Observed variability and trends in extreme rainfall indices and Peaks-Over-Threshold series

H. Saidi¹, M. Ciampittiello¹, C. Dresti¹, and G. Ghiglieri²

¹The National Research Council – Institute of Ecosystem Study, Largo Tonolli 50, 28922 Verbania, Italy
²Department of Earth Science, Cagliari University, Via Trentino 51, 09126 Cagliari, Italy, Desertification Research Group (NRD), Sassari University, Italy

Received: 24 April 2013 – Accepted: 3 May 2013 – Published: 15 May 2013

Correspondence to: H. Saidi (h.saidi@ise.cnr.it)

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

Intensification of heavy precipitation as discussed in climate change studies has become a public concern, but it has not yet been examined well with observed data, particularly with data at short temporal scale like hourly and sub-hourly data.

In this research we digitalized sub-hourly precipitation recorded at the stations of Vercelli (since 1927), Bra (since 1933), Lombriasco (since 1939) and Pallanza (since 1950) in order to investigate historical change in extreme short precipitations. These stations are located in the northwest of Italy.

Besides seasonal and yearly maximum of precipitation we adopted two indices of extreme rainfall: the number of events above an extreme threshold (extreme frequency), and the average intensity of rainfall from extreme events (extreme intensity).

The results showed a statistically significant increase of the extreme frequency index and spring maximum precipitation for Bra and Lombriasco. The extreme intensity index presented by the means of events above 95th percentile is decreasing for Bra regarding hourly precipitation and increasing for Lombriasco regarding 20 min extreme events. In Pallanza, we noticed only a positive trend of the extreme frequency and extreme intensity indices of events with duration of 30 min.

For the analyses presented in this paper, a peak-over-threshold approach was chosen. Investigation presented showed that extreme events have risen in the last 20 yr only for short duration. Here it cannot be said that in our study area recent sub-hourly and hourly precipitation have become unprecedently strong or frequent for all the stations and for all the extreme events duration.

1 Introduction

Trend detection is an active area of interest for both hydrology and climatology in order to investigate climate changes scenario and improve climate impact research. The assumption of stationarity seems to be invalid as a result of anthropogenic influence and
the natural variability of the climate system (Karpouzos et al., 2010). Therefore, trend detection in precipitation time series is crucial for planning regional water resources management and civil defence.

Climate simulations indicate that a warmer climate could result in an increase in the proportion of precipitation occurring in extreme events (Karl et al., 1995). It seems to be generally accepted that the expected climatic changes are not necessarily associated with a higher intensity of extreme values, but rather with a higher frequency of the occurrence of extreme values. Recent studies (Easterling et al., 2000) analyzed the changes in observed heavy precipitation, based on daily precipitation, resulting in the detection of an increase in heavy precipitation at many parts of the world, as well as a decrease at some parts of the world. Due to the limitation of available digitalized records daily precipitation, as mentioned above, has been the major material for analysis so far.

Based on an overview made by Brunetti et al. (2001), we can mention that some studies relating to the variation of heavy and extreme events were performed for the USA (Karl et al., 1995; Karl and Knight, 1998; Trenberth, 1998; Kunkel et al., 1999), Japan (Iwashima and Yamamoto, 1993), eastern and northeastern Australia (Suppiah and Hennessy, 1998; Hennessy et al., 1999; Plummer et al., 1999), South Africa (Mason et al., 1999), the UK (Osborn et al., 2000) and Italy (Brunetti et al., 2004, 2006).

