Chapter 6
Quantifying Extreme Risks

Vicky Fasen, Claudia Klüppelberg, and Annette Menzel

Understanding and managing risks caused by extreme events is one of the most demanding problems of our society. We consider this topic from a statistical point of view and present some of the probabilistic and statistical theory, which was developed to model and quantify extreme events. By the very nature of an extreme event there will never be enough data to predict a future risk in the classical statistical sense. However, a rather clever probabilistic theory provides us with model classes relevant for the assessment of extreme events. Moreover, specific statistical methods allow for the prediction of rare events, even outside the range of previous observations. We will present the basic theory and relevant examples from climatology (climate change), insurance (return periods of large claims) and finance (portfolio losses and Value-at-Risk estimation).

Keywords Climate risk · Extreme risk analysis · Extreme value distribution · Financial risk · Peaks over thresholds

Mathematics Subject Classification (2010) 60G70 · 62G32

The Facts

- Modern risk measures like Value-at-Risk and Expected Shortfall are defined by high quantiles, such that the probability of a large loss is small.

V. Fasen
Institute of Stochastics, Department of Mathematics, Karlsruhe Institute of Technology, Kaiserstr. 89, 76133 Karlsruhe, Germany

C. Klüppelberg (✉)
Chair of Mathematical Statistics, Center for Mathematical Sciences, Technische Universität München, Boltzmannstr. 3, 85748 Garching bei München, Germany
e-mail: cklu@tum.de

A. Menzel
Chair of Ecoclimatology, Center of Life and Food Sciences Weihenstephan, Technische Universität München, Hans-Carl-von-Carlowitz-Platz 2, 85354 Freising-Weihenstephan, Germany

C. Klüppelberg et al. (eds.), *Risk – A Multidisciplinary Introduction*,
DOI 10.1007/978-3-319-04486-6_6,

- Poisson's classic theorem on rare events (also called the law of small numbers) is the basis for extreme value statistics, because it says that the Poisson distribution is the limit of binomial distributions with very small success probabilities.
- The distribution of maxima of large samples can only be a Generalized Extreme Value (GEV) distribution. This is one of the most fundamental results of extreme value theory. On this basis methods to estimate far out tails and high quantiles were developed.
- Another method to estimate far out tails and high quantiles is the Peaks-Over-Threshold (POT) method using the fact that exceedances over high thresholds for large samples follow a Generalized Pareto distribution (GPD).
- We quantify extreme events for three data examples:
  - yearly temperature maxima from 1879–2008;
  - claim sizes of a Danish fire insurance;
  - daily returns of the Standard and Poors 500 Index.

1 Introduction

Extreme risks accompany our lives. Although every single person hopes that she does not suffer any losses, some lose a fortune in a financial crises, some others lose their property in a hurricane, or they have to leave their homes because of a nuclear accident, another person may even lose her life in a car accident or because of a terrorist attack. Whereas our ancestors took dangers and risks as God-given, nowadays we trace the occurrence of most types of risk back to the actions of men. This implies that risk is precisely calculable (an assumption that is mostly wrong), and that somebody has to be responsible. This applies to technical risk, where safety measures are implemented in order to prevent disasters, which still happen occasionally. We even try to adapt to risk of natural catastrophes, when we develop strategies like, for instance, building dikes or simply sign an insurance contract.

In a society guided by such believes it is natural to require formulas from Mathematics and Statistics for risk assessment. It is within this framework that *extreme value theory* and *extreme value statistics* find their natural place. However, the modeling and the assessment of extreme events is not so simple and cannot be gained with standard methods.

We illustrate the problem with a classical example.

Illustration 1.1 (Determine the Height of a Dike) In the Netherlands, where substantial parts of the country are below sealevel, dikes of appropriate height are of vital importance as protection against floods. The dikes have to be built higher than a wave height, which happens at most every 10,000 years. How high has the dike at least to be? Or formulated otherwise, how does one estimate the height of the highest wave in 10,000 years, if one has only measurements of some hundred years available? The problem is to estimate the probability of an event which is more extreme than any recorded to date. This requires a special method, which is provided by statistical methods based on extreme value theory.

Extreme value theory is a fundamental mathematical theory, which can be transferred to statistical methods. It was developed during the last 50 years and is not undebated. Extreme value theory allows (under appropriate conditions) to predict rare events, which are not included in the previous observations because of their rareness. Based on extreme data (later they will be yearly temperature maxima, large insurance claims and large changes in a financial time series) it is possible to extrapolate the data for the prediction of events, which cause higher temperatures, insurance claims or financial losses than have ever been observed before. Naturally it is easy to criticize this extrapolation out of the sample data and it is clear that extrapolation is unreliable by nature. However, extreme value theory provides a solid mathematical basis, and no other reliable alternative has been suggested. We cite the following assessment of Professor Richard Smith (http://www.unc.edu/~rls/), who has substantially contributed to the development of extreme value statistics: "There is always going to be an element of doubt, as one is extrapolating into areas one doesn't know about. But what extreme value theory is doing is making the best use of whatever you have about extreme phenomena".

We emphasize that the statistical treatment of rare events as the far-out tail behavior can only succeed with specific methods, which implement probabilistic results of extreme value theory into the estimation procedure and, hence, compensate for the insufficient amount of data. This will be the topic of Sects. 3 and 4. Parts of this chapter have corresponding parts in Fasen and Klüppelberg [29].

2 Extreme Risks

2.1 Climate Risk

Fire, water, air—these three basic elements cause climate or weather-related natural disasters. They comprise meteorological hazards (such as storm, hail, lightning), hydrological (flooding, mass movement), and climatological ones (such as extreme temperatures, heat waves, drought, forest fire). Apart from devastating earthquakes in Chile, Haiti (2010) and Japan, New Zealand (2011), making 2011 the costliest year ever, the natural catastrophe losses in the last few years were dominated by weather-related catastrophes, such as devastating floods in Pakistan (2010) and Thailand (2011), the Winter Storm Xynthia in western Europe (2010), Hurricane Sandy in the US (2012), wildfires in Russia (2010) and the summer drought in the US (2012) (see also Sect. 2.3 Insurance Risks). According to Munich Re data, there is an increasing trend of these natural disasters in respect to intensities, frequencies, damages and losses. The Intergovernmental Panel on Climate Change (IPCC) concluded in its last report in 2007 (see [6]) that in past records the dominant signal was significantly increased in the values of exposure at risk. However climate change has likely altered and will virtually certainly alter also the occurrence of extreme events dramatically: frequency and magnitude of extreme events are strongly linked to anthropogenic induced climate change.

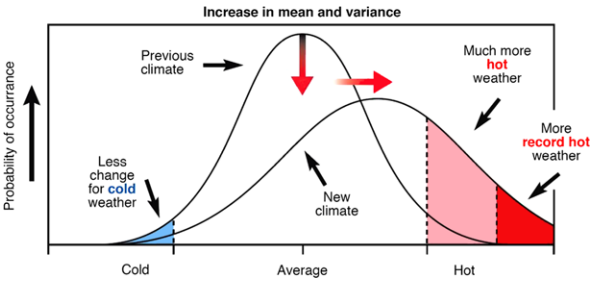

Fig. 1 Illustration of the consequences of an increase of temperature in mean and variance

The latest IPCC report confirmed a 100-year linear trend (1906–2005) of 0.74 °C, more precisely, eleven of the last twelve years (1995–2006) ranked among the 12 warmest years in the instrumental record of global surface temperature since 1850. Most of the observed warming since the mid-20th century is very likely due to the observed increase in anthropogenic greenhouse gas concentrations. Linked to this climate change are marked observed changes in extreme events, much more intense and longer droughts since the 1970s, particularly in the tropics and subtropics, higher frequency of heavy precipitation events, or widespread changes in extreme temperatures. For the latter one, a human contribution to the observed trends is likely. Also future trends have been assessed by simulation of different scenarios with strong impacts on extreme events, e.g., increase in intense tropical cyclone activity or incidence of extreme high sea level are likely at the end of the 21th century. Due to the importance of extreme events the IPCC published a Special Report Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation (SREX) in 2012.

Many important research questions are linked to this increase in weather related extreme events. First of all, is climate becoming more extreme under climate change conditions? This question has traditionally been answered by fitting Gaussian distributions to temperatures. Figure 1 displays how an increase in mean and variance of temperature causes more hot and more record hot weather. However, Gaussian distributions do not provide a good fit for the distribution tails of high temperature measurements.

Second, if there are changes in extremes, which vulnerability of humans is to be expected? Not all extreme events end in disasters. The most recent World Risk Report of 2012, published by the BündnisEntwicklungshilfe in cooperation with the United Nations University (UNU-EHS) (http://www.weltrisikobericht.de), summarizes the risk by natural hazards to nations with different vulnerability, starting with

(1) the likelihood of extremes to occur (exposition),
(2) the vulnerability of societies with respect to infrastructure, housing, food, poverty, economy,
(3) the coping capacity based on governance, catastrophe precautions, medical situation, social networks, insurances, and
(4) the adaptation capacity linked to education, environmental protection, projects and investments.

Similarly, it is a question of tremendous importance how the occurrence of physical extreme events translates to extreme biological impacts or hazards which threaten the fitness and survival of ecosystems more than any change in mean conditions (cf. Hegerl, Hanlon, and Beierkuhnlein [5], Menzel, Seifert, and Estrella [8]). Not all rare climatological events translate into extreme impacts: the responses in nature may be non-linear, the species may be resilient, resistant, recover fast, or are well adapted by management. Due to this variation in response, always more and more data on impacts of extreme events are needed. The goal is to bridge the gap between extreme events and extreme impacts, especially for climatological hazards, such as temperature extremes, heat waves, cold spells, frost events, drought or fire. They impact primarily agricultural and forest ecosystems, however, as combined, longer lasting events their proper statistical modeling and assessment is a scientific challenge.

2.2 Financial Risks

The Basel Committee for Banking Supervision (http://www.bis.org/bcbs/) recommends for insurance companies and financial institutions the building of capital reserves to hedge against unpredictable risks. This is in Germany explicitly required by the regulatory authorities, the BAFIN (Bundesanstalt für Finanzdienstleistungsaufsicht, http://www.bafin.de/) in the framework of "Basel II" for banks (http://www.bis.org/publ/) and in the framework of "Solvency II" for insurance companies (http://ec.europa.eu/internal_market/insurance/). The risk management department of every company is responsible for the respective calculations of the required capital reserves and their administration, which requires a mathematical-statistical training.

The focus of Basel II, which was initially published in June 2004, was to manage and measure *credit risks*, *operational risks* and *market risks*. In this chapter we will only pay attention to market risk, the risk that a value of a portfolio will change due to movements in the market risk factors as, e.g., interest rates, foreign exchange rates, equity prices and commodity prices.

In the Basel framework the capital requirement for market risk is based on the so-called *Value-at-Risk*, which is the p-quantile of the portfolio risk, and is defined as follows.

Let X be the financial risk in terms of the *daily losses*, defined as the negative profit/loss of the market portfolio. To be precise, if Z_t for $t = 1, 2, \ldots$ denote the daily market prices of the portfolio, then the losses X_t represent the daily negative log-returns defined as $X_t = -(\log Z_t - \log Z_{t-1}) \approx -(Z_t - Z_{t-1})/Z_{t-1}$, approximating the negative relative price changes for each day.

