

Caractérisation et évolution des précipitations extrêmes horaires en France à partir d'un modèle régional de climat à convection profonde résolue

Nicolas Decoopman
UGA M2 SSD

Juliette Blanchet
CNRS, IGE

Antoine Blanc
RTM

Résumé

Le changement climatique provoque un réchauffement global (+1,1°C), plus marqué en France métropolitaine (+1,7°C) et dans les Alpes françaises (+2°C) depuis l'ère préindustrielle. L'air plus chaud contient davantage d'humidité, ce qui favorise théoriquement l'augmentation des précipitations extrêmes, bien que cette tendance varie selon les régions et les circulations atmosphériques. Les modèles climatiques classiques (GCM et RCM) sont limités pour représenter les précipitations extrêmes à résolution infra-journalière en raison de leur faible résolution et de la paramétrisation de la convection. Les modèles à résolution kilométrique (CP-RCM), comme CNRM-AROME (2,5 km), permettent désormais une meilleure représentation explicite de la convection profonde. Le stage vise à analyser les tendances des précipitations extrêmes horaires en France (1959-2022) grâce aux données CP-RCM et pluviométriques, en appliquant des modèles de valeurs extrêmes (GEV) stationnaires et non stationnaires.

Mots-clés: Changement climatique, Précipitations extrêmes, Clausius-Clapeyron, Convection profonde, Modèles climatiques régionaux (RCM), Modèles de climat à résolution kilométrique (CP-RCM), CNRM-AROME, Théorie des valeurs extrêmes (GEV), Tendances non stationnaires.

1. Introduction et contexte

Le changement climatique entraîne un réchauffement de l'air à la surface de la planète, plus marqué sur les continents que sur les océans [9]. L'augmentation est de $+1^{\circ}\text{C}$ à l'échelle mondiale, $+1,7^{\circ}\text{C}$ à l'échelle de la France métropolitaine et $+2^{\circ}\text{C}$ à l'échelle des Alpes françaises depuis l'ère préindustrielle. Par ailleurs, la relation de Clausius-Clapeyron montre que l'air chaud contient plus d'humidité ($+7\%/^{\circ}\text{C}$) [5]. L'ascension de l'air chaud dans l'atmosphère entraîne son refroidissement adiabatique, provoquant la condensation de la vapeur d'eau qui se transforme en précipitations, telles que la pluie [11]. Ainsi, en réponse au réchauffement climatique il existe une augmentation *théorique* des précipitations. Cette augmentation est variable suivant les changements de circulations atmosphériques [1].

Dans l'objectif de dégager des tendances durables (changement climatique) en opposition aux variations naturelles (variabilité climatique), il convient d'utiliser des données issues de modèles de climat sur de longues périodes. Ces dernières années, l'utilisation des modèles régionaux de climat à haute résolution (CP-RCM pour Convection-Permitting Regional Climate Models, avec une résolution de 1 à 3 km) a marqué une avancée significative. Ces modèles simulent explicitement la convection profonde, ce qui leur permet de capturer des échelles spatio-temporelles plus fines. Cette approche ouvre de nouvelles perspectives prometteuses pour l'analyse climatique [14].

L'objectif de cette étude est de caractériser et d'étudier les tendances de précipitations extrêmes horaires en France. Les précipitations extrêmes sont généralement définies de façon statistique, non par un seuil fixe, mais par leur caractère exceptionnel au regard de la climatologie locale. Une méthode courante consiste à considérer les maxima pour une période et un pas de temps donnés mesurés par les réseaux de pluviomètres [2].

2. Méthodologie

2.1. Données utilisées

Dans cette étude, le modèle numérique Application of Research to Operations at MEscale (AROME) a été choisi [4]. C'est un modèle CP-RCM de résolution spatiale 2,5km et temporelle 1h, forcé par réanalyse ERA5 [8] offrant des données de précipitations de 1959 à 2022. En parallèle, l'étude utilise aussi les données de précipitations issues d'observation Météo-France [12] au pas de temps journalier (1959-2022) et horaire (1990-2022). L'ensemble des analyses qui suivent se font indépendamment entre les données modélisées par AROME et les données observées par les stations.

2.2. Statistiques descriptives

Par extraction, transformation et chargement, les statistiques descriptives suivantes sont générées : nombre de jour de pluie (seuil fixé à 1mm/j), cumul de précipitations, maximum de précipitations. Ceci en chaque point de grille AROME et chaque station Météo-France à partir des données journalières de 1959 à 2022 et 1990 à 2022 ; et des données horaires de 1990 à 2022, pour chaque année et chaque saison.

Les saisons sont définies par **SON** pour septembre (**SEP**) octobre (**OCT**) novembre (**NOV**), **DJF** pour décembre (**DEC**) janvier (**JAN**) février (**FEV**), **MAM** pour mars (**MAR**) avril (**AVR**) mai (**MAI**), et **JJA** pour juin (**JUI**) juillet (**JUILL**) août (**AOU**). L'année hydrologique (**HYDRO**) est définie comme la période allant du 1er septembre de l'année N au 31 août de l'année N+1.

