Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

Federal Department of Home Affairs FDHA
Federal Office of Meteorology and Climatology MeteoSwiss

Technical Report MeteoSwiss No. 255

MeteoSwiss extreme value analyses: User manual and documentation

Sophie Fukutome and Anne Schindler

ISSN: 2296-0058

Technical Report MeteoSwiss No. 255

MeteoSwiss extreme value analyses: User manual and documentation

Sophie Fukutome and Anne Schindler

Recommended citation:

Fukutome S, Schindler A : 2015, MeteoSwiss extreme value analyses:
User manual and documentation, *Technical Report MeteoSwiss*, 255, 50 pp.

Editor:

Federal Office of Meteorology and Climatology, MeteoSwiss, © 2015

MeteoSwiss
Operation Center 1
8058 Zürich-Flughafen
T +41 58 460 91 11
www.meteoswiss.ch

Abstract

Extreme value analyses of meteorological quantities, such as precipitation, or temperature, are of great importance in design engineering and risk management. At the end of 2013, MeteoSwiss has made extreme value analyses of precipitation publicly available for 27 stations of its observational network. The present report provides users with a manual for understanding and interpreting these analyses; it documents the data used and the statistical methods employed. It is meant for engineers and other experts familiar with basic statistical concepts. The report is structured in three parts. The first describes and explains the contents of the analysis sheets, the second documents the underlying data, and the third is dedicated to the theoretical background.

Zusammenfassung

Extremwertanalysen von meteorologischen Grössen, wie Niederschlag oder Temperatur, sind von grossem Interesse für Dimensionierung und Risikomanagement. MeteoSchweiz hat Ende 2013 Extremwertanalysen von Niederschlag, gemessen an 27 Stationen des meteorologischen Netzwerks der MeteoSchweiz, der Öffentlichkeit zur Verfügung gestellt. Dieser Bericht liefert eine "Bedienungsanleitung", um diese Analysen zu verstehen und zu intepretieren und dokumentiert die zugrundeliegenden Daten und Methoden. Er richtet sich an Ingenieure und andere Experten, die statistisches Grundlagenwissen besitzen. Der Bericht ist dreiteilig gegliedert: Das erste Kapitel beschreibt und erklärt die Inhalte der Analyse-Blätter. Das zweite Kapitel dokumentiert die zugrundeliegenden Daten und das dritte Kapitel befasst sich mit den nötigen theoretischen Grundlagen.

Contents

1 Introduction

Switzerland has witnessed a number of weather events causing devastating floods in the last decades, as for instance in August 2005, when a large part of the Swiss northern Alpine rim was affected. Such events raise questions regarding adequate planning of infrastructure based on reliable statistical estimates. The most recent official documents on the matter were compiled by the Federal Institute for Forest, Snow and Landscape Research (WSL) in the 70s and 80s, and are outdated, both in terms of precipitation data and of statistical methods.

As a federal office operating a broad network of meteorological observation stations, MeteoSwiss has a public responsibility to provide information on the severity of rare events. In a first step, a web page (http://www.meteoswiss.ch/home/klima/vergangenheit/extrem-niederschlaege.html) was set up on the website of MeteoSwiss at the end of 2013, presenting extreme value analyses of 1-day, 2-day, and 3-day precipitation at 27 stations of the MeteoSwiss observational network. On the instigation and with the support of the Federal Office for the Environment (FOEN), a more comprehensive platform is underway with analyses at more than 350 (60) stations for daily (subdaily) precipitation, in which state-of-the-art statistical methods will be applied.

The extreme value analyses made available in 2013 consist of a 2-page sheet in pdf-format. They are intended for engineers and other experts requiring quantitative information for design and risk management. The present report explains how to read and interpret these analyses, and documents the underlying data and methods. The methods employed here are long established, and climatological aspects are not addressed. Thus, the report must be seen as a "user manual" for a person familiar with basic statistical concepts, wishing to understand the contents of the analyses, and how they came into being. The exact mathematical expressions have been added for completeness in separate boxes, but can be ignored, as they are not essential for understanding the text.

The report is divided into three parts. The first provides a detailed description (chapter 2) of the contents of the analysis sheet. The second is dedicated to the data used (chapter 3), their measurement, eventual manipulation, and quality issues. Finally, the third part (chapter 4) describes the theoretical background: the relevant aspects of extreme value statistics, the assumptions that need to be made, and when these may be considered appropriate in this particular context.

This report will be updated as more sophisticated methods are introduced for the new web platform.

"Graphical" table of contents (pages 8 - 9):
Example of a 2-page extreme value analysis sheet (as appears on the web page).

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Eidgenössisches Departement des Innern EDI
Bundesamt für Meteorologie und Klimatologie MeteoSchweiz

Zürich / Fluntern: 556m, 47.38N, 8.57E

Extreme Value Analysis

1-day precipitation, 5:40-5:40 UTC

1961 - 2013 (number of missing years: 0)

Block Maxima (GEV). **Reliability of results: questionable.**

Short summary
Interpretation ..... p. 14

Plot of return levels and their uncertainty (ordinate) for a given return period (abscissa).

The best estimate of the return levels is colored blue. The return level 95% confidence intervals are in green.

The dots represent observations (in the ordinate) to which empirical return periods (abscissa) have been assigned. These empirical return periods depend solely on the sample size. Here, for instance, the largest observation is automatically assigned a return period of approx. 53 years.

Return level plot
Interpretation ..... p. 10
Definition .... p. 37

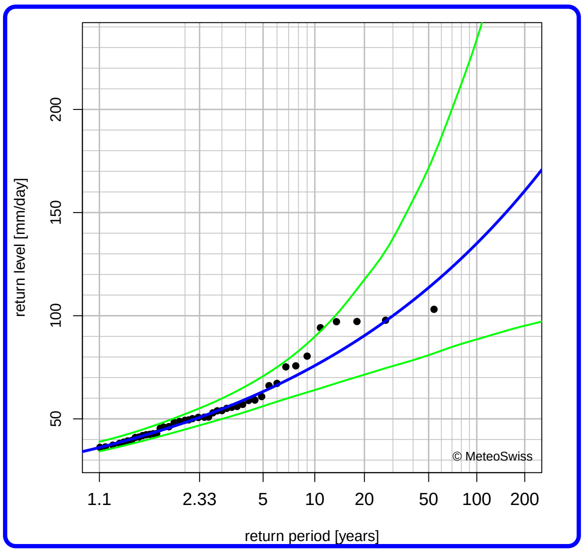


Table of the largest annual extrema in the period analysed.

If two large events occur in the same year, only the largest will appear in this table. The return periods are estimated with the fitted GEV distribution.

date	precipitation [mm/day]	estimated return period [years]
1968-09-21	103.1	34
2007-08-08	97.8	27
1978-08-07	97.2	27
1973-06-23	97.1	27
1999-05-12	94.2	24

Table of return levels for a selection of return periods.

The 95% confidence intervals are shown in brackets.

return period [years]	return value [mm/day]	confidence interval [mm/day]
2	48.2	(44.6 - 52.6)
5	63.2	(56.3 - 71.5)
10	75.8	(64.5 - 90.9)
20	90.4	(72.3 - 117.6)
30	100.0	(76.2 - 136.8)
50	113.6	(81.4 - 169.5)
100	135.1	(88.7 - 233.5)

Krähbühlstrasse 58, 8044, Zürich
rre150yx; xs_2.3.1; June 10, 2015

Contact: kundendienst@meteoswiss.ch

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Eidgenössisches Departement des Innern EDI
Bundesamt für Meteorologie und Klimatologie MeteoSchweiz

Distribution Function and estimation methods

- The Generalized Extreme Value distribution (GEV) is fitted to the yearly extrema.
- The distribution parameters are estimated with Maximum Likelihood.
- The confidence intervals are estimated with parametric resampling.

Method information
Documentation ... p. 27

Data and data quality

- The raw data is quality-checked, but not homogenized.
- Missing data: None.

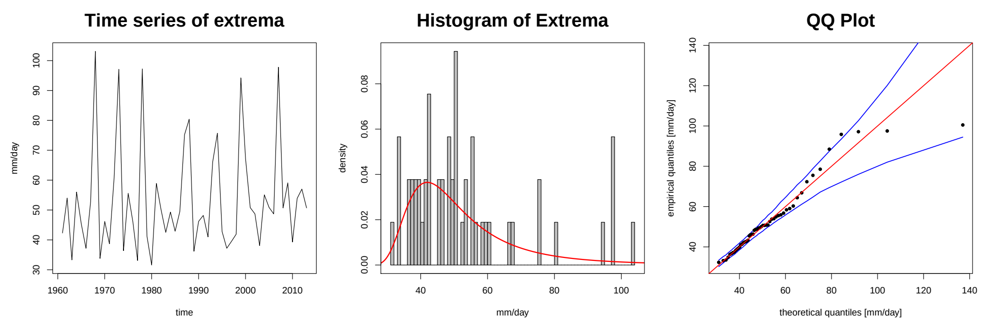
Data information
Interpretation ...... p. 15
Documentation ... p. 19

Distribution Parameters

- Location: 44.25
- Scale: 10.39
- Shape: 0.25

Distribution information
Interpretation ..... p. 15

Additional information



Left: **Time series of extrema**.

Middle: **Histogram of extrema**. Red line: fitted GEV density distribution.

Right: **QQ plot**. Plot of empirical vs. theoretical quantiles of the observations. The theoretical quantiles are estimated with the fitted GEV. Should the dots align on the diagonal (red line), the fit would be perfect. The QQ-plots of 1000 samples randomly drawn from the fitted GEV have a 95% chance of being within the blue lines.

Krähbühlstrasse 58, 8044, Zürich
rre150yx; xs_2.3.1; June 10, 2015
Contact: kundendienst@meteoswiss.ch

2 Interpretation

The station-based extreme value analysis is presented on two pages. The first page (shown on page 8 and again in figure 3 for another station, left) states briefly the data employed and the fit quality, and displays the return level plot, a table of return levels for a selection of return periods, and a table of return values for the five yearly maxima in the data series. The second page (shown on page 9 for Zurich and again in figure 3 for Geneva, right) contains more detailed information on the data, the distribution and the estimation.

The extreme value analysis presented on the webpage of MeteoSwiss is based on the generalized extreme value (GEV) distribution (see section 4), which is used here to describe the behaviour of the yearly maxima of daily precipitation.

2.1 Return level plots

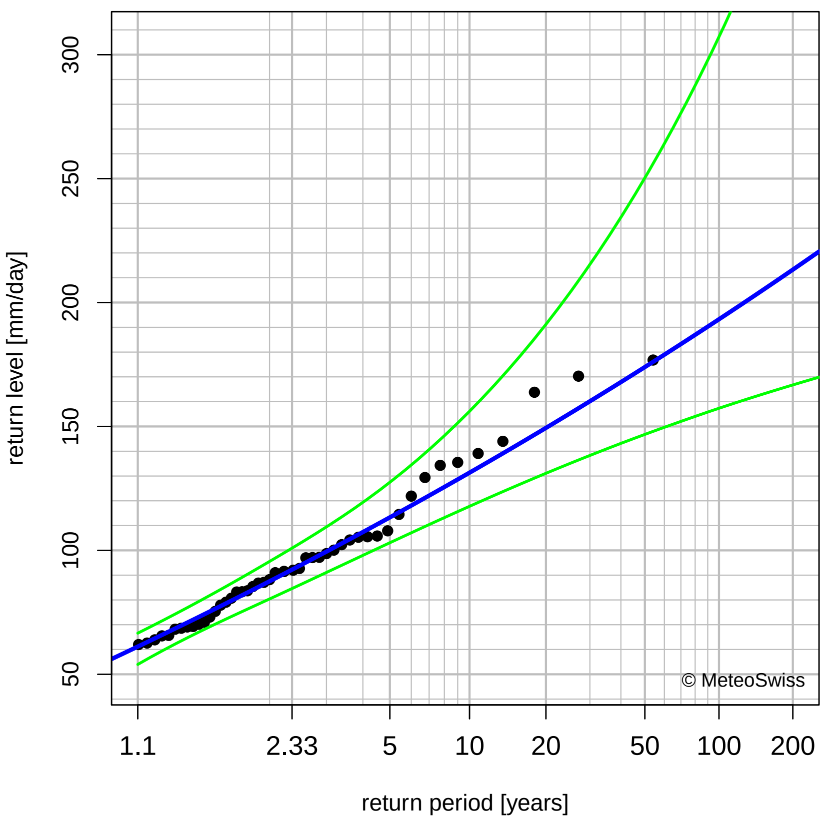


Figure 1: Return level plot as presented in the extreme value analysis information of 1-day precipitation at station Lugano. The blue line represents the return level estimates based on the estimated GEV-distribution; the green lines represent their 95% confidence interval; and the y-coordinate of the black points represent data used for estimation of the GEV-distribution; the plotting position (x-coordinate) of the points is solely determined by the length of the underlying record (and is not robust).

Return level plots (figure 1) display the probability that a given value is exceeded. The probability is expressed in terms of years. Thus a value that has a probability of 1% of being exceeded in any given year is - in the (very) long term average - expected to be exceeded once in a 100 years. Then, the value, or “return level”, is said to have a “return period” of 100 years. In fact, return levels are extreme quantiles: the 100-year return value has a probability of 0.99 of not being exceeded, and therefore corresponds to the customary 0.99-quantile (see figure 2).

To focus the attention on the behavior of the rarest events, return periods are transformed in such a way that the relation to the return levels will appear as a straight line for a Gumbel distribution (see section 4) of the yearly maxima.

