# CALIBRATION OF THE BULK AND EXTREMES OF SPATIAL DATA

## (*Invited Paper*)

Authors:   MARIA ANTÓNIA AMARAL TURKMAN [iD] *
– Centro de Estatística e Aplicações, Faculdade de Ciências,
Universidade de Lisboa, Portugal
maturkman@fc.ul.pt

KAMIL FERIDUN TURKMAN [iD]
– Centro de Estatística e Aplicações, Faculdade de Ciências,
Universidade de Lisboa, Portugal
kfturkman@fc.ul.pt

PATRÍCIA DE ZEA BERMUDEZ [iD]
– Departamento de Estatística e Investigação Operacional and
Centro de Estatística e Aplicações, Faculdade de Ciências,
Universidade de Lisboa, Portugal
pcbermudez@fc.ul.pt

SORAIA PEREIRA [iD]
– Centro de Estatística e Aplicações, Faculdade de Ciências,
Universidade de Lisboa, Portugal
sapereira@fc.ul.pt

PAULA PEREIRA [iD]
– Escola Superior de Tecnologia de Setúbal, Instituto Politécnico Setúbal and
Centro de Estatística e Aplicações, Faculdade de Ciências,
Universidade de Lisboa, Portugal
paula.pereira@estsetubal.ips.pt

MIGUEL DE CARVALHO [iD]
– School of Mathematics, University of Edinburgh, United Kingdom and
Centro de Estatística e Aplicações, Faculdade de Ciências,
Universidade de Lisboa, Portugal
Miguel.deCarvalho@ed.ac.uk

Abstract:

• In an environmental framework, extreme values of certain spatio-temporal processes, for example wind speeds, are the main cause of severe damage in property, such as electrical networks, transport and agricultural infrastructures. Typically, as is the case of wind speeds, data are available at few stations with many missing observations and consequently simulators are often used to augment information. However, simulated data often mismatch observed data, particularly at tails, therefore calibrating and bringing it in line with observed data may offer practitioners more reliable and richer data sources. In this work we will concentrate on calibrating the bulk and the extremes of data, simultaneously, avoiding methods that rely on the choice of a threshold. We propose and describe in detail a specific conditional quantile matching calibration method and exemplify it with wind speed data. We also briefly suggest how calibration should be extended specifically to data coming from the tails of simulated and observed data, using asymptotic models and methods suggested by extreme value theory.

---

*Corresponding author.

Keywords:

AMS Subject Classification:

## 1. INTRODUCTION

Extreme values of certain spatio-temporal processes, such as wind speeds, are the main cause of severe damage in property, from electricity distribution grid to transport and agricultural infrastructures. Accurate assessment of causal relationships between environmental processes and their effects on risk indicators, are highly important in risk analysis, which in return depends on sound inferential methods as well as on good quality informative data. Often, information on the relevant environmental processes comes from monitoring networks, as well as from numerical-physical models (simulators) that typically solve a large set of partial differential equations, capturing the essence of the physical process under study (see, for example, Skamarock *et al.* 2008 [18], Cardoso *et al.* 2013 [6]). In general, monitoring networks are formed by a sparse set of stations, whose instrumentation are vulnerable to disruptions, resulting in data sets with many missing observations. On the other hand, simulated data from numerical simulators typically supply average yield of the process in grid cells of pre-specified dimensions, often at high resolutions, spanning large spatial domains, with no missing observations. However, simulated data typically mismatch and misaligned observed data, therefore calibrating it and bringing it in line with observed data may supply modellers with more reliable and richer sources of data. Data assimilation methods, namely combining data from multiple sources, are well known in environmental studies, with data often being used to generate initial boundary conditions for the numerical simulators (Kalnay 2003 [11]). There is a very rich statistical literature on data assimilation and data fusion with the objective of enriching the information for inference (Fuentes and Raftery 2005 [9], Berrocal *et al.* 2012 [4], Zidek *et al.* 2012 [23], Berrocal *et al.* 2014 [5], McMillan *et al.* 2010 [12]). These statistical methods are often based on Bayesian hierarchical models for space-time data (see Banerjee *et al.* 2004 [1]) and are constructed around the idea of relating the monitoring station data and the simulated data using spatial linear models with spatially varying coefficients (see Berrocal, 2019 [3]). Since these relations involve data measured at different spatial resolutions, the models are often called downscaler models (see Berrocal *et al.* 2012 [4]). The principal objective of these downscaler models is to relate observations measured at different space resolutions using spatial linear models.

The motivation behind this present work has its roots in a consulting work done for a major electricity producer and distributor. The electricity grid constantly faces disruptions due to damages in the distribution system, with heavy economic losses. These damages and consequent disruptions occur due to a combination of many factors such as topography and precipitation, however extreme winds and storms are the main cause of such damages. Risk maps that indicate likely places of costly disruptions in electric grids are important decision support tools for administering the power grid and are particularly useful in deciding if costly corrective actions should be taken to improve structures. It is natural that these risk maps should be based primarily on observed wind speeds among other factors and it was decided that daily maximum wind speeds should be used as proxy information. Hence, such risk maps can be interpreted as vulnerability maps of electricity grid to extreme wind speeds, expressed in terms of probability. However generating such maps depends on reliable wind data at fairly high spatial and temporal resolutions.

The data set available for this particular study corresponds to simulated wind speeds from a simulator (The WRF model, version 3.1.1) at a regular grid of 81ksq grid cell size,

obtained at 10 minutes interval from 2006–2013; however only daily maximum wind speeds will be used. Observed daily maximum wind speeds are also available during the same period of time, from 117 stations in Portugal mainland, but missing observations reach to 90% in some stations. Only around one third of the stations have less than 30% missing observations. There is an additional challenge: although simulated and observed data match in the bulk of the distribution, they quite often mismatch at extreme values. Therefore, adequate methods of data fusion and calibration can be used to combine these two different sources of data, providing information which is more reliable from a spatial point of view and producing more accurate probability maps showing the spatial distribution of damage risks. Since electricity grid damages are ultimately caused by extreme wind speeds, the aim should be to develop statistical methods for data fusion and calibration that can extrapolate beyond the range of observed data into the tails of a distribution in line with extreme value theory.

We propose and describe, in detail, a specific conditional quantile matching calibration method for the bulk and the extreme observations of the data, based on models proposed by Naveau *et al.*, 2016 [14]. The outline of the paper is as follows: In Section 2, we report a new approach for calibration through a conditional quantile matching calibration method (Pereira *et al.*, 2019 [16]), using an extended generalized Pareto distribution (Papastathopoulos and Tawn 2013 [15], Naveau *et al.* 2016 [14]), adequate for calibrating simultaneously the bulk and the tails of the distribution. In Section 3, we built a Bayesian hierarchical model to implement this calibration strategy for spatio-temporal data. In Section 4, this method will be exemplified using a wind speed data. Finally, further discussion and conclusions are given in Section 5.

