A semi-parametric distribution stitch based on the Berk-Jones test for French daily precipitation bias correction

Philippe Ear, Elena Di Bernardino, Thomas Laloë, Magali Troin, Adrien Lambert

A semi-parametric distribution stitch based on the Berk-Jones test for French daily precipitation bias correction

Ear Philippe[1,2*], Di Bernardino Elena[2], Laloë Thomas[2], Troin Magali[1], Lambert Adrien[1]

[1] Hydroclimat, Aubagne, 13100, France.
[2]Université Côte d'Azur, Laboratoire J.A. Dieudonné, UMR CNRS 7351, Nice, 06108, France.

*Corresponding author(s). E-mail(s): philippe.ear@hydroclimat.com;
Contributing authors: elena.di_bernardino@unice.fr; thomas.laloe@unice.fr; magali.troin@hydroclimat.com; adrien.lambert@hydroclimat.com;

Abstract

Modeling daily precipitation data in a large territory is a complex issue due to its asymmetric distribution with few and spatially sparse extremes. Most parametric distributions fail to model rainfall correctly over a large area, and many impact studies use the non-parametric empirical distribution instead of parametric ones, preferring the robustness of the model on the observed data to the extrapolation to unobserved extremes. In the present paper, we built a distributional semi-parametric model for the bias correction of the ERA5-Land reanalysis using the CERRA-Land reanalysis. The proposed inference procedure is constructed as follows. Firstly, we fit an Extended Generalized Pareto (EGP) distribution to the data. These EGP models give a Generalized Pareto distribution in the upper tail while allowing greater flexibility in the lower one. Secondly, for each location, using an adapted version of the Berk-Jones (BJ) statistical test, we propose to replace a portion of the EGP distribution with either the empirical distribution or an eventually lighter-tail parametric distribution such as the Exponentiated Weibull (ExpW) distribution. The final obtained model is a stitch between the EGP, ExpW and the empirical distributions. The proposed semi-parametric stitch model has been evaluated in a bias correction context against classical pure parametric quantile mapping based on Gamma, ExpW and EGP distributions.

Comparisons to other classical models show a reduction of the mean absolute and extreme error metrics, especially by removing outliers.

Keywords: Bias correction, Extreme value theory, Goodness-of-Fit, Parametric distribution, Precipitation modeling

1 Introduction

1.1 On the bias correction literature: a focus on heavy-tail precipitations

Precipitation is one of the main input variables of hydrological models and processes, where heavy rainfall produces catastrophic events with large economic and human impact (Carrió et al., 2022; Costache and Tien Bui, 2020). Extreme Value Theory (EVT) is well suited for modeling extreme hydrological events, especially when it comes to flood events (Katz, 2002). EVT allows for distributions to fit block maxima data in random samples for fixed intervals and to model heavy-tailed data (more details on EVT can be found in De Haan and Ferreira (2006)). Accurate and high-resolution daily precipitation data, especially on extremes, are crucial to correctly calibrate models, forecast floods and contain natural catastrophes (Alfieri and Thielen, 2015; Sangati and Borga, 2009).
However, most available global or continental datasets do not have a high enough resolution (Soares and Cardoso, 2018) for most local impact studies (Henckes et al., 2018), or to represent correctly the extremes (Prein et al., 2016) due to the presence of significant bias, and need to be downscaled and bias corrected (Şan et al., 2023; Cucchi et al., 2020; Xu et al., 2015).
The most popular method for bias correction is the quantile mapping (Lafon et al., 2013) based on the pure empirical distribution in a univariate context. The non-parametric distribution is widely adopted since it does not require any distributional model assumption. However, the corrected values are limited to the reference-period observed ones, thus new extreme events can not be extrapolated or predicted correctly because they result from arbitrary workarounds (Déqué, 2007). Let $x_{\texttt{obs}}$ be the reference time series, considered to have no bias and $x_{\texttt{mod}}$ a time series produced by a biased model that must be corrected. The empirical quantile mapping method can then be described as:

$$x_{\texttt{mod}}^{corr} = F_{n,\texttt{obs}}^{-1}(F_{n,\texttt{mod}}(x_{\texttt{mod}})), \tag{1.1}$$

where we use the notation $F_{n,\texttt{data}}$ for the empirical cumulative distribution function built on the sample `data` of size n, $F_{n,\texttt{data}}^{-1}$ for the associated quantile function (`data` can be referring to `obs` and `mod` here) and $x_{\texttt{mod}}^{corr}$ is the bias corrected model data.

Parametric distributions have also been studied extensively for daily precipitation modeling, with the Gamma distribution (Martinez-Villalobos and Neelin, 2019), a mixture of Generalized Pareto (GP) distribution (Li et al., 2021) or the Extended

Generalized Pareto (EGP) distribution (Naveau et al., 2016) (see Equation (B.1) for a precise definition). The light-tailed Gamma distribution fails to represent rainfall in locations with heavy-tail precipitations, better fitted for instance by an Exponentiated Weibull (ExpW) or a GP distribution which can both belong in the maximum Fréchet domain of attraction given the appropriate parameters (Blain and Meschiatti, 2015; Vlček and Huth, 2009). In particular, the GP distribution is the classical model used in the peak-over-threshold models (*i.e.*, based on exceedances) (Rootzén and Tajvidi, 1997; McNeil et al., 1997) and is also often used in extreme precipitation and flood modeling (Yue et al., 2022; Acero et al., 2011). However, the need for a data-driven automatic threshold selection makes it very computationally intensive when the area of interest becomes large or with a highly-resoluted spatial data set. The selection of the threshold is the main issue regarding the GP distribution. Setting the threshold to a unique threshold for all locations have also been considered in Langousis et al. (2016) and Mamalakis et al. (2017). A review of multiple threshold calibration methods has been done in Langousis et al. (2016). To overcome this threshold calibration issue, the EGP distribution class has been developed and allows to model both lower and upper heavy tails (Gamet and Jalbert, 2022; Tencaliec et al., 2019; Naveau et al., 2016) without the need to select a threshold. While this distribution class may appropriately fit in some locations, it is not satisfactory everywhere when the study covers a geographically large and highly-resoluted area.

1.2 The proposed semi-parametric stitch method

In the present work, we propose a method to produce a semi-parametric distributional model based on an adapted version of the Berk-Jones (BJ) statistical test (Berk and Jones, 1979; Moscovich et al., 2016). The proposed inference procedure on the considered daily French precipitations will produce a final stitch between the heavy-tailed EGP, the lighter-tailed ExpW and the empirical distribution function. More precisely:

(i) We fit the EGP distribution on each location;
(ii) Based on the adapted BJ statistical test, the upper and lower tails are eventually replaced by the ExpW distribution;
(iii) Finally, if neither parametric distribution was satisfactory enough, the empirical distribution F_n is used.

This mixture of non-parametric and parametric distribution is often called a semi-parametric approach in the literature while combining multiple parametric distributions is referred to as a stitch or a spliced distribution. In this article, we will refer to such distributions as semi-parametric stitch distributions. Multiple study on semi-parametric distributions have been done in the past by Langousis et al. (2016) with an empirical/Generalized Pareto stitch, by Mamalakis et al. (2017) with an Exponential/Generalized Pareto stitch or, more recently, by Holthuijzen et al. (2022) with an empirical/linear stitch. All these methods used a unique stitching threshold for every locations.

Then, in this study, we provide the semi-parametric stitch precipitations using the ERA5-Land and CERRA-L and distributional fitting for Equation (1.1). The derived

bias correction application (see Section 1.3 for details) is mostly an illustration of the performance of the proposed stitch distribution instead of the empirical one. A deeper bias correction comparison, for longer time series, is discussed as an interesting future perspective in Section 5.

The proposed model shows various benefits, which will be further discussed in Section 5.1. i) A parametric distribution is fitted on most locations on the upper tails, which allows for possible extreme extrapolation. ii) The use of the EGP as the main distribution removes the need to find an appropriate threshold as one would need in the case of a GP. iii) Removing the need of a threshold allows for an application over large non-spatial stationary areas of study. iv) Stitch and cutting indices are defined pixel-by-pixel allowing us to adapt to spatially heterogeneous data. v) Rejection using the proposed cutting indices is based on the BJ test which has theoretical guarantees.

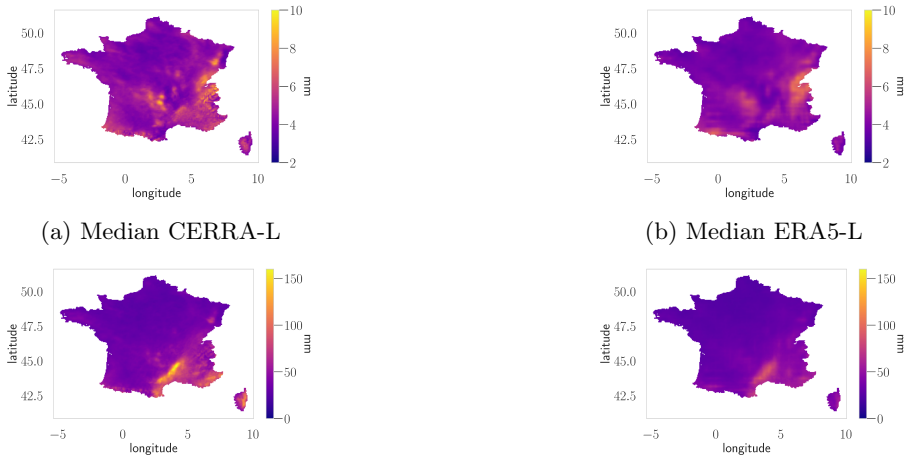
As it will be discussed in Section 5.2, the proposed stitch model faces to several limitations: i) the use of the empirical distribution on some locations' upper tail which result in bias-corrected extremes limited to the range of the observation data and are dependent on the considered period. This empirical upper-tails modeling issue occurs 9.21% and 6.34% of the times for respectively ERA5-L and CERRA-L (for details see Table 1) on the whole considered locations. ii) Our inference methodology assumes the stationarity of the considered time-series data, which is debatable in a climate change context. iii) Finally, in this paper, we only consider wet days, with a threshold at 1 mm.

1.3 Bias correction importance on the considered daily precipitations using the ERA5-Land and CERRA-Land datasets

The whole of metropolitan France is used in this daily precipitation study which covers around 550000 km^2. To perform bias correction and assess the performance of the proposed new method, a pseudo-`obs`ervational and a to-be-corrected-`mod`el dataset are needed, respectively corresponding to $x_{\texttt{obs}}$ and $x_{\texttt{mod}}$ in Equation (1.1). The $x_{\texttt{mod}}$ considered is the ERA5-Land dataset (Muñoz-Sabater et al., 2021), while $x_{\texttt{obs}}$ is the Copernicus regional reanalysis for Europe (CERRA-Land) dataset (Verrelle et al., 2021). Datasets can be downloaded via the link in the *Data availability* section. Both datasets cover Europe on a large range of climate variables and are produced by the European Centre for Medium-Range Weather Forecast (ECMWF). The ERA5-L reanalysis is a global reanalysis with a $0.1° \times 0.1°$ (approx. 11×11 km) resolution, spanning from 1950 to 2021, while the CERRA-L database is a high-resolution reanalysis at 5.5×5.5 km resolution from January 1984 to July 2021. In this study, we considered 36 common years from 01/01/1985 to 31/12/2020. The CERRA-L uses the ERA5 (C3S, 2018) reanalysis for assimilation while ERA5-L is a spatially enhanced version of ERA5 over the dry land areas. The CERRA-L data being on a regular kilometres grid has been interpolated to a regular longitude-latitude grid using the linear scattered interpolant (`scatteredInterpolant`) from `MATLAB` (The MathWorks Inc., 2022). Then, the ERA5-L data has been interpolated to the same

grid as a simple mapping step using `Python`'s library `SciPy CloughTocher2D` interpolator. While the ERA5-L reanalysis has been extensively used for climate studies (Yang et al., 2023; Malaekeh et al., 2022), the CERRA-L dataset is still new and has not seen many published studies based on its data (Monteiro and Morin, 2023), and none, to the best of our knowledge, for precipitation modeling and bias correction.

The CERRA-L reanalysis is chosen here as *ground-truth* for its higher resolution and availability over a large time-period. However, it is almost certain that it is not an unbiased dataset and it only serves as an illustration for the fitting and bias-correction results in Sections 3 and 4. A visual analysis at the median and 99.5th quantile of daily precipitation over the considered period shows the similarity as well as the differences between both reanalysis when it comes to daily rainfall. Figure 1 (a,b) clearly shows similar median precipitations over France with differences mainly focused on high-altitude areas (Massif Central and Alps region).



(a) Median CERRA-L

(b) Median ERA5-L

(c) 99.5th quantile CERRA-L

(d) 99.5th quantile ERA5-L

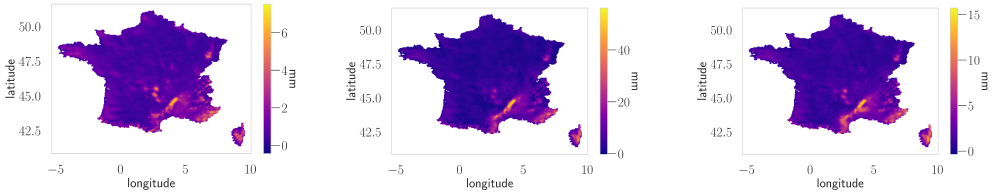
Fig. 1: Median (panels a and b) and 99.5th quantile (panels c and d) daily precipitation in mm over the considered period for CERRA-L reanalysis data (first column) and ERA5-L reanalysis data (second column)

While most of France has a daily median precipitation of around 3 to 5mm per day, the Massif Central region receives close to 10mm of precipitation for CERRA-L shown in Figure 1 (a), but only 7mm for ERA5-L in (b). These differences are much more noticeable when looking at the extremes in the Cévennes region reaching 160mm for CERRA-L in Figure 1 (c,d), while the same region does not exceed 100mm for ERA5-L. The impact of higher resolution over strong orography regions is clear for extreme daily precipitations. Values from ERA5-L are smoothed compared to CERRA-L and interpolating the data to a higher resoluted grid cannot bring relevant variability,

thus the clear need for bias correction.

In this paper, we only consider wet days of the given time series, with a threshold for wet days at 1 mm. This means no correction over the dry days' proportion is done and the corrected data will keep the same number of dry and wet days. The number of wet days differs from point to point, but the range is from 2627 to 7274 days for ERA5-L and 1759 to 7641 days for CERRA-L (for a total of 13149 days).

Since both reanalysis were produced from the ECMWF and are influenced by a common reanalysis ERA5, a first naive approach could be to avoid the bias correction phase, *i.e.*, not apply Equation (1.1). However, this would result in erroneous models and impact studies. To illustrate this error, the Mean Absolute Error (MAE), Mean Absolute Error over the 95th quantile (MAE95sup) and the Root Mean Square Deviation (RMSE) metrics (see Appendix A for details) have been computed without any bias correction (see Figure 2).



(a) MAE (b) MAE95sup (c) RMSE

Fig. 2: Error metrics in mm (see Appendix A) between the CERRA-L and ERA5-L reanalysis without bias correction.

Both maps in Figure 2 show significant bias in the uncorrected interpolated ERA5-L reanalysis. The median MAE is around 0.60mm with peaks at 7.02mm, the median MAE95sup is at 3.43mm with peaks at 55.61mm and for the RMSE, the median error is at 1.16mm with peaks at 15.21mm. Then, bias correction in Equation (1.1) seems to be a necessary step to correctly use those datasets in the impact studies. Based on these considerations, we aim in this paper to build a flexible distributional model to fit both reanalysis precipitation datasets to apply Equation (1.1).