The distribution function of the daily portfolio loss X is given by $F(x) = \mathbb{P}(X \le x)$ for $x \in \mathbb{R}$. We define the *quantile function* of F or *Value-at-Risk* as

$$\mathrm{VaR}_p(X) = F^{-1}(p) = \inf\{x : F(x) \ge p\}, \quad p \in (0, 1). \tag{2.1}$$

(Note that for strictly increasing F this is simply the analytic inverse.) Hence, $\mathrm{VaR}_p(X)$ is the smallest number such that the probability of a loss larger than $\mathrm{VaR}_p(X)$ does not exceed $1-p$. Then for a large value of p (usually $p=0.95$ or larger) $\mathrm{VaR}_p(X)$ is a prominent risk measure.

Depending on the specific risk, choices are $p=95\ \%$ (0.95) or $p=99\ \%$ (0.99) or even $p=99.9\ \%$ (0.999). In the case of market risks $p=99\ \%$.

By the perception and experiences gained through the financial crises, which started in 2007, the Basel Committee on Banking Supervision decided a reformation of Basel II to strengthen the regulation, supervision and risk management of the banking sector in September 2010. This revision had to be implemented until 31 December 2011 [16] and introduced—as a response to the crises—a *stressed Value-at-Risk* requirement taking into account a historic one-year observation period relating to significant losses, which must be estimated in addition to the classical Value-at-Risk based on the recent one-year observation period. Basel III [15] now aims at raising the resilience of the banking sector by strengthening the risk coverage of the capital reserves. It suggests reforms of capital requirements for counterparty credit risk using stressed inputs, addresses the systemic risk arising from the interconnectedness of banks and other financial institutions, and supplements the risk-based capital requirement to constrain too high leverage (details to the changes in market risk can be found in http://www.bis.org/publ/bcbs193.htm). The implementation of Basel III will start in 2013.

Typical methods to estimate the Value-at-Risk in practice are *historical simulations*, the *variance-covariance method* and *Monte Carlo simulation*.

The "historical simulation method" simply estimates $\mathrm{VaR}_p(X)$ by the corresponding empirical quantile based on the required one year of data. For instance, $\mathrm{VaR}_{0.99}(X)$ is estimated as the largest 1 % of daily losses. Alternatively, a weighted estimation scheme is used, which gives higher weights to those data near to the current date and lower to the more distant data. Criticism of this method is obvious: reliable estimation of high quantiles like $\mathrm{VaR}_{0.99}(X)$ requires a large amount of high losses, but 1 % of the required one year of data provides no reliable estimator. Consequently, the estimated $\mathrm{VaR}_{0.99}(X)$ depends very much on the present market situation and estimates can differ substantially almost from day to day. We shall analyse the Standard and Poors 500 Index data during 1990–2004, abbreviated as S&P500. Moreover, $\mathrm{VaR}_{0.99}(X)$ is supposed to predict future high losses, which may be substantially higher than losses of the previous year and requires extrapolation outside the observations.

For the "variance-covariance method" the risk factors are assumed to be multivariate normal distributed. Then the distribution function of the portfolio X is a one-dimensional normal distribution with mean $\mu \in \mathbb{R}$ and variance $\sigma > 0$ determined by the portfolio weights, the means and variances of the components and the pairwise correlations of the components. The loss distribution F of X is given by

$$F(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{x} \mathrm{e}^{-\frac{(y-\mu)^2}{\sigma^2}} \mathrm{d}y \quad \text{for } x \in \mathbb{R}. \tag{2.2}$$

Then $\mathrm{VaR}_{0.99}(X) = \mu + \sigma z_{0.99}$, where $z_{0.99}$ is the 0.99-quantile of the standard normal distribution. It is particularly easy to estimate and to update, when the estimates

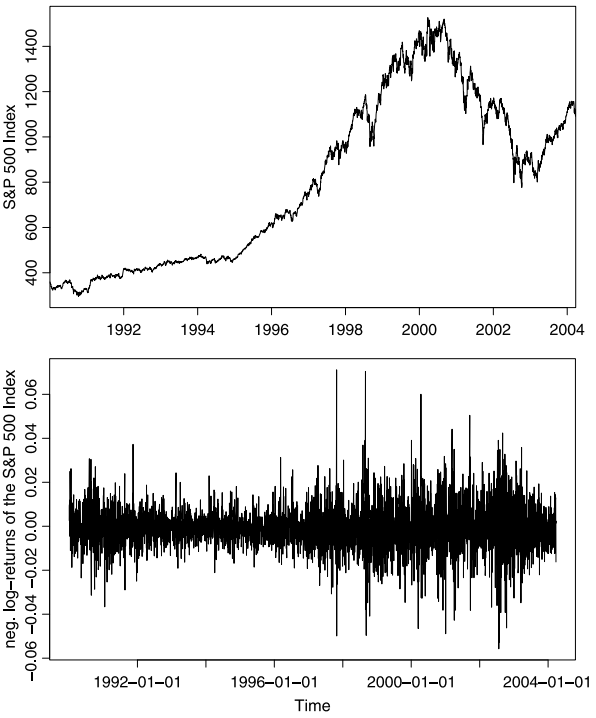

Fig. 2 The S&P500 (*top*) and the corresponding losses (*bottom*) during 1990–2004

for μ and σ change in time. In Fig. 2 we see the S&P500 (left) and its losses (right) during 1990–2004.

From this example we see that the normal model is completely inadequate: The histogram (empirical density) of the daily losses of the S&P500 and the normal density with mean and standard deviation estimated from the data are depicted in Fig. 3. The histogram clearly shows that the daily losses of the S&P500 have more mass in the tails than the normal distribution; i.e. for ± 0.03 and larger/smaller the histogram exhibits more large/small values than is likely for the normal distribution. This mismatch leads to an underestimation of the required capital reserve. The fact that the empirical distribution and the normal distribution differ around 0 is for risk management based on high quantiles irrelevant. Moreover, financial loss data are usually negatively skewed and leptokurtic, again properties which can not be captured by a Gaussian distribution.

The third VaR estimation method is the "Monte Carlo simulation". Here a more sophisticated parametric distributional model is fitted to the daily losses, its parameters are estimated, and then large numbers of random samples of arbitrary length are simulated, its VaR estimated for each sample, and then the average VaR is taken as an estimate. This method can be made more efficient by variance reduction methods (Glasserman [33], Korn [38]), and estimates VaR for a given model with arbitrary precision. However, the estimate depends on the chosen model (as it does for the

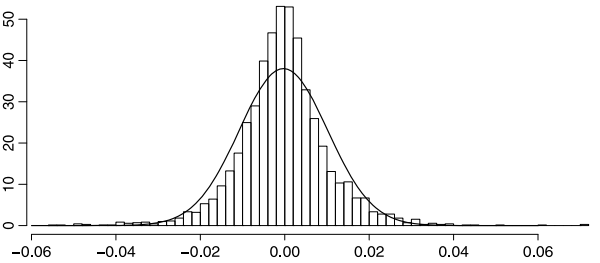

Fig. 3 Histogram of the daily losses of the S&P500 in comparison to the density of the normal distribution. The mean μ and the variance σ^2 have been estimated by their empirical versions

normal model in the variance-covariance method), so model risk can be considerable; cf. Chap. 10, Bannör and Scherer [14].

Remark 2.1 (i) In the Basel II market risk framework the calculation of the capital reserves requires as risk measure the Value-at-Risk for a holding period of 10 days at a confidence level 0.99 %. A standard method in practice to calculate the Value-at-Risk for a holding period of 10 days is to calculate the Value-at-Risk for a holding period of one day and scale it by $\sqrt{10}$. This scaling factor is based on the scaling property of the normal distribution and can be completely wrong.

(ii) In the amendments to the Basel II accord, which have been incorporated into Basel III ([15]), the $\mathrm{VaR}_{0.99}$ has been extended to incorporate so-called stressed periods like the financial crises during 2007/2008. Let X denote the loss of a market risk portfolio (over the next 10 days) and $\mathrm{VaR}_{0.99,avg}(X)$ the average of the estimated VaR values of the preceding 60 business days. Then the new capital requirement has to be calculated according to

$$\begin{aligned} &\max\left\{\mathrm{VaR}_{0.99}(X), m_c \mathrm{VaR}_{0.99,avg}(X)\right\} \\ &\quad + \max\left\{\mathrm{SVaR}_{0.99}(X), m_s \mathrm{SVaR}_{0.99,avg}(X)\right\} \end{aligned} \tag{2.3}$$

where m_c and m_s are multiplication factors, which are not smaller than 3 (and are related to the ex-post performance of the bank's model). The quantity SVaR is the Value-at-Risk of the loss portfolio estimated from historical data of a 12-month period of significant financial stress; e.g the financial crises 2007/2008.

(iii) Finally, we argue that the Value-at-Risk is not an appropriate risk measure. It is appropriate for the dike height of Illustration 1.1, for financial risk however, the situation is different. If a flood with waves higher than the dike happens, the dike usually breaks and nothing can be done for salvation. The land behind the dike disappears under water. For financial risks, however, it is extremely relevant to know also the amount of resulting losses. This quantity is taken into account, when using the *Average Value-at-Risk* as an alternative risk measure, which describes the expected losses given a loss larger than the Value-at-Risk happens. It is given as

$$\mathrm{AVar}_p(X) = \frac{1}{1-p}\int_p^1 \mathrm{VaR}_\gamma d\gamma$$

(cf. Chap. 5, Biagini, Meyer-Brandis, and Svindland [19] for a detailed introduction into risk measures). If X has continuous distribution function F, then $\mathrm{AVar}_p(X) =$

$\mathbb{E}(X \mid X > \mathrm{VaR}_p(X))$, which represents exactly the expected losses, given an extreme loss occurs. A second drawback of the Value-at-Risk is that it is in general not subadditive, i.e. $\mathrm{VaR}_p(X+Y) \leq \mathrm{VaR}_p(X) + \mathrm{VaR}_p(Y)$ may not hold for risks X, Y. Subadditivity reflects the diversification effect. It is better to have a portfolio of risks than several individual risks. However, if for example X and Y are independent with distribution $F(x) = 1 - \frac{1}{1+x}$ for $x \geq 0$, then $\mathrm{VaR}_p(X+Y) > \mathrm{VaR}_p(X) + \mathrm{VaR}_p(Y)$ and there is no chance for risk diversification. In contrast, the Average Value-at-Risk is a subadditive risk measure. Although there were serious attempts to communicate to regulators that the Average Value-at-Risk may be a more appropriate risk measure (cf. Danielsson et al. [24]), this academic initiative was not successful. The lobby work of the banks has prevented this: the capital reserves calculated on the basis of Expected Shortfall would be substantially larger than the Value-at-Risk.

2.3 Insurance Risks

Insurance companies take over the risks of their customers. Typical insurance risks are health problems, death, accidents, burglary, floods and fire. With the acquisition of an insurance contract customers transfer their risk to an insurance company, which is then financially liable to insurance claims. Also the insurance company does not know the risk for a claim to happen to a customer, but by selling a large number of policies, it subsumes customers with similar risk in a portfolio and takes advantage of the fact that in a large portfolio with similar and independent risks the total claim amount is constant in mean. In probability theory this fact is proved and is called the *Law of Large Numbers*. For the insurance company this makes the risk of a portfolio of similar and independent risks calculable. Random fluctuations in the portfolio are hedged by reserves. In this context insurance companies have to evaluate the frequency as well as the severity of risks. To do this they have to suggest appropriate risk models and estimate the model parameters, they have to analyze the model statistically and test it under extreme conditions. But they also have to calculate the premiums and reserves. As capital reserves of insurance companies are substantial, it is also subject to capital regulations like Basel II. Taking the total insurance business into account, new regulations are being implemented under Solvency II, following the very same ideas as the Basel framework. We do not want to explain these ideas in detail, but instead want to present the very traditional concept of the *return period*, which is used universally to describe extreme events and serves as a risk measure, in particular, for abnormally large insurance claims.