Karl et al. (1995) and Karl and Knight (1998) observed a significant positive trend in the frequency of extreme rainfalls (greater than 50 mm per day) over the last few decades in the USA. In Australia, Suppiah and Hennessy (1996) and Hennessy et al. (1999) showed a significant increase in the 90th and 95th percentiles, while Hennessy et al. (1999) and Plummer et al. (1999) showed increases in the 99th percentile. Iwashima and Yamamoto (1993) found that, in Japan, more stations recorded their highest precipitation events in recent decades. Brunetti et al. (2004, 2006) confirmed a strong decrease in precipitation trends over Italy, with a rainfall reduction of about 135 mm in the southern regions during the last 50 yr.
Groisman et al. (1999) performed a study on heavy precipitation over a wide area comprising Canada and Norway (for the period 1900–1995), the USA and Australia (spanning the period 1910–1999), the former Soviet Union (1936–1994), Mexico, China, Alaska and Poland (whose data are available for post-World War II). They found an increase both in summer rainy days and in heavy precipitation frequency over the past century for the USA, Norway and Australia, but they found no significant trend for any other country where the series are shorter and/or have many missing data. In most of the analysed areas, the positive trend observed in rain intensity is generally associated with an increase in total precipitation. Groisman et al. (1999) studied the relationship between the increase in total precipitation and the frequency of heavy rain events.

Typical temporal scales of precipitation phenomena may suggest the need to analyze precipitation records of shorter resolution than a day. Kanae et al. (2004) digitalized and investigated the hourly precipitation measured at the Tokyo observatory since 1890, and proved an upward trend in heavy precipitation over Japan. They report that “many hourly heavy precipitation events (above 20 mm h\(^{-1}\)) occurred in the 1990s compared with the 1970s and the 1980s”.

In the Alpine regions the evidence is growing stronger that climate warming is accompanied by an increase in frequency of intense precipitation events (Frei and Schär, 1998, 2000). Some previous studies (Saidi et al., 2012) took the advantage of utilizing sub-daily precipitation data for analysis of regional changes in precipitation in the Lake Maggiore watershed where the Alps cause heavy and extreme events. The time periods of the utilized data were generally limited to a few decades due to the availability of digital data. Recent progress on digitalization method of old climatic data, point out to us the importance of studying the changes in hydrological cycle with a longer record.

The purpose of this work is to investigate the variability of precipitation data collected in 4 different sites in the Piedmont region in Italy. The historical extreme rainfall series with high-resolution from 5 min to 30 min and above: 1, 2, 3, 6, and 12 h collected at different gauges have been computed to perform a statistical analysis to determine
whether the recent changes in frequency and magnitude of the rainfall extremes can be considered statistically significant. Trends are analysed both on an annual and a seasonal scale. The implications of changes on the seasonal scale are particularly significant for water resource management processes related to seasonal cycles.

2 Data

Understanding climate change demands attention towards changes in climate variability and extremes, but knowledge of the behaviour of these variables has been limited by the lack of long-term high-resolution data. The extending spread of new technological techniques at meteorological stations has made it possible to have high temporal resolution data (i.e. hourly and sub-hourly) and to collect them automatically.

Very few of such stations were operated before 1980. Hence the time period available with sub-daily rainfall totals is quite short for conducting analyses on long-term changes. We decided to transform the oldest data available, currently in paper format, to digital format. It was possible to have a more complete view of evolution and trend of extreme events.

Given that still now we have few stations with sufficient data record length to justify a trend detection study, so we decided to concentrate our research only on some of them situated in the Piedmont region (Italy): Vercelli, Bra, Lombrirasco and Pallanza to verify and calibrate better the model analysis of long time series of extreme events. All the station are situated in plain and good allocated in the flat total area. Details of the location and the period of observation are summarized in Table 1.

As a matter of fact, several types of software dedicated to data digitization are available, which allow the transcription of the paper-recorded data into text files, after the acquisition of the tracks by a scanner as an image file (Fig. 1).

We used two types of software, Plot2data (Leonardi et al., 2006) and GetData Graph Digitizer (2012, http://getdata-graph-digitizer.com), dedicated to the automatic reading
of scanned images of records of precipitation. The same software provides the immediate storage of the resulting data in text files.

The conservation of the cartograms represents a critical factor. As years pass by, the paper gets dusty, spoiled and worm-eaten, while the tracks fade and become less and less readable. The longer the delay in digitizing the cartograms, the more difficult it is to retrieve the recorded data and meteorological information stored therein.

The strip chart is converted into a file by means of a scanner as a true-colour (24 bit) image with a resolution of 150 to 200 dpi depending on the size and quality of the cartogram.

