temps, l’entièreté des stations météorologiques (51 au total) disponibles dans la partie « Weerstations » ont été considérées, en extrayant les données de cumul quotidien des précipitations (RH) du 1^{er} janvier 1951 au 31 décembre 2024. Pour la suite de ce mémoire, les précipitations sont toujours exprimées dans l’unité initiale de la variable RH, c’est-à-dire en dixièmes de millimètre (ou [0.1 mm]). Les données ont été extraites dans deux jeux de données. Le premier jeu de données contient les données de géolocalisation (latitude, longitude, altitude) des stations ainsi que leurs noms et numéros respectifs (STN), auquel a été ajouté la distance à la mer de chaque station au moyen des packages R `sf` et `rnaturalearth` [82, 83]. Le second jeu de données contient les quantités de précipitations journalières du 1^{er} janvier 1951 au 31 décembre 2024 pour chacune des stations (représentées par leur numéro STN).

4.2 Disponibilité, sélection et prétraitement des données

Les données ont ensuite été filtrées de manière à obtenir, pour chaque station, une série quotidienne continue couvrant la période la plus longue possible, avec un minimum de 40 années d’observations afin de garantir une taille d’échantillon suffisante pour appliquer la théorie des valeurs extrêmes par blocs annuels. La figure 2 illustre la présence et l’absence de données dans le jeu de données RH pour chaque station. Les segments noirs correspondent aux périodes où des mesures quotidiennes sont disponibles, tandis que les segments gris indiquent des périodes sans données (valeurs manquantes). Sur base de cette figure, treize stations ont été retenues pour l’application des méthodes : les stations 235, 240, 260, 270, 275, 280, 290, 310, 344, 350, 370, 375 et 380. En effet, seules ces stations disposent d’une série continue de mesures couvrant une période de 41 ans, du 1^{er} octobre 1984 au 31 mars 2025.

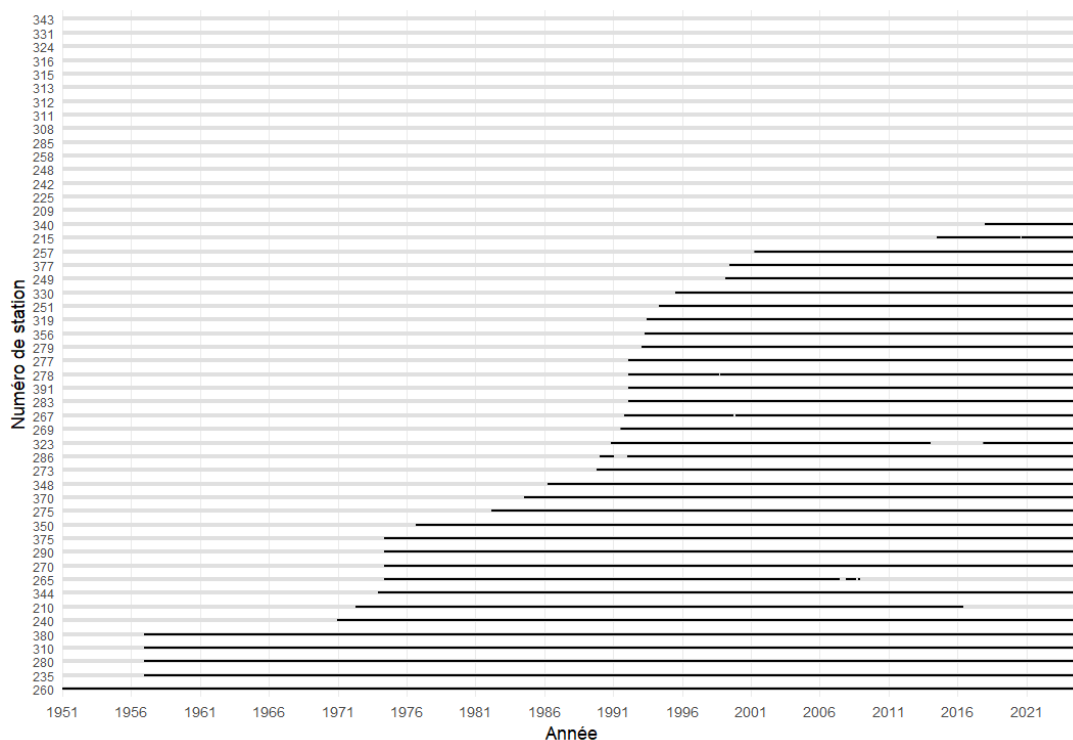


FIGURE 2 – Présence journalière des données de pluie par station.

Les données annuelles ont également été restreintes à la période octobre–mars. En effet, les préci-

pitations extrêmes de cette saison proviennent d’un mécanisme météorologique relativement homogène (précipitations frontales et stratiformes), ce qui permet de mieux satisfaire l’hypothèse de données indépendantes et identiquement distribuées (i.i.d.) sous-jacente à la théorie des valeurs extrêmes [84, 85]. Inclure les mois d’été introduirait des événements issus d’un régime pluviométrique différent (pluies convectives), ce qui pourrait modifier la structure de la queue de distribution et conduire à une estimation biaisée ou instable du paramètre de forme ξ . De plus, Vannitsem et Naveau [55] affirment que les extrêmes hivernaux apparaissent plus spatialement cohérents que les extrêmes estivaux, et seront donc plus faciles à ajuster. Le choix de la période a été guidé par le concept de « winter half-year », défini dans un rapport du KNMI publié par Van der Wiel et al. [86].

Cependant, la non-considération des mois d’été n’influencera pas l’interprétation de la période de retour : elle pourra toujours être exprimée en année. Il faudra néanmoins préciser, lors de l’interprétation des périodes de retour, que celles-ci s’appliquent uniquement aux événements extrêmes survenant entre octobre et mars, puisque l’ajustement de la GEV repose sur les maxima issus de cette période. Un niveau de retour z_T représente donc une quantité de précipitation qui est dépassée en octobre–mars tous les T ans en moyenne. Dans la suite du rapport, les années seront exprimées en années de saison : pour une saison hivernale du 1 octobre de l’année N au 31 mars de l’année $N+1$, l’année de saison est N .

Après cette sélection saisonnière, les quelques valeurs manquantes restantes ont été traitées. Le jeu de données retenu ne contenait que sept valeurs manquantes, toutes associées à la station 290. Ces valeurs concernaient les dates du 11, 12 et 13 décembre 2021, ainsi que les dates du 25, 26, 27 et 28 mars 2022. Comme ces observations manquantes sont très peu nombreuses et concernent une seule station, elles ont été remplacées par la valeur 0. Cela ne crée donc pas de maximum artificiel et n’influence pas les séries de maxima observées, puisque les maxima saisonniers de précipitations restent déterminés par les valeurs positives les plus élevées de chaque saison.

Le jeu de données final des maxima saisonniers a ensuite été construit à partir des données quotidiennes prétraitées. Comme expliqué plus tôt, les observations appartenant aux mois d’octobre, novembre et décembre ont été associées à l’année de saison correspondant à leur année calendaire, tandis que les observations de janvier, février et mars ont été associées à l’année de saison précédente. Les données ont ensuite été restreintes aux années de saison comprises entre 1984 et 2024. Pour chaque station et pour chaque année de saison, le maximum journalier observé entre octobre et mars a été extrait. Cette opération conduit à un tableau de maxima saisonniers contenant 41 observations par station, pour 13 stations, soit un total de 533 maxima hivernaux. Le jeu de données final possède donc 41 lignes, correspondant aux années de saison, et 13 colonnes, correspondant aux stations retenues. Le tableau complet des maxima saisonniers utilisés dans l’analyse est présenté en annexe A.

4.3 Analyse descriptive

4.3.1 Répartition des stations météorologiques

Le tableau 1 présente les caractéristiques physiques des 13 stations pluviométriques considérées dans ce mémoire. Il reprend, pour chaque station, son identifiant (STN), son nom, sa position géographique, son altitude ainsi que sa distance à la côte. Ces informations permettent de décrire la structure spatiale du réseau de stations et seront utilisées par la suite pour construire ou interpréter certaines covariables spatiales.

La figure 3¹ complète cette description en représentant la localisation géographique des stations retenues. Les stations couvrent une grande partie du territoire des Pays-Bas, avec des points situés aussi bien à proximité immédiate du littoral qu’à l’intérieur du pays. Cette répartition spatiale permettra d’étudier l’influence de la distance géographique, de l’altitude, de la distance à la côte et de l’orientation entre stations sur la structure des extrêmes pluviométriques.

La carte montre également que le réseau de stations n’est pas parfaitement homogène dans l’espace. Certaines stations sont relativement isolées, notamment aux extrémités nord, sud et est du domaine, tandis que d’autres sont situées dans des zones plus centrales. Cette configuration sera donc une problématique à prendre en compte dans ce mémoire : les stations situées en bordure du domaine disposent de moins de voisins proches dans toutes les directions, ce qui peut influencer la qualité des estimations obtenues par interpolation spatiale.

1. Toutes les cartes présentées dans ce mémoire ont été obtenues en utilisant le package R `leaflet` [87]

TABLEAU 1 – Caractéristiques physiques des 13 stations pluviométriques utilisées

STN	Nom	Latitude [°N]	Longitude [°E]	Altitude [m]	Distance à la côte [km]
235	De Kooy	52.928	4.781	1.2	1.9
240	Schiphol	52.318	4.790	-3.3	18.8
260	De Bilt	52.100	5.180	1.9	52.9
270	Leeuwarden	53.224	5.752	1.2	11.6
275	Deelen	52.056	5.873	48.2	98.4
280	Eelde	53.125	6.585	5.2	32.9
290	Twenthe	52.274	6.891	34.8	109.9
310	Vlissingen	51.442	3.596	8.0	0.5
344	Rotterdam	51.962	4.447	-4.3	19.3
350	Gilze-Rijen	51.566	4.936	14.9	45.6
370	Eindhoven	51.451	5.377	22.6	74.9
375	Volkel	51.659	5.707	22.0	100.0
380	Maastricht	50.906	5.762	114.3	109.2

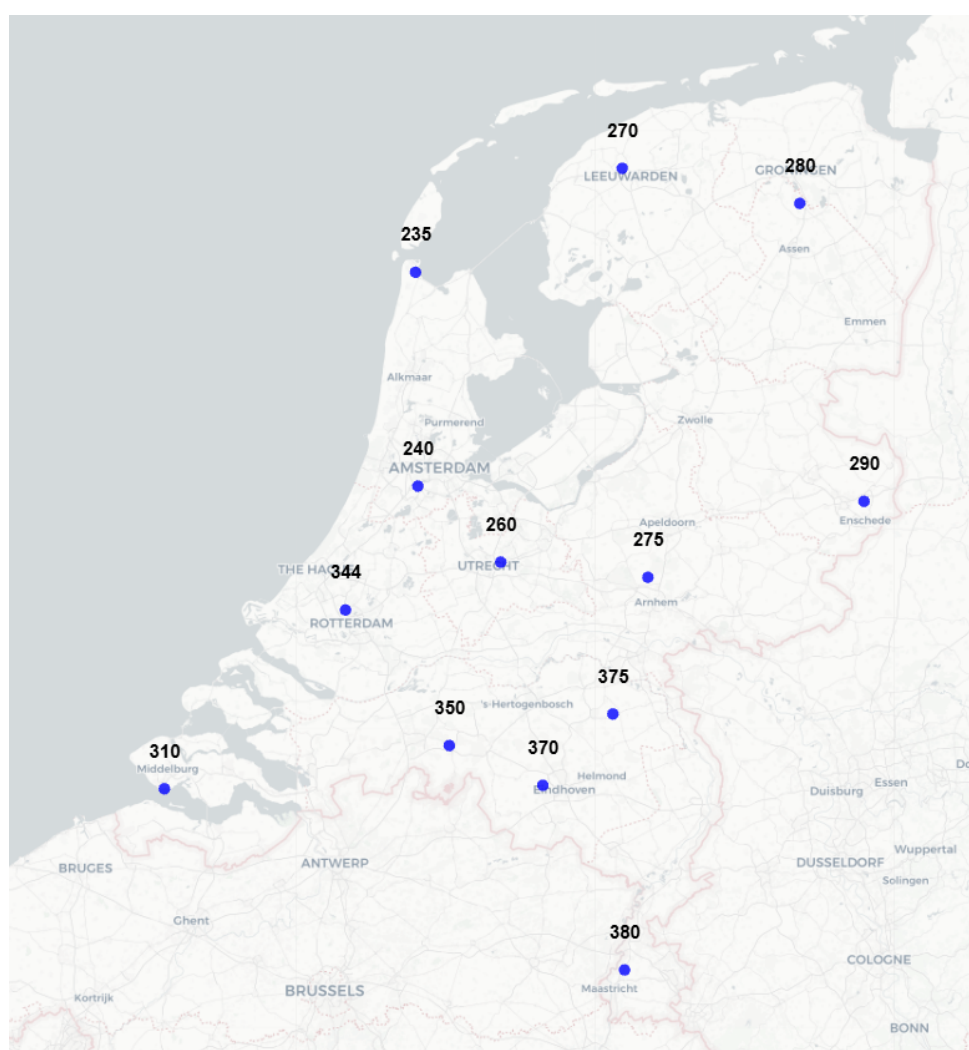


FIGURE 3 – Localisation géographique des 13 stations de relevés météorologiques retenues

4.3.2 Distribution des précipitations journalières hivernales

Avant d'étudier directement les maxima saisonniers, il est utile de décrire la distribution des précipitations journalières hivernales. Cette analyse permet de caractériser la structure générale des observations brutes utilisées pour construire les séries de maxima.

La figure 4 représente la distribution globale des précipitations journalières hivernales, toutes stations confondues. Cette distribution est très asymétrique : elle est fortement dominée par les précipitations faibles ou nulle, tandis que seules quelques observations atteignent des niveaux beaucoup plus élevés. Cette forme est cohérente avec la nature des données pluviométriques journalières, qui comprennent une proportion importante de jours sans pluie ou avec de faibles quantités de pluie, ainsi qu'un nombre limité d'événements intenses. La présence d'une queue droite marquée justifie donc l'intérêt d'une approche centrée sur les valeurs extrêmes pour analyser les fortes précipitations.

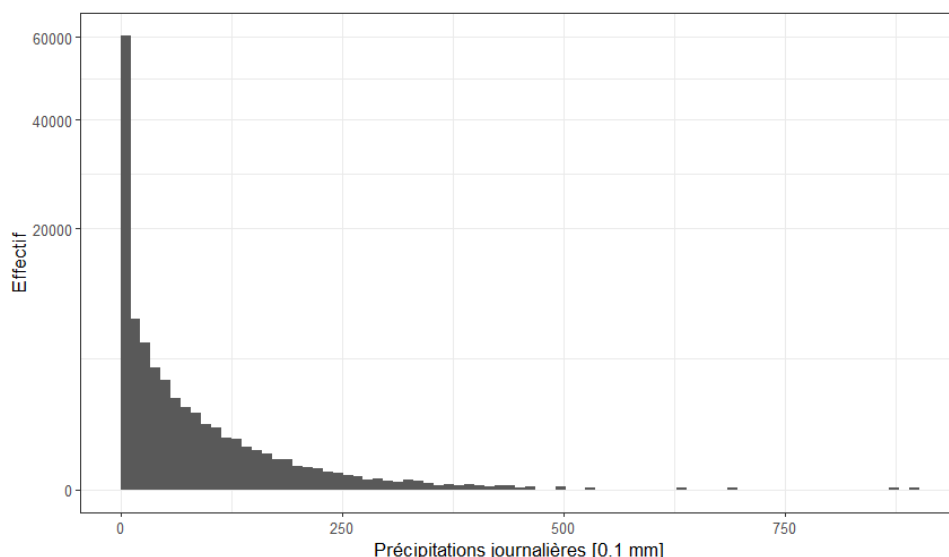


FIGURE 4 – Distribution globale des précipitations journalières hivernales, toutes stations confondues.

La figure 5 représente la distribution des précipitations journalières hivernales station par station. Les boxplots présentent des structures globalement similaires entre les stations, qui rejoignent la distribution globale : les boîtes sont concentrées sur les faibles valeurs, tandis que les observations les plus extrêmes apparaissent comme des points isolés au-dessus de la distribution centrale. Les stations présentent globalement des comportements journaliers globalement comparables, même si certaines différences apparaissent dans l'intensité des valeurs les plus fortes.

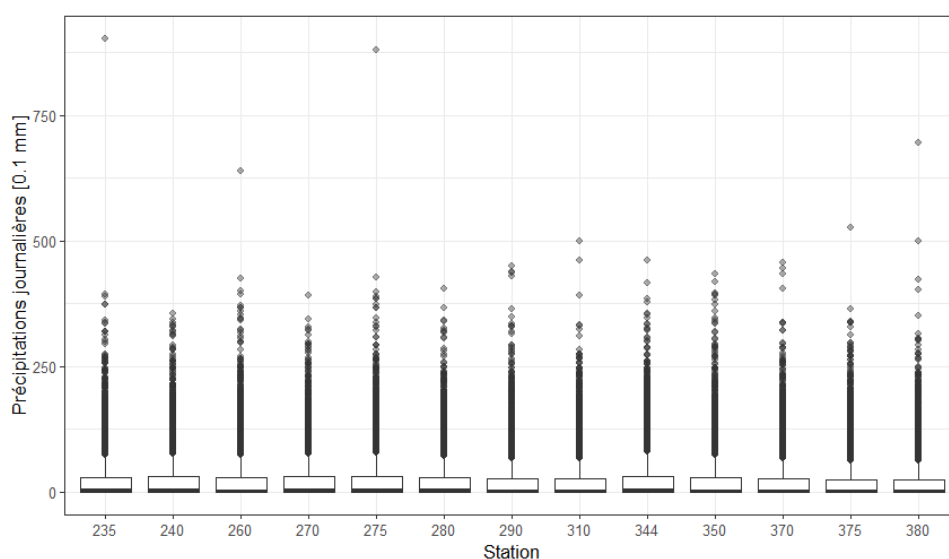


FIGURE 5 – Distribution des précipitations journalières hivernales par station.

Les deux figures montrent donc que la distribution des pluies est dominée par les quantités faibles

ou nulles de précipitations, avec quelques événements de forte intensité situés dans la queue droite. Cela motive l'utilisation de la théorie des valeurs extrêmes si l'on veut étudier ces événements extrêmes. Dans la suite, l'analyse portera donc directement sur les maxima de précipitations annuels pour chaque station. C'est cette étape d'extraction des maxima par blocs qui permet ensuite d'utiliser la loi GEV comme modèle pour la modélisation de la distribution des pluies extrêmes.

4.3.3 Distribution des maxima de précipitations hivernaux

Après avoir exploré les données brutes de précipitation, il convient ensuite d'analyser les maxima annuels de précipitations, obtenus pour chaque station et pour chaque année de saison entre octobre et mars. Chaque station possède donc une série de 41 maxima hivernaux, correspondant aux années de saison 1984 à 2024. Les valeurs sont exprimées dans l'unité initiale de la variable RH, c'est-à-dire en dixièmes de millimètre.

La figure 6 représente la distribution des maxima saisonniers pour chacune des 13 stations retenues. Les distributions présentent une structure globalement comparable entre les stations : les médianes se situent dans des ordres de grandeur proches et les intervalles interquartiles sont relativement similaires. Cette similarité indique que les maxima hivernaux observés aux différentes stations appartiennent à un même contexte pluviométrique général.

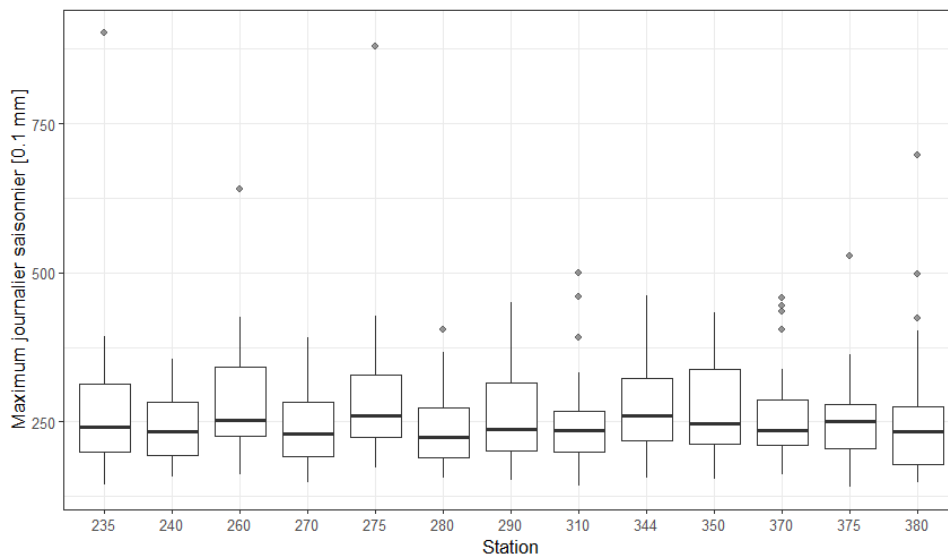


FIGURE 6 – Distribution des maxima saisonniers de précipitations par station.

Les boxplots montrent toutefois que les stations ne sont pas parfaitement équivalentes. Certaines présentent une dispersion plus importante ou des observations extrêmes isolées, visibles sous forme de points situés au-dessus de la distribution centrale. Ces valeurs élevées sont précisément celles qui motivent l'utilisation d'une approche par valeurs extrêmes. Elles indiquent également que, même si les distributions sont comparables dans leur structure générale, les paramètres des lois GEV ajustées pourront varier d'une station à l'autre.

Les médianes visibles sur la figure 6 sont représentées spatialement sur la carte de la figure 7, ce qui permet d'observer leur répartition géographique. La carte montre que les médianes des maxima saisonniers ne sont pas réparties de manière parfaitement uniforme entre les stations. On observe une zone regroupant les médianes les plus élevées au centre de la carte. Certaines stations proches ont également des médianes assez différentes, ce qui suggère que la distance seule ne suffit pas à expliquer les différences entre les stations.

Cette carte doit cependant être interprétée comme une analyse descriptive. La médiane permet de visualiser les différences de niveau entre stations, mais ne renseigne pas directement sur la variabilité ou sur le comportement de la queue de distribution. Les différences spatiales présentes justifient néanmoins l'intérêt d'une approche d'interpolation, puisque les quantités liées aux extrêmes ne semblent pas identiques d'une station à l'autre.

