

# Extreme Precipitation in Madagascar Part II: Establishing Frequency Isohyets by Spatial Interpolation (Kriging)

Justin Ratsaramody

Laboratoire d'Hydraulique, Ecole Supérieure Polytechnique, Université d'Antsiranana, Madagascar

**Abstract** In this second part of the studies on extreme rainfall in Madagascar, the central question is that of spatial interpolation in order to produce isohyet maps corresponding to different return periods with processing the series of maximum annual daily rainfall from 1970 to 2023 across 132 observed sites. The results showed that, contrary to traditional practice in Madagascar, the data from all the sites could not be modelled solely by a Gumbel distribution but that a significant proportion (28.7%) had to be modelled by a GEV distribution to take account of climate change. The results also showed the presence of a spatial trend in the form of an affine surface that had to be removed before any spatial interpolation. Normalization with the Box-Cox transformation of the raw data was made necessary by the non-stationary nature of the raw data. There is geometric anisotropy in the data, but this had no significant influence on the spatial interpolation results. Compared through cross-validation and diagnostic metrics, it was found that the Ordinary Kriging interpolation model was better than Universal Kriging even though the variances of the prediction residuals of these two models were almost identical. Cross-validation showed that the two models both overestimated predictions beyond a threshold value, but that these deviations were negligible. The spatial interpolation results obtained from this work using Ordinary Kriging, which was then used to draw the frequency isohyet maps, are extremely accurate because the variances are of the order of a millimetre and these differences are generally outside the Madagascar area.

**Keywords** Spatial structure, Trend surfaces, Box-Cox, Ordinary Kriging, Universal Kriging, Frequency isohyets

## 1. Introduction

In the first part of this series of two articles [1], the process of frequency analysis of extreme daily maximum rainfall data for Madagascar using the Block Maxima method was studied for the 1970-2023 and 2000-2023 time series, and the results for the six provincial capitals were presented as an example. In particular, these results showed that stationary GEV models produced unrealistic results and that the return periods for the same extreme event were different because they depended on the size of the time series considered. In addition, despite the fact that the two time series were nested, it was also found that the trends, reflecting climate change, were different and even in the opposite direction in some cases.

For better statistical modelling, the longest series (1970-2023) had then been studied and, to take account of climate change, non-stationary GEV distribution models were applied to this time series with variation as a function

of time of the position parameter  $\mu(t) = \mu_0 + \mu_1 t$  and the scale parameter  $\sigma(t) = \sigma_0 + \sigma_1 t$  but keeping the shape parameter  $\xi$  constant. This resulted in eight possible non-stationary models depending on the values of the intercepts  $\mu_0$  and  $\sigma_0$  and the values of the slopes  $\mu_1$  and  $\sigma_1$ . These models also included the special case of the Gumbel distribution when the slopes were zero, i.e.  $\mu_1 = 0$  and  $\sigma_1 = 0$ . This led to much more realistic results, but in which the one-to-one relationship between the return period and the return level was called into question by the presence of a third parameter interpreted as the risk level.

Even if this process of frequency analysis were absolutely necessary to determine the return period and/or the return level at a given spatial point, it is extremely cumbersome and tedious and may not be within everyone's reach, particularly for engineers in charge of establishing design flows for hydraulic and road infrastructures, etc. In this case, a map of frequency isohyets is much more practical because, for a return period and an estimated level of risk, all you have to do is identify the spatial point concerned on the map and read the value of the corresponding return period. This would considerably reduce the time spent on this activity in projects and studies in various fields such as flood risk assessment and mitigation, environmental impact studies,

\* Corresponding author:

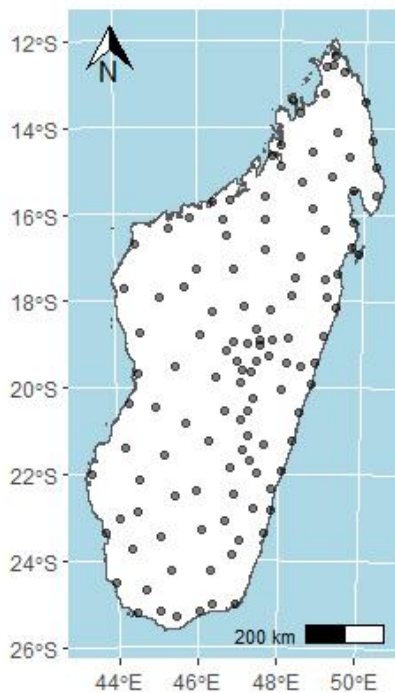
justinratsaramody@yahoo.fr (Justin Ratsaramody)

Received: Jun. 1, 2024; Accepted: Jun. 19, 2024; Published: Jun. 22, 2024

Published online at <http://journal.sapub.org/re>

land-use planning, river restoration and the design of hydraulic structures, etc. [2].

Such maps are produced using spatial interpolation methods because of the presence of ungauged areas, which is the case for virtually all of Madagascar, and the entire frequency analysis process described above was carried out beforehand on the 132 measurement points shown in Figure 1. These points were chosen at random but so that there was at least 1 measurement point in all 119 districts of Madagascar.



**Figure 1.** Locations of rainfall data sites on the island of Madagascar

As already indicated in the first article, the overall aim of the study is to update this type of isohyet map, whose most recent values for the whole of Madagascar, to our knowledge, date from 1993 [3]. There are practically no publications on Madagascar either on isohyets or on spatial interpolation and the only recent publication we have found on this subject is an article dating from 2018 which deals with the use of historical normals to improve the modern monthly climate normal surfaces for Madagascar [4].

The literature describes several possible spatial interpolation methods, including purely geometric methods such as Thiessen polygons, deterministic methods such as Inverse Distance Weighting and Thin Plate Spline surfaces, and especially geostatistical methods such as the various forms of kriging. The vast majority of these studies have generally concluded that kriging methods are superior, in various fields such as groundwater studies, extreme or non-extreme rainfall, etc., and with various time steps, monthly or annual [5-16]. In line with this conclusion, this article considers only kriging interpolation methods.

As already indicated above, the use of non-stationary GEV distribution models brings into play, for each return

period, a third parameter interpreted as being the notion of risk [17] and which corresponds to the uncertainty limits of the return levels derived from the a posteriori probabilities [18]. As such, these uncertainty limits vary from 0% to 100% and it is not possible to draw up isohyet maps for a return period for the full extent of these uncertainty limits. In this article, therefore, we will confine ourselves to calculating and then plotting the frequency isohyets for two risk levels: the median level (probability = 50%) and the very high level (probability = 90%), even though some authors suggest that the probability to be assigned to a given return period for a very high risk level is the third quartile (probability = 75%) for the final return level [17].

The isohyet maps envisaged for this work concern return periods of 10, 20, 50, 100 and 200 years, but with two risk levels (M: median and VH: very high) per return period, this leads to 10 isohyet maps. However, for lack of space, only the procedures and results for the 10-year return period VH will be presented in this article, but the procedures are strictly identical for the other return periods, with obviously different results. The 10-year return period was chosen because it is generally the contractual return period for engineering works in rural Madagascar.

## 2. Materials

### 2.1. Rainfall Data

The rainfall data are those used in the first article, i.e. daily rainfall data from 1970 to 2023, combined with GLDAS (1970 to 2000), TRMM (2001 to 2019) and GPM (2020 to 2023) data. They were obtained from gridded data, but reduced to the measurement points (Figure 1) and downloaded from the <https://giovanni.gsfc.nasa.gov/website>.

