



High-resolution analysis of 1 day extreme precipitation in Sicily

M. Maugeri^{1,2}, M. Brunetti², M. Garzoglio¹, and C. Simolo²

¹Department of Physics, Università degli Studi di Milano, Milan, Italy

²Institute of Atmospheric Sciences and Climate, CNR, Bologna, Italy

Correspondence to: M. Maugeri (maurizio.maugeri@unimi.it)

Received: 27 February 2015 – Published in Nat. Hazards Earth Syst. Sci. Discuss.: 2 April 2015

Revised: 2 October 2015 – Accepted: 6 October 2015 – Published: 15 October 2015

Abstract. Sicily, a major Mediterranean island, has experienced several exceptional precipitation episodes and floods during the last century, with serious damage to human life and the environment. Long-term, rational planning of urban development is indispensable to protect the population and to avoid huge economic losses in the future. This requires a thorough knowledge of the distributional features of extreme precipitation over the complex territory of Sicily.

In this study, we perform a detailed investigation of observed 1 day precipitation extremes and their frequency distribution, based on a dense data set of high-quality, homogenized station records in 1921–2005. We estimate very high quantiles (return levels) corresponding to 10-, 50- and 100-year return periods, as predicted by a generalized extreme value distribution. Return level estimates are produced on a regular high-resolution grid (30 arcsec) using a variant of regional frequency analysis combined with regression techniques. Results clearly reflect the complexity of this region, and show the high vulnerability of its eastern and northeastern parts as those prone to the most intense and potentially damaging events.

The statistical analysis of hydrological extremes is usually based on extreme value theory (EVT, see, e.g. Leadbetter et al., 1983; Coles, 2001 and references therein), which sets a theoretical basis for the inference of an asymptotic distribution and the extrapolation of relevant properties thereof. An example is the estimation of very high distribution quantiles occurring e.g. once every 100 years or even more (return levels, RLs) that cannot be extracted from the observed frequencies because they are extremely rare – or even never recorded – but yet potentially damaging. Since long time series of e.g. daily or sub-daily precipitation records are hardly available, in practice we often deal with rather small data samples. This complicates the application of EVT approaches, such as fitting a generalized extreme value (GEV) distribution to block data (e.g. annual maxima), and points to the need for efficient methods of parameters estimation with small samples (Martins and Stedinger, 2000).

The use of a generalized Pareto (GP) distribution to fit partial-duration data, and more generally the use of peaks-over-threshold (POT) approaches (Davison and Smith, 1990; Madsen et al., 1997a), has the advantage over the block data approach of including more data in the analysis, although it brings an additional uncertainty in the fitting procedure, related to the choice of the threshold. Further, since partial-duration data are supposed to cluster in time, some selection criteria for identifying independent events are generally applied (Katz et al., 2002). Nevertheless, potential benefits and drawbacks from the use of any EVT model should be carefully evaluated, in connection with the nature of the problem investigated, the region examined and the available data.

Several techniques have been conceived to overcome the problem posed by the inherent rareness of extreme events, either for improving fits of EVT parameters to small data samples – e.g. methods using linear (L-) moments (Landwehr et

1 Introduction

Sensitivity of the Mediterranean land areas to heavy precipitation and floods, and their responsiveness to climate change, is well documented (see e.g. Luterbacher et al., 2006; Ulbrich et al., 2012). Partly due to the excessive exploitation of natural resources and soil, the exposure of the Mediterranean area to hydrometeorological hazards is supposed to increase in future scenarios (Giorgi and Lionello, 2008; Toreti et al., 2013), and requires in-depth research of present-day characteristics of precipitation extremes over this region.

al., 1979; Hosking, 1990) – or for magnifying fitting samples themselves. This strategy is used in the regional frequency analysis (RFA), originally developed in the context of hydrology (see, e.g. Hosking and Wallis, 1997; Buishand, 1991; Madsen et al., 1997b), to reduce the large errors on fit parameters and any resulting quantity. RFA prescribes to enlarge fitting samples by pooling extreme data from more than one measurement site, provided that observational records from the gathered sites show the same distributional features after a convenient rescaling. A common choice for the scaling factor, named index flood (IF), is the site-specific sample mean of extreme data, though any indicator of their central tendency may work (Hosking and Wallis, 1997). An important advantage of RFA is that it can be used to calculate RLs at an ungauged site, provided that the corresponding IF can be estimated: RLs are obtained by first estimating the distributional features of the rescaled data by means of the neighbouring measuring sites and second by using the IF to restore the physical dimensions of the RLs (Hosking and Wallis, 1997).

The basic idea of RFA has been implemented into many variants, and in conjunction with L-moments. For instance, the regional L-moment algorithm (Hosking and Wallis, 1997) prescribes computing regional EVT parameters from weighted averages of L-moment-derived station parameters across many similar sites, rather than performing a straightforward fit of a pooled sample of extreme data (station-year method, Buishand, 1991). Some questionable aspects of RFA have been raised, concerning residual spatial heterogeneities and non-negligible correlations within regions, even though advantages from regional over at-site procedures have generally emerged (Hosking and Wallis, 1988; Madsen and Rosbjerg, 1997). RFA methods have frequently been used in regional studies (e.g. Fowler and Kilsby, 2003; Kjeldsen and Jones, 2009; Hanel et al., 2009; Svensson and Jones, 2010; Roth et al., 2012, 2014; Jones et al., 2013), for extracting RLs of a GEV/GP distribution fitted to extreme precipitations from both observations and models, either in stationary conditions or non-stationary – i.e. with EVT parameters adjusted for temporal evolution.

Several studies have attempted assessments of present-day precipitation extremes and their possible changes over the Mediterranean area, either by empirical or (mainly at-site) EVT approaches (see e.g. Brunetti et al., 2001, 2004; Alpert et al., 2002; Kostopoulou and Jones, 2005; Moberg et al., 2006; Brunet et al., 2007; Kioutsioukis et al., 2010; Rodrigo, 2010; Toreti et al., 2010, van den Besselaar et al., 2013). The temporal tendencies are not fully coherent throughout the region (Ulbrich et al., 2012) and rather conditioned by the specific site, the used approach and the period examined (Brugnara et al., 2012; Brunetti et al., 2012).

As far as Sicily is concerned, a first extensive study of yearly maxima of 1 day and sub-daily (1, 3, 6 and 12 h periods) precipitation was performed by Cannarozzo et al. (1995) using a RFA approach based on a two-component extreme value distribution (Rossi et al., 1984). Then the same anal-

ysis was applied to an extended database by D'Asaro and Grillone (2008) which considered also GEV distribution using a unique set of parameters over the entire region of Sicily and for all periods from 3 to 24 h. This latter approach, in our opinion, does not fully capture the spatial variability of extreme precipitation over Sicily, which represents an outstanding example of complex terrain for orographic and land–sea contrasts, with highly space- and time-localized heavy precipitation. Both Canarozzo et al. (1995) and D'Asaro and Grillone (2008) provide spatial maps of the IF at ungauged sites that allow estimating the RLs for any point of Sicilian territory. These maps however do not represent the high spatial variability of the IF over Sicily, as they are based on interpolation techniques that do not take into account the dependence of precipitation on elevation.

In this work, we address this issue using 1 day precipitation only, as daily data have better availability (231 records with average length of 56 years) than sub-daily ones. Moreover, 1 day precipitation is available from the yearbooks of the Italian Hydrographic Service with daily resolution, whereas for sub-daily periods, just yearly maxima are available. The aim of our study is to provide high-resolution maps of the sensitivity of Sicily to the most destructive events.

