

Calculating and interpolating the T-100 year rainfall event in Austria

Final Seminary Work
Statistics of Extreme Events and Geostatistics

Gauthier Heroufousse, Valentino Mascherini

2023-01-15

Contents

1	Introduction	2
2	Data set description	2
2.1	HISTALP	2
2.2	E-OBS	3
3	Methodological approach	3
3.1	Extreme value selection	3
3.2	Parameter estimation	4
3.3	Assessment methods	5
3.4	Geostatistical analysis	6
3.5	Software and packages	7
4	Results	7
4.1	Method comparison	7
4.2	Geostatistical analysis	8
5	Discussion	11
	References	12

1 Introduction

Extreme value theory is surely one of the most influential statistical branches of the last decades and the upcoming future. Its applications range from finance markets to earthquake, and all have in common the characteristic of dealing with scarce, extreme data. The very nature of such values implies a magnification in the scale of the processes, that can only be achieved by extrapolation, using asymptotic arguments. In the field of meteoric phenomena, it is even a bigger challenge to predict accurately the return levels of high magnitude, due to climate change. Therefore, the more robust methods and approaches to tackle said issues have to be identified and tried out, in order to improve the chances at the scientist’s disposal.

This study strives to add another piece to the extreme value theory puzzle, applied to climate science. The variable of interest is precipitation, arguably one of the most important characteristics of the weather, as it affects agriculture by direct rainfall and through balancing water reservoirs, and it constitutes a possible hazard as well to people, settlements and infrastructures. After all, extreme value theory had been originally applied to civil engineering, to build emergency defiant structures.

In this research, it has been compared which method of extreme value selection is more suitable for precipitation time series. Within that, a benchmarking of parameter estimation techniques, such as maximum likelihood estimation and L-moments, took place. Moreover, two different, novel threshold selection approaches have also been employed with the scope of gaining more understanding in regard of their applicability and individual performance. The assessment of the aforementioned methods has been carried out through algebraic and visual measures, both at model level and at method level. Given the data of various rainfall measuring stations, the most appropriate model has been detected, and a count of the success over the others has been recorded to select the best one through the entire region. Through additional data, recording the years following the first dataset, an additional comparison has been performed in order to quantify the predictions previously obtained. Furthermore, through geospatial interpolation techniques, the predictions have been extended to the whole territory and visually compared to the second dataset.

Thus, the scope of this study is to discuss which methods better suit extreme rainfall event modeling through model selection and validation techniques.

Please find all code and additional figures and tables in the Appendix.

2 Data set description

2.1 HISTALP

The Historical Instrumental Climatological Surface Time Series Of The Greater Alpine Region (HISTALP) data set comprises over 200 stations spread over the so called Greater Alpine Region, ranging from 4 to 19°E, and from 43 to 49°N. The territory covered thus comprises many different countries, but in this research only the Austrian region has been selected, the reason being that it is the only one with access to daily data. The time series, for most stations, start from 1950 and end in 2009, and mainly record precipitation and temperature. It has to be noted how the data are homogenised through a QC procedure, and many non-climatic breaks and outliers have been removed (Auer et al. 2007). After the selection of the stations befitting the purpose of this study, only 54 remained to be analysed, a few of which presenting data gaps up to 1 - 2 year of length. The map of the area of interest is shown in Figure 1. It was expected that the estimation of the T-100 event level would be affected by such gaps, but the biggest problem lies in the limited number of stations to perform the kriging procedure on. Table 1 shows the main descriptive statistics for the data set, and in the Appendix another table with the single station statistics can be found.

Table 1: Descriptive statistics of the entire HISTALP dataset.

