# Temperature Extremes Analysis in Beznau

Mattia Mariantoni and Jaka Strmcnik

January 8, 2021

**Abstract**

The aim of this project is to analyze and asses the risks of temeperature extremes at the nuclear installation in Beznau, Switzerland, based on monthly maxima data from the past 25 years. The final goal is to estimate the quantiles of temperature extremes at probability levels $0.99, 0.999$ and $0.9999$ for the year 2030. To go further, data from Beznau is then analyzed against the data for temperature montly maxima in a nearby city Mühleberg, for which there is 100 years of data at disposal. Finally, the dependence of extremes between the two sets of data is modeled in order to estimate possible dependence between maxima for the two sites.

## 1 Introduction and Data Description

The first data set considered in this paper is a collection of 25 years of monthly maxima of daily maximum temperatures in Beznau, Switzerland. The events of extreme temperatures, like some other weather extremes, such as heavy rainfall, extreme river water temperature and heavy snowfall, for example, can present significant hazards for a nuclear, such as in Beznau. Especially after the Fukushima catastrophe in 2011, nuclear regulators have been re-assesing the risk due to these natural hazards.

The focus potential hazard of this paper will be extremely high temperatures. To begin with, Figure 1 represents the data collected. The monthly maxima behaviour will later on be compared with the annual maxima, which is why the last two data entries, meaning the maxima from January and February 2013 were removed from consideration, because the data was not collected for all the months until the end of the year. As we can see there seems to not be any significant upward or downward trend, although the data does seem to reflect a sense seasonality. We can clearly observe that annual maxima are achieved during the summer, which is obviously in accordance with the fact that Beznau lies on the northern hemisphere.

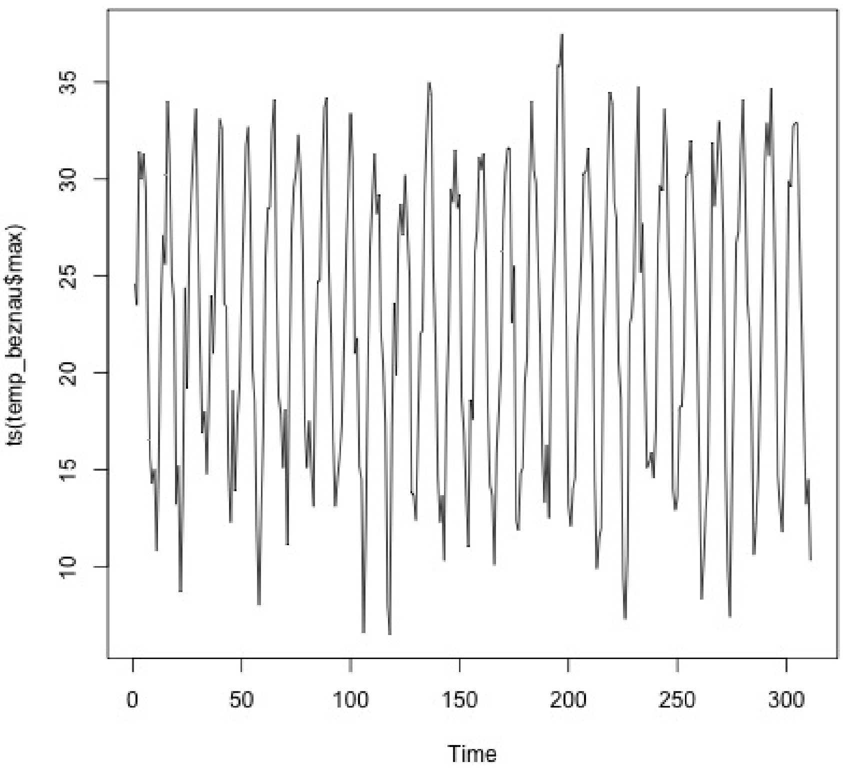


Figure 1: Beznau data plot.