### 2.2. Digital Elevation Model (DEM)

The DEM used for this study was a set of DEMs produced by SRTM (Shuttle Radar Topography Mission) and available at <https://earthexplorer.usgs.gov/>. After being assembled, the various tiles were projected onto the CSR EPSG: 29702, which is the CSR covering the entire island of Madagascar.

### 2.3. Spatial Interpolation Grid

The spatial interpolation grid used for this study was a grid with square cells of dimensions 5000 m × 5000 m, which provides sufficient accuracy given the large size of the country (1700 km × 500 km) while maintaining a certain fluidity in the calculations.

### 2.4. Calculation Tools

All calculations relating to this study were carried out using the free software R 4.3.2 (<https://www.r-project.org/>) with the corresponding packages, mainly the **gstat** package [19] and its dependencies. Other packages were also used for the various operations required.

### 3. Methodology

#### 3.1. Preliminary Analysis

In order to determine the return levels corresponding to the different return periods  $T_m$ ,  $m = 10, 20, 50, 100$  and 200 years, the entire procedure described in the first article was applied to each of the 132 measurement points shown in Figure 1 with non-stationary GEV distribution models.

#### 3.2. Exploratory Analysis and Spatial Structure of the Data

In this analysis, the aim was to examine both the statistical distribution of the various values of  $T_m$  (24-hour rainfall with a return period of  $m$  years) and the spatial distribution of these values in order to detect a spatial trend. Indeed, as we propose to carry out an interpolation by kriging, this is only possible if the process that generated the values at the observation points is a random process following a normal distribution and if, spatially, they have a stationary structure (Gaussian fields) [10]. To verify the normality of the data, summary statistics were calculated to account for the asymmetry and dispersion of the data. The presence of a spatial trend was also highlighted using a map of observed points weighted by their values, as well as the polynomial regression of  $T_m$  values against X and Y coordinates.

#### 3.3. Normalisation of the Data

For geostatistical modelling to be possible, the data must be (approximately) normally distributed and stationary, i.e. the mean and variance must not vary significantly in space. To make the  $y$ -values of  $T_m$  approximate to a normal distribution, the well-known Box-Cox transformation was used [20,21]:

$$y^* = f(y, \lambda) = \begin{cases} \frac{y^\lambda - 1}{\lambda} & (\lambda \neq 0) \\ \log y & (\lambda = 0) \end{cases} \quad (1)$$

Following this transformation, the  $y$  data was replaced by  $y^*$ , which follows a normal distribution. The parameter  $\lambda$  was determined using the maximum likelihood (ML) method. These transformed  $y^*$  data were then used for spatial interpolation. To return to the initial data  $y$ , the inverse transformations are immediate by inverting equations (1):

$$y = \begin{cases} (1 + \hat{\lambda} y^*)^{1/\hat{\lambda}} & \hat{\lambda} \neq 0 \\ \exp(y^*) & \hat{\lambda} = 0 \end{cases} \quad (2)$$

where  $\hat{\lambda}$  is the estimated value of the parameter  $\lambda$ .

#### 3.4. Spatial Interpolation

Generally speaking and whatever the method used, the objective of spatial interpolation is to determine the value  $\hat{z}$  at a non-gauged location,  $u_0$ , knowing the values at  $n$  locations  $z(u_1), z(u_2), \dots, z(u_n)$ , where  $u_i = (x_i, y_i)$  is a location and  $x_i$  and  $y_i$  are the coordinates in geographical space. Thus, at an ungauged point  $u_0$ , this value  $\hat{z}$  is given by

$$\hat{z}(u_0) = \sum_{i=1}^n \omega_i(u_0) \cdot z(u_i) \quad (3)$$

where  $\omega_i$  is the weight for gauged station  $i$ .

However, as already indicated above, the spatial interpolation considered in this study is spatial interpolation by kriging, despite the fact that there are other forms of spatial interpolation, such as the very popular IDW [12,22], which we did not use because of the subjectivity involved in choosing the power parameter to be applied to the inverse distance. The kriging interpolation method differs in that it provides an unbiased estimate by considering the autocorrelation of all available data on the basis of various semivariograms [23].

##### 3.4.1. Empirical Semivariogram

Under the assumption of intrinsic stationarity, it is assumed that the process which generated the samples at the measured points  $\mathbf{u}$  is a random function  $Z(\mathbf{u})$  composed of a mean  $m$  and a residual  $\varepsilon(\mathbf{u})$  [24,25]:

$$Z(\mathbf{u}) = m + e(\mathbf{u}) = \mathbb{E}[Z(s)] + \varepsilon(\mathbf{u}) \quad (4)$$

as well as a variogram:

$$\gamma(h) = \frac{1}{2} \mathbb{E}[Z(\mathbf{u}) - Z(\mathbf{u} + \mathbf{h})]^2 \quad (5)$$

Under this assumption, the variance of  $Z$  is constant and the spatial correlation of  $Z$  does not depend on the location  $\mathbf{u}$ , but only on the separation distance  $\mathbf{h}$ . It is then possible to form several pairs  $\{z(\mathbf{u}_i), z(\mathbf{u}_j)\}$  whose separation vectors  $\mathbf{h} = \mathbf{u}_i - \mathbf{u}_j$  are (almost) identical and estimate the correlation from these. The empirical semivariogram obtained from the transformed data  $y^*$  in equation (1) is then defined according to equation (6):

$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [z(u_i) - z(u_i + h)]^2 \quad (6)$$

where  $\hat{\gamma}(h)$  is the measured semivariance for distance  $h$ ,  $N(h)$  is the number of data pairs separated by distance  $h$  and  $z(u_i)$  is the value of the variable at location  $u_i$ .

In addition to what has been done in the exploratory analysis, the empirical semivariogram is a valuable exploratory tool for assessing the presence of spatial correlation in the data. It should also be noted that, under the assumption of intrinsic second-order stationarity, the empirical semivariogram expressed by equation (6) is an unbiased estimator of the theoretical semivariogram expressed by equation (7) below because the unknown mean of the random function is not involved in the expression of the empirical semivariogram in equation (6).

##### 3.4.2. Theoretical Semivariogram Models

As the empirical semivariogram only covers the observed points, it has to be adjusted by a theoretical model to cover the entire space of points to be interpolated. Under the assumption of second-order stationarity and the intrinsic assumption, the semivariogram shows how the dissimilarity

between  $z(\mathbf{u} + \mathbf{h})$  and  $z(\mathbf{u})$  changes with distance  $\mathbf{h}$  [26]:

$$\begin{aligned}\gamma(h) &= \frac{1}{2} \text{Var}[z(\mathbf{u} + \mathbf{h}) - z(\mathbf{u})] \\ &= \frac{1}{2} \mathbb{E}[(z(\mathbf{u} + \mathbf{h}) - z(\mathbf{u}))^2]\end{aligned}\quad (7)$$

During this study, after a visual examination of the empirical semivariogram, three of the most common theoretical models were tested:

- the exponential model:

$$\gamma(|\mathbf{h}|) = m \left[ 1 - \exp\left(-\frac{|\mathbf{h}|}{a}\right) \right] \quad (8)$$

where  $m = C(\mathbf{0})$  is the value of the semivariogram when it reaches the sill, and  $a$  is the range. For this model, the practical range  $a'$  is defined as the distance for which

$$\gamma(a') = m \left[ 1 - \exp\left(-\frac{a'}{a}\right) \right] = 0.95m \quad (9)$$

i.e. the distance above which the covariance represents only 5% of the value recorded at the origin. This means that  $a' \cong 3a$ .

