Extreme Event Modelling

Liwei Wu

Abstract

In the **Theory and Methods** Part of this paper, we discussed the basic theory and methods of Extreme Value Modelling. Both the Generalized Extreme Value (GEV) distribution and the Generalized Pareto Distribution (GPD) are introduced. Correspondingly, there are two methods to model the Extremes, i.e. the Block Maxima method and the Threshold method. Then in the **Application** Part, we applied these methods to fit the data to obtain the corresponding return level. Firstly, we applied the Block Maxima Method to the simulated normal data. Then we applied the Threshold Method to the Dow Jones Index data and the Hong Kong climate data and reached informative conclusion based on the return levels we obtained. *KEY WORDS*: Extreme value

theory; Generalized Pareto distribution; Dow Jones Index; Hong Kong climate data; Threshold method.

The Generalized Extreme Value Distribution

If there exist sequences of constants $\{a_n > 0\}$ and $\{b_n\}$ such that

$$P\left(\frac{M_n - b_n}{a_n} \le z\right) \to G(z) \text{ as } n \to \infty \quad (1)$$

for a non-degenerate distribution function G, then G is a member of the GEV family

$$G(z) = \exp\left\{-\left[1 + \xi\left(\frac{z-\mu}{\sigma}\right)\right]^{-1/\xi}\right\}$$

defined on the set $\{z : 1 + \xi(z-\mu)/\sigma > 0\}$, where $-\infty < \mu < \infty$, $\sigma > 0$ and $-\infty < \xi < \infty$.

The Generalized Pareto Distribution

Denote an arbitrary term in the X_i sequence by X, and suppose that F satisfies Theorem 2.2, so that for large n,

$$P(M_n \le z) \approx G(z),$$

where

$$G(z) = \exp\left\{-\left[1 + \xi\left(\frac{z-\mu}{\sigma}\right)\right]^{-1/\xi}\right\}$$

for some $\mu, \sigma > 0, \xi$. Then, for large enough u, the distribution function of $(X - u)$, conditional on $X > u$, is approximately