2.3. Modélisation statistique

Définitions

Si on note x une réalisation de la variable aléatoire X , représentant le maximum annuel de précipitations en un point spatial donné, alors la loi GEV (loi des valeurs extrêmes généralisée) est une loi de probabilité continue paramétrée par le triplet $\theta = (\mu, \sigma, \xi)$ — respectivement la position, l'échelle (strictement positive) et la forme dont la fonction de répartition cumulative (CDF) de la GEV est définie comme suit :

$$F(x; \mu, \sigma, \xi) = \exp \left\{ - \left[1 + \xi \left(\frac{x - \mu}{\sigma} \right) \right]^{-\frac{1}{\xi}} \right\}$$

Covariable temporelle

On dispose d'une série temporelle de n maxima annuels indépendants de précipitations pour un point géographique. Ces observations sont notées $\{x_1, x_2, \dots, x_n\}$ où chaque x_i est un maximum annuel de précipitations observé à l'année t_i et qui suit une loi GEV de paramètre θ . On transforme l'année t_i en une covariable normalisée notée \tilde{t}_i . Cette transformation est simplement réalisée pour des raisons numériques mais elle ne change rien au résultat théorique.

$$\tilde{t}_i = \frac{t_i - t_{\min}}{t_{\max} - t_{\min}} \quad \text{avec} \quad \begin{cases} t_{\min} = \min_i t_i \\ t_{\max} = \max_i t_i \end{cases}$$

On crée également une covariable temporelle avec point de rupture noté t_+ tel que :

$$\tilde{t}_i^* = \begin{cases} 0 & \text{si } t_i < t_+ \\ \frac{t_i - t_+}{t_{\max} - t_+} & \text{si } t_i \geq t_+ \end{cases}$$

Ce codage permet d'appliquer une pente temporelle seulement après la date de rupture, avec une covariable encore normalisée sur $[0, 1]$ dans la portion post-rupture.

Modèles utilisés

Soit la covariable temporelle $t \in \mathbb{N} \mid t_{\min} \leq t \leq t_{\max}$. Le **modèle stationnaire** est défini par :

$$M_0(\theta_0) \quad \text{et} \quad \theta_0 = (\mu_0, \sigma_0, \xi_0) \quad \begin{cases} \mu(t) = \mu_0 \\ \sigma(t) = \sigma_0 \\ \xi(t) = \xi_0 \end{cases}$$

Les **modèles non stationnaires** sont définis par :

$$M_1(\theta_1) \quad \text{et} \quad \theta_1 = (\mu_0, \mu_1, \sigma_0, \xi_0) \quad \begin{cases} \mu(t) = \mu_0 + \mu_1 \cdot t \\ \sigma(t) = \sigma_0 \\ \xi(t) = \xi_0 \end{cases}$$

$$M_2(\theta_2) \quad \text{et} \quad \theta_2 = (\mu_0, \sigma_0, \sigma_1, \xi_0) \quad \begin{cases} \mu(t) = \mu_0 \\ \sigma(t) = \sigma_0 + \sigma_1 \cdot t \\ \xi(t) = \xi_0 \end{cases}$$

$$M_3(\theta_3) \quad \text{et} \quad \theta_3 = (\mu_0, \mu_1, \sigma_0, \sigma_1, \xi_0) \quad \begin{cases} \mu(t) = \mu_0 + \mu_1 \cdot t \\ \sigma(t) = \sigma_0 + \sigma_1 \cdot t \\ \xi(t) = \xi_0 \end{cases}$$

Lorsqu'un point de rupture noté t_+ est introduit, on note :

$$t^* = t \cdot \mathbb{1}_{t > t_+} \quad \text{avec} \quad t_+ \in \mathbb{N}$$

Les modèles M_1 , M_2 et M_3 deviennent respectivement M_1^* , M_2^* et M_3^* . Sur cete même notation θ_i devient θ_i^* avec $i \in \{1, 2, 3\}$. Dans cette étude, on réalise les modélisations stationnaire et non-stationnaires avec pour covariable l'année et un effet temporel sur μ ou σ ou μ et σ . ξ est choisi comme constant. Sur la base bibliographique, on choisi $t_+ = 1985$ [2].

Niveau de retour

Le niveau de retour (ou quantile d'ordre $1 - \frac{1}{T}$) dans une loi GEV correspond à une valeur seuil z_T que l'on dépasse, en moyenne, une fois tous les T ans. Soit $X \sim \text{GEV}(\mu, \sigma, \xi)$, alors en notant F^{-1} la fonction quantile de la GEV, on obtient :

$$\mathbb{P}(X > z_T) = \frac{1}{T}, \quad \text{soit} \quad z_T = F^{-1}\left(1 - \frac{1}{T}\right) = \begin{cases} \mu + \frac{\sigma}{\xi} \left[(-\log(1 - \frac{1}{T}))^{-\xi} - 1\right] & \text{si } \xi \neq 0 \\ \mu - \sigma \log(-\log(1 - \frac{1}{T})) & \text{si } \xi = 0 \end{cases} \quad (\text{Gumbel})$$

