INTERNATIONAL JOURNAL OF CLIMATOLOGY
Int. J. Climatol. **29**: 1883–1893 (2009)
Published online 22 April 2009 in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/joc.1921

Review

A comprehensive analysis of changes in precipitation regime in Tuscany

Simone Fatichi and Enrica Caporali*
Department of Civil and Environmental Engineering, University of Florence, Via S. Marta 3, 50139 Firenze, Italy

ABSTRACT: The analysis of changes in precipitation is nowadays of great interest because rainfall space-time distribution is considered one of the most important indexes in the natural climatic cycle. This study proposes an investigation of precipitation time series recorded in the region of Tuscany (Central Italy) in the period 1916–2003. Forty indexes are defined to evaluate changes in the precipitation patterns and trend detection is performed through statistical tests to assess the significance of temporal and spatial changes in the precipitation regime. The presented analysis does not show strong evidence of nonstationarity. A few indexes in a given number of stations of the analysed territory do show a slight trend, but the significance of those trends is lost once the multiple location testing issue is considered. The complexity of the climate in central Italy, i.e. the presence of numerous feedbacks might in fact distort or remove the consequences of global warming on the precipitation regime. Copyright © 2009 Royal Meteorological Society

KEY WORDS stationarity; climate change; trend; extreme rainfall; daily rainfall; spatial analysis; multiple testing

Received 3 August 2008; Revised 31 January 2009; Accepted 15 March 2009

1. Introduction

The prevailing studies in the field of climatology basically follow two main paths. The first is the development and application of the General Circulation Models (GCMs), aimed at forecasting climate evolution in the next decades or centuries by means of simulations. The second is the analysis of long-range data series in order to ascertain whether signals of climatic change related to the increasing greenhouse effect can already be detected. As pointed out by Piervitali *et al.* (1997), attention is focused primarily on climatological variables such as air temperature, precipitation amount, ice coverage, pressure, sea level, and others. Nevertheless, the most investigated variables are global and local temperatures and rainfall space–time distribution (Burlando and Rosso, 2002). Interest in the time series analysis of those variables is nowadays significant because it is widely recognized that the natural and anthropogenic production of greenhouse gases induces changes in the natural climatic cycle. In the last few years, several reports on climate change, compiled by governmental and scientific institutions and widely diffused in the media, have had a strong impact on public opinion worldwide. The most popular of these are the technical report of the European Environment Agency (2007); the communication of the Commission of the European Communities (2007); and especially the Fourth Assessment Reports of the Intergovernmental Panel on Climate Change (2007). All the above-mentioned reports confirm the negative effects of climate change on environment and society, particularly with reference to water resource availability, increased desertification, loss of biodiversity, rise in sea level, and changes in agricultural productivity.

At the beginning of climate change studies, greater attention was paid to the broader effects of climate change (Manabe and Wetherald, 1987). Nowadays, interest has been transferred to both local and global effects; thus a lot of researchers have investigated potential climate change in local areas, and these studies are considered an important step towards knowledge of climate trends.

Special attention has been paid to the potential changes in the extreme events that could accompany global climate change; these constitute a primary concern in estimating the impacts of climate change. Extreme events, such as heat waves, heavy rain, hailstorms, snowfall, and droughts, are in fact responsible for a large proportion of climate-related damages (Meehl *et al.*, 2000), and their impact is of great concern for the community and stakeholders (Easterling *et al.*, 2000). A number of theoretical modelling and empirical analyses have also suggested that notable changes in the frequency and intensity of extreme events, including floods, may occur even when there are only small changes in climate (Katz and Brown, 1992; Wagner, 1996). A comparison of nine global GCMs shows a tendency towards an increase in

---

* Correspondence to: Enrica Caporali, Department of Civil and Environmental Engineering, University of Florence, Via S. Marta 3, 50139 Firenze, Italy. E-mail: enrica.caporali@unifi.it.

extreme events (Tebaldi *et al*., 2006). Scientific observations also indicate that an increase in extreme rainfall events is observable in many areas all around the world, including northern and southern Italy (Hopkins and Holland, 1997; De Michele *et al*., 1998; Mason *et al*., 1999; Brunetti *et al*., 2000; Fowler and Kilsby, 2003; Groisman *et al*., 2004; Bonaccorso *et al*., 2005). Even in Tuscany, a region located in central Italy, some authors have identified signals of climate change by analyzing extreme rainfall and hailstorms (Crisci *et al*., 2002; Piani *et al*., 2005).

Besides the great interest in extreme events, daily precipitation and drought indexes are also often investigated, as confirmed by the numerous references (Naidu *et al*., 1999; De Luìs *et al*., 2000; Brunetti *et al*., 2001; González-Hidalgo *et al*., 2003; Fujibe, *et al*., 2006).

Although there is evidence that it is easier to identify changes in extreme events rather than changes in the mean values of climatic variables (Storch and Zwiers, 1988), both the extreme and the average precipitation indexes are analysed here. Statistical tests to assess the significance of temporal changes are used to evaluate trends in the indexes derived from the recorded time series of the precipitation regime. A comprehensive analysis of the results obtained for the region of Tuscany in Italy is provided in the following paragraphs. It includes on-site analysis for all the indexes and a particular spatial analysis for the Total Annual Precipitation (TAP). Trend analysis of hydroclimatic variables is in fact considered a useful tool for effective water resource planning, design, and management, since it provides helpful information on the possible tendencies of time series and thus enables experts to choose the proper mitigation procedures.

2. Size of the dataset

The analysed territory is the region of Tuscany in central Italy. This area is expected to suffer significantly from global climate change (Burlando and Rosso, 2002).