We performed therefore a detailed study of RLs in 1 day precipitation extremes on a high-resolution grid (30 arcsec), to identify the parts of Sicily especially exposed to the most intense events. Specifically, we used a GEV approach with RFA to fit annual maxima from observational precipitation records, assuming stationarity. RFA was implemented in the station-year variant to enlarge fit samples at any grid point, after rescaling of the annual maxima of each station by their site-specific median (the IF). RLs corresponding to 10-, 50-, and 100-year return periods (RL10, RL50 and RL100, respectively) with related uncertainties were extrapolated from the GEV in the rescaled form. Finally, the absolute precipitation amounts (mm) of grid-point RLs were obtained by exploiting the strict connection between IFs and mean annual totals, which are available from Brunetti et al. (2014) at a 30 arcsec resolution.

Compared to the above-mentioned studies on the same issue, some improvements contained in the present work should be pointed out. These mainly concern the availability of a wider database that allows a detailed check of quality and homogeneity of records, the use of a new pooling procedure that captures the spatial variability of GEVs parameters better and the estimation of the spatial distribution of the IF from a high-resolution climatology for Sicily, which has been developed by exploiting the dependence of precipitation on orography (Brunetti et al., 2009, 2014). From a purely scientific perspective, the most relevant aspect of this study consists of combining well-established EVT approaches with regression techniques for the spatial interpolation of the station IFs at a high resolution. This is a rather innovative topic that,

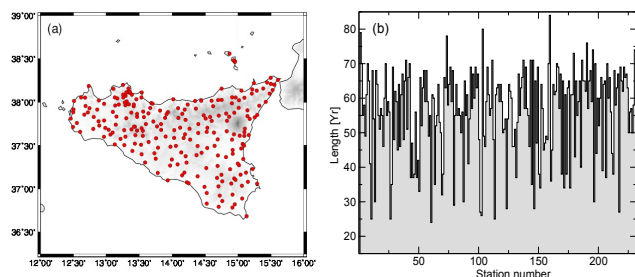


Figure 1. (a) Location of station sites. (b) Number of years with at least 90 % of valid data per station record.

to our knowledge, has been rarely addressed (see e.g. Carreau et al., 2013).

Observations and their processing are described in Sect. 2. Details of the methods are given in Sect. 3. Some intermediate results and the high-resolution maps of RLs are presented in Sect. 4. Finally, conclusions are drawn in Sect. 5.

2 Data processing

The observational network for this study integrates daily precipitation series in 1921–2005 from various sources, i.e. the Italian National Air Force, the Agricultural Research Council, historical observatories and for the most part, the “Osservatorio delle Acque” (the former Hydrografic Service), for a total of 325 station records endowed with validated geographic coordinates. In case of multiple records from the same measuring site the longest and/or the most complete series of data was preferred.

Since extremal analysis is expected to be highly sensitive to potential errors in daily records, raw data have been thoroughly processed to isolate anomalously large precipitation amounts (outliers) and spurious long dry spells, by exploiting local coherence of precipitation episodes.

Specifically, as with outliers, single daily events of each station record were routinely compared with the average values of the events measured at the 10 nearest stations in a 3 day window about that day to account for possible 1 day lags between neighbouring sites. Sensible bounds imposed on the differences and ratios between the test and the 3 day reference value were: (i) absolute difference from each of the reference values in the 3 day window higher than 60 mm; (ii) ratio to each of the reference values in the 3 day window higher than 6 and (iii) ratio to the absolute maximum across all values in the 3 day window from the individual 10 closest stations (i.e. not averaged) higher than 3. These bounds should be conceived as a trade-off between the risk of retaining unreliable data and that of selecting an excessive number of potential outliers. The quality check uncovered about 300 questionable events across the entire data set, i.e. with out-of-bounds precipitation amounts. Once individually examined with auxiliary tools (e.g. yearbooks, weather maps, old lo-

cal newspapers), fewer than 10 % of the uncertain events remained unconfirmed and were therefore cancelled. Yet, few of these values were identified as monthly amounts, thereby forcing cancellation of entire years (6 years in three station records).

Likewise, dry spells with more than 60 consecutive days in one station record were marked as doubtful whenever no track of similar dry episodes (i.e. at least 90 % of dry days during that time interval and precipitation below 5 mm for the rest) appeared at the 10 nearest measuring sites. Several blocks of missing data masked by dry days were uncovered, and therefore either consecutive months or years were cancelled (59 months and 12 years in 20 station records).

Next, we assessed the homogeneity level of quality-controlled records covering at least 24 years, and, when required, we subjected them to homogenization. The lower bound on the series length was set to balance the need for enough data to perform reliable corrections and the risk of wasting valuable information for extremal analysis. Further, only genuinely homogeneous series were retained if their record length was below 30 years.

Artificial discontinuities in the daily series, generally originated by station relocation and/or instrument malfunctions, were distinguished from real climatic signals using a multiple application of the Craddock test (Craddock, 1979) described in detail in previous works (Brunetti et al., 2004). In practice, records in subgroups of 10 elements were mutually controlled for both precipitation amount and number of wet days. Inhomogeneous periods detected in one precipitation amount record were corrected by a scaling factor appropriately derived from the precipitation amount of neighbouring station records that were previously marked as homogeneous. Conversely, station records with either several intractable inhomogeneities in precipitation amount or unrecoverable inhomogeneities in the number of wet days (regardless of the homogenization of the corresponding precipitation amount) were either entirely or partly dismissed.

Ultimately, 231 station records passed both the length-based selection and the homogeneity control. Of these, 133 are homogenized records, with generally one or two readjusted periods. The spatial distribution of the selected records is shown in Fig. 1a, whereas their data availability is shown in Fig. 1b. Specifically, Fig. 1b gives evidence of the number of years with at least 90 % of valid data in each station record: they are the years we consider in this paper. About 56 nearly complete years per station are available on the mean, albeit a major gap in the years 1942–1950 interrupts nearly all records.

The question of whether the homogenization process could spoil the estimation of extreme quantiles – by either exaggerating heavy precipitation or downsizing true extremes – required an in-depth examination. This issue was investigated a posteriori, by comparing GEV-estimated RL50s at a given station estimated from a pooled sample of records from the nearest sites – excluding that station – with those esti-

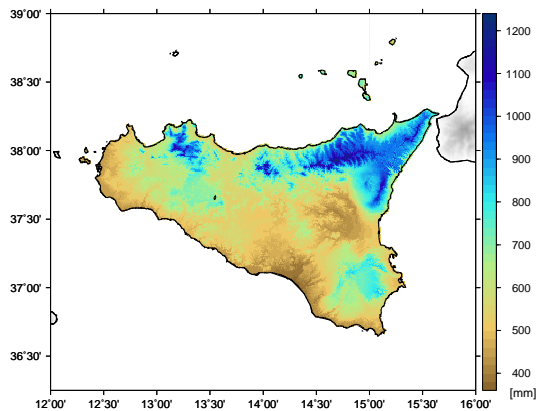


Figure 2. Climatic normals (1961–1990) of annual precipitation totals for Sicily at 30 arcsec resolution.

mated from both the homogenized station record alone and its original counterpart. Results are discussed in Sect. 4.2.

