# Rebuttal of Review 2

**The manuscript addresses the analysis of precipitation across spatially-related sites with a focus on modeling zero-inflated data. It places significant emphasis on the multivariate nature of the problem, demonstrated through a five-dimensional application. While the topic is interesting and relevant, the manuscript requires greater detail in several areas to adequately convey the value and robustness of the proposed methodology. Below are specific comments and suggestions for improvement:**

**General CommentsComputational Complexity and Data Requirements:**

**The model estimates $2^d$ distributions, where d is the number of sites. This naturally entails considerable computational costs. Additionally, since each model is estimated on a subset of the data (e.g., Group 32 is only modeled when all variables are non-zero), the approach presupposes access to a large dataset. These aspects warrant discussion and explicit acknowledgment within the manuscript.**

While the estimation of $2^d$ distributions may suggest high computational costs, the parameter estimation process is efficient. Modern statistical packages, such as *VineCopula*, incorporate optimized algorithms that handle this task with minimal computational expense, even in high-dimensional settings. As a result, this step remains computationally feasible without requiring further simplifications. However, the main computational cost arises from estimating the critical hypersurface, which entails a high computational cost. To address this complexity, we employed Gaussian Process Regression (GPR).

Regarding data availability, the need for sufficiently large sample sizes is a well-known challenge in the literature (Brunner 2023; Serinaldi 2013). A widely used strategy to overcome this issue is the generation of synthetic data, which improves the representation of less frequent events while preserving key statistical dependencies. In this study, we applied this approach by fitting multiple copula families to the observed dataset, selecting the best-fitting model based on statistical tests, and generating additional samples accordingly. To ensure that the synthetic data accurately reflected the dependence structure of the original dataset, we compared and validated their correlation coefficients, as discussed in Section 3.3.2.

To strengthen the manuscript, we will explicitly address these aspects, highlighting the computational efficiency of the model and the approach taken to address data limitations.

**Detailed Comments**

**Equation (1):**

**Greater clarity is required for the parameters p0,p1,..., etc, to ensure they constitute a valid probability model. Specifically, do these parameters sum to 1? Providing this information would strengthen the theoretical foundations of the model.**

The parameters $p_0, p_1, ...p_n$ represent the probabilities associated with each group, where each group corresponds to a specific pattern of rainfall across the analyzed stations. Since these probabilities describe mutually exclusive and collectively exhaustive events, their sum is equal to 1, ensuring that Equation (1) defines a valid probability distribution (further details on the group configurations are provided in the *Response to Figure 1*).

To address this point, we will revise the manuscript to clearly articulate this condition, ensuring a more transparent representation of the framework.

**Model Descriptions (Page 5):**

**Model 1 (Gaussian Copula Without Intermittency):**

**Does this model fit a copula to zero-inflated data without accounting for the discrete component of the marginal distributions? If so, is it appropriate to apply copulas to non-continuous data? Addressing this issue would clarify the legitimacy of the approach.**

The model applies a Gaussian copula directly to the data without explicitly accounting for the discrete component associated with zero intermittency in the marginal distributions. This choice was intentional, allowing us to assess the impact of treating the data as fully continuous, particularly in extreme event scenarios.

The use of copulas in non-continuous data with zero inflation can bias dependence estimation by assuming continuity. To address this limitation, the other approaches explored in this study (lines 125-138) adapt the methodology of Serinaldi (2009). However, we intentionally use the Gaussian copula as a reference to illustrate the impact of ignoring the discrete-continuous nature of the data on the estimation of extreme dependencies.

In this work, the Gaussian copula was included as a benchmark to evaluate its performance against more flexible alternatives, such as R-vine copulas, which better capture dependencies in zero-inflated data. Our results show that ignoring intermittency leads to an underestimation of extreme dependencies, aligning with findings in the literature and highlighting the limitations of this approach. While applying copulas to zero-inflated data without adjustments introduces biases, this model serves as a useful reference to demonstrate the importance of incorporating intermittency explicitly.

