Modeling moderate and extreme urban rainfall at high spatio-temporal resolution

Chloé Serre-Combe [1] & Nicolas Meyer [2] & Thomas Opitz [3] & Gwladys Toulemonde [4]

[1] *Univ. Montpellier, CNRS, IMAG, Inria, France, chloe.serre-combe@umontpellier.fr*
[2] *Univ. Montpellier, CNRS, IMAG, Inria, France, nicolas.meyer@umontpellier.fr*
[3] *INRAE, BioSP, Avignon, France, thomas.opitz@inrae.fr*
[4] *Univ. Montpellier, CNRS, IMAG, Inria, France, gwladys.toulemonde@umontpellier.fr*

Résumé. La modélisation des précipitations est d'un grand intérêt pour l'analyse des risques d'inondation. Nous proposons de modéliser la distribution des précipitations urbaines mesurées à haute résolution spatiale et temporelle par le réseau de pluviomètres de l'Observatoire Urbain de Montpellier sur quatre années de mesures. Nous les combinons avec des données de réanalyse radar pour étendre notre analyse à une période plus longue avec une résolution moins fine. Pour cette modélisation, nous considérons simultanément les pluies modérées et intenses en utilisant l'*Extended Generalized Pareto Distribution (EGPD)* pour éviter la sélection d'un seuil, souvent délicat en statistiques des extrêmes, mais aussi pour réduire la complexité de l'estimation des paramètres. Nous modélisons également la dépendance spatio-temporelle en incorporant l'advection à travers un processus spatio-temporel de Brown-Resnick. Nous utilisons des indices d'autocorrélation extrême, pour montrer sa variabilité entre les différents sites et à différents temps des mesures. Nous mettrons en évidence l'importance de considérer l'advection en comparant le modèle obtenu à un modèle séparable plus simple.

Mots-clés. Théorie des valeurs extrêmes, modélisation de précipitation, EGPD, haute résolution spatio-temporelle, modélisation de la dépendance, processus de Brown-Resnick, advection

Abstract. Precipitation modeling is of great interest for flood risk analysis. We propose to model the distribution of urban precipitation measured at high spatial and temporal resolution by the Montpellier Urban Observatory rain gauge network over four years of measurements. We combine them with radar reanalysis data to extend our analysis to a longer period with less fine resolution. For our modeling approach, we simultaneously consider moderate and intense rainfall by using the Extended Generalised Pareto Distribution (EGPD) to avoid explicit threshold selection, often tricky in extreme statistics, and to reduce the complexity of parameter estimation. We also model the spatio-temporal dependence by incorporating advection through a spatio-temporal Brown-Resnick process. We use indices of extreme autocorrelation, to show its variability between locations in relation to their spatial distances and to the temporality of the measurements. We will highlight the importance of including advection by comparing it with a simpler separable model.

Keywords. Extreme value theory, rainfall modeling, EGPD, high spatio-temporal resolution, dependence modeling, Brown-Resnick process, advection

1 Study

Managing flood risk requires a precise understanding of precipitation patterns. Severe flooding can indeed occur after extreme rainfall events, but also after moderate ones, for example if the latter last a long time or if the ground is already saturated with water. The risk is even greater in urban areas where water absorption is reduced by impervious surfaces. This shows the importance of understanding the behaviour of both extreme and moderate rainfall events and their spatio-temporal variability.

We focused our work on a specific area in Montpellier, south of France. Montpellier is known to be frequently exposed to significant rainfall events known as Mediterranean episodes, especially during the autumn season. This type of event is characterised by a large amount of rainfall in a short period of time and is very localised, causing local urban flooding. This encourages us to build a model which takes into account this high spatio-temporal resolution. Our study relies on 17 rain gauges in Montpellier in the water catchment of the Verdanson, a tributary of the Lez (see Figure 1). The rainfall measurements from these stations are provided by the urban Observatory of the HydroScience Montpellier (OHSM) (see Finaud-Guyot et al. (2023)). They cover the period from 2019 to 2022 and are given with a high temporal resolution, recorded minute by minute. In order to minimise measurement errors and to avoid very strongly discretized values, the data are aggregated in 5-minute intervals, which allows us to maintain a high temporal resolution. This fine resolution implies that we obtain a very small proportion of non-zero values: they represent only 1.2% of the data. In terms of spatial granularity, we also have a high resolution with a distance between two stations ranging from 77 to 1531 meters. We want to model the spatio-temporal characteristics of these rainfall data to better understand the rainfall behaviour over this area.