mean	max	sd	count
2.9	233.2	1.899	54

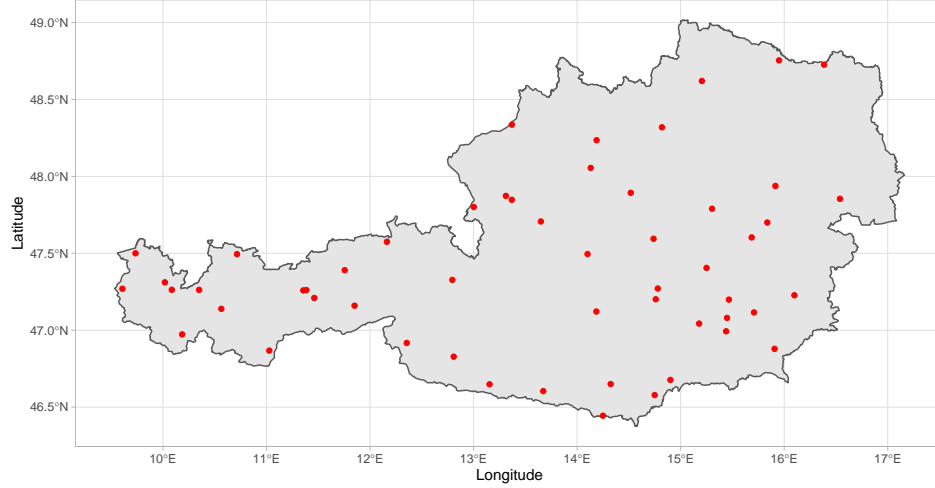


Figure 1: Map of the 54 HISTALP stations in Austria selected for the research. There are some stations that are quite close to each other, while leaving some areas bare, like in the region of Vienna.

2.2 E-OBS

The E-OBS is an ensemble data set that stems from the stations of the ECA&D initiative and contains a number of climatic variables like precipitation, temperature, sea level pressure for Europe (Cornes et al. 2018). It has to be taken into account that the data presented are in fact spatially interpolated in form of a grid. The dataset is sparse and the possible presence of outliers can produce unrealistic values where just few stations are near, which is usually at the borders of the dataset. Since the region the study is focused in is Austria, it is to be expected to have a robust estimation of such data, even considering that precipitation is more sensible to that issue than other variables (Morice et al. 2012). A plot of such dataset is shown in figure 5a.

3 Methodological approach

Hydrology is supported by statistical tools in a variety of ways, and that is also true for the field of extreme events. Extreme value theory is underlying to calculations such as the T-100 year events, and is based upon probability distribution models (Bras and Rodriguez-Iturbe 1993). Those are fit to the subset of data that is obtained after an extreme value selection. Here are discussed the methods employed in this study. The implementation of those methods is automated by the script that can be found in the Appendix.

3.1 Extreme value selection

3.1.1 Annual maxima

This method is the most common of the block maxima approach, which selects the maximum value among n observations within a time period, which is described by a so called “block”. This means that there is no need for declustering the data, but at the same time a significant portion of them is lost. The generalised extreme value distribution (GEV) based on Fisher and Tippet (1928) is considered to be a good fit for it, and its cumulative distribution function (CDF) is thus described:

$$e^{-t(x)} \quad (1)$$

where

$$t(x) = \begin{cases} (1 + \xi(\frac{x-\mu}{\sigma}))^{-1/\xi} & \xi \neq 0 \\ e^{-(x-\mu)/\sigma} & \xi = 0 \end{cases} \quad (2)$$

with μ as the location parameter, $\sigma > 0$ as the scale parameter and ξ as the shape parameter.

Based on ξ we differentiate the GEV into three families: Gumbel, Weibull, Frechet. In our case, the Gumbel distribution is the more appropriate, since it is not limited by an upper boundary. At any rate, there is no need of subjectively choose between those three options a priori, as through inference on ξ , the data themselves determine the most appropriate type of tail behavior (Coles et al. 2001).

3.1.2 Peak over threshold

Contrary to the annual maxima, the method of POT cannot be implemented with ease. The question of threshold selection is one that is unsolved to this day, and there are many different approaches to tackle it (Caeiro and Gomes 2015), without a definitive best method. Mean excess and parameter stability plots have been deemed scarcely interpretable (Wadsworth 2016) and more objective methods have been developed. Those are also able to be automatised without the necessity of visualising plots.

The subset of data obtained by the threshold selection is then fit with a general Pareto distribution (GPD), whose CDF is so defined:

$$F(z) = \begin{cases} 1 - (1 + \xi z)^{-1/\xi} & \xi \neq 0 \\ 1 - \exp(-z) & \xi = 0 \end{cases} \quad (3)$$

where

$$z = \frac{x - \mu}{\sigma} \quad (4)$$

with μ as the location parameter, $\sigma > 0$ as the scale parameter and ξ as the shape parameter.

As in the case of the GEV, the GPD is most influenced by the shape parameter. The scale parameter is the one that is determined by the threshold selection (Coles et al. 2001). The advantage of using POT is most notably the increase of available observations to fit the model with, with the trade-off of increasing subjectivity bias due to the threshold selection. In the case of this study, even if the threshold is automatically selected, upper and lower limits for it have been decided upon. More in the Appendix section.