To clarify this aspect in the manuscript, we will expand the discussion on these limitations and explicitly compare this approach with models that account for zero inflation, ensuring that its role within our analysis is well understood.

**Model 3 (Vine Gaussian Copulas):**

**Are vine Gaussian copulas equivalent to traditional Gaussian copulas? When bivariate Gaussian copulas are assigned to the edges of a vine, the resulting multivariate density corresponds to a Gaussian density parameterized by a partial correlation vine rather than a standard correlation matrix. Clarification on this point would enhance understanding.**

Vine Gaussian copulas differ from traditional Gaussian copulas in their parametrization and dependency structure. The distinction lies in how dependence is structured: while Gaussian copulas rely on a single correlation matrix to describe dependencies, vine copulas decompose the joint distribution into a cascade of pair-copulas, allowing for more flexible and localized dependency modeling. This hierarchical structure enables vine Gaussian copulas to better capture complex dependence patterns while maintaining analytical tractability and computational efficiency in high-dimensional settings (Aas et al. 2009).

In our study, the vine Gaussian copula served as a baseline for evaluating how different copula structures represent multivariate dependencies. To ensure precision, we will revise the manuscript to explicitly differentiate vine Gaussian copulas from traditional Gaussian copulas.

**Model 4 (Vine Extreme Copula Model):**

**The terminology for this model may be misleading. It appears to be a vine copula where pair-copulas can be selected from different classes, including those that describe asymptotic tail dependence. This does not necessarily qualify it as an extreme-value copula. A reconsideration of the terminology is recommended to avoid confusion.**

The concern regarding the terminology is valid. The model employs an R-vine structure that allows for a flexible selection of bivariate copulas, including t-Student, Clayton, Gumbel, Joe6, BB1, BB6, BB7, BB8, and independent copulas, as well as rotated versions of Archimedean copulas to capture negative dependencies. While several of these copulas can model asymmetries and tail dependencies, we recognize that the term "Vine Extreme" could be misinterpreted as implying that the entire model is an extreme-value copula.

To avoid potential ambiguity, we will revise the terminology in the manuscript. We propose replacing *Vine Extreme Copula Model* with *Flexible Vine Copula Model*, which better reflects the model's composition without implying adherence to extreme-value theory.

**Figure 1:**

**It appears that model estimation is conducted independently within each group. Is this correct? Based on Equation (1), the model suggests conditional independence when conditioned on the rain/no-rain status for each site. As a result, the copula for Group 32 is entirely independent from the copula for Group 31. Is this assumption realistic? Further discussion on this matter is recommended.**

Model estimation is indeed conducted independently within each group, as illustrated in Figure 1 and defined in Equation (1). Each group contains all the events that generate precipitation in some stations, but not in others -except for Group 32, in which all events generate precipitation in all the stations-. Each group is represented as a binary sequence, where 1 indicates that all the events belonging to the group generate rainfall for that station, and 0 denotes the opposite behavior. The following sequences illustrate different distribution patterns:

- **Group 1** = [1,0,0,0,0] → Includes events where precipitation occurs only at the last station.
- **Group 2** = [1,1,0,0,0] → Includes events where precipitation occurs at the last two stations.
- **Group 3** = [1,1,1,0,0] → Includes events where precipitation occurs at the last three stations.
- …

- **Group 32** = [1,1,1,1,1] → Includes events where precipitation occurs at all stations.

However, while estimation is performed separately for each group, the joint cumulative distribution function (JCDF) integrates all groups collectively. This step is crucial for defining the critical hypersurface, as it captures the overall complexity of the hydrological system by combining probabilities across all possible precipitation configurations. The JCDF is not solely dependent on within-group dependencies but rather reflects the accumulated probability of extreme event occurrences across the full range of rainfall scenarios.