La figure 8 présente l'évolution temporelle des maxima saisonniers pour l'ensemble des stations. Ce

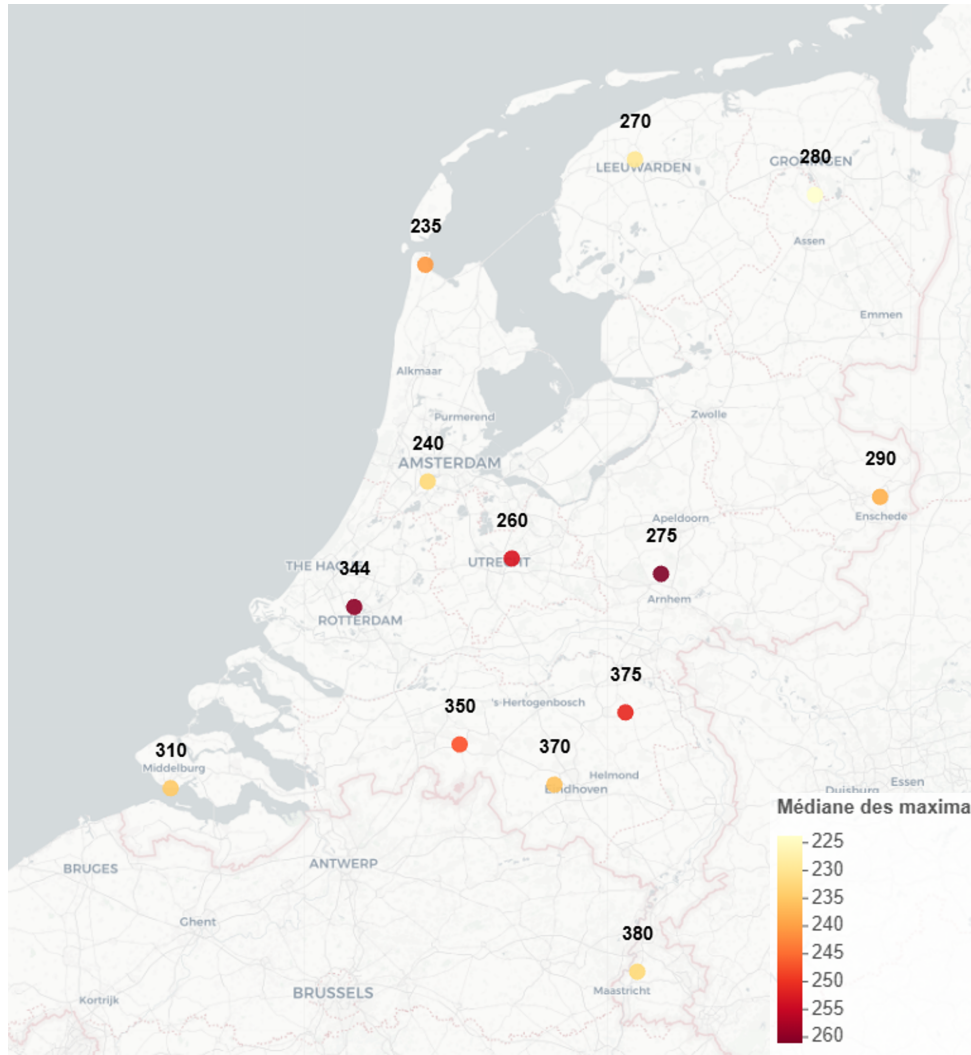


FIGURE 7 – Répartition spatiale des médianes des maxima saisonniers de précipitations [0.1mm].

graphique permet de repérer les années durant lesquelles certains maxima particulièrement élevés ont été observés. Ces pics ne concernent pas toujours les mêmes stations, ce qui illustre la variabilité spatiale des événements extrêmes. La courbe moyenne, représentée en noir, ne met pas en évidence de tendance temporelle nette sur la période étudiée. Les maxima saisonniers semblent donc globalement compatibles avec une hypothèse de stationnarité temporelle sur la période retenue, ce qui est nécessaire pour l'ajustement de lois GEV stationnaires aux séries de maxima.

En résumant, l'analyse descriptive des données met en évidence trois informations importantes pour la suite du mémoire. D'abord, les précipitations journalières hivernales présentent une queue droite marquée, ce qui justifie l'intérêt d'une approche par valeurs extrêmes. Ensuite, les maxima saisonniers ont des distributions globalement comparables entre stations, tout en conservant des différences locales de niveau et de dispersion. Enfin, ces différences possèdent une dimension spatiale, liée à la position des stations dans le domaine étudié. Ces trois éléments justifient donc le recours à la théorie des valeurs extrêmes dans ce mémoire, dont l'objectif est de comparer plusieurs méthodes d'estimation et d'interpolation des niveaux de retour associés aux précipitations extrêmes.

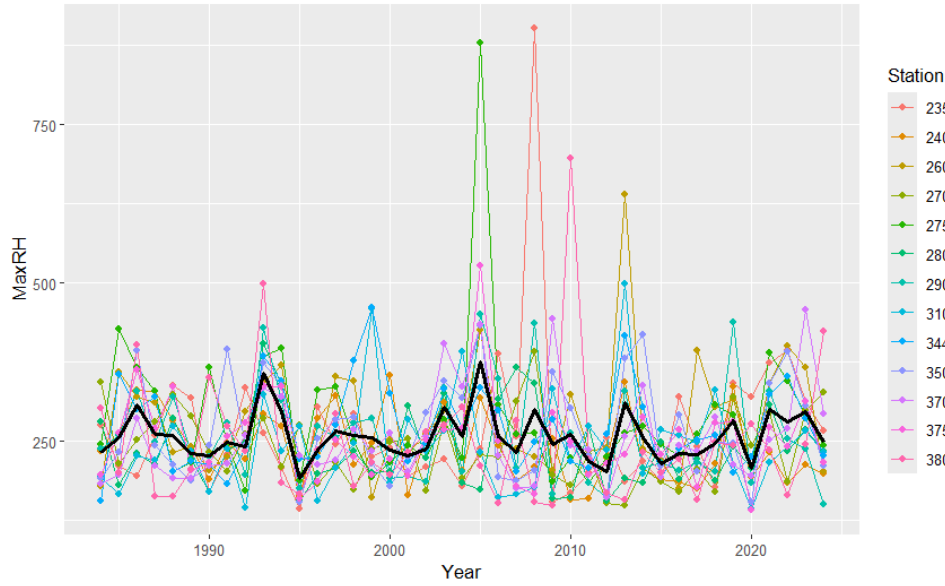


FIGURE 8 – Évolution temporelle des maxima saisonniers de précipitations entre 1984 et 2024 [0.1mm].

5 Méthodologie

5.1 Objectif général de l'étude

L'objectif général de cette étude est de comparer trois méthodes d'estimation des niveaux de retour de précipitations extrêmes en des stations temporairement retirées du jeu de données, pour comparer les capacités d'interpolation des trois méthodes. Le principe est d'ajuster les méthodes à partir des stations conservées, puis de prédire les niveaux de retour aux stations retirées afin de comparer les estimations obtenues.

Les trois méthodes étudiées reposent sur des concepts d'interpolation différents. La méthode 1a ajuste des GEV marginales à chacune des stations conservées puis interpole les niveaux de retour par pondération inverse de la distance (IDW). La méthode 1b suit la même logique d'ajustement marginal, mais utilise des poids fondés sur une mesure de dépendance entre stations. Enfin, la méthode 2 repose sur un modèle GEV spatial ajusté sur l'ensemble des stations conservées, dans lequel les paramètres de la loi GEV dépendent de covariables spatiales.

La comparaison est d'abord menée sur données simulées, car ce cadre permet de répéter l'expérience et de comparer les prédictions à des niveaux de retour théoriques connus. Deux cadres de simulation sont utilisés. Le premier repose sur un modèle GEV spatial ajusté aux données réelles. Il génère une structure spatiale dans les paramètres marginaux, mais ne modélise pas explicitement la dépendance entre stations. Ce cadre peut toutefois influencer la comparaison entre les méthodes. D'une part, il peut favoriser la méthode 2, puisque les données simulées sont générées à partir d'un modèle GEV spatial et que cette méthode repose elle-même sur l'ajustement d'un modèle de même nature. D'autre part, il peut défavoriser la méthode 1b, car aucune dépendance explicite n'est introduite entre les stations lors de la simulation, alors que cette méthode cherche précisément à exploiter cette information pour construire ses poids d'interpolation.

Le second cadre repose sur un modèle max-stable ajusté aux données réelles. Celui-ci conserve des marges GEV dépendant des mêmes covariables spatiales, mais introduit aussi une dépendance extrême entre les stations. Il permet donc de compléter le premier cadre en évaluant les méthodes, en particulier la méthode 1b, dans un contexte où une structure de dépendance est effectivement présente.

Enfin, les trois méthodes sont appliquées au jeu de données réel. Cette application permet d'examiner leur comportement sur les observations disponibles, sans dépendre directement d'un mécanisme de simulation.

5.2 Notation et validation par retrait de stations

On note $\mathcal{S} = \{s_1, \dots, s_p\}$ l'ensemble des stations disponibles, où p désigne le nombre total de stations. Pour une station $s_i \in \mathcal{S}$, on note $X_t(s_i)$ le maximum hivernal observé à l'année de saison t , avec $t = 1, \dots, n$. Le niveau de retour de période T à la station s_i est noté $z_T(s_i)$. Son estimation par une méthode m donnée sera notée $\hat{z}_{T,m}(s_i)$, avec $m \in \{\text{idw}, \text{dep}, \text{spa}\}$.

Afin d'évaluer les performances des différentes méthodes, une procédure de validation par retrait de stations est mise en place. Dans le cadre des simulations, deux stations — toujours les mêmes — sont exclues du jeu de données avant l'ajustement des modèles, ce choix permettant d'accélérer le calcul par rapport à une validation *leave-one-out* complète. Pour l'application aux données réelles, une approche par retrait successif d'une station (*leave-one-station-out*) sera en revanche utilisée, afin de pallier l'absence de plusieurs jeux de données indépendants.

L'objectif est de comparer la capacité des méthodes à prédire les niveaux de retour en des localisations non utilisées lors de l'estimation. Pour les simulations, les stations retirées sont la station 260, située au centre du domaine, et la station 310, située davantage en périphérie du domaine, comme présenté sur la figure 3. Ce choix permet d'évaluer les performances des méthodes dans deux configurations spatiales différentes : une station entourée par plusieurs stations voisines et une station située dans une position plus périphérique.

On note $\mathcal{S}_{test}$ l'ensemble des stations retirées, avec ici pour les simulations

$$\mathcal{S}_{test} = \{260, 310\},$$

et $\mathcal{S}_{train}$ l'ensemble des stations conservées pour l'ajustement des méthodes :

$$\mathcal{S}_{train} = \{235, 240, 270, 275, 280, 290, 344, 350, 370, 375, 380\}.$$

Les différentes méthodes sont donc ajustées uniquement à partir des stations appartenant à $\mathcal{S}_{train}$.

Pour chaque méthode m , l'objectif est donc d'obtenir une prédiction $\hat{z}_{T,m}(s_0)$ pour chaque station retirée, avec $s_0 \in \mathcal{S}_{test}$ pour les simulations, et $s_0 \in \mathcal{S}$ pour les données réelles. Ces prédictions sont ensuite comparées à une valeur de référence. Dans les simulations, cette référence correspond au niveau de retour théorique $z_T(s_0)$ issu du modèle générateur. Dans l'application aux données réelles, le vrai niveau de retour n'étant pas connu, la référence correspond au niveau de retour estimé directement à partir des observations disponibles à la station retirée, noté $\hat{z}_{T,\text{ref}}(s_0)$.

5.3 Cadres de simulation

Les simulations sont utilisées afin de comparer les trois méthodes dans un cadre contrôlé. Les niveaux de retour théoriques $z_T(s_i)$ utilisés comme référence découlent directement du modèle générateur. Il devient alors possible d'évaluer plus systématiquement les erreurs commises par chaque méthode aux stations retirées appartenant à $\mathcal{S}_{test}$.

Deux cadres de simulation sont considérés dans cette étude. Le premier repose sur un modèle GEV spatial ajusté aux données réelles, tandis que le second cadre repose sur un modèle max-stable. Afin que les deux cadres restent comparables, les covariables utilisées pour modéliser les paramètres marginaux des distributions sont d'abord déterminées à partir du modèle GEV spatial. Les mêmes covariables sont ensuite utilisées dans le modèle max-stable pour les marges GEV. Ce choix permet de conserver une structure spatiale marginale similaire dans les deux mécanismes de simulation. La différence entre les deux cadres est ainsi uniquement la présence ou non d'une dépendance extrême explicite entre les stations.

5.3.1 Simulation à partir d'un modèle GEV spatial

Le premier cadre de simulation repose sur un modèle GEV spatial ajusté aux données réelles présentées dans la section 4 et en annexe A. L'objectif est d'obtenir un modèle générateur dont les paramètres marginaux varient spatialement selon les caractéristiques géographiques des stations. Pour ce faire, un modèle GEV spatial est ajusté au moyen de la fonction `fitspatgev()` du package `SpatialExtremes` [25].

Les covariables considérées pour chaque station sont la latitude, la longitude, l'altitude et la distance à la côte. Les paramètres de position et d'échelle peuvent dépendre de ces covariables spatiales, tandis que le paramètre de forme est supposé constant sur l'ensemble du domaine (voir section 2.1.5). Le modèle complet initial peut donc être représenté de manière générale par

$$\mu(s_i) = f_\mu(\text{lat}_i, \text{lon}_i, \text{alt}_i, d_{\text{coast}_i}),$$

$$\sigma(s_i) = f_\sigma(\text{lat}_i, \text{lon}_i, \text{alt}_i, d_{\text{coast}_i}),$$

et

$$\xi(s_i) = \xi,$$

où $\mu(s_i)$, $\sigma(s_i)$ et $\xi(s_i)$ désignent respectivement les paramètres de position, d'échelle et de forme de la distribution GEV à la station s_i .

À partir de ce modèle complet, une procédure de sélection basée sur le critère TIC (section 2.1.5) est ensuite appliquée afin d'obtenir un modèle plus parcimonieux. Cette sélection est réalisée sous la contrainte que toute covariable présente dans le paramètre d'échelle doit également être présente dans le paramètre de position. Autrement dit, si l'on note $\mathcal{C}_\mu$ l'ensemble des covariables retenues pour $\mu(s_i)$ et $\mathcal{C}_\sigma$ l'ensemble des covariables retenues pour $\sigma(s_i)$, la sélection impose la contrainte

$$\mathcal{C}_\sigma \subset \mathcal{C}_\mu.$$

Cette contrainte permet d'éviter qu'une covariable explique la variabilité de l'échelle sans être également prise en compte dans la position.

Les paramètres estimés du modèle GEV spatial retenu sont ensuite utilisés pour générer $N = 100$ jeux de données simulés au moyen de la fonction `rgev` du package R `evd` [88]. Pour chaque simulation, $n = 50$ maxima hivernaux sont générés aux stations considérées à partir des distributions GEV définies par les paramètres estimés. Le nombre de maxima n a été choisi pour rester proche des données réelles disponible ($n = 41$). Les stations simulées possèdent donc des distributions marginales spatialement structurées, puisque les paramètres de ces distributions dépendent des covariables associées à chaque station.

Ce cadre de simulation permet ainsi de produire des niveaux de retour $z_T(s_i)$ présentant une structure spatiale induite par les covariables, pour trois périodes de retour : période courte $T = 10$ ans, période intermédiaire $T = 50$ ans et période longue $T = 100$ ans. Cette variation de période de retour permettra de comparer les performances des méthodes à différents horizons de prédiction. La période intermédiaire correspond approximativement à la période de temps utilisée dans notre jeu de données initial (41 ans). Dans ce contexte, la période de retour $T = 10$ ans permettra d'évaluer les performance à des horizons inférieurs à cette dernière, et la période de retour $T = 100$ ans permettra d'évaluer les capacité d'extrapolation des méthodes par rapport à la période de temps utilisée pour créer le jeu de données.

Ce cadre de simulation ne modélise toutefois pas explicitement la dépendance entre les stations : conditionnellement aux paramètres marginaux estimés, les valeurs simulées aux différentes stations ne sont pas générées à partir d'une loi jointe de dépendance extrême. Ce point est important pour l'interprétation des résultats, en particulier pour la méthode 1b, qui repose précisément sur l'utilisation d'une information de dépendance entre stations.

Il est important de préciser que le modèle générateur est initialement ajusté à partir des données des treize stations, y compris les stations 260 et 310, avant la séparation entre $\mathcal{S}_{\text{train}}$ et $\mathcal{S}_{\text{test}}$. Ce choix permet d'estimer la structure spatiale des paramètres marginaux sur l'ensemble du domaine et d'obtenir un mécanisme de simulation représentatif des données observées. Cela introduit néanmoins une forme de fuite d'information, puisque les observations des stations retirées lors des simulations participent à la calibration du modèle générateur. Étant donné le faible nombre de stations disponible pour établir la structure spatiale des paramètres, la conservation des stations auxquelles seront testées les performances des différents modèles est justifiable, dans le but d'obtenir un modèle générateur le plus adapté possible à la réalité. De plus, étant donné que le même modèle générateur et les mêmes valeurs de référence sont utilisés pour les trois méthodes, cette calibration commune ne favorise pas directement l'une d'elles lors de la comparaison relative. Les résultats seront donc interprétés comme une évaluation conditionnelle à un scénario simulé calibré sur l'ensemble du domaine, et non comme une validation externe totalement indépendante.

5.3.2 Simulation à partir d'un modèle max-stable

Le second cadre de simulation repose sur l'ajustement d'un modèle max-stable. Contrairement au modèle GEV spatial décrit précédemment, ce modèle ne vise pas uniquement à représenter la variation spatiale des paramètres marginaux, mais également à introduire une structure de dépendance extrême entre les stations.

Dans un premier temps, les données sont transformées vers des marges unit Fréchet. La loi unit Fréchet, de fonction de répartition

$$F(z) = \exp(-1/z), \quad z > 0,$$

est couramment utilisée comme marge standard dans les modèles max-stables. En imposant la même distribution marginale à chaque station, cette transformation permet de retirer l'effet des marges GEV propres à chaque station et de se concentrer sur la structure de dépendance entre les maxima. Autrement dit, les différences de niveau, d'échelle ou de forme entre stations sont temporairement neutralisées afin que le modèle max-stable soit ajusté sur la dépendance spatiale des extrêmes plutôt que sur les différences marginales.

Un modèle de dépendance est ensuite ajusté à ces données transformées à l'aide de la fonction `fitmaxstab()` du package `SpatialExtremes` [25]. Le modèle retenu est un modèle max-stable de Brown–Resnick. Ce type de modèle est couramment utilisé pour représenter la dépendance spatiale des extrêmes [80, 85, 89]. Il repose sur l'idée que les maxima observés en différents sites appartiennent à un même champ spatial extrême, dont la dépendance décroît avec la distance entre les stations. Dans ce cadre, la structure de dépendance est notamment décrite par deux paramètres : le paramètre de portée, ou **range**, qui contrôle l'étendue spatiale de la dépendance, et le paramètre de régularité, ou **smooth**, qui contrôle la forme de la décroissance de cette dépendance avec la distance.

Les paramètres de dépendance estimés à partir de ce modèle sont ensuite utilisés pour générer les jeux de données simulés. La simulation est réalisée en deux étapes. Premièrement, $N = 100$ jeux de données à marges unit Fréchet sont générés à l'aide de la fonction `rmaxstab()` du package `SpatialExtremes` ($n = 50$ maxima comme pour le modèle de GEV spatiale). Ces simulations reproduisent la structure de dépendance spatiale estimée par le modèle max-stable, mais elles sont encore exprimées sur l'échelle unit Fréchet.

Deuxièmement, les simulations unit Fréchet sont transformées vers des marges GEV à l'aide de la fonction `frech2gev()` du même package. Cette transformation est réalisée en utilisant les mêmes covariables marginales que celles retenues dans le modèle GEV spatial optimal. Ainsi, les paramètres de position, d'échelle et de forme des distributions GEV simulées conservent la même structure spatiale marginale que dans le cadre de simulation fondé sur le modèle GEV spatial.

Ce second cadre permet ainsi de générer des jeux de données qui présentent à la fois des distributions marginales spatialement structurées et une dépendance spatiale entre les extrêmes. Il fournit donc un contexte plus adapté à l'évaluation de la méthode 1b, puisque cette méthode repose précisément sur l'utilisation d'une information de dépendance pour construire les poids d'interpolation.

5.4 Méthode 1a : ajustements marginaux et interpolation IDW

La méthode 1a combine un ajustement marginal des valeurs extrêmes à chaque station conservée et une interpolation spatiale des niveaux de retour par pondération inverse de la distance. Elle est appliquée séparément à chacun des $N = 100$ jeux de données issus des deux cadres de simulation.

Pour chaque jeu simulé, une distribution GEV est ajustée indépendamment aux maxima hivernaux de chaque station $s_i \in \mathcal{S}_{train}$. Les paramètres de position, d'échelle et de forme sont ainsi estimés séparément pour chaque station. La procédure d'estimation des paramètres de la GEV est présentée dans la section 2.1.3, et celle-ci est effectuée aux moyens de la fonction `gev.fit` du package `ismev` [90].

À partir des paramètres estimés, les niveaux de retour $\hat{z}_T(s_i)$ sont calculés pour chacune des stations de $\mathcal{S}_{train}$ (voir formule section 2.1.3). Trois périodes de retour sont considérées afin d'évaluer les performances de la méthode pour différents degrés d'extrapolation : $T = 10$ ans, $T = 50$ ans et $T = 100$ ans. Les niveaux de retour ainsi estimés constituent ensuite les valeurs à interpoler vers les stations retirées.

Pour chaque station cible $s_0 \in \mathcal{S}_{test}$ et chaque période de retour $T \in \{10, 50, 100\}$, le niveau de retour $\hat{z}_{T,idw}(s_0)$ est obtenu par interpolation IDW à partir des niveaux de retour estimés aux stations conservées (méthode présentée dans la section 2.2.1). Les distances entre les stations sont calculées à partir de leurs coordonnées géographiques au moyen de la distance de Haversine (section 2.2.2).