The tracks impressed on graph paper, once transformed into image files with the use of a scanner, are converted rapidly and accurately into numeric data files of a format chosen by the user adoption sampling times as low as 5 min.

By means of this software we obtained observations every 5 min interrupted by a brief gap every Monday, corresponding with the time taken by the manual procedure necessary to replace the paper form on the rotating drum.

Then it is possible to recover within reasonable times the vast information stored in the voluminous paper archive from chosen stations: Vercelli, Lombriasco and Pallanza. Rainfall data from Bra station are manually extracted from the charts. In the last case (Bra) the highest temporal resolution for the manual work is 1 h.

Quality checks were done on the digitalized dataset, including calculating daily totals from the dataset and comparing them with another independent daily precipitation dataset. The above procedure was carried out to prevent mistyping and to avoid mistakes.

3 Methods

The definition of what constitutes an extreme event has been debated. An extreme event may be selected based on frequency, intensity or threshold exceedance and physical expected impacts (Ntegeka and Willems, 2008). It depends on the intended
use in design or future planning. Afterwards we explained which kind of indices we adopted in this study.

### 3.1 Indices of climate extremes

In addition to yearly and seasonal maxima, two indices of extreme rainfall (Haylock and Nicholls, 2000) were calculated for each year in the period: the number of events above the long-term 95th percentile, referred to as the extreme frequency and the average intensity index of rain falling in the highest events, referred to as the extreme intensity.

The extreme frequency index examines changes in the number of extreme events. In calculating this index, the authors selected to use the mean 95th percentile (which varies for each station), rather than following the method of Karl et al. (1995) involving a fixed threshold for all stations. A fixed threshold is impractical for our study area with a high spatial variation in rainfall intensity. The index is calculated by counting the number of events in the year with intensities above this threshold. This approach is similar to the one used by Karl and Knight (1998) who considered changes in frequencies or probability of events above specified long-term percentiles. They proved that increases in total precipitation are strongly affected by increases in both frequency and intensity of heavy extreme precipitation events, and the proportion of total annual precipitation is derived from heavy and extreme precipitation events which have increased relative to more moderate precipitation (not heavy precipitation).

The extreme intensity index incorporates changes in all events above the upper percentile. This index was calculated using two different methods: averaging the highest four events for each year and averaging all events above the long-term 95th percentile.

### 3.2 Mann–Kendall test

All trends have been calculated using the statistic Mann–Kendall test. Where a trend is indicated as “significant”, it has at least 95% significance using this test.
The Mann–Kendall test (Kendall, 1962; Sneyers, 1990) is based on the comparison between the observed number of increases and decreases (jumps) and the values expected from random series. The occurrence of a trend is suggested if the null hypothesis of no trend is rejected when the level of significance is below a given threshold (here set at value $\alpha = 0.05$).

In this test, for each element $y_i$ the number $n_i$ of element $y_j$ preceding it ($i > j$) is calculated such that $y_i > y_j$.

The test statistic $t$ is then given by the equation:

$$t = \sum n_i$$

And its distribution function, under the null hypothesis is asymptotically normal, with mean and variance:

$$E(t) = \frac{n(n-1)}{4} \quad \text{and} \quad \text{var}(t) = \frac{n(n-1)(2n+5)}{72}$$

The null hypothesis must therefore be rejected for high values of $|u(t)|$ with:

$$u(t) = \frac{[t - E(t)]}{\sqrt{\text{var}(t)}}$$

In particular, if the probability $\alpha_1$ is determined using a standard normal distribution table like this one

$$\alpha_1 = P(|u| > |u(t)|)$$

The null hypothesis is accepted or rejected at the $\alpha_0$ level, depending on whether we have $\alpha_1 > \alpha_0$ or $\alpha_1 < \alpha_0$.

When the values of $u(t)$ are significant, an increasing or decreasing trend can be observed depending on whether $u(t) > 0$ or $u(t) < 0$.
3.3 Peaks-Over-Threshold

One of the commonly used extreme value sampling is to pick the highest value per year, hence it generates annual maximum series whose sample size is identical with the number of years (Ny). It does not include all extreme values because any second highest would be dropped out of the Ny samples. The other procedure chosen is called Peaks-Over-Threshold (POT).