## 2 GEV Modelling.

In order to model the maxima, GEV distribution was used. Firstly, the data was modelled with a stationary GEV distribution, and later with a non-stationary one, both via the maximum likelihood analysis. The latter one proved to be a better fit according to AIC. The non-stationarity was modelled by the following regression forms for the GEV parameters:

$$\mu(t) = \mu_0 + \mu_1 (t - t_0)/(100 \times 365.25) + \sum_{k=1}^{K} [\mu_{c,k} cos\{\frac{2k\pi(t - t_0)}{365.25}\} + \mu_{s,k} sin\{\frac{2k\pi(t - t_0)}{365.25}\}],$$

$$\sigma(t) = \sigma_0 + \sigma_1 (t - t_0)/(100 \times 365.25) + \sum_{k=1}^{K} [\sigma_{c,k} cos\{\frac{2k\pi(t - t_0)}{365.25}\} + \sigma_{s,k} sin\{\frac{2k\pi(t - t_0)}{365.25}\}],$$

$$\xi(t) = \xi_0 + \xi_1 (t - t_0)/(100 \times 365.25) + \sum_{k=1}^{K} [\xi_{c,k} cos\{\frac{2k\pi(t - t_0)}{365.25}\} + \mu_{s,k} sin\{\frac{2k\pi(t - t_0)}{365.25}\}],$$

where $t$ is the day of the year after $t_0$, which was chosen to be January 1st 2000.

The main question of modelling non-stationarity was choosing $K$ such that the regression forms would best fit the data. To evaluate models, AIC criterium was used. Table 1 shows values of AIC for given $K$ values in the regression form of parameter $\mu$, ranging from 1 to 6.

| K | AIC |
|---|---|
| 1 | 1465.574 |
| 2 | 1477.454 |
| 3 | 1540.382 |
| 4 | 1659.974 |
| 5 | 1585.941 |
| 6 | 1520.427 |

Table 1: AIC for different values of $K$ for the regression form of parameter $\mu$.

It is visible from Table 1 that $K = 1$ seems to be the best fit. For $\sigma$ and $\xi$ it turned out to be the best fit if they are kept constant. Furthermore, because the maximum likelihod estimate of parameter $\mu(t)$ turned out to be within the 95% confidence interval from zero, we can say that the linear term $\mu_1$ in the regression form for $\mu(t)$, equals zero. Similarly we can confirm seasonality because some of the coefficients infront of sine and cosine are nonzero at the 95% confidence interval. Thus we have constant $\sigma$ and $\xi$ and

$$\mu(t) = \mu(t_0) + \mu_{c,1} cos\{\frac{2k\pi(t - t_0)}{365.25}\} + \mu_{s,1} sin\{\frac{2k\pi(t - t_0)}{365.25}\}.$$

To insert explicit numbers, we get the following:

$$\sigma = 2.4891818_{0.1069678}, \xi = -0.2411608_{0.0296711}$$

and

$$\mu(t) = 21.0404568_{0.1547317} - 1.1557169_{0.2021324} cos\{\frac{2k\pi(t-t_0)}{365.25}\} - 10.0557440_{0.2065436} sin\{\frac{2k\pi(t-t_0)}{365.25}\}.$$

Diagnostic plots of this model are presented in Figure 2, Figure 3 and Figure 4.

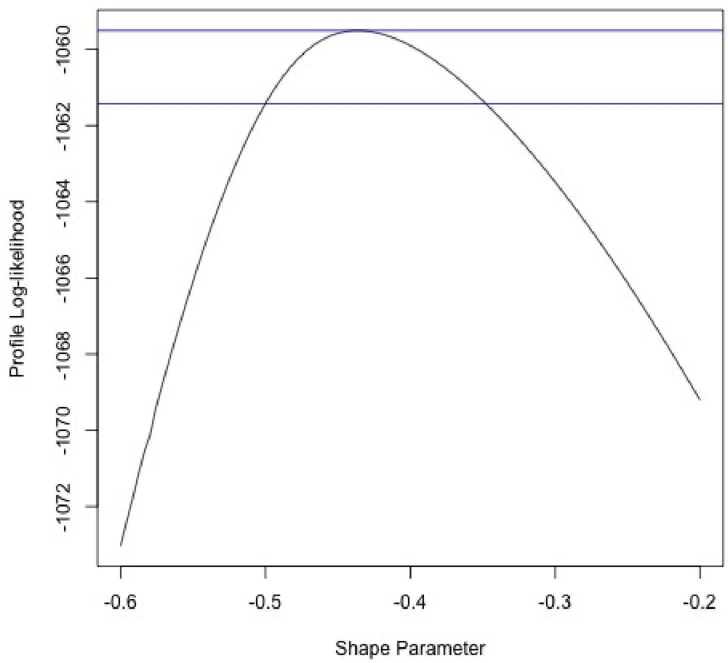


Figure 2: Profile log-likelihood of shape parameter.