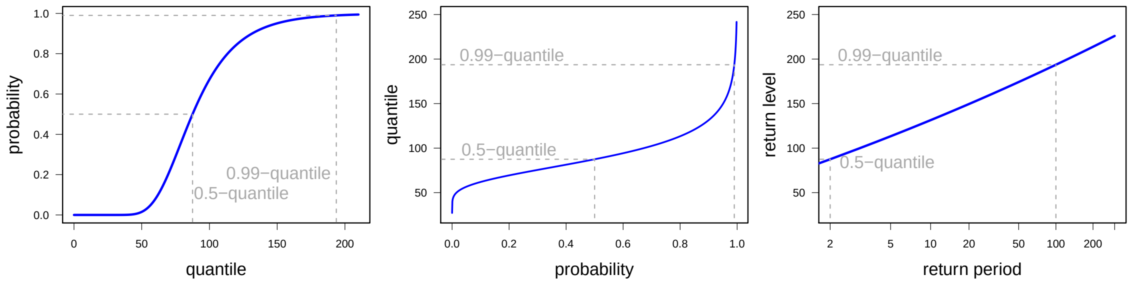


Figure 2: Correspondence between cumulative probability distribution function (CDF, left), quantile function (center) and return level plot (right). The distribution shown here was estimated for Lugano 1-day precipitation extremes (cf. figure 1).

In real life applications, the true probability distribution describing the behavior of extremes is unknown, and must be estimated from the observations. Due to the finite size of the sample, this estimation is inherently uncertain, but the uncertainty can be quantified, and translates to uncertainty of the estimated return levels (green lines, figure 1).

Remark. It is important to keep in mind that the plotted uncertainties only describe the uncertainty in return **level** estimates. No statement can be made regarding an uncertainty in return **period** based on the uncertainties represented in figure 1.

The return level plot often displays so-called “plotting points” (black points, figure 1). The position of these plotting points on the y-axis corresponds to the values of the observed sample maxima. Since the true return period of these events is unknown, however, the position on the x-axis is determined empirically from the sample size. Thus, for a sample size of 50 years, for instance, the largest observation will be assigned a return period of approximately 50 years, the second largest a return period of approximately 25 years, and so on. In other words, it does not represent the local behavior of the quantity of interest (e.g. daily precipitation) at the particular station under consideration, since the return period would be the same regardless of the geographical location of the station. Some conclusions can, however, be drawn from the comparison between plotting points and best estimate (blue line, figure 1): any substantial or systematic disagreement between empirical and GEV estimates, after allowance for sampling error, suggests an inadequacy of the GEV model.

For a more thorough analysis of the fit, it is preferable to examine the agreement between empirical and modeled estimates in a quantile-quantile plot, which is easier to interpret correctly (see section

4.3.2).

2.2 Reliability

Each extreme value analysis is tested for its reliability, and the verdict is stated on the analysis sheet. This verdict does not pertain to the inherent uncertainty of the estimation due to the limited sample size. Rather, it describes the possibility that the assumptions motivating the choice of the statistical model may not be fulfilled. If the reliability is poor, for instance, the model does not represent well the observed extreme values. Three degrees of reliability have been defined: poor, questionable, and good. Their meaning and what to do about them is shown in the following table.

Remark. "The only justification for extrapolating an extreme value model is the asymptotic basis on which it is derived." (*Coles*, 2001, page 3)

"However, if a model is found to perform badly in terms of its representation for extreme values that have already been observed, there is little hope of it working well in extrapolation." (*Coles*, 2001, page 3)

Verdict	Interpretation for the user	Action
poor	Observed extreme values are **badly** represented by the model.	Do not use statistical model! Use largest observed events or closest reliable station instead.
questionable	Observed values are **not well** represented by the model. Underlying assumptions might be violated.	Careful assessment necessary. Use visual guides (sec. 4.3.1) to decide: if points line up - proceed as usual; if not - treat fit as poor.
good	Observed extreme values are **well** represented.	The statistical model can be used.

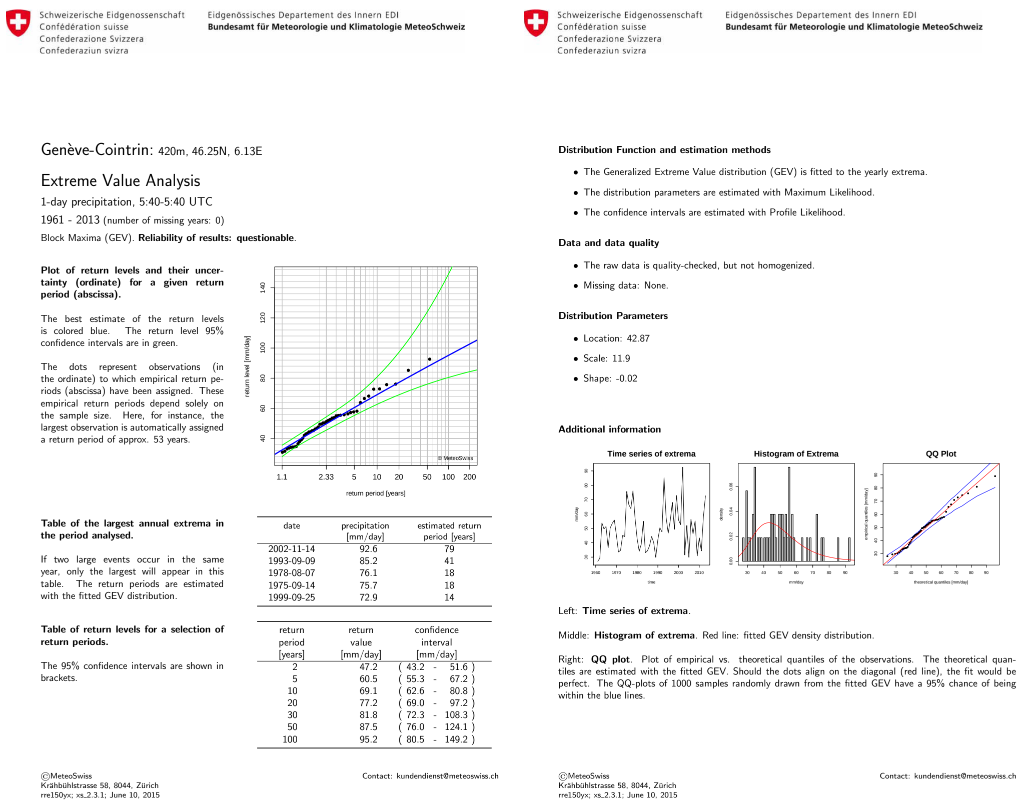
2.3 Example stations

In the following, we illustrate the interpretation of a typical extreme value analysis product for two stations, Genève-Cointrin (abbreviation GVE) and Zürich / Fluntern (abbreviation SMA), that are subject to different climatic conditions in terms of heavy precipitation, and therefore lend themselves well to highlight concepts and problems in interpreting extreme value analysis.

Geneva is located at the south-western tip of Switzerland, in a flat plain, between the Jura mountains and the foothills of the Alpine arch, while Zurich lies in the north-eastern part of the Swiss Plateau. Heavy precipitation is generally due to summer thunderstorms on the background of a weather configuration characterized by moist, south-westerly flow. Geneva is comparatively dry because of its protected position in the lee of the Jura mountains. Although it is not of direct relevance in the present context, it is interesting to note that the seasonal cycle of daily precipitation is very modest in Geneva, with the largest monthly maxima in September; in Zurich on the other hand, it is very pronounced and peaks in August.

The differences in precipitation regime between Genève-Cointrin and Zürich/Fluntern are manifest in the statistical characterization of their extremal behavior. Average heavy precipitation is of similar magnitude at both stations: we can expect events of 50mm per day to be exceeded on average every other year, with similar ranges of deviation from this average. Yet, they differ in the behavior of the very rare events, and thus illustrate two different types of extremal behavior, both typical for one-day precipitation extrema.

As it turns out, the estimates at both stations are of questionable reliability. As we shall see, however, the information generated by the statistical model can be used with some confidence in the case of GVE and not SMA.



Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Eidgenössisches Departement des Innern EDI
Bundesamt für Meteorologie und Klimatologie MeteoSchweiz

Genève-Cointrin: 420m, 46.25N, 6.13E

Extreme Value Analysis

1-day precipitation, 5:40-5:40 UTC

1961 - 2013 (number of missing years: 0)

Block Maxima (GEV). **Reliability of results: questionable.**

Plot of return levels and their uncertainty (ordinate) for a given return period (abscissa).

The best estimate of the return levels is colored blue. The return level 95% confidence intervals are in green.

The dots represent observations (in the ordinate) to which empirical return periods (abscissa) have been assigned. These empirical return periods depend solely on the sample size. Here, for instance, the largest observation is automatically assigned a return period of approx. 53 years.

Table of the largest annual extrema in the period analysed.

If two large events occur in the same year, only the largest will appear in this table. The return periods are estimated with the fitted GEV distribution.

date	precipitation [mm/day]	estimated return period [years]
2002-11-14	92.6	79
1993-09-09	85.2	41
1978-08-07	76.1	18
1975-09-14	75.7	18
1999-09-25	72.9	14

Table of return levels for a selection of return periods.

The 95% confidence intervals are shown in brackets.

return period [years]	return value [mm/day]	confidence interval [mm/day]
2	47.2	(43.2 - 51.6)
5	60.5	(55.3 - 67.2)
10	69.1	(62.6 - 80.8)
20	77.2	(69.0 - 97.2)
30	81.8	(72.3 - 108.3)
50	87.5	(76.0 - 124.1)
100	95.2	(80.5 - 149.2)

©MeteoSwiss
Krähbühlstrasse 58, 8044, Zürich
rre150yx; xs_2.3.1; June 10, 2015

Contact: kundendienst@meteoswiss.ch

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Eidgenössisches Departement des Innern EDI
Bundesamt für Meteorologie und Klimatologie MeteoSchweiz

Distribution Function and estimation methods

- The Generalized Extreme Value distribution (GEV) is fitted to the yearly extrema.
- The distribution parameters are estimated with Maximum Likelihood.
- The confidence intervals are estimated with Profile Likelihood.

Data and data quality

- The raw data is quality-checked, but not homogenized.
- Missing data: None.

Distribution Parameters

- Location: 42.87
- Scale: 11.9
- Shape: -0.02

Additional information

Left: **Time series of extrema**.

Middle: **Histogram of extrema**. Red line: fitted GEV density distribution.

Right: **QQ plot**. Plot of empirical vs. theoretical quantiles of the observations. The theoretical quantiles are estimated with the fitted GEV. Should the dots align on the diagonal (red line), the fit would be perfect. The QQ-plots of 1000 samples randomly drawn from the fitted GEV have a 95% chance of being within the blue lines.

©MeteoSwiss
Krähbühlstrasse 58, 8044, Zürich
rre150yx; xs_2.3.1; June 10, 2015

Contact: kundendienst@meteoswiss.ch

Figure 3: Station-based extreme value analysis information for 1-day precipitation at station Genève-Cointrin.

2.3.1 Genève-Cointrin

Figure 3 shows the extreme value analysis information of MeteoSwiss for 1-day precipitation recorded at the station Genève-Cointrin.

Short summary To define the product at hand, the first first few lines provide succinct information regarding the data and statistical method (figure 4):

- The station name, here **Genève-Cointrin**, its altitude (**420m**) and geographical coordinates (**46.25N, 6.13E**).

Genève-Cointrin: 420m, 46.25N, 6.13E
Extreme Value Analysis
1-day precipitation, 5:40-5:40 UTC
1961 - 2013 (number of missing years: 0)
Block Maxima (GEV). **Reliability of results: questionable.**

Figure 4: Short summary of extreme value analysis.

- The MeteoSwiss product name, here **extreme value analysis**.
- The variable analyzed, here **1-day-precipitation**, recorded daily at **5:40 UTC**, observed from **1961 to 2013 with no missing years** (see section 3).
- The extreme value approach employed: The **block maxima** approach, where annual maxima are assumed to follow a **generalized extreme value** distribution (see section 4).
- A summary of the reliability of the results based on the quality of the fit, here **questionable** (see section 4.3.2 and page 37).

Return level plot The return level plot summarizes the extremal behaviour at a glance (figure 5). A range of 1-day-precipitation amounts (y-axis) is plotted against their exceedance probabilities expressed in terms of average number of years (x-axis). The extremal behavior can be deduced from the return level plot (figure 5): The return level estimate (blue line) is a straight line. This means the probability that a large value will be exceeded decays exponentially. Thus, while any amount of daily precipitation has a non-zero probability of being exceeded, this probability decreases very rapidly as the amount increases. Keep in mind that the return level plot has been transformed in such a way that an exponential decay looks like a straight line.

The strong positive curvature of the upper confidence bound (green line, figure 5) reveals that this extremal behavior is subject to uncertainty: a slower (polynomial) decay of the exceedance probability is also quite possible. This means that very severe events would have a higher probability of being exceeded than in the exponential case. The lower confidence bound is slightly negatively curved - a bounded distribution is therefore possible but rather unlikely. Which would suggest that beyond a given daily precipitation amount, the probability that an amount should be exceeded is zero.

The green lines on the return level plot (figure 5) show the 95% confidence intervals. Note that these confidence intervals only describe the uncertainty of the return levels, expressed in mm.

The dots on the figure are called plotting points. They represent the observed yearly maxima (ordinate) to which have been assigned an empirical return period (abscissa). This empirical return period depends solely on sample size (here the number of yearly maxima), and is the same for all stations, regardless of their geographical location.

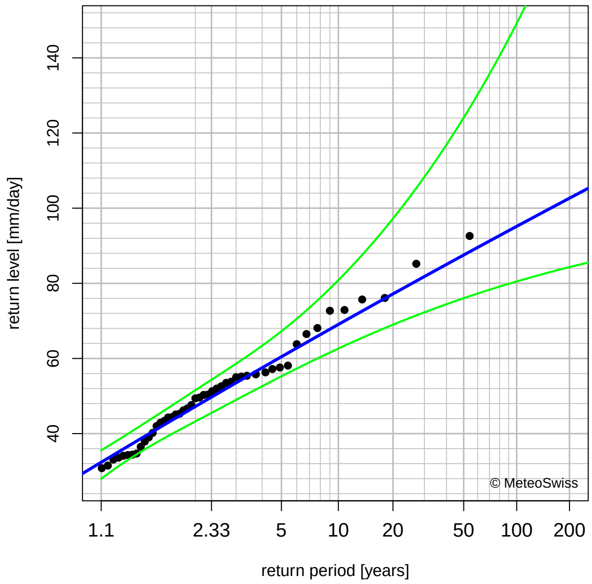


Figure 5: Return level plot as presented in the Extreme Value Analysis information of 1-day precipitation in Geneva. The blue line represents the return level estimates based on the estimated GEV-distribution; the green lines represent their 95$\%$ confidence interval; and the y-coordinate of the black points represent data used for estimation of the GEV-distribution; the plotting position (x-coordinate) of the points is solely determined by the length of the underlying record (and is not robust).

