Nat. Hazards Earth Syst. Sci. Discuss., 3, 2247–2281, 2015
www.nat-hazards-earth-syst-sci-discuss.net/3/2247/2015/
doi:10.5194/nhessd-3-2247-2015

Natural Hazards and Earth System Sciences
Open Access
Discussions

High-resolution analysis of 1 day extreme precipitation in Sicily

M. Maugeri[1,2], M. Brunetti[2], M. Garzoglio[1], and C. Simolo[2]

[1]Department of Physics, Università degli Studi di Milano, Milan, Italy
[2]Institute of Atmospheric Sciences and Climate, CNR, Bologna, Italy

Received: 27 February 2015 – Accepted: 8 March 2015 – Published: 2 April 2015

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

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

Sicily, the major Mediterranean island, experienced several exceptional precipitation episodes and floods during the last century, with dramatic consequences on human life and environment. A long term, rational planning of urban development is mandatory for protecting population and avoiding huge economic losses in the future. This requires a deep knowledge of the distributional features of extreme precipitation over the complex territory of Sicily.

In the present study, we address this issue, and attempt 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 extrapolate 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 make evident the high vulnerability of its eastern and northeastern parts as those prone to the most intense and potentially damaging events. This analysis thus provides an operational tool for extreme precipitation risk assessment and, at the same time, is an useful basis for validation and downscaling of regional climate models.

1 Introduction

Sensitivity of the Mediterranean land areas to heavy precipitation and floods, and their responsiveness to climate change is documented by a vast literature (e.g., Ulbrich et al., 2012; Luterbacher et al., 2006). Partly due to the excessive exploitation of natural resources and soil, the exposure of the Mediterranean area to the hydro-meteorological hazard is supposed to get even higher in the future scenarios (Giorgi and Lionello,

NHESSD
3, 2247–2281, 2015

2008; Toreti et al., 2013), and imposes to strengthen the knowledge of present-day characteristics of precipitation extremes over this region.

The statistical analysis of hydrological extremes is traditionally rooted in the extreme value theory (EVT, see, e.g., Leadbetter at 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 of concern here 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 so rare – or still never recorded –, but yet potentially damaging.

In practice, one often deals with poor data samples, for long time series of, e.g., daily or sub-daily precipitation records are hardly available and even their quality might not be adequate to this scope. This seriously complicates 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 peaks-over-threshold (POT) approaches (Davison and Smith, 1990; Madsen et al., 1997a), have the advantage over the block data approach of including more events in the analysis, although at the cost of bringing some additional uncertainties in the fitting procedure. These are related, for instance, to the choice of a high threshold and its possible seasonality. Further, since partial-duration data are supposed to cluster in time, violation of data independence could be an issue for POT approaches, and some selection criteria for identifying distinct events are generally applied (Katz et al., 2002).

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 al., 1979; Hosking, 1990) – or for magnifying fitting samples themselves. This strategy is assumed in the regional frequency analysis (RFA), originally developed in the context of hydrology

NHESSD
3, 2247–2281, 2015

(see, e.g., Hosking and Wallis, 1997; Buishand, 1991; Madsen et al., 1997b), for reducing 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 possess 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).

In fact, 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) recommends to compute regional EVT parameters from weighted averages of L-moments across many similar sites, rather than to perform 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 were frequently used in regional studies (e.g., Fowler and Kilsby, 2003; Kjeildsen 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, in particular, 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. Even though results are not fully coherent throughout the region and rather conditioned by the specific site and period examined, the prevailing picture is of either decreasing or stationary precipitation extremes (e.g., Kostopoulou and Jones, 2005; D'Asaro and Grillone, 2008; Rodrigo, 2010; Toreti et al., 2010).

This work focuses on Sicily, the major Mediterranean island, which represents an outstanding example of a complex terrain for orographic and land–sea contrasts, with highly localized heavy precipitations both spatially and temporally. We performed a detailed study of RLs in 1-day precipitation extremes on a high-resolution grid (30 arcsec), for identifying the parts of this region 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 re-scaling of annual maxima 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.

Observations and their processing is described in Sect. 2. Details on 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 coming 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 pre-

NHESSD
3, 2247–2281, 2015

cipitation amounts (outliers) and spurious long dry-spells, by exploiting local coherence of precipitation episodes.

Specifically, as with outliers, single daily events of any station records 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 neighboring 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; (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. 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 local newspapers), less than 10 % of the uncertain events remained unconfirmed and thus canceled. Yet, few of these values were identified as monthly amounts, thereby forcing cancellation of entire years (6 years in 3-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 canceled (59 months and 12 years in 20 station records).

Next, quality-controlled records with at least 24 years remaining underwent assessment of their homogeneity level, and when required the homogenization process. The lower bound on the series length was set to balance the need of 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 re-location and/or instruments 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., 2012). In practice, records in sub-groups 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 neighboring 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 1 or 2 re-adjusted periods. The spatial distribution of selected records is shown in Fig. 1a. Their data availability during 1921–2005 is shown in Fig. 1b in terms of the number of years comprising at least 90 % of valid data for any station record, since realistic annual maxima can be extracted from sufficiently complete years only. 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 RLs at a given station estimated from a pooled sample of records from the nearest sites – excluding that station – with those estimated from both the homogenized station record alone and its original counterpart. As discussed in detail in Sect. 4.2, results from homogenized data are superior to those from original data, thereby proving the beneficial effect of the homogenization practice in the context of extremal analysis.