## 2.    CALIBRATION METHODS FOR BULK AND TAILS

We denote by $Y(s,t)$ and $X(s,t)$, respectively the observed and simulated wind speeds at location $s \in \mathbb{R}^2$ and time $t$. To simplify notation, often we will use $Y$ and $X$ for observed and simulated wind speeds when data are used without any space-time reference. Typically $X$ are simulated over a regular grid, say $B$, often represented by points $s_B$ which correspond to the centroid of the grid cells, whereas $Y$ are observed in stations located at different spatial points $s$.

For the time being, if we ignore totally space-time variations and dependence structures, calibration can be seen as a simple scaling making use of marginal distributions fitted corresponding to $X$ and $Y$ (CDF transform method, Michelangeli *et al.* 2009 [13]), as we explain below.

Suppose we have a set of $n$ observed $y_i$ and simulated $x_i$, $i = 1, ..., n$ data. Let $F_Y$ and $F_X$ be, respectively, the distribution functions of $Y$ and $X$. Then the new calibrated (scaled) data $x_i^*$ is defined as

(2.1) $$x_i^* = F_Y^{-1}(F_X(x_i)), \quad i = 1, ..., n.$$

Since

$$P(X^* \leq z) = P(F_Y^{-1}(F_X(X) \leq z) = P(U \leq F_Y(z)) = F_Y(z), \qquad \text{with} \quad U \sim U(0,1),$$

calibrated data have the same distribution as the observed data. Note that if $F_X = F_Y$ then $x_i^* = x_i$. Figure 1 depicts the result of applying this calibration method when $Y$ follows a Student distribution with 3 degrees of freedom, and $X$ follows a standard normal distribution.
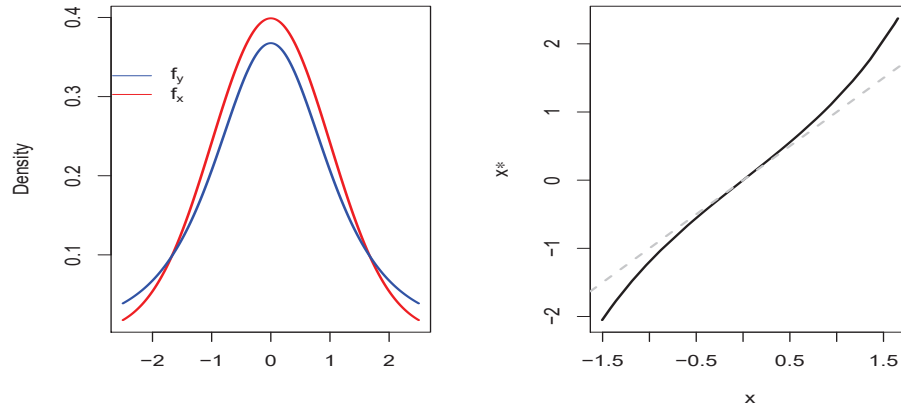


**Figure 1**: Illustration of the quantile matching approach.

This calibration method depends on the marginal distributions of the random variables involved $Y$ and $X$ and hence it does not take into consideration the expected strong dependence between the two sets of data. Thus, an ideal calibration should involve the joint distribution of $Y$ and $X$ defined in some way. One possibility is the use of a conditional quantile matching approach, which will be described in Section 2.2. Further, in the same section, we also introduce an extension to cover space-time non-homogeneity by scaling (calibrating) the data from

$$(2.2) \qquad x^*(s,t) = F_{Y(s,t)}^{-1}(F_{X(s,t)}(x(s,t)),$$

assuming marginal distributions of $Y(s,t)$ and $X(s,t)$ for every $s$ and $t$. This calibration method will take into consideration the strong space-time dependence structures expected in the data and consequently these distributions will be estimated by fitting them to data and considering the parameters as smooth functions of spatially and temporally varying covariates and space-time latent processes as in Section 4. Notice that, in this case, $x^*(s,t)$ as defined in (2.2) will depend on unknown parameters and hence calibrated data have to be estimated.

Pereira *et al.* (2019) [16] develop a covariate-adjusted version of the quantile matching-based approach as in (2.1) where the distributions of simulated and real data change along a covariate. At the same time they suggest a regression method that simultaneously models the bulk and the (right) tail of the distributions involved, using the extended generalized Pareto distribution (EGPD) (Naveau *et al.*, 2016 [14]) as a model for both the simulated and observed data. In their approach, Pereira *et al.* (2019) [16] do not take into consideration any strong spatio-temporal variations and dependence structures that may exist both in the simulated as well as in the observed data. In what follows, we propose an extension of this conditional quantile matching calibration for the bulk and tails, taking into consideration spatio-temporal variations and dependence structures, thus extending their results significantly.

de Carvalho *et al.* 2020 [7] also work on covariate adjusted version of the extended generalized Pareto distribution (EGPD) (Naveau *et al.*, 2016 [14]) for the conditional bulk and conditional tail of a possibly heavy-tailed response. However, their objective, contrary to ours, is not calibration, but to learn the effect of covariates on an extreme value setting via a Lasso-type specification.

Under fairly general conditions, according to the asymptotic theory of extremes, the generalized Pareto distribution (GPD) appears as a natural model for the right tail of a distribution, by focusing on the excesses over a high but fixed threshold. Here, the choice of this threshold plays a very important role in inference, ignoring the part of the data that lie below this threshold. See, for example, Beirland *et al.* (2004) [2]. The EGPD modelling strategy suggested by Naveau *et al* (2016) [14] avoids this selection problem, as we will see in next section.

## 2.1.  Naveau *et al.* (2016) EGPD models

Naveau *et al.* (2016) [14] suggest an extension of generalized Pareto model tailored for both the bulk and tails, and — contrarily to most methods for extremes — does not require a threshold to be selected. The objective of this extension is to generate a new class of distributions with GPD type tails consistent with extreme value theory, but also flexible enough to model efficiently the main bulk of the observed data without complicated threshold selection procedures.

Let $Y$ be a positive random variable with cumulative distribution function defined as:

$$F_Y(y \mid \theta) = G\bigg( H(y \mid \xi, \sigma) \bigg),$$

where $H$ is the cumulative distribution function of a generalized Pareto distribution (GPD) and $G$ is a function obeying some general assumptions, so that a Pareto-type tail is ensured and the bulk is driven by the carrier $G$. (see Naveau *et al.* 2016 [14] and de Carvalho *et al.* 2020 [7]), that is

$$H(y \mid \xi, \sigma) = \begin{cases} 1 - \left(1 + \dfrac{\xi}{\sigma} y\right)_+^{-1/\xi}, & \xi \neq 0. \\[2mm] 1 - \exp\left(-\dfrac{y}{\sigma}\right), & \xi = 0 \,. \end{cases}$$

with $a_+ = \max(a, 0)$, $\sigma > 0$, and $y > 0$ if $\xi \geq 0$ and $y < -\frac{\sigma}{\xi}$ if $\xi < 0$. The parameter $\sigma$ is a dispersion parameter while $\xi$ is a shape parameter controlling the rate of decay of the right tail of a distribution (e.g. de Zea Bermudez and Kotz [8]).