This strategy balances computational efficiency with the need to preserve the system's spatial and temporal complexity. We propose the following modification to the manuscript discussion, as this aspect deserves further attention:

"*While model estimation is conducted independently for each group, considering specific configurations of rainfall occurrence or absence across the stations, the construction of the joint cumulative distribution function (JCDF) is based on the integration of all modeled groups. This approach allows for capturing the complexity of the hydrological system by combining the probabilities associated with each configuration within a unified framework.*

*Segmenting by groups facilitates detailed modeling of particular scenarios, while the JCDF, calculated by summing the conditional probabilities of each group, preserves the system's integrity by considering all possible scenarios. This approach is particularly relevant in defining the critical surface, where it is necessary to evaluate the accumulated probability of extreme event occurrence across all possible combinations of rainfall and no rainfall among the stations.*

*Although this method assumes conditional independence between groups during estimation, the final integration into the joint CDF enables the evaluation of dependencies at a global level. This combination of segmented modeling and joint analysis aims to balance computational efficiency with an accurate representation of the system's dependency structure.*"

**Page 11 (Marginal GEV Estimation):**

**The manuscript briefly mentions that marginal GEV estimation is conducted using Bayesian techniques. However, details regarding the Bayesian approach are sparse. Please provide a more comprehensive description of the estimation procedure.**

The estimation of Generalized Extreme Value (GEV) distribution parameters was conducted using Bayesian techniques, implemented via the Stan package in Python, which leverages Markov Chain Monte Carlo (MCMC) algorithms to model complex statistical structures. This approach accounts for parameter uncertainty, providing a more robust representation of variability in extreme events.

For each station, non-informative priors were assigned to the location, scale, and shape parameters, ensuring that posterior estimates were primarily data-driven. The No-U-Turn Sampler (NUTS) algorithm, integrated into Stan, was used for MCMC simulations, optimizing parameter space exploration and improving sampling convergence.

Recognizing that the main manuscript provides only a brief mention of this procedure, we will expand the description in the Supplementary Information, detailing:

- Bayesian Framework
- The rationale for the choice of prior distributions.
- Bayesian Parameter Estimation and MCMC Implementation for GEV Models
- Convergence validation criteria and goodness-of-fit metrics.

This addition will enhance methodological transparency and facilitate reproducibility for researchers applying Bayesian techniques in extreme value modeling.

**Line 310 (Computational Complexity of GPR Technique):**

**Given the computational intensity of calculating copula values from vine specifications, more details should be provided, including an algorithm if possible, to clarify the practical implementation of the Gaussian Process Regression (GPR) technique.**

The comment on the computational complexity of Gaussian Process Regression (GPR), particularly in relation to calculating copula values from vine specifications, is well taken. A more detailed explanation of its implementation and role in reducing computational costs will be included in the manuscript.

As previously noted, GPR is central to our methodology, providing an efficient surrogate model for approximating the joint distribution function while avoiding repeated, computationally intensive evaluations of vine copula structures. This approach effectively manages the high-dimensional complexity of copula-based models, particularly in calculating the critical layer associated with joint return periods.

To ensure transparency and reproducibility, a detailed description of the GPR methodology will be included in the Supplementary Information, covering:

- Theoretical foundations of GPR in the context of copula modeling.
- Detailed algorithmic steps, including data preparation, training, and validation.
- Hyperparameter selection and kernel optimization processes.
- Application to JCDF Estimation.

**Line 330 (Upper Tail Dependence):**

**The statement regarding the superior fit of copulas with upper tail dependence requires clarification. The model is five-dimensional, making it unclear how tail dependence is conceptualized and evaluated. Further elaboration is necessary.**

The reviewer raises an important point regarding the conceptualization and evaluation of tail dependence in a five-dimensional setting. In our study, tail dependence is assessed at the pairwise level by calculating the tail dependence coefficient for each station pair. These coefficients are then analyzed collectively to infer patterns in the multivariate space. This approach enables us to identify consistent upper tail dependence structures across multiple locations rather than relying on a single aggregated metric for the entire five-dimensional model.