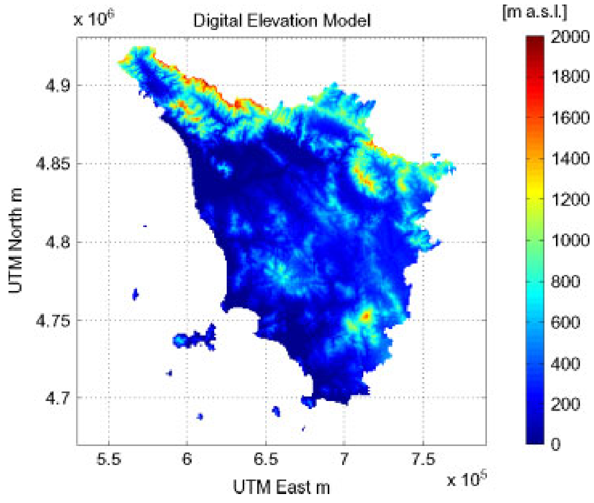

Figure 1. Geographical location and digital elevation model of Tuscany region. This figure is available in colour online at www.interscience.wiley.com/ijoc

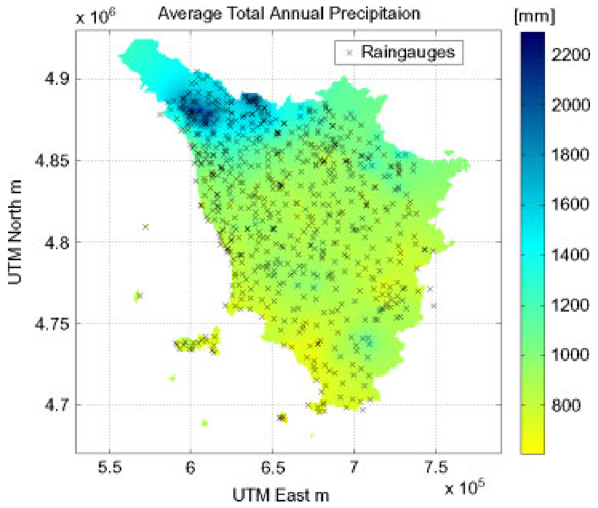

Figure 2. Average total annual precipitation and rain gauges distribution (daily precipitation dataset). This figure is available in colour online at www.interscience.wiley.com/ijoc

The region, covering an area of about 23 000 km^2, is characterized by a climate that ranges from temperate to Mediterranean maritime and by a complex physical topography. It presents plain areas near the sea and around the main river valleys, hilly internal zones, and the mountainous area of the Apennines (Figure 1).

The precipitation regime is greatly influenced by the topography. TAP ranges from 530 to 2600 mm (Figure 2). Heavy storms mainly occur in autumn following dry summers. Several areas of Tuscany have suffered in the past from many severe hydrogeological events.

The dataset available is subdivided into two different sets of time series. The time series of the precipitation recorded in 785 rain gauges were used to analyse the average rainfall regime (Figure 3(a)). The measurements cover the period 1916–2003. The available series length ranges from 1 to 84 years. The selected gauges cover almost the whole of Tuscany. Only the rain gauges with more than 40 years of data (a total of 200 rain gauges) are considered for the analysis. The time series of annual rainfall maxima for five different durations (1, 3, 6, 12, and 24 h), extracted from the historical data series recorded at every rain gauge, are instead considered to analyse the extreme rainfall. This dataset contains 445 stations covering the period 1928–2002. Stations are selected on the basis of completeness of records and are chosen only when at least 50 years of records are available. After this screening, 60 rain gauges remain for the analysis (Figure 3(b)).

3. Methodology

The methodology proposed by the authors follows the procedure employed by other researchers (Fujibe *et al*., 2006; Ramos and Martinez-Casanovas, 2006). The precipitation variable is investigated by means of indexes (the simplest index being the hydrological variable itself), and the time series of the indexes are tested to detect

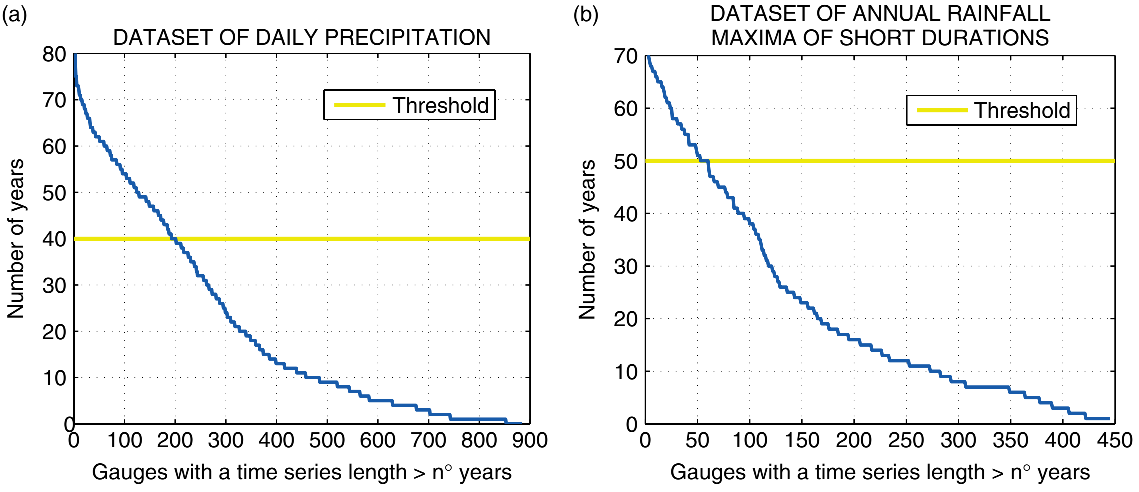

Figure 3. Size of the dataset of daily precipitation (a) and annual rainfall maxima of short durations (i.e. 1 h, 3, 6, 12, and 24 h) (b). This figure is available in colour online at www.interscience.wiley.com/ijoc

trends. The types of indexes and the statistical tests adopted are described in the following sections.