3.2 Parameter estimation

3.2.1 Maximum Likelihood Estimation

A few different methods exist to estimate the parameters of a distribution. One of the most common in the Maximum Likelihood method. If we take a GEV distribution as an example; the basic idea is to find the parameters $\hat{\theta}$ from the GEV (where $\theta = \{\mu = \text{location}; \sigma = \text{scale}; \alpha = \text{shape}\}$) that maximize the probability of the observed data given the parameters. This probability is known as the likelihood. We first need to define the likelihood function that describes the probabilities of the observed data given the parameters values :

$$L(\theta) = \prod f(x_i; \theta)$$

Then, an algorithm is used to maximise the Likelihood function, with a given $\hat{\theta}$ that is the best estimator of the θ set of parameters.

3.2.2 L-moments

The L-moment method is another method to estimate the parameters of a distribution. It is based on the first four linear moments λ_1 to λ_4 that are defined as:

$$\lambda_1 = \mathbb{E}(X_{(1:1)})$$

$$\lambda_2 = \frac{1}{2}\mathbb{E}(X_{(2:2)} - X_{(1:2)})$$

$$\lambda_3 = \frac{1}{3}\mathbb{E}(X_{(3:3)} - X_{(2:3)} + X_{(1:3)})$$

$$\lambda_4 = \frac{1}{4}\mathbb{E}(X_{(4:4)} - 3X_{(3:4)} + 3X_{(2:4)} - X_{(1:4)})$$

Once the sample L-moments are computed, the parameters of the distribution (GEV, GPD) are estimated by solving a set of equations that are derived from the relation between the L-moments and those parameters (J. Hosking and Wallis 1993; Huard, Mailhot, and Duchesne 2010).

It is also common to standardize moments of higher order to make them independent from the lower moment:

$$\text{L-coefficient of variation : } \tau_1 = \lambda_2/\lambda_1$$

$$\text{L-skewness : } \tau_3 = \lambda_3/\lambda_2$$

$$\text{L-kurtosis : } \tau_4 = \lambda_4/\lambda_3$$

3.3 Assessment methods

Model selection is another somewhat problematic step in the process. Visual interpretation is unreliable and not applicable in case of large datasets, thus two indicators of model assessment have been implemented following the paper of Schlögl and Laaha (2017). Those are the CRMSE and the CMAE, conditional variants of the root mean square error and the mean absolute error (Weibull, 1939; Makkonen, 2006), and are so described:

$$CRMSE_{T^*} = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n_{T^*}}} \quad \forall \quad y_i : \left[-\frac{1}{\ln(\frac{m}{N+1})} \right] \geq T^* \quad (5)$$

$$CMAE_{T^*} = \frac{\sum_{i=1}^n |(\hat{y}_i - y_i)|}{n_{T^*}} \quad \forall \quad y_i : \left[-\frac{1}{\ln(\frac{m}{N+1})} \right] \geq T^* \quad (6)$$

where \hat{y}_i is the model prediction and y_i is the observed value. N and m are statistical rankings, m being the minimum and N the maximum. n_{T^*} is the number of the elements that are above the threshold T^* .

The main issue with those indicators is that they both carry some subjective bias, as a return period (T^*) has to be selected in order to calculate them. A $T^* = 5$ has been selected since the dataset cover about 50 years. Thus, the observations to be employed have been reduced to a dozen, in most cases. In this research the $CRMSE_5$ and $CMAE_5$ are employed to compare the HISTALP observations with the model fits for every method.

Another model assessment approach has been implemented, that considers all the available stations. Comparing the T-13 return level to the data from the E-OBS data set, it is possible to assess how the models have performed in predicting the events that have been occurring from 2010 to the end 2022, that is the period that isn't covered by the HISTALP dataset. To do so, firstly it is each E-OBS tile that covers each HISTALP station is located, obtaining its maximum value, then compared to the T-13 return level.

3.4 Geostatistical analysis

3.4.1 Boxcox transformation

As the resulting T-100 was not normal, a boxcox transformation was used to make the data normal, in order to realize a kriging afterwards. The technique involves a power transformation of the form :

$$y = \frac{(x^\lambda - 1)}{\lambda} \text{ for } \lambda \neq 0 \quad (7)$$

and

$$y = \log x \text{ for } \lambda = 0 \quad (8)$$

where x is the original data and y is the transformed data. The value of λ is chosen such that the transformed data is as close to normal as possible. \

Moreover, as the *eyefit* function used to fit the variogram doesn't let the user choose all the parameters when using a boxcox transformation with a $\lambda < 0$, a comparison between the results with the optimal value of λ and with $\lambda = 0$ will be done.