The statement regarding the superior fit of copulas with upper tail dependence is supported by both visual and statistical evidence. While KDE density plots (Figure 9) provide an intuitive representation of tail concentration, statistical validation was conducted using the non-parametric estimator proposed by Schmidt and Stadtmuller (2006), known for its ability to detect tail dependencies in multivariate contexts, as detailed in lines 203–205. This method, also applied in Brunner, Furrer, and Favre (2019), formally quantifies tail dependence and confirms its presence in the dataset. However, we are aware of the limitations of this method, as acknowledged in the manuscript by citing Serinaldi, Bárdossy, and Kilsby (2015), who note that the estimator can be sensitive to data sparsity in extreme tails, potentially affecting the stability of the estimates. Therefore, the results were interpreted with caution, taking into account both the statistical limitations of the method and prior evidence supporting the existence of upper tail dependence in extreme precipitation events (Serinaldi 2013; Evin, Favre, and Hingray 2018).

To address this comment, we will revise the manuscript to explicitly define how tail dependence is evaluated within the five-dimensional framework and refine the statement regarding the fit of copulas with upper tail dependence to ensure alignment with this explanation.

**Section 3.5.2 (Likelihood vs. Probability):**

**It may be beneficial to distinguish between likelihood and probability, as is customary in the literature. This would ensure greater terminological precision and alignment with established conventions.**

**I hope these suggestions prove helpful in strengthening the manuscript.**

The distinction between "likelihood" and "probability" is well noted. Referring to the algorithm as "Maximum Likelihood Estimation (MLE)" may be misleading, as it does not maximize a likelihood function in the classical sense but rather applies gradient descent to identify the highest-density point within the critical layer.

To enhance clarity, we will specify that, while the algorithm is named *"MLE"* following the convention used in the Spotpy library, its application in this context does not adhere to the statistical definition of likelihood. Rather, it is employed to optimize the joint probability density function and identify the most representative events within the defined space.

Proposed modification in the manuscript (*Section 3.5.2*):

*"Although referred to as the Maximum Likelihood Estimation (MLE) algorithm, following the naming convention used in the Spotpy library, this method does not perform classical likelihood maximization. Instead, it functions as a gradient descent optimization technique aimed at maximizing the joint probability density function. This distinction is essential, as the term 'likelihood' here reflects the algorithm's label rather than the statistical definition. This approach was selected due to its computational efficiency in identifying the most probable event within the critical layer."*

## References

Aas, Kjersti, Claudia Czado, Arnoldo Frigessi, and Henrik Bakken. 2009. "Pair-Copula Constructions of Multiple Dependence." *Insurance: Mathematics and Economics* 44 (2): 182–98. https://doi.org/10.1016/j.insmatheco.2007.02.001.

Brunner, Manuela I. 2023. "Floods and Droughts: A Multivariate Perspective." *Hydrology and Earth System Sciences* 27 (13): 2479–97. https://doi.org/10.5194/hess-27-2479-2023.

Brunner, Manuela I., Reinhard Furrer, and Anne-Catherine Favre. 2019. "Modeling the Spatial Dependence of Floods Using the Fisher Copula." *Hydrology and Earth System Sciences* 23 (1): 107–24. https://doi.org/10.5194/hess-23-107-2019.

Evin, Guillaume, Anne-Catherine Favre, and Benoit Hingray. 2018. "Stochastic Generation of Multi-Site Daily Precipitation Focusing on Extreme Events." *Hydrology and Earth System Sciences* 22 (1): 655–72. https://doi.org/10.5194/hess-22-655-2018.

Schmidt, Rafael, and Ulrich Stadtmuller. 2006. "Non-Parametric Estimation of Tail Dependence." *Scandinavian Journal of Statistics* 33 (2): 307–35. https://doi.org/10.1111/j.1467-9469.2005.00483.x.