$$H(y) = 1 - \left(1 + \frac{\xi y}{\tilde{\sigma}}\right)^{-1/\xi} \quad (2)$$

defined on $\{y : y > 0 \text{ and } (1 + \xi y/\tilde{\sigma}) > 0\}$, where

$$\tilde{\sigma} = \sigma + \xi(u - \mu)$$

Data points plot

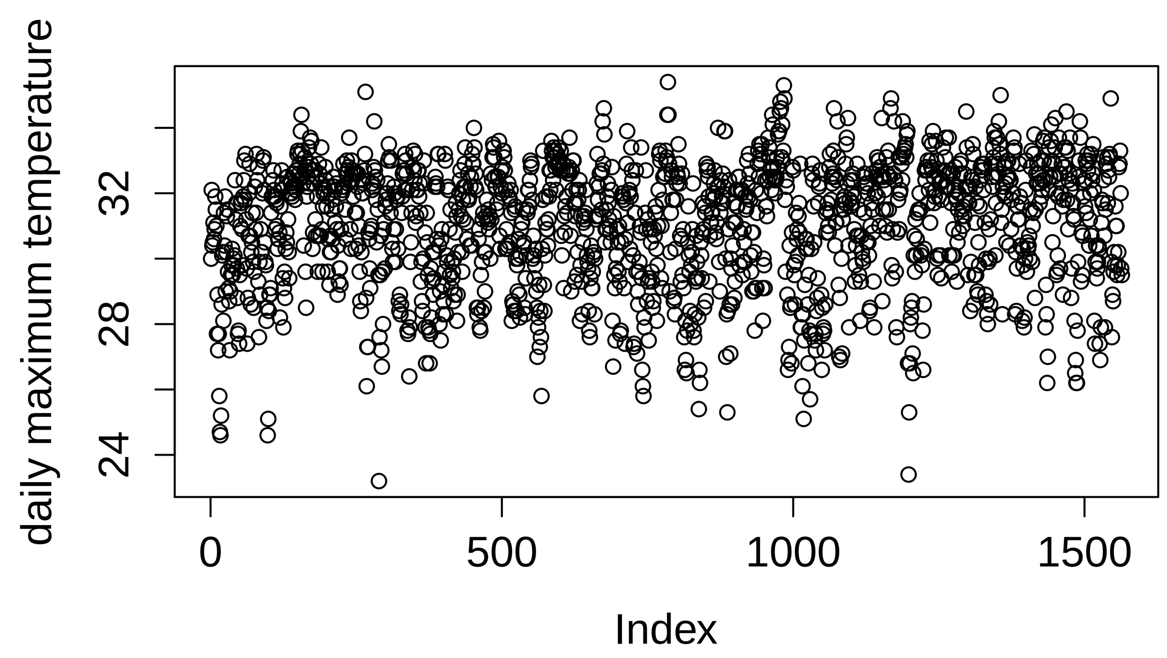

Figure : data points of daily maximum temperature of Hong Kong climate data

Selecting the proper threshold

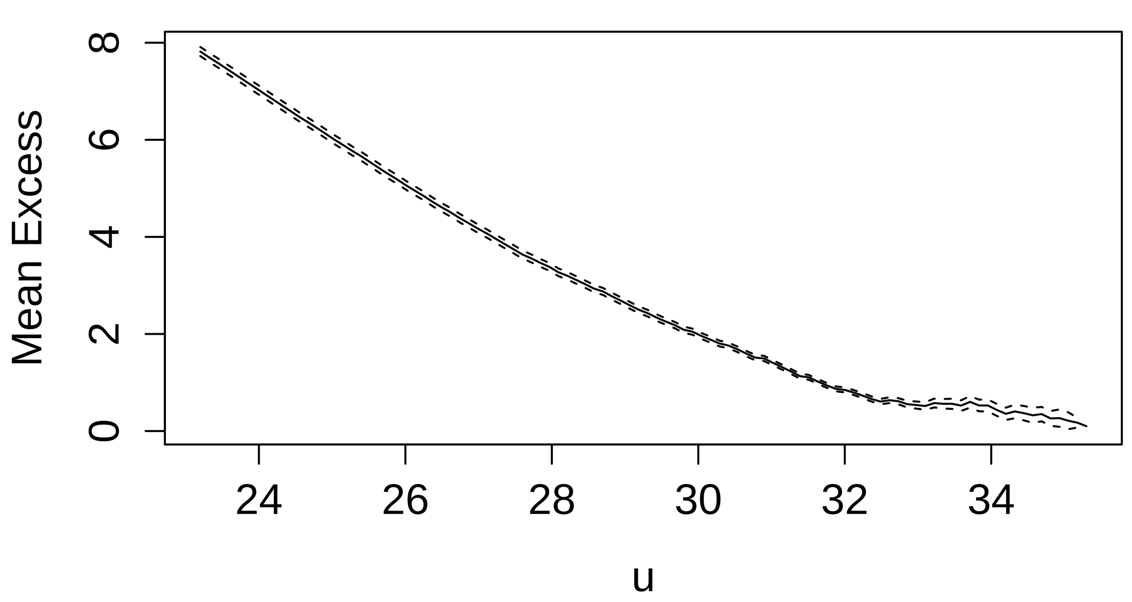

Figure : mean residue life plot

Model Checking

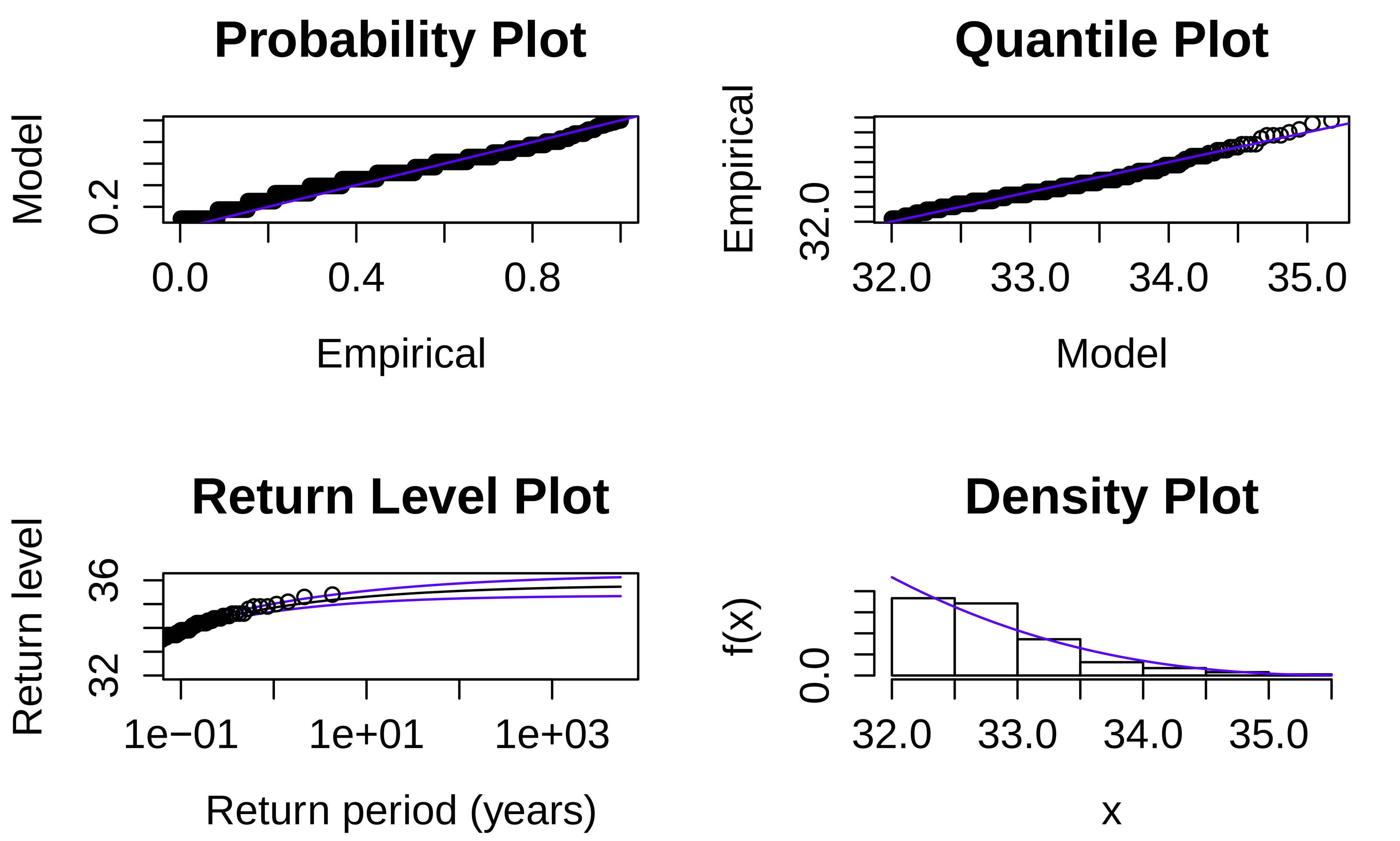

Figure : Diagnostic plots

Return level plot

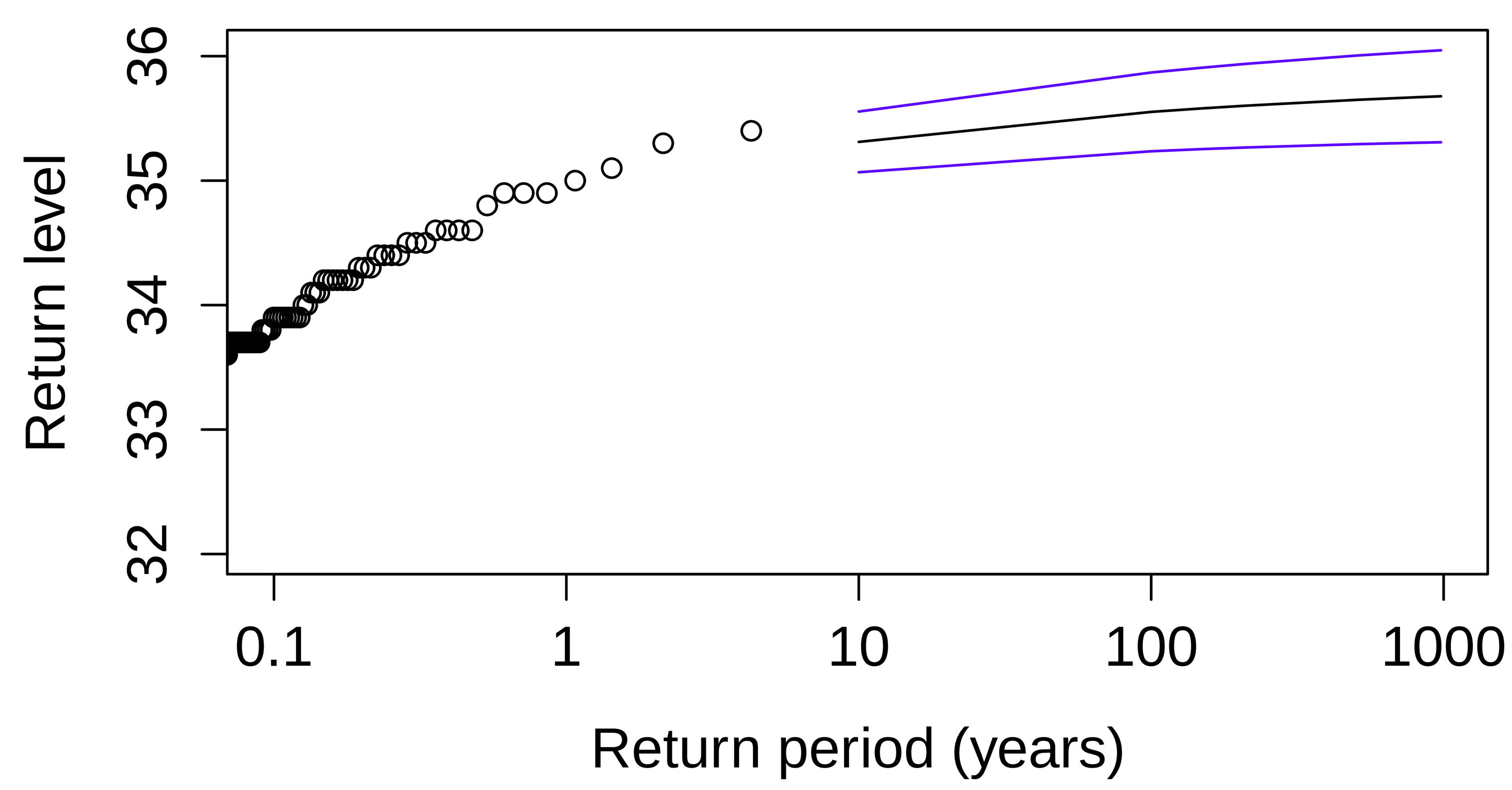

Figure : return level plot

Conclusion

- we are 95% confident to say that the maximum daily air temperature within 100 years will fall between 35.23559 and 35.86844.