Tables The table of the largest annual extrema and the table of return levels contain the same information as the return level plot. The first (table 1, left) "reads" the return periods corresponding to the 5 largest maxima, the second (table 1, right) "reads" the precipitation amounts corresponding to given return periods. The relation between the two is dictated by the estimated GEV.

date	precipitation [mm/day]	estimated return period [years]
2002-11-14	92.6	79
1993-09-09	85.2	41
1978-08-07	76.1	18
1975-09-14	75.7	18
1999-09-25	72.9	14

return period [years]	return value [mm/day]	confidence interval [mm/day]
2	47.2	(43.2 - 51.6)
5	60.5	(55.3 - 67.2)
10	69.1	(62.6 - 80.8)
20	77.2	(69.0 - 97.2)
30	81.8	(72.3 - 108.3)
50	87.5	(76.0 - 124.1)
100	95.2	(80.5 - 149.2)

Table 1: Largest 1-day precipitation events with estimated return period (left) and return level estimates to a selection of return periods (right) as presented in the extreme value analysis information of 1-day precipitation at Genève.

Distribution information The estimated GEV distribution is provided via the three estimated parameters: the location parameter, here 42.87 mm; the scale parameter, here 11.9 mm; and the shape parameter, here -0.02. The corresponding probability density function (PDF) is shown in figure 6 as the red line of the right panel.

Data information The time series of annual maxima (figure 6, left) informs visually about missing data or possible trends or cycles. Here, there could be a decadal oscillation or a trend towards higher values at the end of the record. Further inquiries would, however, be necessary to rule out random effects. The histogram (figure 6, right) shows the frequency with which the annual maxima have been recorded, and can be understood as the empirical PDF.

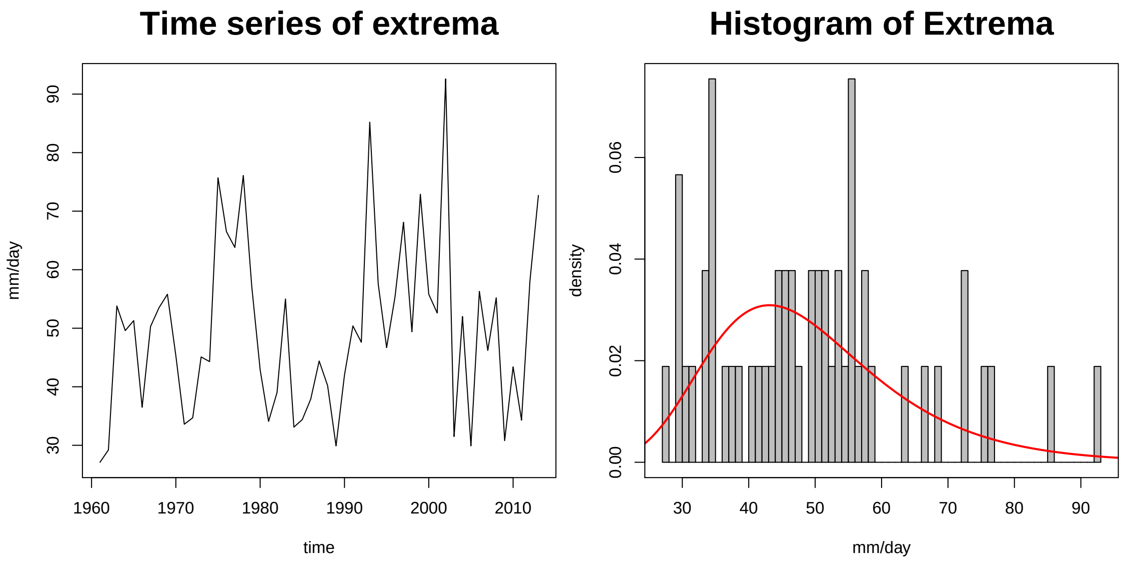


Figure 6: Time series (left) and histogram (empirical density) (right) of annual maxima with estimated density function (red line) as presented in the extreme value analysis information of 1-day precipitation at station Genève-Cointrin.

Estimation information Genève-Cointrin is an excellent example for the problems experienced when applying extreme value statistics: the estimated distribution does not fit well, either for the lower tail or the heavier (but not heaviest, here around 60mm) events, and the fit turns out to be questionable. It is possible that the lower daily precipitation maxima (perhaps below 30 or 40mm) are not really "extreme". As we shall see in section 4, the maximum should in principle represent the largest of an infinite number of independent values, and a year may be too short. Thus, fitting the distribution to 2-year maxima, rather than 1-year maxima, could eventually yield more convincing results. The GEV is also only a valid approximation if there is no systematic change in process within a year (stationarity assumption). This assumption is most likely violated by the seasonal cycle in daily precipitation. In such a case, restricting the analysis to one particular season may improve the results.

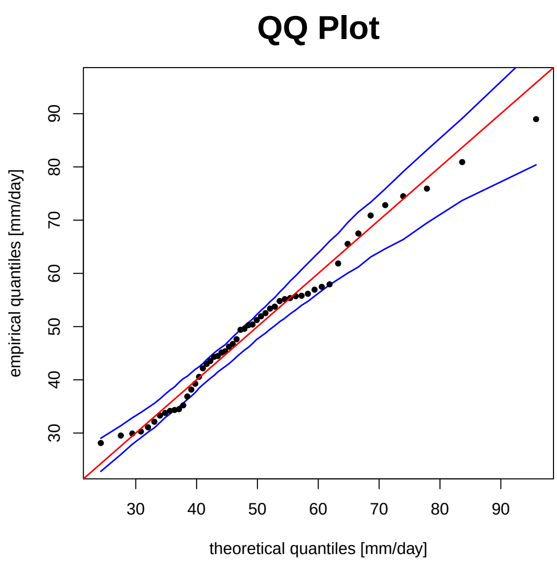


Figure 7: Quantile-quantile plot as presented in the extreme value analysis information of 1-day precipitation at station Genève-Cointrin.

Remark. What have we learned:
at Genève-Cointrin, the probability of very extreme precipitation decays exponentially with the severity of events. Statistically speaking, the data is consistent with a polynomial (therefore slower) decay of the tail of the GEV distribution. Thus, for each precipitation amount, there exists a non-negligible probability of exceeding this amount (there exists no upper bound).

The reliability of the fit is questionable. Thus, further investigation would be necessary to find a statistical model for which the assumptions are less violated. Extrapolation to unobserved levels, i.e., 100-year and higher return levels, is also not advisable. However, the information can be put to use as a pragmatic guide providing only very rough orders of magnitude. Given that the fit is questionable, the uncertainties themselves are not entirely reliable.

The largest 1-day precipitation event (92.6mm) was recorded November 14th, 2002, which corresponds roughly to an 80-year event.

2.3.2 Zürich / Fluntern

The extreme value analysis for 1-day precipitation observed in Zürich / Fluntern is shown on pages 8 and 9. Here, we will highlight differences to the statistical and extremal behavior of the extreme value analysis for Genève-Cointrin.

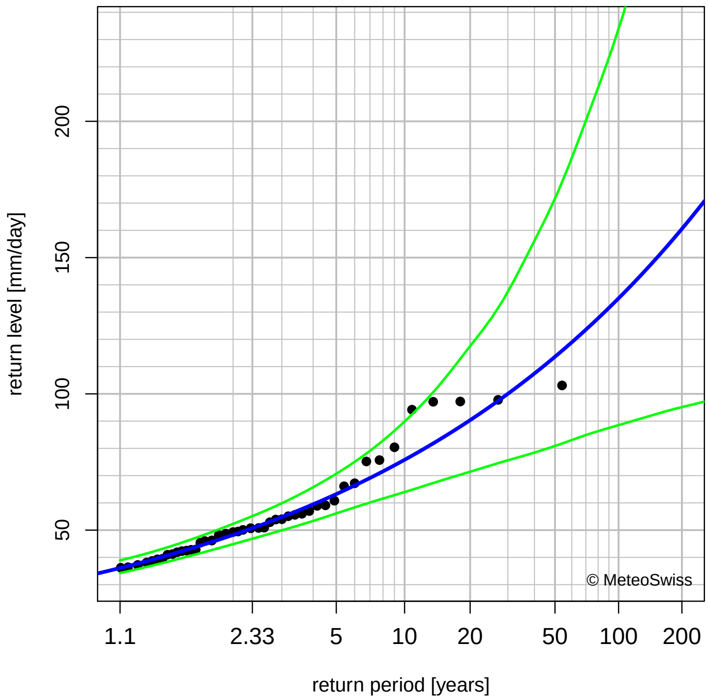


Figure 8: Return level plot as presented in the Extreme Value Analysis information of 1-day precipitation in Zürich. The blue line represents the return level estimates from the estimated GEV-distribution; the green lines represent the 95% confidence interval of the estimation; and the y-coordinate of the black points represent data used for estimation of the GEV-distribution; the plotting position (x-coordinate) of the points is solely determined by the length of the underlying record (and is not robust).

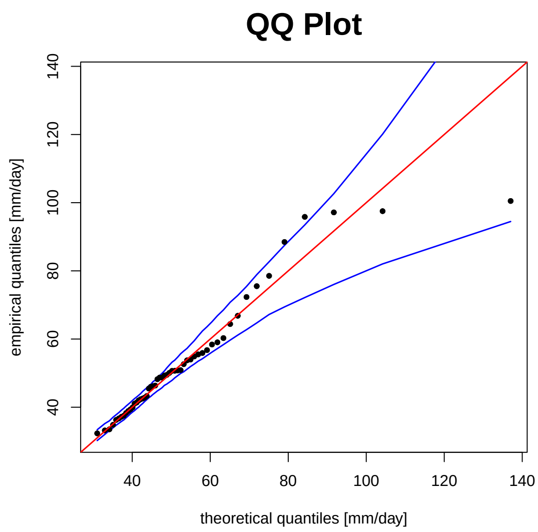


Figure 9: Quantile-quantile plot as presented in the Extreme Value Analysis information of 1-day precipitation at station Zürich / Fluntern.

Remark. In Zürich, the largest 1-day precipitation event was recorded September 21st, 1968 with 103.1mm. Three events of similar amount (97.8mm, 97.2mm, 97.1mm) have been recorded 08/2007, 08/1978 and 06/1973.

In Zürich, the statistical model does not represent the largest recorded 1-day precipitation events well, and is therefore unlikely to perform better as regards extremal behavior. The use of any return level larger than a 5-year event is not advisable. As a first step, a statistical model that takes into account the annual cycle of heavy precipitation would improve the estimated information.

Meanwhile, as a pragmatic approximation to the extremal behavior, the information about the largest observed annual maxima can be used.

3 Data

3.1 Stations

For the MeteoSwiss web-page, the observations used were taken at stations of the Swiss National Basic Climatological Network (NBCN), which groups the most important climatological stations within the observational network of MeteoSwiss (*Begert et al.*, 2007). Most of these 28 stations have very long records: data at some of these stations date back to the 19th century. The NBCN stations cover all of Switzerland (figures 10, 11), and should be to some extent representative of the regional climatic variations in monthly precipitation.

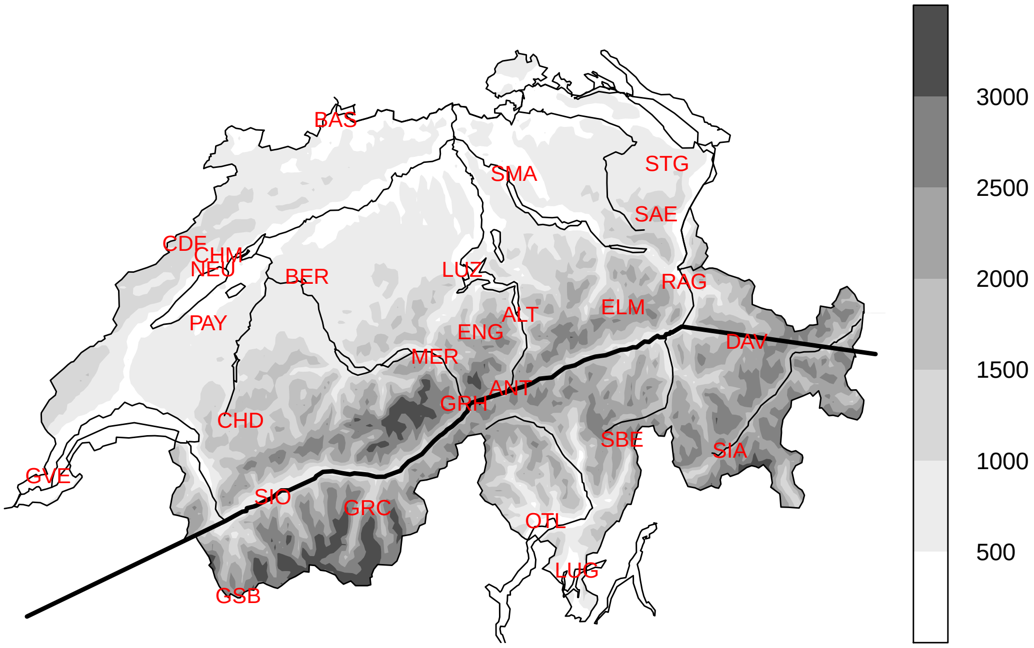


Figure 10: Map of NBCN stations. A list of station names, geographical location, and altitude can be found on page 42. Black line: line used to define the vertical cross-section across the Alpine ridge, defined as the Rhone and the beginning of the Rhine valleys, connected to each other and prolonged on either side by straight lines. To the east of the Rhine valley, the straight line is made to pass through the station Scuol.