Naveau *et al.* [14] consider four forms of $G(u)$ resulting in four different classes of distributions. Although the theory below can be easily extended to any of the forms of the $G$ function, in what follows we use one of the forms, namely, $G(u) = u^\kappa$, the canonical form of the EGPD (de Carvalho *et al.* 2020 [7]), where $\kappa$ is a parameter controlling the shape of the lower tail. It is clear that smaller the $\kappa$ more the distribution is concentrated near zero. The EGPD will have then three parameters, and we will refer to it as a $EGPD(\kappa, \xi, \sigma)$.

## 2.2. Spatio-temporal conditional quantile matching calibration for the bulk and tails

Let us assume that both random variables $X$ and $Y$ are space-time dependent and we want to calibrate $X$ based on $Y$. The calibrated data are given as in (2.2). Now assume further that both random variables are distributed as a canonical $EGPD(\kappa, \xi, \sigma)$ where the parameters are indexed by the correspondent random variables. In order to better accommodate for the situation $\xi < 0$ we make a transformation $\delta = -\frac{\sigma}{\xi}$. Hence, for $\xi_x \neq 0$

$$(2.3) \qquad F_{X(s,t)}(x(s,t) \mid \delta_x(s,t), \xi_x, \kappa_x) = \left(1 - \left(1 - \frac{1}{\delta_x(s,t)} x(s,t)\right)_+^{-1/\xi_x}\right)^{\kappa_x},$$

for $x > 0$ if $\xi_x > 0$ and $x < \delta_x$ if $\xi_x < 0$.

Assuming as well $\xi_y \neq 0$

$$(2.4) \qquad F_{Y(s,t)}(y(s,t) \mid \delta_y(s,t), \xi_y, \kappa_y) = \left(1 - \left(1 - \frac{1}{\delta_y(s,t)} y(s,t)\right)_+^{-1/\xi_y}\right)^{\kappa_y},$$

for $y > 0$ if $\xi_y > 0$ and $y < \delta_y$ if $\xi_y < 0$.

Although it is assumed that these random variables are conditional independent, a dependence structure is introduced through the transformed space-time dependent parameters $\delta_x, \delta_y$ by modelling them as a function of a common latent spatio-temporal process, in a Bayesian hierarchical modelling framework. Here, we are mainly interested in modelling the right tail as function of space and time. de Carvalho *et al.* 2020 [7], consider a Bayesian hierachical modelling of the EGPD for the case $\xi > 0$, where both bulk and tail are covariate adjusted.

As an exemplification of this modelling strategy, in the next section, we will built a Bayesian hierarchical model for the wind speed data.

## 3. BAYESIAN HIERARCHICAL MODEL FOR WIND SPEED DATA

A preliminary data analysis of the wind speed data used in this study, shows that observed and simulated data are consistent with the case $\xi < 0$ and hence, the distributions for $X$ and $Y$ will have an end-point characterized by the respective parameter $\delta$.

Let $T$ be length of the time period under study, $N$ the number of stations with complete observed data during that period and $N_s$ the total number of stations ($N < N_s$).

The observed data $\mathbf{Y}(\mathbf{s}, \mathbf{t}) = \{Y(s_i, t_j), i = 1, ..., N; j = 1, ..., T\}$, are assumed to follow a distribution as in (2.4), with parameters $\kappa_y$, $\xi_y$ and $\delta_y(i,j)$, such that $\delta_y(i,j) \sim \text{Exp}(\lambda_y(i,j))$, $\delta_y(i,j) > \max(y)$, i.e., follows a shifted exponential distribution, with

$$(3.1) \qquad \log(\lambda_y(i,j)) = \beta_y + W(s_i) + Z(t_j),$$

where $W \sim MVN(0, \tau_W \Sigma_W)$ follows a Multivariate Gaussian process, defined on the space, as in Thomas *et al.* (2014) [21], with $\tau_W$ a precision parameter and the matrix $\Sigma_W$ with diagonal elements equal to 1 and off-diagonal elements, $\Sigma_{i\ell} = f(d_{i\ell}; \alpha)$, where

$$f(d_{i\ell}; \alpha) = 2/\pi * \left\{ \cos^{-1}(d_{i\ell}/\alpha) - [(d_{i\ell}/\alpha)(1 - (d_{i\ell}^2/\alpha^2))]^{1/2} \right\}, \quad d_{i\ell} < \alpha,$$

is a function of $d_{i\ell}$, the centroids' distance of every two stations $s_i$ and $s_\ell$, and $\alpha$ a parameter representing the radius of the 'disc' centred at each $s$. The parameter $\alpha$ controls the rate of decline of correlation with distance.

For the temporal random process $\{Z(t_j)\}$ we assume a random walk process of order 1, $Z \sim MVN(0, \tau_Z \Sigma_Z)$, where $\tau_Z$ is a precision parameter and $\Sigma_Z$ is a matrix with a structure reflecting the fact that any two increments $z_i - z_{i-1}$ are independent (Rue and Held, 2005 [17]).

For the simulated data $\mathbf{X}(\mathbf{s}, \mathbf{t}) = \{X(s_i, t_j), i = 1, ..., N_s; j = 1, ..., T\}$ we assume a distribution as in (2.3) with $N_s$ total number of stations, where the model for $\delta_x(i, j) \sim \text{Exp}(\lambda_x(i, j))$, $\delta_x(i, j) > \max(x)$, shares the same latent processes $W$ and $Z$ as the model for the observed data (3.1), i.e.,

$$\log(\lambda_x(ij)) = \beta_x + W(s_i) + Z(t_j).$$

Let $\theta$ be a vector containing all model parameters including the latent Gaussian models $W$ and $Z$. Assuming conditional independence, the likelihood $L(\theta \mid \mathbf{y}(\mathbf{s}, \mathbf{t}), \mathbf{x}(\mathbf{s}, \mathbf{t}))$ is a product of individual terms $L_{ij}(\theta \mid y(s_i, t_j))$ and $L_{\ell j}(\theta \mid x(s_\ell, t_j))$.

To complete the Bayesian hierarchical model we consider the following prior specification for the parameters and hyperparameters of the models not yet specified:

$$
\begin{aligned}
\beta_y, \beta_x \ & \text{i.i.d.} \ N(0, 100), \\
\kappa_y, \kappa_x \ & \text{i.i.d.} \ Ga(0.05, 0.05), \\
\xi_y, \xi_x \ & \text{i.i.d.} \ U(-0.5, 0), \\
\tau_W, \tau_Z \ & \text{i.i.d.} \ Ga(1, 0.1), \\
\alpha \ & \sim \ U(0.1, 0.9).
\end{aligned}
$$

(3.2)

Further, all these parameters are assumed to be *a priori* independent and hence the prior distribution $h(\theta)$ is the product of the individual priors. In Subsection 4.1 an explanation is given for the choice of these priors.