La procédure est appliquée successivement aux stations 260 et 310 appartenant à $\mathcal{S}_{test}$. Pour chaque jeu simulé, elle fournit donc une prédiction $\hat{z}_{T,idw}(s_0)$ pour chaque combinaison de station retirée et de période de retour. Ces estimations sont ensuite comparées aux niveaux de retour théoriques associés au modèle générateur afin d'évaluer les performances de la méthode.

La méthode 1a est appliquée de manière identique aux données simulées par le modèle GEV spatial et par le modèle max-stable. Elle ne tient compte ni des covariables propres aux stations ni de la dépendance extrême entre leurs maxima : seule la distance géographique intervient dans la construction des poids.

Lors de l'application aux données réelles, le même protocole est utilisé. Les GEV marginales sont ajustées aux observations des stations gardées, puis les niveaux de retour associés à $T = 10$, $T = 50$ et $T = 100$ ans sont interpolés vers la station retirée à chaque étape du *leave-one-station-out*.

5.5 Méthode 1b : ajustements marginaux et interpolation par dépendance

La méthode 1b reprend le principe général de la méthode 1a, mais remplace les poids fondés uniquement sur la distance géographique par des poids construits à partir de la dépendance extrême estimée entre les stations. Elle combine ainsi un ajustement marginal des maxima à chaque station conservée, une modélisation de la dépendance entre les paires de stations de $\mathcal{S}_{train}$, puis une interpolation des niveaux de retour vers les stations de $\mathcal{S}_{test}$.

La procédure est appliquée séparément à chacun des $N = 100$ jeux de données issus des deux cadres de simulation. Comme dans la méthode 1a, une distribution GEV est d'abord ajustée indépendamment aux maxima hivernaux de chaque station $s_j \in \mathcal{S}_{train}$, pour chaque jeu simulé. À partir des paramètres marginaux estimés, les niveaux de retour $\hat{z}_T(s_j)$ sont calculés pour les trois périodes de retour considérées, soit $T \in \{10, 50, 100\}$.

Parallèlement à ces ajustements marginaux, la dépendance extrême est estimée entre toutes les paires de stations conservées aux moyens du F-madogramme. Comme expliqué dans la section 2.3.1, il est d'abord nécessaire de transformer les observations sur une échelle marginale commune. Pour chaque station $s_j \in \mathcal{S}_{train}$ et chaque saison $t = 1, \dots, n$, une pseudo-observation uniforme est obtenue à partir du rang du maximum observé :

$$U_t(s_j) = \frac{\text{rang}\{X_t(s_j)\}}{n+1}.$$

La division par $n+1$, plutôt que par n , évite d'obtenir des valeurs exactement égales à 1. Cette transformation permet de comparer les positions relatives des maxima dans leurs distributions respectives, indépendamment des différences de position, d'échelle ou de forme entre les stations.

Toutes les paires non ordonnées (s_j, s_k) de stations appartenant à $\mathcal{S}_{train}$ sont ensuite construites, avec $j < k$. Puisque $\mathcal{S}_{train}$ contient onze stations, la modélisation repose sur

$$\binom{11}{2} = 55$$

paires de stations. Pour chaque paire, le F-madogramme empirique (la formule du F-madogramme théorique est exposée dans la section 2.3.1) est calculé selon

$$d_{jk} = \frac{1}{2n} \sum_{t=1}^n |U_t(s_j) - U_t(s_k)|.$$

La quantité d_{jk} mesure la dissimilarité entre les comportements extrêmes des stations s_j et s_k . Dans le cadre max-stable, elle appartient à l'intervalle $[0, \frac{1}{6}]$ (voir section 2.3.1). Une valeur proche de 0 traduit une forte dépendance extrême, tandis qu'une valeur proche de $\frac{1}{6}$ correspond à une situation proche de l'indépendance.

Afin de relier cette dissimilarité à des caractéristiques spatiales des paires de stations, quatre covariables sont calculées. La première, notée h_{jk} , correspond à la distance de Haversine exprimée en kilomètres entre les stations s_j et s_k . La deuxième est la différence absolue d'altitude,

$$\Delta_{alt,jk} = |\text{alt}(s_j) - \text{alt}(s_k)|.$$

Les deux dernières covariables décrivent l'orientation spatiale du segment reliant les stations. On note ϕ_{jk} le *bearing* de la station s_j vers la station s_k , mesuré dans le sens horaire à partir du nord et exprimé en radians. L'orientation est représentée par les deux composantes

$$\cos(2\phi_{jk}) \quad \text{et} \quad \sin(2\phi_{jk}).$$

Le facteur 2 permet de traiter l'orientation comme une direction non orientée. En effet, les directions ϕ_{jk} et $\phi_{jk} + \pi$, qui correspondent au même axe parcouru dans les deux sens opposés, fournissent les mêmes valeurs de $\cos(2\phi_{jk})$ et de $\sin(2\phi_{jk})$. Ces covariables, décrites plus en détail dans la section 2.3.2, permettent de prendre en compte une éventuelle anisotropie de la dépendance extrême.

Le F-madogramme étant borné entre 0 et $\frac{1}{6}$, il ne peut pas être utilisé directement comme réponse dans une régression linéaire sans risquer d'obtenir des prédictions situées en dehors de cet intervalle. Il est donc d'abord ramené sur l'intervalle $[0, 1]$ par la transformation $6d_{jk}$, puis transformé par la fonction logit. La variable réponse utilisée dans la régression est ainsi définie par

$$\psi_{jk} = \log \left(\frac{6d_{jk}}{1 - 6d_{jk}} \right).$$

Un modèle linéaire complet est ensuite ajusté aux 55 paires de stations conservées :

$$\psi_{jk} = \beta_0 + \beta_h h_{jk} + \beta_{alt} \Delta_{alt,jk} + \beta_c \cos(2\phi_{jk}) + \beta_s \sin(2\phi_{jk}) + \varepsilon_{jk}.$$

Ce modèle vise à expliquer la variation de la dissimilarité extrême à partir de la distance géographique, de la différence d'altitude et de l'orientation entre les stations. Par exemple, un coefficient positif associé à h_{jk} indique qu'une augmentation de la distance tend à augmenter ψ_{jk} et donc le F-madogramme, ce qui correspond à une diminution de la dépendance extrême.

À partir du modèle complet, une sélection pas à pas dans les deux directions est réalisée selon l'AIC. Cette sélection est répétée séparément pour chaque jeu de données simulé. La structure retenue peut ainsi varier d'une simulation à l'autre selon la dépendance estimée entre les stations de $\mathcal{S}_{train}$.

Une fois le modèle sélectionné, les covariables sont calculées entre chaque station conservée $s_j \in \mathcal{S}_{train}$ et chaque station cible $s_0 \in \mathcal{S}_{test}$. Ces valeurs sont ensuite introduites dans le modèle sélectionné afin de prédire

$$\hat{\psi}_{j0}.$$

La transformation logit est ensuite inversée afin de revenir à l'échelle du F-madogramme. La dissimilarité prédite entre la station conservée s_j et la station cible s_0 est donnée par

$$\hat{d}_{j0} = \frac{1}{6} \frac{\exp(\hat{\psi}_{j0})}{1 + \exp(\hat{\psi}_{j0})}.$$

Cette transformation garantit que $\hat{d}_{j0}$ appartient à l'intervalle $[0, \frac{1}{6}]$. La dissimilarité prédite est ensuite convertie en une mesure de dépendance comprise entre 0 et 1 :

$$W_j(s_0) = 1 - 6\hat{d}_{j0}.$$

Une valeur de $\hat{d}_{j0}$ proche de 0 conduit à un poids brut $W_j(s_0)$ proche de 1, ce qui signifie que la station s_j présente une forte dépendance extrême prédite avec la station cible. À l'inverse, lorsque $\hat{d}_{j0}$ est proche de $\frac{1}{6}$, le poids brut devient proche de 0, ce qui correspond à une dépendance faible ou à une situation proche de l'indépendance.

Les poids bruts sont enfin normalisés séparément pour chaque station cible :

$$w_j(s_0) = \frac{W_j(s_0)}{\sum_{s_k \in \mathcal{S}_{train}} W_k(s_0)} = \frac{1 - 6\hat{d}_{j0}}{\sum_{s_k \in \mathcal{S}_{train}} (1 - 6\hat{d}_{k0})}.$$

Ils vérifient ainsi

$$w_j(s_0) \geq 0 \quad \text{et} \quad \sum_{s_j \in \mathcal{S}_{train}} w_j(s_0) = 1.$$

Le niveau de retour à la station retirée est alors obtenu par une moyenne pondérée des niveaux de retour marginaux estimés aux stations conservées :

$$\hat{z}_{T,dep}(s_0) = \sum_{s_j \in \mathcal{S}_{train}} w_j(s_0) \hat{z}_T(s_j), \quad T \in \{10, 50, 100\}.$$

La procédure fournit donc, pour chaque jeu simulé, une estimation $\hat{z}_{T,dep}(s_0)$ pour chacune des stations $s_0 \in \mathcal{S}_{test}$ et chacune des périodes $T \in \{10, 50, 100\}$. Elle est appliquée de manière identique aux données simulées à partir du modèle GEV spatial et à celles issues du modèle max-stable. Dans le premier cadre, les dépendances estimées proviennent principalement de la variabilité d'échantillonnage, puisque les stations sont simulées sans dépendance explicite. Dans le second, la méthode peut exploiter la structure de dépendance introduite par le processus max-stable. L'interprétation des effets associés à la différence d'altitude et à l'orientation doit toutefois rester limitée. Le processus de Brown–Resnick utilisé étant isotrope et sa structure de dépendance dépendant uniquement de la distance entre les stations, aucune influence de l'altitude ou de la direction n'est donc introduite dans le mécanisme générateur [80]. Leur

interprétation sera dès lors surtout pertinente lors de l'application aux données réelles, tandis que les conséquences de leur inclusion dans les simulations seront discutées plus en détail lors de la discussion des résultats.

Lors de l'application aux données réelles, le même protocole est suivi. Les pseudo-observations, les F-madogrammes et le modèle de dépendance sont estimés à partir des stations gardées. À chaque étape du *leave-one-station-out*, les poids prédits entre ces stations et la station retirée sont ensuite utilisés pour interpoler les niveaux de retour correspondant à $T = 10$, $T = 50$ et $T = 100$ ans. Dans ce cas, il sera intéressant d'étudier les comportements de dépendance en fonction de la différence d'altitude et du bearing.

5.6 Méthode 2 : modèle GEV spatial

La méthode 2 repose sur l'ajustement d'un modèle GEV spatial aux maxima hivernaux des stations conservées. Contrairement aux méthodes 1a et 1b, elle ne consiste pas à ajuster séparément une distribution GEV à chaque station avant d'interpoler les niveaux de retour. Les paramètres des distributions GEV sont directement modélisés comme des fonctions de covariables spatiales, puis prédits aux stations retirées.

Pour chaque jeu de données simulé, le modèle est ajusté uniquement à partir des stations $s_i \in \mathcal{S}_{train}$. La procédure d'ajustement et de sélection est pratiquement identique à celle utilisée pour obtenir le modèle générateur dans le cadre GEV spatiale sans dépendance, décrite dans la section 5.3.1. La différence principale est que le modèle générateur est ajusté une seule fois à partir des treize stations afin de calibrer les simulations (voir section 5.3.1 pour la justification de l'utilisation de l'entièreté des stations), tandis que la méthode 2 réajuste séparément le modèle à chacun des $N = 100$ jeux simulés en utilisant uniquement les stations de $\mathcal{S}_{train}$.

Comme pour le modèle générateur, les covariables candidates sont la latitude, la longitude, l'altitude et la distance à la côte. Un modèle complet est d'abord ajusté à l'aide de la fonction `fitspatgev()` du package `SpatialExtremes` [25]. Les paramètres de position et d'échelle sont autorisés à dépendre des quatre covariables, tandis que le paramètre de forme est maintenu constant sur l'ensemble du domaine, conformément au modèle non stationnaire présenté dans la section 2.1.5.

La simplification du modèle suit également la procédure décrite dans la section 5.3.1. Les différentes structures candidates sont comparées à l'aide du TIC, et la sélection est réalisée sous la contrainte que toute covariable retenue pour le paramètre d'échelle soit également présente dans le paramètre de position. Pour chaque jeu simulé, le modèle retenu est donc celui qui minimise le TIC parmi les structures respectant cette contrainte. La sélection étant répétée indépendamment pour chaque simulation, les covariables retenues peuvent varier d'un jeu de données à l'autre.

Une fois le modèle sélectionné, les valeurs des covariables des stations $s_0 \in \mathcal{S}_{test}$ sont introduites dans les modèles estimés afin d'obtenir les paramètres prédits $\hat{\mu}(s_0)$, $\hat{\sigma}(s_0)$ et $\hat{\xi}$. Les niveaux de retour sont ensuite calculés à partir de ces paramètres, pour les mêmes trois périodes de retour $T \in \{10, 50, 100\}$.

La méthode 2 ne réalise donc pas une interpolation directe des niveaux de retour estimés aux stations voisines. Elle estime d'abord une relation commune entre les covariables spatiales et les paramètres de la GEV à partir des stations de $\mathcal{S}_{train}$, puis utilise cette relation pour prédire la distribution des maxima aux stations retirées.

La méthode est appliquée de la même manière aux deux cadres de simulation. Dans le cadre fondé sur le modèle GEV spatial, elle réajuste un modèle de même nature que celui utilisé pour générer les données, ce qui peut lui conférer un avantage. Dans le cadre max-stable, elle reste adaptée à la modélisation des marges GEV et de leur variation spatiale, mais elle ne tient pas explicitement compte de la dépendance extrême introduite entre les stations. Cette distinction sera prise en compte dans la discussion des résultats.

Lors de l'application aux données réelles, la même procédure est suivie. La station retirée est exclue de l'ajustement, le modèle complet est estimé à partir des stations gardées, puis simplifié selon le TIC sous la contrainte $\mathcal{C}_\sigma \subseteq \mathcal{C}_\mu$. Les covariables de la station retirée sont enfin utilisées pour prédire ses paramètres GEV et calculer ses niveaux de retour aux trois périodes de retour.

5.7 Critères d'évaluation

Les performances des trois méthodes sont évaluées à partir de l'erreur quadratique moyenne, ou *Root Mean Squared Error* (RMSE), calculée entre les niveaux de retour prédits et les niveaux de retour utilisés comme référence. La procédure d'évaluation diffère légèrement entre les simulations, pour lesquelles les

niveaux de retour théoriques sont connus, et l'application aux données réelles, pour laquelle ils doivent être estimés.

5.7.1 Évaluation dans les cadres de simulation

Dans chaque cadre de simulation, les méthodes sont évaluées aux deux stations appartenant à $\mathcal{S}_{test}$, c'est-à-dire les stations 260 et 310. Pour chaque jeu simulé $b = 1, \dots, N$, avec $N = 100$, chaque méthode $m \in \{\text{idw}, \text{dep}, \text{spa}\}$ fournit une estimation $\hat{z}_{T,m}^{(b)}(s_0)$ du niveau de retour à la station retirée s_0 , pour les périodes $T \in \{10, 50, 100\}$.

Dans les simulations, le niveau de retour théorique $z_T(s_0)$ est connu à partir des paramètres du modèle générateur. Pour une méthode m , une station retirée s_0 et une période de retour T , le RMSE est donc calculé sur les N jeux simulés selon

$$\text{RMSE}_{m,T}(s_0) = \sqrt{\frac{1}{N} \sum_{b=1}^N \left[\hat{z}_{T,m}^{(b)}(s_0) - z_T(s_0) \right]^2}.$$

Ce critère combine le biais et la variabilité des estimations produites au cours des simulations. Une valeur faible indique que les niveaux de retour prédits sont, en moyenne, proches de la valeur théorique et relativement stables d'un jeu simulé à l'autre. À l'inverse, une valeur plus élevée correspond à une estimation des niveaux de retour plus éloignée du niveau de retour théorique, et/ou à une instabilité des estimations d'un jeu simulé à l'autre.

Le RMSE est calculé séparément pour chaque station de $\mathcal{S}_{test}$, afin de distinguer les performances obtenues au centre du domaine, à la station 260, de celles obtenues dans une position plus périphérique, à la station 310. Il est également calculé séparément pour chacune des trois périodes de retour, puisque l'incertitude et les erreurs d'estimation sont susceptibles d'augmenter lorsque la période de retour considérée devient plus longue.

Cette procédure est répétée dans les deux cadres de simulation. Les performances des méthodes sont ainsi comparées séparément pour les données générées à partir du modèle GEV spatial et pour celles générées à partir du modèle max-stable. Pour chaque cadre, les résultats comprennent donc une valeur de RMSE pour chaque combinaison de méthode, de période de retour et de station retirée. Le RMSE constitue le principal critère utilisé pour comparer les performances des trois méthodes dans les simulations, bien que d'autres éléments seront apportés dans les résultats pour étayer la discussion.

5.7.2 Évaluation sur les données réelles

Pour l'application aux données réelles, l'évaluation repose sur une procédure de validation par retrait successif d'une station, ou *leave-one-station-out*. Contrairement aux simulations, l'objectif n'est plus uniquement de prédire les niveaux de retour aux stations 260 et 310, mais d'évaluer les méthodes successivement à chacune des treize stations disponibles.

Pour une station $s_i \in \mathcal{S}$, celle-ci est temporairement retirée du jeu de données. La méthode m considérée est alors ajustée à partir des stations restantes $\mathcal{S}_{-i}$, puis utilisée pour prédire le niveau de retour $\hat{z}_{T,m}(s_i)$ à la station retirée. La procédure est répétée jusqu'à ce que chacune des treize stations ait été utilisée une fois comme station test.

Contrairement au cadre simulé, le vrai niveau de retour $z_T(s_i)$ n'est pas connu. La valeur de référence, notée $\hat{z}_{T,\text{ref}}(s_i)$, est donc obtenue par l'ajustement direct d'une GEV marginale aux maxima observés à la station s_i . Cette estimation est utilisée uniquement pour évaluer la prédiction produite lorsque la station est retirée; ses observations ne participent pas à l'ajustement de la méthode évaluée.

Pour une méthode m et une période de retour T , le RMSE sur les données réelles est calculé sur l'ensemble des $p = 13$ stations selon

$$\text{RMSE}_{m,T} = \sqrt{\frac{1}{p} \sum_{i=1}^p \left[\hat{z}_{T,m}(s_i) - \hat{z}_{T,\text{ref}}(s_i) \right]^2}.$$

Cette procédure fournit, pour chaque période de retour, trois valeurs de RMSE, soit une pour chacune des méthodes 1a, 1b et 2. Comme trois périodes de retour sont étudiées, neuf valeurs de RMSE sont obtenues au total pour l'application aux données réelles.

Afin de compléter ce critère et de limiter l'influence potentielle d'une station présentant une erreur particulièrement élevée, l'erreur absolue moyenne (*Mean Absolute Error*, MAE) est également calculée.

Contrairement au RMSE, qui accorde un poids important aux écarts les plus élevés, le MAE mesure directement l'amplitude moyenne des erreurs absolues. Pour une méthode m et une période de retour T , il est défini par

$$\text{MAE}_{m,T} = \frac{1}{p} \sum_{i=1}^p |\hat{z}_{T,m}(s_i) - \hat{z}_{T,\text{ref}}(s_i)|.$$

L'utilisation conjointe du RMSE et du MAE permet ainsi de distinguer une méthode présentant des erreurs globalement modérées d'une méthode dont la performance moyenne serait fortement dégradée par une ou quelques stations particulièrement mal prédites.

Les erreurs relatives sont également examinées pour chaque station, méthode et période de retour. Elles sont définies par

$$e_{i,m,T}^{\text{rel}} = 100 \frac{\hat{z}_{T,m}(s_i) - \hat{z}_{T,\text{ref}}(s_i)}{\hat{z}_{T,\text{ref}}(s_i)} \quad [\%].$$

Le signe de celles-ci permet de distinguer les cas de sous-estimation et de surestimation. Les erreurs relatives sont également cartographiées pour chacune des méthodes et des périodes de retour, afin d'identifier les stations pour lesquelles les prédictions sont les moins précises et d'évaluer dans quelle mesure certaines d'entre elles contribuent aux valeurs globales du RMSE et du MAE.

Ces critères complémentaires visent ainsi à nuancer l'interprétation du RMSE. Dans les simulations, les valeurs de référence correspondent aux niveaux de retour du modèle générateur et les performances peuvent être évaluées sur un grand nombre de répétitions, ce qui fournit une mesure relativement stable de l'erreur. Sur les données réelles, le RMSE repose au contraire sur seulement treize erreurs et sur des niveaux de retour de référence eux-mêmes estimés à partir des observations. Il peut dès lors être plus sensible à une erreur particulièrement importante sur une station et constituer un indicateur moins stable de la performance générale d'une méthode. L'analyse conjointe du RMSE, du MAE et de la distribution spatiale des erreurs relatives permet donc de fournir une évaluation plus complète et plus robuste des performances prédictives.

La validation *leave-one-station-out* permet d'utiliser successivement toutes les stations pour l'évaluation et fournit ainsi une comparaison des méthodes plus globale que le seul retrait des stations 260 et 310. Les résultats doivent néanmoins être interprétés avec prudence, puisque les valeurs de référence utilisées sur les données réelles sont elles-mêmes des estimations et non les vrais niveaux de retour inconnus.

6 Résultats et discussion des simulations

La section méthodologique a présenté les trois méthodes d'estimation et d'interpolation des niveaux de retour, ainsi que les deux cadres de simulation retenus pour évaluer leurs performances. Cette section expose et discute les résultats obtenus dans ces deux cadres. Les performances prédictives sont évaluées au moyen de l'indicateur RMSE. Les valeurs résultantes de RMSE pour chacune des deux stations retirées sont présentées pour les trois méthodes et pour les périodes de retour de 10, 50 et 100 ans. Pour les méthodes 1b et 2, le pourcentage de présence des covariables utilisées lors de l'ajustement et de la sélections des modèles est également présenté et discuté.