Vraisemblance et maximum de vraisemblance

Soit la fonction de vraisemblance $\mathcal{L}(\theta; x) : \theta \mapsto f(x; \theta)$. La log-vraisemblance $\ell(\theta) = \log \mathcal{L}(\theta)$ s'écrit après développement (Annexes 1.1) :

$$\ell(\theta) = - \sum_{i=1}^n \left[\log \sigma + \left(1 + \frac{1}{\xi}\right) \log z_i + z_i^{-\frac{1}{\xi}} \right] \quad \text{avec} \quad z_i(\theta) = 1 + \xi \frac{x_i - \mu}{\sigma} \quad (1)$$

On obtient alors :

$$\begin{aligned} \ell_{M_0}(\mu_0, \sigma_0, \xi_0) &= - \sum_{i=1}^n \left[\log \sigma_0 + \left(1 + \frac{1}{\xi_0}\right) \log \left(1 + \xi_0 \frac{x_i - \mu_0}{\sigma_0}\right) + \left(1 + \xi_0 \frac{x_i - \mu_0}{\sigma_0}\right)^{-1/\xi_0} \right] \\ \ell_{M_1}(\mu_0, \mu_1, \sigma_0, \xi_0) &= - \sum_{i=1}^n \left[\log \sigma_0 + \left(1 + \frac{1}{\xi_0}\right) \log \left(1 + \xi_0 \frac{x_i - (\mu_0 + \mu_1 \cdot \tilde{t}_i)}{\sigma_0}\right) + \left(1 + \xi_0 \frac{x_i - (\mu_0 + \mu_1 \cdot \tilde{t}_i)}{\sigma_0}\right)^{-1/\xi_0} \right] \\ \ell_{M_2}(\mu_0, \sigma_0, \sigma_1, \xi_0) &= - \sum_{i=1}^n \left[\log(\sigma_0 + \sigma_1 \tilde{t}_i) + \left(1 + \frac{1}{\xi_0}\right) \log \left(1 + \xi_0 \frac{x_i - \mu_0}{\sigma_0 + \sigma_1 \tilde{t}_i}\right) + \left(1 + \xi_0 \frac{x_i - \mu_0}{\sigma_0 + \sigma_1 \tilde{t}_i}\right)^{-1/\xi_0} \right] \\ \ell_{M_3}(\mu_0, \mu_1, \sigma_0, \sigma_1, \xi_0) &= - \sum_{i=1}^n \left[\log(\sigma_0 + \sigma_1 \tilde{t}_i) + \left(1 + \frac{1}{\xi_0}\right) \log \left(1 + \xi_0 \frac{x_i - (\mu_0 + \mu_1 \tilde{t}_i)}{\sigma_0 + \sigma_1 \tilde{t}_i}\right) + \left(1 + \xi_0 \frac{x_i - (\mu_0 + \mu_1 \tilde{t}_i)}{\sigma_0 + \sigma_1 \tilde{t}_i}\right)^{-1/\xi_0} \right] \end{aligned} \quad (1')$$

Les vraisemblances de M_1^*, M_2^*, M_3^* sont obtenues en remplaçant \tilde{t}_i par \tilde{t}_i^* dans les expressions ci-dessus.

En pratique, les paramètres (μ, σ, ξ) sont inconnus et estimés à partir des données par un estimateur $\hat{\theta} = (\hat{\mu}, \hat{\sigma}, \hat{\xi})$ obtenu par maximum de vraisemblance (MLE) via une optimisation numérique tel que $\hat{\theta} = \arg \max_{\theta} \ell(\theta)$. Il n'existe pas de formule explicite des paramètres.

L'estimateur du niveau de retour \hat{z}_T s'écrit alors $\hat{z}_T = F_{\hat{\theta}}^{-1}(1 - \frac{1}{T})$. Le MLE classique donne un point estimé, mais pas d'intervalle. On souhaite aussi connaître l'incertitude autour de l'estimation de \hat{z}_T . Pour cela, on utilise la vraisemblance profilée.

Vraisemblance profilée et intervalle de confiance

z_T peut se réécrire sous la forme $\mu = z_T - \frac{\sigma}{\xi} \left[(-\log(1 - \frac{1}{T}))^{-\xi} - 1 \right]$. La combinaison des paramètres temporels de la loi GEV conduit à une expression linéaire en t :

$$z_T(t) = z_{T,0} + z_{T,1} \cdot t$$

En développant les paramètres soumis à un effet temporel et on regroupant terme à terme on peut montrer que (Annexes 1.2) :

$$\begin{aligned} \mu_1(z_{T,1}) &= z_{T,1} - \frac{\hat{\sigma}_1}{\hat{\xi}_0} \left([-\log(1 - \frac{1}{T})]^{-\hat{\xi}_0} - 1 \right) \\ \sigma_1(z_{T,1}) &= \frac{\hat{\xi}_0 (z_{T,1} - \hat{\mu}_1)}{[-\log(1 - \frac{1}{T})]^{-\hat{\xi}_0} - 1} \end{aligned}$$