Serinaldi, Francesco. 2009. "Copula-Based Mixed Models for Bivariate Rainfall Data: An Empirical Study in Regression Perspective." *Stochastic Environmental Research and Risk Assessment* 23 (5): 677–93. https://doi.org/10.1007/s00477-008-0249-z.

———. 2013. "An Uncertain Journey Around the Tails of Multivariate Hydrological Distributions: Multivariate Frequency Analysis with Uncertainty." *Water Resources Research* 49 (10): 6527–47. https://doi.org/10.1002/wrcr.20531.

Serinaldi, Francesco, András Bárdossy, and Chris G. Kilsby. 2015. "Upper Tail Dependence in Rainfall Extremes: Would We Know It If We Saw It?" *Stochastic Environmental Research and Risk Assessment* 29 (4): 1211–33. https://doi.org/10.1007/s00477-014-0946-8.

# Return period of high-dimensional compound events. Part II: Analysis of spatially-variable precipitation - Supplemantary information

Diego Urrea Méndez[1], Manuel Del Jesus[1], and Dina Vanessa Gomez Rave[1]

[1]IHCantabria - Instituto de Hidráulica Ambiental de la Universidad de Cantabria, Santander, Spain.

**Correspondence:** Manuel Del Jesus (manuel.deljesus@unican.es)

## 1 Generalized extreme value distribution

The Generalized Extreme Value (GEV) distribution, rooted in the work of Fisher and Tippett (1928), was formalized by Jenkinson (1955), who unified the Gumbel, Fréchet, and Weibull distributions into a single framework. Widely applied in hydrology, climatology, and engineering, the GEV is defined by three parameters: location ($\mu$), scale ($\sigma$), and shape ($\xi$), which control its position, spread, and tail behavior. Its flexibility in modeling light-tailed, heavy-tailed, and bounded distributions makes it a key tool for analyzing extreme events.

The probability density function (pdf) of the GEV is defined in Eq. (S1) as:

$$f(x|\mu,\sigma,\xi) = \frac{1}{\sigma}\left[1+\xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-\left(\frac{1}{\xi}+1\right)} \exp\left\{-\left[1+\xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-\frac{1}{\xi}}\right\} \tag{S1}$$

where $1+\xi((x-\mu)/\sigma) > 0$.

The GEV encompasses the following classical extreme value distributions. Gumbel ($\xi = 0$) for exponentially tailed events, Fréchet ($\xi > 0$) for heavy-tailed events nad Weibull ($\xi < 0$) for bounded-tailed events.

## 2 Bayesian Framework

The modeling of extreme events has become increasingly relevant in various scientific domains, including hydrology, climatology, and finance. Understanding and predicting rare but impactful events such as floods, heatwaves, and financial crashes is essential for effective risk management. The Generalized Extreme Value (GEV) distribution has emerged as a fundamental tool in this context due to its flexibility in modeling block maxima and its solid theoretical foundation derived from the Fisher-Tippett-Gnedenko theorem.

The Bayesian framework offers a robust approach to estimating GEV parameters, allowing the incorporation of prior knowledge and a natural quantification of uncertainty. This approach has been widely applied across different fields. In urban

planning, Balbi et al. (2016) applied a spatial Bayesian network model to assess the benefits of early warning systems for urban flood risk, improving risk management strategies in complex hydrological environments.

These diverse applications highlight the versatility and robustness of the Bayesian GEV approach in both environmental and economic contexts, underscoring its value as a key tool for decision-making under uncertainty.