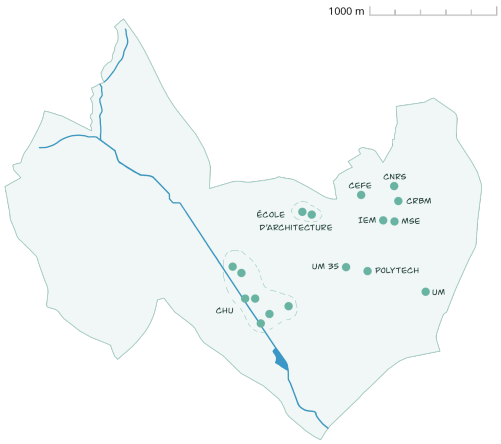


Figure 1: Rain gauges location

In order to refine our modeling, we consider another dataset: the French COMEPHORE[1] Mosaic from Météo France (see Tabary et al. (2012)). These reanalysis data, which combine radar and rain gauge measurements over France, have a lower resolution than the previous OHSM data. The rainfall measurements are provided with a spatial resolution of 1 km^2 per pixel. Each pixel represents the cumulative rainfall over one hour, and the time period ranges from 1997 to 2021. We extract measurements pixels on the Montpellier area.

Notations: Let $\mathcal{S} \subset \mathbb{R}^2$ be our spatial domain and let $\mathcal{T} \subset \mathbb{R}_+$ be our temporal domain with equidistant time points.

[1]COmbinaison en vue de la Meilleure Estimation de la Précipitation HOraiRE, dataset is available on AERIS platform (https://radarsmf.aeris-data.fr/)

2 Univariate modeling

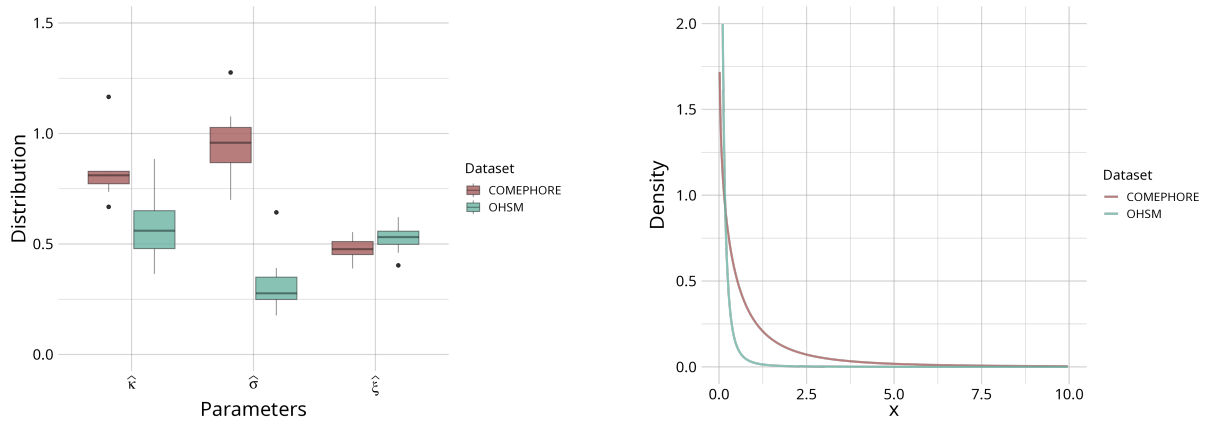
In this section, we focus on characterizing the behaviour of moderate and extreme precipitation at individual locations. In a univariate framework, let X_s be the random variable representing the amount of rainfall in millimeters measured at a given location $s \in \mathcal{S}$.