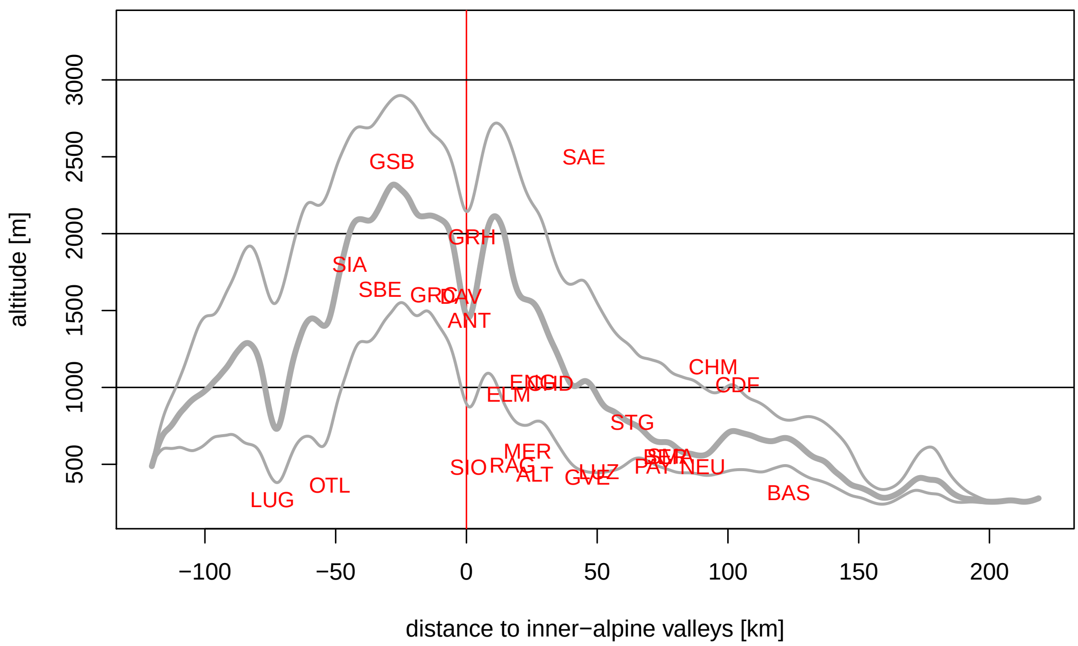


Figure 11: Profile of NBCN stations. The vertical cross-section across the Alpine ridge is defined as the distance to the inner-Alpine valleys (black line in figure 10). The profile of the Alps is computed with the USGS GTOPO30 (http://eros.usgs.gov) digital elevation model, and is the minimum air-line distance of each grid-point center to the inner-Alpine valleys. The thick (thin) grey line(s) represent(s) the smoothed median (10% and 90% quantiles) of the distances in 100m bins.

3.2 Data quality

The MeteoSwiss observational network encompasses different types of stations. NIME stations observe precipitation only, and measurements are carried out manually, at 7:30 local time every day. Climate Stations and SwissMetNet measure a wide range of meteorological parameters. At Climate Stations, manual precipitation measurements take place daily at 6:00 and 18:00 UTC. At SwissMetNet stations, precipitation is measured automatically in 10-minute intervals, starting on the hour. Note that for manual measurements, there is a time window of approximately 70 minutes for the observer to carry out the reading. The NBCN stations used for the extreme value analyses of precipitation on the MeteoSwiss WebSite are all Climate Stations.

The measuring instruments of the MeteoSwiss observational network have changed over time (table 2). The first ombrometers measured precipitation by means of an open cylindrical receptacle with a horizontal cross-section of 400cm^2. In the early 20th century, these were replaced with Hellmann rain gauges with a horizontal cross-section of 200cm^2. Both of these systems require manual reading. At the beginning of the 1980s, about two thirds of the NBCN stations were equipped with automatic (heated) tipping bucket rain gauges. Since 2012, a further automatic ombrometer has been introduced that uses a weighing mechanism.

measuring instrument	time period	reading time
ombrometer 400 cm^2	19th / early 20th century	6.00 UTC or 7.30 local time (m)
Hellmann rain gauge 200 cm^2	early 20th century till today	6.00 UTC or 7.30 local time (m)
tipping bucket rain gauges	since 1980s	every 10min., starting on the hour (a)
weighing precipitation gauge	starting 2012	every 10min., starting on the hour (a)

Table 2: Precipitation measuring instruments of the MeteoSwiss observational network with manual (m) and automatic (a) reading.

All of these ombrometers have their weaknesses, and precipitation measurements suffer from both systematic and non-systematic errors. The Hellmann rain gauges require manual reading at regular times, and the measurements may have been carried out at incorrect reading times. If the reading time has been skipped, an accumulated value is provided for the following day, and the amount of precipitation belonging to the proper time interval is unknown. The tipping bucket systems systematically underestimate (overestimate) precipitation at high (low) intensities, while the weighing rain gauges sometimes yield ghost precipitation due to pressure fluctuations caused by wind, or gradients in temperature.

3.3 Heavy precipitation data

The data analyzed consists of 1-day, 2-day, and 3-day precipitation. The aggregation of 1-day precipitation to 2 and 3-day precipitation ignores missing data, i.e. a missing daily sum is treated as zero precipitation and simply does not contribute to the aggregation. For extreme value analysis with the block maxima approach (see section 4 for details), only the maximum value per year is considered. The yearly maximum is computed from the monthly maxima and accepts no missing data, i.e. one missing month results in a missing year. Monthly maxima of daily precipitation sums are computed if there are no more than two non-consecutive missing days. In other words, two consecutive, or three non-consecutive, missing days will result in a missing year.

Since long records are essential for the credibility of extreme value analyses, only stations with a minimum of 40 years between 1961 and 2013 were selected (see table 3).

Remark. Requirements regarding observations of 1-day precipitation used for estimating extreme value distribution.

- at least 40 years of observations
- quality checked
- still operated today
- non-homogenized data
- well-documented station history
- at stations with known problems such as SAE, GSB, PIL, WFJ, MLS, COV (abbreviations are given on page 42) no extreme value distribution is fitted

Station abbreviation	1-day record (date)	1-day record (amount) in [mm]	2-day record (date)	2-day record (amount) in [mm]	3-day record (date)	3-day record (amount) in [mm]
ALT	31.07.1874	147.2	31.07.1874	194.7	04.05.2002	204.3
ANT	04.07.1916	185.3	03.05.2002	224.4	16.11.2002	269.5
BAS	25.05.1872	94.8	26.05.1872	139.4	26.05.1872	150.8
BER	14.08.2010	90.3	03.10.1888	139.4	03.10.1888	141.4
CDF	25.08.2002	100.8	19.01.1910	170.4	20.01.1910	206.7
CHD	28.12.1947	96.9	19.01.1910	167.3	20.01.1910	207.6
CHM	25.09.1987	102.9	26.09.1987	175.8	27.09.1987	175.8
DAV	18.11.1874	95.0	19.11.1874	143.0	20.11.1874	198.0
ELM	29.08.1890	147.3	04.07.1891	187.4	15.02.1990	205.6
ENG	31.07.1874	153.8	31.07.1874	226.5	01.08.1874	226.5
GRC	02.11.1968	170.3	15.10.2000	198.4	15.10.2000	226.8
GRH	21.12.1991	151.4	22.12.1991	236.2	22.12.1991	292.8
GSB	12.11.1996	159.6	12.11.1996	273.2	13.11.1996	321.5
GVE	14.11.2002	92.6	09.09.1993	111.3	09.09.1993	129.2
LUG	21.08.1911	262.8	21.08.1911	342.5	22.08.1911	368.4
LUZ	06.06.2002	111.8	06.06.2002	123.3	27.07.1976	133.2
MER	22.08.2005	110.9	22.08.2005	205.0	22.08.2005	218.0
NEU	08.10.1949	116.1	26.09.1987	125.7	20.01.1910	126.1
OTL	26.09.1991	317.9	22.09.1981	377.8	23.09.1981	416.4
PAY	26.09.1987	78.4	26.09.1987	106.8	27.09.1987	106.8
RAG	29.08.1890	175.0	29.08.1890	217.0	31.08.1890	241.0
SAE	22.08.2005	186.7	03.11.1921	290.0	04.11.1921	321.4
SBE	07.08.1978	175.5	15.11.2002	258.5	16.11.2002	367.7
SIA	03.11.2000	108.0	15.11.2002	159.3	16.11.2002	223.0
SIO	14.02.1990	79.3	14.02.1990	138.5	15.02.1990	170.2
SMA	11.06.1876	171.5	12.06.1876	244.5	12.06.1876	272.5
STG	01.09.1881	250.0	02.09.1881	309.1	03.09.1881	338.9

Table 3: NBCN stations with more than 40 available years of daily precipitation data with the records in mm of the 1-day, 2-day, and 3-day precipitation.

4 Theoretical background: Extreme Value Statistics

Univariate extreme value statistics offers three different approaches to describe extreme events. The approaches differ in the way they characterize extreme events: via the largest/smallest values; via the excess of a pre-defined threshold value; or, more generally, via the count of events falling into a pre-defined set of values. They are called block maxima (BM), peaks over threshold (POT) or point process (PP) approach, respectively. In the following we will describe the BM approach closely following *Coles* (2001).

Here, we first present some general properties of the generalized extreme value distribution (GEV), the “extremal” distribution that - so extreme value theory tells us - represents the behavior of the maxima. Then, we explain different methods to infer this distribution from the data at hand, and assess the goodness of the estimated fit. Finally, we close the section with the notion of return levels, a means to illustrate the behavior of extreme values.

4.1 The GEV distribution

Using the maxima of a block of observations to characterize the behavior of rare events find its justification in extreme value theory. *Fisher and Tippett* (1928) and *Gnedenko* (1943) prove that the distribution of the largest values of any[1] random variable converges towards one of three types of limit distributions, and that these three types are the only possible limit distributions. This law is known as the extremal types theorem or Fisher-Tippett-Gnedenko theorem. Strictly speaking, the maxima must first be appropriately re-scaled for the limit distribution to be non-degenerate (see *Coles*, 2001, p.46 for details).

Remark. The extremal types theorem - as does the central limit theorem for sample means - justifies the practice of approximating the distribution of sample maxima via extreme data only; thus, it is not necessary to estimate the population’s distribution F in order to describe the behavior of the population’s rare events.

Remark. *Let M_n be a sequence of maxima, a_n and b_n appropriate re-scaling sequences of constants and G the limit distribution of $(M_n - a_n)/b_n$, then*

$$\Pr\{M_n \leq z\} \approx G^*(z)$$

[1] see remark 4.1 and *Leadbetter et al.* (1983)

with G^* another member of the GEV family.

In practice, we do not need to estimate the re-scaling sequences, a_n and b_n, but only the distribution G^*.

Remark. For certain distributions F, a non-degenerate limit distribution function does not exist; this is the case, for instance, for the Poisson and geometric distribution, which both describe discrete counting processes. For such distributions, the maxima cannot be described by a GEV.

The three types of limit distributions are the Weibull, Gumbel, and Fréchet families of distributions. They can be summarized in one parameterization of the limit model - in exchange for an additional parameter (*von Mises*, 1954; *Jenkinson*, 1955). This generalized limit distribution, known as the GEV distribution (definition 1), will be used for the subsequent analyses.

Definition 1 (GEV distribution)**.** The family of distributions defined on the set $\{z : 1+\xi(z-\mu)/\sigma > 0\}$ via

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

with $-\infty < \mu < \infty$, $\sigma > 0$, and $-\infty < \xi < \infty$, is called the generalized extreme value family of distributions (GEV distribution). The parameter μ is a location parameter, the parameter σ is a scale parameter and ξ is a shape parameter. The subset of the GEV family with $\xi = 0$ is interpreted as the limit as $\xi \to 0$.

Example (GEV distribution)**.** Examples of

Weibull ($\mu = 45$, $\sigma = 11$, $\xi = -0.1$),
Gumbel ($\mu = 45$, $\sigma = 11$, $\xi = 0$), and
Fréchet ($\mu = 45$, $\sigma = 11$, $\xi = 0.25$)

probability density functions. These distributions and their samples will be used throughout the document to illustrate extremal behavior, estimation behavior, uncertainties, and exceedance probabilities.
They loosely resemble the 1-day-accumulated precipitation annual maxima distributions at Fribourg (not shown), Genève, and Zürich-Fluntern, respectively.

The cumulative distribution function (CDF) and probability density function (PDF) of a Weibull (red), a Gumbel (black) and a Fréchet (blue) distribution are shown in figure 12. All three have a location parameter $\mu = 45$, and a scale parameter $\sigma = 11$. They differ by their shape parameter ξ, which is negative (in this example $\xi = -0.1$) for a Weibull, zero for a Gumbel, and positive (in this example $\xi = 0.25$) for a Fréchet distribution (see generalized parameterization of definition 1).

Figure 13 shows the PDFs of the three example distributions seen in figure 12, here in separate plots, with the histograms of the respective samples appearing in figure 14. Figure 14 shows random samples (of equal length) drawn from these three distributions: these can be imagined as time-series of annual maxima of daily precipitation.

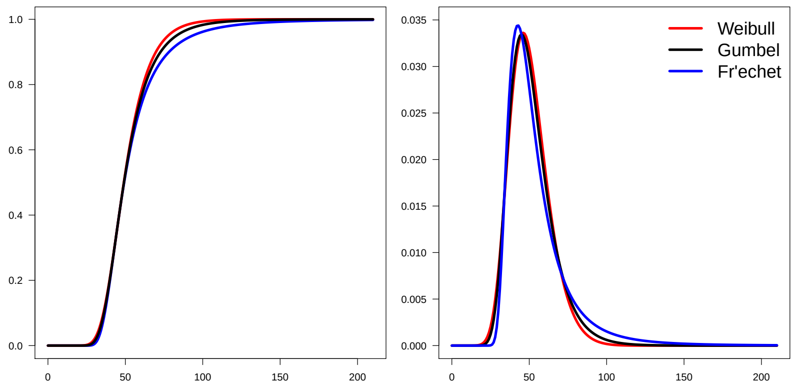


Figure 12: Three extreme value distribution functions (left: cumulative distribution function (CDF); right: probability density function (PDF)) belonging to the Weibull (red; $\xi = -0.1$), Gumbel (black; $\xi = 0$), and Fréchet (blue; $\xi = 0.25$) distribution families, respectively. All three have the same location parameter ($\mu = 45$) and scale parameter ($\sigma = 11$).