If the original monthly maxima dataset were adapted to annual maxima datasets, and we would use stationary GEV distribution as a model based on maximum likelihood, the following parameters would be obtained:

$$\sigma = 1.4571765_{0.2269814}, \xi = -0.1535069_{0.1262831}, \mu = 32.5346628_{0.3232727}.$$

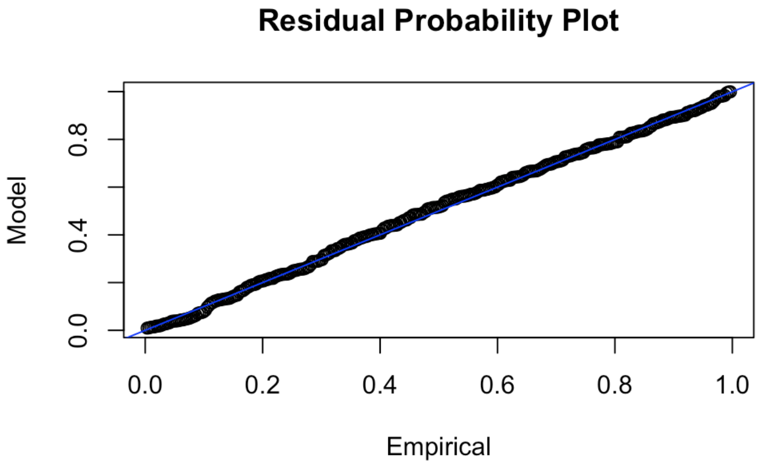


Figure 3: Residual probability plot.

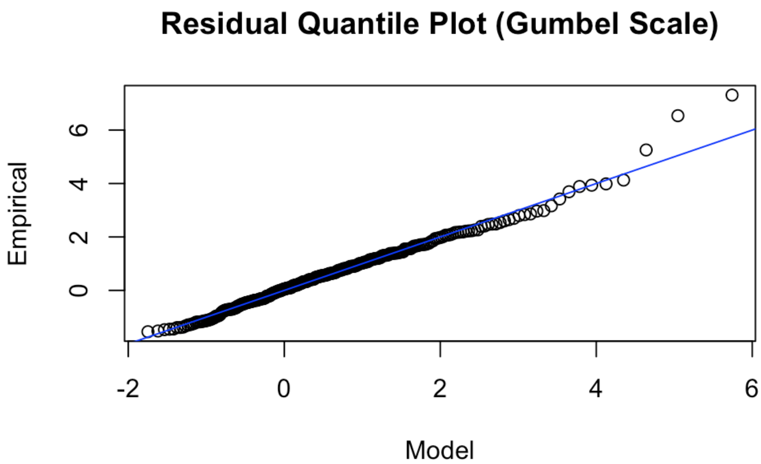


Figure 4: Residual quantile plot.

# 3 Quantiles and Return Levels

Using the GEV model of monthly maxima described above, this section estimates the quantiles of temperature extremes at probability levels 0.99, 0.999 and 0.9999, which correspond to 100, 1000 and 10000 year return levels. The quantiles were calculated for the 15th day of each month in the year 2030 and are shown in Table 2.

| | 100 | 1000 | 10000 |
|---|---|---|---|
| 1 | $17.929_{0.662}$ | $19.381_{0.891}$ | $20.213_{1.069}$ |
| 2 | $20.014_{0.662}$ | $21.466_{0.891}$ | $22.98_{1.069}$ |
| 3 | $23.823_{0.662}$ | $25.276_{0.891}$ | $26.107_{1.069}$ |
| 4 | $29.094_{0.662}$ | $30.546_{0.891}$ | $31.378_{1.069}$ |
| 5 | $33.909_{0.662}$ | $35.362_{0.891}$ | $36.193_{1.069}$ |
| 6 | $37.245_{0.662}$ | $38.698_{0.891}$ | $39.529_{1.069}$ |
| 7 | $38.022_{0.662}$ | $39.475_{0.891}$ | $40.306_{1.069}$ |
| 8 | $36.075_{0.662}$ | $37.528_{0.891}$ | $40.306_{1.069}$ |
| 9 | $31.874_{0.662}$ | $33.326_{0.891}$ | $34.158_{1.069}$ |
| 10 | $26.758_{0.662}$ | $28.210_{0.891}$ | $29.042_{1.069}$ |
| 11 | $21.815_{0.662}$ | $23.268_{0.891}$ | $24.099_{1.069}$ |
| 12 | $18.646_{0.662}$ | $20.098_{0.891}$ | $20.930_{1.069}$ |

Table 2: Return levels with standard errors for every month in 2030.