The posterior distribution

$$h(\theta \mid \mathbf{y}(\mathbf{s}, \mathbf{t}), \mathbf{x}(\mathbf{s}, \mathbf{t})) \propto L(\theta \mid \mathbf{y}(\mathbf{s}, \mathbf{t}), \mathbf{x}(\mathbf{s}, \mathbf{t})) h(\theta),$$

is analytically intractable and hence one has to resort to the use of computational methods, such as MCMC methods.

Now, let $F_X(x(s_i, t_j) \mid \kappa_x, \xi_x, \delta_x(i, j)) = p_{ij}$, for $i = 1, ..., N_s$ and $j = 1, ..., T$. Then, since, for any $p \in (0, 1)$ the inverse function of $F_Y$ is

$$\delta_y \left[ 1 - \left( 1 - p^{1/\kappa_y} \right)^{-\xi_y} \right],$$

the calibrated values defined in (2.2), as a function of the model parameters, are given by

$$(3.3) \qquad x^*(s_i, t_j \mid \theta_{ij}) = \delta_y(s_i, t_j) \left[ 1 - \left( 1 - p_{ij}^{1/\kappa_y} \right)^{-\xi_y} \right],$$

where $\theta_{ij} = (\kappa_x, \xi_x, \delta_x(i,j), \kappa_y, \xi_y, \delta_y(i,j))$.

Hence, following Bayesian methodology, calibrated data are estimated as the mean of the calibrated function, defined in (3.3), with respect to the posterior distribution $h(\theta \mid \mathbf{x}, \mathbf{y})$, that is

$$x^*(s_i, t_j) = \int \delta_y(s_i, t_j) \left[ 1 - \left( 1 - p_{ij}^{1/\kappa_y} \right)^{-\xi_y} \right] h(\theta \mid \mathbf{x}, \mathbf{y}) d\theta.$$

In what follows we will call this posterior mean as the *calibrated data*.

Computation of calibrated data is achieved through MCMC methods, by simulating a sample of size $M$ from the posterior distribution and approximating the integral as

$$x^*(s_i, t_j) \approx \frac{1}{M} \sum_{k=1}^{M} \delta_y(s_i, t_j)^{(k)} \left[ 1 - \left( 1 - (p_{ij}^{(k)})^{1/\kappa_y^{(k)}} \right)^{-\xi_y^{(k)}} \right],$$

where $\delta_y(s_i, t_j)^{(k)}, -p_{ij}^{(k)}, \kappa_y^{(k)}, \xi_y^{(k)}$ are simulated values at the $k$th iterate after convergence is achieved. A sample of size $M$ from the posterior distribution of the calibrated function given by

$$\delta_y(s_i, t_j)^{(k)} \left[ 1 - \left( 1 - (p_{ij}^{(k)})^{1/\kappa_y^{(k)}} \right)^{-\xi_y^{(k)}} \right], \quad k = 1, ..., M$$

will allow the computation of any relevant summary statistics, such as, $\gamma\%$ credible intervals.

---

## 4. APPLICATION TO WIND SPEED DATA

We used observed and simulated wind speed data from 01/01/2013 to 28/02/2013, so $T = 59$. There are $N = 51$ stations where we have both observed and simulated daily maximum wind speeds. Additionally we have extra 66 stations with simulated values for the maximum daily wind speeds, so that $N_s = 117$. In Figure 2 we depict the median of observed and simulated wind speeds for the 51 stations together with the 2.5% and 97.5% empirical quantiles (referred to on the figure as the 95% IQR).

The model was implemented in `R2OpenBUGS` (Sturtz *et al.* 2005 [19]). In Table 1 we show the summary statistics for the marginal posterior distributions of the parameters of the model, based on a sample of size 40000, after a burn-in period of 20000 iterates. Convergence was assessed for the parameters of the model specified in Section 4.1.

We observe that the posterior mean of $\kappa_y$ has a much smaller value than the posterior mean of $\kappa_x$ which is consistent with the fact that, in general, simulated data are shifted to the right in relation to the observed data, indicating the possible existence of some bias in the simulated data. The posterior mean of the precision (inverse of the variance) parameters for the space model ($\tau_W$) and for the temporal model ($\tau_Z$) suggest that time dependence is stronger than space dependence. The posterior mean for $\beta_y$ is slightly smaller than the posterior mean for $\beta_x$.
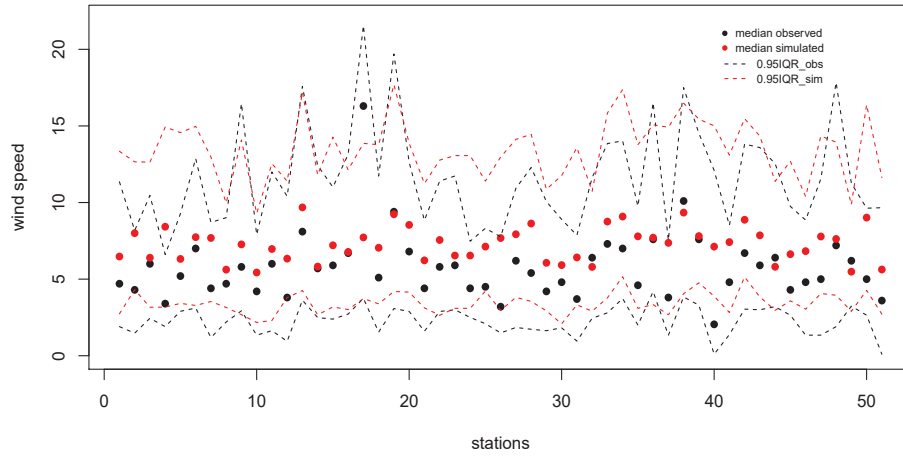
**Figure 2**: Median of observed and simulated wind speeds for the 51 stations, and the 95% IQR wind speeds by station (dashed lines).