- the spherical model:

$$\gamma(|\mathbf{h}|) = \begin{cases} m \left[ 1.5 \frac{|\mathbf{h}|}{a} - 0.5 \left( \frac{|\mathbf{h}|}{a} \right)^3 \right] & \text{si } |\mathbf{h}| \leq a \\ m & \text{si } |\mathbf{h}| > a \end{cases} \quad (10)$$

with the same definitions for the sill  $m$  and the range  $a$ .

- the Gaussian model:

$$\gamma(|\mathbf{h}|) = m \left[ 1 - \exp\left(-\frac{|\mathbf{h}|^2}{a}\right) \right] \quad (11)$$

with the same definitions for the sill  $m$  and the range  $a$ .

### 3.4.3. Anisotropy

When the theoretical models given by equations (5), (7) and (8) are isotropic, the overall omnidirectional semivariogram depends only on the length of vector  $\mathbf{h}$ , i.e.  $h = \|\mathbf{h}\| = \sqrt{h_x^2 + h_y^2}$ ,  $h_x$  and  $h_y$  being the components of vector  $\mathbf{h}$  projected onto X and Y. However, anisotropy has an important effect on kriging weights because, if it exists, the weight of points located in the main direction of the anisotropy ellipse will increase: it is therefore essential to analyse the presence of anisotropy in the data. When geometric anisotropy exists in the stationary case, it implies that the thresholds of the directional semivariograms, although equal, are reached with two different distances (ranges). The anisotropy ratio is defined as follows [12]:

$$R = \frac{r_{\min}}{r_{\max}} \quad (12)$$

where  $r_{\max}$  is the largest range in the direction of the major axis and  $r_{\min}$  is the smallest span in the direction of the minor axis. The spatial variation between any pair  $(x_1, y_1)$  and  $(x_2, y_2)$  is then calculated by the distance  $h'$  in the maximum continuity axis variogram model according to the rotation angle  $\varphi$  and the anisotropy ratio  $R$  where  $h'$  is given by equation (11) below [12,27]:

$$h' = \sqrt{\frac{[(x_1 - x_2) \cos \varphi + (y_1 - y_2) \sin \varphi]^2}{\frac{1}{R^2} [(y_1 - y_2) \cos \varphi + (x_1 - x_2) \sin \varphi]^2}} \quad (13)$$

Instead of  $\hat{\gamma}(h)$ , the omnidirectional empirical semivariogram  $\hat{\gamma}(h')$  will then be used to be fitted by the best of one of the theoretical models (equations (7), (9) and (10)) for kriging interpolation.

### 3.4.4. Ordinary Kriging

We have Ordinary Kriging (OK) when, in equation (4) above, the mean  $m$  is constant (or zero) and the kriging is then done on the residuals  $\varepsilon(\mathbf{u})$ . For Ordinary Kriging (OK), the weights  $\omega_i(u_0)$  in equation (3) are obtained by solving the following system of linear equations [12,28]:

$$\begin{cases} \sum_{i=1}^n \omega_i^{OK}(u_0) \gamma(u_i, u_j) + \mu = \gamma(u_j, u_0) & \text{for all } j \\ \sum_{i=1}^n \omega_i^{OK}(u_0) = 1 \end{cases} \quad (14)$$

where  $\gamma(u_i, u_j)$  is the value of the theoretical semivariogram function for the distance between  $u_i$  and  $u_j$ ;  $\gamma(u_j, u_0)$  is the value of the theoretical semivariogram of the distance between  $u_j$  and  $u_0$ ; and  $\mu$  is the Lagrange parameter. The second equation in (14) ensures that the interpolation is unbiased.

### 3.4.5. Universal Kriging (UK)

When, in equation (4), the mean  $m$  of the measured data is no longer stationary but depends on the X and Y coordinates, this confirms the presence of a trend that needs to be removed, and interpolation then needs to be performed on the remaining residuals. Interpolation by kriging can only be performed on random fields that are substantially Gaussian [10]. The predictions  $\hat{z}$  at the ungauged locations  $u_0$  are then obtained from:

$$\hat{z}(u_0) - \hat{m}_{UK}(u_0) = \sum_{i=1}^n \omega_i^{UK}(u_0) [z(u_i) - \hat{m}_{UK}(u_i)] \quad (15)$$

The Universal Kriging equations are then identical to those of Ordinary Kriging (equations (14)) but with the weights  $\omega_i^{OK}$  replaced by  $\omega_i^{UK}$  from equation (15). As for the mean  $\hat{m}_{UK}$  which includes the spatial tendency to be removed, it is generally expressed in the form of a first or second degree polynomial [25,29]:

$$m_{UK}(u) = \beta_0 + \beta_1 X + \beta_2 Y + \varepsilon'(u) \quad (16)$$

or

$$\begin{aligned} m_{UK}(u) &= \beta_0 + \beta_1 X + \beta_2 Y + \beta_3 X^2 \\ &\quad + \beta_4 Y^2 + \beta_5 XY + \varepsilon'(u) \end{aligned} \quad (17)$$

where  $\beta_j$  are regression coefficients to be determined and  $\varepsilon'(u)$  are regression residuals. In the case of equation (16), the trend surface is an affine surface while in the case of equation (17), the trend surface is a quadratic surface.

### 3.5. Assessment of Model Quality

Given the large amount of modelling to be carried out in this work, it was necessary to have criteria and/or procedures for evaluating the quality of these models in order to choose the best model. The performance of the final theoretical semivariogram model, and hence of the kriging predictions, is assessed using leave-one-out cross-validation [30], a procedure in which the observed values of the data are eliminated one by one and then predictions are made for each of these data using the remaining data [31,32]. The predicted values are then compared with the observed values.

The quality of the interpolation fit was also assessed by comparing the following diagnostic metrics [22,33,34]: the mean error ME, which measures bias, the mean absolute error MAE and the root mean square error RMSE:

$$ME = \frac{1}{n} \sum_{i=1}^n (z_i - \hat{z}_i) \quad (18)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |z_i - \hat{z}_i| \quad (19)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (z_i - \hat{z}_i)^2} \quad (20)$$

where  $z_i$  and  $\hat{z}_i$  are the observed and predicted values, respectively, at the same point  $i$ ,  $n$  is the number of observation points. The closer the values of ME, MAE and RMSE are to zero, the better the fit.

## 4. Results and Discussion

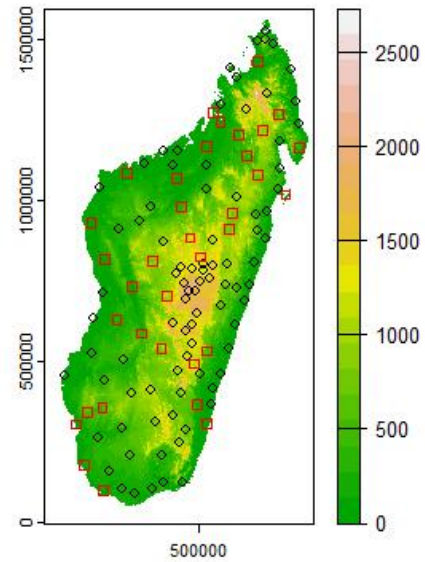
As already indicated, the results shown in the figures are for the 10-year return period T10VH (Very High), but the tables show all the results.

### 4.1. Results of the Preliminary Analysis

In a non-stationary context and after applying the procedures described in the first article for the 1970-2023 series of annual maxima, the distributions found for the 132 data points are shown in Figure 2. For the GEV distributions, the shape coefficient was such that  $\xi \neq 0$  while for the Gumbel distributions, we had  $\xi = 0$  for this shape coefficient.

Traditionally, the Gumbel distribution had always been the distribution used in Madagascar. However, the results show that this is not entirely correct, since 28.7% of the measurement points here follow a GEV distribution, even though the vast majority (71.3%) follow a Gumbel distribution.