Large claims are rare events with very high costs for an insurance company. They include natural catastrophes like earth quakes, fire, storms or floods, which are typical events where large claims occur (cf. Fig. 4), but also so-called *man-made claims* from large industrial structures. In 2010 the earth quake in Chile and the sinking of the drilling rig "Deepwater Horizon" were large claims, in 2011 the event in Fukushima, which combined natural catastrophe with man-made disaster, and the hurricane Sandy was a major catastrophe in 2012. It is common practice that an insurance company insures itself against large claims by a contract with a

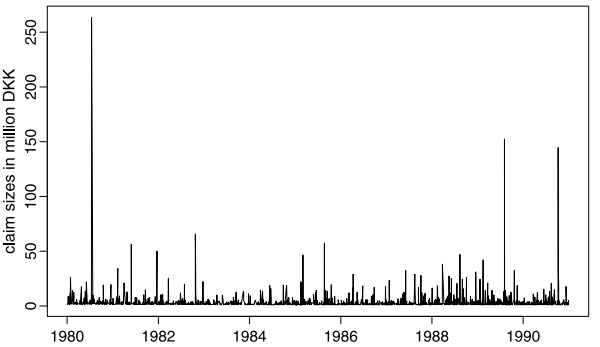

Fig. 4 Claim sizes of a Danish fire insurance during 1980–1990 in million Danish Krone (DKK)

reinsurance company. To-date the hurricane Katrina in 2005 is the most expensive insurance claim in history with about 76.25 billion US-Dollar, followed by the earth quake and the tsunami in Japan by 35.7 billion US-Dollar, hurricane Sandy in 2012 with about 35 billion US-Dollar, hurricane Andrew in 1992 with about 26.1 billion US-Dollar and the terror attack to the World Trade Center in 2001 with about 24.3 billion US-Dollar (the data are going back to http://de.statista.com/).

It is a common feature of large claims that they happen rarely, and hence little data are available to allow for reliable statistical prediction. But obviously, an insurance company and, even more so, a reinsurance company has to prepare for extreme events. Certain quantities can help to assess the frequency and severity of large claims. In the following we denote by $X_1, X_2, \ldots$ the accumulated claims per year of an insurance or reinsurance company (X_k is the total claim amount in year k) and we assume that these yearly claim amounts are independently and identically distributed (shortly i.i.d.) with distribution function F. We further assume that $F(0) = 0$ (a claim can only be positive) and that $F(x) < 1$ for all $x \in \mathbb{R}$ (claims can be arbitrarily large, which has been proved over and over by reality). We denote by $\overline{F}(x) = 1 - F(x)$ for $x \geq 0$ the so-called *tail of* F. We want to determine now the distribution of the first year in the future, where the yearly total claim exceeds a fixed yearly reserve u for the first time. This year is determined by

$$Z(u) = \min\{k \in \mathbb{N} : X_k > u\}.$$

Setting

$$q := \mathbb{P}(X > u) = \overline{F}(u), \tag{2.4}$$

the random variable $Z(u)$ is geometrically distributed with parameter q, i.e. the probability that $Z(u)$ takes the value k is given by

$$\mathbb{P}\big(Z(u) = k\big) = (1-q)^{k-1} q \quad \text{for } k \in \mathbb{N}$$

(in $k-1$ years we experience no excess, but then in year k there is an excess). The *return period* is now the mean waiting time until a yearly total claim amount

exceeds the threshold u (denoted by $\mathbb{E}(Z(u))$), where $\mathbb{E}$ is the mathematical symbol for expectation or mean. The expectation is then

$$\mathbb{E}\big(Z(u)\big) = \sum_{k=1}^{\infty} k\mathbb{P}\big(Z(u)=k\big) = q\sum_{k=1}^{\infty} k(1-q)^{k-1}$$
$$= \frac{1}{q} = \frac{1}{\mathbb{P}(X>u)} = \frac{1}{\overline{F}(u)}. \tag{2.5}$$

This provides now a trick to estimate the expectation. The standard way to estimate the expectation is by the arithmetic mean (the sum of all observation values divided by the number of all observations). Note however that, in order to do this, one would need many years, where exceedances have happened. Since the events we are interested in are rare, this classical statistical method can not be applied simply by lack of data. However, estimation via the right hand side of (2.5) is also not straightforward: the problem has been shifted now to the estimation of the tail $\overline{F}(u)$. Also for this tail estimation only few data are available. However, we can now compensate the lack of data by using clever methods from extreme value theory. We will explain this in detail in Sects. 3 and 4.

But also the inverse problem is of great interest. The insurance company wants to calculate premiums and reserves such that a yearly total claim amount larger than u should happen with a probability 0.1 at most every 50 years, which means that $\mathbb{P}(Z(u) \le 50) \le 0.1$. Since

$$\mathbb{P}\big(Z(u) \le 50\big) = q\sum_{i=1}^{50}(1-q)^{i-1} = 1-(1-q)^{50},$$

we have $1-(1-q)^{50} = 0.1$. This implies that $q = 0.002105$. Hence the return period in this example is $1/q = 475$ years. For the calculation of premiums and reserves we need now also the threshold u, and this requires the estimation of the quantile of the distribution function F. With the definition of the p-quantile in (2.1) we conclude with (2.4) that $u = x_{1-q}$ holds. We come back to this in Sect. 4.

3 Basic Extreme Value Theory

In the following we present the most important concepts for realistic modeling and quantification of rare events. The precise mathematical background as well as many application examples can be found in Beirlant et al. [1], Coles [3], Embrechts, Klüppelberg, and Mikosch [4], McNeil, Frey, and Embrechts [7], Reiss and Thomas [9], Stephenson [43] gives an excellent overview on extreme events in climatology.

Figure 3 presents a rather typical figure in many statistical applications areas. The normal distribution is often wrongly applied to extreme risk problems. This can only

be explained by the fact that everybody with a basic statistical education has learnt about the normal distribution. Moreover, the sum of normally distributed random variables is again normally distributed, and the mean and the standard deviation of this sum are easy to calculate.

There is no doubt that the normal distribution is a very important distribution in probability theory and statistics: it is the limit distribution for sums. For a sequence of i.i.d. random variables $X_1, X_2, \ldots$ (under the weak condition of a finite variance), we have

$$\frac{1}{\sqrt{n}} \sum_{k=1}^{n} \big(X_k - \mathbb{E}(X_k)\big) \xrightarrow{d} \mathcal{N}\big(0, \sigma^2\big) \quad \text{as } n \to \infty,$$

where the random variable on the right hand side is normally distributed with distribution function as in (2.2). The symbol $\xrightarrow{d}$ stands for convergence in distribution; i.e. the distribution functions of the random variables on the left hand side converge to the normal distribution function with mean 0 and variance σ^2. This is the so-called *Central Limit Theorem*. Because of this very basic result the normal distribution is an excellent model for random variables, which can be approximated by a sum of many small random effects. The great German mathematician Carl Friedrich Gauß (1777–1855) has derived it in his book [32].

It has long been known that the normal distribution is unrealistic for risk considerations. But which model is a good model for extreme events? The answer to this question has been given by the great French mathematician Siméon [40] (1781–1840), which we formulate nowadays as follows.

Theorem 3.1 (Poisson Theorem, [40]) *A statistical experiment with possible outcome E_n is repeated independently n times. The probability that the event E_n happens in one of the n trials is $\mathbb{P}(E_n) = p_n$. If $\lim_{n\to\infty} np_n = \tau$ holds for some $0 < \tau < \infty$, then*

$$\lim_{n\to\infty} \mathbb{P}\,(\textit{in exactly } m \textit{ of the } n \textit{ trials we have outcome } E_n)$$
$$= \lim_{n\to\infty} \binom{n}{m} p_n^m (1-p_n)^{n-m} = \mathrm{e}^{-\tau} \frac{\tau^m}{m!} \quad \textit{for } m = 0, 1, 2, \ldots, \tag{3.1}$$

where $\binom{n}{m} = \frac{n!}{m!(n-m)!}$ with $0! = 1$ and $m! = 1 \cdot 2 \cdots m$.

In honor of Poisson, the distribution on the right hand side of (3.1) is called Poisson distribution with parameter τ, abbreviated by $\mathrm{Poi}(\tau)$. The distribution on the left hand side of (3.1) (before the limit is taken) is the binomial distribution $\mathrm{Bin}(n, p_n)$, which for large n and small p_n approximates the Poisson distribution (cf. Fig. 5). Note that $\lim_{n\to\infty} np_n = \tau > 0$ implies obviously that $\lim_{n\to\infty} p_n = 0$. Hence the events E_n happen with vanishing probability, when the number of trials n is getting large. For this reason the Poisson distribution is also called the distribution of rare events. We want to present some ideas concerning the applicability of the

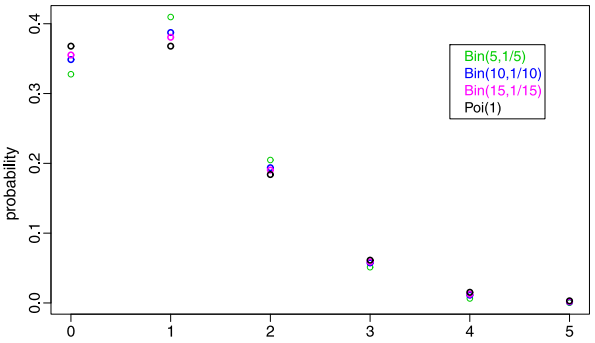

Fig. 5 Counting density of Bin(5, 1/5)-, Bin(10, 1/10)-, Bin(15, 1/15)-distribution and the Poi(1)-distribution. Note that for all parameters of the binomial distributions presented $np = 1$ holds

Poisson distribution, which leads to the two essential statistical concepts of extreme value theory. The first statistical method is called the *blocks method*, and the second one the *Peaks-Over-Thresholds (POT) method*. Which method to use depends on the question posed and on the data at hand. We will come back to both statistical methods in Sect. 4.

In the following we present the necessary mathematical results to understand the concepts. Let $X_1, \ldots, X_n$ be a sample of random variables; think for instance of yearly total claim amounts of an insurance company or losses of a financial asset. We assume that $X_1, \ldots, X_n$ are i.i.d. having the same distribution function as the random variable X; we denote it again by $F(x) = \mathbb{P}(X \leq x)$ for $x \in \mathbb{R}$.