Again, we could perform the same task for the data of annual maxima. When we do so, we obtain the return levels annual maxima shown in Table 3.

| 100 | 1000 | 10000 |
|---|---|---|
| $37.342_{1.071}$ | $38.740_{1.914}$ | $39.719_{2.801}$ |

Table 3: Return levels of annual extremes with standard errors for the year 2030.

This data seems to go along fine with the fact that the yearly extremes are mostly to be expected in the summer months, so the annual return levels seem to nicely correlate with monthly return levels of the summer months, depicted in Table 2.

# 4 Modelling dependence to the Mühleberg dataset

This section takes into account the monthly maxima data for the city of Mühleberg. This data series has a lot more entries than the one for Beznau, 100 years worth of data, to be precise. Also this time, the two entries for the year 2013 remain omited. Figure 5 shows the monthly maxima for daily maximum temperatures in Mühleberg.

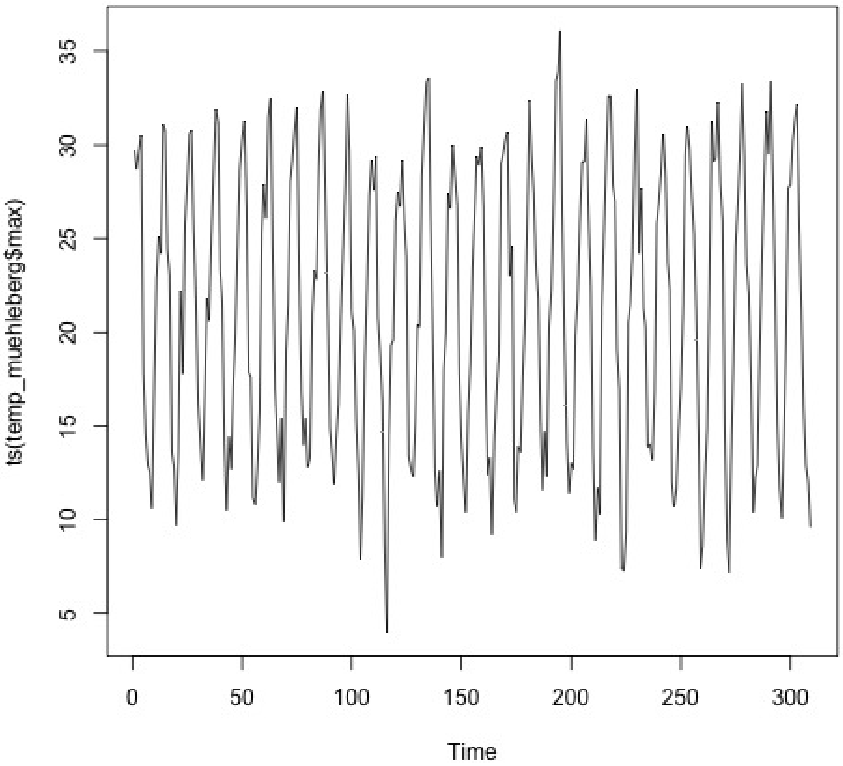


Figure 5: Mühleberg data plot.

Next, an apropriate bivariate model had to be established. Firstly, we applied the same GEV maximum likelihood model to the Mühleberg data, as we did before to the Beznau dataset, since it proved to be the best fit. We obtained the following estimates for the GEV parameters:

$$\sigma = 2.3322674_{0.09799278}, \xi = -0.2560641_{0.02643195}$$

and

$$\mu(t) = 19.8644542_{0.14347152} - 1.3413196_{0.19504398} cos\{\frac{2k\pi(t-t_0)}{365.25}\} - 9.9855231_{0.18738734} sin\{\frac{2k\pi(t-t_0)}{365.25}\}.$$

We observe that the 95% confidence intervals for the parameter $\xi$ overleap in the model for Mühleberg data and the model for Beznau data, therefore, we can assume the same shape parameter for both of them. Furthermore, when we calculate the correlation between the two datasets, we get a result of 0.9888295, which suggests that the datasets are highly correlated.

As mentioned already, a bivariate extreme model hat to be fitted. Upon comparrison via the AIC, we established that among the logistic, negative logistic and the Coles-Tawn model, the latter was the best fit. To examine dependence and assymptotic dependence, the chi plots were plotted and are shown in Figure 6 and Figure 7. Since values on both plots approach 1 as the quantile value approaches to 1, we can consider the two datasets to be dependent.