Figure 2 also shows that there is no preferred location for the nature of the distribution (GEV or Gumbel) either in relation to coastal areas or in relation to elevation.



**Figure 2.** Distribution models on a terrain relief background (elevation in metres). Black circles: Gumbel model (87), red squares: GEV model (35)

### 4.2. Results of the Exploratory Analysis and the Spatial Structure of the Data

The skewness coefficients for the data leading to the 10 maps considered are shown in Table 1.

**Table 1.** Asymmetry coefficients

| Data  | Skewness | Data   | Skewness |
|-------|----------|--------|----------|
| T10M  | 0.790183 | T50VH  | 0.704151 |
| T10VH | 0.490097 | T100M  | 0.814530 |
| T25M  | 0.630158 | T100VH | 0.953927 |
| T25VH | 0.546100 | T200M  | 1.077816 |
| T50M  | 0.650116 | T200VH | 1.255942 |

The spatial structure of the data for the return period T10VH is shown in Figure 3 below. Figure 3, left-hand side, shows that there is an overall increase in T10VH values in the west-east direction and also in the south-north direction, although there are a few points that do not conform to this observation. This is proof of the presence of a spatial trend, which is confirmed by the polynomial regressions on the right-hand side of Figures 3 and 4. These regression curves show that along X, the polynomial regression curves of order 2 and order 3 are practically identical, whereas along Y, the regressions of order 1 and order 2 are identical. The results of the evaluations of this spatial trend are given below.

### 4.3. Results of Trend Surface Calculations

As the presence of a trend was highlighted above in Section 4.2, it is important to confirm its existence in the data. Note that if the raw data contains a trend, the same will also be true for the Box-Cox transformed data (BC data). The  $\beta_i$  parameters of the trend surface equations (16) and (17) had been determined by ordinary least squares and are shown in Tables 2 and 3.

**Table 2.** Parameters of the 1st order trend surfaces of the raw data

| Raw Data | $\beta_0$ | $\beta_1$ | $\beta_2$ |
|----------|-----------|-----------|-----------|
| T10M     | 72.27     | 7.96E-05  | 6.90E-06  |
| T10VH    | 84.92     | 1.05E-04  | 1.08E-05  |
| T25M     | 91.06     | 1.02E-04  | 2.37E-06  |
| T25VH    | 106.35    | 1.38E-04  | 5.24E-06  |
| T50M     | 107.43    | 1.22E-04  | -3.87E-06 |
| T50VH    | 125.29    | 1.67E-04  | -2.75E-06 |
| T100M    | 126.37    | 1.47E-04  | -1.33E-05 |
| T100VH   | 147.45    | 2.03E-04  | -1.50E-05 |
| T200M    | 148.65    | 1.76E-04  | -2.67E-05 |
| T200VH   | 173.78    | 2.45E-04  | -3.26E-05 |

Graphically, and for example for T10VH data, these trend surfaces are shown in Figures 4 and 5.

The maps for the other data also show similar patterns (but with different values).

To determine which of the two models to adopt, the residuals from each of these two models were tested for independence (Durbin-Watson test), normality (Shapiro-Wilk test) and homoscedasticity (Breusch-Pagan test) of variances.

For all the data, both models passed the independence and homoscedasticity tests, but only the 1st order models passed the normality tests, while the 2nd order models did not. Consequently, the 1st order models were adopted. This is confirmed by Table 3, which shows that the  $\beta_i$  coefficients,  $i = 4$  to 5, of the second-order terms are all less than  $10^{-9}$  for the second-order surface. This decision is consistent with the principle of parsimony in statistical modelling.

Visually, if we refer to Figures 4 and 5, we can see that the high variances are found outside the island, which is normal since these are non-sampled areas.

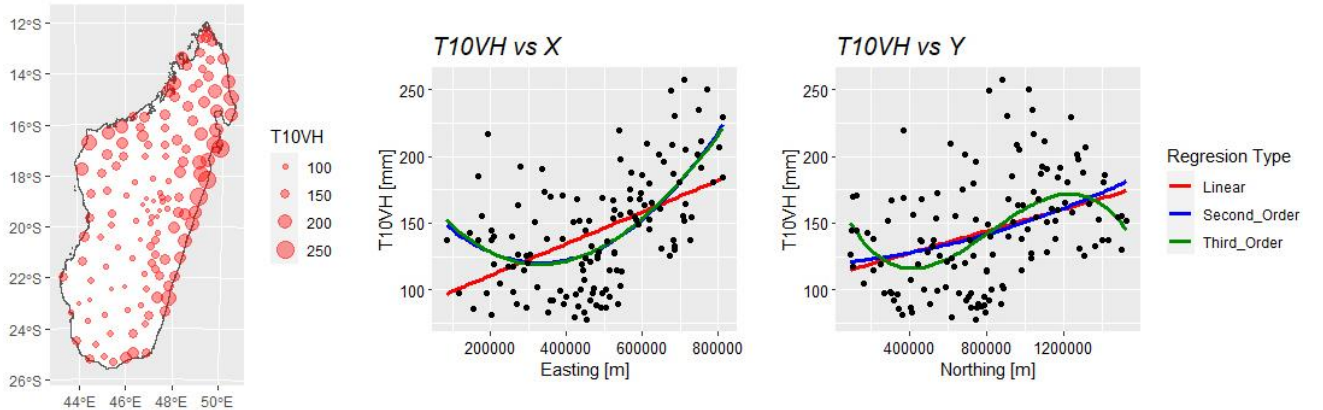
#### 4.4. Results of Data Normalisation

Given the results in Table 1, a Box-Cox transformation was used to normalise the data according to equations (1) and the  $\lambda$  values found are summarised in Table 4.

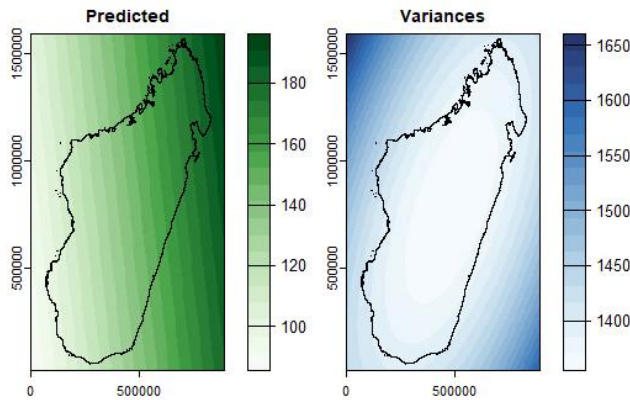
It is noting that many of these values of  $\lambda$  are fairly close to zero, which means that a logarithmic transformation could have been performed [35]. However, in the present study, these values of  $\lambda$  were kept unchanged and it was the first form of equations (1) and (2) that was used during the transformation and back-transformation.