**Table 1**:     Summary statistics for the marginal posterior distributions.

| Parameter | Mean | Standard deviation | 2.5% quantile | Median | 97.5% quantile |
|:---:|:---:|:---:|:---:|:---:|:---:|
| $\alpha$ | 0.6691 | 0.1007 | −0.4668 | −0.5028 | −0.8564 |
| $\beta_y$ | −1.1552 | 0.1702 | −1.4760 | −1.4250 | −0.8063 |
| $\beta_x$ | −0.9274 | 0.1614 | −1.2300 | −1.1830 | −0.5957 |
| $\kappa_y$ | −5.2951 | 0.1976 | −4.9020 | −4.9740 | −5.7020 |
| $\kappa_x$ | 18.7384 | 0.7467 | 17.2900 | 17.5100 | 20.3000 |
| $\tau_W$ | −3.5699 | 0.7274 | −2.3280 | −2.4840 | −5.1520 |
| $\tau_Z$ | −0.4240 | 0.1081 | −0.2457 | −0.2678 | −0.6675 |
| $\xi_y$ | −0.0703 | 0.0018 | −0.0739 | −0.0675 | −0.0670 |
| $\xi_x$ | −0.0806 | 0.0014 | −0.0834 | −0.0782 | −0.0777 |
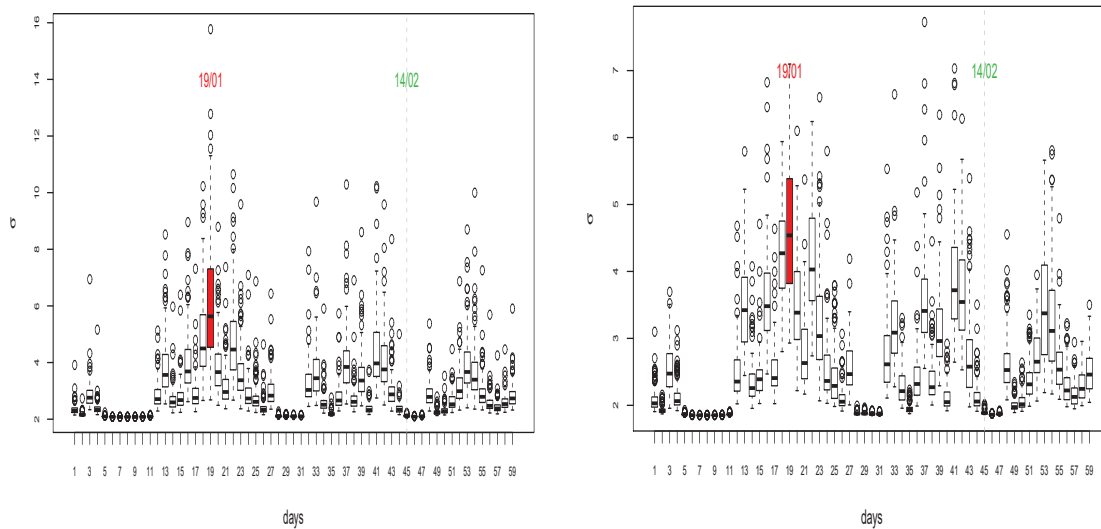


**Figure 3**: Boxplot of the posterior means of $\sigma_y(i,j)$ (left) and $\sigma_x(i,j)$ (right).

This naturally contributes for higher values for $\sigma_y(i,j)$ relatively to $\sigma_x(i,j)$ and with greater dispersion, as it can be seen in Figure 3 where we show daily boxplots of the posterior means of the parameters $\sigma(i,j), \forall j$ for both models. In that figure it is marked two dates, 19 of January, a day where it was observed a storm with heavy winds (storm GONG, maximum observed wind 29.6m/s), particularly in regions close to the littoral, and 14th of February, a very mild day all over the country (Valentine's day; maximum observed wind 8.20m/s). The variation observed along the days is consistent with the fact that on windy days the maximum wind speed along the stations varies much more than on mild days. Also the temporal dependence is clear in these pictures.

These two days were studied, in particular, for exemplification of the conditional quantile calibration method proposed. For the purpose of exemplification of the results we represent in Figures 4 and 5, on the left, a kernel density estimation (considering all the stations) for the observed and simulated maximum wind speed on that day, together with the mean of the posterior distribution of the calibrated data as defined in (3.3). On the right side, we represent the observed and simulated maximum wind speed on that day for each station, together with the calibrated data with the corresponding 95% credible interval.
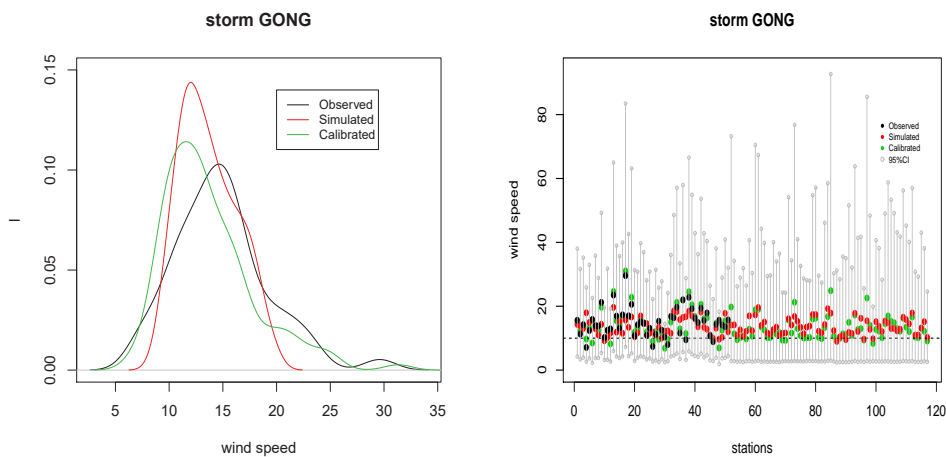


**Figure 4**: Kernel density estimation (left), observed and simulated maximum wind speed for each station, together with the mean of the posterior distribution for the calibrated data and 95% credible interval, for a storm day.
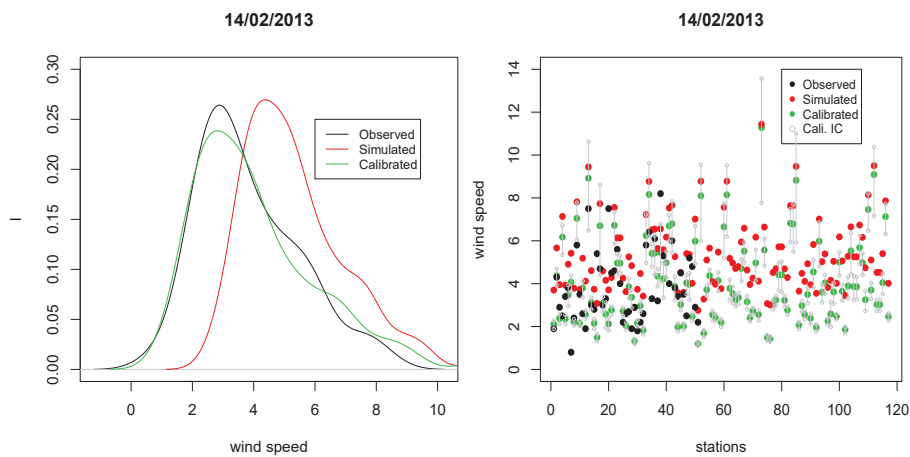


**Figure 5**: Kernel density estimation (left), observed and simulated maximum wind speed for each station, together with the mean of the posterior distribution for the calibrated data and 95% credible interval, for a mild day.