On cherche l'intervalle de confiance sur $z_{T,1}$ donc pour chaque valeur candidate $z_{T,1}$ dans une grille (autour de l'estimateur $\hat{z}_{T,1}$), on maximise les log-vraisemblances (1') qui deviennent des log-vraisemblances profilées ℓ^p :

$$\begin{aligned}
\ell_{M_1}^p(z_{T,1}; \hat{\mu}_0, \hat{\sigma}_0, \hat{\xi}_0) &= - \sum_{i=1}^n \left[\log \hat{\sigma}_0 + \left(1 + \frac{1}{\hat{\xi}_0}\right) \log \left(1 + \hat{\xi}_0 \frac{x_i - (\hat{\mu}_0 + \mu_1(z_{T,1}) \cdot \tilde{t}_i)}{\hat{\sigma}_0}\right) + \left(1 + \hat{\xi}_0 \frac{x_i - (\hat{\mu}_0 + \mu_1(z_{T,1}) \cdot \tilde{t}_i)}{\hat{\sigma}_0}\right)^{-1/\hat{\xi}_0} \right] \\
\ell_{M_2}^p(z_{T,1}; \hat{\mu}_0, \hat{\sigma}_0, \hat{\xi}_0) &= - \sum_{i=1}^n \left[\log(\hat{\sigma}_0 + \sigma_1(z_{T,1}) \cdot \tilde{t}_i) + \left(1 + \frac{1}{\hat{\xi}_0}\right) \log \left(1 + \hat{\xi}_0 \frac{x_i - \hat{\mu}_0}{\hat{\sigma}_0 + \sigma_1(z_{T,1}) \cdot \tilde{t}_i}\right) + \left(1 + \hat{\xi}_0 \frac{x_i - \hat{\mu}_0}{\hat{\sigma}_0 + \sigma_1(z_{T,1}) \cdot \tilde{t}_i}\right)^{-1/\hat{\xi}_0} \right] \\
\ell_{M_3}^p(z_{T,1}; \hat{\mu}_0, \hat{\sigma}_0, \hat{\sigma}_1, \hat{\xi}_0) &= - \sum_{i=1}^n \left[\log(\hat{\sigma}_0 + \hat{\sigma}_1 \tilde{t}_i) + \left(1 + \frac{1}{\hat{\xi}_0}\right) \log \left(1 + \hat{\xi}_0 \frac{x_i - (\hat{\mu}_0 + \mu_1(z_{T,1}) \tilde{t}_i)}{\hat{\sigma}_0 + \hat{\sigma}_1 \tilde{t}_i}\right) + \left(1 + \hat{\xi}_0 \frac{x_i - (\hat{\mu}_0 + \mu_1(z_{T,1}) \tilde{t}_i)}{\hat{\sigma}_0 + \hat{\sigma}_1 \tilde{t}_i}\right)^{-1/\hat{\xi}_0} \right]
\end{aligned} \tag{2}$$

On cherche donc :

$$\hat{z}_{T,1} = \arg \max_{z_{T,1}} \ell_{M_\bullet}^p(z_{T,1}; \hat{\theta}_\bullet) \quad \text{avec} \quad \hat{\theta}_\bullet = \begin{cases} \hat{\theta}_1^p = (\hat{\mu}_0, \hat{\sigma}_0, \hat{\xi}_0) & \text{pour } M_1 \\ \hat{\theta}_2^p = (\hat{\mu}_0, \hat{\sigma}_0, \hat{\xi}_0) & \text{pour } M_2 \\ \hat{\theta}_3^p = (\hat{\mu}_0, \hat{\sigma}_0, \hat{\sigma}_1, \hat{\xi}_0) & \text{pour } M_3 \end{cases}$$

On trace ainsi pour chaque modèle M_\bullet la fonction $\mathcal{L}_{M_\bullet} : z_{T,1} \mapsto \ell_{M_\bullet}^p(z_{T,1}; \hat{\theta}_\bullet)$

L'intervalle de confiance de $\hat{z}_{T,1}$ pour un modèle M_\bullet au seuil $(1 - \alpha)$ basé sur le profil de vraisemblance est donné par :

$$\text{IC}_{M_\bullet}^{(1-\alpha)}(\hat{z}_{T,1}) = \left\{ z_{T,1} : 2[\ell_{M_\bullet}^p(\hat{z}_{T,1}; \hat{\theta}_\bullet) - \ell_{M_\bullet}^p(z_{T,1}; \hat{\theta}_\bullet)] \leq \chi_{1,1-\alpha}^2 \right\}$$

où $\chi_{1,1-\alpha}^2$ est le quantile d'ordre $1 - \alpha$ d'une loi du χ^2 à un degré de liberté. On fixe ici $\alpha = 0,10$.

2.4. Choix du meilleur modèle

Test du rapport de vraisemblance (LRT)

En tout point géographique on dispose de $M_0, M_1, M_2, M_3, M_1^*, M_2^*$ et M_3^* . Notons k_j le nombre de paramètres du modèle M_j . Pour chaque modèle non stationnaire $j \neq 0$ et chaque point i :

$$\Lambda_{ij} = 2(\ell_{ij} - \ell_{i0}) \stackrel{H_0}{\sim} \chi_{k_j - k_0}^2.$$

La p -valeur est $p_{ij} = \mathbb{P}(\chi_{k_j - k_0}^2 \geq \Lambda_{ij})$.