3.4.2 Kriging

The chosen method to realize the geographical interpolation is the ordinary kriging. The kriging method is an interpolation method that is based on the spatial correlation between measurements. The semivariogram is first computed to express the dissimilarity of the two points in regards to the distance h for the whole dataset :

$$\gamma = \frac{(z(x+h) - z(h))^2}{2}$$

Based on that , you estimate the best function $\gamma(h)$ that can represent the covariance of the points depending on the distance h . The variogram model used in our simulation is defined as follow :

$$\gamma(h) = b_0 * nug(h) + b_1 * exp(h, a_1)$$

The weights are then calculated to minimise the variance estimation. We end up with the Best Linear Unbiased Estimator (BLUE). The prediction is also exact as $Z^*(x_o) = Z(x_\alpha)$ if $x_o = x_\alpha$.

The kriging requires a 2nd-order stationary in order to be performed. That means that the mean is constant on the study field and the covariance depends only on the distance.

3.4.3 Validation

In order to validate the results, a comparison with real precipitation data will be done using data from the E-OBS dataset. From this dataset, only the years from 2010 to 2022 have been selected and the maximum values have been extracted. Thus, the subset of data can be used to perform model validation for the extreme value series models, by comparing it to the T-13 year event.

3.5 Software and packages

This research has been carried out using *R* version 4.0 via *RStudio* (R Core Team 2020) and specific packages to perform statistics. The *extRemes* package (Gilleland and Katz 2016) is used to perform extreme value analysis, the model fitting is done by the *fevd* function, which supports the many EV distributions, with the GEV and GPD used in this study as well. Thresholds and preferred parameter estimation methods can be selected when fitting the model. In the Appendix the fitting of the models is shown. To estimate L-moments the *lmom* package (J. R. M. Hosking 2022) and the *lmomco* (Asquith 2022) were employed, in the first POT threshold selection method.

To perform the geostatistical analysis, the package *geoR* was used. The fitting of the variogram was done using the *eyefit* function, used as its name suggest, to fit an empirical variogram “by eye”. Then, the kriging was realised using the *krige.conv* function. A few other geo-related packages were used to preprocess the data, such as *MASS*, *Raster* or *sf*.

4 Results

4.1 Method comparison

Due to the CRMSE₅ and CMAE₅, findings can presented in regard to which method is the best according to those two indicators. The Table 2 shows that the best overall method seems to be the GPD/POT-WAD/MLE with the threshold selection performed by the script developed by Wadsworth (2016), with 23% of success which is slightly more than the 20% of the GEV/AM/lmom and GPD/POT/MLE. For CRMSE₅, GPD/POT/MLE with threshold selection via L-moments (Silva Lomba and Fraga Alves 2020) is the best by far with 26% of success rate, while for CMAE₅ the second approach dominates with 31% in the GPD/POT-WAD/MLE. Figure 2 is an example of how the different models compare to each other. For simplicity, the methods will be called from now on AM, POT and POT-WAD respectively.

As for the T-13 return level comparison, every model has a mean of about 17-18% of difference with the data from E-OBS. Table 5 shows the sum of absolute errors for each model. Moreover, 28% of the times the actual rainfall exceeded the simulated one, in 11% by 10 mm or more with a maximum of 27 mm or 26% of the observed value. The result confirm what Schlögl and Laaha (2017) reported in regard to the fact that POT techniques tend to overestimate the return levels when dealing with non extreme events, with low return period.

Table 2: Comparison between the methods used to obtain the T-100 level. The first table reports the success percentage in regard to the CRMSE, the second one to the CMAE. POT indicates the GPD/POT method with threshold selection using L-moments, POT-WAD using MLE, according to the methods included in the Appendix section 3.

	MLE	Lmom	Sum		MLE	Lmom	Sum
AM	15	17	31	AM	4	22	26
POT	26	15	41	POT	15	6	20
POT_WAD	15	13	28	POT_WAD	31	22	54
Sum	56	44	100	Sum	50	50	100

Table 3: Table showing the total success percentage among the CRMSE and CMAE.

	MLE	Lmom	Sum
AM	9	20	29
POT	20	10	30
POT_WAD	23	18	41
Sum	52	48	100

Table 4: Summary statistics of the thresholds obtained through the two different methods described in section 3 of the Appendix. The ones obtained with MLE are lower overall, and do not present unreasonable maxima. It is not to be confused the step of threshold selection (MLE or Lmom) with the one of model parameter estimation, which uses the same mathematical approaches.