Finally, when going from the station-level to gridded analysis we used the 30 arcsec resolution digital elevation model (DEM) GTOPO30 provided by the United States Geological Survey (USGS, 1996) and restricted it to the frame 12.4–15.7° E to 36.5–38.5° N. As described in Sect. 3, annual precipitation normals on the same grid were required as an intermediate step of the grid-point estimation of RLs. Gridded normals were obtained by means of a local weighted linear regression of precipitation versus elevation, as described in detail in Brunetti et al. (2009, 2014), and are shown in Fig. 2. The interpolation procedure used here is based on the PRISM (parameter elevation regression on independent slopes model) conceptual framework. Complete documentation on PRISM is available on the website of the PRISM climate group (<http://www.prism.oregonstate.edu>) as well as in a number of papers (see e.g. Daly et al., 1994, 2008).

3 Method

Grid-point RL10, RL50 and RL100 are the final outcome of the algorithm we used in this work. It consists of four steps:

- i. division of annual maxima from single-station records by an appropriate, site-specific IF (*rescaling*);
- ii. assignment to each grid point of an enhanced sample of rescaled maxima drawn from the station records falling within the grid-point neighbourhood, the so-called region of influence (*pooling*);
- iii. fit of a unique GEV distribution to each grid-point sample and extrapolation of non-dimensional RLs for the given return periods (*fitting*);
- iv. interpolation of site-specific IFs on the high-resolution grid by regression with annual mean precipitation totals and estimation of dimensional RLs (*spatialization*).

i. As described above, the annual maximum of 1 day precipitation was extracted only from nearly complete years (at least 90 % of valid data) to obtain a representative sequence of block maxima for any given station site, although with several gaps. The above sequences were divided by their own median, taken as the IF, because the median is a robust estimator of central tendency in extreme data samples. This was done to help compare frequency distributions of precipitation maxima from different sites. Indeed, rescaling may unmask similarities in subsets of data that could be exploited for improving extremal analysis. According to this technique, the whole procedure was carried out in terms of rescaled data except for the very end of the operations chain when physical dimensions of grid-point RLs were restored.

- ii. Each grid point over Sicily was set as the centre of a search area with a 25 km radius. Rescaled maxima from the station sites falling within that grid-point neighbourhood were merged into a unique data sample for the subsequent operations, according to the logic of RFA in the station-year variant.

A key point here is defining the optimal size of the region of influence of each given site, as this should balance the scope of reducing fitting errors and the need of maintaining similarities in data distributions within the same region. We investigated this issue by considering the agreement of the single-station and the pooled records at the station sites. Specifically, the upper bound on the pooling distance was selected after comparison of the end results – RL50 with 95 % confidence intervals – obtained by single-station and pooled-station records with varying radius in a 5–50 km range and a 5 km step. For a fair comparison, RL50 from pooled records was computed in the “leave-one-out” approach to exclude the influence of the given station itself. Performance at each step was evaluated by the pooled uncertainty measure (PUM) (Kjeldsen and Jones, 2009) that quantifies the average deviation of the pooled estimate from the respective single-station value over all stations, i.e.

$$\text{PUM} = \left(\frac{\sum h_i (\ln z_i - \ln \hat{z}_i)^2}{\sum h_i} \right)^{1/2}, \quad h_i = \frac{n_i}{1 + n_i/16}. \quad (1)$$

Here, z_i and \hat{z}_i are the single-station and the pooled-station quantile under study, respectively, n_i the station record length (years) and the index i runs over all stations. The PUM minimum occurring at about 25 km (see Fig. 3a) identified the optimal upper bound on the pooling distance. This was used in the following to define the region of influence of each grid point.

Nevertheless, one caveat is that serial correlations actually limit the pooling enhancement of the sample size and may further bias fit results if many replicas of the

same event from nearby locations occur in the pooled sample. Thus, a minimum distance between station sites (lower bound) should also be kept to obtain an independent sample. The appropriate lower bound on the pooling distance was defined, again at the station level, using the collective measure given by Eq. (1). Specifically, at any step in the search range 0–15 km, only different-year maxima from any couple of stations with below-threshold distance were retained, so as to enhance data independence in the pooled sample. Setting the lower band at 5 km was the best compromise both to keep PUM and mean error band low (amplitude of the confidence interval at the 95 % level, CI95) and to maintain mean sample size high enough (Fig. 3b, d). Higher values of the lower band would entail sensible loss of data; conversely, values below 5 km would not reduce errors considerably. In addition, the above choice well agrees with the decay distance of common variance between series of annual maxima, as the latter falls below 0.5 for distances starting as low as 2.5 km.

The optimal grid-point sample for the fit of a GEV distribution was finally given by all rescaled maxima from the station sites falling within the upper bound, excluding replicas from sites below the lower bound. On the mean, the fit samples were magnified with respect to the single station records by about a factor 10, with more than 700 data per region with very few exceptions concerning quite isolated areas – e.g. some coastal sites with fewer than 300 data and the north-eastern islets with only about 70 data.

Prior to the fit, pooled samples were further investigated at the station level to check for distributional similarities between merged records, evaluate the impact of data homogenization and their temporal stationarity. All these findings are relevant for a correct interpretation of the end results and are detailed in Sect. 4.

- iii. Extrapolation of very high quantiles from the observed frequency of precipitation maxima goes through the definition of a GEV model to fit the data, enabling predictions of the rarest events even though they are not yet observed. The GEV distribution function has the well-known form:

$$G(z) = \exp \left\{ - \left[1 + \xi \left(\frac{z - \mu}{\sigma} \right) \right]^{-1/\xi} \right\}, \quad (2)$$

for all z such that $1 + \xi(z - \mu)/\sigma > 0$. The time-independent (see Sect. 4.3) shape parameter ξ , scale parameter σ and location parameter μ were directly estimated for any grid point on the pooled sample of rescaled maxima defined as above, using the maximum likelihood method. The 95 % confidence interval for each parameter was obtained by constraining its deviance function below the 0.95 quantile of a

χ_1^2 distribution with 1 degree of freedom (Coles, 2001). Then, the RL z_p associated with a small probability p or, equivalently, with a return period $T = 1/p$ was obtained by inverting the GEV in Eq. (2), i.e. $z_p = G^{-1}$, where $G = 1 - p$ and the corresponding confidence interval was obtained as for the GEV parameters (Coles, 2001).

GEV is a well-established distribution for extremal analysis because it has generally high performance in the fit of block maxima. Indeed, a goodness-of-fit test based on the sample moments of the frequency distribution (Z test, see Hosking and Wallis, 1997) yielded a 95 % acceptance rate, even on single-station records. Some detailed results about the fit performance of the GEV model and the behaviour of parameters from both single-station and pooled records are discussed further in Sect. 4.4.

Finally, notice that the so-called regional L-moment algorithm (Hosking and Wallis, 1997) can be used alternatively to the straightforward fit of a unique GEV to the pooled sample. In fact, the L-moment algorithm and the station-year method yield nearly equal results in the present case. Indeed, the GEV parameters obtained with the L-moment algorithm fall well within the error band of the station-year parameters, with difference between the two sets generally below one half of the CI95 amplitude. Yet, this might not be true in all situations, and the appropriateness of the inference method should be evaluated case by case.