NHESSD
3, 2247–2281, 2015

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 to the frame from 12.4 to 15.7° E and from 36.5 to 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 vs. elevation, as described in detail in Brunetti et al. (2014), and are shown in Fig. 2.

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. divide annual maxima from single station records by an appropriate, site-specific IF (*rescaling*);

ii. assign to any grid-point an enhanced sample of rescaled maxima drawn from the station records falling within a definite grid-point neighborhood, so-called region of influence (*pooling*);

iii. fit an unique GEV distribution to any grid-point sample and extrapolate non-dimensional RLs for the given return periods (*fit*);

iv. interpolate site-specific IFs on the high-resolution grid by regression with annual mean precipitation totals and retrieve dimensional RLs (*spatialization*).

i. As already noted, 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. Dividing the above sequences by their own median – taken as the IF for is a robust estimator of central tendency in extreme data samples – facilitates comparison between frequency distributions of precipitation maxima at different sites. Indeed, rescaling may unmask similarities in subset of data that could be exploited for improving extremal analysis. In this view, 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. Any of the nearly 38 000 Sicily grid-points was set as the center of a search area with a 25 km radius. Rescaled maxima from the station sites falling within that grid-point neighborhood were merged into an 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 about any 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 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 were computed in the leave-one-out approach to exclude the influence of the given station itself. Performances at any step were evaluated by the Pooled Uncertainty Measure (PUM) (Kjeildsen and Jones, 2009) that quantifies the average deviation of the pooled estimate from the respective single-station value over all stations, i.e.,

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

Here, z_i and $\hat{z}_i$ are respectively the single-station and the pooled-station quantile under study, 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

NHESSD
3, 2247–2281, 2015

optimal upper bound on the pooling distance. This was used in the following to define the region of influence about any 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 should also be kept to ensure correlation-free fits. In a way similar to the above, the appropriate lower bound on the pooling distance was defined once for all, 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 bound at 5 km appeared a reasonable choice for both the PUM and the mean error band (amplitude of the confidence interval at the 95 % level, CI95) stay quite low while the mean sample size is still high (Fig. 3b and d). More restrictive choices would entail sensible loss of data and, on the contrary, relaxing threshold below 5 km would not pull-down 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 yet above 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 as a total per region, very few exceptions concerning quite isolated areas – e.g., some coastal sites with less than 300 data and the north-eastern islets with down to about 70 data.

Prior to the fit, pooled samples were further investigated at the station level to ascertain 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.

NHESSD
3, 2247–2281, 2015

iii. Extrapolation of very high quantiles from the observed frequency of precipitation maxima goes through the definition of a GEV model that well fits the data, enabling predictions about the rarest – and even not yet observed – events. The GEV distribution function of block maxima has the well-known form

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

for all z such that $1+\xi(z-\mu)/\sigma > 0$. The time-independent (see Sect. 4.3) shape ξ, scale 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 embedded in the suite of functions of the R-package "extRemes" (Gilleland and Katz, 2011). The error band for any 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), i.e., for any $i = 1{,}2{,}3$

$$CI95_i = \{\theta_i : D(\theta_i) < c_{0.05}\}, \quad D(\theta_i) = 2\{l(\theta_0) - l_p(\theta_i)\} \tag{3}$$

Here, θ_i is a parameter, $D(\theta_i)$ and $l_p(\theta_i)$ are respectively the deviance function and the profile log-likelihood for θ_i, and $l(\theta_0)$ the maximized log-likelihood over all parameters. Finally $c_{0.05}$ is the 0.95 quantile of a χ_1^2, which asymptotically approximates the sampling distribution of the deviance function. This procedure generally yields more accurate confidence intervals than those computed from the usual standard errors based on the approximate normality of the likelihood estimators.

Then, the RL z_p associated with a small probability p or, equivalently, with a return period $T = 1/p$ is obtained by inverting the GEV in Eq. (2), i.e., $z_p = G^{-1}$, where $G = 1 - p$. Confidence intervals are obtained from the profile log-likelihood of z_p and Eq. (3) as above, after re-expressing one of the model parameters in terms of z_p (Coles, 2001).

NHESSD
3, 2247–2281, 2015

The GEV is a well-established approach to extremal analysis for its 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 yet on single-station records. Some detailed results about the fit performance of the GEV model and the behavior of parameters from both single-station and pooled records are further discussed 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 an unique GEV to the pooled sample. In that variant of RFA, regional GEV parameters are obtained by the method of moments. The first three L-moments used in the fit are obtained as weighted means of respective moments estimated from single-station data in any region, with weights given by the length of station records. As seen in Fig. 4a–c, this method yields nearly equal results. Indeed, the GEV parameters obtained as above 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.

Nevertheless, this approach slightly complicates the estimation of confidence intervals, since the maximum likelihood principle with Eq. (3) cannot be applied in this context, and some bootstrap errors should be used instead. Since no real improvement of the fit comes from this latter variant of RFA, the above discussed station-year approach is assumed in what follows.

iv. Restoring physical dimensions of grid-point RLs requires re-multiplication by the respective median of precipitation maxima – the IF – evaluated at the grid-point. Hence, at-sites IFs were interpolated on the regular grid by a local, linear regression with mean annual precipitation totals (normals, see Fig. 2), that well correlate spatially with the former quantities. Specifically, for any grid-point, a fit was performed between station normals and medians using only the bulk of data from the nearest sites, within a search area of varying radius from 10 to 25 km, until a minimum number of 7 stations is reached. 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. 5. Here, the pairs of estimated and observed values are seen to line up on the bisector, except for 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 where the number of high-elevation stations is critically low and the fit underestimates precipitation maxima. The total uncertainty on the grid-point IFs can be measured by the root mean square error (RMSE) over all sites, that amounts to about 8 mm, i.e., the 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

Before going into details on 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 about any grid-point have been defined, it is desirable to verify 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 their temporal stationarity have been explored fully (Sects. 4.2 and 4.3), along with the behavior of GEV parameters from both the single-station and the pooled-station fits (Sect. 4.4). Finally, high-resolution maps of RL10, RL50, RL100 are discussed in detail (Sect. 4.5).