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
MLE selection	3.6	7.925	10.35	12.07404	13.675	34.97835
Lmom selection	3.4	7.375	12.45	14.87155	18.375	59.20057

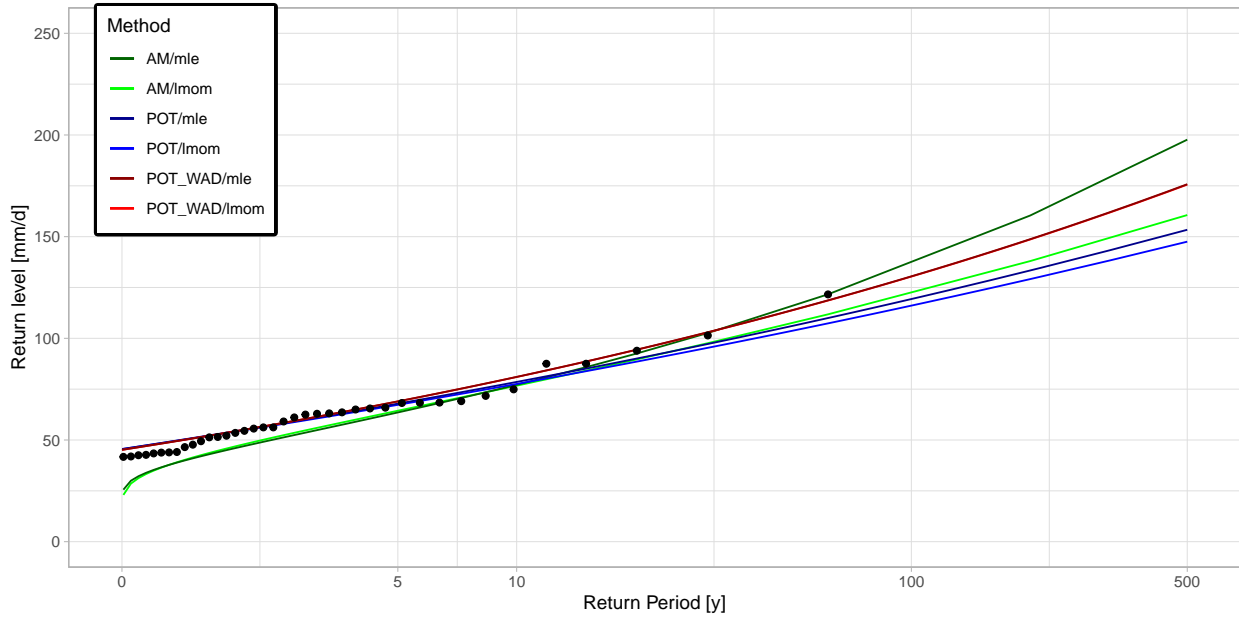


Figure 2: Plot of the station 46 Reichenau Rax showing a comparison of the methods employed. In this case both the CRMSE and CMAE are lowest with the POT-WAD method, and its MLE and Lmom models yield the same result, overlapping.

Table 5: Comparison between the methods used to obtain the T-13 level with the eOBS data set maximum values from year 2010 to 2022. The indicator used is the sum of absolute errors, which doesn't show a great deal of difference but still validates what has been proposed in previous papers.

	MLE	Lmom
AM	818	815
POT	841	835
POT_WAD	845	840

4.2 Geostatistical analysis

The results suggest that the POT with the Maximum Likelihood method perform globally better than all the other combinations. Thus, all the geostatistical analysis was done using the return period resulting from this statistical approach.

Table 6: Type and parameters of fitted variogram.

Model	Sill	Phi	Nugget	PracticalRange
exponential	0.062	0.61	0.0077	1.83

4.2.1 Statistical distribution and transformations

It is clear that the initial T-100 rain event data set is particularly not normal. Some stations present extreme T-100 rain events that drive the distribution on the right. After a logarithmic transformation, the phenomenon is attenuated, but still present. The optimal boxcox transformation result in a much more centered data set, with almost all the values in the qqplot confidence interval. Regarding the geographical distribution, the optimal fitting of the variogram using $\lambda = 0$ resulted in the parameters seen in 6.