In this case, extremes are extracted from a series by applying a threshold (in Sect. 3.1 the threshold was 95% long term percentile), which implies that the analysis is valid only for those values above a certain return period.

There is no standard method for the selection of the threshold and many subjective techniques can work well. Lang et al. (1999) proved that threshold selection is tightly linked to the over-threshold distribution and to the hypothesis of independence. Following the extreme value theory of Pickands (1975), the Generalized Pareto Distribution (GPD) models independent extremes extracted from a univariate series after applying a threshold.

For that reason we need to ascertain that the threshold is high enough to be in the asymptotic limit of the distribution of exceedances.

In the present study, the independency criterion is based on a procedure for extracting Peaks-Over-Threshold (POT) values for rainfall which is similar to that of extracting Peaks-Over-Threshold values for flows (Ntegeka and Willems, 2008). The independency criterion for rainfall events states that two consecutive events are independent if the occurrence of one event does not affect the occurrence of the other event. Ntegeka and Willems (2008) proposed for extreme value analysis based on rainfall series a minimum of 12 h inter-event time considering two events happening within the same day or night as one event.

The theoretical background of the POT method is based on the following fact: excesses over a high threshold $u$ can be modelled by a generalized Pareto distribution with the following distribution function:
\begin{align*}
G_{\xi,\beta}(x) &= 1 - \left(1 + \frac{\xi x}{\beta}\right)^{-\frac{1}{\xi}} \quad \text{if } \xi \neq 0 \\
G_{\xi,\beta}(x) &= 1 - \exp\left(-\frac{x}{\beta}\right) \quad \text{if } \xi = 0
\end{align*}

The procedure consists of choosing the subsequence \(\{X_j\}\) from the basic sequence that exceeds a threshold \(u\), calculating the values \(\{X_j - u\}\) for those values that exceed the threshold \(u\) and estimating parameters \(\xi\) and \(\beta\) either by the \(L\)-moment method (Hosking and Wallis, 1997).

\section{Results and discussion}

\subsection{Mann–Kendall test}

As mentioned above we tried to adopt several indices for heavy precipitation analysis, since there is not a single index to clarify the changes in the time series of heavy precipitation. Seasons were defined according to the standard meteorological definition: winter (DJF), spring (MAM), summer (JJA), autumn (SON).

\subsubsection{Annual maximum}

The application of the Mann–Kendall test demonstrates that annual maximum precipitation increased only for Bra with a time scale of 12 h (Fig. 2). For all the other stations and for all aggregation levels we do not have a statistical significant change.

\subsubsection{Seasonal maximum}

Regarding short duration (sub-hourly precipitation), trends are very weak and not significant. For long duration (from 1 h to 12 h), there is significant positive trend in spring maximum precipitation for Bra and Lombriasco (Table 2, Fig. 3).
4.1.3 Extreme frequency index

It can be stated that 30 min is the only time scale that shows a significant increase for the station of Pallanza. All other stations did not record significant change in the extreme frequency index for short duration.

The frequency of hourly heavy precipitation in the case of the station of Vercelli is increasing (Fig. 4).

Need to be highlighted that, although statistically not significant, the test values, regarding many aggregation levels less than 1 h (10 min, 15 min, 20 min and 30 min), indicated an increase of intensity and a decrease of frequency of this kind of extreme events registered in the station of Vercelli.

The stations of Bra and Lombriasco are the ones where we recorded similar trends regarding the extreme frequency index probably related to the geographic proximity of the two stations (Fig. 5).

4.1.4 Extreme intensity index

For the station of Vercelli we don’t have any significant trend for all aggregation levels. The station of Pallanza showed a significant increase of extreme intensity index for short duration, like 30 min: the same result as the extreme frequency indicator showed above. This increase in the intensity and frequency of rainfall events with short-time scale in the station of Pallanza is due to the fact that this station is situated is in the watershed of Lake Maggiore characterized by an escalation of extreme precipitation events which are very short in duration (Saidi et al., 2012).