NHESSD
3, 2247–2281, 2015

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.

First, following Hosking and Wallis (1997), a discordancy test was recursively applied at any 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 neighboring 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 fulfill 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 verify 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 neighboring 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 taken with a grain of salt. 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 – only weakly satisfied. Rather, the above critical thresholds have been here conceived as a benchmark for further inspecting threshold-exceeding regions in conjunction with more physical arguments. Indeed, poor H test results mainly gather in 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 val-

NHESSD
3, 2247–2281, 2015

ues 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.

4.2 Data homogenization

As already mentioned in Sect. 2, benefits and/or drawbacks of data homogenization were thoroughly assessed from its impact on the GEV RL50s, taken as a benchmark for they are intermediate quantiles in the investigated range. 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 latter case, RL50s were obtained using again both original and homogenized data, and the RFA about any station sites as described in Sect. 3. Since results from pooled homogenized data (HP) are largely consistent with those from pooled original data (OP), only the latters were used for direct comparison with RL50s from both original (OS) and homogenized single-station records (HS). As seen in Fig. 6 HS RL50s generally compare better than OSs (when they differ) with OP results, indicating that homogenization does remove gross errors and restores the true expected behavior 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.

4.3 Stationarity

The prerequisite for a stationary GEV 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 abruptly drops during the decade 1942–1950. As shown in Fig. 7a, less than 15 % of the station series of rescaled maxima exhibits a significant trend of decreasing extreme precipitation, from the starting to the ending year. This tendency disappears when the analysis is carried out from 1952 to the end (Fig. 7b), with only 6 % of the series showing trends, either positive or negative evenly.

To some degree, the temporal discontinuity in data availability reflects on the overall quality level of records that, regardless of their detailed pre-processing, remains slightly lower in the early decades. Indeed, both the quality controls and homogenization point to remove 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 quite large areas. 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.

4.4 GEV parameters from single-station and RFA fits

Here, we discuss the behavior of relevant GEV parameters, i.e., the scale and shape, when switching from a single-station analysis to pooled-stations in the perspective of RFA. Indeed, as already noted, the aim of RFA is to constrain parameters' values by reducing their errors considerably. This technique fixes one basic drawback of the EVT approaches – that, by nature, rely on poor sample sizes with ensuing instability of distributional parameters –, thereby making the theory more informative and useful for practical applications.

A GEV distribution is fitted to rescaled annual maxima from any single-station (SS) record 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. 8a and c. Notice again that the alternative discussed in Sect. 3 to the straightforward parameters' fit to pooled-station samples is

NHESSD
3, 2247–2281, 2015

totally equivalent. As seen in Fig. 8b and d, RFA provides a net reduction of confidence intervals, up to 80 % in many cases. Likewise, central (best guess) values of RFA parameters – both scale and shape – 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., a Frechèt-type, with positive shape in the notation of Eq. (2). Conversely, RFA estimates stay always positive within the error bars.

By way of illustration, Fig. 9a and b shows a couple of cases where SS fits of the GEV 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, Kjeildsen and Jones, 2009) obtained from both SS and RFA fits 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 distribution of the scale and shape parameters from RFA fits to rescaled maxima at the station sites are shown on a map in Fig. 10a 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 to the most intense events are assigned to these zones.

4.5 RL estimates

RL10, RL50 and RL100 from RFA fits to rescaled maxima at the station sites are shown on a map in Fig. 11a–f together with respective uncertainties drawn from CI95s. Indeed, given the asymmetry of the likelihood function (see Sect. 3), the right-hand amplitude

– the highest value of the CI95 minus the best guess – is the largest, and thus safely taken as a measure of uncertainty.

As explained in Sect. 3, the above RLs express very high GEV quantiles in IF units, and are here displayed with a IF scale adapted to the respective range of values for resolving pattern details. Except for the magnitudes – the lowest the event probability the largest the RL –, the geographic pattern of RL10, RL50 and RL100 is quite the same in the three cases and resembles closely the spatial distribution of the shape parameter. 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, RL100 are shown in Fig. 12a–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 northeastern 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 precipitations with the highest RL100 (more than 450 mm), owing to the largest annual normals observed in the region, coupled with quite large GEV parameters, both scale and 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.

NHESSD
3, 2247–2281, 2015

5 Conclusions

A dataset 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 that have been used to investigate, by means of the GEV distribution, extreme precipitation over Sicily.

The main results highlighted by the analyses are:

i. A small, but significant number of outliers due to errors in the data were identified by the quality check procedure. This procedure turned therefore out to be relevant, as GEV parameters are rather sensitive to outliers.

ii. Homogenisation 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 – here used in the station-year variant and adopting the median of the station yearly 1-day precipitation maxima as IF – turned out to be a very effective method to reduce the errors of GEV parameters and RLs.