1.4 Structure of the paper

Section 2 describes the proposed semi-parametric distributional model. In Section 3, we apply our method in a case study on the considered reanalysis ERA5-Land (ERA5-L) and CERRA-Land (CERRA-L). Fit and bias correction results are presented in Section 4 with the bias correction of ERA5-L using CERRA-L. Conclusion and key points are discussed in Section 5. The considered metrics are presented in Appendix A and used classical parametric distributions are recalled in Appendix B. Details on the monotonicity correction of the proposed stitch distributional model

are given in Appendix C. Supplementary figures are available in Appendix D.

2 The proposed semi-parametric stitch distributional model

2.1 An adapted goodness-of-fit statistical test

The well-known Kolmogorov-Smirnov (KS) (Massey Jr., 1951), Cramer-von Mises (CvM) (Anderson, 1962) and the Anderson-Darling (AD) (Anderson and Darling, 1954) are the most commonly used statistical tests to evaluate the adequacy of a distribution and eventually discriminate between multiple models (Laio, 2004; Vlček and Huth, 2009). However, the KS and CvM tests tend to underestimate the deviation that occurs in the tail of the distribution (Steinskog et al., 2007) and while the AD test includes a weight function able to increase the sensitivity on the distribution tails and standardizes the variance, the AD test still mishandles the tails (Moscovich et al., 2016). Instead of looking at the deviations of the distributions, the Berk-Jones (BJ) statistical test (Berk and Jones, 1979) uses the most statistically significant deviation. In this paper, we will use an adapted version of the consistent BJ test to determine the best-fitting models for a considered time series. Let us introduce some notations:

- Let $\tilde{x} := \{\tilde{x}_1, ..., \tilde{x}_n\}$ be an $i.i.d$ sample from a random variable X. Let X follow a continuous cumulative distribution function (cdf) F. Then, under the null hypothesis $\mathcal{H}_0 : X \sim F$. The alternative hypothesis will be $\mathcal{H}_1 : X \not\sim F$.
- Let $x := \{x_{(1)}, ..., x_{(n)}\}$ be empirical quantiles of the sample $\tilde{x}$, with $x_{(1)} < x_{(2)} < ... < x_{(n)}$ which are the order statistics of x (we can safely assume that none of the $i.i.d$ samples are equal).
- We also denote the quantile of a distribution F corresponding to the i-th ordered quantile as $F^{-1}(\frac{i-1}{n}) = q_{(i),F}$, with F the candidate cdf fitted on x. Let $q := \{q_{(1),F}, ..., q_{(n),F}\}$.

The random variable $F(X)$ follows a uniform distribution $U[0, 1]$ by the probability integral transform result. Consider $u_{i,F} = F(x_{(i)})$ for $i = 1, \dots, n$ an $i.i.d$ sample from the $U[0, 1]$ distribution and the corresponding order statistics $u_{i,F}$. We know that under $\mathcal{H}_0, u_{i,F}$ is a random sample from u_i which has a known distribution:

$$u_i \sim \text{Beta}(i, n - i + 1), \quad \text{for } i = 1, \dots, n, \tag{2.1}$$

where u_i follows a Beta distribution of parameters i and $n - i + 1$. The BJ test then computes the following p-values:

$$p^*_{i,F} := \mathbb{P}(u_i < u_{i,F}). \tag{2.2}$$

For each i, we can then define:

$$p_{i,F} = \min(p^*_{i,F}, 1 - p^*_{i,F}). \tag{2.3}$$

In Moscovich and Nadler (2017), the M_n statistic is defined by:

$$M_n := \min_{1 \leq i \leq n} p_{i,F} \in [0, 1], \tag{2.4}$$

where small values of M_n describe a bad fit of the candidate distribution F. The p-value associated to α, denoted $p_{\alpha,n}$, is given by $\mathbb{P}(M_n \leq p_{\alpha,n}) = \alpha$. Theoretical properties including the asymptotic consistency of the BJ statistics and algorithms to explicitly compute the p-values in Equation (2.3) can be found in Moscovich et al. (2016); Moscovich and Nadler (2017); Moscovich (2023).

The BJ statistical test can detect deviations in both tails of the distributions. In the case of skewed distributions, such as the ones used for precipitation modeling, most of the weight is concentrated in the lower tail, making it harder to detect deviation from the upper tail. However, heavy upper tail behaviour is a crucial issue if one aims to model extremely large precipitation events. The way the BJ test is constructed allows us to associate each quantile (or rank) with a $p-$value. We can then use this $p-$value to determine if the considered quantile deviation would have been enough to give a rejection or not.

In practice, this translates to not severe enough statistics produced by the extreme deviations, and errors of 50mm or even 150mm may not be considered significant enough by the statistical test. Indeed the rejection threshold for such events, even for $\alpha = 15\%$ may be 10 or 20 times lower than the associated $p_{i,F}$ in Equation (2.3). To increase the sensitivity of the BJ test on the upper tails, we introduce a weight using the prediction error between $x_{(i)}$, the ordered statistic of rank i, and the modelled corresponding quantile from the candidate distribution $q_{(i),F}$. We define the *weighted-level* $k_{i,F}$ such that

$$k_{i,F} := \frac{p_{i,F}}{\max(|x_{(i)} - q_{(i),F}|, 1)}, \qquad \text{for } i = 1, \ldots, n. \tag{2.5}$$

The $p_{i,F}$ assesses the goodness of fit for the rank i. The same consideration could be applied on the weighted-level $k_{i,F}$ in (2.5) on the extremely large precipitation quantile levels. For daily precipitation data, a high prediction error will typically be magnitudes greater than 1mm. Thus, in these cases, the weighted-level as defined in Equation (2.5) will be reduced compared to the corresponding $p_{i,F}$, making it much easier to detect. The weighted-levels are penalizing large deviations in the tails by reducing the corresponding $p_{i,F}$. Such deviations should mostly occur when $\mathcal{H}_0$ is `False`, especially for large enough sample sizes. While no theoretical proof is given for the guarantee of the test significance level or power, we are currently working on a complete and exhaustive simulation study. This ongoing work is out the scope of the present paper and may be discussed in a future work. However, the preliminary results appear to confirm an increase of the test power for almost all sample sizes and without loss of significance.

2.2 Cutting indices for the semi-parametric model

Sometimes, a single parametric distribution is not flexible enough to describe daily rainfall time series on a broad range of locations. The selected model may only be adequate up to (or starting from) a specific quantile and be misfitted otherwise.
To overcome this, we introduce a semi-parametric distribution based on the modified BJ test introduced in Section 2.1 and on three distinct distribution types:

- A main heavy-tailed distribution F_{heavy},
- A secondary lighter-tailed distribution $F_{lighter}$,
- The empirical distribution of the data F_n.

The idea is quite simple: replace the badly fitted portion of the heavy-tailed distribution with the lighter-tailed distribution using the previously adapted BJ test, and finally, if needed, replace the upper and lower tails with the empirical distribution. The proposed method presents the same limitations as the empirical distributions on locations where the latter is used on the upper tail (9.21% and 6.34% of pixels for ERA5-L and CERRA-L data-set respectively, for details see Table 1). Indeed, in this case, no value can be modelled outside the range of the dataset used for the fitting of our method. The range is also heavily dependent on the dataset as discussed in Section 5.
Precipitation modeling via semi-parametric distributions has been explored by a few authors. The most common version is a stitch between a first distribution modeling low and moderate precipitations and a Generalized Pareto distribution for the extremes, also known as spliced distribution (Langousis et al., 2016; Castro-Camilo et al., 2019; Kim et al., 2019). The main issue with semi-parametric and hybrid distribution is the difficult selection of an appropriate threshold. This latter needs to be high enough in the case of the Generalized Pareto so the Pickands–Balkema–de Haan theorem can be applied (see Pickands (1975) and Balkema and Haan (1974)).
To implement our stitch model, we introduce the cutting indices for a distribution F from the $p-$values and weighted values $p_{i,F}$ and $k_{i,F}$ from Equations (2.3) and (2.5), as follows.

Definition 1 (**Cutting indices**). *Let $x_{(i)}$ and $q_{(i)}$ be respectively the empirical and modelled quantiles from a time series as described in Section 2.1. Let $p_{i,F}$ and $k_{i,F}$ be the $p-$values and weighted levels induced from the BJ test as in Equations (2.3) and (2.5). Let $p_{\alpha,n}$ be the rejection threshold at level α for the BJ test for a sample size n, i.e., such that $\mathbb{P}(M_n \leq p_{\alpha,n}) = \alpha$. Let `lag` be a strictly positive natural number.*

1. ***Lower cutting index**. The lower tail is considered adequate if and only if there are no rejections at level $p_{\alpha,n}$ for the next `lag` indices. We define the lower cutting index $i_{\ell,F}$ as equal to 0 if no rejection is detected at level $p_{\alpha,n}$ for the first `lag` ordered quantiles. Conversely, if a rejection is detected, we consider the first rank not rejected at level $p_{\alpha,n}$, such that the next `lag` $p-$values are also not rejected at*

the same level. Formally we can write

$$i_{\ell,F} = i - 1 \begin{cases} \forall\, 0 \leq j < i,\ \exists\, 0 \leq k \leq \texttt{lag}, \text{ such that } p_{(j+k),F} < p_{\alpha,n}, \text{ and} \\ \forall\, 0 \leq j \leq \texttt{lag},\ p_{(i+j),F} \geq p_{\alpha,n}. \end{cases} \tag{2.6}$$

2. ***Upper cutting index*** *The upper tail is considered adequate if and only if there are no rejections at $p_{\alpha,n}$ for the previous* `lag` *indices. More precisely, we define the upper cutting index $i_{u,F}$, as the first index $n-i-1$, for $0 \leq i < n$, satisfying the following two conditions: the corresponding weighted level is not rejected at $p_{\alpha,n}$ and the previous* `lag` *weighted levels are also not rejected at $p_{\alpha,n}$.*

$$i_{u,F} = n - i - 1 \begin{cases} \forall\, i < j \leq n,\ \exists\, 0 \leq k \leq \texttt{lag}, \text{ such that } p_{(j-k),F} < p_{\alpha,n} \text{ and} \\ \forall\, 0 \leq j \leq \texttt{lag}, k_{(n-i-j),F} \geq p_{\alpha,n}. \end{cases} \tag{2.7}$$

An illustrative example is available in Appendix D, see Figure 23. Notice that $i_{u,F}$ can be equal to $n-1$ (resp. $i_{\ell,F}$ can be equal to 0), which indicates that no cut in the upper (resp. lower) tail of the considered distribution is needed.
From the upper and lower cutting indices defined in Definition 1 (see Equations (2.6) and (2.7)), we can list five types of possible rejections, denoted $\texttt{typeRej}_F$ in the following.

Definition 2 (**Types of rejection**).

$\texttt{typeRej}_F == 1$ *No rejection: $i_{\ell,F} = 0$ and $i_{u,F} = n-1$, neither the lower nor the upper tail have been cut and the distribution F is kept as initially fitted.*
$\texttt{typeRej}_F == 2$ *Left rejection: $i_{\ell,F} > 0$, only the lower tail of the distribution has been rejected at some degree and will be replaced.*
$\texttt{typeRej}_F == 3$ *Right rejection: $i_{u,F} < n-1$, only the upper tail of the distribution has been rejected at some degree and will be replaced.*
$\texttt{typeRej}_F == 4$ *Double rejection: $i_{\ell,F} > 0$ and $i_{u,F} < n-1$, both a left and right rejection occurs and the distribution will be replaced for both tails.*
$\texttt{typeRej}_F == 5$ *Total rejection: If the rejection on the upper tail or the lower tail is too strong, we decide to completely reject the distribution. In particular for the upper tail: if $i_{u,F} < l_{upper}$, the whole distribution is rejected. For the lower tail: if $i_{\ell,F} > l_{lower}$ and at least one point is rejected in the upper tail ($i_{u,F} < n-1$), the whole distribution is rejected.*

Notice that l_{upper} and l_{lower} are scalar hyperparameters used to adjust how much of the distribution we replace before we discard it completely.
In the previous Definitions 1 and 2, `lag`, l_{upper} and l_{lower} are all hyperparameters to be calibrated. In our procedure, we evaluate $i_{\ell,F}$ and $i_{u,F}$ in Equations (2.6)-(2.7) and the rejection types $\texttt{typeRej}_F$ in Definition 2 with F being either F_{heavy} or $F_{lighter}$. Secondly, three booleans variables are defined as follows.

Definition 3 (**Rejection booleans**). *Let us define* $r_{\texttt{right}}$ *as a boolean variable which will be* `True` *if we replaced the upper tail of* F_{heavy} *with the upper tail of* $F_{lighter}$ *(a lighter-tailed distribution). Then*

$$r_{\texttt{right}} = (\texttt{typeRej}_{F_{lighter}} == 1 \text{ or } \texttt{typeRej}_{F_{lighter}} == 2) \text{ and } \texttt{typeRej}_{F_{heavy}} \geq 3$$

In this situation, the last $i_{u,F_{heavy}}$ *of* F_{heavy} *quantiles will be replaced with those of* $F_{lighter}$*.*
Let us define $r_{\texttt{left}}$ *as a boolean variable which will be* `True` *if* $F_{lighter}$ *the lower tail of* F_{heavy}*. Then*

$$r_{\texttt{left}} = \begin{cases} (\texttt{typeRej}_{F_{lighter}} == 1 \text{ or } \texttt{typeRej}_{F_{lighter}} == 3) & \text{and} \\ (\texttt{typeRej}_{F_{heavy}} == 2 \text{ or } \texttt{typeRej}_{F_{heavy}} == 4) \end{cases}$$

In this case, the $i_{\ell,F_{heavy}}$ *first quantiles of* F_{heavy} *will be replaced by the corresponding quantiles of* $F_{lighter}$*.*
Let us define $r_{\texttt{all}}$ *as a boolean variable which will be* `True` *if* $F_{lighter}$ *completely replaces* F_{heavy}*. Then*

$$r_{\texttt{all}} = \begin{cases} \texttt{typeRej}_{F_{lighter}} == 1 \text{ and } \texttt{typeRej}_{F_{heavy}} == 4 & \text{or} \\ \texttt{typeRej}_{F_{lighter}} \leq 4 \text{ and } \texttt{typeRej}_{F_{heavy}} == 5 \end{cases}$$

In this case, $F_{lighter}$*'s quantiles will completely replace* F_{heavy}*'s quantiles.*
The resulting model will be called below the Semi-Parametric Stitch Berk-Jones based model (Stitch-BJ model or Stitch-BJ distribution in the rest of this paper).
Combining two distribution functions or stitching their quantile functions should be done carefully to preserve classical properties. An analytical version of the resulting distribution can be found in Appendix C as well as a proposed process to guarantee the monotonicity of the resulting quantile function.
Notice that in the proposed method the heavy-tailed distribution is favoured to the lighter-tailed one in the sense that a better fitted F_{ligher} is only used if F_{heavy} is rejected, with no assessment on the fit error produced.
In the next section, we apply the proposed stitch semi-parametric distributional model to the ERA5-L and CERRA-L reanalysis datasets.

3 Results of the Stitch-BJ model fitted on daily precipitation reanalysis over France

In this section, the proposed Stitch-BJ model is fitted on the ERA5-L and CERRA-L reanalysis. We chose here (see Appendix B for details):

- F_{heavy} = Extended Generalized Pareto distribution (EGPD),
- $F_{lighter}$ = Exponentiated Weibull distribution (ExpWD).