Remark. The three different families of distributions are distinguishable by their tail behavior, and in particular by the rate of decay in the tail:

- The Weibull family, i.e., the subset of GEV distributions with $\xi < 0$, has a finite upper end-point (figure 12 red line - also called type III or bounded). Thus, there is 0 probability for values above that end-point to occur.
- The Gumbel, family, i.e., the subset of GEV distributions with $\xi = 0$, has an infinite upper end-point. The tail of the probability density distribution decays exponentially (figure 12, black line - also called type I or light-tailed). Thus, regardless the value, there is a non-vanishing probability for a higher value to occur.
- The Fréchet family, i.e., the subset of GEV distributions with $\xi > 0$, has an infinite upper end-point. The tail of the density distribution decays polynomially (figure 12, blue line - also called type II or heavy-tailed). Here too, whatever the amplitude of the event, a larger one has a non-negligible probability to occur.

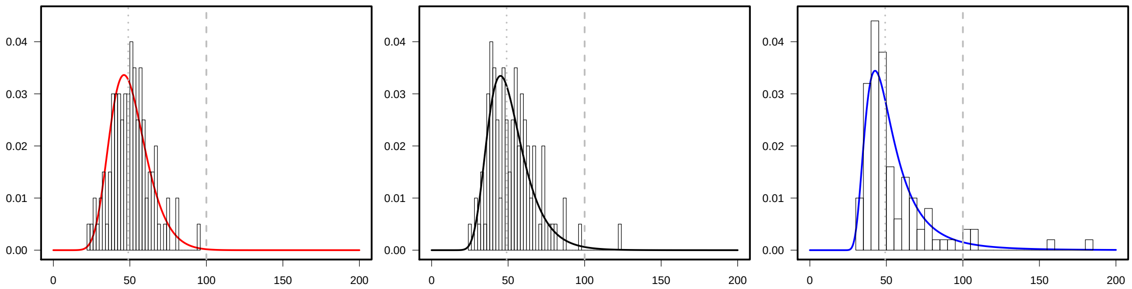


Figure 13: Probability density functions of figure 12 (solid lines) of the Weibull (left, red, $\xi = -0.1$), Gumbel (center, black, $\xi = 0$), and Fréchet (right, blue, $\xi = 0.25$) distribution, with a location parameter of 45 and a scale parameter of 11. The histograms (with 50 breaks) show the empirical densities of the three samples of length 100 seen in figure 14, drawn from the three respective distributions. The vertical dotted line represents the sample median, while the dashed vertical line at 100 is meant to ease comparison between panels, and with figure 14.

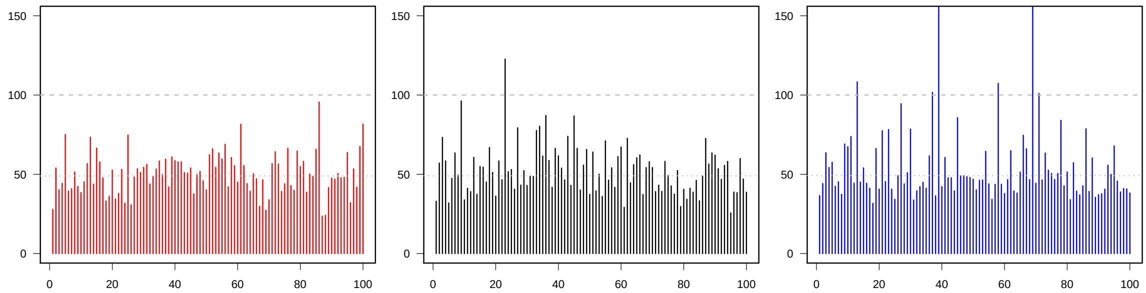


Figure 14: Three samples of length 100 drawn from the Weibull (top, red, $\xi = -0.1$), Gumbel (center, black, $\xi = 0$), and Fréchet (bottom, blue, $\xi = 0.25$) distribution, with a location parameter of 45 and a scale parameter of 11. The horizontal lines indicate, as in figure 13, the sample median (dotted line) and the value 100 (dashed line).

4.2 Inference for the GEV distribution

Inference is - one could say - an educated guess of the distribution, based on the available sample. Several methods exist to estimate the underlying GEV distribution from the data at hand. Here, we only present two of them: the Maximum Likelihood (ML) method and the method of L-moments. Both approaches are used for estimation of the analyses: The default routine is ML estimation. In case no global maximum is found for the ML estimation, the L-moments estimation is used.

Given the finite size of the sample, the estimates are associated with uncertainties. These describe a range of values that would also be consistent with the data at hand. For some estimation methods, theoretical results allow quantification of the estimation uncertainty. Alternatively one can apply bootstrapping methods to obtain confidence intervals of the estimated parameters or derived quantities.

Remark. As it was customary at the time, the approach of *Röthlisberger et al.* (1991) was to first opt for one of the extremal distribution families, and then estimate the scale and location parameter using regression methods.

In effect, this means selecting a particular paper, with a pre-drawn graph in either log or log-log scale. Thus, the estimation hinges on the preliminary (educated) guess of the appropriate family.

“But”, as pointed out by (*Coles*, 2001, page. 47) “there are two weaknesses: first, a technique is required to choose which of the three families is most appropriate for the data at hand; second, once such a decision is made, subsequent inferences presume this choice to be correct, and do not allow for the uncertainty such a selection involves, even though this uncertainty may be substantial.”

With the adoption of the generalised extreme value distribution parametrization, “the uncertainty in the inferred value of ξ measures the lack of certainty as to which of the original three types is most appropriate for a given dataset.“ (*Coles*, 2001, page. 48)

Notation. Given a sequence of independent random variables $X_1, \ldots, X_n$ with a common distribution function F, we denote with $M_n = \max\{X_1, \ldots, X_n\}$ the maximum of the sequence. For example, M_{365} would be the annual maximum of daily observations.
Note that the daily observations should not display seasonal behaviour or dependence on the

observations of previous days.
To avoid multiple indices we denote with Z_i the maximum M_n of block i, and omit the information on the block length n. You can imagine $\{Z_1, \ldots, Z_m\}$ as the set of m annual maxima of daily precipitation.

4.2.1 Maximum likelihood estimation

Maximum likelihood estimation looks for the (distribution) parameters that are most likely to have led to the sample at hand, assuming that the largest values of the process of interest are indeed described by the chosen distribution family, here the GEV distribution.

The likelihood of drawing the recorded values from one particular distribution of the GEV family can be represented as an n-dimensional surface, where n is the number of parameters of the distribution. Maximizing this likelihood function or its logarithm (the log-likelihood) yields an estimate of the parameters.

Parameter estimation The log-likelihood function, ℓ, derives from the distribution function. In our analyses, the numerical optimization for obtaining the minimum of $-\ell$ is done with Nelder-Mead at initial values estimated by L-moments with a fixed $\xi = 0.5$ (see section 4.2.2 or *Hosking* (1990) for details).

Definition 2. Log-likelihood function of the GEV, at $z_1, \ldots, z_m$,

$\ell(\mu, \sigma, \xi | z_1, \ldots, z_m) =$

$$-m \log \sigma - (1 + \frac{1}{\xi}) \sum_{i=1}^{m} \log \left[1 + \xi \left(\frac{z_i - \mu}{\sigma} \right) \right] - \sum_{i=1}^{m} \left[1 + \xi \left(\frac{z_i - \mu}{\sigma} \right) \right]^{-\frac{1}{\xi}}$$

Notation. To distinguish between theoretical and estimated parameters, one uses the notation of a small hat above the parameter name, e.g., ξ denotes the theoretical shape parameter, while $\hat{\xi}$ denotes the estimated shape parameter.

For each of the samples in figure 14, ML estimation was used to estimate the distribution of the population as if the sample were a set of maxima of blocks of observations. Figure 15 shows the histograms of our samples with the true (solid lines) distribution as in figure 13, as well as the estimated (dashed lines) distribution of the population. The histogram can be understood as the empirical "PDF“, i.e., without the assumption of a model. The dashed lines show the distribution estimated by maximum-likelihood from the samples. The vertical lines represent the median of the theoretical (solid line) and of the estimated (dashed line) distribution. Keep in mind that this is a sandbox example in which we know the distribution of the population the sample was drawn from (solid lines). In real life, of course, the population is unknown. We call the estimated distributions (dashed lines) fitted models, because they are adapted to the observed data.

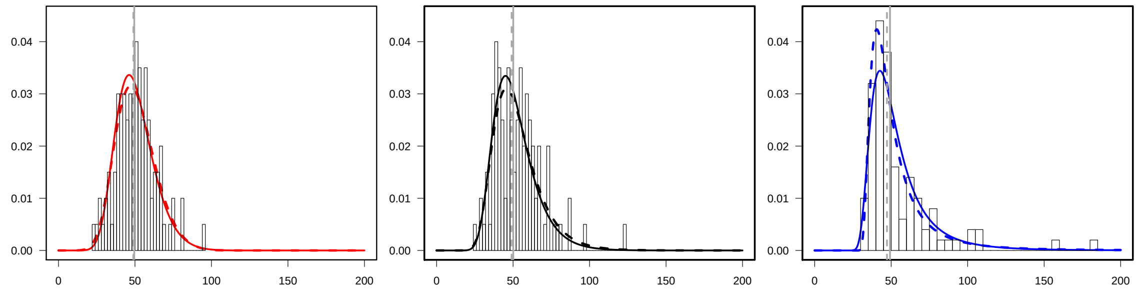


Figure 15: Histograms of the Weibull ($m = 100$, left, red), Gumbel ($m = 100$ center, black), and Fréchet sample data ($m = 100$, right, blue) seen in figure 14. Overlaid are the PDFs fitted to the sampled data (dashed lines), and the theoretical PDFs (solid lines, see figure 13) of the true distributions from which the data was sampled.

Both the PDFs and the medians of the distributions differ (figure 15), the more so for the heavy-tailed distribution (right panel). Thus, inferring the underlying distribution from a small sample is difficult, and the inferred information is inherently uncertain, because it depends on the particular sample we used. In the setting of maximum likelihood estimation, the uncertainty of the estimates can be reduced by increasing the number of independent maxima.

For the interpretation of uncertainties it is very important to be precise if one talks about uncertainty of the estimation, e.g., the uncertainty in the assessment if a coin is fair based on a sample of coin tosses, or uncertainty about the “prediction” of an event, e.g., the uncertainty if the outcome of tossing a coin is heads or tails.

The example (figure 15) illustrates the inherent uncertainty of inference with few observed extremes. Here, we have estimated the distribution from $m = 100$ years; we can be certain that, by construction, all assumptions (data quality, representativeness (but sample size), Z_i iid, model choice) are fulfilled, and only the estimation uncertainty still remains (and stochastic uncertainty, which does not play a role in this section). Of course, this is not the case in real life.

Even though all our assumptions are correct, the fitted model deviates from the true distribution. This remaining uncertainty is the estimation uncertainty.

Estimation uncertainty ML estimation provides a theoretically founded approximation for the estimation uncertainties. Standard errors of maximum likelihood estimates can be calculated easily. Nevertheless, using the profile likelihood function is recommended, as it allows for asymmetric confidence intervals (CIs).

Definition 3. Profile log-likelihood confidence intervals for $\hat{\xi}$, $\hat{\sigma}$ or $\hat{\mu}$.
For uncertainty estimates of $\hat{\xi}$, one fixes $\xi = \xi_0$ for a range of ξ_0, e.g. from -1.5 to 2, and for each ξ_0 maximizes the log-likelihood function with respect to μ and σ, given the observations. Thus one obtains a concave function of ξ with a maximum at $\hat{\xi}$. Approximate confidence intervals can be derived based on the fact that two times the difference in the log likelihood function at ξ_0 and $\hat{\xi}$ follows a χ^2_1 distribution (see *Coles*, 2001, section 2.6.6 p.35 for details). The same can be done for $\hat{\mu}$ and $\hat{\sigma}$.

In figure 16 we see the profile log-likelihood functions (black solid lines) for all parameters. The vertical solid lines represent the true parameters and the three dashed vertical lines indicate the ML estimate, with the 95% confidence intervals ($1.96 \times \text{se}_{mle}$) on either side. The profile likelihood CIs are represented by the dotted vertical lines. They mark the intersection of the profile log-likelihood function with the critical value of the χ_1^2 distribution (lower blue horizontal line).