Les résultats obtenus sous le cadre GEV spatial sont d'abord présentés et interprétés, avant d'examiner ceux issus du cadre max-stable. Une comparaison transversale des deux cadres est ensuite menée afin de déterminer dans quelle mesure les conclusions relatives aux différentes méthodes dépendent du cadre de simulation considéré.

6.1 Résultats sous le cadre GEV spatial

6.1.1 Modèle générateur

La procédure décrite à la section 5.3.1 a conduit à sélectionner un modèle dans lequel le paramètre de position dépend de l'altitude et de la distance à la côte, tandis que les paramètres d'échelle et de forme sont constants dans l'espace. La latitude et la longitude ne sont donc pas retenues dans le modèle générateur final.

Les covariables ayant été centrées et réduites avant l'ajustement, le modèle retenu s'écrit comme suit :

$$\mu(s_i) = 220.373 - 8.977 \text{ alt}_i^* + 7.816 d_{\text{coast}_i}^*,$$

$$\sigma(s_i) = 52.940,$$

et

$$\xi(s_i) = 0.1293,$$

où alt_i^* et $d_{\text{coast}_i}^*$ désignent respectivement l'altitude et la distance à la côte centrées et réduites de la station s_i .

TABLEAU 2 – Paramètres du modèle GEV spatial retenu comme modèle générateur

Paramètre	Terme	Coefficient
Position μ	Intercept	220.373
	Altitude	−8.977
	Distance à la côte	7.816
Échelle σ	Intercept	52.940
Forme ξ	Intercept	0.1293

Ce modèle est ensuite utilisé pour générer les maxima hivernaux aux différentes stations dans le premier cadre de simulation.

6.1.2 Présentation des résultats

Le tableau 3 présente les performances obtenues par les trois méthodes dans le cadre de simulation fondé sur le modèle GEV spatial. Pour chaque jeu de données simulé, les stations 260 et 310 ont été retirées avant l'ajustement des modèles. Les niveaux de retour prédits à ces deux stations ont ensuite été comparés à leurs valeurs théoriques afin de calculer le RMSE sur l'ensemble des répétitions de simulation. Les résultats sont rapportés séparément pour les périodes de retour de 10, 50 et 100 ans.

TABLEAU 3 – RMSE obtenus pour les trois méthodes sous le cadre de simulation GEV spatial [0.1mm]

Station	Période de retour T (années)	Méthode 1a RMSE	Méthode 1b RMSE	Méthode 2 RMSE
260	10	11.7	17.1	8.7
	50	26.4	39.4	19.2
	100	38.4	55.8	26.6
310	10	11.5	16.2	16.3
	50	29.8	40.2	30.7
	100	44.7	58.7	39.5

Les RMSE obtenus sous le cadre GEV spatial augmentent systématiquement avec la période de retour, quelle que soit la méthode ou la station considérée. Les différences entre méthodes deviennent également plus marquées pour les périodes de retour élevées.

À la station 260, située au centre du domaine, la méthode 2 fournit les RMSE les plus faibles pour les trois périodes de retour, devant la méthode 1a, tandis que la méthode 1b présente les erreurs les plus importantes. À la station périphérique 310, les performances des méthodes 1a et 2 sont davantage comparables : la méthode 1a obtient les plus faibles erreurs pour $T = 10$ et $T = 50$, alors que la méthode 2 devient la plus performante pour $T = 100$. La méthode 1b reste globalement moins performante. La comparaison entre stations montre par ailleurs que les performances des méthodes 1a et 1b varient relativement peu entre 260 et 310, tandis que la méthode 2 se dégrade plus nettement à la station périphérique.

Les fréquences de sélection des covariables au cours des répétitions de simulation ont également été examinées pour les méthodes 1b et 2. Elles correspondent au pourcentage d'ajustements pour lesquels chaque covariable a été conservée à l'issue de la procédure de sélection du modèle.

Le tableau 4 présente les fréquences de sélection des covariables utilisées pour modéliser la dépendance spatiale dans la méthode 1b.

Les fréquences de sélection des covariables du modèle de dépendance de la méthode 1b sont faibles et proches les unes des autres. La distance géographique h est la variable la plus fréquemment retenue,

TABLEAU 4 – Fréquence de sélection des covariables du modèle de dépendance de la méthode 1b sous le cadre GEV spatial

Covariable	Fréquence de sélection
Distance géographique h	20 %
Différence d'altitude Δ_{alt}	18 %
$\cos(2\phi)$	15 %
$\sin(2\phi)$	19 %

mais seulement dans 20 % des ajustements, contre 18 % pour la différence d'altitude, 15 % pour $\cos(2\phi)$ et 19 % pour $\sin(2\phi)$. Aucune covariable ne semble donc constituer un prédicteur stable de la dépendance dans ce cadre de simulation.

Pour la méthode 2, la sélection des covariables est réalisée séparément pour les paramètres de localisation μ et d'échelle σ du modèle GEV spatial. Les fréquences de sélection obtenues sont présentées dans le tableau 5.

TABLEAU 5 – Fréquence de sélection des covariables du modèle GEV spatial de la méthode 2 sous le cadre GEV spatial

Covariable	Localisation μ	Échelle σ
Longitude lon	21 %	5 %
Latitude lat	29 %	8 %
Altitude alt	92 %	17 %
Distance à la côte d_{coast}	77 %	17 %

La structure sélectionnée par la méthode 2 est beaucoup plus stable pour le paramètre de position. L'altitude et la distance à la côte sont retenues dans respectivement 92 % et 77 % des ajustements, contre 21 % pour la longitude et 29 % pour la latitude. Les fréquences de sélection sont nettement plus faibles pour le paramètre d'échelle, avec un maximum de 17 %, ce qui indique qu'une structure plus simple est généralement retenue pour ce paramètre.

6.1.3 Discussion des résultats

Une première tendance commune aux trois méthodes est l'augmentation du RMSE avec la période de retour. Elle traduit la difficulté croissante d'estimer des niveaux de retour correspondant à des événements de plus en plus rares : lorsque T augmente, la prédiction repose davantage sur l'extrapolation dans la queue de la distribution et devient donc plus sensible aux erreurs d'estimation des paramètres.

La position de la station dans le domaine influence également les performances. La station centrale 260 est entourée de stations d'apprentissage dans toutes les directions, alors que la station 310 se situe en périphérie et dispose principalement d'informations provenant d'une seule direction. Cette asymétrie géographique peut avoir un effet limité pour un niveau de retour modéré, mais devenir plus pénalisante lorsque les différences entre les niveaux de retour sont amplifiées pour des périodes de retour plus élevées. C'est notamment le cas pour la méthode 1a. D'un autre côté, la méthode 2 semble davantage pâtir de ces différences de localisation, et ce pour chacune des périodes de retour.

La méthode 2 présente les meilleurs résultats à la station 260 pour les trois périodes de retour. Ce constat doit toutefois être interprété en tenant compte du modèle générateur. Les données de ce cadre sont simulées à partir d'un modèle GEV spatial dont la structure appartient directement à la famille de modèles utilisée par cette méthode. Elle bénéficie donc d'un avantage de spécification puisqu'elle cherche à reconstruire une relation entre les paramètres marginaux de la GEV et les covariables géographiques similaire à celle utilisée pour générer les données. Cet avantage ne garantit cependant pas une meilleure prédiction dans toutes les configurations : à la station 310, la méthode 1a reste légèrement plus performante pour les périodes de retour de 10 et 50 ans, avant que la méthode 2 ne devienne la meilleure pour $T = 100$. La position périphérique de la station semble ainsi pouvoir compenser en partie l'avantage structurel de la méthode 2.

La méthode 1b présente les RMSE les plus élevés pour les deux stations et les trois périodes de retour. Ce résultat s'explique principalement par l'absence de dépendance spatiale introduite dans ce

premier cadre de simulation. Après transformation des marges en pseudo-observations uniformes, deux stations indépendantes possèdent théoriquement un F-madogram égal à $1/6$. Les poids bruts utilisés par la méthode 1b, définis par

$$W_{jk} = 1 - 6\hat{d}_{jk},$$

sont donc théoriquement proches de zéro lorsque les stations sont indépendantes. Dans ce contexte, les écarts observés entre les F-madogrammes empiriques proviennent essentiellement de la variabilité d'échantillonnage, et la normalisation de poids faibles et instables peut amplifier ces fluctuations plutôt que fournir une information pertinente pour la prédiction. Les faibles fréquences de sélection des covariables du modèle de dépendance sont cohérentes avec cette situation : aucune structure spatiale stable ne peut être identifiée lorsque le mécanisme générateur n'en contient pas.

À l'inverse, les fréquences de sélection observées pour la méthode 2 montrent que la procédure parvient généralement à retrouver la structure du modèle générateur. L'altitude et la distance à la côte, qui interviennent dans le paramètre de position du modèle simulé, sont nettement plus souvent sélectionnées que la longitude et la latitude. Pour le paramètre d'échelle, qui est constant dans le modèle générateur, les fréquences de sélection restent faibles. Les sélections occasionnelles d'autres covariables peuvent être attribuées à la variabilité d'échantillonnage et aux corrélations entre caractéristiques géographiques.

Ce premier cadre met donc principalement en évidence trois éléments : l'incertitude des prédictions augmente avec la période de retour, la position spatiale de la station peut modifier les performances, et l'adéquation entre le modèle générateur et la méthode d'estimation influence fortement les résultats. Par contre, il ne permet pas d'évaluer la pertinence d'une interpolation fondée sur la dépendance spatiale des extrêmes, puisque cette dépendance est absente du processus générateur. L'analyse du cadre max-stable permettra précisément d'examiner les méthodes lorsque cette composante est explicitement introduite dans les simulations.

6.2 Résultats sous le cadre max-stable

6.2.1 Modèle générateur

Le modèle générateur retenu pour ce second cadre de simulation est un processus max-stable de Brown-Resnick. Il a été obtenu en suivant la méthodologie présentée dans la section 5.3.2. La structure de dépendance a été ajustée aux maxima hivernaux observés après transformation des données de chaque station vers des marges de Fréchet unitaires. Cette transformation permet d'estimer la dépendance spatiale indépendamment des différences marginales entre les stations.

L'ajustement conduit aux paramètres de dépendance suivants :

$$\text{range} = 39.46 \text{ km} \quad \text{et} \quad \text{smooth} = 1.34.$$

Le paramètre de portée, ou *range*, définit l'échelle spatiale caractéristique de la dépendance. La valeur obtenue indique que les extrêmes observés dans des stations séparées par quelques dizaines de kilomètres conservent une dépendance significative, celle-ci diminuant progressivement lorsque la distance augmente. Cette valeur ne constitue pas un seuil strict, la dépendance ne devient pas nulle au-delà de 39.5 km.

Le paramètre de régularité, ou *smooth*, contrôle la forme de la décroissance de la dépendance avec la distance. La valeur de 1.34 correspond à une décroissance relativement régulière : les stations proches présentent des extrêmes fortement associés, puis la dépendance diminue à mesure que leur éloignement augmente.

Dans un premier temps, les jeux de données sont simulés sur des marges de Fréchet unitaires à partir de cette structure de dépendance. Ils sont ensuite transformés vers des marges GEV en utilisant les paramètres marginaux retenus dans le cadre GEV spatial présenté dans le tableau 2 de la section 6.1.1.

Les deux cadres de simulation possèdent ainsi exactement la même structure marginale. Ils se distinguent uniquement par la présence, dans le second cadre, d'une dépendance spatiale explicite entre les extrêmes. Ce choix permet de limiter les différences entre les deux mécanismes générateurs et d'attribuer plus directement les variations de performance des méthodes à l'ajout de cette dépendance, plutôt qu'à une modification des distributions marginales ou de leurs relations avec les covariables.

6.2.2 Présentation des résultats

Le tableau 6 présente les performances obtenues par les trois méthodes sous le cadre de simulation max-stable. Comme dans le cadre GEV spatial, les stations 260 et 310 sont retirées de chaque jeu de

données simulé avant l’ajustement des méthodes. Les niveaux de retour prédits à ces deux stations sont ensuite comparés à leurs valeurs théoriques afin de calculer le RMSE sur l’ensemble des répétitions de simulation. Les résultats sont présentés pour les périodes de retour de 10, 50 et 100 ans.

TABLEAU 6 – RMSE obtenus pour les trois méthodes sous le cadre de simulation max-stable [0.1mm]

Station	Période de retour T (années)	Méthode 1a RMSE	Méthode 1b RMSE	Méthode 2 RMSE
260	10	18.3	18.5	15.3
	50	45.3	45.3	36.4
	100	66.1	66.3	51.3
310	10	17.5	21.7	24.2
	50	45.0	57.4	49.6
	100	65.8	85.0	66.5

Comme dans le cadre GEV spatial, les RMSE augmentent avec la période de retour pour les trois méthodes. Les différences entre méthodes dépendent toutefois fortement de la station considérée.

À la station centrale 260, la méthode 2 fournit les erreurs les plus faibles pour les trois périodes de retour. Son avantage s’accroît lorsque la période de retour augmente, tandis que les méthodes 1a et 1b présentent des performances presque identiques. À la station périphérique 310, la hiérarchie est différente : la méthode 1a obtient systématiquement les RMSE les plus faibles, la méthode 2 présente des résultats intermédiaires et la méthode 1b est la moins performante. Cette dernière est également la méthode dont les performances se dégradent le plus entre les stations 260 et 310.

Les fréquences de sélection des covariables ont également été examinées pour les méthodes 1b et 2. Elles correspondent au pourcentage d’ajustements au cours desquels chaque covariable a été conservée dans le modèle final.

Le tableau 7 présente les fréquences de sélection des covariables utilisées pour modéliser la dépendance spatiale dans la méthode 1b.

TABLEAU 7 – Fréquence de sélection des covariables du modèle de dépendance de la méthode 1b sous le cadre max-stable

Covariable	Fréquence de sélection
Distance géographique h	99 %
Différence d’altitude Δ_{alt}	33 %
$\cos(2\phi)$	27 %
$\sin(2\phi)$	27 %

La structure sélectionnée pour modéliser la dépendance par la méthode 1b est beaucoup plus stable que dans le cadre GEV spatial. La distance géographique h est retenue dans 99 % des ajustements, contre 33 % pour la différence d’altitude et 27 % pour chacune des deux composantes décrivant l’orientation des paires de stations. La distance constitue donc clairement le principal prédicteur de la dépendance dans ce cadre.

Pour la méthode 2, les fréquences de sélection sont calculées séparément pour les paramètres de position μ et d’échelle σ du modèle GEV spatial. Les résultats sont présentés dans le tableau 8.

TABLEAU 8 – Fréquence de sélection des covariables du modèle GEV spatial de la méthode 2 sous le cadre max-stable

Covariable	Position μ	Échelle σ
Longitude lon	30 %	10 %
Latitude lat	26 %	7 %
Altitude alt	88 %	7 %
Distance à la côte d_{coast}	62 %	6 %

Pour la méthode 2, l’altitude et la distance à la côte restent les covariables les plus fréquemment sélectionnées.

tionnées pour le paramètre de position, dans respectivement 88 % et 62 % des ajustements. La longitude et la latitude sont moins souvent retenues, avec des fréquences respectives de 30 % et 26 %. Comme dans le premier cadre, les fréquences de sélection associées au paramètre d'échelle restent faibles, comprises entre 6 % et 10 %.

6.2.3 Discussion des résultats

L'introduction d'une dépendance spatiale modifie surtout les performances relatives des méthodes selon la localisation de la station. À la station 260, la méthode 2 conserve les meilleures performances, tandis que les méthodes 1a et 1b fournissent des RMSE presque identiques. La présence d'une dépendance spatiale ne suffit donc pas, dans cette configuration, à donner un avantage à la méthode 1b par rapport à une interpolation fondée uniquement sur la distance. La structure de dépendance simulée par le modèle de Brown–Resnick est relativement simple et principalement gouvernée par la proximité géographique. Une méthode telle que l>IDW peut dès lors exploiter indirectement une grande partie de cette information en accordant davantage de poids aux stations voisines.

La situation est différente à la station 310, où la méthode 1b se dégrade nettement. Sa position périphérique implique que les stations disponibles se trouvent principalement d'un même côté du site à prédire et à des distances généralement plus importantes. De plus, le modèle de dépendance est estimé à partir d'une structure globale entre les stations d'apprentissage, qui ne garantit pas que les relations estimées soient également informatives localement pour une station située en bordure du domaine. La prise en compte explicite de la dépendance n'entraîne donc pas nécessairement une meilleure prédiction lorsque la configuration spatiale fournit peu d'information autour du site considéré, qui plus est si la dépendance générée est uniquement fonction de la distance géographique. La plus grande simplicité la méthode 1a, réduisant sa variance, peut expliquer sa meilleure performance à la station périphérique. Cela ne signifie pas que la méthode 1b serait systématiquement désavantageuse, mais que son gain potentiel est limité lorsque la dépendance peut déjà être convenablement approchée par une simple fonction de la distance.

Les fréquences de sélection du modèle de dépendance confirment cependant que la méthode 1b identifie correctement la principale structure introduite par le générateur. La sélection quasi systématique de la distance est cohérente avec le caractère isotrope du modèle de Brown–Resnick utilisé, dans lequel la dépendance est principalement déterminée par l'éloignement entre les stations. Les sélections plus occasionnelles de la différence d'altitude et des variables d'orientation peuvent provenir de la variabilité d'échantillonnage ou de relations existant entre ces caractéristiques et la configuration géographique des stations.

La méthode 2 retrouve également une structure marginale cohérente avec celle utilisée pour générer les données, puisque l'altitude et la distance à la côte dominent à nouveau la sélection du paramètre de position et que peu de covariables sont retenues pour l'échelle. Les fréquences de sélection de l'altitude et de la distance à la côte sont néanmoins légèrement plus faibles que dans le cadre GEV spatial. La dépendance introduite entre les maxima augmente la complexité des données simulées et peut rendre l'identification de la structure marginale moins stable, sans pour autant modifier le mécanisme marginal utilisé pour leur génération.

Ce second cadre montre ainsi que l'introduction d'une dépendance spatiale réaliste ne conduit pas à une amélioration systématique des performances de la méthode 1b, qui la modélise explicitement. Son intérêt dépend également de la forme de cette dépendance et de la configuration des stations disponibles autour du site à prédire. La comparaison directe avec le cadre GEV spatial permettra maintenant de différencier les conclusions qui restent stables entre les deux simulations de celles qui dépendent du mécanisme générateur.

6.3 Comparaison des deux cadres de simulation

La comparaison des deux cadres permet de distinguer les comportements robustes des méthodes de ceux directement liés au processus générateur. Le premier constat commun concerne l'augmentation des erreurs avec la période de retour, observée pour toutes les méthodes et dans les deux cadres. Les RMSE sont toutefois généralement plus élevés sous le cadre max-stable, ce qui traduit la difficulté supplémentaire de prédiction, introduite par la dépendance entre stations. Les performances dépendent également de la position du site à prédire, les différences entre méthodes étant généralement plus marquées à la station périphérique 310 qu'à la station centrale 260.

La méthode 1a présente le comportement le plus stable entre les deux cadres. Ses performances se dégradent lorsque la dépendance est introduite, comme pour les autres méthodes, mais sa position relative varie peu : elle reste compétitive à la station 260 et devient la meilleure méthode à la station 310 sous le cadre max-stable. Ce résultat montre qu’une interpolation fondée uniquement sur la distance géographique peut rester efficace lorsque la dépendance spatiale des extrêmes est elle-même principalement structurée par la distance. Dans le modèle de Brown–Resnick utilisé, les stations proches présentent en effet une dépendance plus forte que les stations éloignées, ce qui permet à la distance exploitée par l’IDW d’extraire pratiquement toute l’information de dépendance ajoutée dans le second cadre de simulation.

L’évolution de la méthode 1b est plus directement liée au changement de cadre. Sous le modèle GEV spatial, l’absence de dépendance entre stations rendait les poids dérivés du F-madogramme peu informatifs. L’introduction d’une dépendance max-stable permet au contraire à la méthode d’identifier une structure beaucoup plus stable, illustrée par la fréquence de sélection de la distance de 99 % lors de l’ajustement du modèle dans le second cadre. Cette amélioration de l’identification de la dépendance ne se traduit cependant pas par une amélioration satisfaisante de la prédiction. À la station 260, ses performances deviennent pratiquement identiques à celles de la méthode 1a, tandis qu’elles restent nettement moins bonnes à la station 310.

La méthode 2 conserve quant à elle un avantage clair à la station centrale 260 dans les deux cadres, particulièrement pour les périodes de retour élevées. La parcimonie du modèle peut contribuer à cette stabilité. Contrairement aux méthodes 1a et 1b, qui estiment un paramètre de forme ξ propre à chaque station, la méthode 2 impose un ξ commun à l’ensemble du domaine. Avec seulement 41 maxima par station, l’estimation séparée de ce paramètre peut être incertaine et se répercuter sur les niveaux de retour, surtout pour les périodes de retour élevées. L’hypothèse d’un comportement de queue relativement homogène sur le domaine néerlandais étant raisonnable, cette contrainte peut constituer un avantage pour la méthode 2 en réduisant la variance d’estimation. Néanmoins, l’explication la plus probable de cette stabilité provient de la construction des simulations : les marges GEV du cadre max-stable sont les mêmes que celles du cadre GEV spatial et restent donc cohérentes avec la structure que la méthode 2 cherche à ajuster. L’introduction d’une dépendance entre stations augmente la complexité des données, mais ne supprime pas cet avantage marginal. Les fréquences de sélection des covariables restent d’ailleurs relativement proches entre les deux cadres, avec une prédominance de l’altitude et de la distance à la côte pour le paramètre de position et peu de covariables sélectionnées pour le paramètre d’échelle. Les performances de la méthode 2 sont néanmoins plus sensibles à la localisation de la station, son avantage étant nettement réduit, voire nul, à la station 310.

Les deux cadres de simulations montrent ainsi qu’aucune méthode ne domine indépendamment du mécanisme générateur et de la configuration spatiale. La méthode 1a semble robuste malgré sa simplicité, la méthode 1b parvient à identifier la dépendance lorsqu’elle est effectivement présente mais n’en tire pas nécessairement un gain prédictif, tandis que la méthode 2 bénéficie fortement du modèle générateur, tout en pouvant être davantage affectée lorsque la station à prédire occupe une position périphérique.