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. A local linear regression of the station IF values vs. the corresponding mean annual precipitation totals allowed estimating the IF onto a high-resolution grid. This allowed producing high-resolution estimates of the RLs, not only as far as the normalised data are concerned, but also for the precipitation absolute values. The methodology presented in the paper turns therefore out to be an operational tool for precipitation risk assessment.

NHESSD
3, 2247–2281, 2015

vi. The high-resolution maps of absolute values of RL10, RL50, RL100 give evidence 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 quite large GEV parameters, both scale and shape. High RLs are however present also 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 also precipitation in shorter time intervals than 1 day by exploiting both the data from the yearbooks of the Italian Hydrographical Service and the data from the automatic stations of Sicily. Moreover, we plan to investigate more in detail the errors of the grid-point RLs. Specifically, we have to better understand how much the confidence intervals we got are influenced by the correlation among the records, to investigate the uncertainty of grid-point annual precipitation totals and to combine the errors of the RLs from the pooling of the normalised 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 (web sites are provided for data access): Italian Air Force ("Servizio Meteorologico dell'Areonautica Militare Italiana", refer to http://clima.meteoam.it/istruzioni.php for data access. We received the data in the frame of 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 since 2003. The data of the previous years have to be request at CRA-CMA), the Palermo Astronomical Observatory (the data are available from Micela G., Granata L., Iuliano V., 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 on-line from http://www.osservatorioacque.it/. Some other data have to be requested).

This study has been carried out in the framework of the EU project ECLISE (265240).

References

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., 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, Italy, 8 pp., available online at: http://www.eclise-project.eu, 2014.

Buishand, T. A.: Extreme rainfall estimation by combining data from several sites, Hydrolog. Sci. J., 36, 345–365, doi:10.1080/02626669109492519, 1991.

Coles, S.: An Introduction to Statistical Modeling of Extreme Values, Springer, London, 2001.

Craddock, J. M.: Methods of comparing annual rainfall records for climatic purposes, Weather, 34, 332–346, 1979.

D'Asaro, F. and Grillone, G.: Regional frequency analysis for extreme rainfall in Sicily, J. Agr. Eng. Res., 39, 11–23, 2008.

Davison, A. C. and Smith, R. L.: Models for exceedances over high thresholds, J. R. Stat. Soc., 52, 393–442, 1990.

Fowler, H. J. and Kilsby, C. G.: A regional frequency analysis of UK 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 T. Am. Geophys. Un., 92, 13–14, 2011.

Giorgi, F. and Lionello, P.: Climate change projections for the Mediterranean Region, Global Planet. 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. R. Stat. Soc., 52, 105–124, 1990.

NHESSD

3, 2247–2281, 2015

High-resolution analysis of Sicilly extreme precipitation

M. Maugeri et al.

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

◀| |▶

◀ ▶

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Hosking, J. R. M. and Wallis, J. R.: Regional Frequency Analysis: an Approach Based on L-Moments, Cambridge Univ. Press, New York, 1997.
Hosking, J. R. M. and Wallis, J. R.: The effect of intersite dependence on regional flood frequency analysis, Water Resour. Res., 244, 588–600, 1988.
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.
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., 155, 1055–1064, 1979.
Leadbetter, M. R., Lindren, G., and Rootzen, H.: Extreme and Related Properties of Random Sequences and Processes, Springer, New York, 1983.
Luterbacher, J., Xoplaki, E., Casty, C., Wanner, H., Pauling, A., Brönnimann, S., Fischer, E., González-Rouco, F. J., Garcìa-Herrera, R., Barriendos, M., Rodrigo, F., Gonzalez-Hidalgo, J. C., Saz, M. A., Gimeno, L., Ribera, P., Brunet, M., Paeth, H., Rimbu, N., Felis, T., Jacobeit, J., Dünkeloh, A., Zorita, E., Guiot, J., Türke, M., Alcoforado, M. J., Trigo, R., Wheeler, D., Tett, S., Mann, M. E., Touchan, R., Shindell, D. T., Silenzi, S., Montagna, P., Camuffo, D., Mariotti, A., Nanni, T., Brunetti, M., Maugeri, M., Zerefos, C., De Zolt, S., and Lionello, P.: Chapter 1. Mediterranean climate variability over the last centuries: a review, in: Mediterranean Climate Variability, edited by: Lionello, P., Melanotte Rizzoli, P., and Boscolo, R., Series: Developments in Earth and Environmental Sciences, Vol. 4., Elsevier, Amsterdam, 27–148, ISBN-10: 0-444-52170-4, 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.

NHESSD
3, 2247–2281, 2015
High-resolution analysis of Sicilly extreme precipitation
M. Maugeri et al.

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.

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.

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., Belušic, 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, 301–346, ISBN: 978-12-416042-2, 2012.

USGS (United States Geological Survey), GTOPO30 Documentation, available at: http://eros.usgs.gov/#/Find_Data/Products_and_Data_Available/gtopo30_info, 1996.

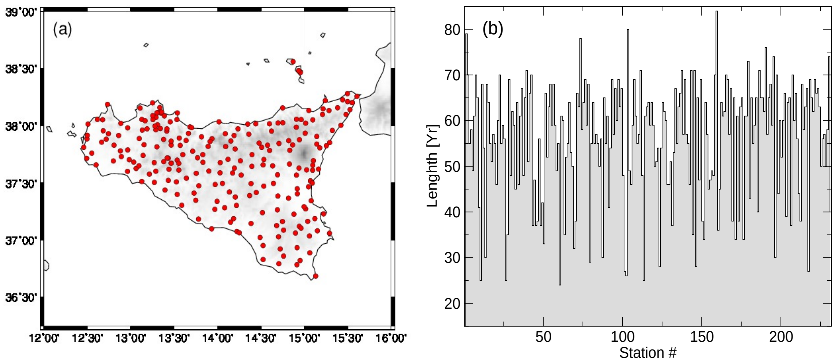

Figure 1. (a) Location of station sites. **(b)** Number of years with at least 90 % of valid data per station record.