3.1. Definition of indexes

3.1.1. Daily precipitation

The indexes chosen to evaluate changes in the precipitation patterns are able to represent a wide variety of precipitation characteristics for both the average regime and the extreme behaviour of the rainfall process. Recent studies (Alexander *et al*., 2006) have pointed out limitations of the use of indexes due to the choice of fixed thresholds, arguing that the indexes have problematic statistical properties and are not as significant as claimed. However, the intent of this study is to document a possible nonstationarity of the precipitation regime, and the introduction of some indexes is unavoidable. The chosen indexes are in fact also used in other studies (Tebaldi *et al*., 2006), and recommended by the World Meteorological Organization and the Commission for Climatology and the Research Programme on Climate Variability and Predictability (Ramos and Martínez-Casanovas, 2006).

The 40 indexes considered in order to test the changes in the daily precipitation series are shown in Table I. Since the Precipitation Concentration Index (PCI) represents how precipitation is distributed over the year, it is worth some further explanation (Oliver, 1980). In the study, the PCI is evaluated with the modified expression of De Luis *et al*. (2000). A value of PCI below 10 suggests uniform distribution of the rainfall during the year, values between 10 and 20 indicate seasonality, and values above 20 indicate marked irregularity of the precipitation over the year. Percentile values of the daily precipitation sample distribution are calculated for each year using all wet days, i.e. daily precipitation (P) greater than 1 mm.

Table I. Daily precipitation (P) indexes.

1	Total annual precipitation (TAP)
2–13	Monthly rainfall precipitation
14	Number of wet days ($P > 1$ mm)
15–26	Monthly number of wet days ($P > 1$ mm)
27	Precipitation concentration index (PCI)
28	1 day precipitation annual maxima
29	5 days precipitation annual maxima
30	10 days precipitation annual maxima
31	The 75th percentile of precipitation
32	The 95th percentile of precipitation
33	The 99th percentile of precipitation
34	Number of days with $P >$ 75th percentile (moderate wet days), evaluate on the whole year
35	Number of days with $P >$ 95th percentile (very wet days), evaluate on the whole year
36	Number of days with $P >$ 99th percentile (extreme events), evaluate on the whole year
37	Fraction of annual total precipitation due to events exceeding the 95th percentile
38	Fraction of annual total precipitation due to events exceeding the 99th percentile
39	Number of days with precipitation larger than 10 mm ($P > 10$ mm)
40	Maximum number of consecutive dry days ($P < 1$ mm), largest dry period within a year

3.1.2. Annual rainfall maxima of short duration

As far as the extreme values of precipitation are concerned, the time series were fitted through two different extreme value distributions, instead of using the raw time series of annual extreme values, in order to smooth the data and test the change in the newly obtained time series. The probability distributions adopted were the extreme value type I distribution and the Generalized Extreme Value (GEV) distribution, as described by many authors (Jenkinson, 1955; Katz *et al*., 2002, 2005) and commonly used for rainfall frequency analysis. GEV distribution in particular has already been employed for climatic variation detection (Mason and Mimmack, 1995; Crisci *et al*., 2002).

Even if the choice of the distribution is not critical, as the study is not aimed at accurately evaluating the precipitation value for a certain return period (RP), but rather at estimating if this value is going to change (Bonaccorso *et al*., 2005), the authors analysed the time series with three different approaches. In the first, the Gumbel distribution was used and the associated parameters were determined with the method of moments. In the second and third approaches we applied the GEV distribution. The three parameters were calculated respectively with the probability method of moments and the L-moment method (Hosking, 1990). Three approaches were employed because, given the large number of fittings, it is not possible to decide *a priori* which one leads to the best approximation of the data.

The precipitations for each rain gauge at all storm durations (i.e. 1 h, 3, 6, 12, 24 h), and with a RP of 10 and 20 years, were derived from the annual rainfall maxima time series with the three above-mentioned approaches, using a 10-year moving time window with 1-year steps. The extreme value analysis with the determination of the precipitation height for a 10-year and 20-year RP for each station was in fact carried out on each moving time window, creating two time series of estimated extreme precipitation (one for 10-year and one for 20-year RP) of a length which is equal to the original time series minus 9 years. These resulting time series were considered as new indexes and were tested to give evidence of changes.

3.1.3. Spatial analysis

A methodology that uses as much data as possible, including the gauges with the shorter time series (i.e. only 1 year long), was proposed to calculate the TAP in Tuscany.

The Tuscany region is divided into a regular 1000 m grid. For each year, from 1916 to 2003, the spatial distribution of TAP was calculated by means of a spatial interpolation on the rain gauges working in that year. The years with less than 100 working stations were excluded. The time series of TAP, for each cell of the grid, could be obtained from this procedure. This includes more information than the traditional method, which considers only the rain gauges with a length of time series superior to a threshold value (40 years in this study). Specifically, the percentage of data considered in the two different methods of calculation varied from 59.4 to 97%, with a remarkable improvement in the second one.

The time series of TAP for each cell and the average value of TAP over the whole Tuscany territory (Figure 4) were investigated as additional indexes to compare the results with those of the other indexes.

Together with the specific spatial analysis of TAP, a common application of the interpolation technique was used to provide further information about the spatial patterns of the changes detected in the indexes. Interpolation maps of the linear correlation coefficient are presented in the following paragraphs.

3.2. Statistical tests

Four different statistical tests were adopted to investigate the presence of trends in the time series of the indexes described earlier.

The first test was a parametric one based on the Pearson linear correlation and the F-Fisher test to compare regression variance and error variance.

The second test was the nonparametric Mann-Kendall (MK) statistical test (Mann, 1945; Kendall, 1975). This test allows us to investigate long-term trends of data without assuming any particular distribution. Moreover, it is less influenced by outliers in the dataset as it is nonparametric. The null hypothesis is trend absence in the analysed variable.