Ces conclusions doivent toutefois rester nuancées par rapport aux cadres de simulation considérés. Le cadre GEV spatial favorise structurellement la méthode 2 et ne permet pas d’évaluer pleinement une méthode fondée sur la dépendance. À l’inverse, le modèle max-stable introduit une dépendance relativement simple, isotrope et principalement gouvernée par la distance, ce qui limite l’avantage potentiel d’une méthode telle que la méthode 1b si la structure de dépendance était plus complexe. Les deux générateurs reproduisent donc certaines caractéristiques des données réelles sans pouvoir en représenter toute la complexité, notamment d’éventuelles structures de dépendance plus locales ou anisotropes.

L’application aux données réelles devient donc intéressante pour compléter cette comparaison, dans un contexte où aucune méthode ne correspond exactement à un mécanisme générateur connu. La validation par retrait successif d’une station permettra d’évaluer les performances sur l’ensemble du réseau, plutôt que sur deux localisations seulement, et d’examiner plus précisément la variabilité spatiale des erreurs. Elle permettra également d’étudier si les covariables du modèle de dépendance autres que la distance, notamment la différence d’altitude et les composantes d’orientation, apportent une information stable lorsque la dépendance est potentiellement plus complexe que celle simulée. De même, la fréquence de sélection des covariables du modèle GEV spatial indiquera si les relations marginales identifiées dans les simulations restent présentes lorsque le modèle n’est plus évalué sur base de données produites selon certaines hypothèses.

7 Application aux données réelles

Après avoir comparé les performances des trois méthodes dans les deux cadres de simulation, leur application aux données réelles permet d'évaluer dans quelle mesure les conclusions précédentes restent valables lorsque le mécanisme générateur des observations est inconnu. Cette analyse vise à confronter les comportements mis en évidence lors des simulations à une situation plus proche de la réalité. L'évaluation repose sur la procédure de validation *leave-one-station-out* décrite dans la section 5.7.2 et porte sur les périodes de retour de 10, 50 et 100 ans. Les performances prédictives des méthodes sont d'abord comparées globalement à l'aide du RMSE et du MAE, avant d'examiner plus précisément la répartition spatiale des erreurs et les structures sélectionnées par les méthodes 1b et 2.

7.1 Performances prédictives globales

Les performances prédictives globales des trois méthodes, obtenues à partir de la procédure de validation *leave-one-station-out*, sont présentées dans le tableau 9. Le RMSE et le MAE sont calculés à partir des erreurs obtenues successivement aux différentes stations retirées, pour chacune des trois périodes de retour considérées.

TABLEAU 9 – Performances prédictives globales des trois méthodes sur les données réelles selon la période de retour [0.1mm]

Méthode	$T = 10$		$T = 50$		$T = 100$	
	RMSE	MAE	RMSE	MAE	RMSE	MAE
1a	31.1	26.8	83.9	69.7	124.1	102.7
1b	34.4	28.3	91.2	76.2	133.9	111.2
2	31.9	27.4	81.0	68.8	119.3	101.0

Comme dans les simulations, les erreurs augmentent fortement avec la période de retour pour les trois méthodes. Les RMSE, compris entre 31.1 et 34.4 pour $T = 10$, atteignent entre 119.3 et 133.9 pour $T = 100$. La même évolution apparaît pour le MAE, dont les valeurs suivent des ordres de grandeurs identiques. Cette augmentation est attendue puisque les niveaux de retour associés aux périodes les plus longues prennent des valeurs plus élevées et reposent sur une extrapolation de plus en plus éloignée dans la queue des distributions.

Les deux critères d'évaluation conduisent par ailleurs au même classement des méthodes pour chacune des périodes de retour. Les RMSE restent systématiquement supérieurs aux MAE, ce qui indique que certaines stations présentent des erreurs suffisamment importantes pour être davantage pénalisées par le RMSE. Toutefois, ces erreurs ne modifient pas la hiérarchie globale obtenue avec le MAE. Leur origine et leur répartition spatiale seront examinées plus précisément dans la section suivante.

Concernant la comparaison des méthodes, les résultats diffèrent selon la période de retour considérée. Pour $T = 10$, la méthode 1a obtient les meilleures performances, avec un RMSE de 31.1 et un MAE de 26.8. La méthode 2 est toutefois très proche, avec respectivement 31.9 et 27.4. Pour une période de retour relativement courte, l'interpolation directe des niveaux de retour à partir de la proximité géographique apparaît donc au moins aussi efficace que la modélisation spatiale des paramètres de la GEV. L'écart entre les deux méthodes étant faible, ces résultats ne permettent cependant pas d'établir une supériorité nette de la méthode 1a.

Lorsque la période de retour augmente, le classement entre ces deux méthodes s'inverse. Pour $T = 50$, la méthode 2 présente un RMSE de 81.0, contre 83.9 pour la méthode 1a, ainsi qu'un MAE de 68.8 contre 69.7. La même tendance est observée pour $T = 100$, avec un RMSE de 119.3 pour la méthode 2 contre 124.1 pour la méthode 1a, et un MAE de 101.0 contre 102.7. L'avantage de la méthode 2 reste modéré, mais il apparaît donc de manière cohérente pour les deux périodes de retour les plus élevées.

La méthode 1b présente quant à elle les valeurs de RMSE et de MAE les plus élevées pour chacune des trois périodes de retour. Contrairement à ce qui pouvait être envisagé lors de la construction de la méthode, l'utilisation d'une pondération fondée sur la dépendance extrême n'améliore donc pas les performances globales par rapport à l'interpolation fondée uniquement sur la distance géographique. Ces résultats ne signifient toutefois pas que la méthode 1b soit dépourvue d'intérêt ou qu'elle ne puisse pas présenter d'avantage dans d'autres configurations. Ses performances dépendent directement de la manière

dont la dépendance extrême est caractérisée, mais également des covariables utilisées pour expliquer cette dépendance entre les stations. Les variables considérées dans ce mémoire, telles que la distance géographique, l'écart d'altitude et les termes directionnels, ne représentent qu'un ensemble possible de facteurs susceptibles d'influencer la structure de dépendance. L'analyse des modèles de dépendance sélectionnés lors des différentes itérations du *leave-one-station-out* apportera également des éléments sur la structure de dépendance identifiée par la méthode et sur sa stabilité lorsque la station retirée varie. Elle permettra ainsi d'évaluer plus précisément dans quelle mesure le modèle utilisé est capable de représenter de manière cohérente la dépendance présente entre les stations, et d'apporter des éléments de réponse quant à la capacité de la méthode 1b à estimer correctement cette structure dans le cadre considéré.

7.2 Répartition spatiale des erreurs

La figure 9 présente les erreurs relatives obtenues à chacune des stations lors de la procédure *leave-one-station-out*. Elle permet de compléter les critères globaux précédents en identifiant les stations qui contribuent le plus aux erreurs et en examinant leur éventuelle organisation spatiale.

Une première caractéristique importante est la forte similitude de la structure des erreurs entre les méthodes et entre les périodes de retour. Les mêmes stations tendent généralement à être surestimées ou sous-estimées, et le signe de l'erreur reste le plus souvent inchangé lorsque T augmente. L'amplitude des erreurs relatives tend également à augmenter avec la période de retour, ce qui indique que l'augmentation du RMSE et du MAE observée précédemment n'est pas uniquement due à l'augmentation des périodes de retour. Cette évolution n'est toutefois pas systématique : certaines stations, notamment 260, 280 et 290, conservent des erreurs relativement faibles ou ne présentent pas d'augmentation nette avec T .

Certaines stations se distinguent particulièrement. La station 240 est systématiquement surestimée par les trois méthodes, et cette surestimation devient importante pour les périodes de retour élevées, surtout avec la méthode 1b. La station 270 présente un comportement similaire, mais de manière moins marquée. À l'inverse, les stations 235 et 380 sont régulièrement sous-estimées. Ces stations contribuent donc fortement aux valeurs du RMSE et du MAE. De plus, la présence de plusieurs stations correctement prédites montre que ces valeurs globales ne traduisent pas une mauvaise capacité de prédiction sur l'ensemble du réseau, mais sont obtenues à partir d'un nombre limité de stations plus difficiles à estimer. Ces erreurs ne sont cependant pas suffisamment différentes entre les méthodes pour modifier leur classement lorsque le RMSE est remplacé par le MAE.

La localisation de ces stations apporte également des informations supplémentaires. Les erreurs les plus importantes concernent presque toutes des stations situées à la périphérie du domaine, comme 235, 240, 270 et 380, tandis que plusieurs stations centrales, notamment 260, 350 et 370, sont relativement bien prédites. Une station située en bordure dispose de moins d'informations dans son voisinage et celles-ci proviennent essentiellement d'une seule direction, ce qui peut rendre son interpolation plus difficile. Cet effet de bord ne suffit cependant pas à expliquer entièrement les erreurs observées. Des stations spatialement proches peuvent en effet présenter des erreurs importantes de signes opposés. La distance géographique seule n'est donc pas suffisante pour représenter les variations locales des niveaux de retour dans la méthode 1a, tandis que les covariables retenues dans les méthodes 1b et 2 pourraient ne pas capturer certains facteurs locaux influençant les précipitations extrêmes.

D'un autre côté, la comparaison entre méthodes confirme les résultats globaux de performance présentés dans la section 7.1. La méthode 1b amplifie généralement les erreurs aux stations déjà mal prédites par les deux autres méthodes, ce qui explique ses valeurs de RMSE et de MAE systématiquement plus élevées. Les différences entre les méthodes 1a et 2 sont en revanche beaucoup moins visibles. La méthode 2 ne modifie pas fondamentalement l'organisation spatiale des erreurs par rapport à la méthode 1a ; son avantage pour $T = 50$ et $T = 100$ semble plutôt provenir d'une réduction modérée de l'amplitude des erreurs aux stations les plus difficiles à prédire. Sa meilleure performance globale ne résulte donc pas d'une amélioration uniforme sur l'ensemble du domaine.

Enfin, le fait que les trois méthodes présentent souvent des erreurs importantes de même signe aux mêmes stations peut également refléter une difficulté liée aux niveaux de retour utilisés comme références. Avant d'approfondir l'analyse des structures spatiales retenues par les méthodes 1b et 2, il convient donc d'examiner l'incertitude associée aux estimations marginales de $\hat{z}_{T,\text{ref}}$, afin d'évaluer dans quelle mesure celle-ci peut contribuer aux erreurs observées, en particulier pour les périodes de retour les plus élevées. L'analyse des structures sélectionnées permettra ensuite de compléter cette discussion en déterminant si les performances des méthodes 1b et 2 peuvent également être liées à une difficulté à identifier une structure spatiale stable lors des différentes itérations du *leave-one-station-out*, ou plutôt aux limites des covariables et des structures de modèles considérées.

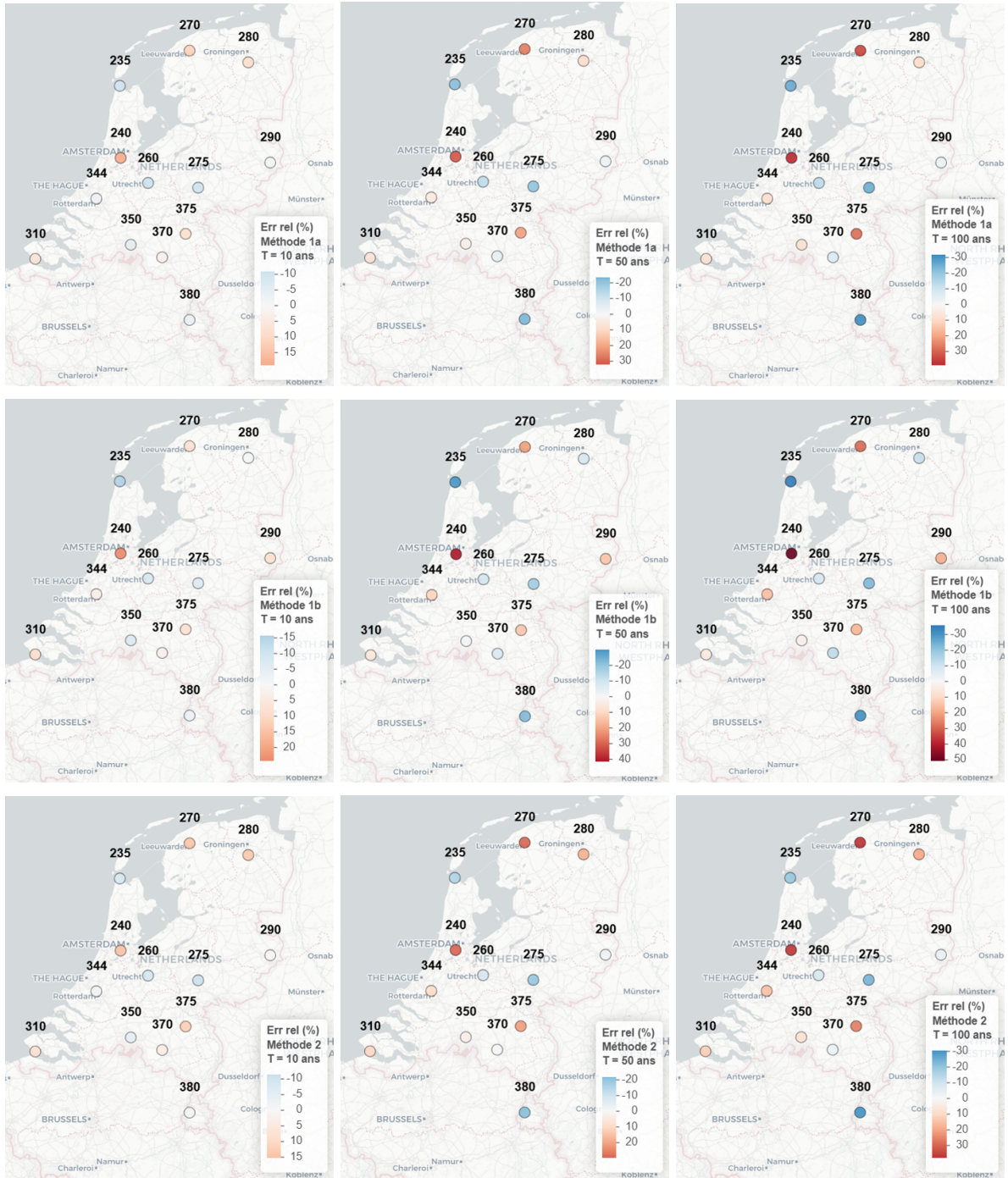


FIGURE 9 – Répartition spatiale des erreurs relatives des trois méthodes pour les périodes de retour de 10, 50 et 100 ans. Les lignes correspondent respectivement aux méthodes 1a, 1b et 2, et les colonnes aux périodes de retour $T = 10$, $T = 50$ et $T = 100$ ans.

7.3 Incertitude sur les niveaux de retour de référence

L'analyse précédente a montré que certaines stations concentrent une part importante des erreurs de prédiction et que les trois méthodes rencontrent souvent des difficultés de prédictions à ces mêmes stations. Une partie de ces erreurs pourrait provenir de l'incertitude associée aux niveaux de retour marginaux utilisés comme références. Afin d'examiner cette hypothèse, le tableau 10 présente, pour les six stations identifiées comme les plus difficiles à prédire, la largeur relative de l'intervalle de confiance à 95 % de $\hat{z}_{T,\text{ref}}$. Ces valeurs sont comparées à la médiane obtenue pour les sept autres stations. Le tableau complet contenant les niveaux de retour de référence, leurs erreurs standards et les bornes des intervalles de confiance est présenté en annexe B.

TABLEAU 10 – Largeur relative des intervalles de confiance à 95 % des niveaux de retour de référence pour les stations présentant les erreurs de prédiction les plus importantes

Station	Largeur relative de l'IC _{95%} (%)		
	$T = 10$	$T = 50$	$T = 100$
235	37.2	73.1	92.7
240	21.7	46.7	60.7
270	20.9	40.8	52.1
275	37.5	80.2	103.9
375	22.9	38.0	46.3
380	50.2	121.1	160.1
Médiane des 7 autres stations	28.3	59.7	76.2

Les résultats ne mettent pas en évidence de relation systématique entre l'incertitude du niveau de retour de référence et la difficulté de prédiction d'une station. En particulier, les trois stations les plus sous-estimées (235, 275 et 380) présentent des intervalles de confiance relativement larges, systématiquement supérieurs à la médiane des autres stations pour les trois périodes de retour. Le cas de la station 380 est particulièrement marqué : la largeur de son intervalle représente 50.2 % de son niveau de retour estimé pour $T = 10$, puis 121.1 % pour $T = 50$ et 160.1 % pour $T = 100$. Pour ces stations, et dans une moindre mesure pour les stations 235 et 275, une partie des écarts observés précédemment peut donc être liée au fait que la valeur utilisée comme référence est elle-même estimée avec une incertitude importante.

À l'inverse, les trois stations les plus surestimées (240, 270 et 375) présentent des largeurs relatives d'intervalles de confiance systématiquement inférieures à la médiane des autres stations. Les fortes erreurs observées à ces stations ne peuvent donc pas être attribuées principalement à une mauvaise précision de leur niveau de retour marginal de référence. Ce résultat est particulièrement important pour les stations 240 et 270, qui figuraient parmi les stations les plus nettement surestimées par les différentes méthodes malgré une incertitude marginale relativement faible. Leurs difficultés de prédiction semblent donc davantage liées à leur représentation par les modèles spatiaux : position en périphérie du domaine, comportement différent des stations voisines, structure spatiale trop simple ou encore absence de covariables permettant de caractériser certaines particularités locales.

L'incertitude des niveaux de retour de référence constitue donc une explication plausible pour certaines stations, mais elle ne permet pas d'expliquer à elle seule la concentration des erreurs observée précédemment. Le fait que les stations sous-estimées soient associées à des incertitudes élevées, tandis que les stations surestimées présentent au contraire des incertitudes plus faibles que la médiane, suggère que d'autres mécanismes interviennent dans les erreurs de prédiction.

L'étape suivante consiste donc à examiner les structures spatiales retenues par les méthodes 1b et 2 au cours des différentes itérations du *leave-one-station-out*. L'analyse de leur stabilité et des covariables sélectionnées permettra de déterminer si une partie des difficultés observées peut être associée à une estimation instable ou insuffisante de la structure spatiale, ou si elles semblent plutôt provenir de particularités locales que les modèles considérés ne parviennent pas à représenter.

7.4 Structures spatiales sélectionnées

Après avoir évalué les performances prédictives des différentes méthodes et examiné dans quelle mesure l'incertitude des niveaux de retour de référence pouvait expliquer les erreurs observées, il reste à analyser les structures spatiales identifiées par les modèles. Cette étape permet de mieux comprendre

les mécanismes sur lesquels reposent les prédictions des méthodes 1b et 2, mais également d'évaluer la stabilité des structures sélectionnées lorsque la station retirée varie au cours de la procédure *leave-one-station-out*. Les modèles de dépendance retenus par la méthode 1b sont d'abord étudiés, avant d'examiner les covariables sélectionnées pour les paramètres de la GEV spatiale de la méthode 2.

7.4.1 Modèle de dépendance de la méthode 1b

L'analyse des modèles sélectionnés lors des différentes itérations du *leave-one-station-out* permet d'examiner la structure de dépendance identifiée par la méthode 1b et sa stabilité lorsque la station retirée varie. Le tableau 11 présente la fréquence de sélection de chacune des quatre covariables candidates ainsi que l'étendue des coefficients estimés. Le détail des modèles et des coefficients obtenus pour chacun des treize folds est présenté en annexe C (tableau 15).

TABLEAU 11 – Fréquence de sélection des covariables et étendue des coefficients estimés pour les modèles de dépendance de la méthode 1b

Covariable	Fréquence de sélection	Étendue du coefficient
Distance h	13/13	[0.066 ; 0.098]
Écart d'altitude Δalt	0/13	–
$\cos(2\phi)$	12/13	[2.030 ; 3.790]
$\sin(2\phi)$	0/13	–

Les modèles sélectionnés apparaissent particulièrement stables entre les différents folds. La distance h est retenue dans les treize ajustements et le terme directionnel $\cos(2\phi)$ dans douze d'entre eux. Seul le fold correspondant au retrait de la station 235 conduit au modèle plus simple $\psi \sim h$. À l'inverse, l'écart d'altitude Δalt et le terme $\sin(2\phi)$ ne sont jamais retenus. Malgré le retrait successif d'une station, la sélection converge donc presque systématiquement vers la même structure,

$$\psi_{jk} = \beta_0 + \beta_h h_{jk} + \beta_{\cos} \cos(2\phi_{jk}),$$

Cette stabilité doit néanmoins être interprétée en tenant compte du fort recouvrement entre les folds, puisque deux ensembles d'apprentissage successifs ne diffèrent que par une station.

L'interprétation des coefficients doit tenir compte de la transformation appliquée au F-madogramme,

$$\psi_{jk} = \log \left(\frac{6d_{jk}}{1 - 6d_{jk}} \right).$$

La transformation étant strictement croissante, une augmentation de ψ_{jk} correspond à une augmentation du F-madogramme d_{jk} et donc à une diminution de la dépendance extrême entre les deux stations. Inversement, une valeur plus faible de ψ_{jk} correspond à une dépendance plus forte qui conduit à un poids d'interpolation plus important et donc à une dépendance plus forte entre les deux stations.

Le coefficient associé à la distance est positif dans les treize folds et varie seulement entre 0.066 et 0.098. Une augmentation de la distance entre deux stations entraîne donc une augmentation de ψ , ce qui implique une diminution de leur dépendance estimée. Ce résultat correspond au comportement attendu d'une dépendance spatiale : les événements extrêmes observés à deux stations proches tendent à être davantage liés que ceux observés à deux stations éloignées. L'étendue relativement limitée des coefficients montre également que l'intensité de cet effet est assez stable entre les folds.

De l'autre côté, la présence presque systématique du terme $\cos(2\phi)$ signifie que la dépendance estimée n'est pas isotrope et varie selon l'orientation séparant deux stations. L'interprétation de ce coefficient peut être détaillée en plusieurs étapes. L'angle ϕ correspond tout d'abord au *bearing* entre les deux stations, avec un angle nul orienté vers le nord. La transformation $\cos(2\phi)$ permet ensuite de considérer des axes plutôt que des directions : une orientation nord-sud possède la même valeur sans tenir compte du sens dans lequel la paire de stations est parcourue.