Règle hiérarchique de sélection

Soit $\alpha = 0,10$ le seuil d'intérêt. Si un des deux modèles M_3 ou M_3^* vérifie $p_{ij} \leq \alpha$, on retient $j = \arg \min_{j \in \{3, 3^*\}} p_{ij}$. Sinon, on compare l'ensemble des six modèles non stationnaires et l'on sélectionne $j = \arg \min_{j \in \{1, 1^*, 2, 2^*, 3, 3^*\}} p_{ij}$. Cette double étape privilégie les formes *simultanément* temporelles sur μ et σ quand elles sont statistiquement justifiées. Cela assure

que la complexité n'est introduite que lorsqu'elle apporte une information statistiquement crédible tout en livrant, pour chaque poste, un modèle non stationnaire.

2.5. Calcul des tendances

A partir des niveaux de retour 10 ans ($T = 10$) la tendance relative (en %) est calculée via :

$$\text{Tendance} = \frac{z_T^{2022} - z_T^{1985}}{z_T^{1985}} \cdot 100$$

Concordance

On évalue l'accord entre les statistiques descriptives et les tendances obtenues à partir des simulations du modèle AROME et celles observées dans la réalité. Pour ce faire, chaque station Météo-France est associée au point de grille AROME (2,5 km × 2,5 km) correspondant à sa localisation géographique. Cette correspondance permet de calculer la corrélation de Pearson (r) ainsi que l'erreur moyenne (ME) entre les valeurs observées et simulées.

3. Résultats

3.1. Évaluation de la climatologie des précipitations simulées par AROME

Afin d'apprécier la capacité d'AROME à restituer le climat pluviométrique les données sont confrontées aux observations de stations suivant : le nombre de jours de pluie (seuil 1 mm/j), le cumul saisonnier des précipitations et la moyenne des maxima saisonnière des précipitations.

Distribution spatiale

Sur l'ensemble de ces critères pour des données journalières de 1959 à 2022, AROME restituent très bien la distribution spatiale de la réalité (Figure 1). On peut voir trois grands régimes spatiaux liés aux influences océaniques, méditerranéennes et orographiques.

Nombre de jours de précipitations annuels

Les massifs montagneux (Alpes, Pyrénées, Massif central, Vosges, Jura) affichent les plus fortes fréquences de jours pluvieux (au-delà de 140–160 jours par an). Le grand Ouest atlantique (Bretagne, Normandie, Pays de la Loire) connaît également un nombre élevé de journées pluvieuses (80–120 jours/an). À l'inverse, la façade méditerranéenne et le pourtour de la Provence sont les plus secs en fréquence, avec souvent moins de 50–70 jours de précipitations par an.

Cumul moyen des précipitations

La bordure atlantique sud-ouest (Pyrénées, Aquitaine) et les Alpes du Nord reçoivent les plus forts cumuls (plus de 5 mm/j). Le Massif central et les reliefs intérieurs (Vosges, Jura)

présentent des cumuls intermédiaires (2,5-4 mm/j). Le pourtour méditerranéen (Languedoc, Provence) reste globalement plus sec ($< 1,5$ mm/j), malgré des pluies intenses ponctuelles.

Moyenne des maxima de précipitations journalières

Les plus fortes pointes quotidiennes moyennes se rencontrent dans les Cévennes et plus généralement sur la face sud-est du Massif central (environ 100–125 mm/j). Les reliefs alpins et pyrénéens montrent aussi des maxima élevés (80–100 mm/j). La façade atlantique et le bassin parisien présentent des maxima plus modérés (30–60 mm/j), tandis que la Provence et la Côte d’Azur, malgré une fréquence moindre, peuvent localement connaître de très gros orages (40–80 mm/j en moyenne).