The third test used was the Cox and Stuart (1955) test for trend that allows us to verify if a variable has a monotonical tendency (rejection of null hypothesis of trend absence), and its variability. The test is very close

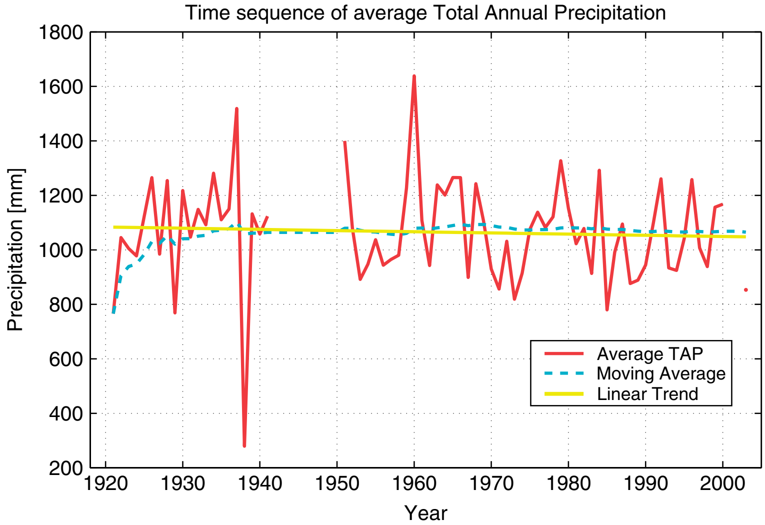

Figure 4. Time sequence of average total annual precipitation on the Tuscany region. This figure is available in colour online at www.interscience.wiley.com/ijoc

DOI: 10.1002/joc

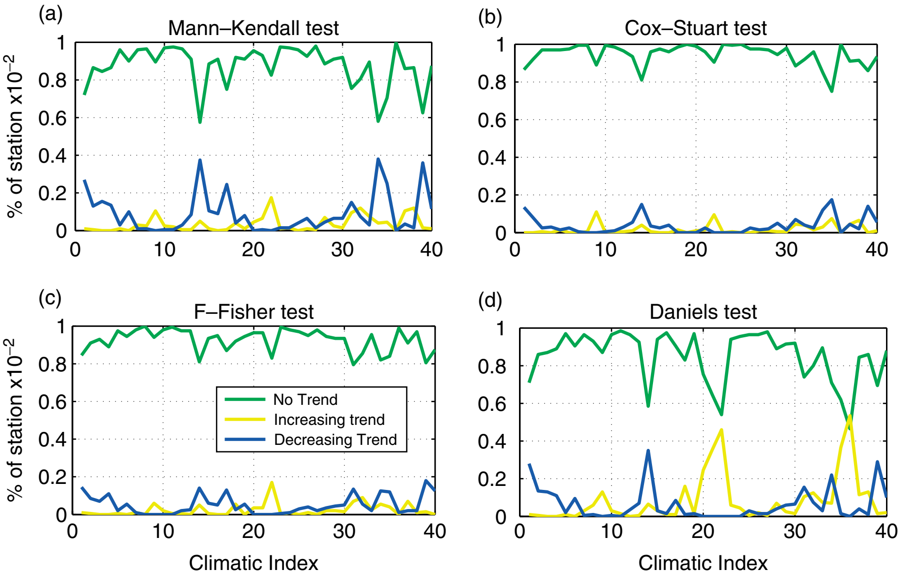

Figure 5. Percentage of gauges with increasing, decreasing, and no trend for the 40 indexes of daily precipitation. FPR hypothesis testing method. (a) Mann-Kendall test, (b) Cox-Stuart test, (c) Fisher test, and (d) Daniels test. This figure is available in colour online at www.interscience.wiley.com/ijoc

to the sign test for two independent samples. More details are given in Conover (1999).

The last test used by the authors was the Daniels (1950) test, which uses the ρ Spearman test to verify if a variable changes monotonically in time, with the hypotheses of both one side and two sides, against the null hypothesis of tendency absence.

The use of so many tests was rendered necessary by the uncertainty of the probability distribution of the analysed variables, and especially by the need to distinguish between the influence of the trend detection methodology adopted and the actual behaviour of the data (Yue *et al.*, 2002).

The above-mentioned tests present different peculiarities. Conover (1999) states that the trend tests based on ρ Spearman and τ Kendall are in general more powerful than the Cox-Stuart test. As previously shown by Stuart (1956), when data distribution is normal, in comparison with the r Pearson linear correlation coefficient, the asymptotic relative efficiency (ARE) of the ρ Spearman test and τ Kendall test is equal to 0.98, whereas for the Cox-Stuart test the ARE is 0.78. The Cox-Stuart test is less accurate than the tests that use nonparametric correlation, because it is based on signs rather than ranks. Nevertheless, the Cox-Stuart test is more general and could also be applied in the presence of strong anomalies and cyclical data. 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. In particular, for each time series the nonexceedance probability of the test statistic value, computed on the sample, is compared with the significance level $\alpha = 0.05$ for a symmetrical two-tailed test. Therefore, the presence of trends for all the indexes and for all the considered rain gauges is detected when the null hypothesis is rejected. The sign of the trend is inferred from the linear regression coefficient. On a regional scale, the ratio between the positive, negative and no trend stations is reported as summary results for each index. In addition, the distribution of the linear correlation coefficient for all the indexes is tested to be normal with mean equal to zero. To highlight which indexes might be considered without trend, two hypothesis tests are applied: the χ^2 test and the Kolmogorov-Smirnov test.