We observe that, on a storm day (Figure 4) the observed winds have longer tails than simulated winds. The calibration method was able to capture both tails of the distribution for the observed data, although it shifted the bulk of the distribution to the left. Notice that the 95% credible intervals are very large. This was to be expected due to the great variability of the simulated values from the posterior distribution of the shape parameter, as it can be observed in Figure 3. Regarding a mild windy day (Figure 5), the distribution of the simulated data is shifted to the right relatively to the distribution of the observed data with longer tails, as it was observed in a preliminary study. This bias is corrected with the calibration method. The 95% credible intervals are, in general, small.

In Figures 6 and 7 there is a spatial representation of the observed, simulated and calibrated values for each of these two days.
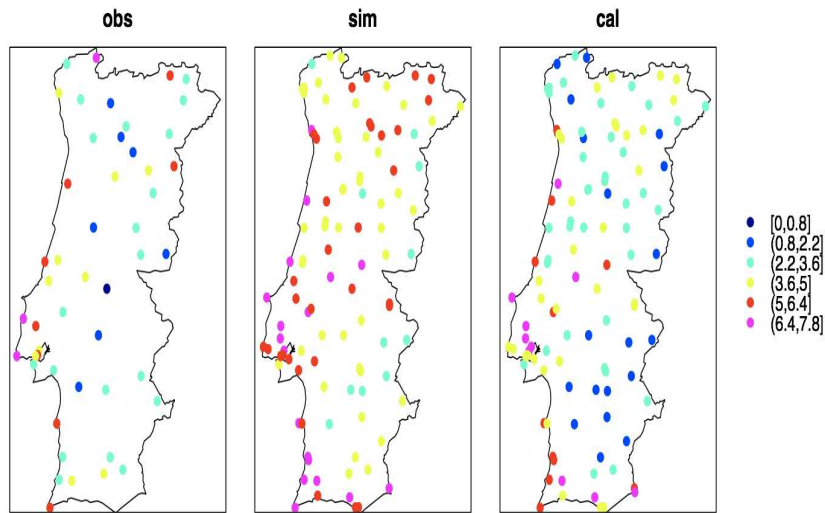


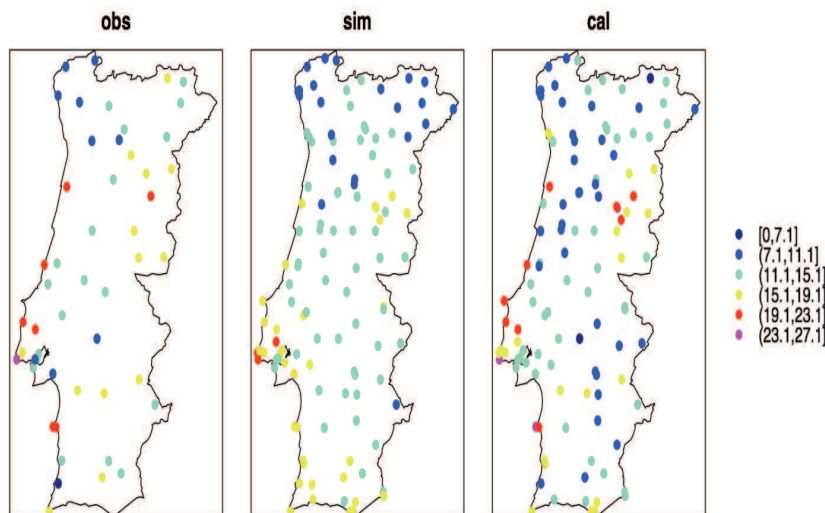**Figure 6**: Storm day: observed, simulated and calibrated maximum wind speeds.



**Figure 7**: Mild day: observed, simulated and calibrated maximum wind speeds.

## 4.1. Choice of prior specification and a sensitivity study

Here a justification is given on the particular choice of priors used in this study. Apart from the parameters $\alpha$ and $\xi$, and in the absence of adequate prior information, we considered relatively vague priors for the other parameters involved, as described in (3.2).

A preliminary data analysis of the wind speed data showed that observed and simulated data were consistent with the case $\xi < 0$. Without further prior information on $\xi$, we considered a uniform prior distribution with support between -0.5 and 0, since it is known from the extreme value theory that estimators with good properties for $\xi$ exist when $\xi > -0.5$.

In order to set a prior for the parameter $\alpha$, that controls the rate of decline of correlation with distance, we follow the suggestion by Thomas *et al.*, 2014 [21]. Accordingly, in the absence of prior information, "a sensible 'default' choice is to consider an upper prior bound equal to a small multiple of the maximum distance in the study region" and the lower prior bound should be larger than the minimum distance between observations. Coordinates of the 117 stations were given in decimal degrees and transformed into km to compute distances.

**Table 2**: Summary statistics for the marginal posterior distributions.

| Upper bound | Parameter | Mean | Standard deviation | 2.5% quantile | Median | 97.5% quantile |
|---|---|---|---|---|---|---|
| 0.5 | $\alpha$ | 0.4552 | 0.0403 | 0.3522 | 0.3733 | 0.4988 |
| 0.9 | $\alpha$ | 0.6691 | 0.1007 | 0.4668 | 0.5028 | 0.8564 |
| 1.2 | $\alpha$ | 0.6744 | 0.1075 | 0.4677 | 0.5020 | 0.8821 |
| 0.5 | $\beta_y$ | $-1.1510$ | 0.1505 | $-1.4390$ | $-1.3970$ | $-0.8641$ |
| 0.9 | $\beta_y$ | $-1.1552$ | 0.1702 | $-1.4760$ | $-1.4250$ | $-0.8063$ |
| 1.2 | $\beta_y$ | $-1.1356$ | 0.1514 | $-1.4250$ | $-1.3820$ | $-0.8353$ |
| 0.5 | $\beta_x$ | $-0.9170$ | 0.1381 | $-1.1820$ | $-1.1400$ | $-0.6549$ |
| 0.9 | $\beta_x$ | $-0.9274$ | 0.1614 | $-1.2300$ | $-1.1830$ | $-0.5957$ |
| 1.2 | $\beta_x$ | $-0.9014$ | 0.1412 | $-1.1590$ | $-1.1240$ | $-0.6171$ |
| 0.5 | $\kappa_y$ | 5.3136 | 0.1926 | 4.9500 | 5.0080 | 5.6970 |
| 0.9 | $\kappa_y$ | 5.2951 | 0.1976 | 4.9020 | 4.9740 | 5.7020 |
| 1.2 | $\kappa_y$ | 5.3056 | 0.1890 | 4.9390 | 5.0050 | 5.7110 |
| 0.5 | $\kappa_x$ | 18.7734 | 0.8037 | 17.2000 | 17.4800 | 20.3100 |
| 0.9 | $\kappa_x$ | 18.7384 | 0.7467 | 17.2900 | 17.5100 | 20.3000 |
| 1.2 | $\kappa_x$ | 18.7919 | 0.7302 | 17.4500 | 17.6200 | 20.2500 |
| 0.5 | $\tau_W$ | 4.1859 | 0.7078 | 2.9510 | 3.1180 | 5.7120 |
| 0.9 | $\tau_W$ | 3.5699 | 0.7274 | 2.3280 | 2.4840 | 5.1520 |
| 1.2 | $\tau_W$ | 3.5600 | 0.7378 | 2.2620 | 2.4460 | 5.1570 |
| 0.5 | $\tau_Z$ | 0.4244 | 0.1034 | 0.2519 | 0.2729 | 0.6535 |
| 0.9 | $\tau_Z$ | 0.4240 | 0.1081 | 0.2457 | 0.2678 | 0.6675 |
| 1.2 | $\tau_Z$ | 0.4101 | 0.0981 | 0.2454 | 0.2665 | 0.6256 |
| 0.5 | $\xi_y$ | $-0.0702$ | 0.0018 | $-0.0738$ | $-0.0673$ | $-0.0668$ |
| 0.9 | $\xi_y$ | $-0.0703$ | 0.0018 | $-0.0739$ | $-0.0675$ | $-0.0670$ |
| 1.2 | $\xi_y$ | $-0.0702$ | 0.0017 | $-0.0736$ | $-0.0674$ | $-0.0670$ |
| 0.5 | $\xi_x$ | $-0.0805$ | 0.0016 | $-0.0836$ | $-0.0778$ | $-0.0774$ |
| 0.9 | $\xi_x$ | $-0.0806$ | 0.0014 | $-0.0834$ | $-0.0782$ | $-0.0777$ |
| 1.2 | $\xi_x$ | $-0.0804$ | 0.0014 | $-0.0831$ | $-0.0781$ | $-0.0776$ |