In order to take the full range of precipitation into account, we use the Extended Generalized Pareto Distribution (EGPD) introduced by Naveau et al. (2016). This distribution provides flexibility in both the bulk and the tail of the distribution, while having a tail behaviour similar to the Generalized Pareto Distribution (GPD) suggested by Extreme Value Theory. This family of distributions allows one to avoid explicit threshold selection and it is particularly efficient for precipitation modeling where we often need the full marginal distribution to model rainfall space-time aggregates leading to flood events. This distribution is a transformation of the classical GPD which corresponds to the limit distribution of threshold exceedances as the threshold u tends to infinity with shape parameter ξ and scale parameter σ_u. In a non-asymptotic setting, for a fixed threshold u, large enough, the corresponding cumulative distribution function is H_ξ and we have the approximation

$$\mathbb{P}\left(X_s - u > y | X_s > u\right) \approx \overline{H}_\xi\left(\frac{y}{\sigma_u}\right) = \begin{cases} \left(1 + \xi \frac{y}{\sigma_u}\right)_+^{-1/\xi} & \text{if } \xi \neq 0\,, \\ e^{-\frac{y}{\sigma_u}} & \text{if } \xi = 0\,, \end{cases}$$

with $a_+ = \max(a, 0)$. The cumulative distribution function of the EGPD is given by adding a transformation G to H_ξ. The efficient and simpler form is $G(x) = x^\kappa\,, \kappa > 0$. The parameter κ parameter controls the lower tail behaviour. To reduce the impact of very low values, a small left censoring is added locally to better fit this distribution to real rainfall data. This censoring was chosen according to the smaller normalized root mean square error at each location (see Haruna et al. (2023)).

For both datasets described above, over similar time periods and similar space locations, we obtain a good fit of the EGPD margins. We have similar estimated parameters for all sites within the same dataset. The boxplots of the estimated EGPD parameters for all the locations and the theoretical EGPD density with mean parameter estimates are given in Figure 2. We have close tail parameter estimates for the two data sets but the hourly one gives a smaller ξ estimate. Indeed, extreme events are smoothed within hourly data. Therefore, the tail of the distribution can appear heavier for OHSM data with a higher ξ parameter. Then, we have a higher shape parameter estimate for the COMEPHORE data. This can be explained by the fact that these rain events are spread over a longer period of time and require a larger scaling parameter to fit the distribution. Finally, there is an expected difference between the κ estimates, which is directly related to the time scale difference. In fact, a resolution of 5 minutes gives more low rainfall values for OHSM data, leading to a lower κ parameter.



(a) Parameter estimates

(b) Mean density estimation

Figure 2: EGPD fit for both datasets over a similar period of 4 years. The COMEPHORE data correspond to 21 pixels around the location of the OHSM rain gauges. (a) boxplots of the parameter estimates for each dataset; (b) theoretical EGPD density with the mean of parameter estimates for each dataset.

3 Spatio-temporal dependence

3.1 Spatio-temporal process for rainfall modeling

Now that we have modeled the marginal behaviour of precipitation in the univariate analysis, the next step is to consider the spatio-temporal dependence. To understand the dependence structure of rainfall events, several standard asymptotic models exist such as max-stable and Pareto processes developed in Davison et al. (2012), Ferreira and De Haan (2014) and Dombry and Ribatet (2015). It seems that the Brown-Resnick process, introduced in Kabluchko et al. (2009), is the most commonly used for spatial and spatio-temporal modeling, especially for precipitation, with its dependence structure flexibility defined in terms of a variogram function. It is a stationary max-stable class of random fields based on Gaussian processes with stronger regularity properties that make it more suitable for modeling extreme events. Thanks to its structure, there are nice links with classical geostatistical properties.

Let $X = \{X_{\boldsymbol{s},t} \mid (\boldsymbol{s},t) \in \mathcal{S} \times \mathcal{T}\}$ be a strictly stationary isotropic Brown-Resnick process representing our spatio-temporal amounts of rainfall. Then, as defined in Buhl and Klüppelberg (2018), we have the following representation

$$X_{\boldsymbol{s},t} = \bigvee_{j=1}^{\infty} \xi_j e^{W^j_{\boldsymbol{s},t} - \gamma(\boldsymbol{s},t)}$$

where ξ_j is a point of a Poisson process with intensity $\xi^{-2}d\xi$. The processes W^j are independent replicates of $W = \{W_{\boldsymbol{s},t} \mid (\boldsymbol{s},t) \in \mathcal{S} \times \mathcal{T}\}$, an underlying Gaussian spatio-temporal process with stationary increments and with a semivariogram γ defined by

$$2\gamma(\boldsymbol{s},t) = Var\left(W_{\boldsymbol{s},t} - W_{\mathbf{0},0}\right).$$