4.2.2 Kriging results

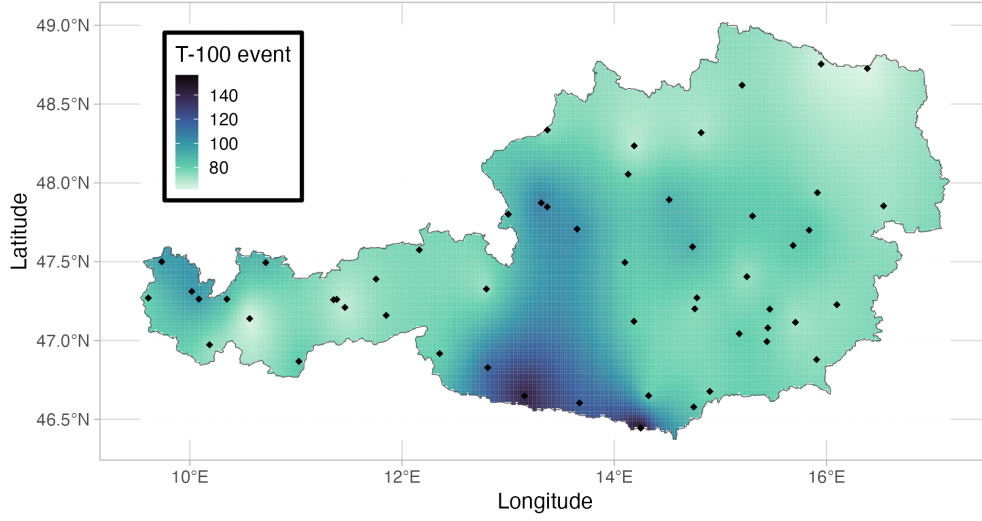


Figure 3: T-100 rain event kriging map

The kriging result gives a relatively non-homogeneous distribution over the country. It is observable that all the East part of Austria is spared from rain in compare to the central part. There is also a hole in the middle of the Tirol region. The second result with the optimal lambda is not displayed here as there is not noticeable difference between the two plots.

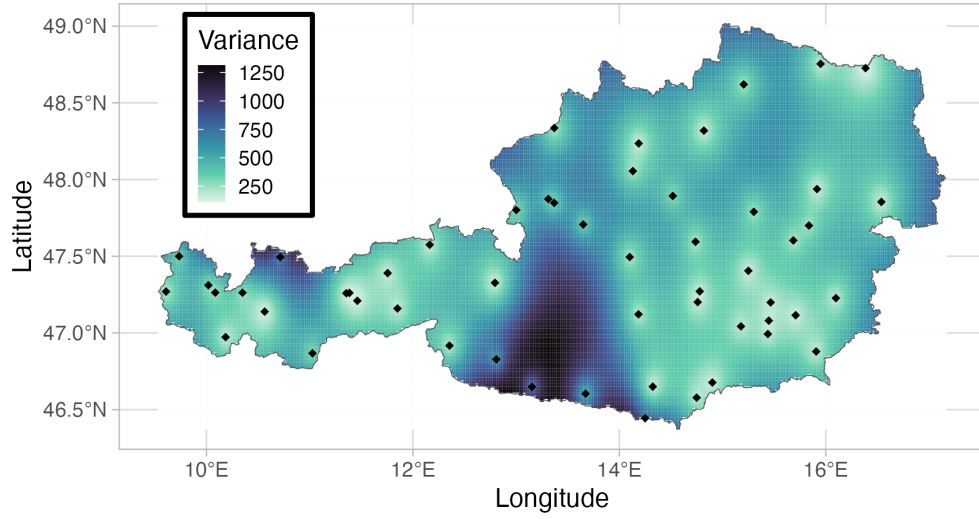


Figure 4: Variance of the T-100 rain event kriging map

You can find here above the variance corresponding to the prediction.