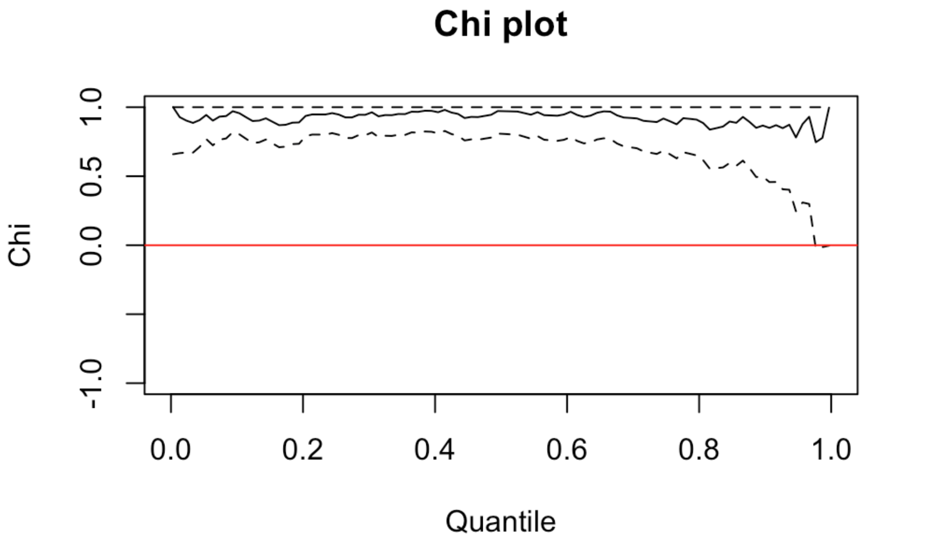


Figure 6: Chi plot.

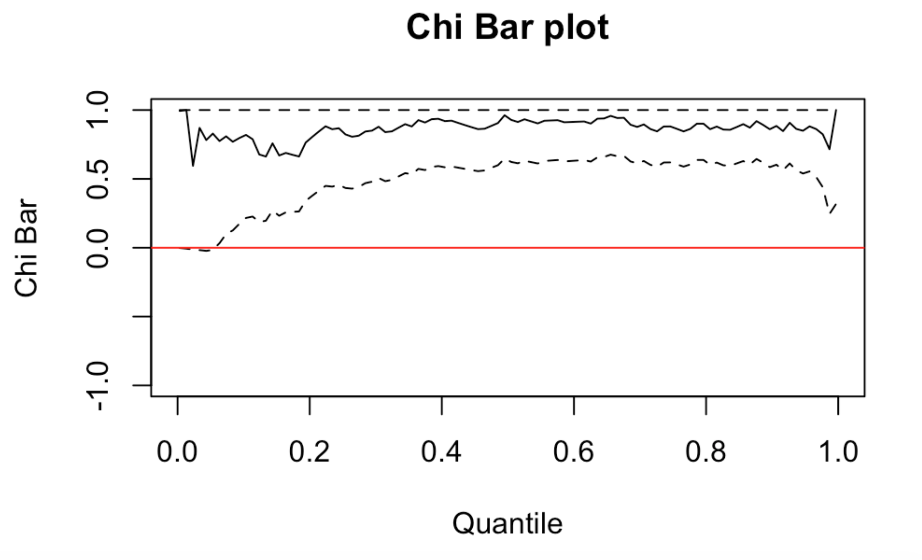


Figure 7: Chi bar plot.

# 5 Conclusion

To conclude, as we have seen, the data of maximal temperatures follows a reverse Gumbell distribution, due to the fact that the $\xi$ parameter proved to be negative for the monthly maxima, as well as for the annual maxima. Furthermore, obvious seasonality is observed, with annual maxima taking place in the summer months, which of course makes complete sense. In the project, the monthly expected levels of return were calculated for the year 2030. As we have seen, the monthly return levels and annual return levels seem to coincide with the fact that the maximum annual temperature is likely to happen in the summer months. Finally, comparing the dataset of Beznau with that of Mühleberg reflects a great deal of correlation. Moreover, the two datasets appear to be dependent as well as asymptotically dependent as portrayed on the chi plots. A possible improvement of the project could include consideration of other possible models for the data, which would perhaps produce even more compelling results.