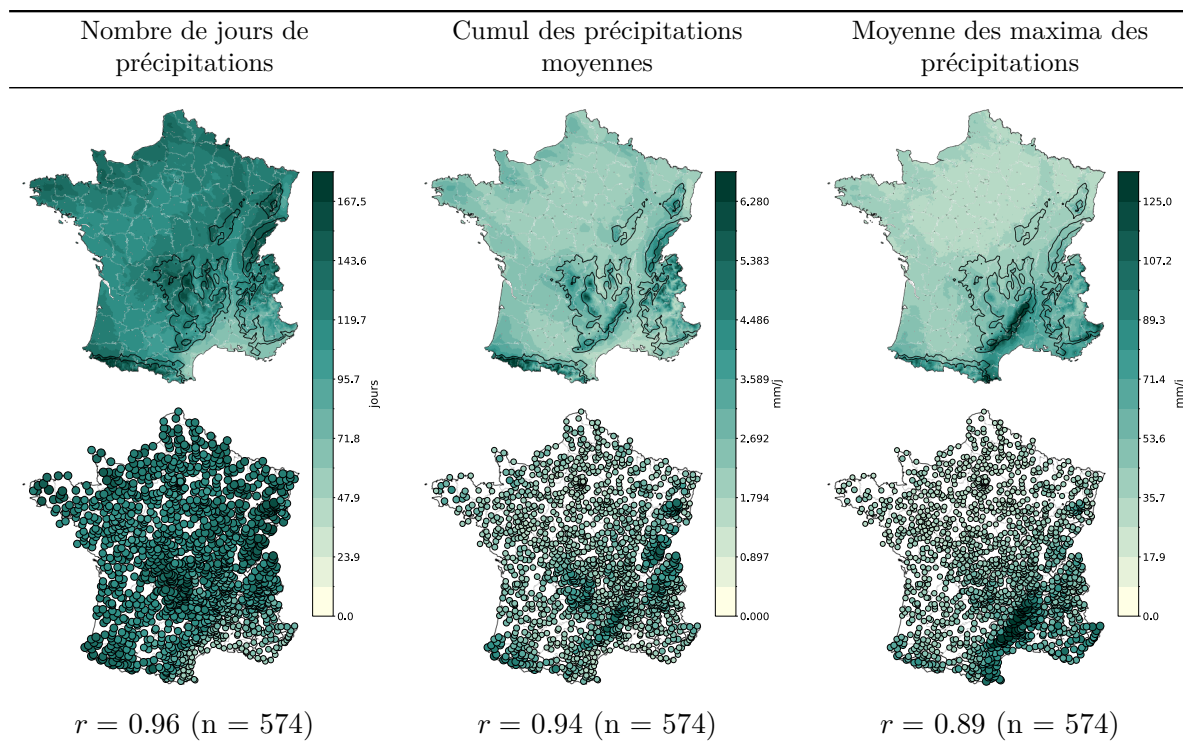


Figure 1: Climatologie et corrélation entre le modèle AROME et les stations Météo-France issues de données journalières allant de 1959 à 2022 pour une année hydrologique. Représentation en trait fin des courbes de niveaux 400 et 800m.

Corrélation entre le modèle AROME et les stations Météo-France (Figure 2)

Quelle que soit la climatologie analysée, le modèle **AROME** reproduit fidèlement les observations, avec une corrélation minimale de **0,70**. Indépendamment de l’échelle temporelle retenue — journalière (1959-2022 ou 1990-2022) ou horaire (1990-2022) — et de la saison, les champs simulés s’accordent très bien avec les données mesurées : la corrélation varie entre **0,92** et **0,98** pour le nombre de jours de pluie et le cumul des précipitations.

Cette performance se maintient pour la moyenne des maxima journaliers (périodes 1959-2022 et 1990-2022) sur l’année hydrologique, l’automne, l’hiver et le printemps, mais elle se dégrade

en été, avec une corrélation de **0,85**. À l'échelle horaire, la qualité de l'estimation des maxima se détériore encore : la corrélation baisse de 0,4 à 0,8 point selon la saison pour l'année hydrologique, l'automne et l'hiver, et chute à environ **0,70** au printemps et **0,69** en été où $ME = -3.753 \text{ mm/h}$ (-26.3%). AROME tend donc à sous-estimer les précipitations extrêmes estivales, ce qui reste vrai aux autres saisons.

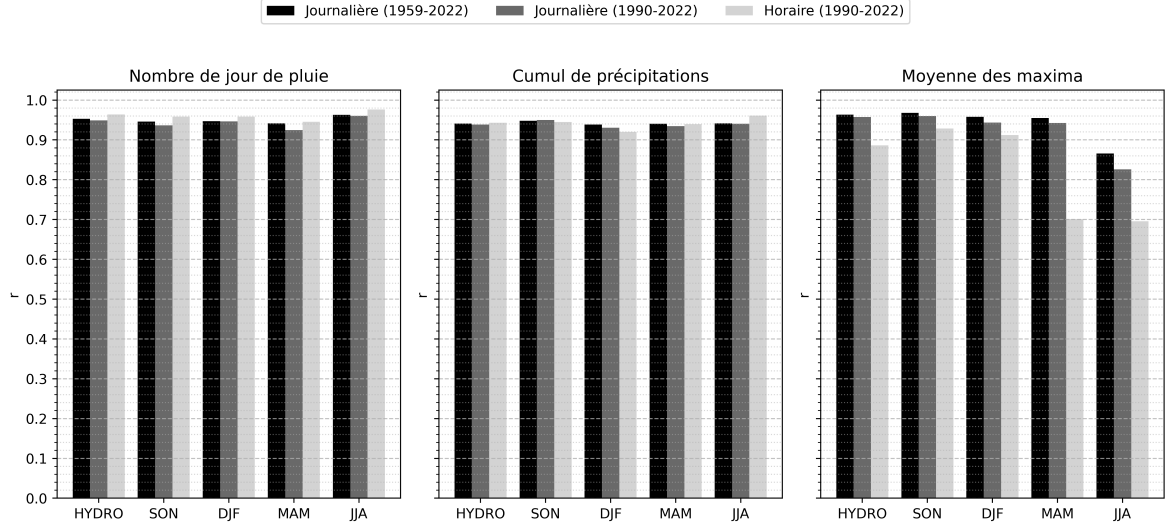


Figure 2: Corrélations des données climatologiques entre le modèle AROME et les stations Météo-France pour chacune des sources de données.