Changes and increase of extreme precipitation frequency coincide with decrease of storm intensity in the case of heavy hourly precipitation registered in Bra (Fig. 6). This may pose a number of problems for water resource managers.

Frequently, the trends of the extreme intensity index calculated through the four highest events and the events exceeding the long-term 95th percentile are similar. The trend
is generally strongest when the index is calculated using the average of events above the 95th percentile.

In most cases, the positive sign of the trend in the extreme intensity index matches the trend in spring maximum precipitation.

4.2 Peaks-Over-Threshold series

In this study we applied the POT model to rainfall data collected in the stations of Vercelli, Lombriasco and Bra (the longest historical time series available, Table 1) in order to investigate changes in growth curve in the last 20 yr comparing the whole long time series which imply temporal changes in extreme storm precipitation. The growth curves parameters are extracted using $L$ moments approach (Hosking and Wallis, 1997).

The peak-over-threshold extremes are extracted using the R software (http://www.r-project.org/).

As this is a time series, we must select independent events above a threshold. First, we fix a relatively low threshold to extract more events. Thus, some of them are not extreme but regular events. This is necessary to select a reasonable threshold for the asymptotic approximation by a GPD.

From Fig. 7 a threshold value of 10 mm should be reasonable for series of 1 h precipitation registered in the station of Vercelli. The selected threshold must be low enough to have enough events above it to reduce variance while not too low as it increases the bias. Thus, in this case we can now re-extract events above the threshold 10 mm.

Figure 8 shows graphic diagnostics for the fitted model. It can be seen that this model seems to be appropriate (10 mm as selected threshold for 1 h precipitation).

Table 3 shows the selected thresholds for the stations of Vercelli. These values are starting point for calculating POT series for every timescale.

POT series extracted from precipitation data collected in the station of Vercelli with high resolution (5, 10, 15, 20, 30 min and 1, 2, 3, 6, 12 h) is used to produce growth curves with an extreme value distribution (GPD in this case).
Figure 9 shows that growth curves have steepened and recent rainfall events have risen during the last 20 yr (1984–2003) of our time series for short durations (less than one hour). This feature is much more evident in the station of Lombriasco (Fig. 9) if compared to the station of Vercelli (Fig. 10).

The more the time scale becomes greater, the more the POT series related to the events recorded in the last 20 yr approaches those obtained from the long time series until arriving to a time scale of 12 h where we noticed a decrease of these recent events in comparison with the past (Fig. 11).

5 Conclusions

The aim of the present study was analysing rainfall time series, as well as detecting potential trends and assessing their significance. It is well known that one of the biggest problems in performing analyses of extreme climate events for most of the globe is a lack of access to high-quality, long-term climate data with the appropriate time resolution for analyzing extreme events.

We adopted an automated recovery of rain data from paper records of tipping-bucket rain gauges regarding four sites situated in the Piedmont region in the north west of Italy: Vercelli, Lombriasco, Bra and Pallanza. We obtained long-term time series of precipitation with high temporal resolution: 5, 10, 15, 20 and 30 min and above: 1, 2, 3, 6 and 12 h.

For intense precipitation in Lombriasco and Bra the trend analysis has yielded substantial evidence of increasing trends in the extreme intensity index of these events. An increase was found also in spring, a season that is characterized by high weather activity.

Globally, we can say that the analysis of extreme short precipitation series from these 4 stations (Vercelli, Lombriasco, Bra and Pallanza) gave the following main results:
– On a yearly basis, maximum precipitation is stationary for all the stations and for all time scales, except for Bra, where we registered an increase of 12 h precipitations.

– On a seasonal basis, there is a significant positive trend in spring maximum precipitation for Bra and Lombriasco which are close to each other.

– In most cases, the positive sign of the trend in the extreme intensity index matches the trend in spring maximum precipitation.

– On a sub-hour scale we noticed a significant increase of the mean of events above 95th percentile of precipitation with duration of 20 min registered in the station Lombriasco.

– For the station of Pallanza, 30 min is the only time scale that presents a significant increase of extreme frequency and intensity indices. This is due to the fact that intense precipitation events registered in this area (Lake Maggiore watershed) are generally very short in terms of duration.