**Table 3.** Parameters of 2nd order trend surfaces of raw data

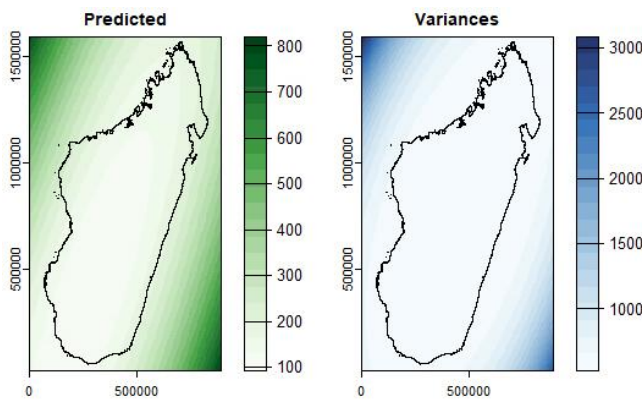
| Raw Data | $\beta_0$ | $\beta_1$ | $\beta_2$ | $\beta_3$ | $\beta_4$ | $\beta_5$ |
|----------|-----------|-----------|-----------|-----------|-----------|-----------|
| T10M     | 113.34    | -2.50E-04 | 1.99E-05  | 8.65E-10  | 1.52E-10  | -5.61E-10 |
| T10VH    | 128.65    | -4.06E-04 | 9.53E-05  | 1.34E-09  | 1.91E-10  | -8.58E-10 |
| T25M     | 152.12    | -3.37E-04 | 1.04E-06  | 1.14E-09  | 2.10E-10  | -7.32E-10 |
| T25VH    | 177.40    | -5.55E-04 | 8.84E-05  | 1.76E-09  | 2.64E-10  | -1.11E-09 |
| T50M     | 186.44    | -4.05E-04 | -2.33E-05 | 1.37E-09  | 2.65E-10  | -8.80E-10 |
| T50VH    | 220.89    | -6.72E-04 | 7.03E-05  | 2.12E-09  | 3.33E-10  | -1.33E-09 |
| T100M    | 226.49    | -4.77E-04 | -5.87E-05 | 1.62E-09  | 3.30E-10  | -1.05E-09 |
| T100VH   | 272.00    | -7.96E-04 | 3.84E-05  | 2.51E-09  | 4.16E-10  | -1.58E-09 |
| T200M    | 273.82    | -5.53E-04 | -1.08E-04 | 1.91E-09  | 4.10E-10  | -1.25E-09 |
| T200VH   | 332.89    | -9.29E-04 | -1.09E-05 | 2.96E-09  | 5.17E-10  | -1.86E-09 |

**Figure 3.** Left: Map of proportional values of T10VH. Right: Polynomial regressions of T10VH as a function of X and Y coordinates





**Figure 4.** Map of the trend surface predicted by the 1st order model and map of the variances of its residuals (T10VH)

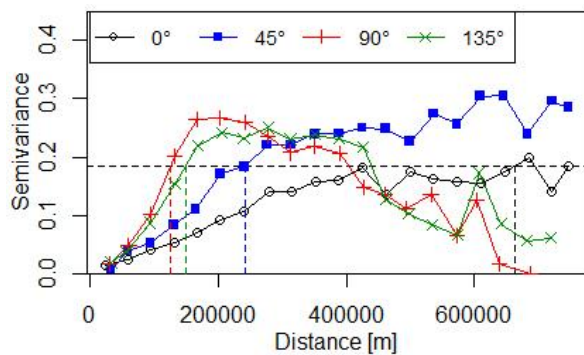


**Figure 5.** Map of the trend surface predicted by the 2nd order model and map of the variances of its residuals (T10VH)

**Table 4.** Values of parameter  $\lambda$  of the Box-Cox transformation

| Data  | $\lambda$    | Data   | $\lambda$  |
|-------|--------------|--------|------------|
| T10M  | -0.005610981 | T50VH  | 0.03300624 |
| T10VH | 0.07651904   | T100M  | -0.1095962 |
| T25M  | 0.1092762    | T100VH | -0.0724485 |
| T25VH | 0.101106     | T200M  | -0.2766751 |
| T50M  | 0.04398702   | T200VH | -0.1825193 |

#### 4.5. Anisotropy Study Results



**Figure 6.** Directional empirical semivariograms of BC data for T10VH. The dotted horizontal line corresponds to the sill (variance) of the BC data

As it is on the BC transformed data that the kriging interpolations will be applied, the examination of the anisotropy

was carried out on this BC transformed data. The directional empirical semivariograms for T10VH and for the four main directions in space, 0°, 45°, 90° and 135°, are shown in Figure 6.

Examination of the ranges of the directional empirical semivariograms in Figure 6 shows that the major axis of the anisotropy ellipse is along the 0° direction (south-north direction) while the minor axis is along the 45° direction (west-east direction). All the other data in the study show patterns similar to Figure 6, but with different values for sills and ranges. The anisotropy ratios (equation (12)) for all the data are summarised in Table 5.

**Table 5.** Anisotropy ratios

| BC data | $R$   | BC data | $R$   |
|---------|-------|---------|-------|
| T10M    | 2.976 | T50VH   | 4.258 |
| T10VH   | 5.298 | T100M   | 3.119 |
| T25M    | 3.867 | T100VH  | 4.160 |
| T25VH   | 5.373 | T200M   | 2.779 |
| T50M    | 3.236 | T200VH  | 3.543 |

Thus, in accordance with equation (13) and with  $\varphi = 0$ , the distance  $h'$  to be considered between any pair  $(x_1, y_1)$  and  $(x_2, y_2)$  in semivariogram calculations should be:

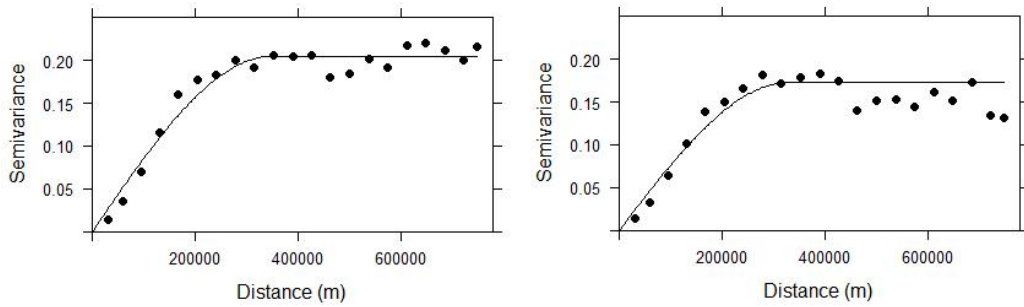
$$h' = R \frac{(x_1 - x_2)}{(y_1 - y_2)} \quad (21)$$

Nevertheless, the application of this distance did not give any significant differences with the omnidirectional semivariograms taken without consideration of the anisotropy. This can be explained by the fact that the anisotropy may be a consequence of the non-stationarity, even intrinsic, of the random function studied and the differences between the different directional semivariograms may be due to a lack of data in certain directions [26]. This lack of data is clearly visible in Figure 1, where 119 of the 132 points observed were centred on the centroids of the district territories, which are much closer together in the central part than in the northern and western parts of the island. As these stationarity problems are supposed to disappear with a Box-Cox transformation and the removal of the trend, anisotropy was not taken into account in the spatial interpolation.

#### 4.6. Results of Theoretical Semivariogram Models

After removing the 1st order trend, by comparing the sums of squared errors, the fits of the theoretical models to the omnidirectional semivariograms of the BC data all gave as the best model a spherical model with a zero nugget among the theoretical models tested (exponential, Gaussian and spherical). The characteristics of these models are shown in Table 6 for Ordinary Kriging and Table 7 for Universal Kriging.

Figure 7 shows, for example, the omnidirectional empirical semivariograms and the fitted theoretical models for BC T10VH data.