The Bayesian approach allows incorporating prior information and quantifying uncertainty in parameter estimation. It is based on Bayes' theorem, presented in Eq. (S2):

$$p(\theta|x) = \frac{L(x|\theta)p(\theta)}{p(x)} \tag{S2}$$

where $p(\theta|x)$ is the posterior distribution, $L(x|\theta)$ is the likelihood function derived from the GEV, $p(\theta)$ represents the prior distributions assigned to the parameters, and $\theta = (\mu, \sigma, \xi)$ represents the vector of unknown parameters. The term $p(x)$, known as the evidence or marginal likelihood, acts as a normalizing constant ensuring that the posterior distribution integrates to one. Although $p(x)$ is independent of $\theta$ and often omitted in practical MCMC implementations, it is essential in the formal definition of Bayes' theorem.

Given the analytical intractability of the posterior distribution, Markov Chain Monte Carlo (MCMC) methods are used for posterior sampling. Algorithms such as Metropolis-Hastings and Hamiltonian Monte Carlo (HMC) are preferred due to their efficiency in exploring parameter spaces. These techniques allow for the estimation of the full posterior distribution of the GEV parameters, enabling a comprehensive uncertainty assessment that is critical for risk analysis in fields such as hydrology and finance.

Significant contributions to Bayesian GEV modeling have shaped the current state of the art. Coles and Powell (1996) laid the foundational work, introducing Bayesian principles into extreme value modeling and demonstrating their advantages over traditional methods. This was further extended by Heffernan and Tawn (2004), who incorporated covariates and hierarchical structures into the Bayesian GEV framework, enabling more complex and context-sensitive analyses.

### S2.1 Bayesian Parameter Estimation and MCMC Implementation for GEV Models

The Bayesian fitting of the Generalized Extreme Value (GEV) distribution relies on carefully chosen prior distributions for its three parameters. In our analysis of rainfall extremes, the priors have been explicitly defined based on prior knowledge and the expected behavior of extreme precipitation events.

#### S2.1.1 Prior Specifications

Location Parameter ($\mu$): We assign a normal prior to $\mu$, formally defined in Eq. (S3):

$$\mu \sim \mathcal{N}(\mu_0, \sigma_0^2) \tag{S3}$$

In our analysis, we set $\mu_0$ to the sample mean of the annual maximum rainfall observations, reflecting our empirical understanding of the central tendency of extreme rainfall events. We choose $\sigma_0^2$ as a large value (e.g., 10 times the sample variance) to ensure that the prior is weakly informative, allowing the data to play a significant role in shaping the posterior estimates.

Scale Parameter ($\sigma$): We assign a log-normal prior to $\sigma$ by applying a log-transformation, formally defined in Eq. (S4):

$$\log(\sigma) \sim \mathcal{N}(\alpha, \beta^2) \tag{S4}$$

In our analysis, we set $\alpha = 0$ to center the distribution and choose $\beta^2$ as a large value (e.g., approximately 10, which corresponds to $\beta \approx 3$) to ensure weak informativeness. This approach guarantees that $\sigma$ remains positive and allows for a wide range of plausible values, enabling the data to have a significant influence on the posterior estimates. Additionally, the log-transformation improves computational stability and chain mixing during the MCMC sampling.

Shape Parameter ($\xi$):We assign a normal prior to $\xi$, formally defined in Eq. (S5):

$$\xi \sim \mathcal{N}(0, \gamma^2) \tag{S5}$$

In our analysis, we set $\gamma = 0.25$, which concentrates most of the prior mass in the interval $[-0.5, 0.5]$. Given the high sensitivity of the shape parameter to prior specifications, this choice reflects our empirical expectations regarding tail behavior in extreme rainfall events and allows the model to adequately capture the three types of tail behavior (Gumbel, Fréchet, and Weibull). This careful selection ensures that the data play a significant role in refining the posterior estimates for $\xi$.

The shape parameter exhibits the highest sensitivity to the prior specification, directly influencing the inference on extreme values. In contrast, the location parameter tends to be robust across different prior choices, while the scale parameter shows moderate sensitivity. For the estimation, MCMC methods are employed; the log-transformation of $\sigma$ improves both computational stability and chain mixing.