The EGPD is supposed to be able to model both light and heavy-tailed data (Naveau et al., 2016). The ExpWD is an extension of the Weibull distribution and has been

chosen for its ability to be both light or heavy-tailed depending on its shape parameter. Moreover, when its shape parameter is equal to 1, the ExpWD becomes a special case of the EGPD.

When taking the quantile at level 1 of parametric distributions, we actually take the quantile $1 - \frac{1}{n+1}$ with n the number of wet days for the selected location.

Hyperparameters values in Definition 2 are fixed in our case study, both for ERA5-L and CERRA-L, as follows:

$$l_{upper} = \lfloor 0.97 \times n \rfloor, l_{lower} = \lfloor 0.5 \times n \rfloor \text{ and } lag = \lfloor 0.01 \times n \rfloor,$$

where $\lfloor \cdot \rfloor$ is the floor number operator and n is the number of wet days of the given time series. The l_{upper} and l_{lower} hyperparameters are quite stable for our study, with results being similar when small changes were applied to these parameters. Here, l_{upper} is chosen high enough so that only extremes may be corrected. A lower value may lead to a distribution with most of its upper tail being empirical, a too high value may increase too strongly total rejections. We were able to set, l_{lower} quite high since the lower tail are usually well modelled by empirical distributions. Replacing a high proportion of the lower tail will not impact production of extremes or the overall performance of the distribution and will help retain parametric upper tail instead of producing a total rejection. The lag parameter allows to adjust the sensitivity of the detection: a high value will make the detection very sensitive to rejections, since a lone rejection happening far from a tail may cause the whole tail to be rejected. This last parameter has been chosen based on trials not shown here. Those parameters were tuned using as criterion the stability of the produced rejection map (see, *e.g.*, Figure 4). One could aim to keep l_{lower} and lag as low as possible, and l_{upper} as high as possible while getting only small changes in the rejection map for neighboring hyperparameters.

Applications 1, 2 and 3 below describe the steps we implemented to obtain a Stitch-BJ model for the ERA5-L and CERRA-L reanalysis.

Application 1 ***Rejections type and cutting indices for the EGPD***

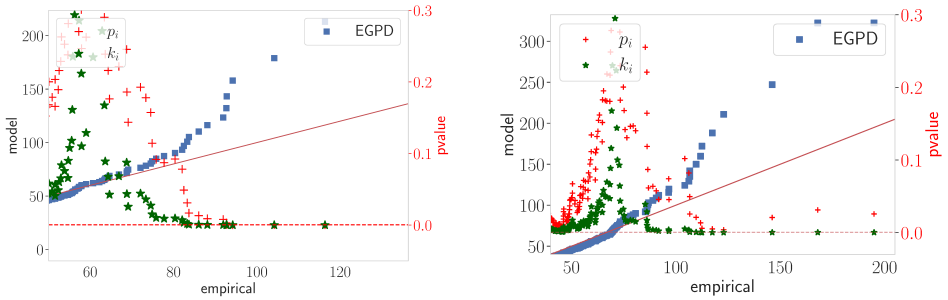
Step (1) Let us consider x and q_{EGP} (see Section 2.1) the empirical quantiles and the EGPD quantiles for a reanalysis data in a specific grid point of the French territory.

Step (2) From Equations (2.2)-(2.3), compute $p^*_{i,EGP}$ and $p_{i,EGP}$, for the considered EGPD in Step (1).

Step (3) Compute the weighted-levels $k_{EGP} = \{k_{i,EGP}, i = 1, ..., n\}$ as in Equation (2.5).

Step (4) Compute the cutting indices and type of rejection $i_{\ell,EGP}, i_{u,EGP}$ and $\texttt{typeRej}_{EGP}$ as described in Definitions 1 and 2.

Steps (2) and (3) of Application 1 with respect to the empirical quantiles x are illustrated in Figure 3. In this figure, p_i's (red crosses) and k_i's (green stars) can be



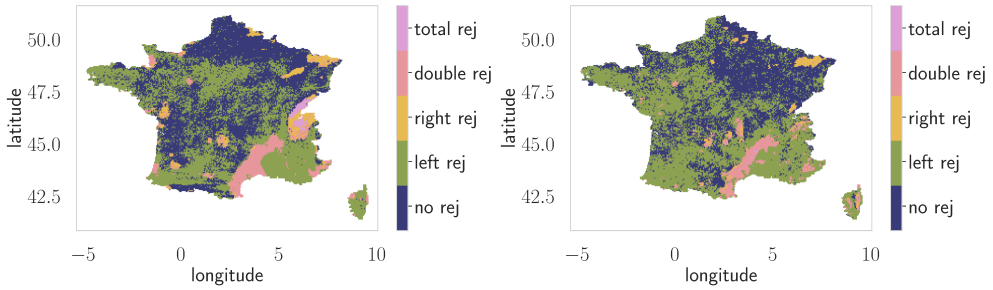
(a) ERA5-L (b) CERRA-L

Fig. 3: Blue squares: Quantile-quantile plot of the EGP model at location (Longitude 2.94, Latitude 42.53). Red crosses: $p_{i,EGP}$; green stars: $k_{i,EGP}$, for $i = 1, \ldots, n$. The horizontal red dashed line is the rejection threshold $p_{\alpha,n}$, for $\alpha = 5\%$. The scale for $p_{i,EGP}$ and $k_{i,EGP}$ is represented in red on the right side

seen as a probability of observing such value at rank i under the considered distribution (here EGPD). The lower their values, the more unlikely it is to observe the given corresponding value at its rank. On both panels of Figure 3, the impact of the weighted-levels $k_{i,EGP}$ in Equation (2.5) on the upper tail is noticeable. For panel (a), the $k_{i,EGP}$ are rejected much faster than the original $p-$values $p_{i,EGP}$, allowing the model to detect misfitted quantiles better. On panel (b), the impact is even more noticeable with the original $p_{i,EGP}$ never crossing the rejection threshold $p_{\alpha,n}$ for the upper tail, while the fit error exceeds 100mm for some quantile points.

On both panels of Figure 4, one can appreciate that the vast majority of locations in the considered French territory are either fully accepted (no rejection) or rejected only on the lower tail (left rejection). Some regions exhibit fit that is not satisfying enough with a high density of rejection of type 4 and 5 (respectively corresponding to a rejection of both tails and a total rejection). Those regions are mainly focused around the Cévennes and Alps regions for the ERA5-L reanalysis and the Cévennes region for the CERRA-L reanalysis. To improve the model on these locations, we replace the misfitted portions of the EGPD with the ExpWD (see Appendix B for details) as described in Application 2 below.

QQ-plots in Figure 5 (panels a and c) show the difference for two chosen locations between the modelled EGP quantiles and the ExpW ones in ERA5-L (panel a) and CERRA-L (panel c) data. Here we decided to replace the upper tail of the EGP with the ExpW one, due to the high errors of the last EGP quantiles. This can specifically be seen on panel (c) where the last quantiles have a fit error of almost 200mm. To illustrate Application 2, in Figure 5 (panels b and d) we also display the corresponding $p_{i,ExpW}$ (red crosses) and $k_{i,ExpW}$ (green stars) in Equations (2.3) and (2.5). We refer to the analysis of Figure 3 for the interpretation of p_i's and $k_i's$. Notice that all $k_{i,ExpW}$ are larger than $p_{\alpha,n}$, meaning that the ExpW upper tail is not rejected and



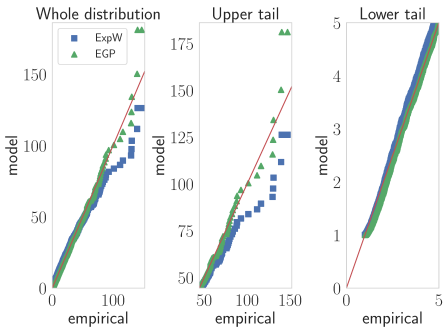
(a) ERA5-L (b) CERRA-L

Fig. 4: Rejection type map for the EGPD over France using $\alpha = 5\%$ for our two reanalysis datasets.

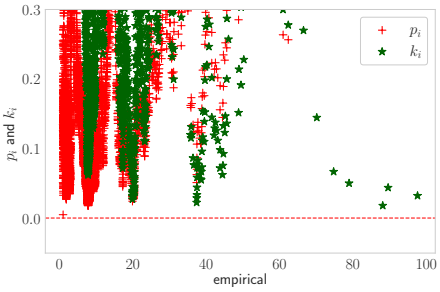
Application 2 Exponentiated Weibull distribution replacement

Step (1) Compute $p_{ExpW}, k_{ExpW}, i_{\ell,ExpW}, i_{u,ExpW}$ and $\texttt{typeRej}_{ExpW}$ as in Application 1 (Step 2 to 4), for the ExpW distribution. Set the boolean variables $r_{\texttt{right}}, r_{\texttt{left}}$ and $r_{\texttt{all}}$ as in Definition 3.
Step (2) If $r_{\texttt{right}}$ is `True`, replace the last $n - 1 - i_{u,EGP}$ quantiles by the corresponding ExpW quantiles.
Step (3) If $r_{\texttt{left}}$ is `True`, replace the first $i_{\ell,EGP}$ quantiles by the corresponding ExpW quantiles.
Step (4) If $r_{\texttt{all}}$ is `True`, replace totally by the ExpW distribution.

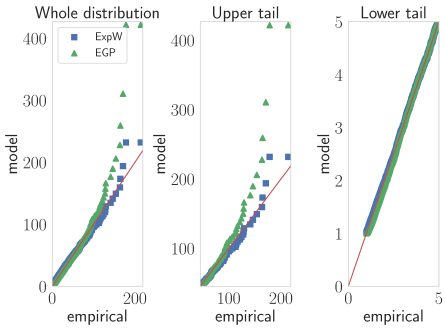
will replace the EGP upper tail for these locations.



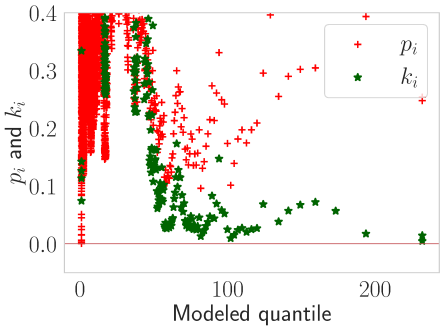
(a) ERA5-L at (Longitude 6.70, Latitude 46.29)



(b) ERA5-L at (Longitude 6.70, Latitude 46.29)



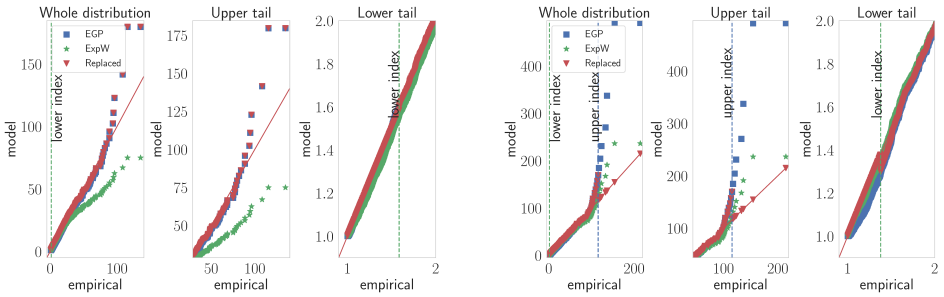
(c) CERRA-L at (Longitude 9.42, Latitude 42.53)



(d) CERRA-L at (Longitude 9.42, Latitude 42.53)

Fig. 5: First row: ERA5-L location. Second row: CERRA-L location. First column: QQ-plot of respectively the whole, upper (above 97th quantile) and lower tail of the EGP (green triangles) and ExpW (blue squares) against the empirical data. Second column: corresponding $p_{i,ExpW}$ (red crosses) and $k_{i,ExpW}$ (green stars). Red horizontal line is the rejection threshold $p_{\alpha,n}$ for $\alpha = 5\%$

While most locations can be modelled via Application 2 using only parametric distributions (see for instance locations in Figure 5), for some locations with complex and particularly skewed distributions, we chose to use the empirical distribution to replace the misfitted parts, as detailed in Application 3 below.



(a) ERA5-L at (Longitude 9.18, Latitude 42.39)

(b) CERRA-L at (Longitude 2.61, Latitude 42.67)

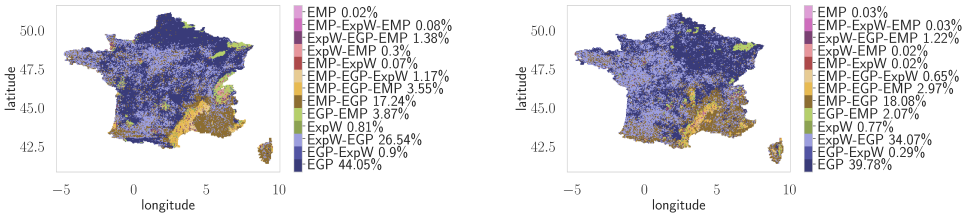
Fig. 6: QQ-plot of the EGP model (blue square), ExpW model (green stars) and the EGP with empirical stitching model (red triangle). Green (resp. blue) dashed line is the lower (resp. upper) cutting index $i_{\ell,EGP}$ (resp. $i_{u,EGP}$) as in Definition 1 for the EGP model.

Application 3 Empirical distribution replacement

Step (1) If $r_{\texttt{left}}$ is `False` and $i_{\ell,EGP} > 0$, replace the first $i_{\ell,EGP}$ quantiles by the empirical ones.
Step (2) If $r_{\texttt{right}}$ is `False` and $i_{u,EGP} < (n-1)$, replace the last $n - i_{u,EGP} - 1$ quantiles by the empirical ones.
Step (3) If $r_{\texttt{all}}$ is `False` and $\texttt{typeRej}_{EGP} == 5$, replace totally by the empirical distribution.
Step (4) If $r_{\texttt{all}}$ is `True` and $i_{\ell,ExpW} > 0$, replace the first $i_{\ell,ExpW}$ quantiles by the empirical ones.
Step (5) If $r_{\texttt{all}}$ is `True` and $i_{u,ExpW} < (n-1)$, replace the last $n - i_{u,ExpW} - 1$ quantiles by the empirical ones.

In Figure 6, neither the EGP nor the ExpW distribution seems to be able to model the upper tail and the lower tail correctly. In this case, the empirical distribution will be used to model both tails on this specific location. Indeed both the ExpW and EGP distributions meet the conditions for a double rejection as defined in Definition 2. In Figure 6 we show the lower and upper index $i_{\ell,EGP}$ and $i_{u,EGP}$. By using Application 3, the final stitch model with the lower and upper tails replaced by the empirical distribution is called `Replaced` in Figure 6 (red triangles).
Moreover, Figure 6 illustrates cases where the monotonicity of the resulting final quantile function may be broken. For instance here, at $i_{u,EGP}$, the last EGP quantile used is at around 180mm while the next quantile taken from the empirical distribution is at around 110mm (see Figure 6 panel b, upper tail QQ-plot). A break in monotonicity can also be detected on the lower tail at the $i_{\ell,EGP}$ cutting point (see Figure 6 panel b, lower tail QQ-plot). We propose a simple method to correct breaks in the monotonicity, which is presented in Appendix C.

Figure 7 shows the final and detailed stitch combination using Applications 1, 2 and 3, on every location for ERA5-L and CERRA-L reanalysis dataset.



(a) ERA5-L (b) CERRA-L

Fig. 7: Final map of the Stitch-BJ model with the EGP (EGPD) and ExpW (WEIB) and empirical (EMP) distributions. We display the proportion of locations for each combination of models for both considered datasets

For both reanalyses, we can see that the vast majority of points are modelled using fully parametric distributions, either with no modification or using a stitch. Interestingly, for ERA5-L and CERRA-L datasets, a few locations (respectively less than 9% and 6.3%) use the empirical distribution to correct the upper tail and less than respectively 0.02% and 0.03% use the empirical distribution to fully model the location's precipitation distribution. For almost 75% of the locations, a full parametric distribution has been used for both datasets.