**Figure 7.** Empirical semivariograms and fitted theoretical model for BC T10VH data. Left: Ordinary Kriging. Right: Universal Kriging

**Table 6.** Characteristics of spherical theoretical models of semivariograms of BC data (nugget = 0) for Ordinary Kriging

| BC data | Partial sill | Range [m] | BC data | Partial sill | Range [m] |
|---------|--------------|-----------|---------|--------------|-----------|
| T10M    | 0.06718      | 396 071   | T50VH   | 0.19760      | 310 015   |
| T10VH   | 0.20468      | 348 107   | T100M   | 0.03915      | 316 154   |
| T25M    | 0.25171      | 368 716   | T100VH  | 0.07511      | 290 709   |
| T25VH   | 0.33504      | 328 100   | T200M   | 0.00783      | 295 650   |
| T50M    | 0.15823      | 340 144   | T200VH  | 0.02592      | 279 952   |

**Table 7.** Characteristics of spherical theoretical models of semivariograms of BC data (nugget = 0) for Universal Kriging

| BC data | Partial sill | Range [m] | BC data | Partial sill | Range [m] |
|---------|--------------|-----------|---------|--------------|-----------|
| T10M    | 0.05405      | 363 958   | T50VH   | 0.17630      | 304 293   |
| T10VH   | 0.17233      | 328 271   | T100M   | 0.03534      | 310 928   |
| T25M    | 0.21312      | 349 530   | T100VH  | 0.06823      | 287 467   |
| T25VH   | 0.29193      | 316 506   | T200M   | 0.00722      | 293 668   |
| T50M    | 0.13883      | 329 696   | T200VH  | 0.02386      | 277 809   |

These results are not entirely surprising, at least for the Gaussian model that has been excluded, since the Gaussian model exhibits parabolic behaviour near the origin, which means that it can be associated with infinitely differentiable second-order stationary functions. However, such regularity is practically non-existent in practical applications, and the Gaussian model is considered unrealistic under normal circumstances [26].

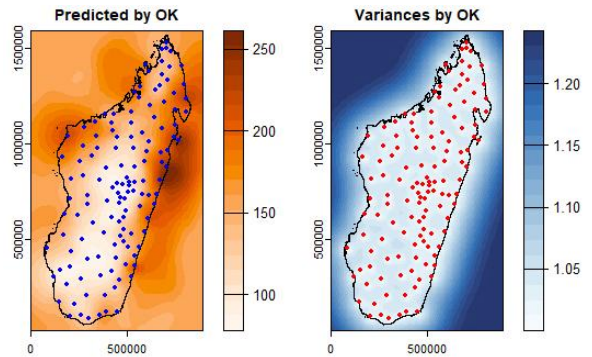
The exponential model, although also having a linear behaviour near the origin like the spherical model, was not retained because its sum of squared errors was higher than that of the spherical model. In addition, the zero value of the nugget reflects a better quality of the data, probably after it had been transformed according to Box-Cox.

#### 4.7. OK and UK Results

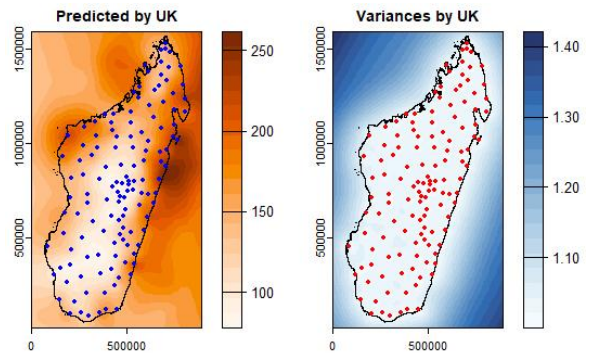
The theoretical semivariogram models found in Section 4.6 had been applied for Ordinary Kriging (OK) and Universal Kriging (UK) on the interpolation grid described in Section 2.3 (5000 m × 5000 m). For both Ordinary Kriging and Universal Kriging, this interpolation was carried out on the residuals of the BC data and, for Universal Kriging, the trend surface (affine) was directly taken into account in the model in the form  $T_m \sim X + Y$  and the interpolation was carried out on the residuals of this model.

For both Ordinary Kriging and Universal Kriging, it can be seen that the variances of the prediction residuals are minimal at the observed points, which clearly confirms the unbiased nature of spatial interpolation by kriging. Furthermore, the maximum values of these variances are found outside the island of Madagascar, which is normal since these are unsampled areas.

With the exception of the values, the mapping for the other data in the study shows similar patterns to Figures 8 and 9 and, to summarise, the maximum variances of these predictions are given in Table 8.



**Figure 8.** Ordinary Kriging for T10VH. Left: predicted values. Right: variances of predicted residuals



**Figure 9.** Universal Kriging for T10VH. Left: predicted values. Right: variances of predicted residuals

Table 8 shows three things: firstly, the maximum variances are lower for interpolation by OK than by UK. Secondly, the values are extremely accurate because, even for the highest value (UK.var = 1.782 mm<sup>2</sup>), the standard deviation is only 1.335 mm. Thirdly and finally, these



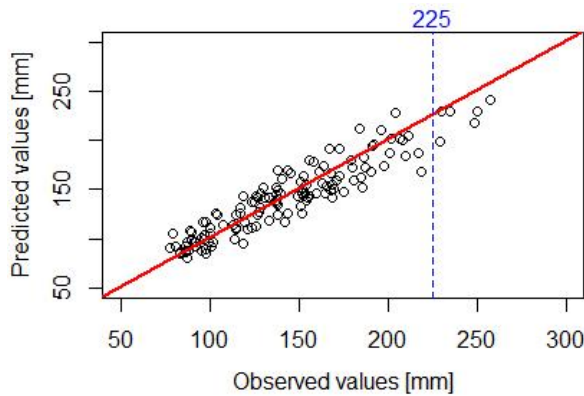
maximum values are outside the area of interest, i.e. in the unsampled part of Madagascar. Figures 8 and 9 show, for example, that within Madagascar, the variances for OK and UK lead to standard deviations of less than 1 mm.

**Table 8.** Maximum prediction variances (mm<sup>2</sup>) for OK and UK

| Data  | OK.var | UK.var | Data   | OK.var | UK.var |
|-------|--------|--------|--------|--------|--------|
| T10M  | 1.030  | 1.030  | T50VH  | 1.232  | 1.421  |
| T10VH | 1.243  | 1.421  | T100M  | 1.043  | 1.074  |
| T25M  | 1.305  | 1.552  | T100VH | 1.083  | 1.144  |
| T25VH | 1.420  | 1.782  | T200M  | 1.008  | 1.014  |
| T50M  | 1.184  | 1.331  | T200VH | 1.028  | 1.047  |

#### 4.8. Quality Assessment of the Two Kriging Models

Using the leave-one-out cross-validation (LOOCV) method, the two kriging models were compared: this method compares the predicted values with the observed values and the conclusions on the variances are similar to those obtained for Table 8 above, i.e. Ordinary Kriging presented the lowest variances at the observed points. In relation to the observed points, the two kriging models show exactly the same prediction behaviour, as shown for example in Figure 10 for the T10VH data.



**Figure 10.** Prediction behaviour of OK and UK models compared with observed values. Red line: model

Figure 10 shows that the model (red line) overestimates the predicted values above the limit, which is 225 mm for the T10VH data example. For all the data, we observed the same overestimation behaviour above a certain limit, but with different values for this limit. These limit values are given in Table 9 below.

**Table 9.** Limit values beyond which the kriging model overestimates the predicted values

| Data  | Limit [mm] | Data   | Limit [mm] |
|-------|------------|--------|------------|
| T10M  | 200        | T50VH  | 400        |
| T10VH | 225        | T100M  | 275        |
| T25M  | 215        | T100VH | 400        |
| T25VH | 265        | T200M  | 330        |
| T50M  | 250        | T200VH | 475        |

Nevertheless, these overestimation differences can be neglected (of the order of a millimetre), as we have seen with the variances of the residuals of each of these two models, which are outside the area of interest. The quality of the two models was also assessed using diagnostic metrics (equations (18), (19) and (20)), the results of which are shown in Table 10.