4. Discussion

Fidélité spatiale de la climatologie simulée

Les résultats confirment qu'AROME reproduit correctement les grands régimes pluviométriques hexagonaux confirmé par la littérature [7], [4], [6], [10]. Il y a un excédent orographique sur les Alpes, les Pyrénées et le Massif central, un gradient atlantique-continentale marqué à l'ouest et un déficit fréquentiel sur le pourtour méditerranéen. Cette cohérence avec la réalité mesurée témoigne d'une représentation satisfaisante des forçages dynamiques (transport d'humidité par les flux d'ouest, soulèvement orographique, circulation de basse couche en Méditerranée).

Variabilité saisonnière et représentation des extrêmes

La capacité d'AROME à restituer la fréquence et la quantité de précipitations se maintient tout au long de l'année, mais la performance chute pour la moyenne des maxima journaliers en été et davantage à l'échelle horaire. On retrouve le fait qu'AROME sous-estime des précipitations d'intensité élevées ($> 40 \text{ mm/h}$ [3]). La convection estivale reste partiellement sous-résolue malgré la résolution spatiale de 2,5 km. Dans cette étude, on a pu montrer (résultats non affichés) que la corrélation augmente lorsque la fenêtre temporelle s'agrandit à 6 ou 9h. Le modèle pourrait reproduire la cellule orageuse en démarrant trop tard ou trop tôt,

en étalant l'intensité sur plusieurs mailles ou en sous-estimant les précipitations maximales. Le pas de 2–3 km est une étape majeure pour représenter la convection sans paramétrage, mais il reste trop grossier pour certaines applications sensibles aux maxima intenses [13]. Il aurait été intéressant d'introduire les données COMEPHORE (1 km, 15 min) de Météo-France dans cette étude mais les réanalyses ne débutent qu'en 1997.

Remerciements

Je tiens à remercier Juliette Blanchet et Antoine Blanc pour l'encadrement rigoureux, stimulant et bienveillant tout au long de ce stage. Leurs conseils avisés, leur disponibilité constante et leurs nombreuses remarques constructives m'ont permis d'approfondir considérablement mes compétences scientifiques et méthodologiques. Leur implication a rendu ce stage particulièrement enrichissant et motivant. Je leur suis reconnaissant pour leur confiance et leur soutien tout au long de ce travail.

References

- [1] Blanchet J, Blanc A, Creutin JD (2021). “Explaining recent trends in extreme precipitation in the Southwestern Alps by changes in atmospheric influences.” *Weather and Climate Extremes*, **33**, 100356. doi:10.1016/j.wace.2021.100356.
- [2] Blanchet J, Creutin JD (2022). “Instrumental agreement and retrospective analysis of trends in precipitation extremes in the French Mediterranean Region.” *Environmental Research Letters*, **17**(7), 074011. doi:10.1088/1748-9326/ac7734.
- [3] Caillaud C, Somot S, Alias A, et al (2021). “Modelling Mediterranean heavy precipitation events at climate scale: an object-oriented evaluation of the CNRM-AROME convection-permitting regional climate model.” *Climate Dynamics*, **56**, 1717–1752. doi:10.1007/s00382-020-05558-y. URL <https://doi.org/10.1007/s00382-020-05558-y>.
- [4] Caillaud C, et al. (2021). “Simulation using CNRM-AROME46t1 (2.5km) CP-RCM performed by CNRM.” *Climate Dynamics*. doi:10.1007/s00382-020-05558-y. URL <https://doi.org/10.1007/s00382-020-05558-y>.
- [5] Clausius R (1850). “Ueber die bewegende Kraft der Wärme und die Gesetze, welche sich daraus für die Wärmelehre selbst ableiten lassen.” *Annalen der Physik*, **155**(3), 368–397. doi:doi.org/10.1002/andp.18501550302.
- [6] Dura V, Evin G, Favre AC, Penot D (2024). “Spatial variability in the seasonal precipitation lapse rates in complex topographical regions – application in France.” *Hydrology and Earth System Sciences*, **28**(12), 2579–2601. doi:10.5194/hess-28-2579-2024. URL <https://hess.copernicus.org/articles/28/2579/2024/>.
- [7] Fumière Q, Déqué M, Nuissier O, Somot S, Alias A, Caillaud C, Laurantin O, Seity Y (2020). “Extreme rainfall in Mediterranean France during the fall: added value of the CNRM-AROME Convection-Permitting Regional Climate Model.” *Climate Dynamics*, **55**, 77–91. doi:10.1007/s00382-019-04898-8.