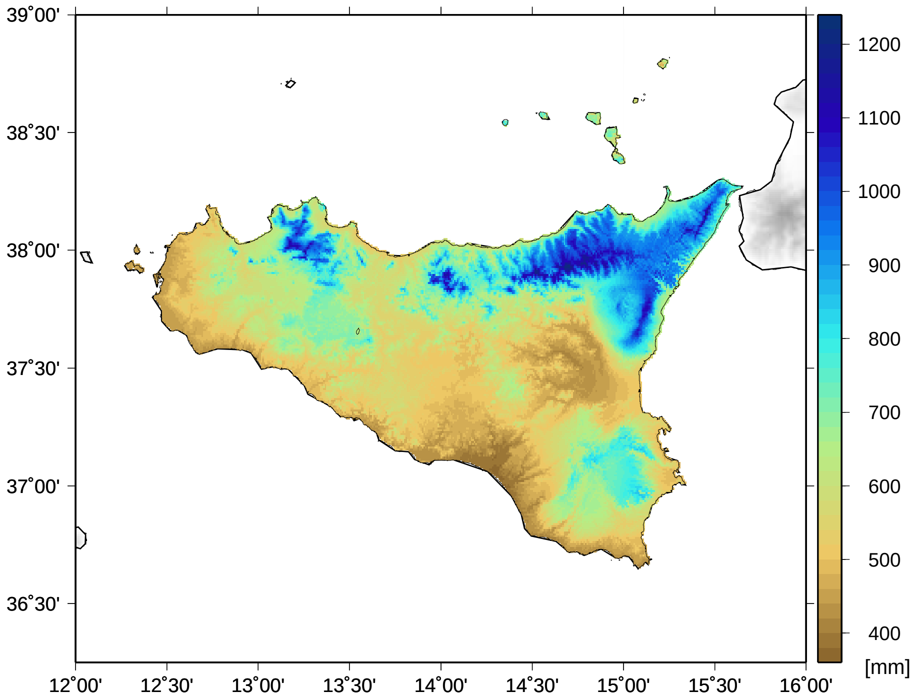

Figure 2. Climatic normals (1961–1990) of annual precipitation totals for Sicily at the 30 arcsec resolution.

NHESSD

3, 2247–2281, 2015

High-resolution analysis of Sicilly extreme precipitation

M. Maugeri et al.

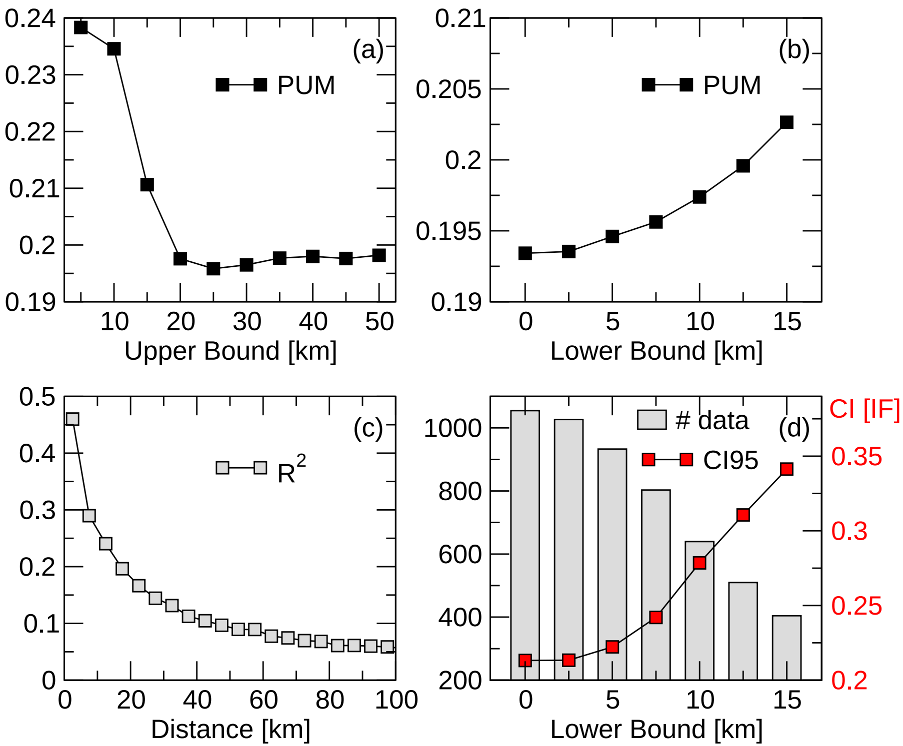

Figure 3. Variation of the RL50 PUM given by Eq. (1), as a function of **(a)** the upper and **(b)** the lower bound to the pooling radius. **(c)** Common variance between series of annual maxima vs. distance of sites, computed at a 5 km step, starting from a minimum distance of 2.5 km. **(d)** Variation 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.