We show first how to use the Poisson Theorem 3.1 for the description of the behavior of the maximum of a sample and investigate in a first step the so-called *partial maxima*

$$M_n = \max(X_1, \ldots, X_n) \quad \text{for } n \in \mathbb{N}.$$

As in real life we assume that risks larger than any we have observed before can continue to occur. This is formulated mathematically by investigating $\mathbb{P}(M_n \leq u_n)$, where the sequence u_n increases with n (and hence with M_n). Then the following fundamental result holds (which one can prove by means of the Poisson Theorem 3.1):

$$\lim_{n\to\infty} n\mathbb{P}(X_1 > u_n) = \tau \quad \Longleftrightarrow \quad \lim_{n\to\infty} \mathbb{P}(M_n \leq u_n) = \mathrm{e}^{-\tau}. \tag{3.2}$$

We want to motivate the implication from the left side to the right side:

Consider a rare event E, for example the event that the loss of a financial asset at a day is larger than a threshold u for large u. The daily losses of an asset constitute again a sample $X_1, \ldots, X_n$. Then

$$p = \mathbb{P}(E) = \mathbb{P}(X > u).$$

Invoking the same argument as Poisson, we find that the probability that the event E within the sample occurs m times is given by

$$\binom{n}{m} p^m (1-p)^{n-m} \quad \text{for } m = 0, \ldots, n;$$

i.e. it is Bin(n, p)-distributed. Now we let u depend on n in the sense that u_n increases with the sample size n. Then p becomes p_n, which converges to 0, and E becomes $E_n = \{X > u_n\}$. When u_n is chosen such that

$$\lim_{n\to\infty} np_n = \lim_{n\to\infty} n\mathbb{P}(X > u_n) = \tau \in (0, \infty),$$

then the Poisson Theorem 3.1 implies

$$\lim_{n\to\infty} \binom{n}{m} p_n^m (1-p_n)^{n-m} = \mathrm{e}^{-\tau} \frac{\tau^m}{m!} \quad \text{for } m = 0, 1, 2, \ldots.$$

In particular,

$$\begin{aligned}\lim_{n\to\infty} \mathbb{P}(M_n \le u_n) &= \lim_{n\to\infty} \mathbb{P}(E_n \text{ never occurs in the } n \text{ trials}) \\ &= \lim_{n\to\infty} \binom{n}{0} p_n^0 (1-p_n)^n = \mathrm{e}^{-\tau}.\end{aligned}$$

Consequently, we have shown how by the Poisson Theorem 3.1 the right hand side follows from the left hand side of (3.2). We shall resist to prove the reverse here.

The following result by [31] dating back to 1928 complements the above result; it describes precisely the possible limit distributions of partial maxima and provides the relevant tools for the estimation of tails and quantiles. For extreme value theory the Theorem of Fisher and Tippett is of equal fundamental importance as the Central Limit Theorem. The English statistician Ronald A. Fisher (1890–1962) has been one of the creators of modern statistics, working in many diverse areas.

Theorem 3.2 (Fisher-Tippett Theorem, [31]) *Let $X_1, X_2, \ldots$ be i.i.d. random variables, and $a_n > 0$ and $b_n \in \mathbb{R}$ appropriate constants. Moreover we assume that*

$$\lim_{n\to\infty} \mathbb{P}\big(\max(X_1, \ldots, X_n) \le a_n x + b_n\big) = G(x) \quad \textit{for } x \in \mathbb{R} \tag{3.3}$$

holds for a distribution function G. Then G belongs to the class $\{G_{\gamma,\sigma,\mu} : \gamma, \mu \in \mathbb{R}, \sigma > 0\}$, where

$$G_{\gamma,\sigma,\mu}(x) = \left\{ \begin{array}{ll} \mathrm{e}^{-(1+\gamma\frac{x-\mu}{\sigma})^{-\frac{1}{\gamma}}}, & \textit{if } \gamma \in \mathbb{R}\backslash\{0\}, \\ \mathrm{e}^{-\mathrm{e}^{-\frac{x-\mu}{\sigma}}}, & \textit{if } \gamma = 0, \end{array} \right\} \quad \textit{for} \quad \left\{ \begin{array}{ll} 1+\gamma\frac{x-\mu}{\sigma} > 0, & \textit{if } \gamma \neq 0, \\ x \in \mathbb{R}, & \textit{if } \gamma = 0. \end{array} \right.$$

The class of distributions $\{G_{\gamma,\sigma,\mu} : \gamma, \mu \in \mathbb{R}, \sigma > 0\}$ is called *generalized extreme value distribution (GEV)*. We recall that the *support* of a distribution function is the set of all $x \in \mathbb{R}$, where $0 < F(x) < 1$. Since $G_{\gamma,\sigma,\mu}(x) = G_{\gamma,1,0}(\frac{x-\mu}{\sigma})$, μ is called *location parameter* and σ is called *scale parameter*. The parameter γ is known as *shape parameter* and defines the type of distribution: if $\gamma > 0$ the distribution $G_{\gamma,\sigma,\mu}$ is a Fréchet distribution with support on $[\mu - \sigma/\gamma, \infty)$; if $\gamma = 0$ the distribution $G_{0,\sigma,\mu}$ is a Gumbel distribution with support on $\mathbb{R}$; if $\gamma < 0$ the distribution is a Weibull distribution with support on $(-\infty, \mu - \sigma/\gamma]$. The Fisher-Tippett

Theorem 3.2 thus states that the limit distribution of maxima are necessarily generalized extreme value distributions (and the normal distribution does obviously not belong to this class).

We want to explain the modelling and statistical consequences of the Fisher-Tippett theorem leading to the so-called *blocks method*. Recall the classical central limit theorem, which ensures that the distributions of sums and means of random variables converge to a normal distribution (for i.i.d. and even weakly dependent variables under the assumption of a finite variance). This motivates the modelling of random variables, which can be regarded as sums or means of random quantities by a normal distribution. Similarly, random variables which represent extreme quantities can be modelled by an extreme value distribution; Sect. 4.1 discusses the typical example of yearly maxima. Underlying this example the measurements consist of daily temperature values, and the maximum over every year is considered. So an extreme value distribution is an appropriate model for these yearly maxima. Moreover, the assumption of independence between the different maxima is also realistic as the time between two of such maxima is several months. We will discuss in Sect. 4.1, if the assumption of those maxima being identically distributed is realistic.

Under the conditions of the Fisher-Tippett Theorem 3.2 much more holds. We denote the class $\{H_{\gamma,\sigma} : \gamma \in \mathbb{R}, \sigma > 0\}$ of distribution functions *Generalized Pareto Distribution functions (GPD)*, which are defined as

$$H_{\gamma,\sigma}(x) = \begin{cases} 1 - (1 + \gamma \frac{x}{\sigma})^{-\frac{1}{\gamma}}, & \text{if } \gamma \in \mathbb{R}\backslash\{0\} \\ 1 - \mathrm{e}^{-\frac{x}{\sigma}}, & \text{if } \gamma = 0 \end{cases} \quad \text{for} \quad \begin{cases} x \geq 0, & \text{if } \gamma \geq 0, \\ 0 \leq x < -\sigma/\gamma, & \text{if } \gamma < 0. \end{cases}$$

Again γ denotes the shape parameter and σ the scale parameter. Indeed the parameter γ here is the same as in the Fisher and Tippett Theorem 3.2. Then the following theorem holds, which was proved independently by Pickands [39] and by Balkema and de Haan [13].

Theorem 3.3 (Pickands-Balkema-de Haan Theorem) *Assume that the conditions of the Fisher-Tippett Theorem 3.2 hold and that F is the distribution function of X. Then there exists a function $\sigma : (0, \infty) \to (0, \infty)$ and some $\gamma \in \mathbb{R}$ such that*

$$\lim_{u \to \infty} \mathbb{P}\big(X > u + \sigma(u)x \mid X > u\big) = \lim_{u \to \infty} \frac{\overline{F}(u + \sigma(u)x)}{\overline{F}(u)} = \overline{H}_{\gamma,1}(x)$$

for x in the support of $H_{\gamma,1}$.

It is now important for the *Peaks-Over-Threshold (POT)* method that for a large threshold u the following approximation holds by Theorem 3.3, where we set $y = \sigma(u)x$ and use that $\overline{H}_{\gamma,1}(y/\sigma(u)) = \overline{H}_{\gamma,\sigma(u)}(y)$:

$$\mathbb{P}(X > u + y \mid X > u) = \frac{\overline{F}(u + y)}{\overline{F}(u)} \approx \overline{H}_{\gamma,\sigma(u)}(y) \quad \text{for } y \geq 0. \tag{3.4}$$

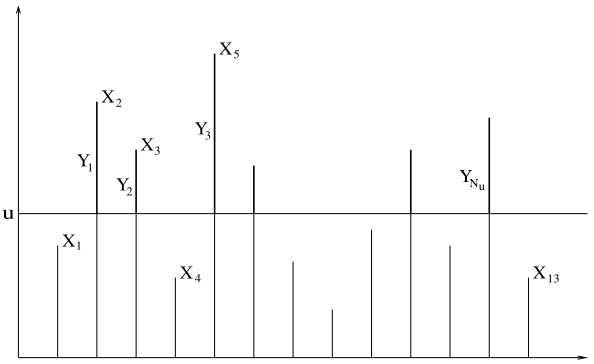

Fig. 6 Data $X_1, \ldots, X_{13}$ with corresponding excesses $Y_1, \ldots, Y_{N_u}$

Note first that an observation larger than $u + y$ is only possible, if the observation is larger than u; this means one needs a so-called *exceedance* of u. Such an observation has then necessarily a so-called *excess* over the threshold u, which is larger than y; cf. Fig. 6. If we investigate the special case that X has distribution $H_{\gamma,\sigma}$, we already have after some calculations that

$$\mathbb{P}(X > u + y \mid X > u) = \frac{\overline{H}_{\gamma,\sigma}(u+y)}{\overline{H}_{\gamma,\sigma}(u)} = \overline{H}_{\gamma,\sigma+\gamma u}(u) \tag{3.5}$$

and $\sigma(u) = \sigma + \gamma u$.

Let now $X_1, X_2, \ldots$ (as illustrated in Fig. 6) be i.i.d. with distribution $H_{\gamma,\sigma}$, then (3.5) means that $Y_1, Y_2, \ldots$, the exceedances of u, namely, $(X - u \mid X > u)$, are $H_{\gamma,\sigma+\gamma u}$ distributed. In the case $\gamma = 0$, where $H_{0,\sigma}$ is the exponential distribution with parameter σ^{-1}, $Y_1, Y_2, \ldots$ are again exponentially distributed with parameter σ^{-1}. This phenomena is well known as *loss-of-memory* property. In the general context of Theorem 3.3 with $X_1, X_2, \ldots$ i.i.d. with distribution function F, (3.4) says that $Y_1, Y_2, \ldots$ are asymptotically generalized Pareto distributed.

In contrast to the Fisher-Tippett Theorem 3.2, which models extreme observations directly, the Pickands-Balkema-de Haan Theorem 3.3 models all large values of a sample, more precisely, all those which exceed a high threshold. This is, where the acronym "Peaks-Over-Thresholds" (POT) originates. Compared to the modelling of yearly extremes (the so-called blocks method) the POT method has a positive and a negative property: on the one hand, taking all exceedances of a sample usually gives more observations, on the other hand, such exceedances can occur in clusters, so that the independence property can be violated. We will apply the POT method in Sect. 4.3.

4 Fundamental Results from Extreme Value Statistics

The books of Beirlant et al. [1], Coles [3], McNeil, Frey, and Embrechts [7], Reiss and Thomas [9] mentioned at the beginning of Sect. 3 provide also their own

software package for analyzing extremal events. An extensive overview on quite a number of R-packages and other extreme statistics software is given in [11]; cf. http://www.ral.ucar.edu/~ericg/softextreme.php and http://www.isse.ucar.edu/extremevalues/extreme.html. In particular, we want to mention the Extremes Toolkit (extRemes) developed in R by Eric Gilleland, which provides a user friendly graphical interface.