Table 10 shows that the ME values that measure the bias are all very close to zero, with the values for Ordinary Kriging being the lowest, and the maximum value being 0.0040 mm (T25VH for UK) which confirms once again that kriging is an unbiased interpolator. The MAE and RMSE values are also very close to zero. Therefore, in addition to the comments on prediction variances already made above, the best model for spatial interpolation is the Ordinary Kriging model and this is the one that was used for the final plotting of the frequency isohyets.

#### 4.9. Frequency Isohyet Plot

Based on the rasters produced by Ordinary Kriging, the frequency isohyet maps, i.e. the contour lines, were drawn using the free software QGIS. A simplified version of these maps for the T10VH data is shown in Figure 11.

**Table 10.** Comparison of diagnostic metrics for OK and UK models

| Metric       | OK     | UK     | Metric        | OK     | UK     | Metric       | OK     | UK     | Metric        | OK     | UK     |
|--------------|--------|--------|---------------|--------|--------|--------------|--------|--------|---------------|--------|--------|
| <b>T10M</b>  |        |        | <b>T10VH</b>  |        |        | <b>T25M</b>  |        |        | <b>T25VH</b>  |        |        |
| ME           | 0.0000 | 0.0010 | ME            | 0.0010 | 0.0030 | ME           | 0.0010 | 0.0020 | ME            | 0.0020 | 0.0040 |
| MAE          | 0.0630 | 0.0640 | MAE           | 0.1310 | 0.1320 | MAE          | 0.1500 | 0.1520 | MAE           | 0.1880 | 0.1920 |
| RMSE         | 0.0860 | 0.0860 | RMSE          | 0.1580 | 0.1610 | RMSE         | 0.1970 | 0.1990 | RMSE          | 0.2400 | 0.2450 |
| <b>T50M</b>  |        |        | <b>T50VH</b>  |        |        | <b>T100M</b> |        |        | <b>T100VH</b> |        |        |
| ME           | 0.0010 | 0.0020 | ME            | 0.0020 | 0.0030 | ME           | 0.0000 | 0.0010 | ME            | 0.0010 | 0.0020 |
| MAE          | 0.1360 | 0.1380 | MAE           | 0.1610 | 0.1650 | MAE          | 0.0760 | 0.0780 | MAE           | 0.1100 | 0.1120 |
| RMSE         | 0.1800 | 0.1830 | RMSE          | 0.2090 | 0.2140 | RMSE         | 0.1020 | 0.1040 | RMSE          | 0.1450 | 0.1480 |
| <b>T200M</b> |        |        | <b>T200VH</b> |        |        |              |        |        |               |        |        |
| ME           | 0.0000 | 0.0000 | ME            | 0.0010 | 0.0010 |              |        |        |               |        |        |
| MAE          | 0.0380 | 0.0380 | MAE           | 0.0700 | 0.0710 |              |        |        |               |        |        |
| RMSE         | 0.0510 | 0.0520 | RMSE          | 0.0930 | 0.0940 |              |        |        |               |        |        |

Figure 11 is shown just to give an idea of the final version of the isohyet maps produced in this work. Full maps with additional graticules and city markers to aid user location are available in PDF format and at larger scales for all data processed. Other useful information is also included on these definitive maps. These maps can be obtained by sending an e-mail to the author.



Figure 11. Overview of the T10VH frequency isohyet map

## 5. Conclusions

This work on extreme rainfall in Madagascar consists of two articles, the final objective of which was to update the frequency isohyet maps, the most recent of which dates back to 1993. The first article dealt mainly with frequency analysis, which led to a completely different view of the notion of return period, with non-stationary models to take account of the effects of climate change.

In this second article, the focus was on spatial interpolation, where a preliminary analysis of the data had shown that the traditional use of the Gumbel distribution to model extreme annual daily rainfall no longer made sense when the effects of climate change were taken into account. Analysis of the spatial structure had also shown the presence of spatial trends in the data, trends modelled by an affine surface or a quadratic surface depending on the coordinates. However,

the various tests showed that this trend was best expressed by an affine surface, which was also in line with the principle of parsimony in statistical modelling.

Spatial interpolation was then carried out on the Box-Cox transformed data, as the raw data showed non-negligible skewness coefficients, and two competing spatial interpolation models were considered: Ordinary Kriging and Universal Kriging. To carry out this geostatistical interpolation, after removing the trend, the directional and omnidirectional empirical semivariograms were analysed and it was shown that, despite the presence of geometric anisotropy, this had no significant influence on the interpolation results. Finally, the best theoretical models fitted to the omnidirectional semivariograms were all spherical models without nuggets, for both Ordinary Kriging and Universal Kriging.

After comparing the variances of the interpolation residuals, leave-one-out cross-validation (LOOCV) and diagnostic metrics, it was found that Ordinary Kriging was better than Universal Kriging and that the final frequency isohyet maps were drawn using the rasters produced by Ordinary Kriging. The cross-validation also showed that the interpolation models found overestimated the predicted values beyond a threshold value, although the deviations were extremely small and outside the area of interest, i.e. outside Madagascar.

In view of recent disasters in various fields (floods, bridge and road failures, etc.), we hope that this work will help to mitigate or even eliminate these disasters by taking into account the correct return levels and/or return periods to be applied in the design of infrastructures or in the management of areas at risk.

## REFERENCES

- [1] Ratsaramody J. (2024). Extreme Rainfall in Madagascar Part I: return levels and return periods with stationary and non-stationary models. *Resources and Environnement*, 2024, 14(1): 21-31.; <https://doi.org/10.5923/j.re.20241401.02>.
- [2] González-Álvarez A., Vilorio-Marimón O. M., Coronado-Hernández O. E., Vélez-Pereira A. M., Tesfagiorgis K., Coronado-Hernández J. R. (2019). Isohyetal Maps of Daily Maximum Rainfall for Different Return Periods for the Colombian Caribbean Region. *Water* 2019, 11, 358; <https://doi.org/10.3390/w11020358>.
- [3] Chaperon P., Danloux J., Ferry L. (1993). *Fleuves et Rivières de Madagascar*. Ed. ORSTOM, Paris (France), (1993) 883 p.
- [4] Stalenberg E., Hutchinson M.F., Foley W.J. (2018). Using historical normals to improve modern monthly climate normal surfaces for Madagascar. *International Journal of Climatology*. 2018; 38: 5746–5765. <https://doi.org/10.1002/joc.5776>.
- [5] Karandish F. & Shahnazari A. (2014). Appraisal of the geostatistical methods to estimate Mazandaran coastal ground water quality. *Caspian Journal of Environmental Sciences* 2014, Vol. 12 No.1 pp. 129-146.
- [6] Abo-Monassar A. & Al-Zahrani M.A. (2014). Estimation of