- iv. Converting non-dimensional grid-point RLs to millimeter values requires re-multiplication of the above quantities by the respective median of precipitation maxima – the IF – that must be evaluated at each grid point. To this end, station IFs were interpolated on the regular grid by a local linear regression with mean annual precipitation totals (normals, see Fig. 2), that correlate well spatially with the former quantities. Specifically, for each grid point, a linear regression was performed between the IFs and precipitation yearly normals using the station values from the nearest sites to the given grid point. The search radius about the grid point extended from 10 to 25 km, until at least seven stations were found. Grid-point IFs were then estimated from the grid-point precipitation normals and the fit coefficients.

The goodness of fit was evaluated by comparing the station IFs with those estimated on the grid points closest to the station sites, in the leave-one-out approach. The observed IFs are nicely reproduced by the fit, as shown in Fig. 4. Here, the pairs of estimated and observed values are seen to line up on the bisector, except for a few far-out points with deviations as large as 30 mm. These errors mainly originate from the sites with the highest observed IFs, lying on the eastern slope of Mount Etna

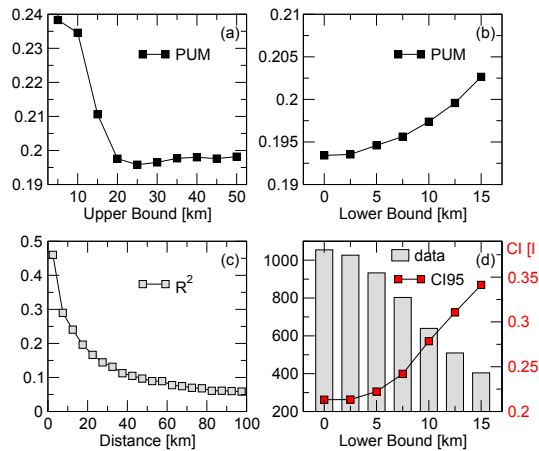


Figure 3. Variation of the RL50 PUM given by Eq. (1), as a function of (a) the upper and (b) the lower bound to the pooling radius. (c) Common variance between series of annual maxima versus distance of sites, computed at a 5 km step, starting from a minimum distance of 2.5 km. (d) Variation of both the mean number of data in the pooled samples and the mean amplitude of the CI95 as a function of the lower bound to the pooling radius.

where the number of high-elevation stations is critically low and the fit underestimates precipitation maxima. The total uncertainty on the grid-point IFs was measured by the root-mean-square error (RMSE) over all sites, that amounts to about 8 mm, i.e. 13 % of the observed mean value. This overall fit error should be then combined with the GEV confidence intervals accompanying rescaled RLs, when the final conversion to physical dimensions is made.

4 Results and discussion

Before going into detail about the estimation of grid-point RLs, we briefly discuss a set of results from preliminary analyses aimed at screening data for their aptitude to be processed with RFA and a GEV approach. Indeed, once the regions of influence of any grid point have been defined, it must be checked that frequency distributions of station maxima in any given region do exhibit the expected similarities (Sect. 4.1). In addition, some issues about data homogenization and temporal stationarity have been fully explored (Sect. 4.2–4.3), along with the behaviour of GEV parameters from both the single-station and the pooled-station fits (Sect. 4.4). Finally, high-resolution maps of RL10, RL50 and RL100 are discussed in detail (Sect. 4.5).

4.1 Tests for distribution similarities

Using appropriate combinations of sample moments, the main distributional features of rescaled annual maxima from station sites assigned to the same region were inspected altogether to assess the degree of similarity between records.

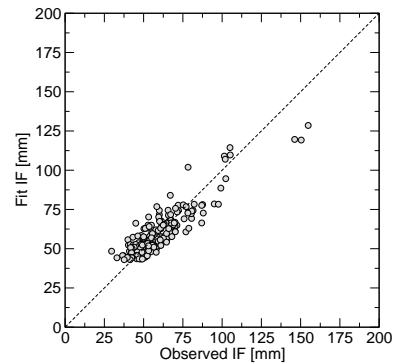


Figure 4. Estimated vs. observed IFs at the station sites.

First, following Hosking and Wallis (1997), a discordancy test was recursively applied at each station site to identify potential discrepancies between the L-moment ratios of any given site and the average L-moment ratios of the related region as a whole, possibly due to untraced errors in the data. Along with an overall agreement between records from neighbouring sites, a few critical cases were pointed out. These are due to exceptional events that truly appeared at one site only, and therefore all sites with outstanding discrepancies were retained.

Then, to satisfy the hypothesis underlying regional analysis, frequency distributions of annual maxima from sites assigned to the same region should be identical up to a site-specific scaling factor. To check this, we used the so-called heterogeneity test (H test, see Hosking and Wallis, 1997 for details), again based on sample L-moment ratios of rescaled maxima from neighbouring sites. Results of the H test recursively applied at any station point are quite satisfying over large areas, since only about 15 % of sites show marked heterogeneity, thereby suggesting region re-definition. Yet, as recommended by Hosking and Wallis (1997), results of the H test should be used with caution. Indeed, critical thresholds for the H -statistics cannot be regarded as strict rejection levels, given some arbitrariness in their definition and the test hypotheses – negligible serial and cross correlations, parent distribution exactly known – being only weakly satisfied. Rather, the above critical thresholds have been conceived here as a benchmark for further inspecting threshold-exceeding regions in conjunction with more physical arguments. Indeed, poor H test results concern mainly the northern part of Sicily where precipitation regimes are strongly influenced by the complex orography that naturally enhances differences between at-site distributional features. Shrinking regions to very few sites in this area is not beneficial for regional analysis since H values still remain high while the number of data per region becomes uselessly small. Keeping this limitation in mind, the size of the northern regions was thus retained as it is in the subsequent analysis.

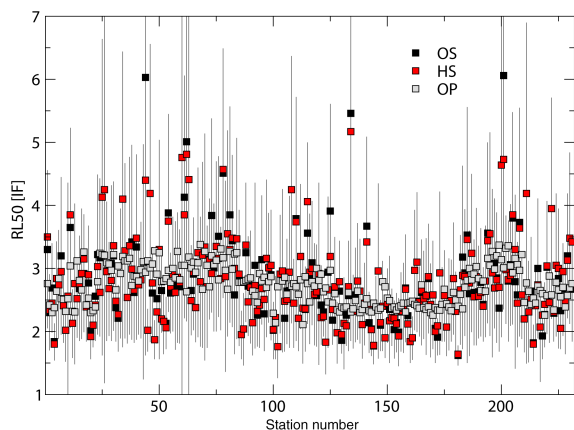


Figure 5. Rescaled RL50s (IF units) for any station site, obtained from both original (OS) and homogenized records (HS) at the single-station level, and compared with expected RL50 from the pooled-station samples (OP). Only error bars of OS RL50 are shown, as they largely overlap with HS errors and completely mask those of the OPs.