4.1 Fitting the GEV to a Sample of Extreme Data (the Blocks Method)

The GEV family can be applied as any other parametric family of distributions, whenever the model is justified by the data. Consequently, the GEV has been used for a sample of i.i.d. random variables, which result from some experiment and justify such a model.

Assume we have given yearly maxima $Y_1, \ldots, Y_n$, which can be assumed to be i.i.d. GEV distributed with distribution function $G_{\gamma,\sigma,\mu}$ and density $g_{\gamma,\sigma,\mu}$ with realizations $y_1, \ldots, y_n$. This means that data are block maxima and every year is a block. Then the maximum likelihood estimator of the parameters is given as

$$(\widehat{\gamma}, \widehat{\sigma}, \widehat{\mu}) = \operatorname*{argmin}_{\gamma,\sigma,\mu} \prod_{t=1}^{n} g_{\gamma,\sigma,\mu}(y_t). \tag{4.1}$$

We will use and slightly extend this concept to assess a possible trend in the location or scale parameter of the data over time.

The next example is classic in this respect: we will fit a GEV to a sample of yearly temperature maxima.

Illustration 4.1 (Climate Risk) Hot days are one of the prominent climatological phenomenon changing. According to IPCC 2007 (cf. [6]), it is very likely that warmer and more frequent hot days over most land areas have occurred in the late 20th century, a human contribution to this trend is likely, and it is—following their likelihood classification—virtually certain that this trend will continue for the 21th century. Daily maximum temperatures for example influence the well-being of humans putting additional stress to the thermal regulation and thus the cardiovascular system. Temperature maxima are very closely linked to average summer temperatures, each degree of warming increasing the maximum temperatures by 1.2 °C in Basel (Switzerland); see Beniston and Diaz [17]. Other projected impacts of more hot days comprise decreasing agricultural and forest yields in warmer environments, reduced energy demand for heating, increased demand for cooling, or declining air quality in cities.

We study long-term changes in daily maximum temperatures recorded at the oldest mountain climate station in the world, the observatory Hohenpeißenberg (977 m above sealevel, south-west of Munich), where regular meteorological observations

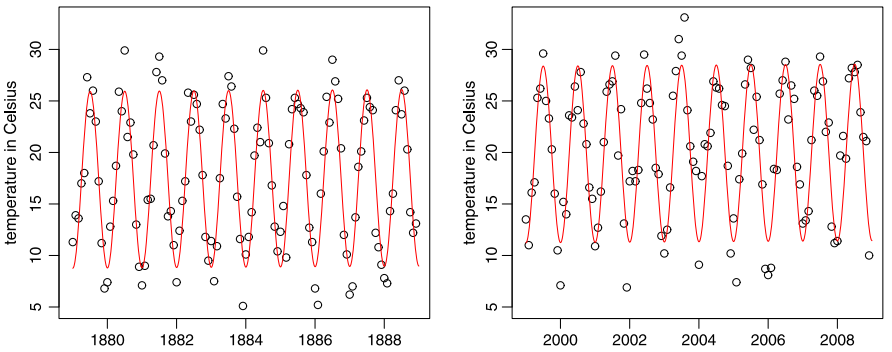

Fig. 7 Two decades of monthly temperature maxima: 1879–1888 and 1999–2008. The *red line* shows the estimated seasonality and trend

started beginning of 1781. We restrict our analysis to the period of 1879–2008, because in 1879 observations started being measured with new instruments under the guidance of the Munich Meteorological Central Station and thus the time series is homogenous. Due to its location on top of a mountain, summer temperatures are 2 °C to 3 °C lower than in the surrounding lowlands, whereas winter inversion layers lead to higher temperatures than in the valleys. The absolute maximum so far was recorded on July 29th in 1947 with 33.8 °C. Figure 7 displays the first (1879–1888) and last decade (1999–2008) of monthly temperature maxima.

It is one of the most demanding problems in environmental statistics to deal with trend and seasonality in data. When we are interested in the development of extreme events, we have to specify the event we want to study. In environmental statistics a usual measure of extremes is the return period as defined in (2.5). We could investigate the return periods of extremes in each month, January to December. Then we could answer, for instance, whether extreme temperatures in winter or summer have changed. Alternatively we could investigate the difference to a long-term mean or some other quantities, which describe extreme events.

In the present paper we will concentrate on a possible long-term trend in high temperatures at the station Hohenpeißenberg. Consequently, our analysis will be based on yearly maxima (see Fig. 8), which we assume to be GEV distributed (in Fig. 9 we shall see that this assumption is justified). Recall, however, that based on the IPCC 2007 report a 130 year temperature time series cannot be regarded as stationary. Thus, we want to incorporate some time-dependence into our model, i.e. a linear warming trend, although we know that there was not a uniform increase in mean temperature, but two periods with particular warming during approximately 1900–1945 and 1975–today.

We will investigate two possibilities to introduce non-stationarity into the model. Recall that classical time series theory (e.g. Brockwell and Davis [2]) suggests for a time series $Y_1, Y_2, \ldots$ either an arithmetic model of the form $Y_t = \Lambda_t + X_t$ or a multiplicative model $Y_t = \Lambda_t X_t$ for $t = 1, 2, \ldots$, where $\Lambda_1, \Lambda_2, \ldots$ models a non-stationary deterministic effect like drift and seasonality, and $X_1, X_2, \ldots$ is a stationary process. If $X_1, X_2, \ldots$ are identically GEV distributed, then we see immediately

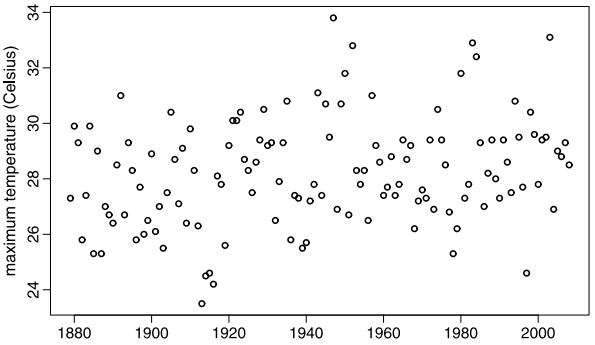

Fig. 8 Maximum yearly temperature over 130 years of data. The highest temperature has been measured in 1947

that $\Lambda_1, \Lambda_2, \ldots$ affect either the location parameter μ (for the arithmetic model) or the scaling parameter σ (for the multiplicative model) of the GEV distribution of $Y_1, Y_2, \ldots$. But the shape parameter γ remains the same under these deterministic location and scale changes. For simplicity, we introduce a linear trend into the location and scale parameter of the yearly maximal temperatures; i.e. we assume that the yearly maximal temperature $Y_1, \ldots, Y_{130}$ are an independent sequence with

$$Y_t \sim G_{\gamma,\sigma(t),\mu(t)} \quad \text{for } t = 1, \ldots, 130,$$

where $\mu(t) = \mu + at$ and $\sigma(t) = \sigma + bt$. Consequently, we will estimate by maximum likelihood estimation and compare the following models:

(1) Model 1: $\mu(t) = \mu$ and $\sigma(t) = \sigma$,
(2) Model 2: $\mu(t) = \mu + at$ and $\sigma(t) = \sigma$,
(3) Model 3: $\mu(t) = \mu$ and $\sigma(t) = \sigma + bt$,
(4) Model 4: $\mu(t) = \mu + at$ and $\sigma(t) = \sigma + bt$.

The estimation results are presented in Table 1.

For a comparison of the four different models, we notice that the negative log-likelihoods indicate already that Models 2 and 4 are better than Models 1 and 3, respectively. Although Model 1 is a special case of Model 3, the likelihood of Model 1 is nearly the same as the likelihood of Model 3. We guess already that the trend in the scale parameter may not be statistically significant, which is indeed true; the fluctuations do not significantly change over time. We have applied *likelihood ratio tests* to all nested pairs of models. Our model pairs are nested, when some of our parameters (a or b) may be zero or not. For details we refer to Coles [3], Sect. 2.6.6.

The tests compare (as we have already done informally) the likelihoods of two models. Rejection is now determined by asymptotic theory. More precisely, assume two (nested) models, say (*I*) and (*II*), with parameter $\theta^{(1)} \in \mathbb{R}^{d-k}$ for $k < d$, in model (*I*) and $\theta^{(2)} = (\theta_1^{(2)}, \theta_2^{(2)}) \in \mathbb{R}^d$ (where $\theta_1^{(2)} \in \mathbb{R}^k, \theta_2^{(2)} \in \mathbb{R}^{d-k}$) in model (*II*) with maximum likelihood estimators $\hat{\theta}^{(1)}$ and $\hat{\theta}^{(2)}$. Then, under some regularity conditions for the maximum likelihood functions $L_1(\hat{\theta}^{(1)})$ and $L_2(\hat{\theta}^{(2)})$, it can be shown that the quantity $-2(\log L_1(\hat{\theta}^{(1)}) - \log L_2(\hat{\theta}^{(2)}))$ is asymptotically χ_k^2-distributed. We present the results of 3 of our tests:

Table 1 Maximum likelihood estimators for μ, a, σ, b, and γ with standard errors in brackets below. The negative log-likelihood corresponding to the estimated models is given in the right-hand column

Parameters	μ	a	σ	b	γ	$-\log L$
Model 1	27.49671	–	1.84122	–	−0.20125	268.9776
	(0.17721)		(0.12203)		(0.05070)	
Model 2	26.65174	0.01320	1.76802	–	−0.19624	264.3865
	(0.32672)	(0.00426)	(0.11814)		(0.05253)	
Model 3	27.21659	–	1.70919	0.00199	−0.18065	268.9581
	(0.18851)		(0.23720)	(0.00377)	(0.06075)	
Model 4	26.65110	0.01321	1.77117	−0.00005	−0.19605	264.3863
	(0.32730)	(0.00426)	(0.22692)	(0.00301)	(0.05332)	

- Model 1 against Model 2: $H_0 : a = 0$ versus $H_1 : a \neq 0$

$$-2\big(\log L_1\big(\widehat{\mu}^{(1)}, \widehat{\sigma}^{(1)}, \widehat{\gamma}^{(1)}\big) - \log L_2\big(\widehat{\mu}^{(2)}, \widehat{a}^{(2)}, \widehat{\sigma}^{(2)}, \widehat{\gamma}^{(2)}\big)\big)$$
$$= 9.1823 > 3.8415 = \chi_1^2(0.95),$$

  i.e. we reject H_0 (p-value $= 0.002444$).
- Model 1 against Model 4: $H_0 : a = b = 0$ versus $H_1 : a \neq 0$ or $b \neq 0$

$$-2\big(\log L_1\big(\widehat{\mu}^{(1)}, \widehat{\sigma}^{(1)}, \widehat{\gamma}^{(1)}\big) - \log L_4\big(\widehat{\mu}^{(4)}, \widehat{a}^{(4)}, \widehat{\sigma}^{(4)}, \widehat{b}^{(4)}, \widehat{\gamma}^{(4)}\big)\big)$$
$$= 9.1826 > 5.9915 = \chi_2^2(0.95),$$

  i.e. we reject H_0 (p-value $= 0.0104$).
- Model 2 against Model 4: $H_0 : b = 0$ versus $H_1 : b \neq 0$

$$-2\big(\log L_2\big(\widehat{\mu}^{(2)}, \widehat{a}^{(2)}, \widehat{\sigma}^{(2)}, \widehat{\gamma}^{(2)}\big) - \log L_4\big(\widehat{\mu}^{(4)}, \widehat{a}^{(4)}, \widehat{\sigma}^{(4)}, \widehat{b}^{(4)}, \widehat{\gamma}^{(4)}\big)\big)$$
$$= 3 \times 10^{-4} < 3.8415 = \chi_1^2(0.95),$$

  i.e. we do not reject H_0 (p-value $= 0.986983$).