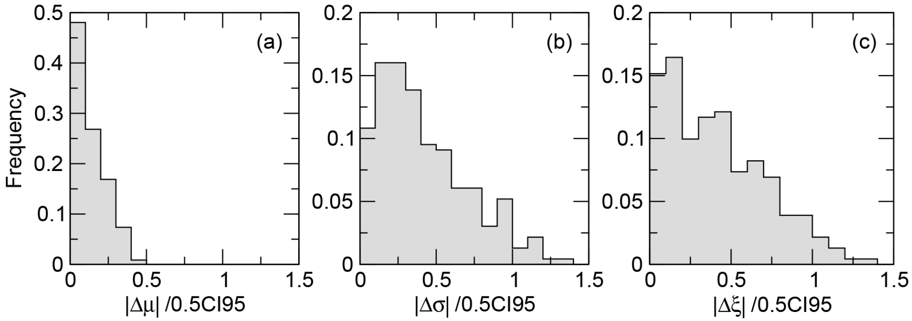

Figure 4. Frequency histograms of discrepancies between GEV parameters from the station-year method and the regional L-moment algorithm, respectively **(a)** location, **(b)** scale and **(c)** shape, expressed as absolute differences relative to one-half of the station-year CI95.

NHESSD
3, 2247–2281, 2015

High-resolution analysis of Sicilly extreme precipitation

M. Maugeri et al.

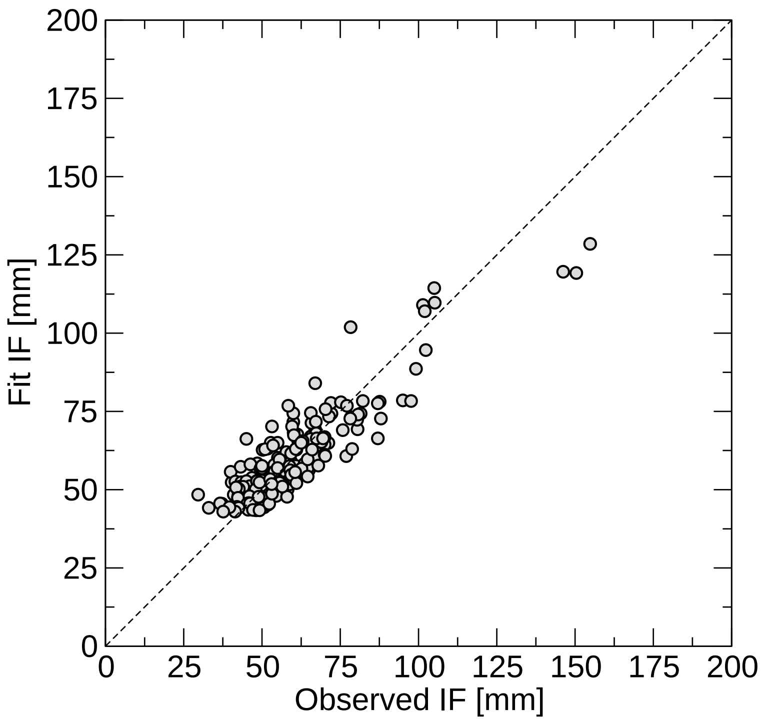

Figure 5. Estimated vs. observed IFs at the station sites.

NHESSD
3, 2247–2281, 2015

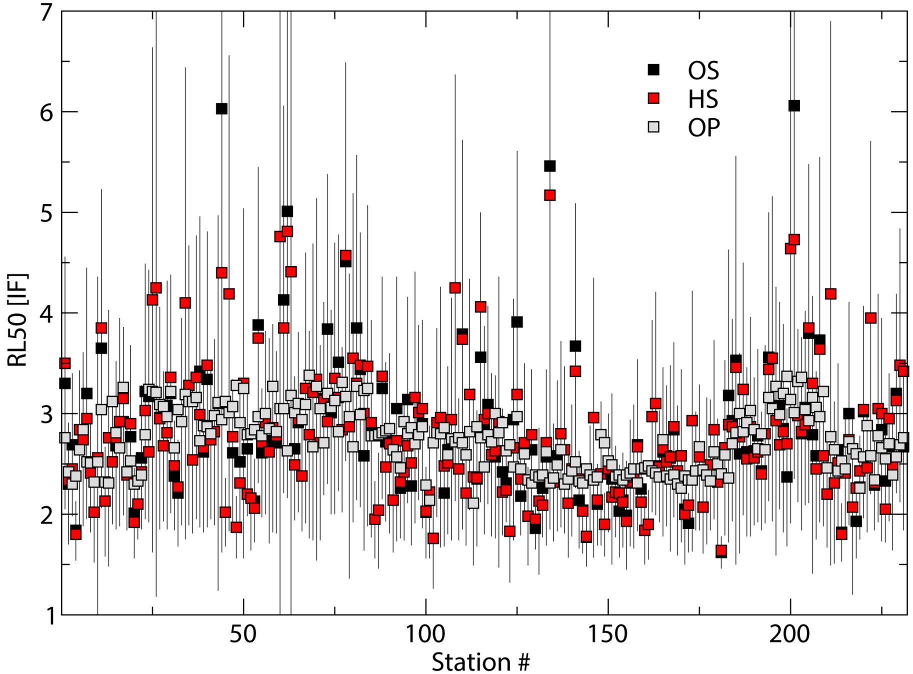

Figure 6. Rescaled RL50s (IF units) for any station site, obtained from both original (OS) and homogenized records (HS) at the single-station level, and compared with expected RL50 from the pooled-station samples (OP). Only error bars of OS RL50 are shown, as they largely overlap with HS errors and completely mask those of the OPs.