4.2 Data homogenization

As already mentioned in Sect. 2, benefits and/or drawbacks of data homogenization were thoroughly assessed according to their impact on the GEV-estimated RL50s. Specifically, these values were computed from the rescaled maxima of any single-station record, both original and homogenized, and compared to those obtained from the pooled sample of the related region deprived of the originating station itself, by a leave-one-out approach. In the pooled case, RL50s were obtained using RFA about any station sites as described in Sect. 3, for both original and homogenized data. Since results from pooled homogenized data (HP) are largely consistent with those from pooled original data (OP), only the latter were used for direct comparison with RL50s from both original (OS) and homogenized single-station records (HS). As seen in Fig. 5, HS RL50s generally compare better than OS RL50s (when they differ) with OP results, indicating that homogenization does remove gross errors and restores the true expected behaviour of the time series. Overall, the discrepancies between the RL50s from single-station and pooled data appear slightly reduced by data homogenization (by about 4%), with few outstanding estimates considerably downsized. Though homogenization proved to be beneficial to extreme data analysis in this context, one should be aware that this is not a universal result. Homogenization at the daily scale is still a matter of investigation, and the convenience of any homogenization technique largely depends on the characteristics of the data (e.g. record length, spatial density and meta-data availability). Therefore, the effect of homogenization on the extreme values is not obvious and should be examined on a case-by-case basis. In our case, one of the most important contributions of data homogenization

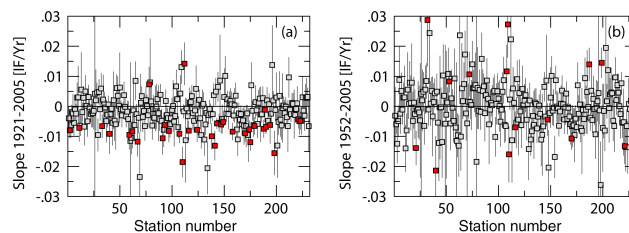


Figure 6. Trend slopes with 1-sigma error bars relative to (a) the full record period and (b) a relevant sub-period (1952–2005), for any station site in both cases. Red symbols denote significant trends (p value < 0.05) according to a Mann–Kendall test.

was the elimination of the initial part of some records that showed clear inhomogeneities in the number of wet days. This is probably because, in some cases, in the oldest yearbooks for Sicily, sequences of days with no rain followed by a rainy day actually correspond to a few days' cumulative precipitation amount.

4.3 Stationarity

The prerequisite for a stationary GEV-based approach is the absence of temporal trends in the station series of annual maxima. This issue was examined in-depth across the full period and a relevant sub-period of records since, as previously noted, data availability drops sharply during the 1942–1950 period. As shown in Fig. 6a, less than 15% of the station series of rescaled maxima exhibit a significant trend of decreasing extreme precipitation, when the full period of the series is considered. This tendency disappears when the analysis is carried out starting from 1952 (Fig. 6b), with only 6% of the series showing trends, either positive or negative evenly.

To some degree, the temporal discontinuity in data availability reflects the overall quality level of the records that, regardless of their detailed pre-processing, remains slightly lower in the early decades. Indeed, both the quality controls and homogenization probably removed only gross errors and breaks, whereas some dubious data gathering in the former period – e.g. possible few days' cumulative precipitation amounts – may persist and induce unexpected trends in a small number of series. In addition, well-known exceptional precipitations have occurred in the years 1931, 1933 and 1951, involving areas that are quite large. These events further enhance the above tendencies, since even a few data may considerably alter the trend analysis of extreme values. All things considered, a picture of essential stationarity of the annual maxima reasonably appears as the most plausible, and suggests the fitting parameters of the GEV to be kept constant in time.

It is interesting to observe that also Calabria, which is very close to Sicily and has data from the same yearbooks we used for Sicily, does not exhibit trends in the number of days and in the precipitation amount corresponding to the

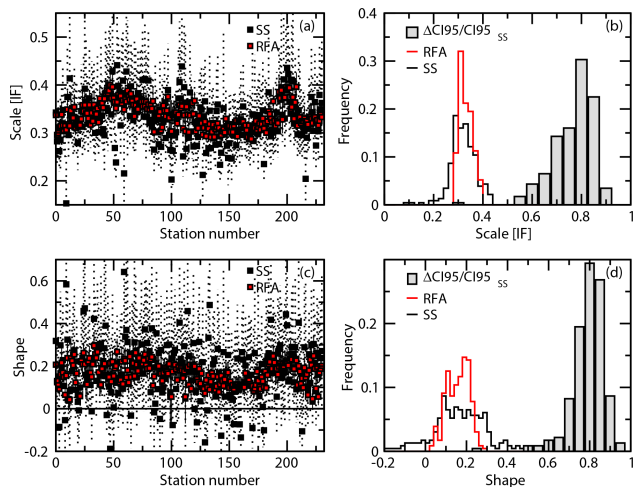


Figure 7. (a, c) Station-by-station GEV scale (IF units) and shape parameters, respectively, with 95 % confidence intervals from fits to single-station (SS) and pooled-station (RFA) data. Dotted bars denote SS errors and solid bars RFA errors. (b, d) Frequency histograms of the SS and RFA estimates of the scale and shape parameters, respectively, across the region, and the fractional reductions of CI95 by RFA (grey-filled bars) for both parameters.

highest precipitation intensity classes after about 1950 (see Figs. 11 and 12 in Brunetti et al., 2012), whereas before 1950 it shows a decrease of these variables.

4.4 GEV parameters from single-station and RFA fits

Here, we discuss the behaviour of the scale and shape parameters, when switching from a single-station analysis to pooled-station samples in the perspective of RFA. This technique tries to overcome the small-sample problem of extremes, thus to get more stable parameters' estimation and provide more useful results for practical applications.

A GEV distribution is fitted to rescaled annual maxima from each single-station record (SS) and the respective pooled-station sample (RFA), as described in Sect. 3. Resulting scale and shape parameters with 95 % confidence intervals in the two cases are compared station-by-station in Fig. 7a, c. Notice again that the alternative discussed in Sect. 3 (L-moment algorithm) to the straightforward parameters' fit to pooled-station samples is totally equivalent. As seen in Fig. 7b, d, RFA provides a net reduction of confidence intervals, up to 80 % in many cases. Likewise, central (best guess) values of RFA scale and shape parameters spread within a narrower band all over the region, compared to the respective SS estimates. As with the shape, in particular, the single-station estimates take on negative values in several cases contradicting the expected property of the GEV of precipitation maxima to be heavy-tailed (Buishand, 1991), i.e., Fréchet-type, with positive shape in the notation

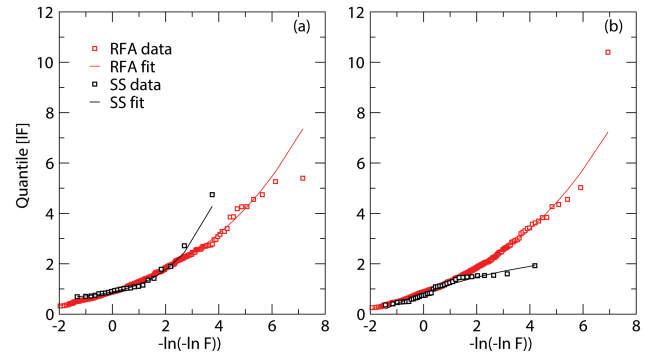


Figure 8. (a, b) Growth curves (IF units) for two exemplifying sites, with SS best guess value of the shape; (a) 0.58 (Diga Ragoletto) and (b) -0.19 (Sigonella). The relative frequency F on the abscissa is defined according to the Gringorten rule (Gringorten, 1963).

of Eq. (2). Conversely, RFA estimates stay always positive within the error bars.

By way of illustration, Fig. 8a, b show a couple of cases where GEV fits of SS give shape parameters far out from a reasonable range, i.e. the highest and the lowest value, respectively. Here, the quantile functions expressed in IF units (growth curves, Kjeldsen and Jones, 2009) obtained both for SS and RFA are compared to their empirical analogues, using a double-log scale for the relative frequencies on the horizontal axis (reduced Gumbel variate). The SS theoretical curves appear overly constrained by very few data, whereas RFA, being less sensitive to isolated data, yields more credible results.