Table 1: Summary of empirical and parametric distributions used for the Stitch-BJ over France

	EMPIRICAL			
	LOWER TAIL	**UPPER TAIL**	**BOTH TAILS**	**FULL**
ERA5-L	18.48 %	5.56 %	3.63 %	0.02 %
CERRA-L	18.75 %	3.31 %	3.00 %	0.03 %
	PARAMETRIC			
	LOWER TAIL	**UPPER TAIL**	**BOTH TAILS**	**FULL**
ERA5-L	5.56 %	18.48 %	0 %	72.30 %
CERRA-L	3.31 %	18.75 %	0 %	74.91 %

A more precise description of the number of locations where the lower (resp. upper) tail has been replaced by an empirical (resp. parametric) distribution is available in Table 1. Note that the proportion of lower and upper tail replacement locations by the empirical or parametric distributions only concern *strictly lower or upper replacements*. If a location has been replaced on both tails or fully by the empirical distribution, it is not taken into account in the lower and upper tail replacement proportion. This explains why one needs to add the proportion of lower tail and upper

tail replacement of either the parametric or empirical distribution to the proportion of both tail replacement and total replacement of both types of distributions to reach 100%. To complete the information provided in Table 1, the distributions (via boxplots) of the cutting indices are available in Appendix D.

In the next section, we discuss the results of the proposed fitting procedure (Applications 1, 2 and 3) on both datasets (see Section 4.2). Then we use it in a bias correction context (see Section 4.3).

4 Results

In this section, we evaluate the fit and the bias correction performance of the proposed Stitch-BJ and we compare it with several classical parametric competitors. For the fitting results (see Section 4.2), distributions are fitted to the wet days time series separately for ERA5-L and CERRA-L data. For the bias correction results (see Section 4.3), we use both ERA5-L and CERRA-L fitted distributions to bias correct ERA5-L time series. Note that the distributions are fitted on the whole period (01/01/1984-31/12/2020) for both datasets as a simple application of a bias correction using the quantile mapping method. Using Equation (1.1), we identify:

- $F^{-1}_{n,obs}$ is the quantile function of the considered distribution fitted on CERRA-L.
- $F_{n,mod}$ is the cdf of the considered distribution fitted on ERA5-L
- x_{mod} is a wet-day time series of ERA5-L.

After bias correction in Equation (1.1), quantiles of $x^{corr}_{\texttt{mod}}$ (the bias-corrected time series) are compared to the quantiles of $x_{\texttt{obs}}$ (the corresponding wet-days time series from CERRA-L).

4.1 Instructions for figures interpretation

In the following, multiple figures illustrate the performance of the proposed method (for fit and bias correction) compared with more classical parametric distributions. Four types of figures are shown:

1. Error maps for a specific metric for considered models.
2. Error differences maps between the proposed method and considered models for a specific metric.
3. Error boxplots for the considered models for a specific metric.
4. QQ-plots for some selected locations.

Type 1 Figures 8, 11, 14 and 18 are the error maps for a given metric. The error is evaluated between the empirical and the fitted quantiles for the fit section, and between the corrected ERA5-L and CERRA-L quantiles for the bias correction section. We limited the values range to reach the 99.9th quantile of the error produced by our Stitch-BJ model for comparison.

Type 2 Figures 9, 12, 15 and 19 show the differences between the Stitch-BJ and the other distributions for a given metric. Contrary to Type 1 figures, the error range

is not fixed and differs for each sub-figure, thus covering the whole range of error differences for each distribution. Here, a positive (resp. negative) value indicates that the competitor distribution performs better (resp. worst) than the Stitch-BJ one. Note that the error range is not symmetric.

Type 3 Figures 10, 13, 16 and 20 are boxplots of Type 1 figure results. To show both the lower and the extreme errors, the y-axis is represented in log scale. The horizontal dashed line represents the median error for the Stitch-BJ method.

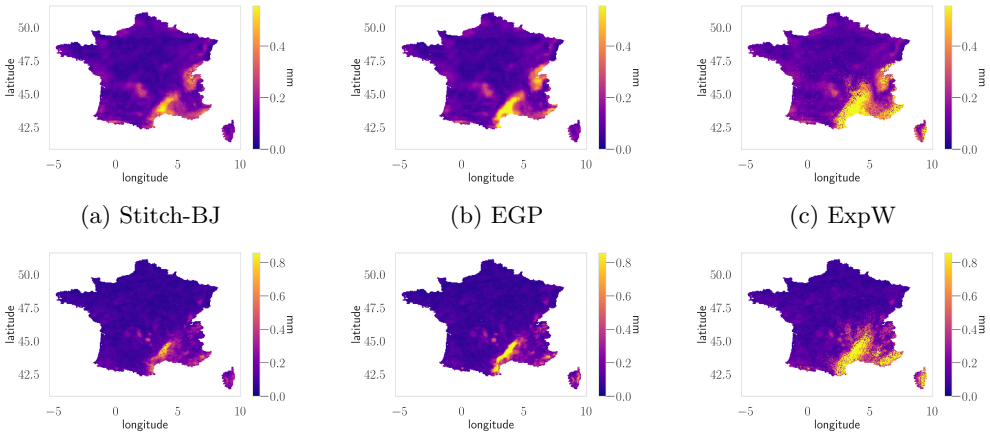
Type 4 Figure 17 are QQ-plots of some selected locations, showing the quantile-per-quantile performance of each method compared to the ground truth. For all Type 4 figures, blue squares represent time series modelled from the Stitch-BJ method, the orange triangles from the EGPD and green crosses from the ExpWD. The red line corresponds to the empirical quantile line. For the fit results, Type 4 figures also include a short description of the resulting Stitch-BJ model by giving the selected distributions of the proposed stitch. For example: `EMP-EGP-ExpW` means the lower tail uses the empirical distribution, the centre is modelled by the EGPD and the upper tail uses the ExpWD.

4.2 Fit results

As explained previously, bias correction requires the same distribution to be fitted twice: once on the starting, biased data (ERA5-L), and once on the target data (CERRA-L). While the two reanalyses show some similarities, differences between the resulting fitting procedure described in Section 3 are expected. We will study the fit for the Stitch-BJ, EGP, ExpW and Gamma distributions globally over France, and provide a local analysis into some selected locations. For all parametric distributions, the parameters are estimated using the associated maximum likelihood estimators. Both datasets used a daily 1mm threshold to remove the drizzle effect (Chen et al., 2021) and distributions are solely fitted on wet days precipitations. A location parameter was available for all distributions except the EGPD. A left shift of 1mm has been applied when fitting to the EGPD (and constructing the Stitch-BJ) to compensate for that. The shift is reintroduced when the quantile function of these distributions is used. For the EGPD, a 3mm left censor has been used. This value was chosen after testing and resulted in the best performance for our application. The interested reader is also referred to Appendix B for details.

4.2.1 Spatialized metrics

Figure 8 display the MAE maps (see Appendix A) for the distributions fitted on respectively ERA5-L and CERRA-L.



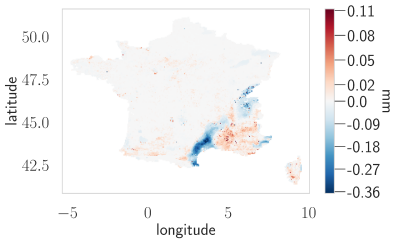
(a) Stitch-BJ (b) EGP (c) ExpW

(d) Stitch-BJ (e) EGP (f) ExpW

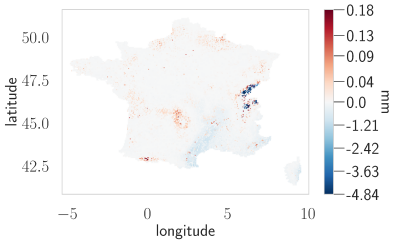
Fig. 8: MAE of considered models fitted to ERA5-L (first row) and CERRA-L (second row), over Metropolitan France

For the ExpW model, the errors' spatial distribution is quite similar to the EGP one, with high errors being more prominent in the Cévènnes and around the city of Nice. In Figure 8, the Stitch-BJ shows differences mainly focused around the Cévènnes and Alps region where it has an improvement with few locations exceeding the 0.5mm MAE for ERA5-L (panels (a,b)) and 0.8mm for CERRA-L (panels (d,e)). Notice that most of the Cévènnes region exceeds this threshold for the EGPD and ExpWD (panels b, c, e and f). The Gamma distribution map is not shown as its errors are large over most of France (see boxplots in Figure 10).

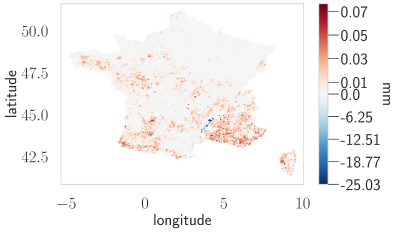
By analyzing the differences maps in Figure 9, the EGPD seems to perform better than the Stitch-BJD in specific locations, but when looking carefully at the actual improvement, the MAE is only slightly reduced. Conversely, the improvements of the Stitch-BJD over both the EGPD and ExpWD (in particular around the Cévènnes region) are large in magnitude.



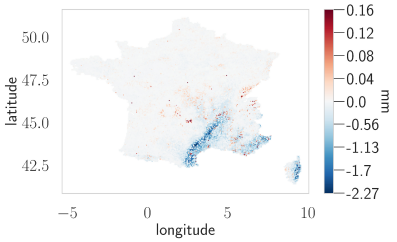
(a) EGPD difference



(b) ExpWD difference



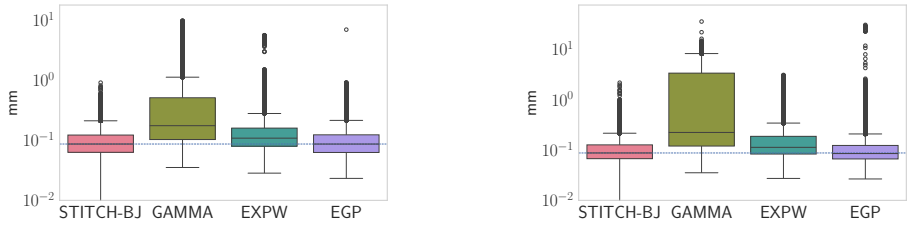
(c) EGPD difference



(d) ExpWD difference

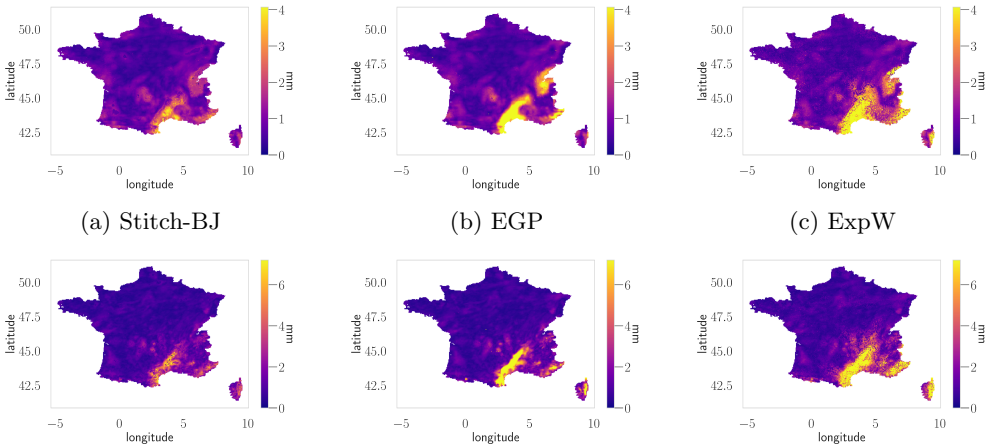
Fig. 9: MAE differences of considered models against the Stitch-BJ distribution fitted to ERA5-L (first row) and CERRA-L (second row) over Metropolitan France. Note that scale are different in each map to highlight the asymmetry in the values' range.

In the MAE in Figure 10, the impact of the Stitch-BJ method on extreme outliers are clear. The ExpW produces less extreme outliers than the EGP on the CERRA-L data, however, the median error and the spread are worse compared to both the EGP and the Stitch-BJ. The median error of the EGP and Stitch-BJ are similar which is expected since the proposed inference procedure keeps as much EGP model in the final Stitch-BJ one as possible. Note that for both boxplots, the y-axis lower limit has been cut at 10^{-2} for readability. These lower values are caused by some sparse locations where the empirical distribution replaces a large part (or fully) the parametric distribution considered, leading to near-zero errors.



(a) ERA5-L (b) CERRA-L

Fig. 10: Boxplots of MAE for considered models fit on ERA5-L and CERRA-L. Note that the y-axis is in log scale.

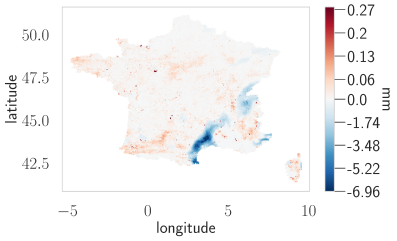


(a) Stitch-BJ (b) EGP (c) ExpW

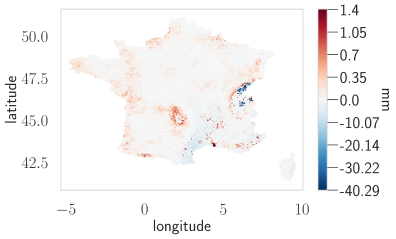
(d) Stitch-BJ (e) EGP (f) ExpW

Fig. 11: MAE95sup of considered models fitted to ERA5-L (first row) and CERRA-L (second row) over Metropolitan France

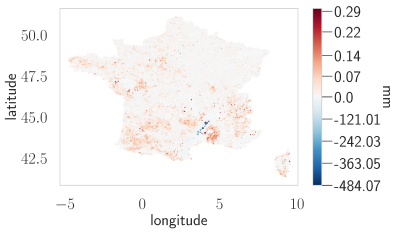
In Figure 11, we display the errors on the upper tail (See Appendix A). The improvements of the Stitch-BJ are further confirmed with a noticeable improvement over the Cévennes compared to all the other distributions, where the colour is not saturated compared to the EGPD and ExpWD for both ERA5-L and CERRA-L data.



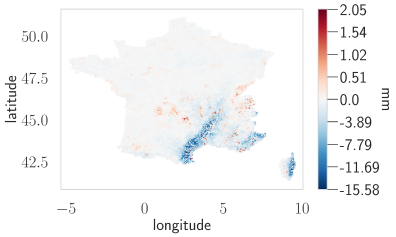
(a) EGPD difference



(b) ExpWD difference



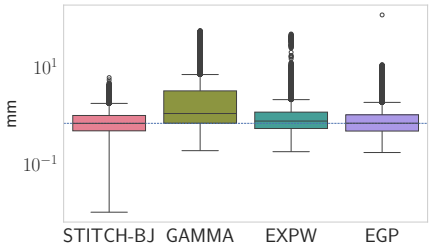
(c) EGPD difference



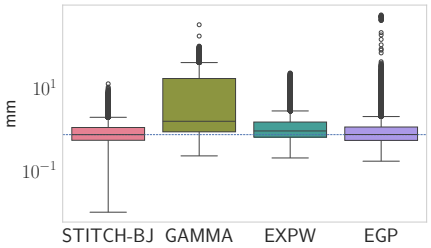
(d) ExpWD difference

Fig. 12: MAE95sup differences of considered models against the Stitch-BJ distribution fitted to ERA5-L panels (a,b) and CERRA-L panels (c,d) over Metropolitan France. Note that scale are different in each map to highlight the asymmetry in the values' range.