NHESSD
3, 2247–2281, 2015

High-resolution analysis of Sicilly extreme precipitation

M. Maugeri et al.

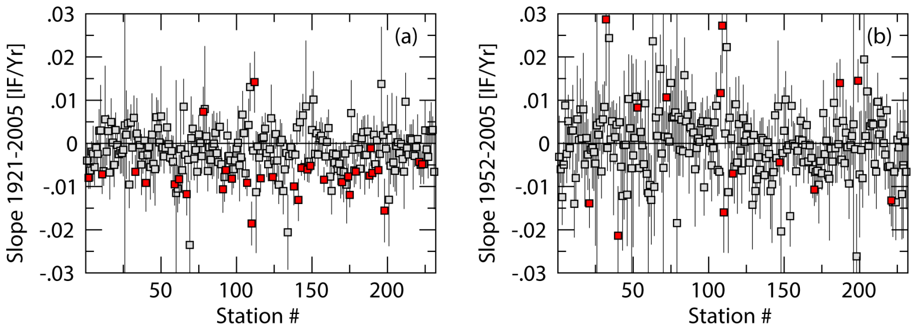

Figure 7. Trend slopes with 1-sigma error bars relative to **(a)** the full record period and **(b)** a relevant sub-period (1952–2005), for any station site in both cases. Red symbols denote significant trends (p value < 0.05) according to a Mann–Kendall test.

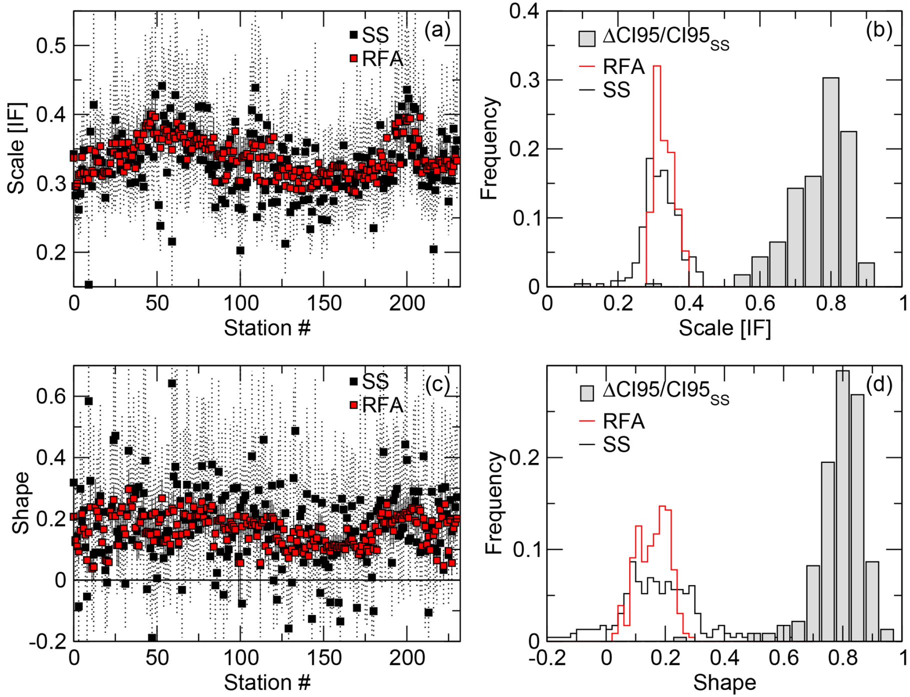

Figure 8. (a, c) Station-by-station GEV scale (IF units) and shape parameters respectively, with 95 % confidence intervals from fits to single-station (SS) and pooled-station (RFA) data. Dotted bars denote SS errors and solid bars RFA errors. **(b, d)** Frequency histograms of the SS and RFA estimates of the scale and shape parameters, respectively, across the region, and the fractional reductions of CI95 by RFA (gray-filled bars) for both parameters.

NHESSD
3, 2247–2281, 2015

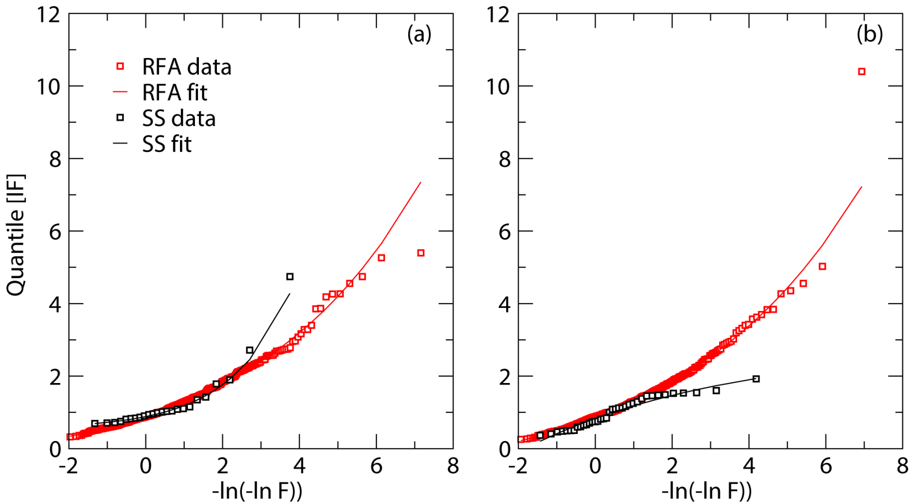

Figure 9. Growth curves (IF units) for two exemplifying sites, with SS best guess value of the shape **(a)** = 0.58 (Diga Ragoleto, CT) and **(b)** = −0.19 (Sigonella, CT). The relative frequency F on the abscissa is defined according to the Gringorten rule (Gringorten, 1963).