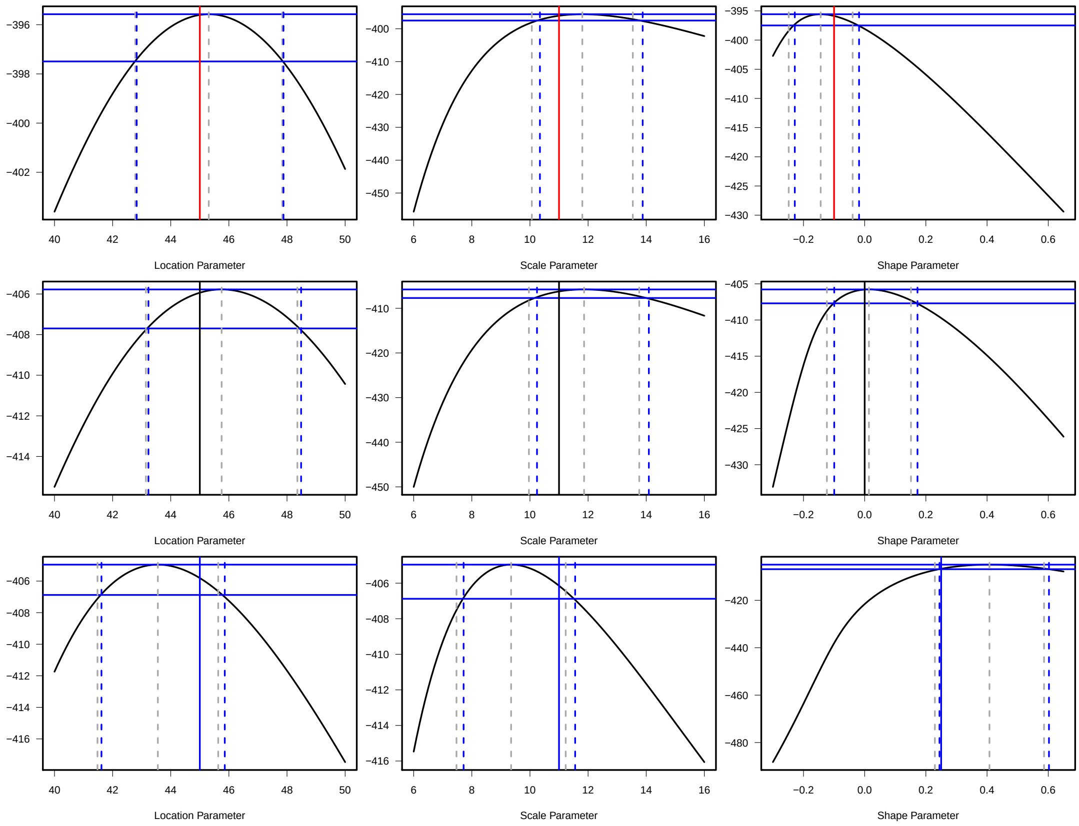


Figure 16: Profile log-likelihood functions (solid black curves) for each parameter - location (left), scale (center), shape (right)- estimated from the Weibull ($m = 100$, red, top), from the Gumbel ($m = 100$, black, center) or from the Fréchet ($m = 100$, blue, bottom) sample data. The blue horizontal solid lines mark the maximum value of the profile log-likelihood function (corresponding to the ML estimate) and the critical value (95%) of the χ_1^2 distribution. The profile likelihood 95% confidence intervals are indicated by the dotted vertical lines, the maximum likelihood 95% estimate and confidence intervals are indicated by the dashed vertical lines and the true value is indicated by the solid vertical lines (Weibull in red, Gumbel in black, and Fréchet in blue).

As expected, the true values for bounded (Weibull) or light-tailed (Gumbel) type samples are well within the CIs (regardless of the approximation method) and the two methods for approximating the CIs hardly make a difference. For the sample of Fréchet type, on the other hand, even though the CIs are larger than in the other two cases, the true value is just barely within the CI based on the standard errors. Thus, taking confidence intervals into consideration in practical applications is essential if 'capturing' the true value is of any importance, and extreme value analyses should mention these inherent uncertainties explicitly.

The width of the CI depends on sample size. The example presented here had a sample size of 100. In real life situations, however, we often have only 40 years of observations at our disposal. For comparison, we have plotted in figure 17 the profile log-likelihood functions for sample size 40 drawn from the same population distribution as in figure 12. Evidently, the confidence intervals for each parameter are considerably larger than with 100 years of data. Thus, shorter data records yield larger CIs.

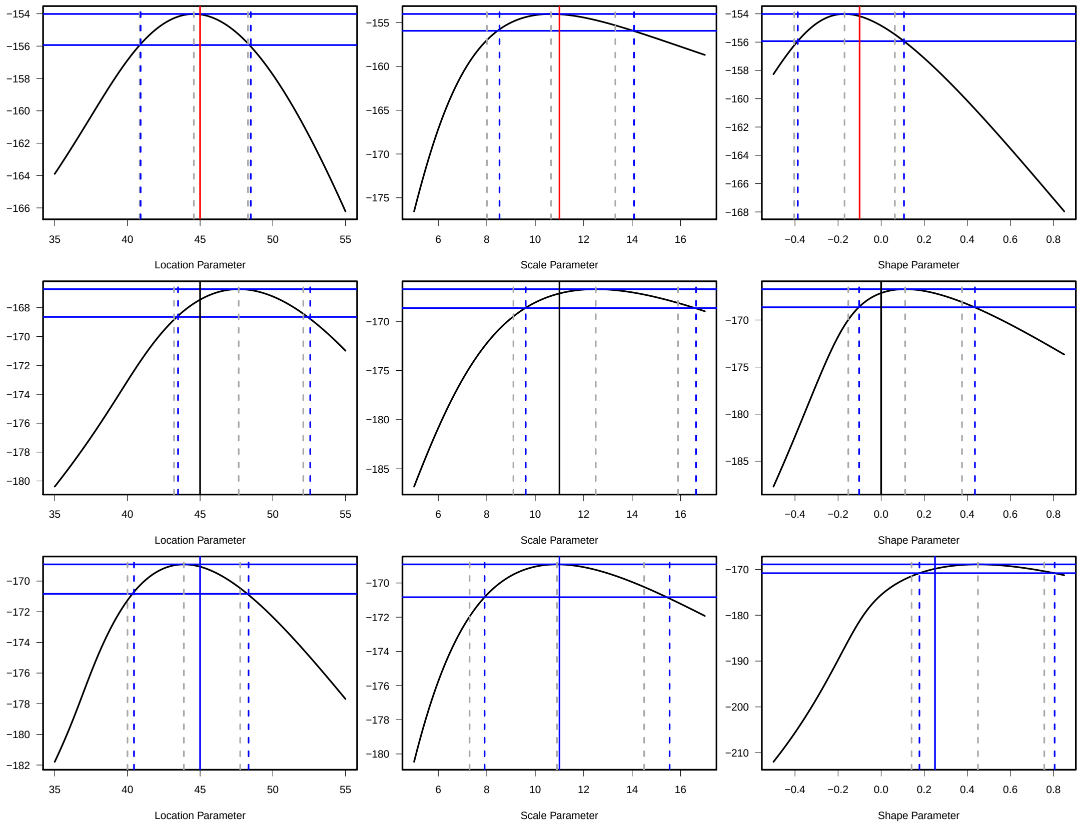


Figure 17: Profile log-likelihood functions (solid black curves) for each parameter - location (left), scale (center), shape (right) - estimated from the Weibull sample data ($m = 40$, top), from the Gumbel sample data ($m = 40$ center), and from the Fréchet sample data ($m = 40$, bottom). The profile likelihood 95% confidence intervals are indicated by the dotted vertical lines, the maximum likelihood 95% confidence intervals are indicated by the dashed vertical lines and the true value is indicated by the solid vertical lines (Weibull in red, Gumbel in black, and Fréchet in blue).

In addition, both the inferred distribution and the approximated confidence intervals depend on the sample at hand. Figure 18 illustrates the fact that subsamples of equal size (m = 40) result in different inferred distributions, although they were drawn from the same population. The difference is particularly striking when drawing from the Fréchet family.

Conclusions regarding the temporal evolution of the behavior of extreme events therefore cannot be based on fitting distributions within different time periods (for instance 1901 - 1950, 1951 - 2000) and comparing the estimates without taking into account the uncertainties. For such a task, rather sophisticated tests designed specifically for extremes are necessary.

4.2.2 L-moments estimation

The L-moments (*Hosking*, 1990) belong to the class of moment estimators, which are quite popular among hydrologists for the estimation of the GEV parameters. One minor drawback of this method is the lack of theoretically derived uncertainty estimates, but this can be circumvented by parametric

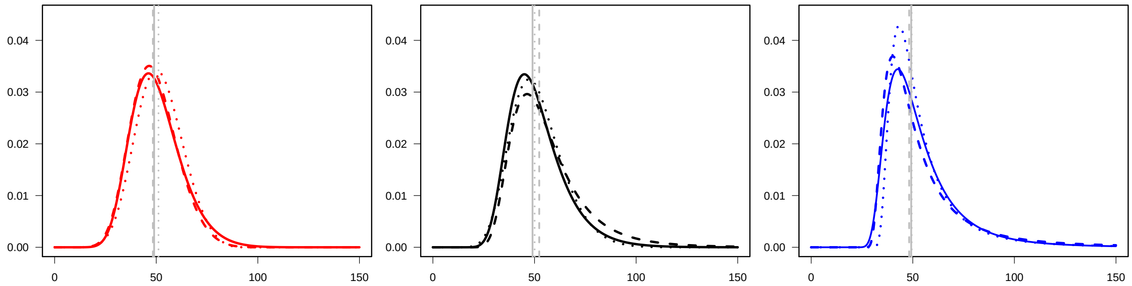


Figure 18: Six fitted PDFs (dashed and dotted lines) estimated from two subsets of Weibull sample data ($m = 40$, left, red), two subsets of the Gumbel sample data ($m = 40$, 6 center, black), and two subsets of the Fréchet sample data ($m = 40$, right, blue). The theoretical PDFs of figure 13 are represented by solid lines and are considered as "truth".

resampling (see below).

Parameter estimation

Definition 4. L-moment estimation of the GEV parameters (*Hosking*, 1990)

$\hat{\xi} \approx -7.8590 \left(\frac{2}{(3+t_3)} - \frac{\log 2}{\log 3} \right) - 2.9554 \left(\frac{2}{(3+t_3)} - \frac{\log 2}{\log 3} \right)^2$ with $t_3 = l_3/l_1$ [a],

$\hat{\sigma} = l_2 \frac{-\hat{\xi}}{(1-2^{\hat{\xi}})} \Gamma(1-\hat{\xi})$ with $l_2 = 1/2 \frac{1}{\binom{n}{2}} \sum\sum_{i>j} (x_{(i)} - x_{(j)})$, and

$\hat{\mu} = l_1 - \frac{\hat{\sigma}}{\hat{\xi}} \left(\Gamma(1-\hat{\xi}) - 1 \right)$ with $l_1 = \frac{1}{n} \sum_i x_i$.

[a] $l_3 = 1/3 \frac{1}{\binom{n}{3}} \sum\sum\sum_{i>j>k} (x_{(i)} - 2x_{(j)} + x_{(k)})$

Here, in figure 19, we see the L-moments estimates in addition to the maximum likelihood estimates. The difference between the estimates is very small.

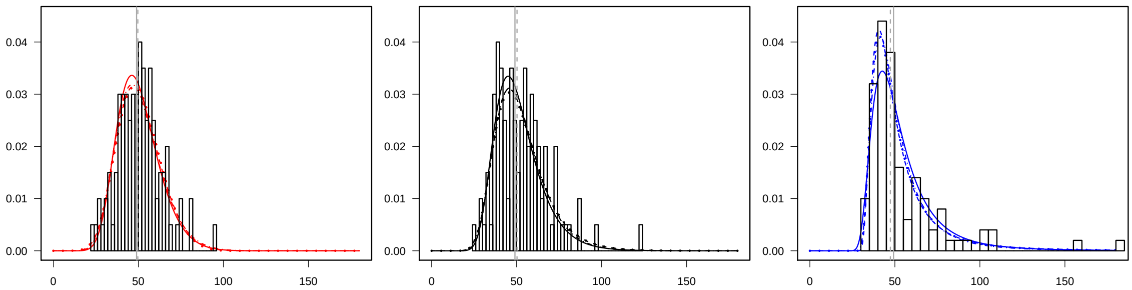


Figure 19: Three fitted PDFs (dashed lines) estimated from the Weibull sample data ($m = 100$, left, red), the Gumbel sample data ($m = 100$, center, black), and the Fréchet sample data ($m = 100$, right, blue). The dotted lines are the maximum likelihood estimates as in figure 15. The theoretical PDFs of figure 13 are represented by solid lines, andare considered as "truth". The samples of figure 14 appear as histograms with 50 breaks.

Parametric resampling For confidence interval estimations, we choose a parametric resampling approach with a sample size of 1000.

The idea is to generate surrogate data from the estimated distribution, representing possible, additional samples. Samples of the same size as the original sample of observations are drawn from the estimated distribution, which is entirely characterozed by the estimated parameters, hence the name "parametric resampling". (Some methods resample the observed data.)

The new, surrogate, samples then serve as starting points for new estimates of the underlying distribution. From the resulting, here 1000, simulated distribution estimates, we construct 2-sided 95% confidence intervals, by taking the 2.5% and the 97.5% quantiles of the estimated quantity.

Remark. Estimation uncertainties are not the only uncertainties of an extreme value analysis. There exist several sources of uncertainty (some of which we cannot quantify):

- the data quality, measurement errors, or the representativeness of the data
- the assumptions that the maxima are independent, identically distributed (i.i.d.)
- the model choice (GEV distribution, sufficient block length)
- the estimation of model parameters
- the interpretation of probabilistic results (stochastic uncertainty / sampling uncertainty)

4.3 Assessing the fit of the model

The validity of the model assumptions usually cannot be checked in a systematic manner. However, it is possible to compare the empirical distribution $\tilde{G}$, that makes no assumptions regarding an underlying model, with the estimated distribution $\hat{G}$.

Notation. $\tilde{G}$ empirical distribution, i.e., for each value of the sample, the proportion of the total number of values in the sample that is below it.
$\hat{G}$ estimated distribution

This can be done with diagnostic plots or with goodness-of-fit tests. Diagnostic plots are visual guides to the validity of model assumptions, and will be introduced in the next subsection. Goodness-of-fit tests quantify the difference between empirical and estimated distribution and are presented in the subsequent subsection.

4.3.1 Visual guides

Probability-probability (empirical vs. estimated probability) and quantile-quantile (empirical vs. estimated quantile) plots are instruments for visual inspection of fit reliability. The empirical and estimated probabilities (quantiles) are plotted against each other. Should the estimated distribution describe the observed data well (up to sampling uncertainty), the points should line up along the line of slope 1 and intercept 0.

Although they offer essentially the same information, probability-probability (PP) plots and quantile-quantile (QQ) plots highlight different areas of the distribution.

PP-plots The probability-probability plot plots the estimated probability against the empirical probability.

Definition 5 (probability-probability-plot.)**.** A probability-probability plot consists of the points

$$\left\{\left(\hat{G}(z_{(i)}), \tilde{G}(z_{(i)}\right) : i = 1, \ldots, m\right\},$$

where $z_(i)$, $i = 1, \ldots, m$, are the ordered observed values, $z_{(1)} \leq \ldots \leq z_{(i)} \leq \ldots \leq z_{(m)}$, $\tilde{G}$ is the empirical distribution function and $\hat{G}$ the estimated distribution function.

In the context of extreme values, the PP plot has a weakness: By definition, both the estimated and the empirical probabilities are bound to approach 1 as the quantile increases. Therefore, the PP plot offers little (or no) information in the region of most interest.