In order to avoid large numbers, we consider 100km as unit to compute distances. The maximum Euclidean distance among the 117 stations was computed as 5.65 and the minimum 0.01. Hence we set a uniform prior between 0.1 and 0.9. However we performed a sensitivity study considering as upper bound 0.5, 0.9 and 1.2, keeping the other priors unchanged. Summary results are displayed in Table 2 and in Figure 8.
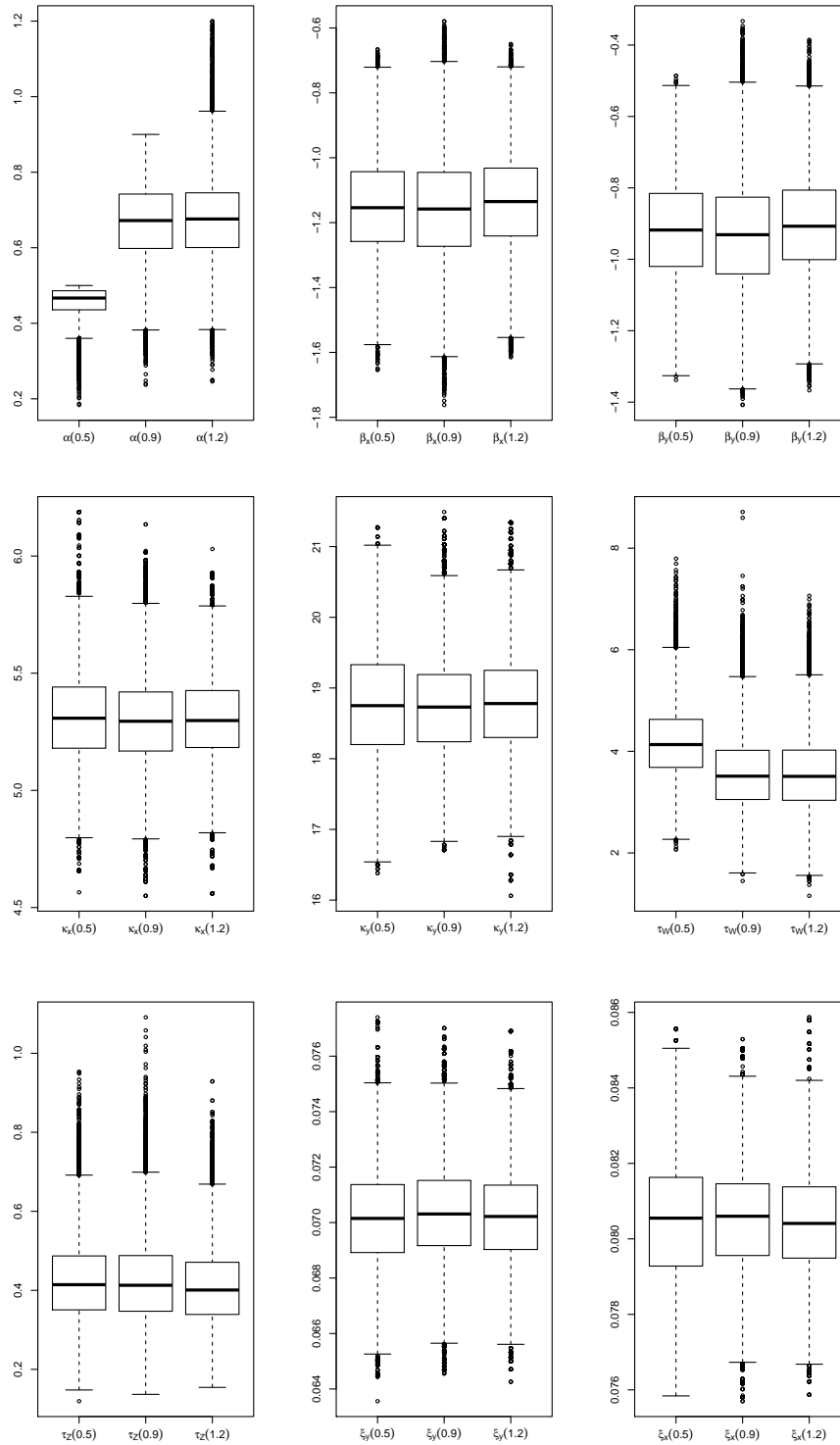


**Figure 8**: Boxplots of the posterior distribution of the parameters for different values for the upper limit of the prior for $\alpha$.

As it can be seen, posterior distribution of $\alpha$ is sensitive to the its prior. The upper bound 0.5 seems not to be adequate since it is clear that there is a concentration of mass near the upper bound. However, when the upper bound is 0.9 or 1.2, the influence of the prior for $\alpha$ on its posterior is much less evident, particularly on the bulk of the posterior distribution. Influence of the prior on $\alpha$ on the other parameters is almost negligible. Basically there is only some influence on the posterior distribution of the parameter $\tau_W$, although this influence is softened while comparing the prior with upper bound 0.9 and 1.2. More important it is that there is no influence of the prior for $\alpha$ regarding the calibrated data (posterior distribution of $F_Y^{-1}(F_X(x(s_i, t_j)))$ at $s_i, i = 1, ..., N_s$ and time $t_j, j = 1, ..., T$).