The semivariogram γ can be used to determine the one of X. In fact, the Brown-Resnick process X is defined by the underlying Gaussian process W, so that the spatio-temporal dependence structure of X is directly determined by the one of W. So our goal is to find a way to model this measure of dependence in order to get information about the overall distributional dependence of X. In order to do this, we look at the space-time extremogram, which will be expressed as a function of the variogram. This measure gives information about the extreme spatio-temporal dependence of a spatio-temporal process by looking at the simultaneous excesses for each pair of locations separated by a spatial lag at two different times separated by a temporal lag. So, let $\Lambda_{\mathcal{S}} \subset \mathbb{R}_+^2$ and $\Lambda_{\mathcal{T}} \subset \mathbb{R}_+$ be sets of spatial and temporal lags respectively. Then, with the definition given in Coles et al. (1999), we can define the extremal dependence measure, in a spatio-temporal context, for $\boldsymbol{h} \in \Lambda_{\mathcal{S}}, \tau \in \Lambda_{\mathcal{T}}$ and $q \in [0, 1[$ as

$$\chi\left(\boldsymbol{h}, \tau\right) = \lim_{q \to 1} \chi_q\left(\boldsymbol{h}, \tau\right) , \quad \text{with} \quad \chi_q\left(\boldsymbol{h}, \tau\right) = \mathbb{P}(X^*_{\boldsymbol{s},\, t} > q \mid X^*_{\boldsymbol{s}+\boldsymbol{h}, t+\tau} > q) ,$$

where for any $\boldsymbol{s} \in \mathcal{S}$ and any $t \in \mathcal{T}$, $X^*_{\boldsymbol{s},\, t}$ is the uniform rank transformation of $X_{\boldsymbol{s},\, t}$. As explained in Davis et al. (2013), the bivariate distribution of the max-stable process X is given by

$$\mathbb{P}\left(X_{\boldsymbol{s},\, t} \leq x_1,\, X_{\boldsymbol{s}+\boldsymbol{h},\, t+\tau} \leq x_2\right) = \exp\left(-V_{\gamma(\boldsymbol{h},\tau)}\left(x_1, x_2\right)\right) , x_1 , x_2 > 0 ,$$

with

$$V_{\gamma(\boldsymbol{h},\tau)}\left(x_1, x_2\right) = \frac{1}{x_1}\phi\left(\frac{\log \frac{x_2}{x_1}}{2\sqrt{\frac{1}{2}\gamma\left(\boldsymbol{h}, \tau\right)}} + \sqrt{\frac{1}{2}\gamma\left(\boldsymbol{h}, \tau\right)}\right) + \frac{1}{x_2}\phi\left(\frac{\log \frac{x_1}{x_2}}{2\sqrt{\frac{1}{2}\gamma\left(\boldsymbol{h}, \tau\right)}} + \sqrt{\frac{1}{2}\gamma\left(\boldsymbol{h}, \tau\right)}\right)$$

where ϕ is the standard normal distribution function. For max-stable processes, this last function $V_\gamma(\boldsymbol{h}\,, \tau)$ has been linked to the extremogram. From Coles et al. (1999), we have

$$\chi\left(\boldsymbol{h}, \tau\right) \;\; = \;\; 2 - V_{\gamma(\boldsymbol{h},\tau)}\left(1, 1\right) .$$

Therefore, we obtain the spatio-temporal extremogram of the max-stable process of Brown-Resnick X, given by

$$\chi(\boldsymbol{h}, \tau) = 2\left(1 - \phi\left(\sqrt{\frac{1}{2}\gamma(\boldsymbol{h}, \tau)}\right)\right) , \; \boldsymbol{h} \in \Lambda_{\mathcal{S}},\, \tau \in \Lambda_{\mathcal{T}}. \tag{1}$$

where ϕ is the standard normal distribution function and γ is a stationary and isotropic variogram. With this expression we can obtain information about the overall distribution dependence over the spatio-temporal semivariogram γ.

3.2 Extremal dependence with additively separable variogram

We first model the spatio-temporal dependence with an additively separable structure, which facilitates parameter inference while providing a relatively simple but useful model that combines spatial and temporal dependence. To do this, we use the strategy developed in Buhl and Klüppelberg (2018).