3.3. Multiple testing issue

To test trends at multiple locations we have to simultaneously evaluate many hypothesis tests. This automatically implies that if the hypothesis test with significance α is performed in n different locations for which the null hypothesis is true (absence of trends in this study), the number of rejections of the null hypothesis (that is true), type I error, is on average $n\alpha$. When the sample is large enough this leads to the detection of a significant number of false positives (trend presence in this study). The problem may be solved by decreasing the α significance level, but the drawback of this procedure is an increase in the acceptance of the null hypothesis when it is false (incapacity to detect trends), type II error, thereby reducing the power of the test. Compromises to avoid the two types of error are a well-known subject in the literature, and several approaches have been developed to tackle the multiple test problem (Livezey and Chen, 1983; Benjamini and Hochberg, 1995; Benjamini and Yekutieli, 2001). For a recent review of different approaches and further insights, see Ventura *et al.* (2004).

In the preliminary study, the classical approach is adopted in which the presence of multiple locations is neglected and the significance level α is fixed. We refer to this method henceforth as false positive rate (FPR) (Ventura *et al.*, 2004). Subsequently, two different

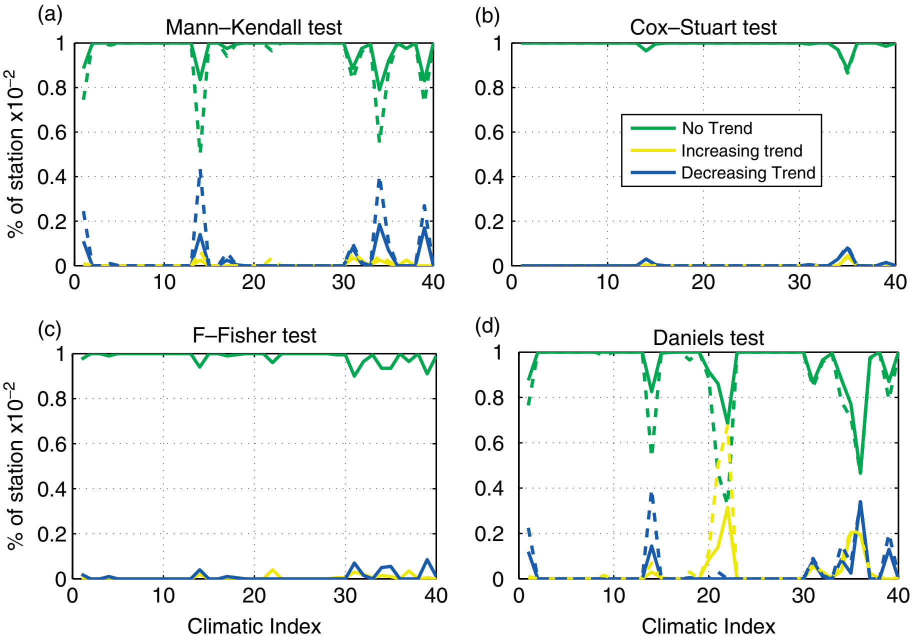

Figure 6. Percentage of gauges with increasing, decreasing, and no trend for the 40 indexes of daily precipitation. FDR hypothesis testing method (solid lines); FDR-mod hypothesis testing method (dashed lines). (a) Mann-Kendall test, (b) Cox-Stuart test, (c) Fisher test, and (d) Daniels test. This figure is available in colour online at www.interscience.wiley.com/ijoc

multiple testing methods are applied: the false discovery rate (FDR) as proposed by Benjamini and Hochberg (1995) and the successive modification introduced by Ventura *et al.* (2004), false discovery rate modified (FDR-mod), which gives the FDR tighter control and increases the power to detect significant changes. Both methods are described in Ventura *et al.* (2004) and also seem to fit well for spatial correlated data like the ones present in this analysis. The adopted nominal FDR rate q is 0.05 and the a_n *a priori* unknown proportion of true alternative hypothesis (FDR-mod test) is evaluated as in Ventura *et al.* (2004).

4. Results and discussion

In the way the analysis is designed, the detection of trends considers each single station independently in the FPR method and considers the distribution of the p value in the FDR and in the FDR-mod methods. What is most interesting in such an analysis is not the trend detection in a single gauge or in few gauges, which may contrast with nearest gauges, but the summary of the results in terms of changes in the precipitation regime over the whole region of Tuscany or in a significant part of this region. The use of several different indexes, numerous tests, and methods of hypothesis testing should avoid the possibility of trends being detected due to the shortness of time series or by chance, rather than due to the actual behaviour of the time series. In the paragraphs below, the authors specifically address the question of change detection in the average precipitation regime, in the seasonality of precipitation, in the drought period length, and in the intensity of extreme precipitation. Finally, a spatial analysis is given to include a larger number of data and to provide insight into the changes in TAP spatial distribution.

4.1. Daily precipitation trend analysis

The tests on the time series of the 40 indexes proposed to investigate the daily precipitation dataset are summarized in Figure 5 for the classical method of FPR hypothesis testing and in Figure 6 for the FDR and FDR-mod methods, where the percentage of stations that present an increasing, decreasing, and absence of trend is shown. The agreement between the four different chosen tests is noticeable for all the proposed hypothesis testing approaches, with a higher amount of refused null hypothesis for the Cox-Stuart and Fisher tests as compared to the MK and Daniels tests.

Since practically all the tests give the same results, these are substantially independent of the specific test adopted, and comments will refer to the general tendency and not to the specific tests.