Looking at Figure 12, the conclusions are similar to Figure 9 with the error range difference being even greater. The improvements by the Stitch-BJ can also be explained by the use of the empirical distribution on locations where the EGP produces inaccurate values.



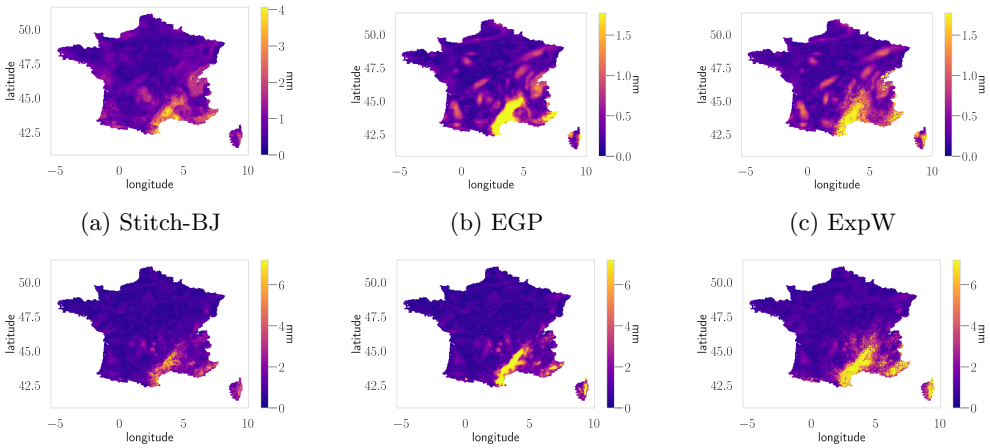
(a) ERA5-L



(b) CERRA-L

Fig. 13: Boxplot of MAE95sup for considered models fit on ERA5-L and CERRA-L. Note that the y-axis is in log scale.

Analogously to Figure 10, we can observe the ability of Stitch-BJ to correct extreme outliers locations errors in boxplots of Figure 13 with a maximum MAE95sup being reduced from 120mm to around 7mm for the models on ERA5-L, and 200mm to 14mm for the models on CERRA-L compared to the EGP. The median for the MAE95sup remains unchanged between the EGP and Stitch-BJ model and the ExpW is noticeably worse than the two previous distributions, especially on the CERRA-L data.

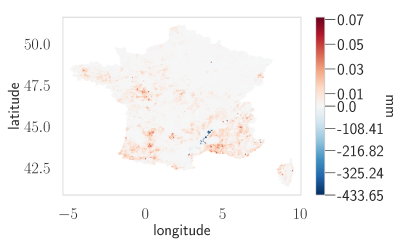


(a) Stitch-BJ (b) EGP (c) ExpW

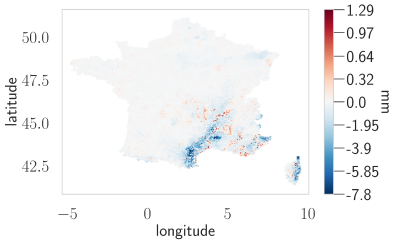
(d) Stitch-BJ (e) EGP (f) ExpW

Fig. 14: RMSE of considered models fitted to ERA5-L (first row) and CERRA-L (second row) over Metropolitan France

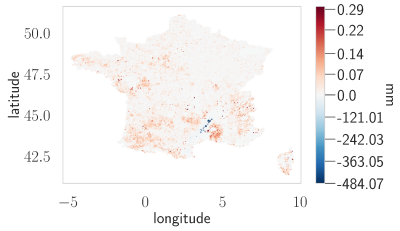
In Figure 14, we display the RMSE (see Appendix A) for the considered models, which adds a quadratic penalty to high errors. Some improvements of the Stitch-BJ are visible with a patchy pattern close to the replacement pattern (see the final map of the Stitch-BJ model in Figure 7). Most of extremes errors around the Cévènnes and PACA regions seem to be contained with the Stitch-BJ model, with respect to the other two chosen parametric models (see panels b, c, e and f in Figure 14). Notice that, for the CERRA-L dataset in Figure 14, some regions around the Cévènnes, which were not rejected by the Berk-Jones test, still show high errors.



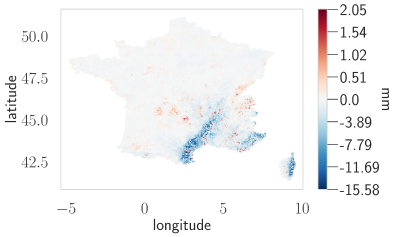
(a) EGPD difference



(b) ExpWD difference



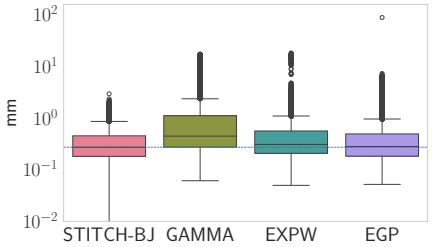
(c) EGPD difference



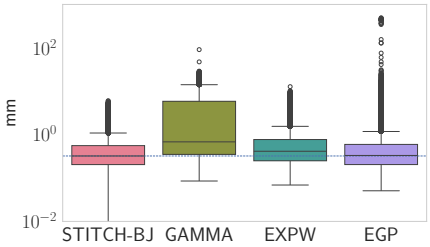
(d) ExpWD difference

Fig. 15: RMSE differences of considered models against the Stitch-BJ distribution fitted to ERA5-L (first row) and CERRA-L (second row) over Metropolitan France. Note that scale are different in each map to highlight the asymmetry in the values' range.

Similarly to Figure 12, in Figure 15 we spatially display the RMSE error ranges. This shows how the RMSE penalizes higher errors, thus producing similar maps to an extreme's focused error such as the MAE95sup.



(a) ERA5-L



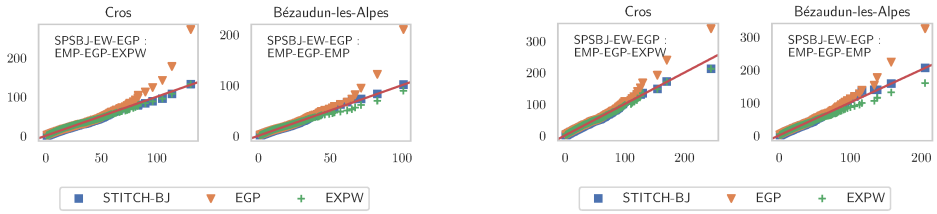
(b) CERRA-L

Fig. 16: Boxplot of RMSE for considered models fit on ERA5-L and CERRA-L. Note that the y-axis is in log scale.

As in Figure 13, boxplots in Figure 16 show the Stitch-BJ seems to realise less extreme outliers than other consider parametric models, and produces lower median error. For instance, the Stitch-BJ's maximum error for ERA5L (resp. CERRAL) is at 2.92mm (resp. 6.03mm) while the EGP's is at 86.68mm (resp. 490.74 mm). Note that for both boxplots from Figure 13, the y-axis lower limit has been cut at 10^{-2} for readability.

4.2.2 Detailed analysis on selected locations

For a more precise study, we selected two locations in the Provence-Alpes-Côte d'Azur region to illustrate various study cases. The selected locations are near Cros (Longitude 3.80, Latitude 44.00) and Bézaudun-les-Alpes (Longitude 7.08, Latitude 43.82).



(a) ERA5-Land (b) CERRA-Land

Fig. 17: QQ-plots of the Stitch-BJ (blue squares), EGP (red triangles) and ExpW (green crosses) fitted on ERA5-L and CERRA-L in the two selected locations

In Figure 17, the QQ-plots for the Stitch-BJ, ExpW and EGP models are shown, fitted for both ERA5-L and CERRA-L, for two different locations. For both ERA5-L and CERRA-L, the first location (Cros) shows a case where the upper tail is replaced by the ExpW, while the EGP is only used for the mid/high precipitation values. The other location, Bézaudun-les-Alpes, is a location where both parametric distributions were unable to correctly model the upper tail which result in the use of the empirical distribution for the extremes. While the ExpW exhibits visually a good fit for the upper tail on this location for both ERA5-L and CERRA-L, the deviation is enough to produce a rejection for the modified BJ test. For both locations and both datasets, neither parametric distribution was able to correctly model the lower tail, therefore it has been modelled using the empirical distribution.

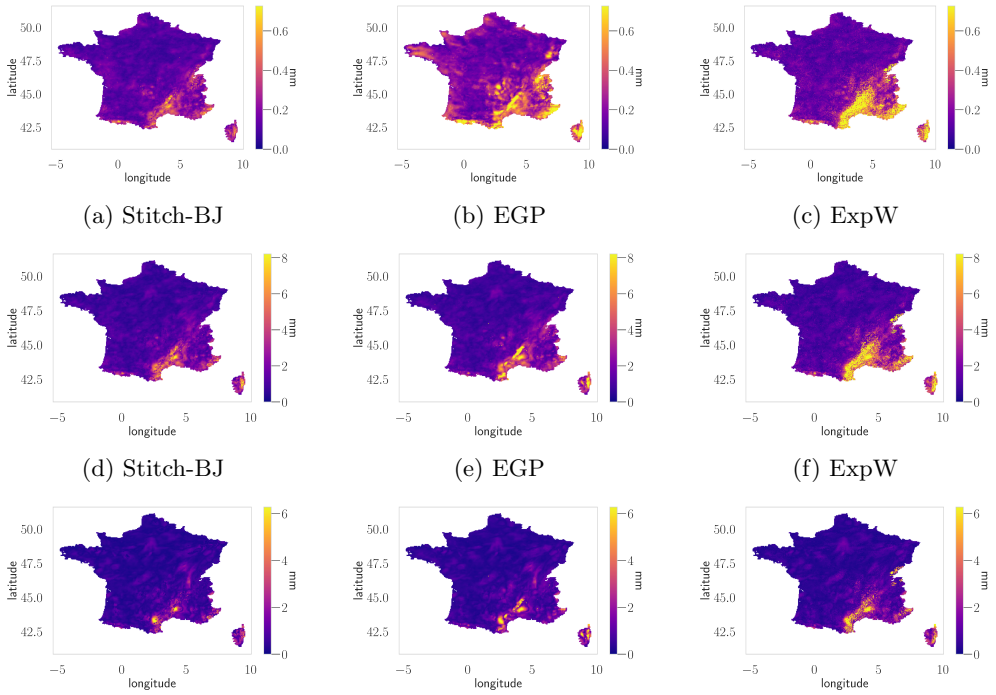
4.3 Bias correction results

After fitting the Stitch-BJ and the other considered distributions on both ERA5-L and CERRA-L, we can now compare the performance of bias correction of the ERA5-L reanalysis against the CERRA-L reanalysis for the same period.

4.3.1 Spatialized metrics

Figure 18 shows metrics (MAE, MAE95sup and RMSE on each row respectively) for the bias correction error of ERA5-Land using CERRA-L. For these three considered

metrics, the Stitch-BJ produces the least extremes, both in number (spatially) and in intensity. The highest MAE is divided by 3 when comparing the EGPD and the Stitch-BJ model while the highest MAE95sup is improved from over 274mm (resp. 62mm) for the EGPD (resp. ExpWD) to 12mm with the Stitch-BJ model. The maximum RMSE for the Stitch-BJ is at 9.27mm while it is at 31.29 and 136.87 for respectively the ExpW and EGP distributions. Two regions are producing most of the errors for the Stitch-BJ around the Cévènnes, which are also present in both the EGP and ExpW maps.



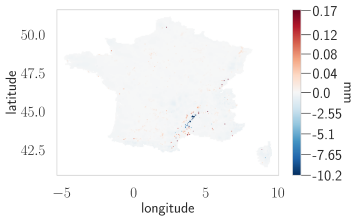
(a) Stitch-BJ (b) EGP (c) ExpW

(d) Stitch-BJ (e) EGP (f) ExpW

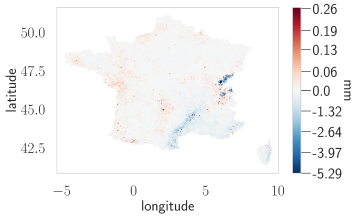
(g) Stitch-BJ (h) EGP (i) ExpW

Fig. 18: MAE (first row), MAE95sup (second row) and RMSE (third row) of considered models for the bias correction of ERA5-L using CERRA-L over Metropolitan France.

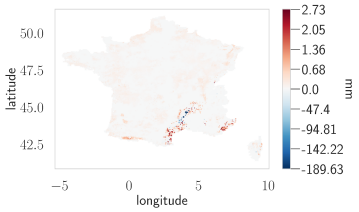
Figure 19 allow us to appreciate the impact of the correction spatially. Most of the Stitch-BJ improvements are located on the Cévènnes and Alps region compared to respectively the EGP and ExpW distributions.



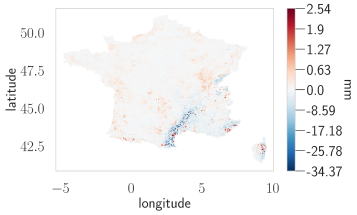
(a) EGPD MAE difference



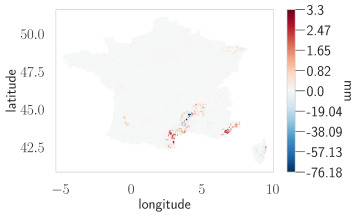
(b) ExpWD MAE difference



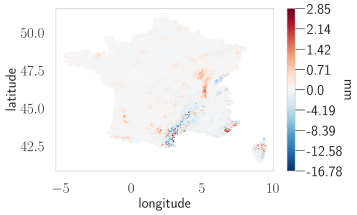
(c) EGPD MAE95sup difference



(d) ExpWD MAE95sup difference

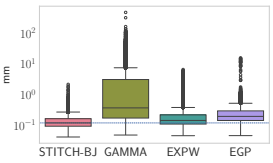


(e) EGPD RMSE difference

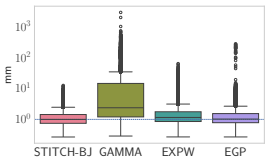


(f) ExpWD RMSE difference

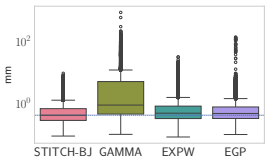
Fig. 19: MAE (first row), MAE95sup (second row) and RMSE differences (third row) of considered models against the Stitch-BJ distribution for the bias correction of ERA5-L using CERRA-L over Metropolitan France. Note that scale are different in each map to highlight the asymmetry in the values' range.



(a) MAE



(b) MAE95sup



(c) RMSE

Fig. 20: From left to right: boxplots of MAE, MAE95sup and RMSE for the bias corrected ERA5-L over CERRA-L for the considered models. Note that the y-axis is in log scale.

Boxplots in Figure 20 confirm the previous observations, with both extremes and median MAE being reduced compared to the other tested distributions. Median MAE95sup is not noticeably improved compared to the EGPD, however extreme outliers are mostly corrected. For the RMSE, both outliers and median errors are lower for the Stitch-BJ when compared to the other considered models. Interestingly, while the ExpW model performed worse than the EGP one in terms of median error in the fit context, it actually outperforms the latter in the bias correction context for the MAE. After bias correction, the maximum MAE (resp. MAE95sup and RMSE) of the Stitch-BJ is more than 3 times (resp. 5 and 3 times) lower than the next model's highest MAE (resp. MAE95sup and RMSE).