## 5.    DISCUSSION AND FURTHER EXTENSIONS

In this article we proposed a hierarchical Bayesian approach to implement a conditional quantile matching calibration (CQCM) using a space-time extended generalized Pareto distribution for both the observed and simulated data.

The performance of the CQCM method was exemplified with two specific days, a storm day and a mild day. In both cases the calibrated data matched well the observed data on the tails, although on the storm day it did not capture well the bulk of the distribution. Also the 95% credible intervals were quite wide for the storm day, which may be an indication that appropriate methods to deal with extreme data should instead be considered to accommodate these rare situations.

Ideally this method should be extended to the grid level, since the simulator produces data at a fine grid level and this is much more interesting if the objective is the construction of a risk map. However this extension is not trivial and some assumptions regarding the model structure have to be assumed.

Damages in electricity grid are basically governed by extreme winds and primarily simulated and observed data coming from the right tail differ. Hence adequate calibration methods must be specifically adapted to extreme observations coming from the right tails and methods and models to be used in calibration should ideally be compatible with extreme value theory. A range of approaches for characterising the extremal behaviour of spatial process have been suggested and a brief comparison of these methods can be found in Tawn *et al.* (2018) [20]. Downscaling method described by Towe *et al.* (2017) [22] — based on the conditional extremes process — is more suitable, with adequate modifications, to calibrate extreme simulated data based on observed wind speeds. Work on this approach is under progress. Alternatively, calibration methods based on bivariate max stable processes (Genton *et al.*, 2015 [10]) can be devised, although this would require substantial computational complications.

## REFERENCES

[1]   BANERJEE, S.; CARLIN, B.P. and GELFAND, A.E. (2004). *Hierarchical Modeling and Analysis for Spatial Data*, Boca Raton, FL: Chapman and Hall.

[2]   BEIRLAND, J.; GOEGEBEUR, Y.; SEGERS, J. and TEUGELS, J. (2004). *Statistics of Extremes: Theory and Applications*, J. Wiley, Chichester.

[3]   BERROCAL, V.J. (2019). *Data assimilation.* In "Handbook of Environmental and Ecological Statistics" (A.E. Gelfand; M. Fuentes; J.A. Hoeting and R.L. Smith, Eds.), Chapman and Hall/CRC, 133–151.

[4]   BERROCAL, V.J.; GELFAND, A.E. and HOLLAND, D.M. (2012). Space-time data fusion under error in computer model output: An application to modeling air quality, *Biometrics*, **68**(3), 837–848.

[5]   BERROCAL, V.J.; GELFAND, A.E. and HOLLAND, D.M. (2014). Assessing exceedance of ozone standards: A space-time downscaler for fourth highest ozone concentrations, *Environmetrics*, **25**, 279–291.

[6]   CARDOSO, R.M.; SOARES, P.M.M.; MIRANDA, P.M.A. and BELO-PERIRA, M. (2013). WRF high resolution simulation of Iberian mean and extreme precipitation climate, *International Journal of Climatology*, **33**, 2591–2608.

[7]   DE CARVALHO, M.; PEREIRA, S.; PEREIRA, P. and DE ZEA BERMUDEZ, P. (2020). An extreme value Bayesian lasso for the conditional bulk and tail, https://arxiv.org/abs/2010.07164

[8]   DE ZEA BERMUDEZ, P. and KOTZ, S. (2010). Parameter estimation of the generalized Pareto distribution, Part I, *Journal of Statistical Planning and Inference*, **140**, 1353–1373.

[9]   FUENTES, M. and RAFTERY, A.E. (2005). Model evaluation and spatial interpolation by Bayesian combination of observations with outputs from numerical models, *Biometrics*, **61**, 36–45.

[10]  GENTON, M.G.; PADOAN, S.A. and SANG, H. (2015). Multivariate max-stable spatial processes, *Biometrika*, **102**, 215–230.

[11]  KALNAY, E. (2003). *Atmospheric Modeling, Data Assimilation and Predictability*, Cambridge University Press, Cambridge.

[12]  MCMILLAN, N.J.; HOLLAND, D.M.; MORARA, M. and FENG, J. (2010). Combining numerical model output and particulate data using Bayesian space-time modeling, *Environmetrics*, **21**, 48–65.

[13]  MICHELANGELI, P.A.; VRAC, M. and LOUKOS, H. (2009). Probabilistic downscaling approaches: Application to wind cumulative distribution functions, *Geophysical Research Letters*, **36**, 1–6.

[14]  NAVEAU, P.; HUSER, R.; RIBEREAU, P. and HANNART, A. (2016). Modeling jointly low, moderate, and heavy rainfall intensities without a threshold selection, *Water Resources Research*, **52**, 2753–2769.

[15]  PAPASTATHOPOULOS, I. and TAWN, J.A. (2013). Extended generalised Pareto models for tail estimation, *Journal of Statistical Planning and Inference*, **143**, 131–143.

[16]  PEREIRA, S.; PEREIRA, P.; DE CARVALHO, M. and DE ZEA BERMUDEZ, P. (2019). *Calibration of extreme values of simulated and real data.* In "Proceedings of International Workshop on Statistical Modelling, Volume I" (L. Meira-Machado and G. Soutinho, Eds.), INE, Guimarães, 147–150.

[17]  RUE, H. and HELD, L. (2005). *Gaussian Markov Random Fields: Theory and Applications*, Monographs on Statistics and Applied Probability, vol. 104, Chapman & Hall, London.

[18] Skamarock, W.C.; Klemp, J.B.; Dudhia, J.; Gill, D.O.; Barker, D.M.; Duda, M.G.; Huang, X.Y.; Wang, W. and Powers, J.G. (2008). A description of the advanced research WRF version 3, *NCAR Technical Notes*, TN-475_STR, 1–125.

[19] Sturtz, S.; Ligges, U. and Gelman, A. (2005). R2WinBUGS: A Package for Running WinBUGS from R, *Journal of Statistical Software*, **12**, 1–16.

[20] Tawn, J.; Shooter, R.; Towe, R. and Lamb, R. (2018). Modelling spatial extreme events with environmental applications, *Spatial Statistics*, **28**, 39–58.

[21] Thomas, A.; Best, N.; Lunn, D.; Arnold, R. and Spiegelhalter, D. (2014). *GeoBUGS User Manual, Version 3.2.3*, MRC Biostatistics Unit, Cambridge.

[22] Towe, R.P.; Sherlock, E.F.; Tawn, J.A. and Jonathan, P. (2017). Statistical downscaling for future extreme wave heights in the North Sea, *Annals of Applied Statistics*, **11**, 2375–2403.

[23] Zidek, J.V.; Le, N.D. and Liu, Z. (2012). Combining data and simulated data for space-time fields: Application to ozone, *Environmental and Ecological Statistics*, **19**, 37–56.