The graphs provided by the canonical FPR hypothesis testing method (Figure 5) are ambiguous since most stations show an absence of trend; however, there is a small number of stations where a significant decrease but also sometimes an increase in the analysed index is detected. This outcome, as seen in Figure 6, is mainly an artefact of the FPR method. In fact, if the problem related to the multiple location test is taken into account, we have further confirmation of the almost total absence of trends in the analysed indexes. As expected, the FDR method is more conservative and therefore less powerful in statistical terms (Ventura *et al.*, 2004) than the FDR-mod method, providing a lower number of null hypothesis refusals; nevertheless, the results of the two

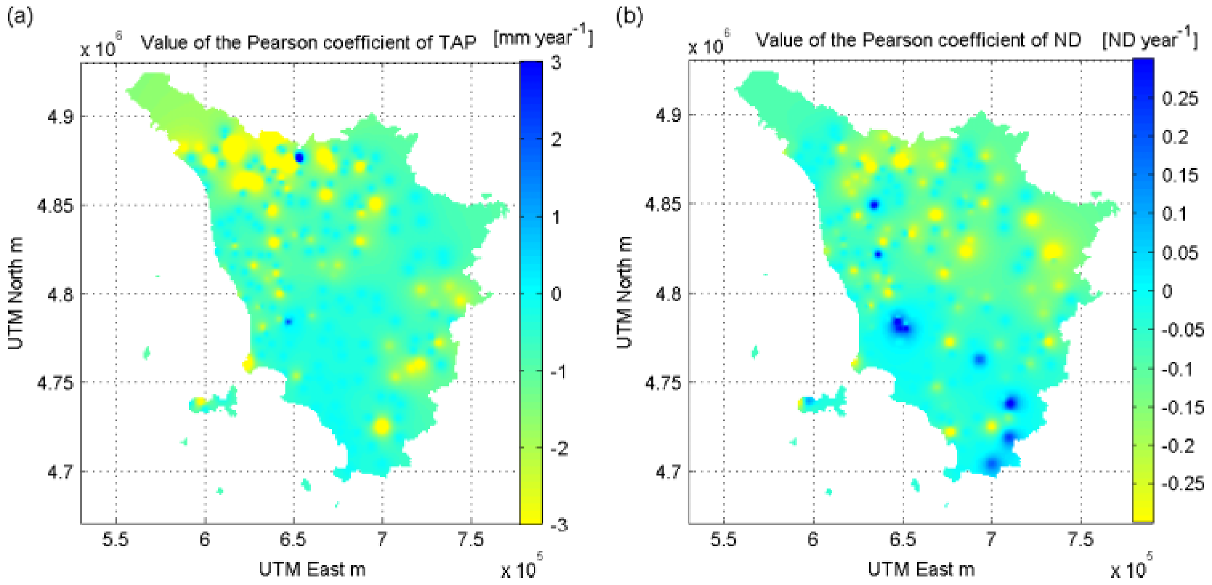

Figure 7. Spatial distribution of the Pearson coefficient, statistically significant (Mann-Kendall test, FPR), for the time series of (a) total annual precipitation (TAP) and (b) number of wet days (ND). Interpolation method: IDW. This figure is available in colour online at www.interscience.wiley.com/ijoc

methods are very close to each other and they contribute to disproving an alteration in the precipitation regime in Tuscany for the analysed period.

Only a very slight general decreasing trend may be identified for a few stations (with the MK and Daniels test), concerning the indexes 1, 14, 31, 34, and 39. A decreasing trend in those indexes refers to a reduction in the number of wet days (indexes 14–34–39) and in the total precipitation (index 1). In any case, this trend detection is not particularly significant since it affects only a few stations, and there is no spatial cluster among the stations, as is shown in the spatial analysis results (Figure 7). A spatial interpolation of the value of the linear correlation coefficient is provided only for the TAP and the total number of wet days (Figure 7). These indexes are the most general and inclusive of the precipitation regime. Precise clusters of stations with the same trend do not seem to be present. The areas of northwestern Tuscany are more affected by a negative linear correlation coefficient of TAP and the number of wet days. The values in central and southern Tuscany are more contradictory. Substantially, there is no clear spatial relationship between the results obtained from the single rain gauge.

The second analysis performed was to test if the distribution of the Pearson coefficient for each index fits a normal distribution with an average equal to zero. The test was carried out by means of two hypothesis tests, the χ^2 test and Kolmogorov-Smirnov test, with a level of significance of $\alpha = 0.05$. These tests allow us to establish if an index presents a linear correlation coefficient distribution with an average equal to zero, hence the predominance of a tendency can be excluded for that index. The contrary is not true because, even if the average of the linear correlation coefficient is close to zero, its distribution might be different from normal and the hypothesis of fitting could be rejected. The outcomes obtained were less significant than the authors' expectations; however, through these tests any tendency is excluded in 9 of the 40 indexes. The results provided by the χ^2 test and the Kolmogorov-Smirnov test are in good agreement.

Therefore, the answer to the question of modification in the daily precipitation dataset in the Tuscany region is that a general nonstationarity or alteration is not identifiable. The average precipitation regime and the length of the drought period did not undergo changes during the analysed period. The seasonality of precipitation seems to be statistically the same, since the analysis of PCI as well as the precipitation and the number of wet days for different months reveals a total absence of trends.

4.2. Annual rainfall maxima trend analysis

Alterations in the intensity of extreme rainfall were investigated through the indexes of the extreme precipitation, i.e. the time series provided by the extreme value analysis were carried out with the methodology described in Section 3.1.2.

An analysis of sensitivity to the distribution fitting and to the RP was also carried out. The results obtained with the three methods of fitting for the extreme value analysis, Gumbel moments, GEV moments, and GEV-L moments are almost the same, confirming the findings of Bonaccorso *et al.* (2005). The choice of the RP was also not critical, as the analysis was focused on change detection rather than on value estimation. The outcomes for the 10-year or 20-year RP are indeed almost the same. The analysis of sensitivity to the chosen time window was not performed in this work, but it is recommended for future applications. The results are shown in Figures 8 and 9 for the GEV distribution – 10-year RP, but must be considered indicative of the whole annual rainfall maxima analysis. The five climatic indexes shown in the figures are in