We assume that we have a variogram with additive separability, so that we can get a linear parameterisation of the theorical extremogram with an appropriate transformation. Thus, we can write our spatio-temporal variogram, with $0 < \alpha_1, \alpha_2 \leq 2$ and $\beta_1, \beta_2 > 0$, as $\frac{1}{2}\gamma(\boldsymbol{h}, \tau) = \beta_1 \|\boldsymbol{h}\|^{\alpha_1} + \beta_2 \tau^{\alpha_2}$. Then, the theorical spatio-temporal extremogram of X is given by

$$\chi\left(\boldsymbol{h}, \tau\right) = 2\left(1 - \phi\left(\sqrt{\beta_1 \|\boldsymbol{h}\|^{\alpha_1} + \beta_2 \tau^{\alpha_2}}\right)\right).$$

From equation (1), we have

$$\frac{1}{2}\gamma(\boldsymbol{h}, \tau) = \left(\phi^{-1}\left(1 - \frac{1}{2}\chi\left(\boldsymbol{h}, \tau\right)\right)\right)^2$$

and this leads to the transformation $\eta\left(\chi\right) = 2\log\left(\phi^{-1}\left(1 - \frac{1}{2}\chi\right)\right)$ which gives the following expressions with the variogram parameters. For the spatial case, by setting the time lag to 0, we obtain

$$\eta\left(\chi\left(\boldsymbol{h}, 0\right)\right) = \log\left(\beta_1\right) + \alpha_1 \log\left(\|\boldsymbol{h}\|\right) =: c_1 + \alpha_1 x_{\boldsymbol{h}}.$$

For the temporal case, by setting the spatial lag to $\mathbf{0}$, we obtain

$$\eta\left(\chi\left(\mathbf{0}, \tau\right)\right) = \log\left(\beta_2\right) + \alpha_2 \log\left(\tau\right) =: c_2 + \alpha_2 x_{\tau}.$$

Therefore, with the separability assumption, we can separate the spatial and the temporal cases into two linear expressions so that we can use a Weighted Least Squares Estimation (WLSE) for parameter inference on the spatial and the temporal empirical extremograms. A summary of the procedure is given by Figure 3.

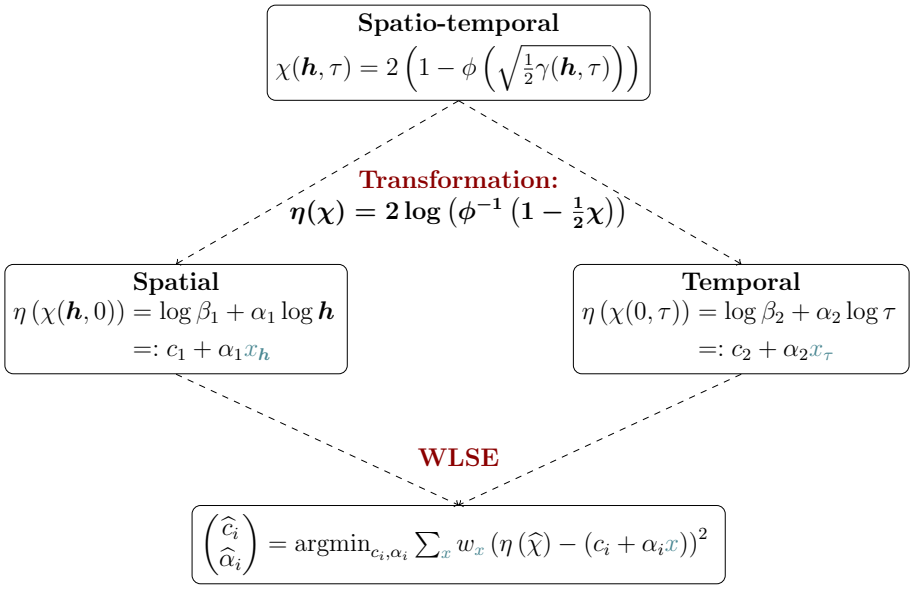


Figure 3: Procedure with the separability assumption on the spatio-temporal variogram

Empirical spatial extremogram

To estimate the spatial extremogram, we first define a spatial structure in an isotropic framework. For the smaller dataset, the OHSM one, we define a set of equifrequent classes $C_{\boldsymbol{h}}$ for each spatial lag $\boldsymbol{h} \in \Lambda_{\mathcal{S}}$ which leads us to the set of all pairs of locations within each distance class as

$$N(\boldsymbol{h}) = \left\{(\boldsymbol{s}_i, \boldsymbol{s}_j) \in \mathcal{S}^2 \;\middle|\; \|\boldsymbol{s}_i - \boldsymbol{s}_j\| \in C_{\boldsymbol{h}}\right\}.$$