4.2.3 Comparison with real rain data

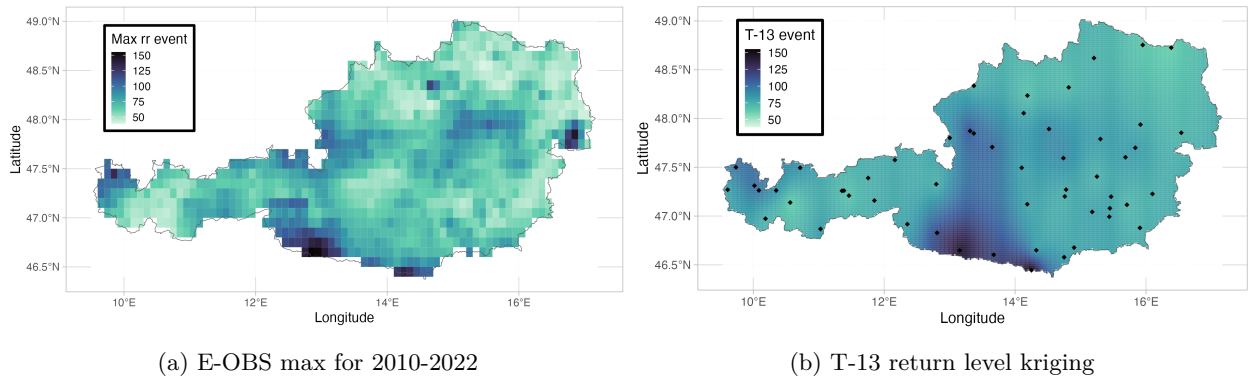


Figure 5: Comparison of T-13 year event. Map A on shows the results of kriging the T-13 return level, while map B represents the maximum value that the E-OBS ensemble database has registered for the years 2010-2022.

The prediction behaves not quite the same as the true data is. That is not that surprising, as the prediction is only a theoretical statement. However, the higher data values are still relatively close to the predicted ones. The biggest difference is seen in the middle of the austrian territory. We clearly see there that the rain values are overestimated.

5 Discussion

This research can confirm the results of Ben-Zvi (2009) and Bezak, Brilly, and Šraj (2014) stating that the GPD/POT is a better method than AM when modeling extreme rainfall values. On the other hand, in regard to which technique of threshold selection is best, the results, which are reported in Table 2 and 3 and are not conclusive. It is interesting to see how POT-WAD, the worse performing method in CRMSE₅, is able to result the overall best due to the CMAE₅ success rate. Taking into consideration the nature of the RMSE and the MAE is the way to get a deeper understanding of the results. In general, MAE doesn't react to variation in noise as much as the RMSE. That means that if the error is distributed or concentrated in a single data point, it doesn't matter as much. RMSE registers more sensibly when a single observation is very skewed and penalizes it more than when the same error spread among the whole sample.

The POT-WAD has been seen selecting lower thresholds (Table 4), with the result of selecting more data points than the POT. That could make it so that it has an overall lower error (CMAE), since it has to accommodate more data points. On the other hand, those few, noise-imbued points carry the most of the unexplained variance, so much that the POT-WAD is the worse in the CRMSE₅ success rate. From this reasoning we can infer that the POT-WAD gives less space to the more noise-imbued points than the other two models, while the POT and the AM get skewed at the expense of the fit of the regular points. That is indeed the effect of employing a larger sample size in modeling. In the graph of Figure 1 one of the stations with lowest CRMSE₅ and CMAE₅ for the POT-WAD method can be seen. The GEV/AM is shown to be really weak in the body of this distribution, but to recover quite well in case of the AM/MLE.

Going back to threshold selection methods, in table 4 the difference between the thresholds do not appear too marked. However, it has to be taken into account how even a single mm of rainfall more can exclude hundreds of observations. The two automatic threshold selecting methods seem to converge to the same results, at least regarding the lower values, showing the robustness of the methods. Still, the L-moments based selection has some clear issues, that make the best threshold approach the upper limit, as in the case of the Schockl station where the threshold value is as high as 59 mm.

The kriging prediction resulting from the geostatistical analysis gives a good view of how could be represented the extreme rain event on the country. However, it is important to keep in mind that this prediction is heavily subject to random effects that are unpredictable. Looking at the variance resulting from the kriging, we see that the whole region of interest is under a high variance. Even when the variance is minimal (250 mm²), the resulting deviation could be up to 15 mm of rain difference. Unfortunately, the variance couldn't be achieved using the optimal lambda from the boxcox transformation using the *geoR* package. It seems to have some troubles dealing with lambda values < 0. A good alternative could be to use the *gstat* package.

Regarding the comparison with the E-OBS dataset, for the POT-WAD model in the majority of the cases there has been an overestimation of the T-13 return level event (72%) with the maximum error being 51 mm or 37%. The POT and AM have performed similarly with the exception of an outlier which puts the maximum percentage error of AM at 82%. It has to be noted how that station is the one that lacks the most data points, missing entire years. That surely affects the AM more negatively than the POT methods. Nonetheless, AM has been proven to be the most accurate as seen in Table 5. Another comparison, this time a visual one, is made between the E-OBS ensemble dataset and the results of the kriging procedure using the T-13 return level values obtained through the POT-WAD method. The plots concerned are shown in figure 5 and show the differences between the two. It seems that the area that has the most variance as seen in 4 is also the one that differs the most between the two maps (Figure 5). Not coincidentally, it is an area lacking measurement stations. Notably, at coordinates 48° N, 17° E, there is a dark blue spot in the E-OBS map that is lacking in the kriging. That also has to do with the lack of measurement stations nearby and being at the edge of the domain. Anyway, here is a reminder that the E-OBS dataset itself is not comprised of observed data, but rather of interpolated ones, that are thus subjected to errors of various nature and should not be considered substitutes to genuinely observed rainfall.