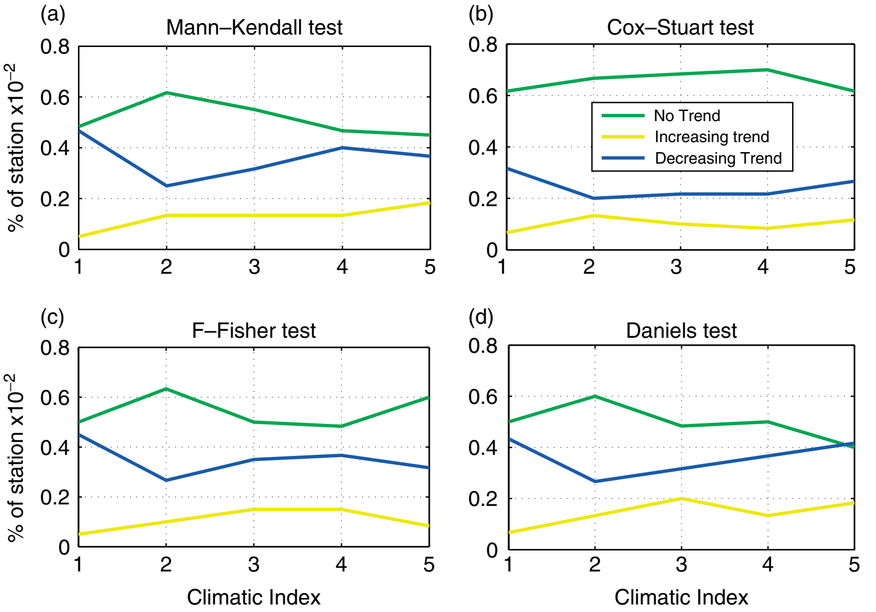

Figure 8. Percentage of gauges with increasing, decreasing, and no trend for the five durations of annual rainfall maxima obtained with a 10-year return period GEV distribution. FPR hypothesis testing method. Climatic indexes (1–5) are the five durations (1-3-6-12-24 h). (a) Mann-Kendall test, (b) Cox-Stuart test, (c) Fisher test, and (d) Daniels test. This figure is available in colour online at www.interscience.wiley.com/ijoc

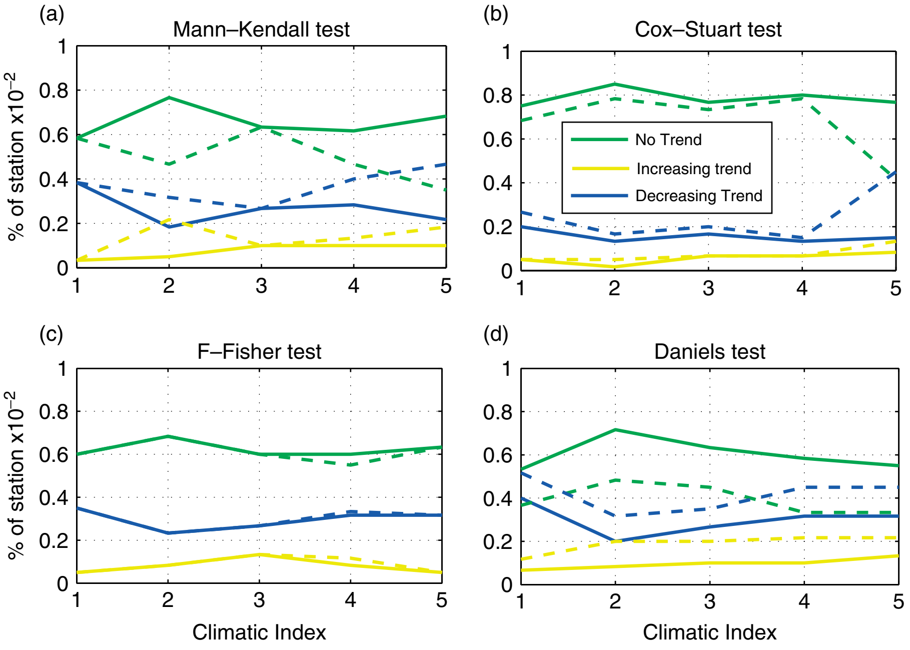

Figure 9. Percentage of gauges with increasing, decreasing, and no trend for the five durations of annual rainfall maxima obtained with a 10-year return period GEV distribution. FDR hypothesis testing method (solid lines); FDR-mod hypothesis testing method (dashed lines). Climatic indexes (1–5) are the five durations (1-3-6-12-24 h). (a) Mann-Kendall test, (b) Cox-Stuart test, (c) Fisher test, and (d) Daniels test. This figure is available in colour online at www.interscience.wiley.com/ijoc

this case the extreme value time series (GEV 10 years) for the five different durations of 1-3-6-12-24 h.

In comparison with the analysis carried out on the daily precipitation, the null hypothesis of no trend was accepted only, on average, for 50% of the total in the FPR multiple hypothesis testing method, and almost 65% of cases with the FDR method, the results of the FDR-mod method lying as expected somewhere between those of the other two methods. The differences between the three methods were smaller than in other cases because the sample consisted of only 40 stations, thereby reducing the problem of multiple site testing.

The detection of a clear tendency on a regional scale is a far from trivial task. The rain gauges presenting negative trends were more than those presenting positive ones; nevertheless, this prevalence was not so high as to indicate a pattern of reduction in the intensity of extreme events. Only for the 1-hour and 24-hour durations of