Plus précisément,

$$\cos(2\phi) = 1$$

pour un axe nord-sud, correspondant à $\phi = 0$ ou $\phi = \pi$, tandis que

$$\cos(2\phi) = -1$$

pour un axe est-ouest, correspondant à $\phi = \pi/2$ ou $\phi = 3\pi/2$. Les orientations diagonales donnent des valeurs intermédiaires, et notamment $\cos(2\phi) = 0$ pour des angles de 45° , 135° , 225° ou 315° . Dans le cas où de tels angles entre stations montreraient des propriétés de dépendance anisotropes, le modèle retiendrait la covariable $\sin(2\phi)$ qui vaut 1 et -1 pour ces angles.

Dans notre cas, seule la covariable $\cos(2\phi)$ est retenue. Le coefficient $\beta_{\cos}$ est positif dans les douze modèles où il est sélectionné, avec des valeurs comprises entre 2.030 et 3.790. À distance identique, une paire orientée selon l'axe nord-sud reçoit donc une contribution de $+\beta_{\cos}$ à ψ , tandis qu'une paire orientée selon l'axe est-ouest reçoit une contribution de $-\beta_{\cos}$. L'écart entre ces deux orientations est ainsi égal à

$$\psi_{N-S} - \psi_{E-O} = 2\beta_{\cos},$$

soit entre 4.060 et 7.580 selon le fold. Une valeur plus faible de ψ correspondant à une dépendance plus forte, le modèle identifie donc une dépendance extrême généralement plus importante suivant l'axe est-ouest que suivant l'axe nord-sud.

Dans le cas où le terme $\sin(2\phi)$ avait été présent, il aurait permis d'orienter l'axe d'anisotropie entre 0° et 90° . Sa non-sélection suggère donc que la composante directionnelle identifiée par le modèle est organisée suivant les axes nord-sud et est-ouest. De même, l'absence de Δalt signifie que, conditionnellement à la distance et à la direction, la différence d'altitude entre deux stations n'apporte pas suffisamment d'information supplémentaire pour être conservée par la procédure de sélection.

La dépendance plus forte suivant l'axe est-ouest pourrait être liée à l'influence de la mer du Nord et aux circulations atmosphériques dominantes dans la région. Manola et al. [91] montrent notamment que les précipitations à Amsterdam sont influencées par des flux d'ouest transportant de l'humidité depuis la mer du Nord vers le continent. De tels déplacements peuvent favoriser le passage d'un même événement de précipitation sur plusieurs stations alignées selon l'axe est-ouest, ce qui pourrait expliquer la dépendance plus importante observée dans cette direction. Cette interprétation doit toutefois rester prudente, car les sources sont peu nombreuses et le pays étudié est relativement petit. De plus, le modèle ne contient aucune variable décrivant directement les vents ou les trajectoires des systèmes météorologiques. L'anisotropie observée constitue donc avant tout une caractéristique empirique des données.

Ces résultats montrent ainsi que les performances plus faibles de la méthode 1b ne semblent pas pouvoir être attribuées à une forte instabilité de la procédure de sélection : la distance, l'effet directionnel et les signes de leurs coefficients sont très reproductibles entre les folds. La structure retenue reste néanmoins une représentation globale de la dépendance sur l'ensemble du domaine et peut ne pas capturer certaines variations locales ou des relations plus complexes que celles décrites par les covariables considérées.

7.4.2 Structure du modèle de GEV spatiale de la méthode 2

L'analyse de la méthode 2 vise de la même manière à déterminer quelles covariables expliquent les variations spatiales des paramètres de la GEV et dans quelle mesure cette structure reste stable lorsque la station retirée varie. Le tableau 12 présente les fréquences de sélection des différentes covariables pour les paramètres de position μ et d'échelle σ . Le paramètre de forme ξ n'est pas présenté, puisque sa structure est fixée constante dans la définition de la méthode 2. Les structures complètes sélectionnées pour chacun des treize folds sont détaillées en annexe C (tableau 16).

TABLEAU 12 – Fréquence de sélection des covariables dans les paramètres de position et d'échelle de la GEV spatiale

Covariable	Position μ	Échelle σ
Latitude	1/13	0/13
Longitude	2/13	0/13
Altitude	12/13	0/13
Distance à la côte	13/13	0/13

Les résultats montrent une structure particulièrement stable pour le paramètre de position. La distance à la côte est sélectionnée dans les treize folds et l'altitude dans douze d'entre eux. À l'inverse, la longitude n'est retenue que dans deux ajustements et la latitude dans un seul. La structure dominante est donc

$$\mu(s) \sim alt + d_{\text{coast}},$$

qui est sélectionnée directement dans dix des treize folds. Les quelques différences apparaissent lorsque les stations 310, 375 et 380 sont retirées : la longitude est ajoutée lors du retrait de la station 310, la

latitude et la longitude lors du retrait de la station 375, tandis que seule la distance à la côte est conservée lorsque la station 380 est retirée.

Cette structure est cohérente avec les résultats obtenus lors de l’ajustement réalisé sur l’ensemble des treize stations. Elle suggère que les différences spatiales de maxima de précipitations sont principalement associées, parmi les covariables considérées, à l’altitude et à la proximité de la côte. La sélection systématique de la distance à la côte suggère une influence maritime sur la répartition spatiale des précipitations extrêmes. L’altitude apparaît également comme une information importante, puisqu’elle n’est absente que dans un seul fold.

La structure du paramètre d’échelle est encore plus stable : aucune covariable n’est sélectionnée dans les treize folds et σ reste donc constant dans l’espace. Dans le cadre de la structure retenue par la méthode 2, les covariables disponibles permettent ainsi principalement de représenter des variations spatiales du paramètre de position, mais pas de mettre en évidence une variation suffisamment nette de la dispersion des maxima extrêmes entre les stations.

Le paramètre de forme ξ n’est pas présenté puisque sa structure est maintenue constante dans la définition de la méthode. Cette contrainte distingue la méthode 2 des méthodes 1a et 1b, pour lesquelles ξ est estimé séparément à chaque station. Elle rend le modèle plus parcimonieux et limite notamment la variabilité associée à l’estimation d’un paramètre de forme distinct à partir d’un nombre relativement faible de maxima.

Comme pour la méthode 1b, cette stabilité ne signifie toutefois pas que la structure sélectionnée représente l’ensemble des variations spatiales présentes dans les données. Certaines stations, notamment en périphérie du domaine, restent difficiles à prédire alors même qu’une structure identique est retrouvée dans la majorité des folds. Des covariables importantes peuvent être absentes du modèle ou leur effet peut être plus local ou plus complexe que la relation globale considérée ici.

Dans l’ensemble, les méthodes 1b et 2 identifient donc des structures spatiales remarquablement stables : une dépendance gouvernée principalement par la distance et l’orientation pour la méthode 1b, et un paramètre de position lié principalement à l’altitude et à la distance à la côte pour la méthode 2. Cette stabilité rend peu probable que les erreurs importantes observées à certaines stations proviennent simplement d’une instabilité de la sélection des modèles. Elle doit cependant être nuancée par la forte similarité des ensembles d’apprentissage du *leave-one-station-out*. Une validation retirant simultanément plusieurs stations imposerait des perturbations plus importantes au réseau et permettrait de tester plus sévèrement la robustesse de ces structures. Les résultats obtenus indiquent donc que les modèles identifient des relations globales cohérentes, sans garantir qu’elles suffisent à représenter toutes les particularités locales du domaine.

7.5 Mise en perspective avec les résultats des simulations

Les résultats obtenus sur les données réelles permettent de nuancer les conclusions issues des deux cadres de simulation. Dans la section 6, les résultats des simulations avaient montré que les performances des méthodes dépendaient en partie du mécanisme utilisé pour générer les données. Le cadre GEV spatial favorisait naturellement la méthode 2, dont la structure était proche de celle du modèle générateur, tandis que l’absence de dépendance entre les stations pénalisait fortement la méthode 1b. L’introduction d’une dépendance max-stable avait permis de réduire ce dernier désavantage et de tester les méthodes dans un cadre différent, sans toutefois reproduire toute la complexité potentielle d’une situation réelle. L’application aux observations réelles permet ainsi d’évaluer les trois approches dans une situation où aucun mécanisme générateur n’est connu et où aucune méthode ne bénéficie directement de la structure ayant produit les données.

Un premier résultat commun aux simulations et aux données réelles concerne l’effet de la période de retour. Dans les deux cas, les erreurs absolues augmentent nettement lorsque T augmente. Sur les données réelles, l’analyse des erreurs relatives montre également que ces dernières tendent à augmenter pour plusieurs stations lorsque la période de retour devient plus élevée. L’augmentation des RMSE et des MAE ne résulte donc pas uniquement de l’augmentation des valeurs des niveaux de retour : l’extrapolation plus éloignée dans la queue de la distribution rend également la prédiction plus difficile. Cette observation renforce l’intérêt d’avoir considéré plusieurs périodes de retour, puisque le classement des méthodes n’est pas totalement indépendant de l’horizon d’extrapolation.

La comparaison des méthodes 1a et 2 confirme ensuite que l’avantage observé pour la méthode 2 dans le cadre GEV spatial devait être interprété avec prudence. Sur les données réelles, les deux approches présentent des performances globales proches : la méthode 1a est légèrement meilleure pour $T = 10$, tandis que la méthode 2 obtient les erreurs les plus faibles pour $T = 50$ et $T = 100$, avec toutefois

des écarts modérés. La méthode 2 semble donc tirer profit de la modélisation de la structure spatiale des paramètres marginaux lorsque l'extrapolation devient plus importante, sans dominer nettement une interpolation beaucoup plus simple.

Les résultats confirment ainsi la robustesse de la méthode 1a déjà observée dans les simulations. Malgré l'utilisation de la seule distance géographique, elle reste compétitive sur l'ensemble des périodes de retour. Cette simplicité constitue un avantage lorsque la proximité spatiale résume suffisamment bien l'information disponible. Les cartes d'erreurs ont cependant montré les limites de cette hypothèse : certaines stations spatialement proches présentent des comportements différents, tandis que les stations situées en périphérie du réseau sont généralement plus difficiles à prédire. La distance ne peut donc pas représenter à elle seule toutes les particularités locales des précipitations extrêmes.

La méthode 1b apporte un résultat différent. Contrairement au premier cadre de simulation, les données réelles présentent une structure de dépendance clairement identifiable, dominée par la distance et par une composante directionnelle, et cette structure est très stable entre les différents folds. Pourtant, cette information supplémentaire ne conduit pas à de meilleures performances prédictives globales. Les résultats du cadre max-stable avaient déjà montré que l'identification correcte d'une dépendance ne garantit pas nécessairement un gain de prédiction lorsque celle-ci est essentiellement liée à la distance. Les données réelles confirment cette observation : même en présence d'une anisotropie identifiable, le passage d'une structure de dépendance estimée globalement sur le réseau à des poids efficaces pour une prédiction par interpolation reste difficile.

Cette absence de gain ne signifie donc pas que la dépendance extrême est dépourvue d'information. Elle indique plutôt que la formulation retenue par la méthode 1b n'exploite pas suffisamment cette information pour arriver à une méthode d'interpolation plus efficace. Le modèle repose sur un nombre limité de covariables et suppose une relation commune à l'ensemble du domaine, alors que certaines variations peuvent être locales ou dépendre de mécanismes météorologiques non représentés. Une structure de dépendance plus flexible ou l'introduction d'autres covariables pourraient éventuellement améliorer cette approche, mais leur intérêt ne peut pas être établi à partir des résultats actuels.

L'analyse de la méthode 2 conduit à une conclusion assez similaire quant à la structure spatiale. L'altitude et la distance à la côte, déjà sélectionnées préférentiellement lors de l'ajustement du modèle générateur sur l'ensemble des stations, sont également retenues presque systématiquement dans les différents folds du *leave-one-station-out*. La structure du paramètre de position reste stable, tandis que le paramètre d'échelle reste constant. Cette cohérence montre que la méthode identifie une tendance spatiale reproductible, mais les erreurs persistantes à certaines stations indiquent que cette structure ne suffit pas à représenter toutes les particularités locales. Les simulations avaient déjà montré que la position dans le domaine pouvait influencer les performances. Les données réelles renforcent cette observation : plusieurs des stations les mieux prédites se trouvent au centre du réseau, alors qu'une grande partie des stations présentant les erreurs les plus importantes se situe à sa périphérie. L'information spatiale disponible autour d'une station semble donc constituer un élément important de la qualité des prédictions, indépendamment de la méthode utilisée.

L'incertitude des niveaux de retour marginaux de référence apporte enfin une nuance supplémentaire qui ne pouvait pas être observée de la même manière dans les simulations, où les valeurs théoriques étaient connues. Pour certaines stations difficiles à prédire, notamment 235, 275 et surtout 380, les niveaux de retour de référence présentent eux-mêmes une forte incertitude. Une partie de l'erreur attribuée aux méthodes peut donc provenir de cette référence. Ce mécanisme ne permet toutefois pas d'expliquer l'ensemble des difficultés observées, puisque les stations 240, 270 et 375 présentent également des erreurs importantes malgré des niveaux de retour de référence relativement plus précis. Les écarts observés sur les données réelles résultent ainsi vraisemblablement de plusieurs sources combinées : incertitude marginale, position de la station dans le réseau, limites des covariables disponibles et incapacité des structures globales à représenter certains comportements locaux.

Dans l'ensemble, l'application aux données réelles confirme donc plusieurs tendances mises en évidence par les simulations. La méthode 2 apparaît comme l'approche la plus intéressante pour les périodes de retour élevées, mais son avantage sur la méthode 1a reste modéré et n'est pas observé pour $T = 10$. La méthode 1a se distingue par sa robustesse et son faible niveau de complexité. La méthode 1b parvient quant à elle à identifier une structure de dépendance stable et plus complexe que celle représentée dans les simulations, mais cette information ne conduit pas à une amélioration des performances prédictives avec les covariables actuellement sélectionnées. Il n'existe donc pas de méthode supérieure sur tous les aspects considérés : chacune des méthodes apporte des informations supplémentaires, dépendant de si l'on s'intéresse à l'horizon d'extrapolation, à l'information spatiale disponible ou à la capacité du modèle à représenter les particularités locales du domaine.

Ces conclusions doivent cependant être nuancées. En effet, la taille réduite du réseau et la procédure de validation employée présente des faiblesses. La forte stabilité des structures sélectionnées par les méthodes 1b et 2 est encourageante, mais les folds du *leave-one-station-out* partagent onze stations sur douze et sont donc très similaires. Une validation retirant simultanément plusieurs stations permettrait de tester plus sévèrement cette stabilité. De même, l’ajout de covariables géographiques ou météorologiques supplémentaires et l’utilisation de structures de dépendance capables de varier localement pourraient modifier les performances observées. Les résultats obtenus permettent néanmoins de dégager une conclusion générale : dans le domaine et avec les données considérées, une interpolation simple par la distance constitue une méthode particulièrement compétitive, tandis que la GEV spatiale apporte un gain modéré mais plus régulier pour les niveaux de retour les plus éloignés. L’exploitation de la dépendance extrême est une méthodologie pertinente, mais nécessite ici une modélisation plus riche pour que l’information supplémentaire qu’elle contient se traduise effectivement en amélioration de la prédiction. Cette dernière apporte cependant des informations complémentaires aux autres méthodes en estimant spatialement la structure de dépendance des données.

8 Conclusion

Ce mémoire avait pour objectif d’étudier et de comparer différentes manières d’exploiter l’information spatiale afin d’estimer des niveaux de retour de précipitations extrêmes en des sites non observés. Cette problématique se situe à l’intersection de la théorie des valeurs extrêmes et de l’interpolation spatiale : il ne s’agit pas uniquement de caractériser les fortes précipitations aux stations où des observations sont disponibles, mais également de déterminer comment cette information peut être transférée vers d’autres localisations. Trois approches de complexité croissante ont ainsi été comparées : une interpolation fondée uniquement sur la distance géographique, une interpolation construite à partir de la dépendance entre les extrêmes et, enfin, une modélisation GEV spatiale intégrant directement des caractéristiques géographiques des stations.

La première partie du mémoire a permis d’établir le cadre théorique nécessaire à cette comparaison. La théorie des valeurs extrêmes fournit, à travers la distribution GEV, un cadre adapté à la modélisation des maxima et à l’estimation de niveaux de retour. Les principes de l’interpolation spatiale ont ensuite montré que la proximité géographique ne traduit pas nécessairement à elle seule la similarité entre les comportements extrêmes de deux stations. L’introduction du F-madogramme et de covariables caractérisant les paires de stations permet ainsi d’envisager une interpolation fondée plus directement sur leur dépendance extrême. La revue de littérature a replacé ces différentes approches dans le contexte de l’estimation spatiale des précipitations extrêmes et a souligné qu’une augmentation de la complexité du modèle ne garantit pas nécessairement une amélioration des performances prédictives.

L’étude empirique a été menée à partir de données de précipitations journalières du KNMI provenant de treize stations néerlandaises et couvrant la période allant de 1984 à 2024. Les maxima hivernaux ont été extraits afin de construire les séries utilisées dans l’analyse. Trois méthodes ont ensuite été comparées pour des périodes de retour de 10, 50 et 100 ans. La méthode 1a repose sur des ajustements GEV marginaux suivis d’une interpolation IDW des niveaux de retour. La méthode 1b remplace la proximité purement géographique par des poids construits à partir d’une structure de dépendance estimée au moyen du F-madogramme. Enfin, la méthode 2 ajuste directement un modèle GEV spatial dans lequel les paramètres marginaux peuvent dépendre de covariables géographiques.

La comparaison a d’abord été menée dans deux cadres de simulation complémentaires. Le cadre GEV spatial reproduisait directement une structure marginale correspondant à celle exploitée par la méthode 2, sans introduire de dépendance explicite entre les stations. Le second cadre conservait cette structure marginale mais ajoutait une dépendance spatiale au moyen d’un processus max-stable de Brown–Resnick. Cette distinction s’est révélée importante pour l’interprétation des résultats : les performances relatives des méthodes dépendent en partie du mécanisme générateur et de son adéquation avec leurs hypothèses. Les simulations ont également mis en évidence deux tendances générales : la difficulté de prédiction augmente avec la période de retour et les stations situées en périphérie du réseau sont généralement plus difficiles à estimer que celles mieux entourées.

Ces simulations ont surtout montré qu’aucune méthode ne pouvait être jugée indépendamment du contexte dans lequel elle est appliquée. La méthode 2 bénéficie logiquement d’une structure marginale conforme à son modèle lorsque les données sont générées sous le cadre GEV spatial. À l’inverse, l’absence de dépendance dans ce cadre limite fortement l’intérêt de la méthode 1b. Lorsque cette dépendance est introduite dans le cadre max-stable, la méthode 1b parvient à identifier une structure spatiale cohérente,

sans pour autant améliorer systématiquement les prédictions. La méthode 1a apparaît quant à elle relativement robuste dans les deux configurations malgré sa simplicité. Ces résultats ont donc justifié la nécessité d'une validation sur les données réelles, pour laquelle aucun mécanisme générateur connu ne favorise a priori une méthode particulière.

Cette comparaison a été réalisée au moyen d'une validation *leave-one-station-out*, chaque station étant successivement retirée avant d'être prédite à partir des douze restantes. Les résultats globaux montrent des performances relativement proches entre les méthodes 1a et 2. Pour une période de retour de 10 ans, la méthode 1a fournit les erreurs les plus faibles, bien que l'écart avec la méthode 2 soit très limité. Pour les périodes de retour de 50 et 100 ans, la méthode 2 devient légèrement plus performante. La méthode 1b présente quant à elle les erreurs globales les plus élevées pour les trois périodes de retour. Ces écarts restent toutefois modérés entre les méthodes 1a et 2 et ne permettent pas de conclure à une domination nette de l'approche la plus complexe.

L'analyse station par station a par ailleurs montré que les performances globales masquent une hétérogénéité spatiale importante. Certaines stations sont bien prédites par les trois méthodes, tandis que d'autres concentrent une part importante des erreurs, notamment en périphérie du domaine. La quantité et la répartition de l'information disponible autour d'un site constituent ainsi des éléments importants de la qualité de la prédiction. L'incertitude des niveaux de retour utilisés comme références nuance également l'interprétation de ces erreurs : pour certaines stations, les estimations marginales présentent elles-mêmes des intervalles de confiance larges, de sorte qu'une partie des écarts observés ne peut être attribuée uniquement aux méthodes d'interpolation.

Au-delà des performances prédictives, les deux méthodes les plus complexes apportent toutefois des informations supplémentaires sur l'organisation spatiale des extrêmes. La méthode 1b met en évidence une structure de dépendance stable principalement liée à la distance entre stations ainsi qu'à leur orientation, suggérant une anisotropie de la dépendance extrême. La méthode 2 identifie quant à elle de manière récurrente l'altitude et la distance à la côte comme principales covariables du paramètre de position de la GEV, tandis que le paramètre d'échelle reste spatialement constant dans les modèles sélectionnés. Ces résultats montrent que les précipitations extrêmes possèdent des structures spatiales qui ne se résument pas entièrement à la seule distance géographique, même si leur prise en compte explicite n'entraîne pas nécessairement une amélioration importante de la prédiction.

Dans le contexte étudié, la méthode 1a apparaît finalement comme l'approche offrant le compromis le plus favorable entre simplicité, robustesse et performances prédictives. Elle ne domine pas systématiquement la méthode 2, notamment pour les périodes de retour élevées, mais les gains obtenus par cette dernière restent modestes au regard de la complexité supplémentaire introduite. La méthode 2 présente néanmoins l'intérêt de modéliser explicitement les variations spatiales des paramètres marginaux, tandis que la méthode 1b fournit une caractérisation complémentaire de la dépendance entre événements extrêmes. Le choix entre ces approches ne doit donc pas reposer uniquement sur leur erreur de prédiction, mais également sur l'objectif poursuivi et sur le type d'information spatiale recherché.