Finally, the spatial distributions of the scale and shape parameters from RFA fits to rescaled maxima at the station sites are shown in Fig. 9a and b, respectively. As can be seen, the largest scale values appear in the eastern part of Sicily, indicating a relatively high frequency of intense precipitations. The largest values of shape – i.e. the heaviest tails of the GEV – are seen in the central and southern part of Sicily and the westernmost coast, meaning that the largest probabilities of the most intense events are assigned to these zones.

The average value of the shape parameter over all pooled records (0.16) is in very good agreement with the shape parameter provided by D'Asaro and Grillone (2008) (0.14), which estimated a unique set of parameters for the entire Sicily territory (the comparison is more difficult for the other parameters as they depend on the IF). Nevertheless, we give evidence of a clear spatial distribution of the GEV parameters over Sicily, with shape parameter values ranging from about 0.0 to about 0.3. This new result sets out the complex spatial pattern of the daily maximum precipitation distributional features over Sicily more clearly.

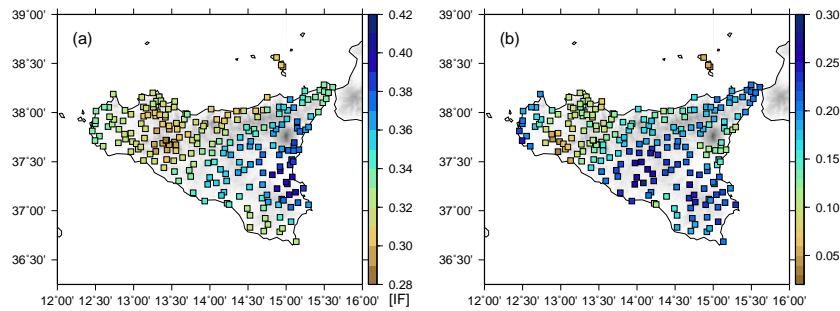


Figure 9. Spatial distribution of best guess values of (a) the scale (IF units) and (b) the shape parameters from RFA fits to rescaled annual maxima at the station points.

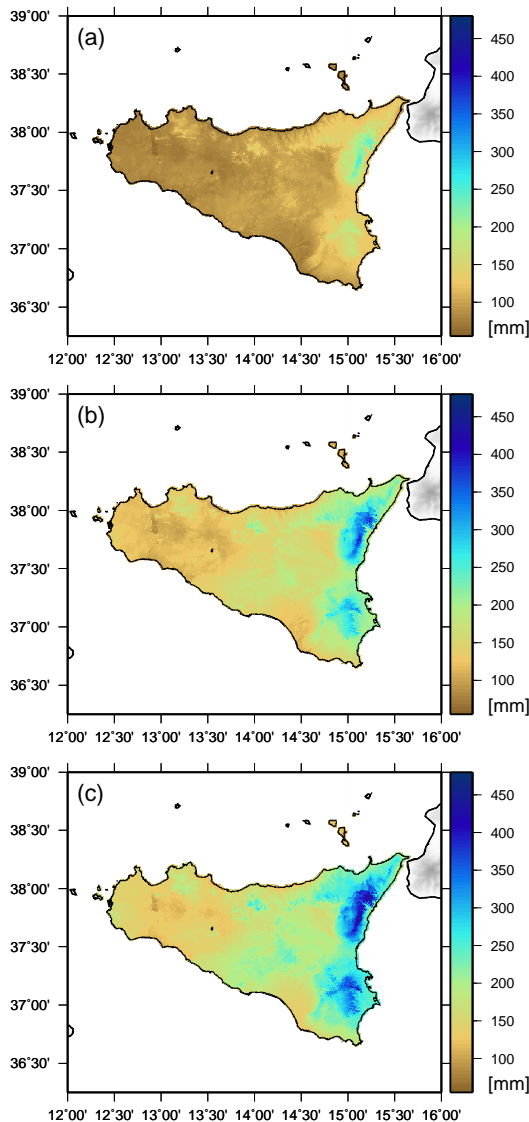


Figure 10. (a, b, c) Grid-point RL10, RL50 and RL100, respectively, expressed as absolute precipitation amounts (mm).

4.5 RFA estimates of RLs

RL10, RL50 and RL100 were first estimated from RFA fits to rescaled maxima at the station sites (not shown) together with respective uncertainties drawn from CI95s.

Except for the magnitudes of RL10, RL50 and RL100 – the lower the event probability, the larger the RL – their geographic pattern is almost the same in the three cases and resembles the spatial distribution of the shape parameter closely. Errors vary from site to site, stay quite moderate, on average, for RL10s (about 10 % of the IF), and steadily increase for RL50s and higher (reaching the magnitude of the IF). The largest errors are generally seen in some coastal sites and the north-eastern islets, where station density is poor and sample enlargement by pooling remains quite low (see Sect. 3). Other large errors can be identified in those areas, such as the Catania Plain, characterized by a high precipitation concentration index, i.e. low annual total precipitation concentrated in few heavy events.

Finally, dimensional grid-point estimates of RL10, RL50 and RL100 are shown in Fig. 10a, b and c, respectively, using an uniform scale of absolute precipitation amount (mm) for enabling relative comparison. As can be seen, the strongest events are expected on the eastern and north-eastern coastal areas, with increasing intensity as the return time of the events increases. These findings reflect the competing effects of a heavy tail of the GEV and of high values of mean annual precipitation totals in these areas. Indeed, as described in Sect. 3, annual normals play the role of regressors in the interpolation of IFs on the grid, ultimately used for recovering dimensional RLs. The north-eastern part of Sicily, in particular, emerges as the most affected by extreme precipitation events, with the highest RL100 (more than 450 mm), owing to the largest annual normals observed in the region, coupled with GEV parameters that are quite large, both in scale and in shape. High RLs are present also in the central and southern parts of eastern Sicily, where the annual normals are much lower than in north-eastern Sicily. The return levels we find for these areas are among the highest in the Mediterranean area (see e.g. Toreti et al., 2010).

5 Conclusions

A data set of 325 quality-checked Sicily daily precipitation records has been set up, collecting data from different sources for the 1921–2005 period. The records have then been homogenized, considering both precipitation amount and frequency and 231 of them have been used to set up records of yearly 1 day precipitation maxima to investigate extreme precipitation over Sicily by means of the GEV distribution.

The main results to be highlighted are as follows.

- i. A small, but significant number of outliers due to errors in the data were identified by the quality-check procedure; thereby this procedure turned out to be relevant, because GEV parameters are rather sensitive to outliers.
- ii. Homogenization turned out to have a beneficial effect on the results of extreme value analysis. This result is not trivial as homogeneity tests do not work on the extremes but on the bulk of precipitation data.
- iii. RFA turned out to be a very effective method to reduce the errors of GEV parameters and RLs. We used it in the station-year variant and adopted the median of the station yearly 1 day precipitation maxima as IF.
- iv. Both GEV parameters and corresponding RLs exhibit strong spatial gradients over Sicily, with increasing differences for higher RLs. These differences are due both to the scale and the shape parameters.
- v. We used a local linear regression of the station IF values versus the corresponding mean annual precipitation totals to estimate the IF on a high-resolution grid. This resulted in high-resolution estimates of the RLs, not only as far as the normalized data are concerned, but also for the precipitation absolute values. Therefore the methodology presented in the paper turns out to be an operational tool to assess precipitation risk.
- vi. The high-resolution maps of absolute values of RL10, RL50 and RL100 show that the north-eastern part of Sicily emerges as the most affected by extreme precipitations with the highest RL100 (more than 450 mm), owing to the largest annual precipitation totals observed in the region, coupled with GEV parameters that are quite large, both in scale and in shape. However, high RLs are also present in the central and southern parts of eastern Sicily, where the annual precipitation totals are much lower than in north-eastern Sicily.