– The only significant decreasing trend was recorded in the station Bra for the precipitation of 1 h duration.

For the analyses presented in this paper, a Peak-Over-Threshold approach was chosen. The POT series from precipitation data collected in different stations has served to the estimate of the growth curve for each station and for each time scale. These growth curves showed that the short precipitation events have increased over the last 20 yr for a very high temporal resolution (from 5 min to 60 min). For the long duration (12 h) these events begin to decrease.

The results obtained are consistent with those provided by Brunetti et al. (2004) for Italy and Burlando (1989) for Florence.

It is anticipated that the research presented will be built upon to further examine the possibilities of:
– Linking trends in rainfall extremes to trends in floods using various case studies.
– Verifying if even weak trends in the mean of a distribution, which can go unnoticed, can cause surprising changes in the probability of exceedance of larger events and, hence, substantial changes in flood risk.

Acknowledgements. This study was carried out as part of the STRADA (climate change adaptation strategies for the management of natural risks in the border region) project, funded under the Cross-border Cooperation Programme INTERREG 2007–2013 Italy–Switzerland. We would like to thank prof. Younes Alila (University of British Columbia – Canada) for his insightful comments in this research.

References


**Table 1.** Main characteristics of meteorological station.

<table>
<thead>
<tr>
<th>Station</th>
<th>Elevation (m a.s.l)</th>
<th>Location UTM: X</th>
<th>Location UTM: Y</th>
<th>Observation period/year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bra</td>
<td>290</td>
<td>409 124</td>
<td>4 950 561</td>
<td>1933–2003</td>
</tr>
<tr>
<td>Vercelli</td>
<td>135</td>
<td>450 886</td>
<td>5 019 210</td>
<td>1927–2003</td>
</tr>
<tr>
<td>Lombriasco</td>
<td>241</td>
<td>392 509</td>
<td>4 966 637</td>
<td>1939–2003</td>
</tr>
<tr>
<td>Pallanza</td>
<td>211</td>
<td>465 025</td>
<td>5 086 015</td>
<td>1950–1991</td>
</tr>
</tbody>
</table>
**Table 2.** Result of the application of Mann–Kendall test to spring maximum precipitation.

<table>
<thead>
<tr>
<th>Timescale (h)</th>
<th>Pallanza</th>
<th>Vercelli</th>
<th>Lombriasco</th>
<th>Bra</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>2</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>3</td>
<td>NS</td>
<td>NS</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>6</td>
<td>NS</td>
<td>NS</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>12</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>+</td>
</tr>
</tbody>
</table>

NS: non significant.
+: significant level greater than 95%.
Table 3. Example of threshold above which POT series were derived: Vercelli.

<table>
<thead>
<tr>
<th>Timescale (h)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>6</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Threshold (mm)</td>
<td>10</td>
<td>13</td>
<td>15</td>
<td>25</td>
<td>34</td>
</tr>
</tbody>
</table>
Fig. 1. Example of rain gauge recorder chart.
Fig. 2. Five year moving average of yearly maximum precipitation for Bra (12 h precipitation).
Fig. 3. Five year moving average of spring maximum precipitation for Bra: (a) 3 h (b) 6 h and (c) 12 h.
Fig. 4. Hourly precipitations registered in the station of Vercelli: (a) Five year moving average of extreme frequency index (b) residuals.
Fig. 5. Five year moving average of extreme frequency index for 3 h precipitation: (a) Bra (b) Lombriasco.
Fig. 6. Five year moving average of extreme intensity index for Bra (mean of events over the 95th percentile: 1 h precipitation).
Fig. 7. Threshold selection for 1 h precipitation (Vercelli).
Fig. 8. Graphic diagnostics for 1 h precipitation (Vercelli).
Fig. 9. Changes in POT series of (a) 5 min and (b) 10 min duration compared to the last 20 yr: Lombriasco.
Fig. 10. Changes in POT series of (a) 5 min and (b) 10 min duration compared to the last 20 yr: Vercelli.
Fig. 11. Decrease in POT series of 12 h duration: (a) Vercelli (b) Bra.