Plusieurs prolongements permettraient de préciser ces conclusions. L'application des méthodes à un réseau comprenant davantage de stations, à des séries temporelles plus longues ou à un domaine géographique plus étendu permettrait d'évaluer leur comportement dans des configurations spatiales plus diversifiées. Pour la méthode 1b, l'introduction de covariables décrivant plus directement les mécanismes météorologiques, ou l'utilisation de structures de dépendance plus flexibles et éventuellement non stationnaires, pourrait mieux représenter certaines particularités locales. Pour la méthode 2, d'autres covariables ou des relations moins restrictives entre celles-ci et les paramètres de la GEV pourraient également être étudiées. Enfin, une validation spatiale plus exigeante, par exemple en retirant simultanément plusieurs stations ou des zones entières, permettrait de mieux évaluer la capacité réelle d'extrapolation des différentes méthodes.

Ce mémoire montre ainsi que la prise en compte d'une information spatiale plus complexe peut enrichir la compréhension de la structure des précipitations extrêmes, mais qu'elle ne se traduit pas automatiquement par de meilleures prédictions. Dans le réseau étudié, la distance géographique constitue déjà une information particulièrement efficace pour l'interpolation des niveaux de retour, tandis que les approches plus complexes apportent principalement une description supplémentaire des mécanismes spatiaux. Leur intérêt dépend dès lors du compromis recherché entre performance prédictive, complexité du modèle et compréhension de la structure spatiale des extrêmes.

Références

- [1] Stuart COLES. « Introduction ». en. In : *An Introduction to Statistical Modeling of Extreme Values*. Sous la dir. de Stuart COLES. London : Springer, 2001, p. 1-17. ISBN : 978-1-4471-3675-0. DOI : [10.1007/978-1-4471-3675-0_1](https://doi.org/10.1007/978-1-4471-3675-0_1). URL : https://doi.org/10.1007/978-1-4471-3675-0_1 (visité le 09/05/2026).
- [2] Miguel DE CARVALHO et al. « Handbook Outline ». In : *Handbook of Statistics of Extremes*. New York : Chapman et Hall/CRC, 2026.
- [3] Myriam CHARRAS-GARRIDO et Pascal LEZAUD. « Extreme value analysis : an introduction ». In : *Journal de la société française de statistique* 154.2 (2013), p. 66-97. ISSN : 2102-6238.
- [4] Anthony DAVISON et Ophelia MIRALLES. « Modeling Univariate Extremes—Why and How ». In : *Handbook of Statistics of Extremes*. New York : Chapman et Hall/CRC, 2026.
- [5] Marielle PINHEIRO et Richard GROTJAHN. « An introduction to extreme value statistics ». In : *Tutorial at University of California at Davis* (2015).
- [6] Stuart COLES. « Classical Extreme Value Theory and Models ». en. In : *An Introduction to Statistical Modeling of Extreme Values*. Sous la dir. de Stuart COLES. London : Springer, 2001, p. 45-73. ISBN : 978-1-4471-3675-0. DOI : [10.1007/978-1-4471-3675-0_3](https://doi.org/10.1007/978-1-4471-3675-0_3). URL : https://doi.org/10.1007/978-1-4471-3675-0_3 (visité le 30/04/2026).
- [7] M. A. Ben ALAYA, F. ZWIERS et X. ZHANG. « An Evaluation of Block-Maximum-Based Estimation of Very Long Return Period Precipitation Extremes with a Large Ensemble Climate Simulation ». In : *Journal of Climate* 33.16 (2020), p. 6957-6970. ISSN : 0894-8755. URL : <https://www.jstor.org/stable/26937998> (visité le 09/05/2026).
- [8] Farabe Khan ALIF et Norhaslinda ALI. « Comparative study of threshold selection methods in the generalised Pareto distribution with application to rainfall datasets ». en. In : *Data Analytics and Applied Mathematics (DAAM)* (mars 2026), p. 50-67. ISSN : 2773-4854. DOI : [10.15282/daam.v7i1.13686](https://doi.org/10.15282/daam.v7i1.13686). URL : <https://journal.ump.edu.my/daam/article/view/13686> (visité le 09/05/2026).
- [9] Stuart COLES, Janet HEFFERNAN et Jonathan TAWN. « Dependence Measures for Extreme Value Analyses ». en. In : *Extremes* 2.4 (déc. 1999), p. 339-365. ISSN : 1572-915X. DOI : [10.1023/A:1009963131610](https://doi.org/10.1023/A:1009963131610). URL : <https://doi.org/10.1023/A:1009963131610> (visité le 09/05/2026).
- [10] Miguel DE CARVALHO et Viviana CARCAISO. « Learning About Extreme Value Distributions from Data ». In : *Handbook of Statistics of Extremes*. New York : Chapman et Hall/CRC, 2026.
- [11] Arvid NAESS. « Classical Extreme Value Theory ». en. In : *Applied Extreme Value Statistics : With a Special Focus on the ACER Method*. Sous la dir. d'Arvid NAESS. Cham : Springer Nature Switzerland, 2024, p. 5-18. ISBN : 978-3-031-60769-1. DOI : [10.1007/978-3-031-60769-1_2](https://doi.org/10.1007/978-3-031-60769-1_2). URL : https://doi.org/10.1007/978-3-031-60769-1_2 (visité le 30/04/2026).
- [12] M. Ivette GOMES et Armelle GUILLOU. « Extreme Value Theory and Statistics of Univariate Extremes : A Review ». en. In : *International Statistical Review* 83.2 (2015). _eprint : <https://onlinelibrary.wiley.com/doi/pdf/10.1111/insr.12058>, p. 263-292. ISSN : 1751-5823. DOI : [10.1111/insr.12058](https://doi.org/10.1111/insr.12058). URL : <https://onlinelibrary.wiley.com/doi/abs/10.1111/insr.12058> (visité le 30/04/2026).
- [13] Diana RYPKEMA et Shripad TULJAPURKAR. « Modeling extreme climatic events using the generalized extreme value (GEV) distribution ». en-US. In : *Handbook of Statistics*. T. 44. Elsevier, jan. 2021, p. 39-71. DOI : [10.1016/bs.host.2020.12.002](https://doi.org/10.1016/bs.host.2020.12.002). URL : <https://www.sciencedirect.com/science/chapter/handbook/abs/pii/S0169716120300511> (visité le 09/05/2026).
- [14] Stuart COLES. « Extremes of Non-stationary Sequences ». en. In : *An Introduction to Statistical Modeling of Extreme Values*. Sous la dir. de Stuart COLES. London : Springer, 2001, p. 105-123. ISBN : 978-1-4471-3675-0. DOI : [10.1007/978-1-4471-3675-0_6](https://doi.org/10.1007/978-1-4471-3675-0_6). URL : https://doi.org/10.1007/978-1-4471-3675-0_6 (visité le 30/04/2026).
- [15] Miguel DE CARVALHO et al. « Regression Models for Extreme Events ». In : *Handbook of Statistics of Extremes*. New York : Chapman et Hall/CRC, 2026.
- [16] Richard W. KATZ. « Statistical Methods for Nonstationary Extremes ». In : *Extremes in a Changing Climate : Detection, Analysis and Uncertainty*. Sous la dir. d'Amir AGHAKOUCHAK et al. Dordrecht : Springer Netherlands, 2013, p. 15-37. ISBN : 978-94-007-4479-0. DOI : [10.1007/978-94-007-4479-0_2](https://doi.org/10.1007/978-94-007-4479-0_2). URL : https://doi.org/10.1007/978-94-007-4479-0_2.

- [17] Phuong LE, Michael LEONARD et Seth WESTRA. « Modeling Spatial Dependence of Rainfall Extremes Across Multiple Durations ». In : *Water Resources Research* 54 (mars 2018). DOI : [10.1002/2017WR022231](https://doi.org/10.1002/2017WR022231).
- [18] Julie CARREAU et Stéphane GIRARD. « Spatial extreme quantile estimation using a weighted log-likelihood approach ». en. In : *Journal de la société française de statistique* 152.3 (2011), p. 66-82. ISSN : 2102-6238. URL : https://www.numdam.org/item/JSFS_2011__152_3_66_0/ (visité le 22/05/2026).
- [19] Alaba BOLUWADE. « Stochastic modeling of spatial dependency structures of extreme precipitation in the Northern Great Plains using max-stable processes ». In : *Journal of Water and Climate Change* 14.9 (août 2023), p. 3131-3149. ISSN : 2040-2244. DOI : [10.2166/wcc.2023.187](https://doi.org/10.2166/wcc.2023.187). URL : <https://doi.org/10.2166/wcc.2023.187> (visité le 21/05/2026).
- [20] Hanbeen KIM et al. « Appropriate model selection methods for nonstationary generalized extreme value models ». In : *Journal of Hydrology* 547 (avr. 2017), p. 557-574. ISSN : 0022-1694. DOI : [10.1016/j.jhydrol.2017.02.005](https://doi.org/10.1016/j.jhydrol.2017.02.005). URL : <https://www.sciencedirect.com/science/article/pii/S0022169417300756> (visité le 11/05/2026).
- [21] Taha B. M. J. OUARDA et Christian CHARRON. « Nonstationary Temperature-Duration-Frequency curves ». en. In : *Scientific Reports* 8.1 (oct. 2018), p. 15493. ISSN : 2045-2322. DOI : [10.1038/s41598-018-33974-y](https://doi.org/10.1038/s41598-018-33974-y). URL : <https://www.nature.com/articles/s41598-018-33974-y> (visité le 11/05/2026).
- [22] Ritei SHIBATA. « Statistical Aspects of Model Selection ». en. In : *From Data to Model*. Sous la dir. de Jan C. WILLEMS. Berlin, Heidelberg : Springer, 1989, p. 215-240. ISBN : 978-3-642-75007-6. DOI : [10.1007/978-3-642-75007-6_5](https://doi.org/10.1007/978-3-642-75007-6_5). URL : https://doi.org/10.1007/978-3-642-75007-6_5 (visité le 11/05/2026).
- [23] Yong LI et al. « A Note on AIC and TIC for Model Selection ». en. In : *Working Papers* (nov. 2024). Number : 202420. URL : <https://ideas.repec.org/p/boa/wpaper/202420.html> (visité le 11/05/2026).
- [24] Scott I. VRIEZE. « Model selection and psychological theory : A discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC) ». In : *Psychological Methods* 17.2 (2012), p. 228-243. ISSN : 1939-1463. DOI : [10.1037/a0027127](https://doi.org/10.1037/a0027127).
- [25] Mathieu RIBATET, Richard SINGLETON et R. Core TEAM. *SpatialExtremes : Modelling Spatial Extremes*. Avr. 2022. URL : <https://cran.r-project.org/web/packages/SpatialExtremes/index.html> (visité le 11/05/2026).
- [26] J. SZOLGAY et al. « Comparison of mapping approaches of design annual maximum daily precipitation ». In : *Atmospheric Research*. 7th International Workshop on Precipitation in Urban Areas 92.3 (mai 2009), p. 289-307. ISSN : 0169-8095. DOI : [10.1016/j.atmosres.2009.01.009](https://doi.org/10.1016/j.atmosres.2009.01.009). URL : <https://www.sciencedirect.com/science/article/pii/S016980950900026X> (visité le 11/05/2026).
- [27] Jin LI et Andrew D. HEAP. « Spatial interpolation methods applied in the environmental sciences : A review ». In : *Environmental Modelling & Software* 53 (mars 2014), p. 173-189. ISSN : 1364-8152. DOI : [10.1016/j.envsoft.2013.12.008](https://doi.org/10.1016/j.envsoft.2013.12.008). URL : <https://www.sciencedirect.com/science/article/pii/S1364815213003113> (visité le 18/05/2026).
- [28] Noel CRESSIE. « Statistics for spatial data ». In : *Terra Nova* 4.5 (1992), p. 613-617. ISSN : 0954-4879.
- [29] Jin LI et Andrew D. HEAP. « A review of spatial interpolation methods for environmental scientists ». In : (2008).
- [30] Donald SHEPARD et VIEW PROFILE. « A two-dimensional interpolation function for irregularly-spaced data ». In : *Proceedings of the 1968 23rd ACM national conference*. ACM Conferences. Jan. 1968, p. 517-524. ISBN : 978-1-4503-7486-6. DOI : [10.1145/800186.810616](https://doi.org/10.1145/800186.810616). URL : <https://dl.acm.org/doi/abs/10.1145/800186.810616> (visité le 18/05/2026).
- [31] Roger W SINNOTT. « Virtues of the Haversine ». In : *Sky and telescope* 68.2 (1984), p. 158. ISSN : 0037-6604.

- [32] E MARIA et al. « Measure distance locating nearest public facilities using Haversine and Euclidean Methods ». en. In : *Journal of Physics : Conference Series* 1450.1 (fév. 2020), p. 012080. ISSN : 1742-6596. DOI : [10.1088/1742-6596/1450/1/012080](https://doi.org/10.1088/1742-6596/1450/1/012080). URL : <https://doi.org/10.1088/1742-6596/1450/1/012080> (visité le 18/05/2026).
- [33] A. A. SABZIPARVAR et al. « Geographical factors affecting variability of precipitation regime in Iran ». en. In : *Theoretical and Applied Climatology* 120.1 (avr. 2015), p. 367-376. ISSN : 1434-4483. DOI : [10.1007/s00704-014-1174-3](https://doi.org/10.1007/s00704-014-1174-3). URL : <https://doi.org/10.1007/s00704-014-1174-3> (visité le 19/05/2026).
- [34] Yangna LEI et al. « Variation Characteristics of Temperature and Rainfall and Their Relationship with Geographical Factors in the Qinling Mountains ». en. In : *Atmosphere* 14.4 (avr. 2023), p. 696. ISSN : 2073-4433. DOI : [10.3390/atmos14040696](https://www.mdpi.com/2073-4433/14/4/696). URL : <https://www.mdpi.com/2073-4433/14/4/696> (visité le 19/05/2026).
- [35] Jacques Célestin Moliba BANKANZA. « Time variation of the effect of geographical factors on spatial distribution of summer precipitation over the Czech Republic ». In : (2011).
- [36] Swathi SHETTY, Pruthviraj UMESH et Amba SHETTY. « Dependability of rainfall to topography and micro-climate : an observation using geographically weighted regression ». en. In : *Theoretical and Applied Climatology* 147.1 (jan. 2022), p. 217-237. ISSN : 1434-4483. DOI : [10.1007/s00704-021-03811-w](https://doi.org/10.1007/s00704-021-03811-w). URL : <https://doi.org/10.1007/s00704-021-03811-w> (visité le 19/05/2026).
- [37] H. W. ter MAAT et al. « Exploring the Impact of Land Cover and Topography on Rainfall Maxima in the Netherlands ». EN. In : *Journal of Hydrometeorology* 14.2 (avr. 2013), p. 524-542. ISSN : 1525-7541, 1525-755X. DOI : [10.1175/JHM-D-12-036.1](https://journals.ametsoc.org/view/journals/hydr/14/2/jhm-d-12-036_1.xml). URL : https://journals.ametsoc.org/view/journals/hydr/14/2/jhm-d-12-036_1.xml (visité le 19/05/2026).
- [38] E. E. DANIELS et al. « Spatial precipitation patterns and trends in The Netherlands during 1951-2009. » eng. In : *International Journal of Climatology* 34.6 (mai 2014), p. 1773-1784. ISSN : 0899-8418. DOI : [10.1002/joc.3800](https://research.ebsco.com/linkprocessor/plink?id=672b2245-f62d-3166-b764-45e6a3cb466a). URL : <https://research.ebsco.com/linkprocessor/plink?id=672b2245-f62d-3166-b764-45e6a3cb466a> (visité le 19/05/2026).
- [39] G. LENDERINK, E. van MEIJGAARD et F. SELTEN. « Intense coastal rainfall in the Netherlands in response to high sea surface temperatures : analysis of the event of August 2006 from the perspective of a changing climate ». en. In : *Climate Dynamics* 32.1 (jan. 2009), p. 19-33. ISSN : 1432-0894. DOI : [10.1007/s00382-008-0366-x](https://doi.org/10.1007/s00382-008-0366-x). URL : <https://doi.org/10.1007/s00382-008-0366-x> (visité le 19/05/2026).
- [40] Dan COOLEY, Philippe NAVEAU et Paul PONCET. « Variograms for spatial max-stable random fields ». In : *Dependence in probability and statistics*. Springer, 2006, p. 373-390.
- [41] Mathieu RIBATET, Clément DOMBRY et Marco OESTING. « Spatial extremes and max-stable processes ». In : *Extreme value modeling and risk analysis : methods and applications* (2016), p. 179-194.
- [42] Elsa BERNARD et al. « Clustering of Maxima : Spatial Dependencies among Heavy Rainfall in France ». EN. In : *Journal of Climate* 26.20 (oct. 2013), p. 7929-7937. ISSN : 0894-8755, 1520-0442. DOI : [10.1175/JCLI-D-12-00836.1](https://journals.ametsoc.org/view/journals/clim/26/20/jcli-d-12-00836.1.xml). URL : <https://journals.ametsoc.org/view/journals/clim/26/20/jcli-d-12-00836.1.xml> (visité le 19/05/2026).
- [43] Justus CONTZEN, Thorsten DICKHAUS et Gerrit LOHMANN. « Regionalization approaches for the spatial analysis of extremal dependence ». en. In : *Extremes* 28.4 (déc. 2025), p. 713-737. ISSN : 1572-915X. DOI : [10.1007/s10687-025-00519-2](https://doi.org/10.1007/s10687-025-00519-2). URL : <https://doi.org/10.1007/s10687-025-00519-2> (visité le 20/05/2026).
- [44] Ivan SOENARIO, Maarten PLIEGER et Raymond SLUITER. *Optimization of rainfall interpolation*. Rapp. tech. Koninklijk Nederlands Meteorologisch Instituut Nederlands, 2010.
- [45] Ana FERREIRA et al. « Estimating space-time trend and dependence of heavy rainfall ». In : *arXiv preprint arXiv :1707.04434* (2017).
- [46] K. R. SAUNDERS, A. G. STEPHENSON et D. J. KAROLY. « A regionalisation approach for rainfall based on extremal dependence ». en. In : *Extremes* 24.2 (juin 2021), p. 215-240. ISSN : 1572-915X. DOI : [10.1007/s10687-020-00395-y](https://doi.org/10.1007/s10687-020-00395-y). URL : <https://doi.org/10.1007/s10687-020-00395-y> (visité le 28/07/2026).
- [47] S BEKTAS. « Direct bearing angles determination on globe ». In : *MOJ Civil Engineering* 5.4 (2019), p. 78-80.

- [48] Robert J. HIJMANS et al. *geosphere : Spherical Trigonometry*. Avr. 2026. URL : <https://cran.r-project.org/web/packages/geosphere/index.html> (visité le 20/05/2026).
- [49] E. G. BEEK, A. STEIN et L. L. F. JANSSEN. « Spatial variability and interpolation of daily precipitation amount ». en. In : *Stochastic Hydrology and Hydraulics* 6.4 (déc. 1992), p. 304-320. ISSN : 1435-151X. DOI : [10.1007/BF01581623](https://doi.org/10.1007/BF01581623). URL : <https://doi.org/10.1007/BF01581623> (visité le 28/07/2026).
- [50] Richard W KATZ, Marc B PARLANGE et Philippe NAVEAU. « Statistics of extremes in hydrology ». In : *Advances in water resources* 25.8-12 (2002), p. 1287-1304. ISSN : 0309-1708.
- [51] Matteo GENTILUCCI et al. « GEV Analysis of Extreme Rainfall : Comparing Different Time Intervals to Analyse Model Response in Terms of Return Levels in the Study Area of Central Italy ». en. In : *Sustainability* 15.15 (jan. 2023), p. 11656. ISSN : 2071-1050. DOI : [10.3390/su151511656](https://doi.org/10.3390/su151511656). URL : <https://www.mdpi.com/2071-1050/15/15/11656> (visité le 20/05/2026).
- [52] Jieling WU. « Extreme Rainfall Analysis and Return Period Estimation Based on Extreme Value Theory ». en. In : *Engineering Proceedings* 128.1 (2026), p. 31. ISSN : 2673-4591. DOI : [10.3390/engproc2026128031](https://doi.org/10.3390/engproc2026128031). URL : <https://www.mdpi.com/2673-4591/128/1/31> (visité le 20/05/2026).
- [53] G. LAZOGLOU et al. « A review of statistical methods to analyze extreme precipitation and temperature events in the Mediterranean region ». In : *Theoretical and Applied Climatology* 136.1-2 (2019), p. 99-117. DOI : [10.1007/s00704-018-2467-8](https://doi.org/10.1007/s00704-018-2467-8). URL : <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85045233810&doi=10.1007%2fs00704-018-2467-8&partnerID=40&md5=3c7adb7f13d363b82366ea97fde19307>.
- [54] V. AHUJA et al. « Statistical modelling of century-long precipitation and temperature extremes in Himachal Pradesh, India : generalized extreme value approach and return level estimation ». In : *Indian Journal of Physics* 98.8 (2024), p. 2653-2664. DOI : [10.1007/s12648-023-03011-4](https://doi.org/10.1007/s12648-023-03011-4). URL : <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85180256404&doi=10.1007%2fs12648-023-03011-4&partnerID=40&md5=7784eececafe5d743d3bb1b74f201be6>.
- [55] S. VANNITSEM et P. NAVEAU. « Spatial dependences among precipitation maxima over Belgium ». English. In : *Nonlinear Processes in Geophysics* 14.5 (sept. 2007), p. 621-630. ISSN : 1023-5809. DOI : [10.5194/npg-14-621-2007](https://doi.org/10.5194/npg-14-621-2007). URL : <https://npg.copernicus.org/articles/14/621/2007/> (visité le 09/05/2026).
- [56] Francesco MARRA et al. « Metastatistical Extreme Value analysis of hourly rainfall from short records : Estimation of high quantiles and impact of measurement errors ». In : *Advances in Water Resources* 117 (juill. 2018), p. 27-39. ISSN : 0309-1708. DOI : [10.1016/j.advwatres.2018.05.001](https://doi.org/10.1016/j.advwatres.2018.05.001). URL : <https://www.sciencedirect.com/science/article/pii/S0309170818302033> (visité le 21/05/2026).
- [57] Simon Michael PAPALEXIOU et Demetris KOUTSOYIANNIS. « Battle of extreme value distributions : A global survey on extreme daily rainfall ». en. In : *Water Resources Research* 49.1 (2013). _eprint : <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2012WR012557>, p. 187-201. ISSN : 1944-7973. DOI : [10.1029/2012WR012557](https://doi.org/10.1029/2012WR012557). URL : <https://onlinelibrary.wiley.com/doi/abs/10.1029/2012WR012557> (visité le 21/05/2026).
- [58] Francesco DE PAOLA et al. « GEV Parameter Estimation and Stationary vs. Non-Stationary Analysis of Extreme Rainfall in African Test Cities ». en. In : *Hydrology* 5.2 (juin 2018), p. 28. ISSN : 2306-5338. DOI : [10.3390/hydrology5020028](https://doi.org/10.3390/hydrology5020028). URL : <https://www.mdpi.com/2306-5338/5/2/28> (visité le 21/05/2026).
- [59] Alexander Lee RISCHMULLER, Benjamin POSCHLOD et Jana SILLMANN. « Bayesian hierarchical modelling of intensity-duration-frequency curves using a climate model large ensemble ». English. In : *Advances in Statistical Climatology, Meteorology and Oceanography* 12.1 (jan. 2026), p. 1-19. ISSN : 2364-3579. DOI : [10.5194/ascmo-12-1-2026](https://doi.org/10.5194/ascmo-12-1-2026). URL : <https://ascmo.copernicus.org/articles/12/1/2026/> (visité le 21/05/2026).
- [60] Michael F. WEHNER et al. « On the uncertainty of long-period return values of extreme daily precipitation ». English. In : *Frontiers in Climate* 6 (mars 2024). ISSN : 2624-9553. DOI : [10.3389/fclim.2024.1343072](https://doi.org/10.3389/fclim.2024.1343072). URL : <https://www.frontiersin.org/journals/climate/articles/10.3389/fclim.2024.1343072/full> (visité le 21/05/2026).