In the future, we plan to consider precipitation in shorter time intervals than 1 day by using data from the Italian Hydrographical Service yearbooks and from the automatic stations in Sicily. Moreover, we plan to investigate the errors of the grid-point RLs in more detail. Specifically, we have to

understand better how much the confidence intervals we obtained are influenced by the correlation among the records, to investigate the uncertainty of grid-point annual precipitation totals more deeply and to combine the errors of the RLs from the pooling of the normalized data with those of the grid-point IF values.

Acknowledgements. We sincerely thank all the institutions that allowed access to the data for research purposes and contributed to set up the 1921–2005 Sicily daily precipitation database. They are (websites are provided for data access): Italian Air Force (“Servizio Meteorologico dell’Aeronautica Militare Italiana”, refer to <http://clima.meteoam.it/istruzioni.php> for data access. We received the data within an agreement between Italian Air Force and the Italian National Research Council), CRA-CMA (“Unità di ricerca per la climatologia e la meteorologia applicate all’agricoltura – Consiglio per la Ricerca e la Sperimentazione in Agricoltura”). The original data are available at <http://cma.entecra.it/homePage.htm> for 2004 onwards. The data of the previous years have to be requested at CRA-CMA), the Palermo Astronomical Observatory (the data are available from Giuseppina Micela, Loredana Granata, Vincenzo Iuliano, 2001: Due secoli di pioggia a Palermo. Università di Palermo – Osservatorio Astronomico di Palermo G. S. Vaiana, ISBN 99-87905-2-09) and the Osservatorio delle Acque (most of the data are available online from <http://www.osservatorioacque.it/>. Some other data have to be requested).

We also kindly acknowledge Ruth Löwenstein for her help in improving the language of the paper.

The analyses presented in the paper were performed using the R-package “extRemes” (Gilleland and Katz, 2011).

This study has been carried out in the framework of the EU FP7 project ECLISE (265240) Enabling CLimate Information Services for Europe, <http://www.eclise-project.eu/>.

Edited by: H. Saaroni

Reviewed by: two anonymous referees

References

- Alpert, P., Ben-Gai, T., Baharad, A., Benjamini, Y., Yekutieli, D., Colacino, M., Diodato, L., Ramis, C., Homar, V., Romero, R., Michaelides, S., and Manes, A.: The paradoxical increase of Mediterranean extreme daily rainfall in spite of decrease in total values. *Geophys. Res. Lett.*, 29, 31-1–31-4, doi:10.1029/2001GL013554, 2002.
- Brugnara, Y., Brunetti, M., Maugeri, M., Nanni T., and Simolo, C.: High-resolution analysis of daily precipitation trends in the central Alps over the last century. *Int. J. Climatol.*, 32, 1406–1422, doi:10.1002/joc.2363, 2012.
- Brunet, M., Sigró, J., Jones, P. D., Saladié, O., Aguilar, E., Moberg, A., Lister, D., and Walther, A.: Long-term changes in extreme temperatures and precipitation in Spain. *Contributions to Science*, 3, 331–342, doi:10.2436/20.7010.01.11, 2007.
- Brunetti, M., Caloiero, T., Coscarelli, R., Gullà, G., Nanni, T., and Simolo, C.: Precipitation variability and change in the Calabria region (Italy) from high resolution daily dataset. *Int. J. Climatol.*, 32, 57–73, doi:10.1002/joc.2233, 2012.