The QQ plot circumvents this weakness by changing the scale.

QQ-plots The quantile-quantile plot associates to each value of the ordered sample (e.g., a maximum over a block) the corresponding quantile of the estimated distribution.

By definition, the cumulative distribution function at a given value z_p gives the probability p that any of the values up to (and including) z_p may be realized. The value z_p is then called the p-quantile: $\Pr\{Y \leq z_p\} = p$. The median of a distribution, for instance, is its $\frac{1}{2}$-quantile or 50%-quantile; if $z_{0.5}$ is the median, there is a 50% chance for a realization of the random variable to be below $z_{0.5}$. If $z_{0.75}$ is the 75%-quantile, there is a probability of 75% that any value smaller or equal $z_{0.75}$ will be be realized. Thus, if a sample value is such that 75% of sample values are below it, it will be associated on the QQ-plot with the 75%-quantile of the estimated distribution.

If the model is appropriate for the data at hand, the empirical quantiles of the data should approximately agree with the estimated quantiles, the more so for a large sample.

For the upper quantiles, empirical quantiles that are systematically smaller than the fitted ones suggest that the tail of the true distribution may decay more slowly than that of the fitted distribution.

Definition 6 (quantile-quantile-plot.)**.** A quantile-quantile plot consists of the points

$$\left\{\left(\hat{G}^{-1}(\frac{i}{m+1}), z_{(i)}\right) : i = 1, \ldots, m\right\},$$

where $z_{(i)}$, $i = 1, \ldots, m$, are the ordered observed values, $z_{(1)} \leq \ldots \leq z_{(i)} \leq \ldots \leq z_{(m)}$, $\hat{G}$ is the estimated distribution function.

4.3.2 Goodness-of-fit criteria

We employ goodness-of-fit methods based on empirical distribution statistics, such as the well-known Kolmogorov-Smirnoff and Cramer-von Mises statistics but also statistics that put more weight on the tail, i.e., on the rarer events of the distribution.

In a first step, we consider the root-mean-square-error (RMSE) between empirical and estimated distribution (see table 5). In addition, we assess the appropriateness of the fit of the model via goodness-

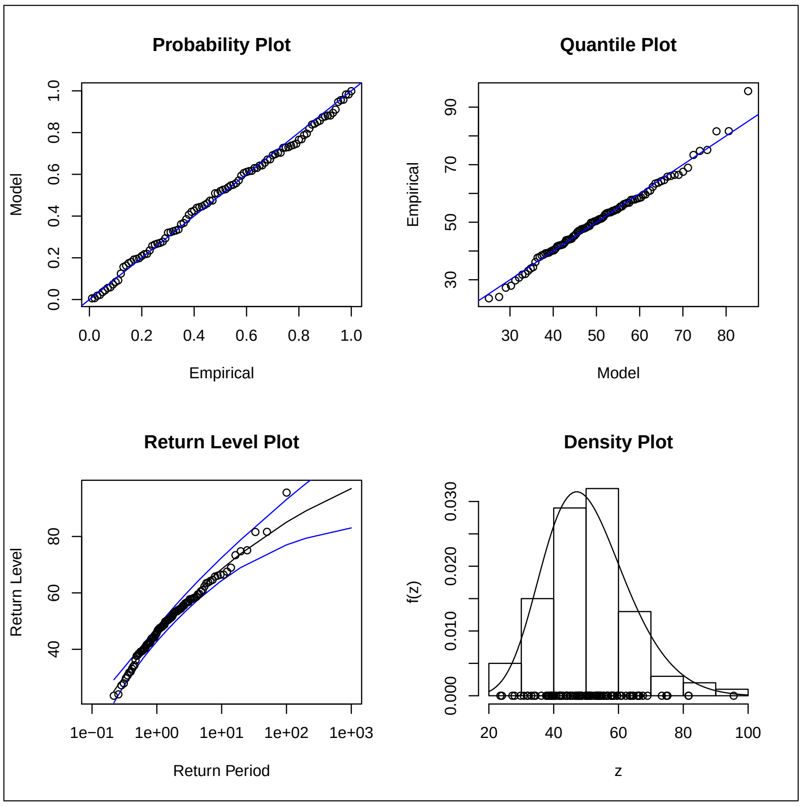


Figure 20: Suit of diagnostic plots for the Weibull sample $(m = 100)$ as provided by the R-package ismev. Note that the CIs are based on the mle standard errors and therefore the upper limit of the CI is underestimated.

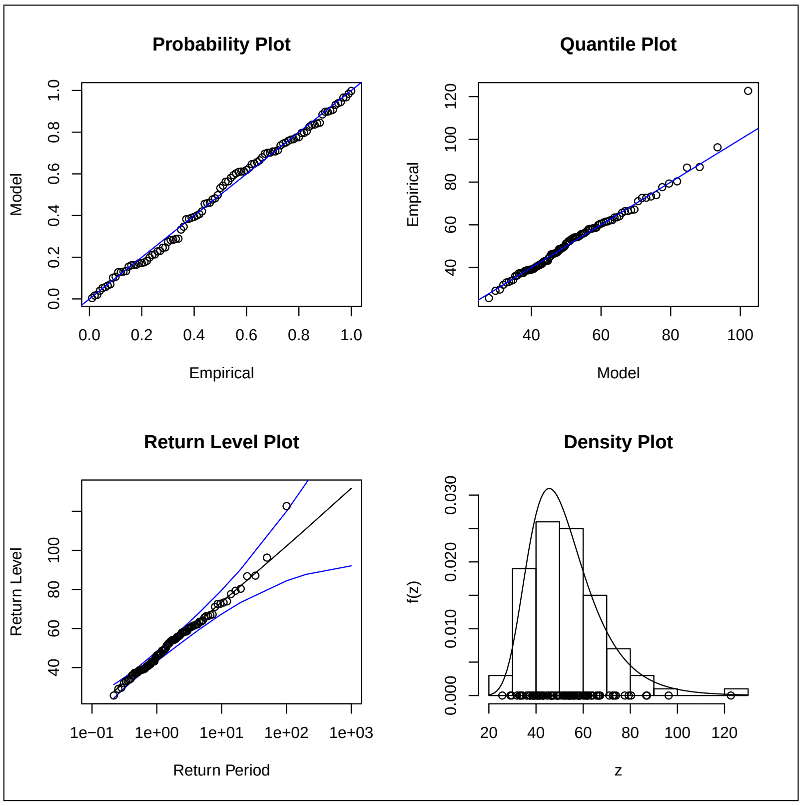


Figure 21: Suit of diagnostic plots for the Gumbel sample $(m = 100)$ as provided by the R-package ismev. Note that the CIs are based on the mle standard errors and therefore the upper limit of the CI is underestimated.

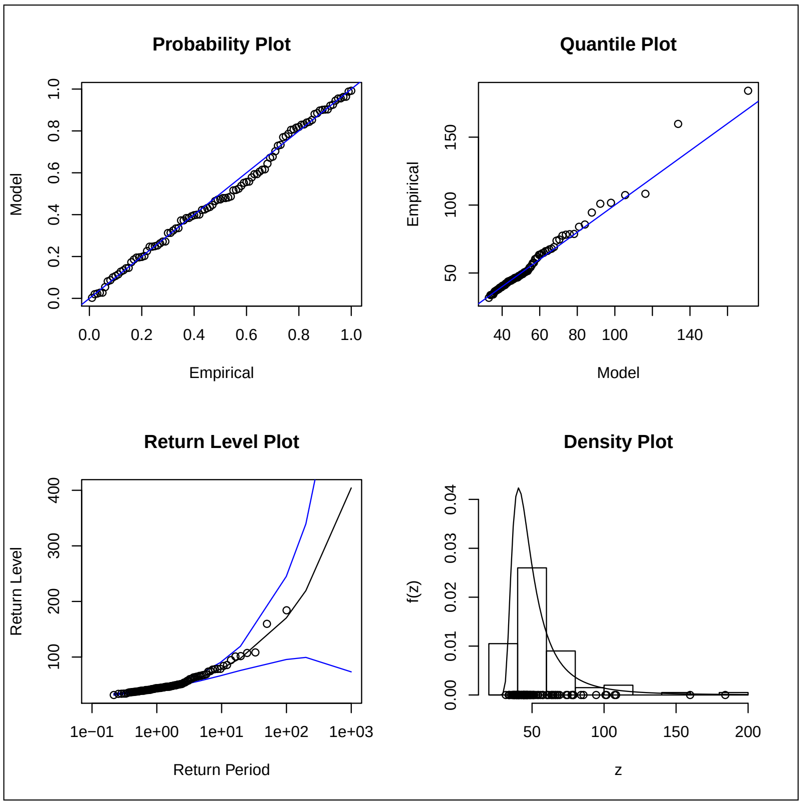


Figure 22: Suit of diagnostic plots for the Fréchet sample ($m = 100$) as provided by the R-package ismev. Note that the CIs are based on the mle standard errors and therefore the upper limit of the CI is underestimated.

of-fit tests. Goodness-of-fit tests test if the overall distance between the empirical distribution and the estimated distribution is significantly different from zero.

As with the visual guides, one can perform the tests on different scales, and / or put weight on areas of the distribution one is interested in. As we would like to draw conclusions regarding the behavior of extremes, we need the right tail of the distribution to be reliable.

To this end, we apply a suit of tests (see table 5 and *Luceno* (2006)). For each test, the critical values are obtained by bootstrapping. The null hypothesis (H_0) is that the empirical and estimated distribution are the same.

For the overall decision whether the fitted distribution is suited for extrapolation outside the observed range, we use the following set of rules:

Verdict	Rule
poor	any H_0 rejected at $\alpha = 5\%$
questionable	at least one H_0 rejected at $\alpha = 20\%$ or at least 3 H_0 rejected at $\alpha = 30\%$ or $\xi < -0.15$ or $\xi > 0.19$
good	otherwise

Table 4: Set of rules defining the verdict. Where H_0 denotes the null hypothesis of any of the tests given in table 5 (i.e., H_0: empirical and estimated distribution are the same).

Root-mean-square-error	$RMSE = \sqrt{1/n \sum \|F(x) - S_n(x)\|^2}$	$RMSE$
Kolmogorov-Smirnoff	$D_n = \sup_x \|F(x) - S_n(x)\|$	KS
Cramer-von Mises	$W_n^2 = n \int_{-\infty}^{\infty} \|f(x) - S_n(x)\|^2 dF(x)$	W_n^2
Anderson-Darling	$A_n^2 = n \int_{-\infty}^{\infty} \frac{\|F(x)-S_n(x)\|^2}{F(x)\{1-F(x)\}} dF(x)$	A_n^2
Right-tail Anderson-Darling	$R_n^2 = n \int_{-\infty}^{\infty} \frac{\|F(x)-S_n(x)\|^2}{\{1-F(x)\}} dF(x)$	R_n^2
Second order right-tail Anderson-Darling	$r_n^2 = n \int_{-\infty}^{\infty} \frac{\|F(x)-S_n(x)\|^2}{\{1-F(x)\}^2} dF(x)$	r_n^2

Table 5: Empirical distribution statistics for goodness-of-fit assessment, based on *Luceno* (2006).

4.4 Return levels

Return levels are the name given to extreme quantiles. A return level to return period m corresponds to a quantile with probability $1 - 1/m$. The equation defining return levels expressed in terms of the GEV parameters is given in definition 7.

Definition 7 (return level)**.** z_p to return period $m = 1/p$ is given by

$$z_p = \begin{cases} \mu - \frac{\sigma}{\xi}[1 - \{-\log(1-p)\}^{-\xi}], & \text{for } \xi \neq 0 \\ \mu - \sigma \log\{-\log(1-p)\}, & \text{for } \xi = 0. \end{cases}$$

where $\Pr(M_n \leq z_p) \approx G(z_p) = 1 - p$, with G a member of the GEV family.

As seen in figures 2 and 12, displaying the probabilities on the original scale of the data does not give sufficient information on the tails of the distribution. Figure 23 illustrates why the return level plot (see definition 8) is an intuitive way to present extreme value statistics. It focuses on the tail behavior of the distribution, by log-log transforming the x-axis. Without any additional tools, one can see if the distribution is bounded or heavy tailed: if the function has a positive curvature (blue line), the distribution is heavy tailed; if the function has a negative curvature (red line) the function is bounded. In case the function is linear (black line), the distribution belongs to the Gumbel family.

Definition 8 (return level plot)**.**

$$\{(log(y_p), z_p) : 0 < p < 1\}$$

with $y_p = -log(1-p)$.

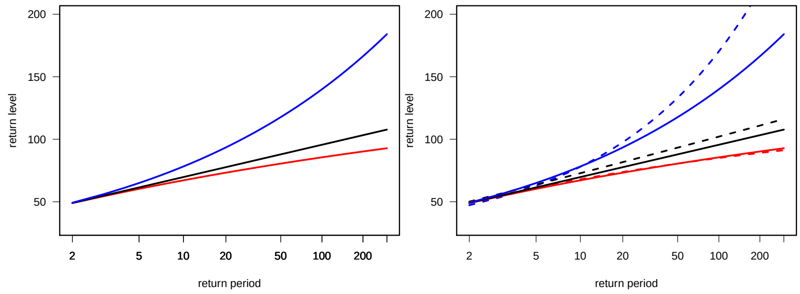


Figure 23: Theoretical return levels (solid lines) of Weibull (red), Gumbel (black), and Fréchet (blue) distributions (figure 13). Right panel: same as left with the estimated return levels (dashed lines - right panel) of the respective fitted distributions (figure 15).

4.4.1 Return level estimation

The estimation of the return levels is straightforward if the parameters of the distribution are already estimated: one simply replaces the parameters of equation 7 with their respective estimates and computes the return level of interest.

In general, $\Pr(M_n \le z_p) \approx 1 - p$ has to be solved for z_p. In some cases, for instance when non-stationarities are modeled explicitly, this must be done numerically as no analytic expression exists.

4.4.2 Uncertainty estimation

As for the parameter uncertainties, the uncertainties of the return level estimates can usually be estimated using the profile likelihood function.