The p-value is an indicator of significance: the p-value of 0.002444 as calculated in the first test ensures that we can reject H_0 for all significance levels larger than this value. So the smaller the p-value, the more justified is a rejection of H_0. The comparison shows that a trend in the location parameter of the GEV model is significant but not the trend in the scale parameter. Model 4 gives no improvement to Model 2. Hence, again with support by statistical theory we conclude that the best model is Model 2, and there is no significant difference between Models 2 and 4, justifying the choice for Model 2.

In order to assess the model fit graphically, we will use a Gumbel probability plot (based on the GEV $G_{0,1,0}$) (PP-plot) and a Gumbel quantile-quantile plot (QQ-plot) for our transformed data set. Therefore, we show that any $G_{\gamma,\sigma(t),\mu(t)}$ distributed random variable Y_t with $\gamma < 0$ (the relevant regime for the temperature example is

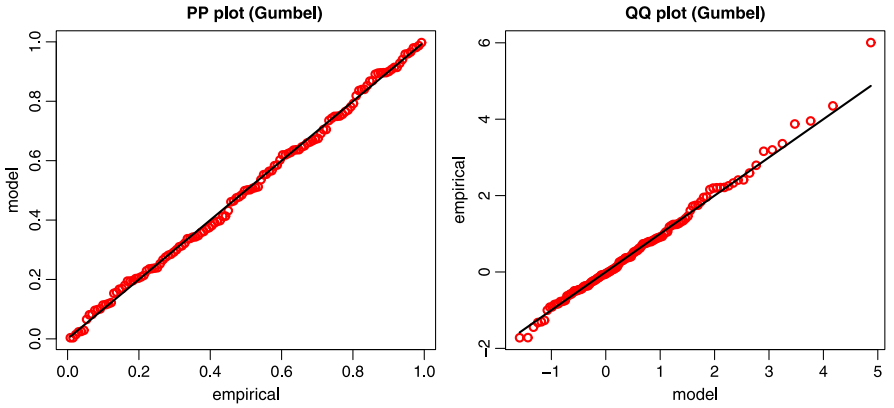

Fig. 9 The linear location Model 2 transformed to standard Gumbel: *PP*-plot and *QQ*-plot

a Weibull GEV distribution) can be transformed to a Gumbel random variable as follows. Afterwards we can use standard software for the plots.

We define

$$Z_t = \frac{1}{\gamma} \ln\left(1 + \gamma \frac{(Y_t - \mu(t))}{\sigma(t)}\right),$$

and prove below that indeed Z_t is standard Gumbel distributed. Note first that the Gumbel distribution has support on the whole of $\mathbb{R}$, whereas the Weibull distribution $G_{\gamma,\sigma(t),\mu(t)}$ has support $(-\infty, \mu(t) - \sigma(t)/\gamma]$; i.e. $G_{\gamma,\sigma(t),\mu(t)}(x) = 1$ for all $x > \mu(t) - \sigma(t)/\gamma$. Then $1 + \gamma(Y_t - \mu(t))/\sigma(t) > 0$ and, hence, Z_t has full support $\mathbb{R}$. Now we calculate

$$\begin{aligned}\mathbb{P}(Z_t \le x) &= \mathbb{P}\left(\frac{1}{\gamma} \ln\left(1 + \gamma \frac{(Y_t - \mu(t))}{\sigma(t)}\right) \le x\right) \\ &= \mathbb{P}\left(Y_t \le \frac{\sigma(t)}{\gamma}\left(\mathrm{e}^{\gamma x} - 1\right) + \mu(t)\right) = \mathrm{e}^{-\mathrm{e}^{-x}} \quad \text{for } x \in \mathbb{R}.\end{aligned}$$

This means that, provided $Y_1, Y_2, \ldots$ are independent Weibull distributed random variables, then $Z_1, Z_2, \ldots$ are independent Gumbel distributed random variables. Consequently, once we have estimated $\mu(t)$, $\sigma(t)$ and γ, we transform our data Y_t to

$$\widehat{Z}_t := \frac{1}{\widehat{\gamma}} \ln\left(1 + \widehat{\gamma} \frac{(Y_t - \widehat{\mu}(t))}{\widehat{\sigma}(t)}\right),$$

which should be close to a Gumbel distribution, provided the data are indeed Weibull GEV distributed with the estimated parameters. Figure 9 assesses the distribution fit by a *PP*-plot and a *QQ*-plot for the estimated parameters of Model 2 with linear location parameter. In the first plot, the *PP*-plot, the empirical distribution of $\widehat{Z}_1, \ldots, \widehat{Z}_{130}$ is plotted against the Gumbel distribution. In the second plot, the

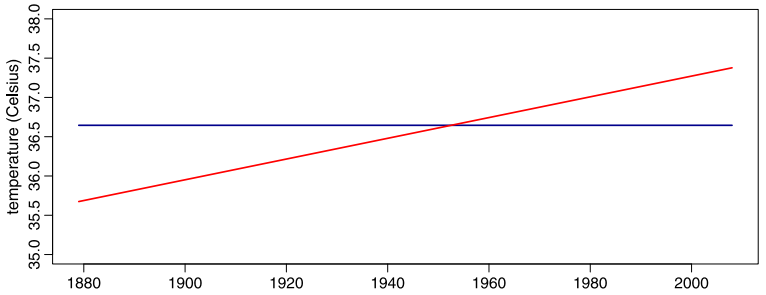

Fig. 10 Estimated right endpoint of the GEV distribution of Model 1 (*blue line*) and the linear trend Model 2 (*red line*). For Model 1 we estimate the constant right endpoint of 36.645 °C. For Model 2 the right endpoint starts at 35.674 °C and ends at 37.377 °C

QQ-plot, the quantiles of the Gumbel distribution are plotted against the empirical quantiles of $\widehat{Z}_1, \ldots, \widehat{Z}_{130}$. Both look very convincing, since they follow a 45° line confirming again Model 2.

For Model 2 we estimated the asymptotic 95 % confidence interval for γ. Let $z_{1-\alpha/2}$ be the $1-\alpha/2$-quantile of the normal distribution and $\widehat{s}_\gamma$ be the estimated standard deviation of $\widehat{\gamma}$. Then by classical likelihood theory (see Smith [10]), at least for $\gamma < 1/2$,

$$(\hat{\gamma} - z_{1-\alpha/2}\,\widehat{s}_\gamma, \hat{\gamma} + z_{1-\alpha/2}\,\widehat{s}_\gamma)$$

denotes the asymptotic $(1-\alpha) \times 100$ % confidence interval for γ. In Model 2 this results in the 95 % confidence interval $(-0.29972, -0.06158)$ for γ. As mentioned after the Fisher-Tippet Theorem 3.2 a negative γ indicates a Weibull distribution with finite right endpoint, meaning that there should be a limit of extreme maximum temperatures, which is not exceeded. Similarly, we obtain for a the 95 % confidence interval

$$(\widehat{a} - z_{1-\alpha/2}\,\widehat{s}_a, \widehat{a} + z_{1-\alpha/2}\,\widehat{s}_a) = (0.0048504, 0.0215496),$$

which reflects that a is positive; we have a statistically significant increase in the location parameter and a trend in the extremal temperatures.

The right endpoint of the Weibull distribution is given by $\mu(t) - \sigma(t)/\gamma$ (representing the maximum yearly temperature), which we can also estimate after having estimated the parameters. Figure 10 visualizes the constant endpoints of Model 1, where we have assumed fixed parameters over the whole time period, and the increase of the endpoint for the linear trend Model 2 caused by the linearity in the location parameter.

From this analysis presented in Fig. 11 we see that the return levels of high temperatures have increased considerably over the last 130 years. This increase is due to an increase of the location parameter of the extreme temperatures, the levels of the return periods have increased. The estimated parameters suggest an increase of $at = 0.01320t = 0.01320 \times 130 = 1.716$ °C over 130 years, corresponding to an

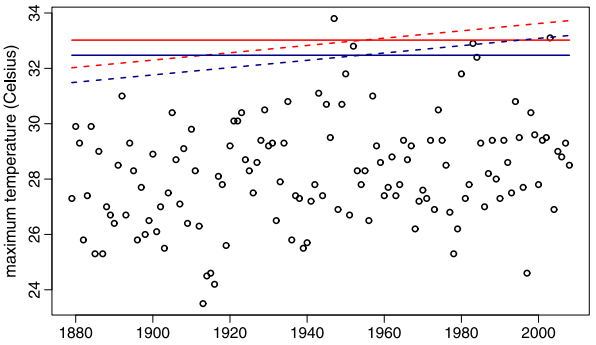

Fig. 11 The *red lines* show the estimated 100-year return level (which is the 99 % quantile), where the *straight line* is based on Model 1 and the *dashed line* on Model 2. Similarly the *blue lines* show the estimated 50-year return level based on Model 1 and Model 2, respectively

increase of 1.32 °C over a century. In contrast, simple least square linear regressions reveal increases in daily mean temperature of 1.472 °C and in daily maximum temperature of 1.515 °C over 130 years at the climate station Hohenpeißenberg, corresponding to an increase of 1.13 °C for the mean and of 1.17 °C for the daily maximum temperature over a century. Compared to these naive estimators the more realistic assessment by EVT methods yields a considerably higher prediction for the daily maximum temperatures in the future.

Prediction could now be based on this analysis. If we believe that the linear trend remains the same over the next 10 years, then we would estimate the value of $37.377 + 0.132 = 37.5090$ for the maximal yearly temperature in 2018. Note however, that such a fixed number is very unlikely. A confidence interval would be needed to give some idea about the statistical variability. By our estimation method we have been able to calculate confidence intervals for every single parameter estimate. However, for a confidence interval of the prediction we would need the whole distribution, which involves all three parameters, and their estimates are dependent. So besides standard errors (based on the estimated variance of the maximum likelihood estimators) also the asymptotic correlations between parameter estimates enter. Such theory, however, goes beyond this introductory paper, and gives rather food for thought.

Apart from this statistical discussion, there is also some doubt on the assumption that future maximum temperatures increase with the same linear drift as the past ones. This also depends on political measures being taken against the threatening climate change.

4.2 The Blocks Method from Scratch

In the previous section we have simply started with maximum yearly temperatures over 130 years, and fitted an extreme value distribution to these data. This model choice was first based on the Fisher-Tippet Theorem 3.2, and later justified by a *PP*-plot and a *QQ*-plot depicted in Fig. 9.

As the name *blocks method* suggests the idea behind it is to divide the data $X_1, X_2, \ldots, X_{nm}$ into m blocks of roughly the same length n and consider the block

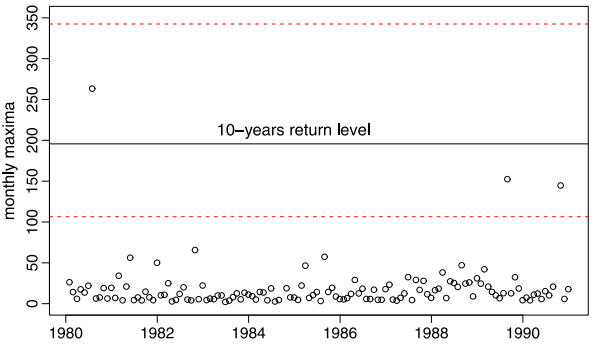

Fig. 12 The largest claims of a Danish fire insurance per month. The *black line* is the 10-year return level, and the *dashed red lines* indicate the 95 %-confidence interval

maxima, i.e. we define $M_{n,j} = \max(X_{(j-1)n+1}, \ldots, X_{jn})$ for $j = 1, \ldots, m$. Recall that on the one hand we want to choose the blocks so small that we get as many block maxima as possible, on the other hand we have to choose them large enough so that we can assume that block maxima follow an extreme value distribution and also that they are independent.