5 Conclusion and discussion

In this study, we introduced a semi-parametric stitch distribution based on the BJ statistical test for daily precipitation modeling, which can automatically deduce cutting indices and stitch our quantile functions and cdf. We then compared its fit and bias correction performance to the EGP, ExpW and Gamma distribution over the ERA5-L and CERRA-L reanalysis on France. We first conducted a short comparison of the ERA5-L and CERRA-L reanalysis in order to justify the need for bias correction. Our semi-parametric stitch distribution allows for a better representation of the extremes, either by using another parametric distribution, or by using the best known data available i.e. the empirical distribution.

The Stitch-BJ distribution is constructed in 4 steps : (i) cutting indices and rejection types are inferred for the EGPD, (ii) same for the ExpWD and replace portions of the distribution accordingly, (iii) if neither parametric distributions are satisfactory, use the empirical distribution and (iv) classical properties (for instance monotonicity) of the cdf and quantile function are restored if they've been broken by the previous steps. Fit results shows that the model is able to find the right combination of distributions, resulting in improvements over the EGP or ExpW model on the extremes, against a trade-off over lower and medium precipitations value. In a bias correction context, the Stitch-BJ method reduced the median MAE by respectively 20% and 40% compared to the ExpWD and EGPD, and the maximum MAE by respectively 67% and 86%. The median MAE95sup was reduced by respectively 14% and 3.5%, while the maximum MAE95sup was reduced by respectively 80% and 95%. Finally, the maximum RMSE was reduced by 70% and 93% respectively, while the median RMSE was reduced by respectively 14% and 12%.

We conclude that the proposed Stitch-BJ distribution is able to improve the extremes against all the parametric distributions tested in this paper, making it a potential contender for rainfall modeling.

However, these results should be considered with care as the use of the empirical distribution on the tails or for the whole distribution in some locations may be responsible for the largest improvements seen. However, as we have seen in Figure 7 and Table 1, the empirical distribution is used in most around 9% of locations in the upper tail, and the replacement is mostly for the last 2 or 3% of the distribution (see Appendix D).

5.1 Benefits

The Stitch-BJ model presented in this study has various benefits compared to more classical parametric distributions, or empirical distributions. Being mainly a fully parametric distribution on its upper tail (90.78% on ERA5-L and 93.66% on CERRA-L, see Table 1), it keeps the ability to extrapolate values on the majority of the study area. Moreover, the main distribution used is the EGP from Naveau et al. (2016), which removes the need to find an appropriate threshold from the GP model, while still keeping the theoretical guarantees on the upper tail. While various methods exist to calibrate this threshold automatically (see Langousis et al. (2016) and Scarrott and MacDonald (2012)), computational performances can be poor which prevents their use over a large study area. Moreover, the stitch and cutting indices being tailored for each pixel, add another layer of flexibility to fit the spatially heterogeneous precipitation distributions. The stitching is also based on the BJ goodness-of-fit test which provides theoretical guarantees, and the modified test was assessed using a simulation study compared to other semi-parametric alternatives that often trades theoretical robustness for reduced complexity (Mamalakis et al., 2017; Trentini et al., 2023).
Thus, the Stitch-BJ can be seen as a compromise between empirical robustness and parametric extrapolation by allowing as much of the model to be parametric as possible before reverting to the empirical model in case of misfitting as well as allowing to combine parametric models when multiple candidates are available.
While the Stitch-BJ has only been applied for daily precipitation in this study, application on climate variables such as temperature or wind speed, or in other fields such as finance or insurance could provide some interesting results in future studies. The methodology is not limited to daily precipitation, though a need for scaling may be needed depending on the variable used.

5.2 Limitations

This work allows us to develop a flexible semi-parametric model for daily precipitation modeling, but many improvements and applications are yet to be done, from applications in a climate change context to modeling other climate variables using different distributions. Indeed, we used the empirical distribution when none of the proposed parametric distributions was able to model the data with enough precision because it is the best-performing model in a bias correction context when considering the same time-period. However, the performance in a climate change context may vary. This opens interesting perspectives to perform a proper comparison of bias correction performance between parametric and empirical quantile mapping in a climate change context, using long enough time series to be able to split them. Moreover, the study was not done using out-of-sample data for the validation process which may skew the results. However, recent work under review using out-of-sample data and seasonal separation shows similar behavior and confirm its applicability even with out-of-sample data.. One may also want to completely remove the implication of the empirical distribution by involving more parametric distributions, or lessen its dependence on potential outliers by using a linear correction as in (Holthuijzen et al.,

2022). The question of stationarity can be addressed by using a non-stationary bias correction method such as the CDFt (Michelangeli et al., 2009) or the Quantile Delta Mapping (Cannon et al., 2015), or by introducing a time dependence on the parametric distributions parameters. In future investigations, data may also be separated by months or seasons to take into account non-stationarity inside the sample data, as in Francois (2022). Finally remark that only a reduced set of concurrent distributions were used as a comparison in the present study, thus performance against other fully parametric (Um et al., 2016; Mamalakis et al., 2017) or semi-parametric distributions (Langousis et al., 2016; Tencaliec et al., 2019; Holthuijzen et al., 2022) is needed in a deeper future study.

Data availability: All data, material, and programming codes used in this study are available upon request. ERA5-Land and CERRA-Land datasets analyzed in the current study are available on the Copernicus Climate Change Service (C3S) Climate Data Store.

Acknowledgments: This work has been partially supported by the French government through the CIFRE funding (CIFRE n°2022/0519) managed by the National Association of Research and Technology and the 3IA Côte d'Azur Investments in the Future project managed by the National Research Agency (ANR-19-P3IA-0002).

References

Anderson, T.W., Darling, D.A.: A Test of Goodness of Fit. Journal of the American Statistical Association **49**(268), 765–769 (1954) https://doi.org/10.1080/01621459.1954.10501232 . Publisher: Taylor & Francis _eprint: https://www.tandfonline.com/doi/pdf/10.1080/01621459.1954.10501232. Accessed 2024-04-23

Acero, F.J., García, J.A., Gallego, M.C.: Peaks-over-Threshold Study of Trends in Extreme Rainfall over the Iberian Peninsula. Journal of Climate **24**(4), 1089–1105 (2011) https://doi.org/10.1175/2010JCLI3627.1 . Publisher: American Meteorological Society Section: Journal of Climate. Accessed 2023-05-16

Anderson, T.W.: On the Distribution of the Two-Sample Cramer-von Mises Criterion. The Annals of Mathematical Statistics **33**(3), 1148–1159 (1962) https://doi.org/10.1214/aoms/1177704477 . Publisher: Institute of Mathematical Statistics. Accessed 2024-04-23

Alfieri, L., Thielen, J.: A European precipitation index for extreme rain-storm and flash flood early warning. Meteorological Applications **22**(1), 3–13 (2015) https://doi.org/10.1002/met.1328 . _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/met.1328. Accessed 2023-05-17

Balkema, A.A., Haan, L.d.: Residual Life Time at Great Age. The Annals of Probability **2**(5), 792–804 (1974) https://doi.org/10.1214/aop/1176996548 . Publisher:

Institute of Mathematical Statistics. Accessed 2024-04-24

Berk, R.H., Jones, D.H.: Goodness-of-fit test statistics that dominate the Kolmogorov statistics. Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete **47**(1), 47–59 (1979) https://doi.org/10.1007/BF00533250 . Accessed 2024-01-23

Blain, G.C., Meschiatti, M.C.: Inadequacy of the Gamma distribution to calculate the Standardized Precipitation Index. Revista Brasileira de Engenharia Agrícola e Ambiental **19**, 1129–1135 (2015) https://doi.org/10.1590/1807-1929/agriambi.v19n12p1129-1135 . Publisher: Departamento de Engenharia Agrícola - UFCG. Accessed 2023-05-16

C3S: ERA5 hourly data on single levels from 1940 to present (2018). https://doi.org/10.24381/CDS.ADBB2D47 . https://cds.climate.copernicus.eu/doi/10.24381/cds.adbb2d47 Accessed 2024-04-05

Castro-Camilo, D., Huser, R., Rue, H.: A Spliced Gamma-Generalized Pareto Model for Short-Term Extreme Wind Speed Probabilistic Forecasting. Journal of Agricultural, Biological and Environmental Statistics **24**(3), 517–534 (2019) https://doi.org/10.1007/s13253-019-00369-z . Accessed 2024-07-08

Chen, D., Dai, A., Hall, A.: The Convective-To-Total Precipitation Ratio and the "Drizzling" Bias in Climate Models. Journal of Geophysical Research: Atmospheres **126**(16), 2020–034198 (2021) https://doi.org/10.1029/2020JD034198 . _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2020JD034198. Accessed 2024-07-08

Carrió, D.S., Jansà, A., Homar, V., Romero, R., Rigo, T., Ramis, C., Hermoso, A., Maimó, A.: Exploring the benefits of a Hi-EnKF system to forecast an extreme weather event. The 9th October 2018 catastrophic flash flood in Mallorca. Atmospheric Research **265**, 105917 (2022) https://doi.org/10.1016/j.atmosres.2021.105917 . Accessed 2023-05-16

Cannon, A.J., Sobie, S.R., Murdock, T.Q.: Bias Correction of GCM Precipitation by Quantile Mapping: How Well Do Methods Preserve Changes in Quantiles and Extremes? Journal of Climate **28**(17), 6938–6959 (2015) https://doi.org/10.1175/JCLI-D-14-00754.1 . Publisher: American Meteorological Society Section: Journal of Climate. Accessed 2023-01-17

Costache, R., Tien Bui, D.: Identification of areas prone to flash-flood phenomena using multiple-criteria decision-making, bivariate statistics, machine learning and their ensembles. Science of The Total Environment **712**, 136492 (2020) https://doi.org/10.1016/j.scitotenv.2019.136492 . Accessed 2023-05-16

Cucchi, M., Weedon, G., Amici, A., Bellouin, N., Lange, S., Müller Schmied, H., Hersbach, H., Buontempo, C.: WFDE5: bias adjusted ERA5 reanalysis data for impact studies (2020) https://doi.org/10.5194/essd-12-2097-2020

De Haan, L., Ferreira, A.: Extreme Value Theory. Springer Series in Operations Research and Financial Engineering. Springer, New York, NY (2006). https://doi.org/10.1007/0-387-34471-3 . http://link.springer.com/10.1007/0-387-34471-3 Accessed 2024-07-09

Déqué, M.: Frequency of precipitation and temperature extremes over France in an anthropogenic scenario: Model results and statistical correction according to observed values. Global and Planetary Change **57**(1), 16–26 (2007) https://doi.org/10.1016/j.gloplacha.2006.11.030 . Accessed 2023-02-06

Francois, B.: Multivariate statistical approaches for bias adjustment of climate simulations and compound events analysis. phdthesis, Université Paris-Saclay (September 2022). https://theses.hal.science/tel-03891297 Accessed 2023-05-03

Gamet, P., Jalbert, J.: A flexible Extended Generalized Pareto distribution for tail estimation. Environmetrics **33**(6), 2744 (2022) https://doi.org/10.1002/env.2744 . _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/env.2744. Accessed 2023-10-17

Holthuijzen, M., Beckage, B., Clemins, P.J., Higdon, D., Winter, J.M.: Robust bias-correction of precipitation extremes using a novel hybrid empirical quantile-mapping method: Advantages of a linear correction for extremes. Theoretical and Applied Climatology **149**(1-2), 863–882 (2022) https://doi.org/10.1007/s00704-022-04035-2 . Accessed 2023-01-23

Henckes, P., Knaut, A., Obermüller, F., Frank, C.: The benefit of long-term high resolution wind data for electricity system analysis. Energy **143**, 934–942 (2018) https://doi.org/10.1016/j.energy.2017.10.049 . Accessed 2023-05-03

Husak, G.J., Michaelsen, J.C., Funk, C.C.: Use of the Gamma distribution to represent monthly rainfall in Africa for drought monitoring applications. International Journal of Climatology **27**(7), 935–944 (2007) https://doi.org/10.1002/joc.1441 . Accessed 2024-06-12

Katz, R.W.: Techniques for estimating uncertainty in climate change scenarios and impact studies. Climate Research **20**(2), 167–185 (2002) https://doi.org/10.3354/cr020167 . Accessed 2024-06-12

Khan, S.A.: Exponentiated Weibull regression for time-to-event data. Lifetime Data Analysis **24**(2), 328–354 (2018) https://doi.org/10.1007/s10985-017-9394-3 . Accessed 2024-01-29

Kim, Y., Kim, H., Lee, G., Min, K.-H.: A Modified Hybrid Gamma and Generalized Pareto Distribution for Precipitation Data. Asia-Pacific Journal of Atmospheric Sciences **55**(4), 609–616 (2019) https://doi.org/10.1007/s13143-019-00114-z . Accessed 2024-07-08

Laio, F.: Cramer–von Mises and Anderson-Darling goodness of fit tests for extreme value distributions with unknown parameters. Water Resources Research **40**(9) (2004) https://doi.org/10.1029/2004WR003204 . _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2004WR003204. Accessed 2024-06-12

Lafon, T., Dadson, S., Buys, G., Prudhomme, C.: Bias correction of daily precipitation simulated by a regional climate model: a comparison of methods. International Journal of Climatology **33**(6), 1367–1381 (2013) https://doi.org/10.1002/joc.3518 . _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/joc.3518. Accessed 2023-05-03

Li, B., Huang, Y., Du, L., Wang, D.: Bias Correction for Precipitation Simulated by RegCM4 over the Upper Reaches of the Yangtze River Based on the Mixed Distribution Quantile Mapping Method. Atmosphere **12**(12), 1566 (2021) https://doi.org/10.3390/atmos12121566 . Number: 12 Publisher: Multidisciplinary Digital Publishing Institute. Accessed 2023-03-28

Langousis, A., Mamalakis, A., Deidda, R., Marrocu, M.: Assessing the relative effectiveness of statistical downscaling and distribution mapping in reproducing rainfall statistics based on climate model results. Water Resources Research **52**(1), 471–494 (2016) https://doi.org/10.1002/2015WR017556 . _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/2015WR017556. Accessed 2024-10-17

Langousis, A., Mamalakis, A., Puliga, M., Deidda, R.: Threshold detection for the generalized Pareto distribution: Review of representative methods and application to the NOAA NCDC daily rainfall database. Water Resources Research **52**(4), 2659–2681 (2016) https://doi.org/10.1002/2015WR018502 . Accessed 2023-09-21

Massey Jr., F.J.: The Kolmogorov-Smirnov Test for Goodness of Fit. Journal of the American Statistical Association **46**(253), 68–78 (1951) https://doi.org/10.1080/01621459.1951.10500769 . Publisher: Taylor & Francis _eprint: https://www.tandfonline.com/doi/pdf/10.1080/01621459.1951.10500769. Accessed 2024-04-23

Mamalakis, A., Langousis, A., Deidda, R., Marrocu, M.: A parametric approach for simultaneous bias correction and high-resolution downscaling of climate model rainfall. Water Resources Research **53**(3), 2149–2170 (2017) https://doi.org/10.1002/2016WR019578 . _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/2016WR019578. Accessed 2024-10-17

Monteiro, D., Morin, S.: Multi-decadal analysis of past winter temperature, precipitation and snow cover data in the European Alps from reanalyses, climate models and observational datasets. The Cryosphere **17**(8), 3617–3660 (2023) https://doi.org/10.5194/tc-17-3617-2023 . Publisher: Copernicus GmbH. Accessed 2024-04-05

Moscovich, A., Nadler, B.: Fast calculation of boundary crossing probabilities for Poisson processes. Statistics & Probability Letters **123**, 177–182 (2017) https://doi.org/10.1016/j.spl.2016.11.027 . Accessed 2024-01-23