- rainfall distribution for the southwestern region of Saudi Arabia. *Hydrological Sciences Journal*, 59:2, 420-431. <https://doi.org/10.1080/02626667.2013.872788>.
- [7] Teng H., Shi Z., Ma Z., Li Y. (2014). Estimating spatially downscaled rainfall by regression kriging using TRMM precipitation and elevation in Zhejiang Province, southeast China. *International Journal of Remote Sensing*, 35:22, 7775-7794. <https://doi.org/10.1080/01431161.2014.976888>.
- [8] Xu W., Zou Y., Zhang G., Linderman M. (2015). A comparison among spatial interpolation techniques for daily rainfall data in Sichuan Province, China. *International Journal of Climatology*, 35: 2898-2907 (2015). <https://doi.org/10.1002/joc.4180>.
- [9] Adhikary S.K., Muttill N., Yilmaz A.G. (2017). Cokriging for enhanced spatial interpolation of rainfall in two Australian catchments. *Hydrological Processes*. 2017; 31: 2143-2161. <https://doi.org/10.1002/hyp.11163>.
- [10] Cecinati F., Wani O., Rico-Ramirez M.A. (2017). Comparing Approaches to Deal With Non-Gaussianity of Rainfall Data in Kriging-Based Radar-Gauge Rainfall Merging. *Water Resources Research*, 53, 8999-9018. <https://doi.org/10.1002/2016WR020330>.
- [11] Pellicone G, Caloiero T, Modica G, Guagliardi I. (2018). Application of several spatial interpolation techniques to monthly rainfall data in the Calabria region (southern Italy). *International Journal of Climatology*. 2018; 38: 3651-3666. <https://doi.org/10.1002/joc.5525>.
- [12] Das S. (2019). Extreme rainfall estimation at ungauged sites: Comparison between region-of-influence approach of regional analysis and spatial interpolation technique. *International Journal of Climatology*. 2019; 39: 407-423. <https://doi.org/10.1002/joc.5819>.
- [13] DeGaetano A.T., Mooers G., Favata T. (2020). Temporal Changes in the Areal Coverage of Daily Extreme Precipitation in the Northeastern United States Using High-Resolution Gridded Data. *Journal of Applied Meteorology and Climatology*. Vol. 59, pp. 551-565. <https://doi.org/10.1175/JAMC-D-19-0210.1>.
- [14] Liu, Y., Zhuo, L., Pregnotato, M., & Han, D. (2021). An assessment of statistical interpolation methods suited for gridded rainfall datasets. *International Journal of Climatology*, 42(5), 2754-2772. <https://doi.org/10.1002/joc.7389>.
- [15] Caloiero T., Pellicone G., Modica G., Guagliardi I. (2021). Comparative Analysis of Different Spatial Interpolation Methods Applied to Monthly Rainfall as Support for Landscape Management. *Applied Science*, 2021, 11, 9566. <https://doi.org/10.3390/app11209566>.
- [16] Ananias D.R.S., Liska G.R., Beijo L.A., Liska G.J.R., Silva de Menezes F. (2021). The assessment of annual rainfall field by applying different interpolation methods in the state of Rio Grande do Sul, Brazil. *SN Applied Sciences* (2021) 3: 687. <https://doi.org/10.1007/s42452-021-04679-1>.
- [17] Hounkpè J., Diekkrüger B., Badou D. F., Afouda A. A. (2015). Non-Stationary Flood Frequency Analysis in the Ouémé River Basin, Benin Republic. *Hydrology* 2015, 2, 210-229; <https://doi.org/10.3390/hydrology2040210>.
- [18] Cheng L., AghaKouchak A., Gilleland E., Katz R.W. (2014). Non-stationary extreme value analysis in a changing climate. *Climatic Change*, 2014, 127, 353-369. <https://dx.doi.org/10.1007/s10584-014-1254-5>.
- [19] Pebesma E.J. (2004). Multivariable Geostatistics in S: The GSTAT Package. *Computers and Geosciences*, 30, 683-691. <https://doi.org/10.1016/j.cageo.2004.03.012>.
- [20] Box, G. E. P., Cox, D. R. (1964). An analysis of transformations (with discussion). *Journal of the Royal Statistical Society – Statistical Methodology* 26:211-252.
- [21] Varouchakis, E.A. (2021). Gaussian Transformation Methods for Spatial Data. *Geosciences* 2021, 11, 196. <https://doi.org/10.3390/geosciences11050196>.
- [22] Fung K. F., Chew K. S., Huang Y. F., Ahmed A. N., Teo F. Y., Ng J. L., Elshafie A. (2022). Evaluation of spatial interpolation methods and spatiotemporal modeling of rainfall distribution in Peninsular Malaysia. *Ain Shams Engineering Journal* 13 (2022) 101571. <https://doi.org/10.1016/j.asej.2021.09.001>.
- [23] Kumari M., Singh C.K., Bakimchandra O., Basistha A. (2017). Geographically weighted regression based quantification of rainfall-topography relationship and rainfall gradient in Central Himalayas. *International Journal of Climatology*, 2017; 37(3): 1299-309.
- [24] Bivand R. S., Pebesma E., Gómez-Rubio V. (2013). *Applied Spatial Data Analysis with R*. Second Edition. Springer, New York 2013, <https://doi.org/10.1007/978-1-4614-7618-4>.
- [25] Webster R., Oliver M.A. (2007). *Geostatistics for Environmental Scientists. Second Edition*. John Wiley & Sons Ltd. ISBN-13: 978-0-470-02858-2 (HB). 332 pages.
- [26] Montero J.M., Fernández-Avilés G., Mateu J. (2015). *Spatial and Spatio-Temporal Geostatistical Modeling and Kriging*. John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781118762387>.
- [27] Cheng K. S., Yeh H. C., Tsai C. H. (2000). An anisotropic spatial modeling approach for remote sensing image rectification. *Remote Sensing of Environment*, 73, 46-54. [https://doi.org/10.1016/S0034-4257\(00\)00079-1](https://doi.org/10.1016/S0034-4257(00)00079-1).
- [28] Wagner P.D., Fiener P., Wilken F., Kumar S., Schneider K. (2012). Comparison and evaluation of spatial interpolation schemes for daily rainfall in data scarce regions. *Journal of Hydrology*, 464-465, 388-400. <https://doi.org/10.1016/j.jhydrol.2012.07.026>.
- [29] Gaetan C., Guyon X. (2010). *Spatial Statistics and Modeling*. Springer, ISBN 978-0-387-92256-0, (308 pages). <https://doi.org/10.1007/978-0-387-92257-7>.
- [30] Adhikary S. K., Yilmaz A. G., Muttill N. (2015). Optimal design of rain gauge network in the Middle Yarra River catchment, Australia. *Hydrological Processes*. 29, 2582-2599 (2015). <https://doi.org/10.1002/hyp.10389>.
- [31] Haddad K, Rahman A, Zaman MA, Shrestha S. (2013). Applicability of Monte Carlo cross validation technique for model development and validation using generalised least squares regression. *Journal of Hydrology* 482: 119-128. <https://doi.org/10.1016/j.jhydrol.2012.12.041>.
- [32] Szolgay J., Parajka J., Kohnová S., Hlavcová K. (2009). Comparison of mapping approaches of design annual maximum daily precipitation. *Atmospheric Research*, 92(3), 289-307. <https://doi.org/10.1016/j.atmosres.2009.01.009>.
- [33] Sanchez-Moreno J. F., Mannaerts C. M., Jetten V. (2014). Influence of topography on rainfall variability in Santiago Island, Cape Verde. *International Journal of Climatology*, 34:

1081 – 1097. <https://doi.org/10.1002/joc.3747>.

- [34] Agarwal S., Mukherjee D., Debbarma N. (2023). Analysis of extreme annual rainfall in North-Eastern India using machine learning techniques. *AQUA - Water Infrastructure, Ecosystems and Society*. Vol 72 No 12, 2201. <https://doi.org/10.2166/aqua.2023.016>.

- [35] Baker H.A., AL-Jawad S.N., Abdulla A.A. (2016). Applied Spatial Data Analysis Technique on Petrophysical Properties of MA Unit of Mishrif Formation / Noor Field. *Iraqi Journal of Chemical and Petroleum Engineering*, Vol.17 No.3 (September 2016) 75- 81, ISSN: 1997-4884.

Copyright © 2024 The Author(s). Published by Scientific & Academic Publishing

This work is licensed under the Creative Commons Attribution International License (CC BY). <http://creativecommons.org/licenses/by/4.0/>