Illustration 4.2 (10-Year Return Period for Danish Fire Data) For the daily losses of the fire insurance portfolio over m months $X_1, X_2, \ldots, X_{nm}$ (i.e., X_k is the loss at the kth day), we determine the maximum losses within a month, respectively. These monthly block sizes are roughly equal, more precisely, n is between 28 and 31 days, and $M_{n,j}$ is the maximum loss during the jth month. As a first ansatz, according to the Fisher-Tippett Theorem 3.2, we exploit the fact that the distribution of $M_{n,j}$ can be approximated by a GEV distribution, so that

$$\mathbb{P}(M_{n,j} \leq u) \approx G_{\gamma,\sigma,\mu}(u),$$

where γ, σ, μ are parameters, which have to be estimated, and the constants a_n and b_n are integrated in σ and μ. We denote by $\widehat{\gamma}, \widehat{\sigma}, \widehat{\mu}$ the respective estimators. Then we approximate

$$\mathbb{P}(M_{n,j} \leq u) \approx G_{\widehat{\gamma},\widehat{\sigma},\widehat{\mu}}(u).$$

The level of the 10-year return period of the largest monthly claim, which happens in mean only once in 10 years can be estimated by means of (2.5). Since also $q = 1/(10 \times 12)$ holds, we obtain

$$\widehat{u} = \widehat{x}_{1-q} = G^{-1}_{\widehat{\gamma},\widehat{\sigma},\widehat{\mu}}\big(1 - (10 \times 12)^{-1}\big). \tag{4.2}$$

For the Danish fire data as depicted in Fig. 4 we estimate 195.7 million Danish Krone as level for extreme monthly claims, which happen in mean every 10 years (see Fig. 12).

4.3 The POT Method

It has been argued that applying the blocks method to data has the drawback of disregarding data, which may contribute information to the statistics of extreme values. Moreover, the blocks method can easily be applied to yearly, monthly or to other blocks-structured data, but what to do, if this is not the case. The *Peaks-Over-Threshold* (POT) method presents a valuable alternative.

The following section is dedicated to the POT method for a sample $X_1, \ldots, X_n$, where we assume for the distribution function F that $F(x) = \mathbb{P}(X \leq x) < 1$ for $x > 0$. We define further for a high threshold u

$$\overline{F}_u(y) := \mathbb{P}(X - u > y \mid X > u) = \frac{\overline{F}(u+y)}{\overline{F}(u)} \quad \text{for } y \geq 0.$$

Consequently, we obtain

$$\overline{F}(u+y) = \overline{F}(u)\overline{F}_u(y) \quad \text{for } y \geq 0. \tag{4.3}$$

How can we use these identities now to estimate tails and quantiles?

If now N_u denotes the number of all $k \in \{1, \ldots, n\}$ satisfying $X_k > u$ given by

$$N_u = \#\left\{k \in \{1, \ldots, n\} : X_k > u\right\},$$

then we denote by $Y_1, \ldots, Y_{N_u}$ the excesses of $X_1, \ldots, X_n$, i.e. the heights of the exceedances of u (cf. Fig. 6). We obtain an estimator for the tail (for values larger than u) by estimating both tails on the right hand side of (4.3). We estimate $\overline{F}(u)$ by the relative frequency

$$\widehat{\overline{F}(u)} = \frac{N_u}{n} \tag{4.4}$$

and approximate $\overline{F}_u(y)$ by the Generalized Pareto Distribution (GPD) of (3.4), where the scale parameter $\sigma(u)$ has to be considered. It is integrated as parameter $\sigma(u)$ into the limit distribution such that

$$\overline{F}_u(y) \approx \left(1 + \gamma \frac{y}{\sigma(u)}\right)^{-1/\gamma} \quad \text{for } y \geq 0, \tag{4.5}$$

where γ and $\sigma(u)$ have to be estimated by some estimators denoted by $\widehat{\gamma}$ and $\widehat{\sigma}(u)$. From (4.3)–(4.5) we obtain a tail estimator of the form

$$\widehat{\overline{F}(u+y)} = \frac{N_u}{n}\left(1 + \widehat{\gamma}\frac{y}{\widehat{\sigma}(u)}\right)^{-1/\widehat{\gamma}} \quad \text{for } y \geq 0. \tag{4.6}$$

Then for given $p \in (0, 1)$ we obtain an estimator $\widehat{x}_p$ for the p-quantile x_p taken from (2.1) by solving the equation

$$1 - p = \frac{N_u}{n}\left(1 + \widehat{\gamma}\frac{\widehat{x}_p - u}{\widehat{\sigma}(u)}\right)^{-1/\widehat{\gamma}}.$$

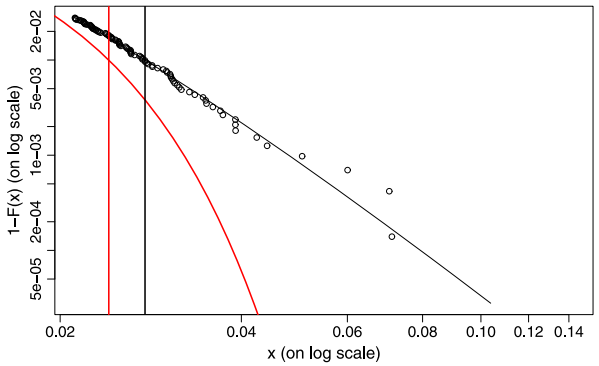

Fig. 13 Estimated tail of the daily losses of the S&P500. The *black curve* shows the tail estimated by the POT method with threshold $u = 0.0212$, $\widehat{\gamma} = 0.193$, $\widehat{\sigma} = 0.00575$ and the *red line* shows the distribution tail estimated under the assumption of a normal distribution for the daily losses. The *vertical black line* indicates the logarithmic $\mathrm{VaR}_{0.99}^{\mathrm{POT}}(X) = 0.028$ estimated by the POT method and the *vertical red line* shows the logarithmic $\mathrm{VaR}_{0.99}^{\mathrm{norm}}(X) = 0.024$ estimated from a normal distribution

This gives

$$\widehat{x}_p = u + \frac{\widehat{\sigma}(u)}{\widehat{\gamma}}\left(\left(\frac{n}{N_u}(1-p)\right)^{-\widehat{\gamma}} - 1\right). \tag{4.7}$$

Illustration 4.3 (Tail and Quantile Estimation) We apply the POT method to the S&P500 loss data using the tail estimate from (4.6) and, for comparison, we also fitted a normal distribution to the data by estimating mean and variance by their empirical versions. Figure 13 depicts both tail estimates in logarithmic scale for a threshold $u = 0.0212$ and $y > 1$. Moreover, $\mathrm{VaR}_{0.99}^{\mathrm{POT}}(X)$ was estimated for the daily losses using the POT estimator (4.7) as well as the normal estimator $\mathrm{VaR}_{0.99}^{\mathrm{norm}}(X) = \widehat{\mu} + \widehat{\sigma} z_{0.99}$, where $z_{0.99}$ is the 0.99-quantile of the normal distribution. Plotted are again the logarithmic quantities; i.e. $\log \mathrm{VaR}_{0.99}^{\mathrm{POT}}(X) = 0.028$ and $\log \mathrm{VaR}_{0.99}^{\mathrm{norm}}(X) = 0.024$, which correspond to $\mathrm{VaR}_{0.99}^{\mathrm{POT}}(X) = 2.795$ and $\mathrm{VaR}_{0.99}^{\mathrm{norm}}(X) = 2.784$; the difference of 0.011 does not look too substantial, but recall that our data are relative losses (i.e. percentage points). Moreover, the standardized S&P500 portfolio value compares only to a standardized bank portfolio, so has to be multiplied by millions to obtain a realistic value.

We clearly see that the normal distribution tail is completely inadequate to estimate the tail of the daily losses of the S&P500. The data are far above its normal tail estimate. Usage of the normal distribution underestimates the risk considerably and yields a completely inadequate risk capital.

In Illustration 4.3 we have estimated the tail and the $\mathrm{VaR}_{0.99}(X)$ for the S&P500 losses and depicted in Fig. 13. The estimation was based on the assumption that the losses (or at least the excesses) are i.i.d. However, modelling of financial data goes

Fig. 14 The empirical standard deviations of the daily losses of the S&P500 during 1991–2004 with estimators based on the previous 250 days, respectively

far beyond marginal distributions. It has been a relevant research area for decades, and we conclude with some facts and references.

Remark 4.4 (i) Dependence between portfolio components are in the normal model given by correlations, which only model linear dependence. Market risk portfolios, however, consist of such different assets as shares, options, and more complex derivatives, which are known to be non-linearly dependent. It is of high importance to have a comprehensive understanding of the influence of the portfolio components to the portfolio loss. Dependence modeling and different dependence measures are discussed in Chap. 9, [37].

(ii) Already from the daily losses depicted in the right plot of Fig. 2 it is clear that the data vary considerably in their structure. We see immediately that a period of low volatility is followed by a period of high volatility (the standard deviation is called volatility in banking jargon). It is certainly not obvious that all observations can be modelled with the same distribution. Recall that (2.3) requires daily estimates based on past year's observations. Figure 14 shows the running empirical estimates of the volatility σ of the daily losses of the S&P500 based on observations of the past one year, respectively. This simple window estimate shows clearly the time-varying volatility, which is typical for most financial time series.

(iii) Until now we have not touched the important questions of time dependence within the time series of daily returns. Financial data show an interesting dependence structure; although most daily returns are uncorrelated, the data do not originate from independent observations. As seen in Fig. 15 the sample autocorrelation function of the daily losses of the S&P500 is almost 0 for all lags, whereas the sample autocorrelation function of the squared returns is substantial, contradicting the independence assumption. The most prominent financial time series model is the GARCH (Generalized AutoRegressive Conditional Heteroskedasticity) model. Volatility is modelled as a stochastic process and can capture a dependence structure as seen in Fig. 15. An excellent overview on discrete-time and continuous-time

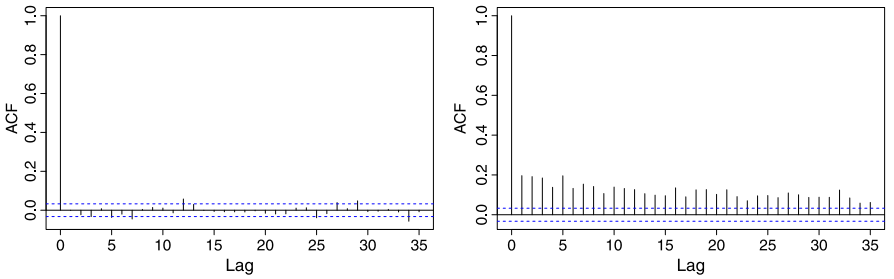

Fig. 15 Sample autocorrelation function of the daily losses (*left*) and the squared daily losses (*right*) of the S&P500

stochastic volatility models in one and multivariate dimensions is the book edited by [12].