Moscovich, A., Nadler, B., Spiegelman, C.: On the exact Berk-Jones statistics and their p-value calculation. Electronic Journal of Statistics **10**(2) (2016) https://doi.org/10.1214/16-EJS1172 . arXiv:1311.3190 [math, stat]. Accessed 2023-11-26

Moscovich, A.: Fast calculation of p-values for one-sided Kolmogorov-Smirnov type statistics. Computational Statistics & Data Analysis **185**, 107769 (2023) https://doi.org/10.1016/j.csda.2023.107769 . Accessed 2024-01-25

Mudholkar, G.S., Srivastava, D.K.: Exponentiated Weibull family for analyzing bathtub failure-rate data. IEEE Transactions on Reliability **42**(2), 299–302 (1993) https://doi.org/10.1109/24.229504 . Conference Name: IEEE Transactions on Reliability. Accessed 2023-11-16

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D.G., Piles, M., Rodríguez-Fernández, N.J., Zsoter, E., Buontempo, C., Thépaut, J.-N.: ERA5-Land: a state-of-the-art global reanalysis dataset for land applications. Earth System Science Data **13**(9), 4349–4383 (2021) https://doi.org/10.5194/essd-13-4349-2021 . Publisher: Copernicus GmbH. Accessed 2023-05-02

Mudholkar, G.S., Srivastava, D.K., Kollia, G.D.: A Generalization of the Weibull Distribution with Application to the Analysis of Survival Data. Journal of the American Statistical Association **91**(436), 1575–1583 (1996) https://doi.org/10.2307/2291583 . Publisher: [American Statistical Association, Taylor & Francis, Ltd.]. Accessed 2024-01-29

Malaekeh, S., Safaie, A., Shiva, L., Tabari, H.: Spatio-temporal variation of hydroclimatic variables and extreme indices over Iran based on reanalysis data. Stochastic Environmental Research and Risk Assessment **36**(11), 3725–3752 (2022) https://doi.org/10.1007/s00477-022-02223-0 . Accessed 2024-04-05

McNeil, A., Saladin, T., Zentrum, E.: The Peaks over Thresholds Method for Estimating High Quantiles of Loss Distributions. Proceedings of 28th International ASTIN Colloquium (1997)

Michelangeli, P.-A., Vrac, M., Loukos, H.: Probabilistic downscaling approaches: Application to wind cumulative distribution functions. Geophysical Research Letters **36**(11) (2009) https://doi.org/10.1029/2009GL038401 . _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2009GL038401. Accessed 2022-10-21

Martinez-Villalobos, C., Neelin, J.D.: Why Do Precipitation Intensities Tend to Follow Gamma Distributions? Journal of the Atmospheric Sciences **76**(11), 3611–3631

(2019) https://doi.org/10.1175/JAS-D-18-0343.1 . Publisher: American Meteorological Society Section: Journal of the Atmospheric Sciences. Accessed 2023-03-28

Nadarajah, S., Choi, D.: Maximum daily rainfall in South Korea. Journal of Earth System Science **116**(4), 311–320 (2007) https://doi.org/10.1007/s12040-007-0028-0 . Accessed 2024-01-29

Naveau, P., Huser, R., Ribereau, P., Hannart, A.: Modeling jointly low, moderate, and heavy rainfall intensities without a threshold selection. Water Resources Research **52**(4), 2753–2769 (2016) https://doi.org/10.1002/2015WR018552 . _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/2015WR018552. Accessed 2023-09-20

Pal, M., Ali, M.M., Woo, J.: Exponentiated Weibull distribution. Statistica **66**(2), 139–147 (2006) https://doi.org/10.6092/issn.1973-2201/493 . Number: 2. Accessed 2024-01-29

Prein, A.F., Gobiet, A., Truhetz, H., Keuler, K., Goergen, K., Teichmann, C., Fox Maule, C., Meijgaard, E., Déqué, M., Nikulin, G., Vautard, R., Colette, A., Kjellström, E., Jacob, D.: Precipitation in the EURO-CORDEX 0.11° and 0.44° simulations: high resolution, high benefits? Climate Dynamics **46**(1), 383–412 (2016) https://doi.org/10.1007/s00382-015-2589-y . Accessed 2023-05-03

Pickands, J.I.: Statistical Inference Using Extreme Order Statistics. The Annals of Statistics **3**(1), 119–131 (1975) https://doi.org/10.1214/aos/1176343003 . Publisher: Institute of Mathematical Statistics. Accessed 2024-04-24

Papastathopoulos, I., Tawn, J.A.: Extended generalised Pareto models for tail estimation. Journal of Statistical Planning and Inference **143**(1), 131–143 (2013) https://doi.org/10.1016/j.jspi.2012.07.001 . Accessed 2023-09-20

Ristić, M.M., Balakrishnan, N.: The Gamma-Exponentiated Exponential distribution. Journal of Statistical Computation and Simulation **82**(8), 1191–1206 (2012) https://doi.org/10.1080/00949655.2011.574633 . Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/00949655.2011.574633. Accessed 2024-01-29

Rootzén, H., Tajvidi, N.: Extreme value statistics and wind storm losses: A case study. Scandinavian Actuarial Journal **1997**(1), 70–94 (1997) https://doi.org/10.1080/03461238.1997.10413979 . Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/03461238.1997.10413979. Accessed 2023-05-03

Sangati, M., Borga, M.: Influence of rainfall spatial resolution on flash flood modelling. Natural Hazards and Earth System Sciences **9**(2), 575–584 (2009) https://doi.org/10.5194/nhess-9-575-2009 . Publisher: Copernicus GmbH. Accessed 2023-05-17

Soares, P.M.M., Cardoso, R.M.: A simple method to assess the added value using high-resolution climate distributions: application to the

EURO-CORDEX daily precipitation. International Journal of Climatology **38**(3), 1484–1498 (2018) https://doi.org/10.1002/joc.5261 . _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/joc.5261. Accessed 2023-05-03

Scarrott, C., MacDonald, A.: A Review of Extreme Value Threshold Estimation and Uncertainty Quantification. REVSTAT-Statistical Journal **10**(1), 33–60 (2012) https://doi.org/10.57805/revstat.v10i1.110 . Number: 1. Accessed 2023-09-21

Sharma, V.K., Singh, S.V., Shekhawat, K.: Exponentiated Teissier distribution with increasing, decreasing and bathtub hazard functions. Journal of Applied Statistics **49**(2), 371–393 (2022) https://doi.org/10.1080/02664763.2020.1813694 . Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/02664763.2020.1813694. Accessed 2024-01-29

Steinskog, D.J., Tjøstheim, D.B., Kvamstø, N.G.: A Cautionary Note on the Use of the Kolmogorov–Smirnov Test for Normality. Monthly Weather Review **135**(3), 1151–1157 (2007) https://doi.org/10.1175/MWR3326.1 . Publisher: American Meteorological Society Section: Monthly Weather Review. Accessed 2024-06-12

Trentini, L., Dal Gesso, S., Venturini, M., Guerrini, F., Calmanti, S., Petitta, M.: A Novel Bias Correction Method for Extreme Events. Climate **11**(1), 3 (2023) https://doi.org/10.3390/cli11010003 . Number: 1 Publisher: Multidisciplinary Digital Publishing Institute. Accessed 2023-05-03

Tencaliec, P., Favre, A.-C., Naveau, P., Prieur, C., Nicolet, G.: Flexible semiparametric Generalized Pareto modeling of the entire range of rainfall amount. Environmetrics **31**(2), 2582–1 (2019) https://doi.org/10.1002/env.2582 . Accessed 2023-10-12

The MathWorks Inc.: MATLAB. The MathWorks Inc., Natick, Massachusetts, United States (2022). https://www.mathworks.com

Um, M.J., Kim, H., Heo, J.H.: Hybrid approach in statistical bias correction of projected precipitation for the frequency analysis of extreme events. Advances in Water Resources **94**, 278–290 (2016) https://doi.org/10.1016/j.advwatres.2016.05.021 . Accessed 2024-10-22

Verrelle, A., Glinton, M., Bazile, E., Moigne, P.L.: CERRA-Land : A new land surface reanalysis at 5.5 km resolution over Europe. Technical Report EMS2021-492, Copernicus Meetings (June 2021). https://doi.org/10.5194/ems2021-492 . Conference Name: EMS2021. https://meetingorganizer.copernicus.org/EMS2021/EMS2021-492.html Accessed 2024-03-21

Vlček, O., Huth, R.: Is daily precipitation Gamma-distributed?: Adverse effects of an incorrect use of the Kolmogorov–Smirnov test. Atmospheric Research **93**(4), 759–766 (2009) https://doi.org/10.1016/j.atmosres.2009.03.005 . Accessed 2023-03-28

Xu, H., Xu, C.-Y., Sælthun, N.R., Zhou, B., Xu, Y.: Evaluation of reanalysis and

satellite-based precipitation datasets in driving hydrological models in a humid region of Southern China. Stochastic Environmental Research and Risk Assessment **29**(8), 2003–2020 (2015) https://doi.org/10.1007/s00477-014-1007-z . Accessed 2024-07-09

Yang, F., Koukoula, M., Emmanouil, S., Cerrai, D., Anagnostou, E.N.: Assessing the power grid vulnerability to extreme weather events based on long-term atmospheric reanalysis. Stochastic Environmental Research and Risk Assessment **37**(11), 4291–4306 (2023) https://doi.org/10.1007/s00477-023-02508-y . Accessed 2024-07-08

Yue, Z., Xiong, L., Zha, X., Liu, C., Chen, J., Liu, D.: Impact of thresholds on nonstationary frequency analyses of peak over threshold extreme rainfall series in Pearl River Basin, China. Atmospheric Research **276**, 106269 (2022) https://doi.org/10.1016/j.atmosres.2022.106269 . Accessed 2023-05-16

Şan, M., Nacar, S., Kankal, M., Bayram, A.: Daily precipitation performances of regression-based statistical downscaling models in a basin with mountain and semi-arid climates. Stochastic Environmental Research and Risk Assessment **37**(4), 1431–1455 (2023) https://doi.org/10.1007/s00477-022-02345-5 . Accessed 2024-07-08

A Considered metrics

To quantify and spatially visualise fit and bias correction errors of our distributional models and data, we selected three metrics: one taking into account the whole distribution (MAE), one focused on the upper tail (MAE95sup, i.e., the MAE over the 5th last quantile), and a sum-of-squares-based error (Root Mean Squared Error, RMSE) to take into account the error variance.

Let $q = \{q_1 = 0, q_2 = \frac{1}{n}, ..., q_n = 1 - \frac{1}{n}\}$, with $n = 1000$. For faster computation times and standardization, we produce an equally spaced quantile representation of the target and prediction data using q. The quantile function used is from Python's `numpy` package, with the linear interpolation method.

The ordered data quantiles are noted y and $\hat{y}$ for respectively $x_{\texttt{obs}}$ and $x_{\texttt{mod}}$, the observed and modelled data.

The Mean Absolute Error is defined as:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_{(i)} - \hat{y}_{(i)}|.$$

The Mean Absolute Error over quantile α (MAEαsup) is derived from the MAE to focus on the upper tail. More precisely, it is the MAE over the quantile α. It is defined as:

$$MAE\alpha sup = \frac{1}{n - \lceil \alpha \times n \rceil}\sum_{i=\lceil \alpha \times n \rceil}^{n}|y_{(i)} - \hat{y}_{(i)}|,$$

with $\lceil x \rceil$ being the ceiling function.
The Root Mean Square Error (RMSE) is defined as :

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_{(i)} - \hat{y}_{(i)})^2}.$$

B Considered parametric distributions

We introduce some parametric distributions: the Extended Generalized Pareto (EGP) distribution, the Exponentiated Weibull (ExpW) distribution and the Gamma distribution. The first two distributions are used to construct our semi-parametric model and the Gamma distribution is used for comparisons.
Note that for precipitation modeling, we used a 1mm threshold to separate between wet and dry days. The support of the following distributions being $]0, +\infty[$, a shift of 1mm has been applied before fitting, since the wet days threshold chosen is also 1mm.

The Gamma distribution

The Gamma distribution is a well-known light-tailed and often used distribution for daily and monthly precipitation modeling (Martinez-Villalobos and Neelin, 2019; Husak et al., 2007). Its cumulative distribution function can be written as:

$$F(x) = \frac{1}{\Gamma(k)}\gamma(\alpha, \beta x),$$

for $x > 0$ and $\alpha > 0$ and $\beta > 0$ respectively the shape and rate parameters.

The Extended Generalized Pareto distribution

An extension of the Generalized Pareto distribution have been introduced in Naveau et al. (2016); Papastathopoulos and Tawn (2013). This extension removes the difficult choice of the threshold for the classical generalized Pareto distribution (GPD).
The distribution has the following form:

$$F_{EGP}(x; G) = \begin{cases} G\left(1 - (1 + \xi x/\sigma)^{-\frac{1}{\xi}}\right), & \xi > 0, \\ G\left(1 - e^{-\frac{x}{\sigma}}\right), & \xi = 0, \end{cases} \tag{B.1}$$

with σ and ξ being the usual parameters of the GPD distribution, and G a continuous cumulative distribution function on $[0, 1]$ to that fulfils the necessary conditions from Naveau et al. (2016).
In the aforementioned article, multiple distributions for G have been presented that satisfy the required conditions. In this paper, we will focus on the first model denoted by EGP, which uses the power law distribution: $G(x) = x^{\kappa}$, for κ and $x > 0$, i.e.,

$$F_{EGP}(x) = \begin{cases} \left(1 - (1 + \xi x/\sigma)^{-\frac{1}{\xi}}\right)^{\kappa}, & \xi > 0, \\ \left(1 - e^{-\frac{x}{\sigma}}\right)^{\kappa}, & \xi = 0. \end{cases}$$

Left censoring can also be applied when fitting the distribution (more information on the censoring can be found in the original article Naveau et al. (2016)). We fixed at 3mm the left censoring in our application in Section 3.

Exponentiated Weibull distribution

The Exponentiated Weibull (ExpW) distribution has been introduced by Mudholkar and Srivastava (1993) and generalizes the Weibull distribution by adding a second parameter shape. Its cumulative distribution function can be written as:

$$F(x; k, \lambda, \alpha) = \left[1 - \exp(-(x/\lambda)^k)\right]^{\alpha},$$

for $x > 0$ and $k, \alpha, \lambda > 0$ being respectively the first and second shape parameter and the scale parameter of the distribution. Taking $\alpha = 1$ gives the usual Weibull distribution and $k = 1$ gives the Exponentiated distribution.
The Exponentiated Weibull distribution has been used historically for failure rates (Khan, 2018; Pal et al., 2006) and survival data modelisation (Mudholkar et al., 1996), but has also been used in some cases to model precipitation data (Nadarajah and Choi, 2007; Sharma et al., 2022; Ristić and Balakrishnan, 2012).

C Monotonicity correction and analytical description of the proposed semi-parametric distributional model

Replacing a portion of the quantile function with the empirical quantile function or the lighter-tailed $F_{lighter}$ quantile function should be done carefully. Since the quantile function is an increasing function, corrections may be needed to guarantee the monotonicity. To solve this issue, we introduce in the next section a modification for all locations.

C.1 Monotonicity correction and application

Let F_1 and F_2 two distribution functions non-decreasing and right-continuous with associated quantile functions F_1^{-1} and F_2^{-1}. We define $\tilde{F}^{-1}$ as the following constructed pseudo-quantile function

$$\tilde{F}^{-1}(p) = \begin{cases} F_1^{-1}(p), & \forall\, p \leq p_{\texttt{stitch}}, \\ F_2^{-1}(p), & \text{otherwise}, \end{cases}$$

for $p_{\texttt{stitch}} \in [0, 1]$. Notice that the monotonicity condition of $\tilde{F}^{-1}$ is not guaranteed.