This explicit framework for prior selection balances theoretical requirements with practical implementation, thereby providing a reliable foundation for extreme value analysis in the context of rainfall extremes.

The likelihood function is formulated based on the GEV density and the observed data, emphasizing the accurate representation of tail behavior, which is crucial in extreme value analysis.

To estimate the posterior distribution, MCMC methods are employed, offering a robust alternative to conventional numerical techniques. In this study, the PyStan library was used for Bayesian inference, utilizing four independent chains of length $N = 1000$, following a burn-in period of 2000 iterations to ensure convergence. The Metropolis-Hastings algorithm was applied to construct the sampling chains, enabling efficient exploration of the parameter space and reducing autocorrelation between samples.

Convergence diagnostics were assessed using the R-hat statistic, where values close to 1 indicate satisfactory convergence across chains. Finally, the posterior analysis focused on estimating key statistics such as the mean, median, and mode for each

parameter, alongside constructing 95% credible intervals to quantify uncertainty. These intervals represent the range within which the true parameter values are expected to lie with 95% probability, offering a comprehensive understanding of the parameter space under the Bayesian framework.

## 3 Gaussian Process Regression (GPR)

Gaussian Process Regression (GPR) is a flexible, non-parametric probabilistic modeling approach widely used for regression tasks involving complex, non-linear relationships. It was formalized for machine learning applications by Rasmussen and Williams (2008), GPR models data as a distribution over functions, defined by a mean function and a covariance function (kernel). Typically, the mean function is set to zero, focusing attention on the choice of the covariance function, which encodes the assumptions about the function's smoothness, periodicity, and overall behavior.

The core of GPR lies in the specification of the kernel function, which determines the covariance between data points. In our study, the *Rational Quadratic kernel* was identified as the most suitable for capturing the underlying relationships within the data. This kernel combines the properties of the Radial Basis Function (RBF) and a scale mixture of Gaussian kernels, allowing it to handle varying length scales effectively. This flexibility was crucial when modeling the complex dependence structures of compound precipitation events across multiple locations.

The Gaussian Process is formally expressed as:

$$f(x) \sim \mathcal{GP}(m(x), k(x, x')) \tag{S6}$$

where $m(x)$ is the mean function (often zero) and $k(x, x')$ is the covariance function, or kernel. Several kernel functions can be used in GPR, each offering different properties in terms of smoothness and flexibility:

1. **Radial Basis Function (RBF) Kernel**: Also known as the Squared Exponential kernel, it assumes smooth and infinitely differentiable functions. It is defined as:

$$k_{\text{RBF}}(x, x') = \sigma^2 \exp\left(-\frac{(x - x')^2}{2l^2}\right) \tag{S7}$$

where $\sigma^2$ is the variance and $l$ is the length scale controlling the smoothness.

2. **Matérn Kernel**: Provides more flexibility by controlling the smoothness through the parameter $\nu$. For $\nu = \frac{3}{2}$ or $\nu = \frac{5}{2}$, the kernel has specific closed forms. The general form is:

$$k_{\text{Matérn}}(x, x') = \sigma^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{\sqrt{2\nu}|x - x'|}{l}\right)^{\nu} K_\nu \left(\frac{\sqrt{2\nu}|x - x'|}{l}\right) \tag{S8}$$

where $K_\nu$ is the modified Bessel function of the second kind, $l$ is the length scale, and $\nu$ controls smoothness.

3. **Rational Quadratic Kernel**: Acts as a scale mixture of RBF kernels with different length scales, offering robustness in modeling varying degrees of smoothness. It is defined as:

$$k_{\text{RQ}}(x, x') = \sigma^2 \left(1 + \frac{(x - x')^2}{2\alpha l^2}\right)^{-\alpha} \tag{S9}$$

where $\alpha$ is a scale-mixture parameter controlling the relative weighting of small and large-scale variations.