- we are 95% confident to say that the maximum daily air temperature within 150 years will fall between 35.25327 and 35.90782.
- Although the range is pretty wide, we can still conclude that the maximum air temperature will most likely never reach 36 degrees in Hong Kong even within 150 years. Therefore, we know what we are up to and take some precautionary measures in the event of extremely high air temperature.

Outlook

- In the HK climate data case, we assume that the data are obtained from one station of the same location. However, in real world application, it is possible that the data are obtained from a number of stations of different locations.
- If so, we can use the Bayesian hierarchical model for spatial extremes to produce a map characterising extreme behaviour across a geographic region.
- By the Bayes rule,

$$p(\theta|Z(x)) \propto p_1(Z(x)|\theta_1)p_2(\theta_1|\theta_2)p_3(\theta_2)$$

- Based on the equation above, we can obtain the posterior distributions of spatially dependent parameters by using MCMC algorithms. Then, the return level posterior distribution as well as the return level maps can be produced accordingly.

Reference

1. Stuart Coles, *An Introduction to Statistical Modeling of Extreme Values* (New York: Springer, 2001).
2. Daniel Cooley, Douglas Nychka & Philippe Naveau (2007) Bayesian Spatial Modeling of Extreme Precipitation Return Levels, Journal of the American Statistical Association, 102:479, 824-840, DOI: 10.1198/016214506000000780

For more information contact liweiwu2-c@my.cityu.edu.hk