- [8] Hersbach H, Bell B, Berrisford P, Hirahara S, Horányi A, Muñoz-Sabater J, Nicolas J, Peubey C, Radu R, Schepers D, Simmons A, Soci C, Abdalla S, Abellan X, Balsamo G, Bechtold P, Biavati G, Bidlot J, Bonavita M, De Chiara G, Dahlgren P, Dee D, Diamantakis M, Dragani R, Flemming J, Forbes R, Fuentes M, Geer A, Haimberger L, Healy S, Hogan RJ, Hólm E, Janisková M, Keeley S, Laloyaux P, Lopez P, Lupu C, Radnoti G, Rosnay P, Rozum I, Vamborg F, Villaume S, Thépaut JN (2020). “The ERA5 global reanalysis.” *Quarterly Journal of the Royal Meteorological Society*, **146**(730), 1999–2049. doi:10.1002/qj.3803.
- [9] Intergovernmental Panel on Climate Change (IPCC) (2021). *Climate Change 2021: The Physical Science Basis*. Cambridge University Press. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change.
- [10] Lucas-Picher P, Brisson E, Caillaud C, *et al.* (2024). “Evaluation of the convection-permitting regional climate model CNRM-AROME41t1 over Northwestern Europe.” *Climate Dynamics*, **62**, 4587–4615. doi:10.1007/s00382-022-06637-y. URL <https://doi.org/10.1007/s00382-022-06637-y>.
- [11] Météo-France (2010). *La météorologie*. 2e édition edition. Éditions Eyrolles, Paris, France. Chapitre 3 : Le cycle de l’eau et les précipitations.
- [12] Météo-France (2024). “Données d’observations météorologiques issues des stations synoptiques et climatologiques.” Open Data on data.gouv.fr. Consulté en juillet 2025, URL <https://www.data.gouv.fr/fr/datasets/donnees-dobservations-meteorologiques-issues-des-stations-synoptiques-et-climatologiques/>.
- [13] Prein A, Langhans W, Fossier G, Ferrone A, Ban N, Goergen K, Keller M, Tölle M, Gutjahr O, Feser F, Brisson E, Kollet S, Schmidli J, van Lipzig N, Leung R (2015). “A Review on Regional Convection-Permitting Climate Modeling: Demonstrations, Prospects, and Challenges.” *Reviews of Geophysics*, **53**(2), 323–361. doi:10.1002/2014RG000475. URL <https://doi.org/10.1002/2014RG000475>.
- [14] Prein AF, Langhans W, Fossier G, Ferrone A, Ban N, Goergen K, Keller M, Tolle M, Gutjahr O, Feser F, *et al.* (2015). “Regional climate modeling with convection-permitting resolutions: a new era in climate prediction?” *Journal of Climate*, **28**(2), 618–636. doi:10.1175/jcli-d-14-00309.1.

Annexes 1 : formules mathématiques

A.1.1. Obtention de (1)

Soit la fonction de vraisemblance $\mathcal{L}(\theta; x) : \theta \mapsto f(x; \theta)$. Alors : $\log \mathcal{L}(\theta; x_1, x_2, \dots, x_n) = \sum_{i=1}^n \log \mathcal{L}(\theta; x_i)$.

Pour $1 + \xi \frac{x - \mu}{\sigma} > 0$, avec $\sigma > 0$:

$$\begin{aligned} \log \mathcal{L}(\theta) &= \sum_{i=1}^n \left[-\log \sigma - \frac{1 + \xi}{\xi} \log \left(1 + \xi \frac{x_i - \mu}{\sigma} \right) - \left(1 + \xi \frac{x_i - \mu}{\sigma} \right)^{-\frac{1}{\xi}} \right] \\ \log \mathcal{L}(\theta) &= -n \log \sigma - \left(1 + \frac{1}{\xi} \right) \sum_{i=1}^n \log \left(1 + \xi \frac{x_i - \mu}{\sigma} \right) - \sum_{i=1}^n \left(1 + \xi \frac{x_i - \mu}{\sigma} \right)^{-\frac{1}{\xi}} \end{aligned}$$

La log-vraisemblance $\ell(\theta) = \log \mathcal{L}(\theta)$ s'écrit alors :

$$\ell(\theta) = - \sum_{i=1}^n \left[\log \sigma + \left(1 + \frac{1}{\xi} \right) \log \left(1 + \xi \frac{x_i - \mu}{\sigma} \right) + \left(1 + \xi \frac{x_i - \mu}{\sigma} \right)^{-\frac{1}{\xi}} \right] \quad (1)$$

A.1.2. Obtention des paramètres

$\mu_1(z_{T,1})$ et $\sigma_1(z_{T,1})$

En développant les paramètres soumis à un effet temporel, on a :

$$\begin{aligned} \mu_0 + \mu_1 t &= z_{T,0} + z_{T,1} t - \frac{\sigma_0 + \sigma_1 t}{\xi_0} \left[\left(-\log \left(1 - \frac{1}{T} \right) \right)^{-\xi_0} - 1 \right] \\ \mu_0 + \mu_1 t &= \left[z_{T,0} - \frac{\sigma_0}{\xi_0} \left(\left[-\log \left(1 - \frac{1}{T} \right) \right]^{-\xi_0} - 1 \right) \right] + \left[z_{T,1} - \frac{\sigma_1}{\xi_0} \left(\left[-\log \left(1 - \frac{1}{T} \right) \right]^{-\xi_0} - 1 \right) \right] t \end{aligned}$$

c'est-à-dire, terme à terme :

$$\begin{aligned} \mu_0 &= z_{T,0} - \frac{\sigma_0}{\xi_0} \left(\left[-\log \left(1 - \frac{1}{T} \right) \right]^{-\xi_0} - 1 \right), \\ \mu_1 &= z_{T,1} - \frac{\sigma_1}{\xi_0} \left(\left[-\log \left(1 - \frac{1}{T} \right) \right]^{-\xi_0} - 1 \right). \end{aligned}$$