For the largest COMEPHORE dataset, we have a spatial grid and we can define a set that contains all pairs of locations for the same spatial lag $\boldsymbol{h} \in \Lambda_{\mathcal{S}}$ as

$$N(\boldsymbol{h}) = \left\{(\boldsymbol{s}_i, \boldsymbol{s}_j) \in \mathcal{S}^2 \;\middle|\; \|\boldsymbol{s}_i - \boldsymbol{s}_j\| = \|\boldsymbol{h}\|\right\}.$$

Then for a fixed $t \in \mathcal{T}$ and for any $(\boldsymbol{s}_i, \boldsymbol{s}_j) \in N(\boldsymbol{h})$, a natural estimator of the spatial extremogram is

$$\widehat{\chi}_q^{(t)}(\boldsymbol{h}, 0) = \frac{\frac{1}{|N(\boldsymbol{h})|}\sum_{i,j\,|\,(\boldsymbol{s}_i,\boldsymbol{s}_j)\in N(\boldsymbol{h})} \mathbb{1}_{\left\{X^*_{\boldsymbol{s}_i,t}>q\,, X^*_{\boldsymbol{s}_j,t}>q\right\}}}{\frac{1}{|\mathcal{S}|}\sum_{i=1}^{|\mathcal{S}|} \mathbb{1}_{\left\{X^*_{\boldsymbol{s}_i,t}>q\right\}}}\,, \boldsymbol{h} \in \Lambda_{\mathcal{S}}$$

where q is a high quantile (99.8%).

Empirical temporal extremogram

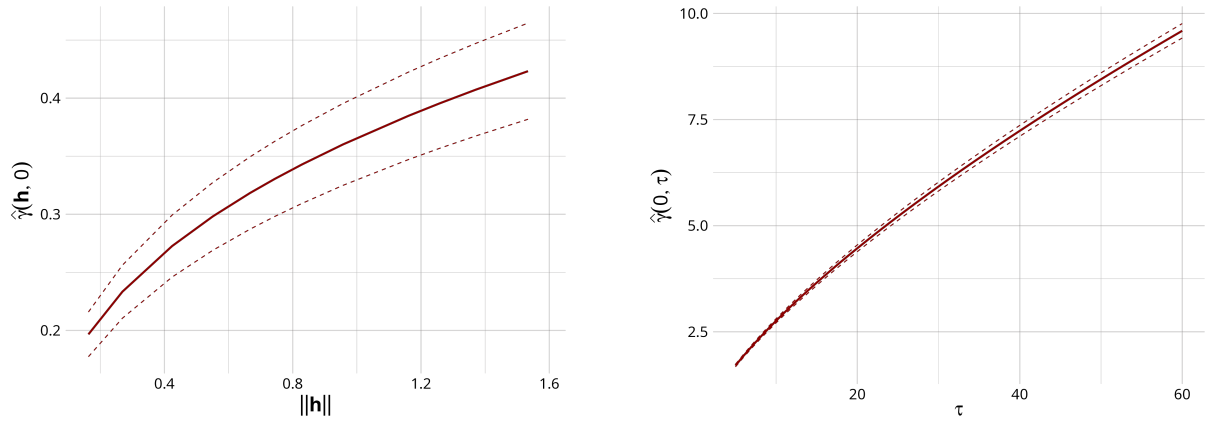
Regarding the temporal case, for a fixed $\boldsymbol{s} \in \mathcal{S}$ and for any $t \in \mathcal{T}$, a natural estimation of the temporal dependence measure is

$$\widehat{\chi}_q^{(\boldsymbol{s})}(0, \tau) = \frac{\frac{1}{T-\tau}\sum_{k=1}^{T-\tau} \mathbb{1}_{\{X^*_{\boldsymbol{s},t_k}>q\,, X^*_{\boldsymbol{s},t_k+\tau}>q\}}}{\frac{1}{T}\sum_{k=1}^{T} \mathbb{1}_{\{X^*_{\boldsymbol{s},t_k}>q\}}}\,, \tau \in \Lambda_{\mathcal{T}}$$

where q is a high quantile (99.8%) and $t_k \in \{t_1, \ldots, t_T\} \subseteq \mathcal{T}$.