5 Food for Thought

Extreme value theory has gone a long way since its beginnings with Fisher and Tippett in 1928. New applications, unheard-of in the 1920-ies have emerged. Climate change, large insurance claims and extreme financial risk are just three of them. Extreme value theory has found its way also into the areas of technical safety and reliability theory, as well as the statistical assessment of environmental quantities like temperatures, floods, droughts, earthquakes and storms.

Concerning the statistical methods we have presented, we want to emphasize the following. In our statistical analyses we have assumed that data are independent and have the same distribution (perhaps enriched by a linear trend, which can easily be implemented). This assumption is often unrealistic. As reported in Remark 4.4 financial time series exhibit in general a very complex dependence structure; for the S&P500 see Figs. 14 and 15. Many data, also insurance claims, are affected by seasonal effects or exhibit some clusters of claim events. Such effects can influence estimation and prediction procedures considerably. Moreover, the one-dimensional case treated above is rather unrealistic. Portfolios of market risks are composed of many components (often several hundreds), and it may be interesting to understand the dependence in the combination of extreme risks. Moreover, risks are often influenced by some latent variables, whose influence would have to be assessed as well. Such problems are hot research topics at the moment and require still a considerable amount of theoretical and practical work.

Extreme value theory has been extended to multivariate data, which is rather demanding, since there exists no finite parameterizations as in the one-dimensional case as seen in the Fisher-Tippett Theorem 3.2. The dependence between different components of a vector is modeled by an integral with respect to some measure and Poisson random measures provide a very powerful tool to deal with such problems; cf. Resnick [41].

Moreover, extreme value theory for time series with marginals ranging from Gaussian to heavy-tailed ones is still a lively research area. The usual picture is that for light-tailed time series models one can more or less ignore the dependence structure, whereas for heavy-tailed models the dependence creeps into the extreme tails (events) by leading to clusters of extremes; cf. Fasen [28], Fasen, Klüppelberg, and Schlather [30].

More recently, also spatial and space-time extreme value models have come into focus in particular for environmental data like heavy rainfall or storms requiring special statistical methods; cf. Davis, Klüppelberg, and Steinkohl [25, 26] for details and further references.

For those interested in the state of the art of extreme value theory research, we recommend to consider the journal "Extremes" (http://www.springer.com/statistics/journal/10687), which is solely devoted to theory and applications of extreme values.

6 Summary

We hope that we have convinced our readers that extreme value theory and extreme value statistics offer an important theory and statistical estimation procedures to assess extreme risks in different applications areas.

We have presented the basic theory and also three estimation procedures to find the distribution and other quantities describing extreme events. The first one was to fit a GEV to extreme data, where we also took care of non-stationarity of the data either in the location parameter (linear trend) or in the scaling parameter (higher fluctuations). The second one was to use the block-maxima method for a sample where only the blocks maxima were distributed according to a GEV distribution. And finally, we introduced the POT method, which models high threshold exceedances.

As a result we obtained for our three examples:

- The climate change data exhibit a higher trend in the yearly maxima over the last century than the mean trend at the corresponding station. The Weibull distribution is the appropriate extreme value distribution, which shows that high temperature is bounded, although the maxima increase.
- Danish insurance claims, which are from a fire insurance portfolio, are very heavy-tailed data, and the model suggests that with a (non-negligible) positive probability the insurance company may experience a claim, which is easily twice as high as they have ever seen before.
- The daily losses of the S&P500 have a 99 % Value-at-Risk of 0.028 % when estimated by the POT method, while based on the normal distribution, it is only 0.024 %. While these numbers look small, in banking business one has to multiply them by millions of Euros, so that the difference becomes substantial. Since capital reserves have to be calculated built on such numbers, the banks are much happier about the smaller numbers coming from the Gaussian distribution.

References

Selected Bibliography

1. J. Beirlant, Y. Goegebeur, J. Segers, J. Teugels, D. De Waal, C. Ferro, *Statistics of Extremes: Theory and Applications* (Wiley, New York, 2004)
2. P. Brockwell, R.A. Davis, *Introduction to Time Series and Forecasting* (Springer, New York, 1996)
3. S.G. Coles, *An Introduction to Statistical Modeling of Extreme Values* (Springer, Berlin, 2001)
4. P. Embrechts, C. Klüppelberg, T. Mikosch, *Modelling Extremal Events for Insurance and Finance* (Springer, Berlin, 1997)
5. G.C. Hegerl, H. Hanlon, C. Beierkuhnlein, Elusive extremes. Nat. Geosci. **4**, 143–144 (2011)
6. IPCC 2007, Climate change 2007: the physical science basis, in *Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, ed. by S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor, H.L. Miller (Cambridge University Press, Cambridge, 2007)
7. A.J. McNeil, R. Frey, P. Embrechts, *Quantitative Risk Management: Concepts, Techniques, and Tools* (Princeton University Press, Princeton, 2005)
8. A. Menzel, H. Seifert, N. Estrella, Effects of recent warm and cold spells on European plant phenology. Int. J. Biometeorol. **55**, 921–932 (2011)
9. R.-D. Reiss, M. Thomas, *Statistical Analysis of Extreme Values with Applications to Insurance, Finance, Hydrology and Other Fields*, 2nd edn. (Birkhäuser, Basel, 2001)
10. R.L. Smith, Estimating tails of probability distributions. Ann. Stat. **15**, 1174–1207 (1987)
11. A. Stephenson, E. Gilleland, Software for the analysis of extreme events: the current state and future directions. Extremes **8**, 87–109 (2005)

Additional Literature

12. T.G. Andersen, R.A. Davis, J.-P. Kreiss, T. Mikosch, *Handbook of Financial Time Series* (Springer, New York, 2009)
13. A.A. Balkema, L. de Haan, Residual life time at great age. Ann. Probab. **2**, 792–804 (1974)
14. K.F. Bannör, M. Scherer, Model risk and uncertainty—illustrated with examples from mathematical finance, in *Risk – A Multidisciplinary Introduction*, ed. by C. Klüppelberg, D. Straub, I. Welpe (2014)
15. Basel Committee on Banking Supervision, Basel III: a global regulatory framework for more resilient banks and banking systems (2011). Revised June 2011
16. Basel Committee on Banking Supervision, Revisions to the Basel II market risk framework (2011). Updated as of 31 December 2010
17. M. Beniston, H.F. Diaz, The 2003 heat wave as an example of summers in a greenhouse climate? Observations and climate model simulations for Basel, Switzerland. Glob. Planet. Change **44**, 73–81 (2004)
18. C. Bernhardt, C. Klüppelberg, T. Meyer-Brandis, Estimating high quantiles for electricity prices by stable linear models. J. Energy Mark. **1**, 3–19 (2008)
19. F. Biagini, T. Meyer-Brandis, G. Svindland, The mathematical concept of measuring risk, in *Risk – A Multidisciplinary Introduction*, ed. by C. Klüppelberg, D. Straub, I. Welpe (2014)
20. K. Böcker, C. Klüppelberg, Operational VaR: a closed-form approximation. Risk **12**, 90–93 (2005)
21. E. Brodin, C. Klüppelberg, Extreme value theory in finance, in *Encyclopedia of Quantitative Risk Assessment*, ed. by B. Everitt, E. Melnick (Wiley, Chichester, 2007)

22. S.J. Brown, J. Caesar, C.A.T. Ferro, Global change in extreme daily temperature since 1950. J. Geophys. Res. **113**, D05115 (2008)
23. R. Cont, Empirical properties of asset returns: stylized facts and statistical issues. Quant. Finance **1**, 223–236 (2001)
24. J. Danielsson, P. Embrechts, C. Goodhart, C. Keating, F. Muennich, O. Renault, H.S. Shin, An Academic Response to Basel II. Financial Markets Group, London School of Economics (2001)
25. R.A. Davis, C. Klüppelberg, C. Steinkohl, Max-stable processes for modelling extremes observed in space and time. J. Korean Stat. Soc. **42**(3), 399–414 (2013)
26. R.A. Davis, C. Klüppelberg, C. Steinkohl, Statistical inference for max-stable processes in space and time. J. R. Stat. Soc., Ser. B **75**(5), 791–819 (2013)
27. S. Emmer, C. Klüppelberg, M. Trüstedt, VaR – ein Mass für das extreme Risiko. Solutions **2**, 53–63 (1998)
28. V. Fasen, Extremes of continuous-time processes, in *Handbook of Financial Time Series*, ed. by T.G. Andersen, R.A. Davis, J.-P. Kreiss, T. Mikosch (Springer, Berlin, 2009), pp. 653–667
29. V. Fasen, C. Klüppelberg, Modellieren und Quantifizieren von extremen Risiken, in *Facettenreiche Mathematik*, ed. by K. Wendland, A. Werner (Vieweg/Teubner, Wiesbaden, 2011)
30. V. Fasen, C. Klüppelberg, M. Schlather, High-level dependence in time series models. Extremes **13**, 1–33 (2010)
31. R.A. Fisher, L.H.C. Tippett, Limiting forms of the frequency distribution of the largest or smallest member of a sample. Proc. Camb. Philos. Soc. **24**, 180–190 (1928)
32. C.F. Gauß, *Theoria Motus Corporum Coelestium in Sectionibus Conicis Solum Ambientium* (Hamburgi Sumptibus F. Perthes et I.H. Besser, Hamburg, 1809)
33. P. Glasserman, *Monte Carlo Methods in Financial Engineering* (Springer, New York, 2004)
34. S. Haug, C. Klüppelberg, L. Peng, Statistical models and methods for dependent insurance data; invited discussion paper. J. Korean Stat. Soc. **40**, 125–139 (2011)
35. IPCC 1990, Climate change: the IPCC scientific assessment, in *Report Prepared for IPCC by Working Group 1*, ed. by J.T. Houghton, G.J. Jenkins, J.J. Ephraums (Cambridge University Press, Cambridge, 1990)
36. R.W. Katz, Statistics of extremes in climate change. Clim. Change **71–76**, 100 (2010)
37. C. Klüppelberg, R. Stelzer, Dealing with dependent risks, in *Risk – A Multidisciplinary Introduction*, ed. by C. Klüppelberg, D. Straub, I. Welpe (2014)
38. R. Korn, E. Korn, G. Kroisandt, *Monte Carlo Methods and Models in Finance and Insurance*. Financial Mathematics Series (Chapman & Hall/CRC Press, London/Boca Raton, 2010)
39. J. Pickands, Statistical inference using extreme order statistics. Ann. Stat. **3**, 119–131 (1975)
40. S. Poisson, *Recherches sur la Probabilité des Jugements en Matière Criminelle et en Matière Civile Précédées des Règles Générales du Calcul des Probabilités* (Bachelier, Imprimeur-Libraire, Paris, 1837)
41. S.I. Resnick, *Extreme Values, Regular Variation, and Point Processes* (Springer, New York, 1987)
42. R.L. Smith, Extreme value analysis of environmental time series: an application to trend detection in ground-level ozone (with discussion). Stat. Sci. **4**, 367–393 (1989)
43. D.B. Stephenson, Definition, diagnosis, and origin of extreme weather and climate events, in *Climate Extremes and Society*, ed. by R. Murnane, H. Diaz (Cambridge University Press, Cambridge, 2008), pp. 11–23
44. G.-R. Walther, E. Post, P. Convey, A. Menzel, C. Parmesani, T.J. Beebee, J.-M. Fromentin, Ecological responses to recent climate change. Nature **416**, 389–395 (2002)