Definition 9. Profile log-likelihood confidence intervals for z_p.
For uncertainty estimates of $\hat{z_p}$, one needs to reparametrize the likelihood function. For instance one can replace μ with $z_p + \frac{\sigma}{\xi}\left[1 - \{-\log(1-p)\}^{-\xi}\right]$ in the likelihood function. Then one maximises the likelihood function with respect to σ and ξ given the observations. Approximate confidence intervals can be derived via the χ^2 distribution.

We see in figure 24 that the CIs are not symmetric around the return level estimates: the smaller the probability, i.e., the longer the return period, the less symmetric the CIs. It is important to keep in mind that the uncertainties can only be read in one direction: the confidence interval describes the uncertainty of the return level, and not of the return period.

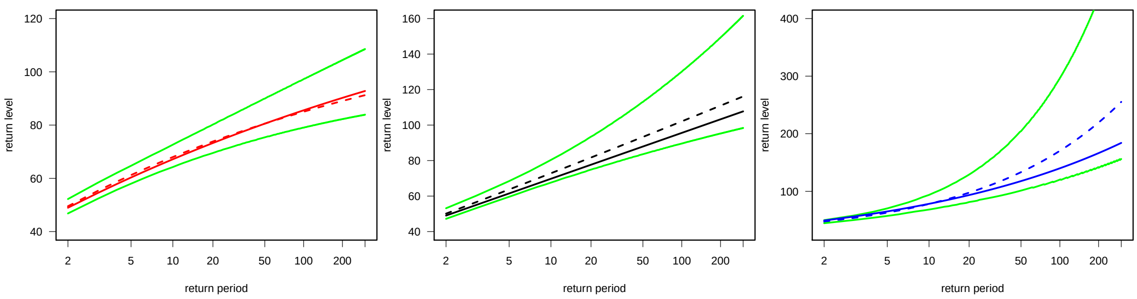


Figure 24: Return level plots of theoretical distribution (solid line), and fitted distribution (figure 15) from samples (figure 14) of size m = 100 (dashed lines) with 95% CI via profile likelihood function (green lines).

Should no analytic expression of the return level exist (as is the case when non-stationarities are modeled explicitly), uncertainty estimates can be obtained by bootstrapping methods.

4.5 Extreme Value Analysis example

Step by step extreme value analysis of 1-day precipitation recorded in Zürich / Fluntern.

Figure 25 shows the daily precipitation recorded during 40 years at Zurich / Fluntern.

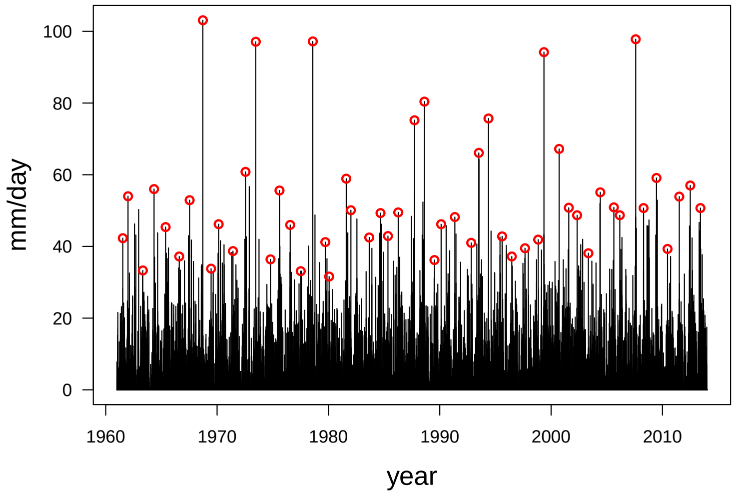


Figure 25: Daily precipitation time-series recorded at station Zurich / Fluntern between 01.01.1961 and 31.12.2013. The largest values per year are circled in red.

For the block maxima approach we need to extract the annual maxima (figure 26).

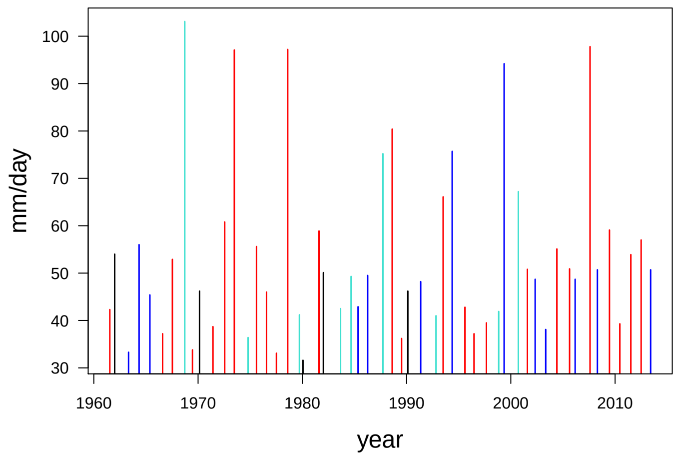


Figure 26: Time-series of annual maxima recorded at station Zürich / Fluntern between 01.01.1961 and 31.12.2013. The climatological season in which the maximum occurs is indicated by color: December, January, February (DJF) - black, March, April, May (MAM) - blue, June, July, August (JJA) - red, and (September, October, November) SON - turquoise.

As we can see, most annual maxima occur during the summer season (table 6).

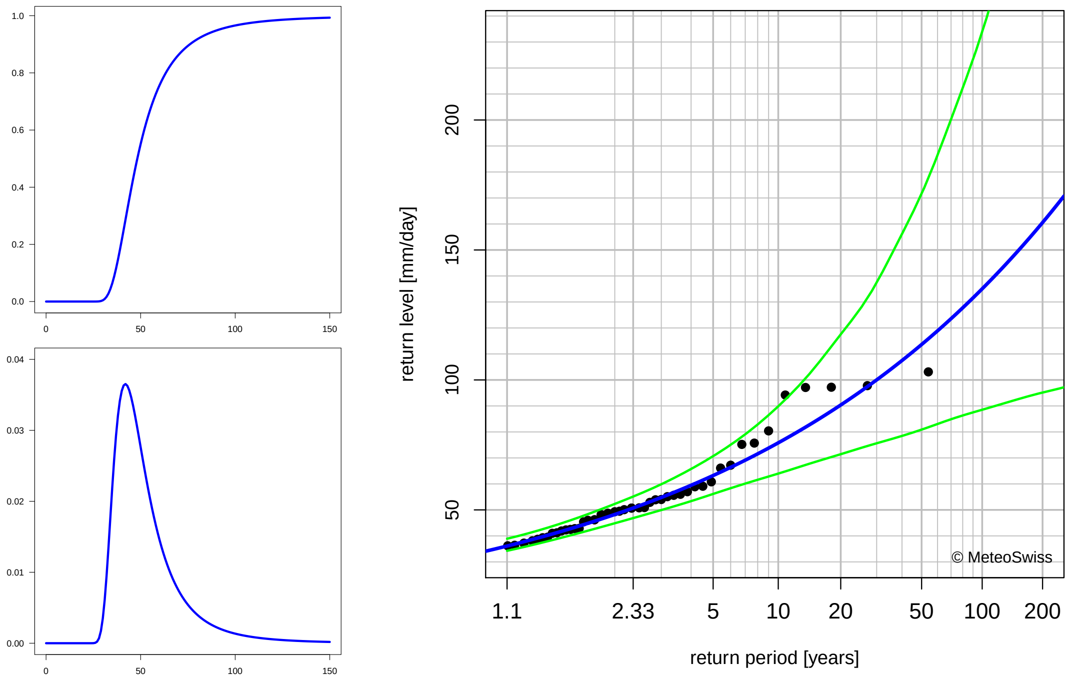
rank	amount	date	rank	amount	date
1	103.1 mm	1968-09-21	6	80.4 mm	1988-08-15
2	97.8 mm	2007-08-08	7	75.7 mm	1987-09-25
3	97.2 mm	1978-08-07	8	67.2 mm	2000-09-20
4	97.1 mm	1973-06-23	9	66.1 mm	1993-07-05
5	94.2 mm	1999-05-12	10	60.8 mm	1972-07-21

Table 6: Top 10 annual maxima with the occurrence date, ordered by decreasing amount. The climatological season in which the maximum occurs is indicated by color: DJF - black, MAM - blue, JJA - red, and SON - turquoise.

To fit the GEV we plug all annual maxima into the negative log-likelihood function and use a numerical optimization routine (in most cases Nelder-Mead) to find the combination of location, scale and shape parameters that maximize the likelihood surface. The negative log-likelihood maximum of **215.4663** is reached at location: **44.25**, scale: **10.39** and shape: **0.26**. These are our maximum likelihood estimates for the GEV parameters. The corresponding standard errors are 1.65, 1.36, and 0.13, respectively.

We have already discussed the goodness-of-fit of the statistical model in chapter 2, where we showed the QQ-plot (figure 9).

The corresponding cumulative probability function and probability density function are shown in the left panel of figure 27a.



(a) Cumulative probability function (top) and probability density function (bottom).

(b) Return level plot as presented in the extreme value analysis information of 1-day precipitation at station Zürich / Fluntern with 95% confidence interval based on the profile likelihood.

As we are interested in the extremes, we calculate return levels (figure 27b) with their confidence intervals.

Coming back to the top 10 events, we can now calculate the probability of exceeding these 10 events. We again express the probability in terms of return periods.

Table 7: Top 10 annual maxima date plus amount and estimated return period.

date	amount	return period	date	amount	return period
1968-09-21	103.1mm	34 years	1988-08-15	80.4 mm	12 years
2007-08-08	97.8 mm	27 years	1987-09-25	75.7 mm	10 years
1978-08-07	97.2 mm	27 years	2000-09-20	67.2 mm	6 years
1973-06-23	97.1 mm	27 years	1993-07-05	66.1 mm	6 years
1999-05-12	94.2 mm	24 years	1972-07-21	60.8 mm	4 years

Abbreviations

CI	confidence interval
EVS	extreme value statistics
EVT	extreme value theory
GEV	generalized extreme value
GPD	generalized Pareto distribution
ℓ	log-likelihood function
mle	maximum likelihood estimation
PP	probability-probability
QQ	quantile-quantile

station abbreviation	station name	longitude	latitude	altitude [m]
ALT	Altdorf	8.62	46.89	438.00
ANT	Andermatt	8.58	46.63	1438.00
BAS	Basel / Binningen	7.58	47.54	316.17
BER	Bern / Zollikofen	7.46	46.99	552.84
CDF	La Chaux-de-Fonds	6.79	47.08	1018.00
CHD	Château-d'Oex	7.14	46.48	1029.00
CHM	Chaumont	6.98	47.05	1136.00
DAV	Davos	9.84	46.81	1594.16
ELM	Elm	9.18	46.92	958.00
ENG	Engelberg	8.41	46.82	1035.66
GRC	Grächen	7.84	46.20	1605.00
GRH	Grimsel Hospiz	8.33	46.57	1980.00
GSB	Col du Grand St-Bernard	7.17	45.87	2472.00
GVE	Genève-Cointrin	6.13	46.25	420.00
LUG	Lugano	8.96	46.00	273.00
LUZ	Luzern	8.30	47.04	454.00
MER	Meiringen	8.17	46.73	588.60
NEU	Neuchâtel	6.95	47.00	485.00
OTL	Locarno / Monti	8.79	46.17	366.75
PAY	Payerne	6.94	46.81	490.00
RAG	Bad Ragaz	9.50	47.02	496.00
SAE	Säntis	9.34	47.25	2502.00
SBE	S. Bernardino	9.18	46.46	1638.70
SIA	Segl-Maria	9.76	46.43	1804.00
SIO	Sion	7.33	46.22	482.00
SMA	Zürich / Fluntern	8.57	47.38	555.95
STG	St. Gallen	9.40	47.43	775.66

List of Figures

List of Tables

References

Begert, M., G. Seiz, N. Foppa, T. Schlegel, C. Appenzeller, and G. Müller (2007), Die Überführung der klimatologischen Referenzstationen der Schweiz in das Swiss National Climatological Network (Swiss NBCN), *Technical Report MeteoSwiss, 215*, 43 pp.

Coles, S. (2001), *An introduction to statistical modelling of extreme values*, Springer Series in Statistics.

Fisher, R., and L. Tippett (1928), On the estimation of the frequency distributions of the largest or smalles member of a sample, *Proceedings of the Cambridge Philosophical Society*, *24*, 180 – 190.

Gnedenko, B. (1943), Sur la distribution limite du terme maximum d'une série aléatoire, *Annals of Mathematics*, *44*, 423 – 453.

Hosking, J. (1990), L-moments: analysis and estimation of distributions using linear combinations of order statistics, *J. R. Statist. Soc. B*, *52*(1), 105–124.

Jenkinson, A. (1955), The frequency distribution of the annual maximum (or minimum) values of meteorological events, *Quarterly Journal ofthe Royal Meteorological Society*, *81*, 158–172.

Leadbetter, M. R., G. Lindgren, and H. Rootzén (1983), *Extremes and related properties of random sequences and processes*, Springer Verlag, New York.

Luceno, A. (2006), Fitting the generalized pareto distribution to data using maximum goodness-of-fit estimators, *Computational Statistics and Data Analysis*, *51*, 904–917.

Röthlisberger, H., J. Zeller, and H. Geiger (1991), Starkniederschläge des schweizerischen alpen- und alpenrandgebietes.

von Mises, R. (1954), *La distribution de la plus grande de n valeurs.*, vol. II, pp. 271–294, American Mathematical Society, Providence, RI.

Acknowledgment

We thank Thomas Schlegel and Stephan Bader for sharing their knowledge of the Swiss Climate in numerous discussions.

MeteoSchweiz
Operation Center 1
CH-8058 Zürich-Flughafen

T +41 58 460 91 11
www.meteoschweiz.ch

MeteoSvizzera
Via ai Monti 146
CH-6605 Locarno Monti

T +41 91 756 23 11
www.meteosvizzera.ch

MétéoSuisse
7bis, av. de la Paix
CH-1211 Genève 2

T +41 22 716 28 28
www.meteosuisse.ch

MétéoSuisse
Chemin de l'Aérologie
CH-1530 Payerne

T +41 26 662 62 11
www.meteosuisse.ch