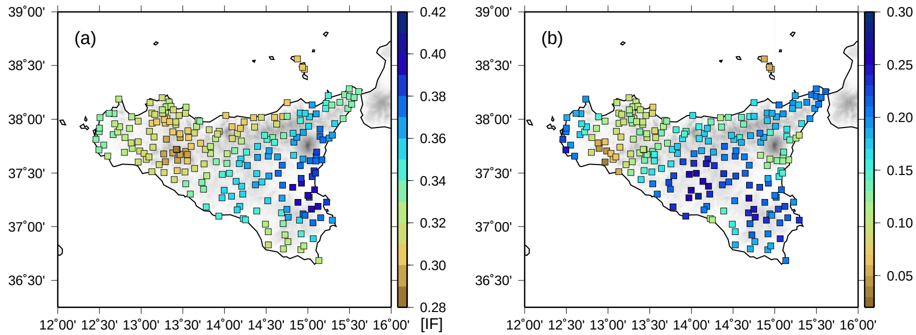

Figure 10. Spatial distribution of best guess values of **(a)** the scale (IF units) and **(b)** the shape parameters from RFA fits to rescaled annual maxima at the station points.

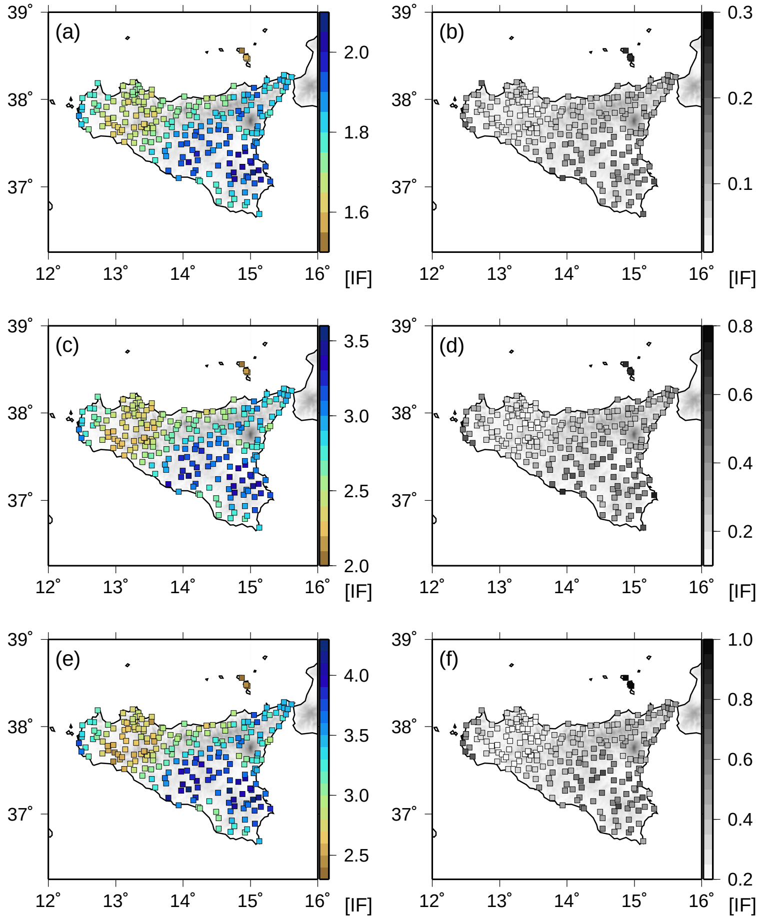

Figure 11. (a, c, e) RL10, RL50 and RL100, respectively, from RFA fits to rescaled maxima at the station sites (IF units). **(b, d, f)** Right-hand amplitude of CI95s for RL10, RL50 and RL100 respectively (IF units).

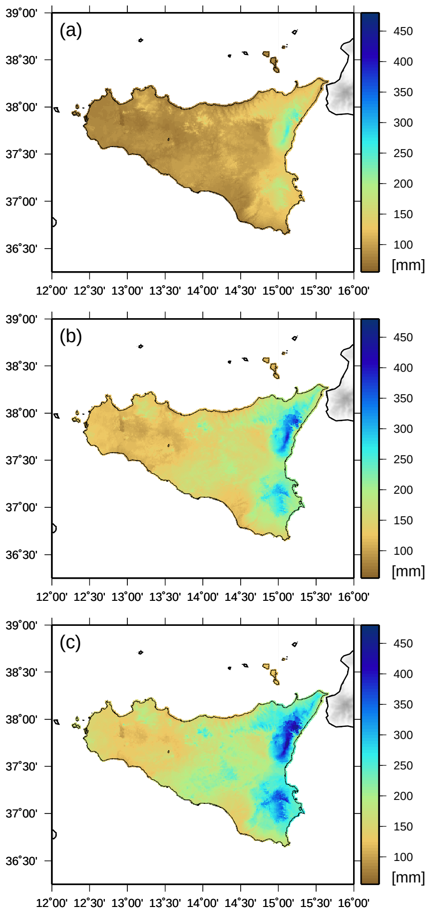

Figure 12. (a, b, c) Grid-point RL10, RL50 and RL100, respectively, expressed as absolute precipitation amounts (mm).

NHESSD

3, 2247–2281, 2015

High-resolution analysis of Sicilly extreme precipitation

M. Maugeri et al.