References

- Asquith, W. H. 2022. *Lmomco—l-Moments, Censored l-Moments, Trimmed l-Moments, l-Comoments, and Many Distributions*.
- Auer, Ingeborg, Reinhard Böhm, Anita Jurkovic, Wolfgang Lipa, Alexander Orlik, Roland Potzmann, Wolfgang Schöner, et al. 2007. “HISTALP—Historical Instrumental Climatological Surface Time Series of the Greater Alpine Region.” *International Journal of Climatology: A Journal of the Royal Meteorological Society* 27 (1): 17–46.
- Ben-Zvi, Arie. 2009. “Rainfall Intensity–Duration–Frequency Relationships Derived from Large Partial Duration Series.” *Journal of Hydrology* 367 (1-2): 104–14.
- Bezák, Nejc, Mitja Brilly, and Mojca Šraj. 2014. “Comparison Between the Peaks-over-Threshold Method and the Annual Maximum Method for Flood Frequency Analysis.” *Hydrological Sciences Journal* 59 (5): 959–77.
- Bras, Rafael L, and Ignacio Rodriguez-Iturbe. 1993. *Random Functions and Hydrology*. Courier Corporation.
- Caeiro, Frederico, and M Ivette Gomes. 2015. “Threshold Selection in Extreme Value Analysis.” *Extreme Value Modeling and Risk Analysis: Methods and Applications*, 69–82.
- Coles, Stuart, Joanna Bawa, Lesley Trenner, and Pat Dorazio. 2001. *An Introduction to Statistical Modeling of Extreme Values*. Vol. 208. Springer.
- Cornes, Richard C, Gerard van der Schrier, Else JM van den Besselaar, and Philip D Jones. 2018. “An Ensemble Version of the e-OBS Temperature and Precipitation Data Sets.” *Journal of Geophysical Research: Atmospheres* 123 (17): 9391–9409.
- Fisher, Ronald Aylmer, and Leonard Henry Caleb Tippett. 1928. “Limiting Forms of the Frequency Distribution of the Largest or Smallest Member of a Sample.” In *Mathematical Proceedings of the Cambridge Philosophical Society*, 24:180–90. 2. Cambridge University Press.
- Gilleland, Eric, and Richard W. Katz. 2016. “extRemes 2.0: An Extreme Value Analysis Package in R.” *Journal of Statistical Software* 72 (8): 1–39. <https://doi.org/10.18637/jss.v072.i08>.
- Hosking, J. R. M. 2022. *L-Moments*. <https://CRAN.R-project.org/package=lmom>.
- Hosking, JRM, and JR Wallis. 1993. “Some Statistics Useful in Regional Frequency Analysis.” *Water Resources Research* 29 (2): 271–81.
- Huard, David, Alain Maillhot, and Sophie Duchesne. 2010. “Bayesian Estimation of Intensity–Duration–Frequency Curves and of the Return Period Associated to a Given Rainfall Event.” *Stochastic Environmental Research and Risk Assessment* 24 (3): 337–47.
- Morice, Colin P, John J Kennedy, Nick A Rayner, and Phil D Jones. 2012. “Quantifying Uncertainties in Global and Regional Temperature Change Using an Ensemble of Observational Estimates: The Had-CRUT4 Data Set.” *Journal of Geophysical Research: Atmospheres* 117 (D8).
- R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Schlögl, Matthias, and Gregor Laaha. 2017. “Extreme Weather Exposure Identification for Road Networks—a Comparative Assessment of Statistical Methods.” *Natural Hazards and Earth System Sciences* 17 (4): 515–31.
- Silva Lomba, Jessica, and Maria Isabel Fraga Alves. 2020. “L-Moments for Automatic Threshold Selection in Extreme Value Analysis.” *Stochastic Environmental Research and Risk Assessment* 34 (3): 465–91.
- Wadsworth, Jennifer L. 2016. “Exploiting Structure of Maximum Likelihood Estimators for Extreme Value Threshold Selection.” *Technometrics* 58 (1): 116–26.