- [61] I. HOSSAIN et al. « Comparison of estimation techniques for generalised extreme value (GEV) distribution parameters : a case study with Tasmanian rainfall ». en. In : *International Journal of Environmental Science and Technology* 19.8 (août 2022), p. 7737-7750. ISSN : 1735-2630. DOI : [10.1007/s13762-021-03693-5](https://doi.org/10.1007/s13762-021-03693-5). URL : <https://doi.org/10.1007/s13762-021-03693-5> (visité le 21/05/2026).
- [62] P. C. D. MILLY et al. « Stationarity Is Dead : Whither Water Management ? » In : *Science* 319.5863 (fév. 2008), p. 573-574. DOI : [10.1126/science.1151915](https://doi.org/10.1126/science.1151915). URL : <https://www.science.org/doi/10.1126/science.1151915> (visité le 21/05/2026).
- [63] H. J. FOWLER et C. G. KILSBY. « A regional frequency analysis of United Kingdom extreme rainfall from 1961 to 2000 ». en. In : *International Journal of Climatology* 23.11 (2003). _eprint : <https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/joc.943>, p. 1313-1334. ISSN : 1097-0088. DOI : [10.1002/joc.943](https://doi.org/10.1002/joc.943). URL : <https://onlinelibrary.wiley.com/doi/abs/10.1002/joc.943> (visité le 21/05/2026).
- [64] Roberto DEIDDA, Matteo HELLIES et Andreas LANGOUSIS. « A critical analysis of the shortcomings in spatial frequency analysis of rainfall extremes based on homogeneous regions and a comparison with a hierarchical boundaryless approach ». In : *Stochastic Environmental Research and Risk Assessment* 35.12 (2021), p. 2605-2628. ISSN : 1436-3240.
- [65] Philomène LE GALL et al. « Improved regional frequency analysis of rainfall data ». In : *Weather and Climate Extremes* 36 (2022), p. 100456. ISSN : 2212-0947.
- [66] H. VAN DE VYVER. « Spatial regression models for extreme precipitation in Belgium ». en. In : *Water Resources Research* 48.9 (2012). ISSN : 1944-7973. DOI : [10.1029/2011WR011707](https://doi.org/10.1029/2011WR011707). URL : <https://onlinelibrary.wiley.com/doi/abs/10.1029/2011WR011707> (visité le 21/05/2026).
- [67] M. FISCHER, H. W. RUST et U. ULBRICH. « A spatial and seasonal climatology of extreme precipitation return-levels : A case study ». In : *Spatial Statistics. Spatio-temporal and geostatistical analysis of hydrological events and related hazards* 34 (déc. 2019), p. 100275. ISSN : 2211-6753. DOI : [10.1016/j.spasta.2017.11.007](https://doi.org/10.1016/j.spasta.2017.11.007). URL : <https://www.sciencedirect.com/science/article/pii/S2211675317300283> (visité le 21/05/2026).
- [68] Hela HAMMAMI et al. « Smooth Spatial Modeling of Extreme Mediterranean Precipitation ». en. In : *Water* 14.22 (jan. 2022), p. 3782. ISSN : 2073-4441. DOI : [10.3390/w14223782](https://doi.org/10.3390/w14223782). URL : <https://www.mdpi.com/2073-4441/14/22/3782> (visité le 21/05/2026).
- [69] Daniel COOLEY, Douglas NYCHKA et Philippe NAVEAU. « Bayesian Spatial Modeling of Extreme Precipitation Return Levels ». In : *Journal of the American Statistical Association* 102.479 (sept. 2007). _eprint : <https://doi.org/10.1198/016214506000000780>, p. 824-840. ISSN : 0162-1459. DOI : [10.1198/016214506000000780](https://doi.org/10.1198/016214506000000780). URL : <https://doi.org/10.1198/016214506000000780> (visité le 21/05/2026).
- [70] C. BRACKEN et al. « Spatial Bayesian hierarchical modeling of precipitation extremes over a large domain ». en. In : *Water Resources Research* 52.8 (2016), p. 6643-6655. ISSN : 1944-7973. DOI : [10.1002/2016WR018768](https://doi.org/10.1002/2016WR018768). URL : <https://onlinelibrary.wiley.com/doi/abs/10.1002/2016WR018768> (visité le 21/05/2026).
- [71] Federica STOLF et Antonio CANALE. « A hierarchical Bayesian non-asymptotic extreme value model for spatial data ». English. In : *Environmetrics* 34.7 (2023). ISSN : 1180-4009. DOI : [10.1002/env.2806](https://doi.org/10.1002/env.2806).
- [72] Sarann LY, Catherine CHARLES et Aurore DEGRÉ. « Different methods for spatial interpolation of rainfall data for operational hydrology and hydrological modeling at watershed scale : a review ». In : *Biotechnologie, agronomie, société et environnement* 17.2 (2013). ISSN : 1370-6233.
- [73] Guillermo Q. TABIOS III et Jose D. SALAS. « A Comparative Analysis of Techniques for Spatial Interpolation of Precipitation ». en. In : *JAWRA Journal of the American Water Resources Association* 21.3 (1985). _eprint : <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1752-1688.1985.tb00147.x>, p. 365-380. ISSN : 1752-1688. DOI : [10.1111/j.1752-1688.1985.tb00147.x](https://doi.org/10.1111/j.1752-1688.1985.tb00147.x). URL : <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1752-1688.1985.tb00147.x> (visité le 22/05/2026).
- [74] Tommaso CALOIERO et al. « Comparative Analysis of Different Spatial Interpolation Methods Applied to Monthly Rainfall as Support for Landscape Management ». en. In : *Applied Sciences* 11.20 (jan. 2021), p. 9566. ISSN : 2076-3417. DOI : [10.3390/app11209566](https://doi.org/10.3390/app11209566). URL : <https://www.mdpi.com/2076-3417/11/20/9566> (visité le 22/05/2026).

- [75] Samiran DAS. « Extreme rainfall estimation at ungauged sites : Comparison between region-of-influence approach of regional analysis and spatial interpolation technique ». en-GB. In : *International Journal of Climatology* 39.1 (jan. 2019), p. 407-423. ISSN : 0899-8418. DOI : [10.1002/joc.5819](https://doi.org/10.1002/joc.5819). URL : <https://eprints.gla.ac.uk/308886/> (visité le 22/05/2026).
- [76] J. BLANCHET et M. LEHNING. « Mapping snow depth return levels : smooth spatial modeling versus station interpolation ». English. In : *Hydrology and Earth System Sciences* 14.12 (déc. 2010), p. 2527-2544. ISSN : 1027-5606. DOI : [10.5194/hess-14-2527-2010](https://doi.org/10.5194/hess-14-2527-2010). URL : https://hess.copernicus.org/articles/14/2527/2010/?utm_source=chatgpt.com (visité le 22/05/2026).
- [77] P. GOOVAERTS. « Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall ». In : *Journal of Hydrology* 228.1 (fév. 2000), p. 113-129. ISSN : 0022-1694. DOI : [10.1016/S0022-1694\(00\)00144-X](https://doi.org/10.1016/S0022-1694(00)00144-X). URL : <https://www.sciencedirect.com/science/article/pii/S002216940000144X> (visité le 22/05/2026).
- [78] Samiran DAS, Dehua ZHU et Yixing YIN. « Comparison of mapping approaches for estimating extreme precipitation of any return period at ungauged locations ». en. In : *Stochastic Environmental Research and Risk Assessment* 34.8 (août 2020), p. 1175-1196. ISSN : 1436-3259. DOI : [10.1007/s00477-020-01828-7](https://doi.org/10.1007/s00477-020-01828-7). URL : <https://doi.org/10.1007/s00477-020-01828-7> (visité le 22/05/2026).
- [79] D. CERESSETTI et al. « Evaluation of classical spatial-analysis schemes of extreme rainfall ». English. In : *Natural Hazards and Earth System Sciences* 12.11 (nov. 2012), p. 3229-3240. ISSN : 1561-8633. DOI : [10.5194/nhess-12-3229-2012](https://doi.org/10.5194/nhess-12-3229-2012). URL : <https://nhess.copernicus.org/articles/12/3229/2012/> (visité le 22/05/2026).
- [80] A. C. DAVISON, S. A. PADOAN et M. RIBATET. « Statistical Modeling of Spatial Extremes ». In : *Statistical Science* 27.2 (2012), p. 161-186. ISSN : 0883-4237. URL : <https://www.jstor.org/stable/41714789> (visité le 22/05/2026).
- [81] Lichao YANG, Christian LE FRANZKE et Wansuo DUAN. « Evaluation and projections of extreme precipitation using a spatial extremes framework ». In : *International Journal of Climatology* 43.7 (2023), p. 3453-3475. ISSN : 0899-8418.
- [82] Edzer PEBESMA et al. *sf : Simple Features for R*. Mai 2026. URL : <https://cran.r-project.org/web/packages/sf/index.html> (visité le 23/05/2026).
- [83] Philippe MASSICOTTE, Andy SOUTH et Koen HUFKENS. *rnaturalearth : World Map Data from Natural Earth*. Jan. 2026. URL : <https://cran.r-project.org/web/packages/rnaturalearth/index.html> (visité le 23/05/2026).
- [84] Geert LENDERINK et al. « Are dependencies of extreme rainfall on humidity more reliable in convection-permitting climate models? » English. In : *Hydrology and Earth System Sciences* 29.4 (mars 2025), p. 1201-1220. ISSN : 1027-5606. DOI : [10.5194/hess-29-1201-2025](https://doi.org/10.5194/hess-29-1201-2025). URL : <https://hess.copernicus.org/articles/29/1201/2025/> (visité le 23/05/2026).
- [85] Oscar E. JURADO, Marco OESTING et Henning W. RUST. « Implications of modeling seasonal differences in the extremal dependence of rainfall maxima ». en. In : *Stochastic Environmental Research and Risk Assessment* 37.5 (mai 2023), p. 1963-1981. ISSN : 1436-3259. DOI : [10.1007/s00477-022-02375-z](https://doi.org/10.1007/s00477-022-02375-z). URL : <https://doi.org/10.1007/s00477-022-02375-z> (visité le 28/07/2026).
- [86] Karin van der WIEL et al. « KNMI'23 Climate Scenarios for the Netherlands : Storyline Scenarios of Regional Climate Change ». en. In : *Earth's Future* 12.2 (2024). _eprint : <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2023EF003983>, e2023EF003983. ISSN : 2328-4277. DOI : [10.1029/2023EF003983](https://doi.org/10.1029/2023EF003983). URL : <https://onlinelibrary.wiley.com/doi/abs/10.1029/2023EF003983> (visité le 23/05/2026).
- [87] Joe CHENG et al. *leaflet : Create Interactive Web Maps with the JavaScript 'Leaflet' Library*. Sept. 2025. URL : <https://cran.r-project.org/web/packages/leaflet/index.html> (visité le 13/08/2026).
- [88] Alec Stephenson Function fbvpot by Chris FERRO. *evd : Functions for Extreme Value Distributions*. Sept. 2024. URL : <https://cran.r-project.org/web/packages/evd/index.html> (visité le 29/06/2026).

- [89] Carolin FORSTER et Marco OESTING. « Non-stationary max-stable models with an application to heavy rainfall data ». en. In : *Extremes* 28.3 (sept. 2025), p. 523-556. ISSN : 1572-915X. DOI : [10.1007/s10687-025-00512-9](https://doi.org/10.1007/s10687-025-00512-9). URL : <https://doi.org/10.1007/s10687-025-00512-9> (visit  le 28/07/2026).
- [90] Eric GILLELAND, Janet E. HEFFERNAN et Alec G. STEPHENSON. *ismev : An Introduction to Statistical Modeling of Extreme Values*. Ao t 2025. URL : <https://cran.r-project.org/web/packages/ismev/index.html> (visit  le 28/07/2026).
- [91] Iris MANOLA et al. « Analysis of urban rainfall from hourly to seasonal scales using high-resolution radar observations in the Netherlands ». en. In : *International Journal of Climatology* 40.2 (2020), p. 822-840. ISSN : 1097-0088. DOI : [10.1002/joc.6241](https://doi.org/10.1002/joc.6241). URL : <https://onlinelibrary.wiley.com/doi/abs/10.1002/joc.6241> (visit  le 14/08/2026).

A Maxima saisonniers de précipitations pour les 13 stations

TABLEAU 13 – Maxima saisonniers de précipitations pour les 13 stations retenues [0.1 mm]. Les stations sont représentées par leurs identifiants STN.

Année	235	240	260	270	275	280	290	310	344	350	370	375	380
1984	275	234	179	343	246	281	238	194	156	192	182	198	302
1985	212	261	359	215	427	181	260	166	356	233	197	263	200
1986	196	331	321	253	367	231	330	228	300	393	286	363	402
1987	260	329	311	281	329	214	248	221	321	244	211	270	164
1988	338	283	232	324	202	286	320	274	202	214	191	336	163
1989	318	240	241	290	217	223	233	235	218	188	191	205	206
1990	171	190	204	233	367	171	226	170	213	244	212	217	350
1991	229	247	226	202	246	238	280	239	182	395	215	257	274
1992	335	222	297	262	173	247	197	145	241	243	260	280	234
1993	263	293	343	287	385	405	430	324	385	374	334	339	498
1994	207	273	371	210	397	340	315	235	345	338	321	336	185
1995	144	163	189	276	189	156	176	273	222	154	227	160	167
1996	304	287	231	182	331	199	273	156	225	254	214	186	237
1997	245	322	353	217	337	207	256	207	275	284	221	253	293
1998	294	213	345	174	251	234	279	247	378	288	268	225	179
1999	226	247	162	239	194	200	286	460	461	215	234	198	217
2000	240	355	253	248	217	203	192	184	326	179	263	222	197
2001	195	165	233	254	244	306	195	284	219	225	202	195	233
2002	210	261	260	172	224	224	187	235	247	295	258	253	265
2003	222	312	299	282	284	326	335	266	332	345	404	269	275
2004	180	204	265	192	224	184	202	392	271	318	336	271	304
2005	239	319	425	230	879	174	450	233	334	433	434	527	211
2006	388	244	248	225	308	317	348	161	298	193	228	251	153
2007	257	189	239	313	261	367	208	167	202	189	175	180	282
2008	902	210	226	391	263	341	436	178	248	204	180	166	154
2009	155	254	205	197	187	160	166	332	285	360	444	296	149
2010	169	158	324	181	224	161	264	218	218	303	244	248	696
2011	200	160	235	230	184	274	184	274	207	227	236	197	222
2012	256	224	240	152	226	156	161	237	261	165	162	215	169
2013	187	344	639	149	263	192	329	499	416	381	258	230	157
2014	218	208	239	208	274	184	207	199	305	419	338	288	233
2015	186	190	220	186	246	249	216	268	194	212	193	230	200
2016	320	187	175	171	228	190	204	259	238	292	244	269	223
2017	241	176	393	206	262	222	216	248	253	203	234	177	158
2018	178	215	305	171	308	188	202	331	260	248	288	279	249
2019	342	336	317	321	291	241	438	213	200	247	214	281	244
2020	320	193	212	243	207	208	185	143	226	155	195	141	277
2021	374	233	342	295	389	307	330	217	323	341	253	272	236
2022	393	185	400	186	346	235	255	242	352	391	271	242	165
2023	313	214	367	299	291	269	238	267	287	306	457	313	246
2024	267	199	203	327	244	235	151	232	228	212	294	217	423

B Incertitude sur les niveaux de retour de référence utilisés pour l'application aux données réelles

TABLEAU 14 – Niveaux de retour marginaux de référence, erreurs standards et intervalles de confiance à 95 %

Station	T	$\hat{z}_{T,\text{ref}}$	Erreur standard	IC _{95%} inf.	IC _{95%} sup.
235	10	382.8	36.3	311.6	453.9
240	10	315.0	17.4	280.8	349.2
260	10	394.1	30.5	334.3	454.0
270	10	315.2	16.8	282.2	348.1
275	10	398.9	38.1	324.2	473.7
280	10	319.0	23.0	274.0	364.0
290	10	364.5	31.4	302.9	426.1
310	10	339.5	24.5	291.5	387.5
344	10	363.3	21.0	322.1	404.5
350	10	374.3	25.1	325.1	423.4
370	10	351.0	29.6	293.1	408.9
375	10	337.9	19.7	299.2	376.5
380	10	368.7	47.3	276.0	461.3
235	50	566.4	105.6	359.4	773.4
240	50	382.6	45.5	293.3	471.9
260	50	533.5	81.3	374.2	692.9
270	50	381.2	39.7	303.5	458.9
275	50	596.4	122.0	357.2	835.5
280	50	419.5	66.1	290.1	549.0
290	50	509.1	94.5	323.9	694.4
310	50	449.8	60.3	331.7	567.9
344	50	449.0	49.3	352.3	545.7
350	50	472.0	67.9	338.9	605.1
370	50	495.3	90.8	317.4	673.3
375	50	424.2	41.1	343.7	504.7
380	50	623.2	192.6	245.8	1000.6
235	100	666.2	157.6	357.4	975.0
240	100	409.9	63.5	285.5	534.2
260	100	600.9	116.8	372.0	829.8
270	100	407.1	54.1	301.2	513.1
275	100	709.6	188.1	340.8	1078.3
280	100	467.5	96.4	278.6	656.4
290	100	582.6	140.7	306.9	858.4
310	100	501.0	84.4	335.6	666.3
344	100	484.5	67.5	352.3	616.8
350	100	512.1	95.5	325.0	699.3
370	100	573.4	137.2	304.5	842.3
375	100	461.0	54.5	354.2	567.8
380	100	787.1	321.5	156.9	1417.3

C Modèles sélectionnés pour les méthodes 1b et 2 lors du procédé *leave-one-station-out* appliqué aux données réelles

TABLEAU 15 – Modèles de dépendance sélectionnés lors des différentes itérations de la procédure *leave-one-station-out* pour la méthode 1b

Station retirée	Modèle sélectionné	β_0	β_h	$\beta_{\cos(2\phi)}$
235	$\psi \sim h$	-2.985	0.074	–
240	$\psi \sim h + \cos(2\phi)$	-2.514	0.073	2.946
260	$\psi \sim h + \cos(2\phi)$	-2.905	0.073	2.541
270	$\psi \sim h + \cos(2\phi)$	-3.322	0.073	2.030
275	$\psi \sim h + \cos(2\phi)$	-2.433	0.069	3.279
280	$\psi \sim h + \cos(2\phi)$	-4.944	0.098	3.237
290	$\psi \sim h + \cos(2\phi)$	-2.217	0.073	2.129
310	$\psi \sim h + \cos(2\phi)$	-2.662	0.066	3.461
344	$\psi \sim h + \cos(2\phi)$	-2.919	0.071	2.249
350	$\psi \sim h + \cos(2\phi)$	-2.203	0.068	2.491
370	$\psi \sim h + \cos(2\phi)$	-2.508	0.072	2.703
375	$\psi \sim h + \cos(2\phi)$	-2.371	0.072	2.807
380	$\psi \sim h + \cos(2\phi)$	-4.146	0.085	3.790

Note : les covariables Δalt et $\sin(2\phi)$ ne sont sélectionnées dans aucune des treize itérations et ne sont donc pas représentées dans le tableau. Le modèle complet utilisé pour la sélection contient h , Δalt , $\cos(2\phi)$ et $\sin(2\phi)$.

TABLEAU 16 – Structures de GEV spatiale sélectionnées lors des différentes itérations de la procédure *leave-one-station-out* pour la méthode 2

Station retirée	Position μ	Échelle σ	Forme ξ (fixée)
235	$\sim alt + d_{\text{coast}}$	~ 1	~ 1
240	$\sim alt + d_{\text{coast}}$	~ 1	~ 1
260	$\sim alt + d_{\text{coast}}$	~ 1	~ 1
270	$\sim alt + d_{\text{coast}}$	~ 1	~ 1
275	$\sim alt + d_{\text{coast}}$	~ 1	~ 1
280	$\sim alt + d_{\text{coast}}$	~ 1	~ 1
290	$\sim alt + d_{\text{coast}}$	~ 1	~ 1
310	$\sim lon + alt + d_{\text{coast}}$	~ 1	~ 1
344	$\sim alt + d_{\text{coast}}$	~ 1	~ 1
350	$\sim alt + d_{\text{coast}}$	~ 1	~ 1
370	$\sim alt + d_{\text{coast}}$	~ 1	~ 1
375	$\sim lat + lon + alt + d_{\text{coast}}$	~ 1	~ 1
380	$\sim d_{\text{coast}}$	~ 1	~ 1

Note : la structure du paramètre de forme ξ est fixée à ~ 1 dans la définition de la méthode 2 et ne fait donc pas l'objet de la procédure de sélection. Seules les structures des paramètres de position μ et d'échelle σ sont sélectionnées parmi les modèles candidats.