Spatial and temporal variogram estimation

We apply the described procedure on OHSM data to infer the variogram parameters. Then we obtain the final estimation of the spatial variogram $\widehat{\gamma}(\boldsymbol{h}, 0) = \widehat{\beta}_1 \|\boldsymbol{h}\|^{\widehat{\alpha}_1}$ and the temporal variogram $\widehat{\gamma}(\boldsymbol{0}, \tau) = \widehat{\beta}_2 \tau^{\widehat{\alpha}_2}$, illustrated in Figure 4. The spatial variogram has an exponential form, indicating a continuous decrease in spatial correlation with distance, without reaching an apparent range. This suggests, as expected, that the rainfall data continue to show some degree of spatial correlation for larger distances. However, even for very small distances, there is variability and a decrease in dependence. The temporal variogram shows an almost linear relationship, indicating temporal variability within the rainfall data. The linear form suggests that the dependence between observations decreases proportionally with the time even over short time periods. This was expected because rain clouds move with time and rain events can be brief. These variograms show the relevance of our model at such a fine spatio-temporal scale.



(a) Spatial variogram

(b) Temporal variogram

Figure 4: Estimation of the spatial and temporal variograms on OHSM data, where $\|\boldsymbol{h}\|$ units are kilometers and τ units are minutes

However, assuming a separable structure is a strong hypothesis in our modeling. We looked at the spatial variogram at different fixed time lags (not shown here) and see that a non-separable structure may be required. So we want to relax this hypothesis to get a better dependence model.

3.3 Adding advection to the dependence model

When dealing with precipitation, it may be relevant to consider the role of advection. In meteorology, this is the horizontal transport of properties such as heat or moisture by the movement of air masses such as wind and clouds. It can therefore influence the behaviour of precipitation. In fact, due to winds, there is a spatial and temporal movement of clouds and precipitation with a specific direction and speed. This can lead to characteristic dependence patterns in precipitation, which we can capture by appropriate parametric structures with the space-time variogram.

Let $\boldsymbol{V} \in \mathbb{R}^2$ be the advection vector. It can be defined by its polar coordinates (S_V, θ_V), where S_V corresponds to a velocity and θ_V to a direction angle. As said in Lepioufle et al. (2012), if we consider the Lagrangian referential L, moving with the air masses, we have a link between the Lagrangian semivariogram and the Eulerian semivariogram with the advection: $\gamma_L(\boldsymbol{h}, \tau) = \gamma(\boldsymbol{h} - \boldsymbol{V}.\tau, \tau)$. Then if we add this quantity in our separable model presented in 3.2, we obtain

$$\frac{1}{2}\gamma(\boldsymbol{h} - \boldsymbol{V}.\tau, \tau) = \beta_1\|\boldsymbol{h} - \boldsymbol{V}.\tau\|^{\alpha_1} + \beta_2\tau^{\alpha_2}.$$

With this expression we have relaxed the previous strong hypothesis of separability and then we cannot use the WLSE approach to infer the 6 parameters of the spatio-temporal variogram. Instead, a parameter optimization method will be used.

Let $\boldsymbol{\Theta}$ be the vector of parameters to be estimated *i.e.* $\boldsymbol{\Theta} = (\beta_1, \beta_2, \alpha_1, \alpha_2, S_V, \theta_V)$ and let m be the total number of space-time pair combinations. Let $E_p = E_{(\boldsymbol{s}_i, \boldsymbol{s}_j), (t_i, t_j)}$ be the indicator

of simultaneous excess for a pair of locations $(\boldsymbol{s}_i, \boldsymbol{s}_j) \in \mathcal{S}^2$ and a pair of times $(t_i, t_j) \in \mathcal{T}^2$. We have $E_p = \mathbb{1}_{\{X^*_{\boldsymbol{s}_i,t_i} > q \,|\, X^*_{\boldsymbol{s}_j,t_j} > q\}}$ and then $\mathbb{E}(E_p) = \chi_{p,\boldsymbol{\Theta}}$. Hence E_p follows a Bernoulli distribution with a probability parameter $\chi_{p,\boldsymbol{\Theta}}$. So we have $\sum_{p=1}^{m} E_p \sim \mathcal{B}(n_p\,;\, \chi_{p,\boldsymbol{\Theta}})$ where n_p is the total number of records for the p-th combination. Hence, with $\underline{E} = (E_p)_{p\in\{1,\ldots,m\}}$, the log-likelihood in order to get parameter estimates is