- Brunetti, M., Colacino, M., Maugeri, M., and Nanni, T.: Trends in the daily intensity of precipitation in Italy from 1951 to 1996, *Int. J. Climatol.*, 21, 299–316, doi:10.1002/joc.613, 2001.
- Brunetti, M., Lentini, G., Maugeri, M., Nanni, T., Simolo, C., and Spinoni, J.: 1961–90 high-resolution Northern and Central Italy monthly precipitation climatologies, *Adv. Sci. Res.*, 3, 73–78, available at: www.adv-sci-res.net/3/73/2009/ (last access: 13 October 2015), 2009.
- Brunetti, M., Maugeri, M., Monti, F., and Nanni, T.: Changes in daily precipitation frequency and distribution in Italy over the last 120 years. *J. Geophys. Res.*, 109, D05102, doi:10.1029/2003JD004296, 2004.
- Brunetti, M., Simolo, C., Maugeri, M., and Pasotti, L.: Report on the impact of climatic change on water availability over the Sicily Region. Project ECLISE (Enabling Climate Information Services for Europe) – RESEARCH REPORT – Work package: WP5 – WATER – TASK 5.3: Water availability in Sicily, Italy, CNR ISAC, available at: http://www.eclise-project.eu/content/mm_files/do_858/D5.3_ECLISE_WP5_Task_5_3_Deliv_5_3_ISAC_CNR.pdf (last access: 13 October 2015), 2014.
- Buishand, T. A.: Extreme rainfall estimation by combining data from several sites, *Hydrolog. Sci. J.*, 36, 345–365, doi:10.1080/02626669109492519, 1991.
- Cannarozzo, M., D’Asaro, F., and Ferro, V.: Regional rainfall and flood frequency analysis for Sicily using the two component extreme value distribution, *Hydrolog. Sci. J.*, 40, 19–41, 1995.
- Carreau, J., Neppel, L., Arnaud, P., and Cantet, P.: Extreme Rainfall Analysis at Ungauged Sites in the South of France : Comparison of Three Approaches, *Journal de la Société Française de Statistique*, 154, 119–138, 2013.
- Coles, S.: An introduction to statistical modeling of extreme values, Springer, London, UK, 208 pp., 2001.
- Craddock, J. M.: Methods of comparing annual rainfall records for climatic purposes, *Weather*, 34, 332–346, 1979.
- Daly, C., Neilson, R. P., and Phillips, D. L.: A statistical-topographic model for mapping climatological precipitation over Mountainous Terrain, *J. Appl. Meteorol.*, 33, 140–158, 1994.
- Daly, C., Halbleib, M., Smith, J. I., Gibson, W. P., Doggett, M. K., Taylor, G. H., Curtis, J., and Pasteris, P. A.: Physiographically-sensitive mapping of temperature and precipitation across the conterminous United States, *Int. J. Climatol.*, 28, 2031–2064, doi:10.1002/joc.1688, 2008.
- D’Asaro, F., Grillone, G.: Regional frequency analysis for extreme rainfall in Sicily, *J. Agr. Eng. Res.*, 1, 11–23, 2008.
- Davison, A. C. and Smith, R. L.: Models for exceedances over high thresholds, *J. Roy. Stat. Soc. B*, 52, 393–442, 1990.
- Fowler, H. J. and Kilsby, C. G.: A regional frequency analysis of United Kingdom extreme rainfall from 1961 to 2000, *Int. J. Climatol.*, 23, 1313–1334, 2003.
- Gilleland, E. and Katz, R. W.: New software to analyze how extremes change over time, *Eos*, 92, 13–14, 2011.
- Giorgi, F. and Lionello, P.: Climate Change Projections for the Mediterranean Region, *Global and Planetary Change*, 63, 90–104, doi:10.1016/j.gloplacha.2007.09.005, 2008.
- Gringorten, I. I.: A plotting rule for extreme probability paper, *J. Geophys. Res.*, 68, 813–814, 1963.
- Hanel, M., Buishand, T. A., and Ferro, C. A. T.: A nonstationary index flood model for precipitation extremes in transient regional climate model simulations, *J. Geophys. Res.*, 114, D15107, doi:10.1029/2009JD011712, 2009.
- Hosking, J. R. M.: L-moments: Analysis and estimation of distributions using linear combinations of order statistics, *J. Roy. Stat. Soc. B*, 52, 105–124, 1990.
- Hosking, J. R. M. and Wallis, J. R.: The effect of intersite dependence on regional flood frequency analysis, *Water Resour. Res.*, 24, 588–600, 1988.
- Hosking, J. R. M. and Wallis, J. R.: *Regional Frequency Analysis: An Approach Based on L-Moments*, Cambridge Univ. Press, New York, USA, 1997.
- Jones, M. R., Fowler, H. J., Kilsby, C. G., and Blenkinsop, S.: An assessment of changes in seasonal and annual extreme rainfall in the UK between 1961 and 2009, *Int. J. Climatol.*, 33, 1178–1194, doi:10.1002/joc.3503, 2013.
- Katz, R. W., Parlange, M. B., and Naveau, P.: Statistics of extremes in hydrology, *Adv. Water Resour.*, 25, 1287–1304, 2002.
- Kioutsoukakis, I., Melas, D., and Zerefos, C.: Statistical assessment of changes in climate extremes over Greece. *Int. J. Climatol.*, 30, 1723–1737, doi:10.1002/joc.2030, 2010.
- Kjeldsen, T. R. and Jones, D. A.: A formal statistical model for pooled analysis of extreme floods, *Hydrol. Res.*, 40, 465–480, 2009.
- Kostopoulou, E. and Jones, P. D.: Assessment of climate extremes in the eastern Mediterranean, *Meteorol. Atmos. Phys.*, 89, 69–85, doi:10.1007/s00703-005-0122-2, 2005.
- Landwehr, J. M., Matalas, N. C., and Wallis, J. R.: Probability weighted moments compared with some traditional techniques in estimating Gumbel parameters and quantiles, *Water Resour. Res.*, 15, 1055–1064, 1979.
- Leadbetter, M. R., Lindren, G., and Rootzen, H.: *Extreme and related properties of random sequences and processes*, Springer, New York, USA, 1983.
- Luterbacher, J., Xoplaki, E., Casty, C. et al.: Chapter 1. Mediterranean climate variability over the last centuries: A review. In: Lionello, P., Melanotte Rizzoli, P., Boscolo, R. (eds), *Mediterranean Climate Variability*, Elsevier, Amsterdam, the Netherlands, ISBN-10: 0-444-52170-4 (Series: *Developments in Earth and Environmental Sciences*, Vol. 4), 2006.
- Madsen, H. and Rosbjerg, D.: The partial duration series method in regional index-flood modeling, *Water Resour. Res.*, 33, 737–746, doi:10.1029/96WR03847, 1997.
- Madsen, H., Rasmussen, P. F., and Rosbjerg, D.: Comparison of annual maximum series and partial duration series methods for modeling extreme hydrologic events: 1. At-site modeling, *Water Resour. Res.*, 33, 747–757, doi:10.1029/96WR03848, 1997a.
- Madsen, H., Pearson, C. P., and Rosbjerg, D.: Comparison of annual maximum series and partial duration series methods for modeling extreme hydrologic events: 2. Regional modeling, *Water Resour. Res.*, 33, 759–769, doi:10.1029/96WR03849, 1997b.
- Martins, E. S. and Stedinger, J. R.: Generalized maximum-likelihood generalized extreme-value quantile estimators for hydrologic data, *Water Resour. Res.*, 36, 737–744, doi:10.1029/1999WR900330, 2000.
- Moberg, A., Jones, P. D., Lister, D. et al.: Indices for daily temperature and precipitation extremes in Europe analysed for the period 1901–2000, *J. Geophys. Res.*, 111, D22106, doi:10.1029/2006JD007103, 2006.

- Rodrigo, F. S.: Changes in the probability of extreme daily precipitation observed from 1951 to 2002 in the Iberian Peninsula, *Int. J. Climatol.*, 30, 1512–1525, doi:10.1002/joc.1987, 2010.
- Rossi, F., Fiorentino, M., and Versace P.: Two Component Extreme Value Distribution for flood frequency analysis, *Water Resour. Res.*, 20, 847–856, doi:10.1029/WR020i007p00847, 1984.
- Roth, M., Buishand, T. A., Jongbloed, G., Klein Tank, A. M. G., and vanZanten, J. H.: A regional peaks-over-threshold model in a nonstationary climate, *Water Resour. Res.*, 48, W11533, doi:10.1029/2012WR012214, 2012.
- Roth, M., Buishand, T. A., Jongbloed, G., Klein Tank, A. M. G., and van Zanten, J. H.: Projections of precipitation extremes based on a regional, non-stationary peaks-over-threshold approach: A case study for the Netherlands and north-western Germany, *Weather and Climate Extremes*, 4, 1–10, 2014.
- Svensson, C. and Jones, D. A.: Review of rainfall frequency estimation methods, *Journal of Flood Risk Management*, 3, 296–313, doi:10.1111/j.1753-318X.2010.01079.x, 2010.
- Toreti, A., Xoplaki, E., Maraun, D., Kuglitsch, F. G., Wanner, H., and Luterbacher, J.: Characterisation of extreme winter precipitation in Mediterranean coastal sites and associated anomalous atmospheric circulation patterns, *Nat. Hazards Earth Syst. Sci.*, 10, 1037–1050, doi:10.5194/nhess-10-1037-2010, 2010.
- Toreti, A., Naveau, P., Zampieri, M., Schindler, A., Scoccimarro, E., Xoplaki, E., Dijkstra, H. A., Gualdi, S., and Luterbacher, J.: Projections of global changes in precipitation extremes from Coupled Model Intercomparison Project Phase 5 models, *Geophys. Res. Lett.*, 40, 4887–4892, doi:10.1002/grl.50940, 2013.
- Ulbrich, U., Lionello, P., Belušić, D., Jacobeit, J., Knippertz, P., Kuglitsch, F. G., Leckebusch, G. C., Luterbacher, J., Maugeri, M., Maheras, P., Nissen, K. M., Pavan, V., Pinto, J. G., Saaroni, H., Seubert, S., Toreti, A., Xoplaki, E., and Ziv, B.: Chapter 5. Climate of the Mediterranean: Synoptic Patterns, Temperature, Precipitation, Winds, and Their Extremes, in: *The Climate of the Mediterranean Region From the Past to the Future*, edited by: Lionello, P., Elsevier, London, UK, ISBN: 978-0-12-416042-2, 301-346, 2012.
- USGS (United States Geological Survey), GTOPO30 Documentation, available at: http://eros.usgs.gov/#/Find_Data/Products_and_Data_Available/gtopo30_info (last access: 13 October 2015), 1996.
- van den Besselaar, E. J. M., Klein Tank, A. M. G., and Buishand, T. A.: Trends in European precipitation extremes over 1951–2010, *Int. J. Climatol.* 33, 2682–2689, doi:10.1002/joc.3619, 2013.