#### S3.0.1 Application to Copula CDF Estimation

In our framework, GPR was employed to approximate the Joint Cumulative Distribution Function (JCDF), a critical step in understanding the joint behavior of multiple variables or, in our case, a single variable measured at different locations. Direct computation of the JCDF, especially for high-dimensional data and complex copula structures like R-vines, is computationally expensive. To address this, GPR served as a surrogate model, offering a balance between computational efficiency and accuracy.

The process begins by generating a representative sample of data points from the JCDF using conventional methods, ensuring sufficient coverage of the distribution's variability. These samples serve as the training dataset for the GPR model. The Rational Quadratic kernel is often preferred for its flexibility in adapting to varying data scales, effectively capturing both local and global dependencies within the JCDF structure.

Once trained, the GPR model was used to predict JCDF values for a much larger dataset, enabling high-resolution analysis without incurring significant computational costs. The predictive mean from the GPR provided the JCDF approximation, while the predictive variance quantified the uncertainty of the estimation.

The hyperparameters of the GPR, including the length scale ($l$), variance ($\sigma^2$), and the mixture parameter ($\alpha$) for the Rational Quadratic kernel, were optimized by maximizing the log-marginal likelihood:

$$\log p(y|X) = -\frac{1}{2} y^T [K(X,X) + \sigma_n^2 I]^{-1} y - \frac{1}{2} \log |K(X,X) + \sigma_n^2 I| - \frac{n}{2} \log 2\pi \tag{S10}$$

This optimization ensured that the GPR achieved a balance between model complexity and data fit, preventing overfitting while capturing essential dependencies.

The predictive performance was evaluated using both relative and absolute errors, with the Rational Quadratic kernel achieving the lowest error rates among the tested kernels. Error values consistently remained below 0.05, validating the effectiveness of the GPR in approximating the JCDF.

By leveraging GPR, we significantly reduced the computational burden of high-dimensional copula analysis, enabling efficient calculation of Joint Return Periods (JRP) and critical layers. This approach provided a robust and scalable framework for extreme value analysis in hydrology.

.

## References


Balbi, S., Villa, F., Mojtahed, V., Hegetschweiler, K. T., and Giupponi, C.: A spatial Bayesian network model to assess the benefits of early warning for urban flood risk to people, Natural Hazards and Earth System Sciences, 16, 1323–1337, https://doi.org/10.5194/nhess-16-1323-2016, publisher: Copernicus GmbH, 2016.

Coles, S. G. and Powell, E. A.: Bayesian Methods in Extreme Value Modelling: A Review and New Developments, International Statistical Review / Revue Internationale de Statistique, 64, 119–136, https://doi.org/10.2307/1403426, publisher: [Wiley, International Statistical Institute (ISI)], 1996.

Fisher, R. A. and Tippett, L. H. C.: Limiting forms of the frequency distribution of the largest or smallest member of a sample, Mathematical Proceedings of the Cambridge Philosophical Society, 24, 180–190, https://doi.org/10/bscxvn, publisher: Cambridge University Press, 1928.

Heffernan, J. E. and Tawn, J. A.: A conditional approach for multivariate extreme values (with discussion), Journal of the Royal Statistical Society: Series B (Statistical Methodology), 66, 497–546, https://doi.org/10.1111/j.1467-9868.2004.02050.x, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-9868.2004.02050.x, 2004.

Jenkinson, A. F.: The frequency distribution of the annual maximum (or minimum) values of meteorological elements, Quarterly Journal of the Royal Meteorological Society, 81, 158–171, https://doi.org/10.1002/qj.49708134804, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/qj.49708134804, 1955.

Rasmussen, C. E. and Williams, C. K. I.: Gaussian processes for machine learning, Adaptive computation and machine learning, MIT Press, Cambridge, Mass., 3. print edn., ISBN 978-0-262-18253-9, 2008.