$$\log\left(L_{\boldsymbol{\Theta}}\left(\underline{E}\right)\right) = \sum_{p=1}^{m}\left[\log\binom{n_p}{k_p} + k_p \log \chi_{p,\boldsymbol{\Theta}} + (n_p - k_p)\log(1-\chi_{p,\boldsymbol{\Theta}})\right],$$

The separability approach, described in 3.2, gives us an initial estimate of $\boldsymbol{\Theta}$ with a zero advection. These would be the initial parameters for the log-likelihood maximisation. We apply this procedure to both datasets to obtain an overall estimate of a constant advection over the study area and hence we obtain a spatio-temporal variogram estimate.

4 Perspectives

Until now, both rainfall data (OHSM and COMEPHORE) were studied separately. The first dataset corresponds to 17 random spatial points with a high spatio-temporal resolution, while in the other we have a regular spatial grid with a lower spatial and temporal resolution. In future work, we want to combine the two datasets to obtain more regular spatial point measurements across the study area, while maintaining the fine spatio-temporal resolution. To do this, we will explore a downscaling approach that will also allow us to compare the two datasets. Downscaling involves dividing each spatial and temporal intervals of the higher resolution dataset into a number of smaller intervals. This can be done by interpolating, or by using statistical or physical models.

In addition, we will look at other datasets, such as the basic climate data from Météo France[2], which provides rainfall amounts with a resolution of 6 minutes for a station in Montpellier. This dataset also provides wind information with magnitude and direction, which may be relevant for our analysis. Indeed, we currently have a dependence model with constant advection, which is a strong hypothesis. Taking this wind data into account in the model could lead to a more accurate estimation of the advection in space and time.

References

Buhl, S., & Klüppelberg, C. (2018). Limit theory for the empirical extremogram of random fields. *Stochastic Processes and their Applications*, *128*(6), 2060–2082.

Coles, S., Heffernan, J., & Tawn, J. (1999). Dependence measures for extreme value analyses. *Extremes*, *2*, 339–365.

Davis, R. A., Klüppelberg, C., & Steinkohl, C. (2013). Max-stable processes for modeling extremes observed in space and time. *Journal of the Korean Statistical Society*, *42*(3), 399–414.

[2]The dataset is available here: https://meteo.data.gouv.fr/datasets/6569ad61106d1679c93cdf77

Davison, A. C., Padoan, S. A., & Ribatet, M. (2012). Statistical modeling of spatial extremes.

Dombry, C., & Ribatet, M. (2015). Functional regular variations, pareto processes and peaks over threshold. *Statistics and its Interface*, *8*(1), 9–17.

Ferreira, A., & De Haan, L. (2014). The generalized pareto process; with a view towards application and simulation.

Finaud-Guyot, P., Guinot, V., Marchand, P., Neppel, L., Salles, C., & Toulemonde, G. (2023). Rainfall data collected by the HSM urban observatory (OMSEV).

Haruna, A., Blanchet, J., & Favre, A.-C. (2023). *Modeling areal precipitation hazard: A data-driven approach to model intensity-duration-area-frequency relationships using the full range of non-zero precipitation in switzerland* (preprint). Preprints.

Kabluchko, Z., Schlather, M., & De Haan, L. (2009). Stationary max-stable fields associated to negative definite functions.

Lepioufle, J.-M., Leblois, E., & Creutin, J.-D. (2012). Variography of rainfall accumulation in presence of advection. *Journal of Hydrology*, *464-465*, 494–504.

Naveau, P., Huser, R., Ribereau, P., & Hannart, A. (2016). Modeling jointly low, moderate, and heavy rainfall intensities without a threshold selection. *Water Resources Research*.

Tabary, P., Dupuy, P., L'Henaff, G., Gueguen, C., Moulin, L., & Laurantin, O. (2012). A 10-year (1997—2006) reanalysis of quantitative precipitation estimation over france: Methodology and first results. *IAHS-AISH publication*, (351), 255–260.