A breakage in the monotonicity of the pseudo-quantile function $\tilde{F}^{-1}$ at probability $p_{\texttt{stitch}}$ is characterized by the following condition:

$$\exists\, \epsilon > 0 \,|\, \tilde{F}^{-1}\left(p_{\texttt{stitch}}\right) > \tilde{F}^{-1}\left(p_{\texttt{stitch}} + \epsilon\right), i.e.,$$
$$F_1^{-1}\left(p_{\texttt{stitch}}\right) > \lim_{\epsilon \to 0^+} F_2^{-1}\left(p_{\texttt{stitch}} + \epsilon\right).$$

We will consider two types of breaks: small breaks (with amplitude smaller than $\eta = 5$ mm) and large breaks with amplitude larger than 5 mm. This amplitude of $\eta = 5$mm was chosen arbitrarily in the present work and can be modified.
In the first case, the correction is a shift on all values higher than the breaking point quantile $F_1^{-1}(p_{\texttt{stitch}})$,

$$\tilde{F}^{-1}(p) = F_2^{-1}(p) + \underbrace{F_1^{-1}\left(p_{\texttt{stitch}}\right) - \lim_{\epsilon \to 0^+} F_2^{-1}\left(p_{\texttt{stitch}} + \epsilon\right)}_{C}, \forall p > p_{\texttt{stitch}}. \tag{C.1}$$

In the second case, the values produced by F_1^{-1} near $p_{\texttt{stitch}}$ are considered as too high and they might significantly change the quantile function if the correction in Equation (C.1) is applied. In this second case, a type of ceiling function is applied to the values produced by F_1 for all probabilities lower than the breakage point:

$$\tilde{F}^{-1}(p) = \min\left\{F_1^{-1}(p), \underbrace{\lim_{\epsilon \to 0^+} F_2^{-1}(p_{\texttt{stitch}} + \epsilon)}_{C_{ceil}}\right\}, \quad \forall\, p \leq p_{\texttt{stitch}}. \tag{C.2}$$

Using these corrections on $\tilde{F}^{-1}$, the resulting quantile function is well-defined.

Application 4 Correction of monotonicity

Step (1) Find the breaking point index i (either $i_{\ell,EGP}, i_{u,EGP}, i_{\ell,ExpW}$ or $i_{u,ExpW}$. If there is more than one break in the pseudo-quantile function, the corrections should be applied in ascending order (lower indices first) and are cumulative.
Step (2) Determine the amplitude of the break η defined by $q_{(i-1)} - q_{(i)}$
Step (3) If $\eta \leq 5$, correct the monotonicity by shifting all values on the right of the breaking point by the amplitude η as described in Equation (C.1).
Step (4) If $\eta > 5$, correct the monotonicity break by mapping those values to $x_{(i)}$ for all indices lower than i as in Equation (C.2).

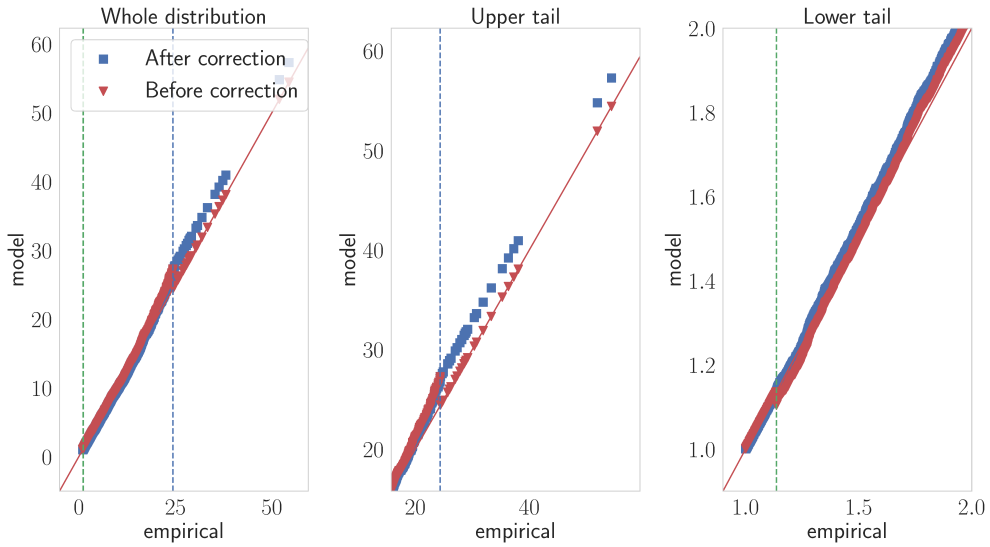


Fig. 21: QQ-plots of the same selected location as in Figure 6 but with the correction of monotonicity

In Figure 21, both types of break in monotonicity have been corrected. In the lower quantiles (right panel), the break have been corrected by shifting all following values using Equation (C.1). In the higher quantiles (center panel), the break has been corrected by putting a ceiling to the previous values as in Equation (C.2) since the range of the break exceeds 5 mm. From the newly corrected quantile function, it is now possible to properly define a proper cdf.
We can now give an analytical description for both the resulting stitched quantile function and associated cdf which will allow for both interpolation and extrapolation on unobserved values where the empirical distribution has not been stitched.

C.2 Analytical model

A semi-continuous description of the Stitch-BJ distribution can be derived using the cutting indices and rejection types $i_{u,F}$, $i_{\ell,F}$ and $\texttt{typeRej}_F$ for F_{heavy} and $F_{lighter}$ as defined in Definitions 1 and 2, and the booleans $r_{\texttt{all}}, r_{\texttt{left}}$ and $r_{\texttt{right}}$ as defined in Definition 3. In Definition 2, we have at most three different quantile functions used to describe infer the Stitch-BJ model. Let us define F_1^{-1}, F_2^{-1} and F_3^{-1}, three potentially identical quantile functions as follows:

$$F_1^{-1} = \begin{cases} F_n^{-1}, & \text{if } \texttt{typeRej}_{F_{heavy}} \text{ and } \texttt{typeRej}_{F_{lighter}} == 2, 4 \text{ or } 5 \\ F_{lighter}^{-1}, & \text{if } r_{\texttt{left}} \text{ is } \texttt{True}, \text{ or } r_{\texttt{all}} \text{ is } \texttt{True} \text{ and } \texttt{typeRej}_{F_{lighter}} \neq 2 \text{ or } 4 \\ F_{heavy}^{-1}, & \text{otherwise} \end{cases} \tag{C.3}$$

$$F_2^{-1} = \begin{cases} F_n^{-1}, & \text{if } r_{\texttt{all}} \text{ is } \texttt{False} \text{ and } \texttt{typeRej}_{F_{heavy}} == 5 \\ F_{lighter}^{-1}, & \text{if } r_{\texttt{all}} \text{ is } \texttt{True} \\ F_{heavy}^{-1}, & \text{if } r_{\texttt{all}} \text{ is } \texttt{False} \text{ and } \texttt{typeRej}_{F_{heavy}} \neq 5 \end{cases} \tag{C.4}$$

$$F_3^{-1} = \begin{cases} F_n^{-1}, & \text{if } \texttt{typeRej}_{F_{heavy}} \text{ and } \texttt{typeRej}_{F_{lighter}} \geq 3 \\ F_{lighter}^{-1}, & \text{if } r_{\texttt{right}} \text{ is } \texttt{True}, \text{ or } r_{\texttt{all}} \text{ is } \texttt{True} \text{ and } \texttt{typeRej}_{F_{lighter}} \leq 2 \\ F_{heavy}^{-1}, & \text{otherwise} \end{cases} \tag{C.5}$$

Let us define i_ℓ and i_u as follow :

- $i_\ell = \begin{cases} 0 & \text{if } r_{all} \text{ is } \texttt{True} \text{ , or } i_{\ell,F_{heavy}} == 0 \\ i_{\ell,F_{heavy}} & \text{if } i_{\ell,F_{heavy}} > 0 \text{ and } r_{all} == \texttt{False} \\ i_{\ell,F_{lighter}} & \text{if } i_{\ell,F_{lighter}} > 0 \text{ and } r_{all} \text{ is } \texttt{True} \end{cases}$
- $i_u = \begin{cases} n-1 & \text{if } r_{all} \text{ is } \texttt{True} \text{ or } i_{u,F_{heavy}} == n-1 \\ i_{u,F_{heavy}} & \text{if } i_{u,F_{heavy}} < n-1 \text{ and } r_{all} == \texttt{False} \\ i_{u,F_{lighter}} & \text{if } i_{u,F_{lighter}} < n-1 \text{ and } r_{all} == \texttt{True} \end{cases}$

A different quantile function is used depending on the considered quantile, i.e.,

$$\tilde{F}^{-1}(p) = \begin{cases} F_1^{-1}(p), & \text{if } p < \frac{i_\ell}{n} \\ F_2^{-1}(p), & \text{if } \frac{i_\ell}{n} \leq p < \frac{i_u+1}{n} \\ F_3^{-1}(p), & \text{if } \frac{i_u+1}{n} \leq p. \end{cases} \tag{C.6}$$

This is a pseudo-quantile function because the monotonicity condition is not yet verified with the resulting $\tilde{F}$ from (C.6).
Let us also define $C1$, $C1_{ceil}$, $C2$ and $C2_{ceil}$ as defined in Equations (C.1) and (C.2), as the corrections coefficient for respectively the lower and upper cutting index.
If the index is such that the correction coefficient cannot be defined (if the index is equal to 0 or $n-1$ for respectively i_ℓ and i_u) , then $C = 0$ and $C_{ceil} = +\infty$.
Since the corrections (C.1) and (C.2) can not happen simultaneously, if $C > 0$, then $C_{ceil} = +\infty$, otherwise we let $C = 0$ if $C_{ceil} \neq +\infty$. The corrected-quantile function can then be defined as:

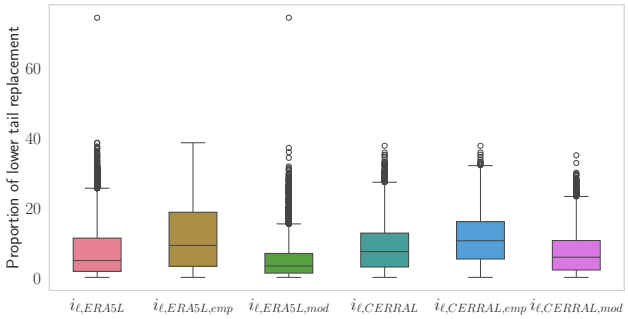
$$F^{-1}(p) = \begin{cases} \min(\tilde{F}^{-1}(p), C1_{ceil}) & \text{if } p < \frac{i_\ell}{n} \\ \min(\tilde{F}^{-1}(p) + C1, C2_{ceil}) & \text{else if } \frac{i_\ell}{n} \leq p < \frac{i_u+1}{n} \\ \tilde{F}^{-1}(p) + C1 + C2 & \text{otherwise.} \end{cases} \tag{C.7}$$

Given the properly quantile function defined in Equation (C.7), one can express the corresponding cdf given F_1, F_2 and F_3 in (C.3)-(C.5), the corrections $C1, C1_{ceil}, C2$ and $C2_{ceil}$ from Equations (C.1)-(C.2), and cutting indices, i.e.,

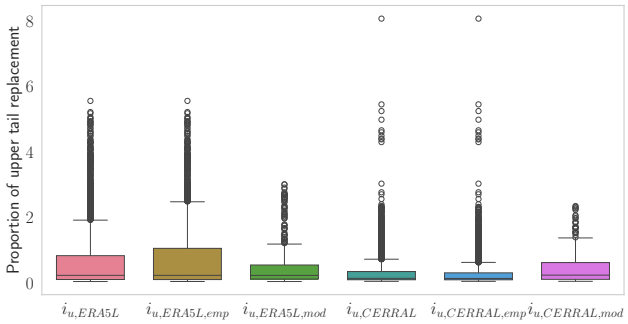
$$F(x) = \begin{cases} F_1(x) & \text{if } x < F^{-1}(\frac{i_\ell}{n}) \\ F_2(x - C1) & \text{if } F^{-1}(\frac{i_\ell}{n}) \leq x < F^{-1}(\frac{i_u+1}{n}) \\ F_3(x - C1 - C2) & \text{if } F^{-1}(\frac{i_u+1}{n}) \leq x. \end{cases}$$

D Complementary figures

Boxplots of the proportion of upper and lower tail replacements are shown in Figure 22. In this figure, $i_{\ell,data,emp}$ and $i_{\ell,data,mod}$ (resp. $i_{u,data,emp}$ and $i_{u,data,mod}$) refer to the lower (resp. upper) cutting indices introduced in Definition 1 of locations where respectively the empirical or the ExpW distributions were used to replace the lower (resp. upper) tail for a given time series. When *emp* or *mod* are not specified, i.e. $i_{\ell,ERA5L}$, $i_{\ell,CERRAL}$, $i_{u,ERA5L}$, $i_{u,CERRAL}$ in Figure 22, it refers to any cutting index of locations where a replacement occurred. We can see that for the upper tail, most replacements are done only on the last part of the upper tail, with at most 0.08% of the distribution replaced on the upper tail by an empirical distribution. On the lower tail, the proportion of the distribution replaced is higher. However, this point is less crucial for this particular study since small precipitations usually have a negligible impact on considered extreme rainfalls.



(a) Proportion of lower tail replaced



(b) Proportion of upper tail replaced

Fig. 22: Proportion of replaced upper and lower tail for the model fitted on ERA5-L and CERRA-L.

Figure 23 illustrates how the lower and upper cutting indices are determined for a rejection threshold $p_\alpha = 0.175$ and $lag = 3$. The star and cross markers are the upper and lower cutting indices, respectively. Starting from the left side, all points below the rejection threshold are rejected. At rank 4, the marker is above the p_α rejection level, however, it is still not accepted since the lag condition is not respected as the next 3 markers are not also above the rejection threshold. Rank 6 satisfies all conditions :

- It is above the p_α threshold;
- All 3 ranks after it are above the p_α threshold;
- It is the first rank from the left satisfying the first two conditions.

A similar reading from the right provides us with the upper index marked with the star. Remark that the rank 10 is not rejected even though it is below the threshold p_α because the indices only use the next three points to evaluate the rejection.

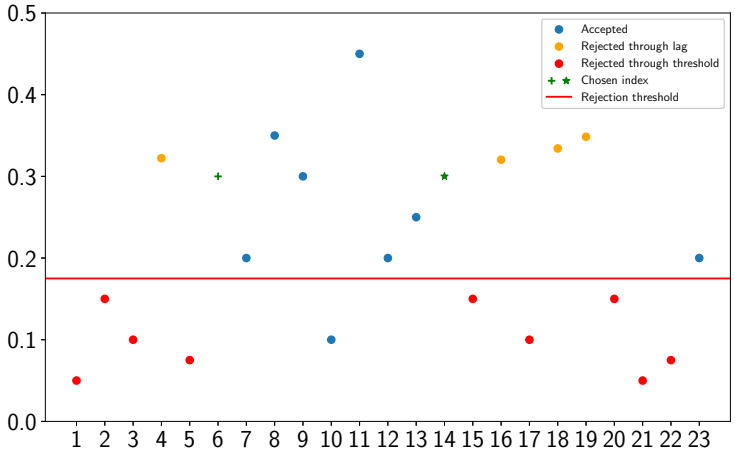


Fig. 23: Illustration of the left and right cutting indices from Definition 1