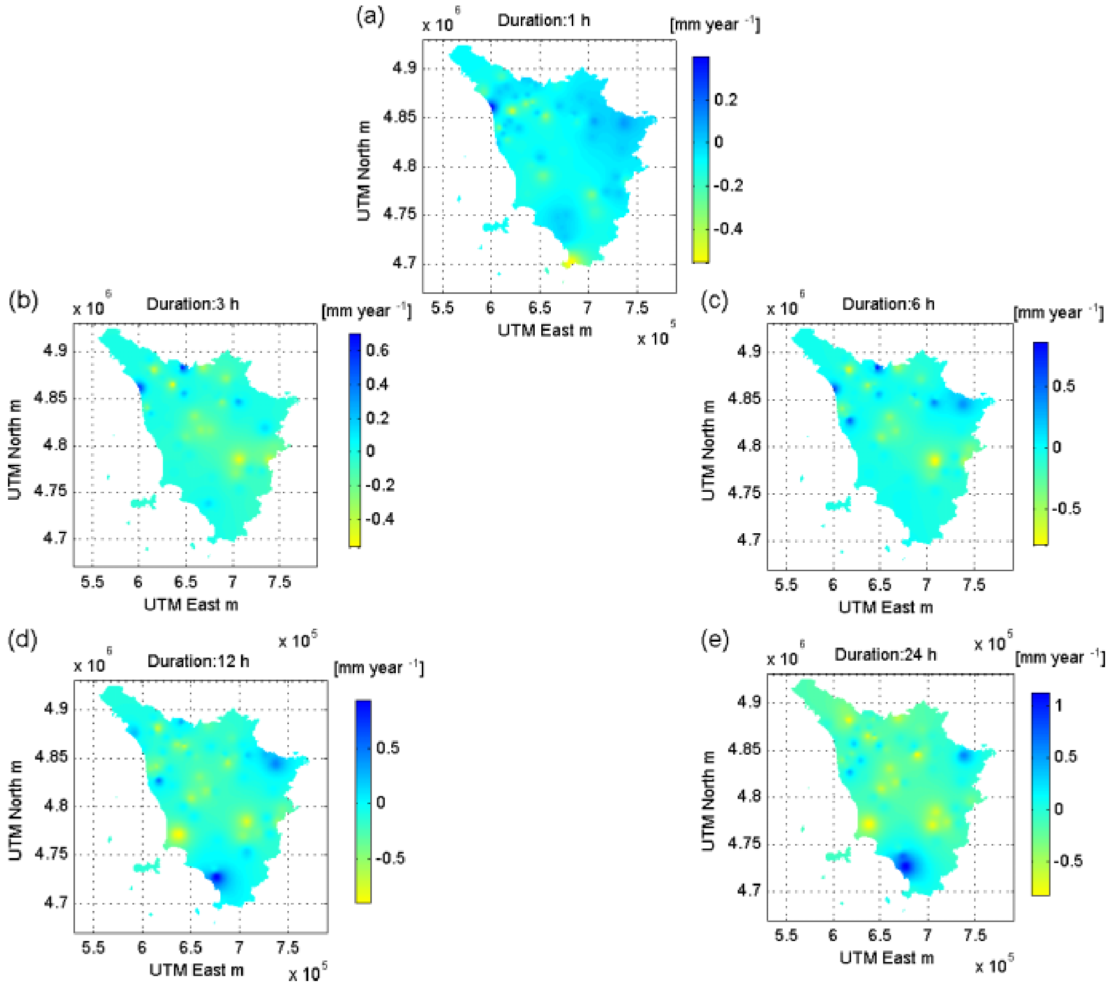

Figure 10. Spatial distribution of the Pearson coefficient, statistically significant (Mann-Kendall test, FPR), for the time series of annual rainfall maxima obtained for a return period of 10 years with the GEV distribution: (a) 1 h, (b) 3 h, (c) 6 h, (d) 12 h, and (e) 24 h. Interpolation method: IDW. This figure is available in colour online at www.interscience.wiley.com/ijoc

might a hypothesis of a decrease in rainfall intensity be taken into consideration. The spatial interpolation (Figure 10) shows that the zone most affected by an extreme reduction is the central part of Tuscany, thus no clear spatial pattern for the Pearson coefficient appears. As expected, the variability and absolute value of the linear correlation coefficient increased with the duration.

As for the daily precipitation indexes, the distribution of the Pearson coefficient was tested for each duration to fit a normal distribution with an average equal to zero. The control, carried out through the two hypothesis tests, i.e. the χ^2 test and the Kolmogorov-Smirnov test ($\alpha = 0.05$), confirmed the indications provided by the other tests, excluding trends for the 3-, 6-, and 12-h durations. This analysis does not provide an unequivocal answer as to whether there is a pattern in the modification of extreme precipitation in Tuscany, even if the preliminary evidence suggests a substantial stationarity also in the extreme precipitation regime.

4.3. Spatial analysis discussion

Following what is described in Section 3.1.3, the time series of TAP produced for each cell (almost 22 000 time series) and the average value of TAP over the entire territory of Tuscany were considered as further indexes to be investigated. The same four tests of trend detection were used. The analysis of sensitivity to the spatial interpolation method (Table II) shows some differences between the nearest-neighbour method and the inverse distance weight (IDW) method, as demonstrated also by the results obtained for the mean of the average value of TAP in Tuscany (Table II). The results obtained with the FPR method, coupled with the nearest-neighbour interpolation, show a decreasing trend for about 12–25% of the territory (depending on the test), and practically no increasing trend (Table III). The IDW method tends to smooth the precipitation distribution and only 1–15% of the territory shows a decreasing trend (Table III). Since in this case the number of time series analysed is very large, the multiple site testing issue plays an important role; several rejections of null hypothesis may be related to a type I error. The same analysis using the FDR and FDR-mod methods provides a complete absence of significant trends throughout the Tuscany region, confirming once more that statistically significant nonstationarity cannot be found for the TAP during the last 70–80 years. The above statement is confirmed by the fact that the time sequence of average TAP in Tuscany is statistically not significant (Figure 4).

The spatial pattern of the Pearson coefficient (Figure 11) shows a general decrease but the most significant outcome is the almost total absence of positive values. An increase in precipitation is present only in limited zones located mainly on the top of the mountain relief in the north of the region. This could suggest an increase in the average thermoconvective activity of the spring–summer storms due to the higher temperature recorded in the last few years (Toreti and Desiato, 2007), without affecting the annual rainfall maxima.

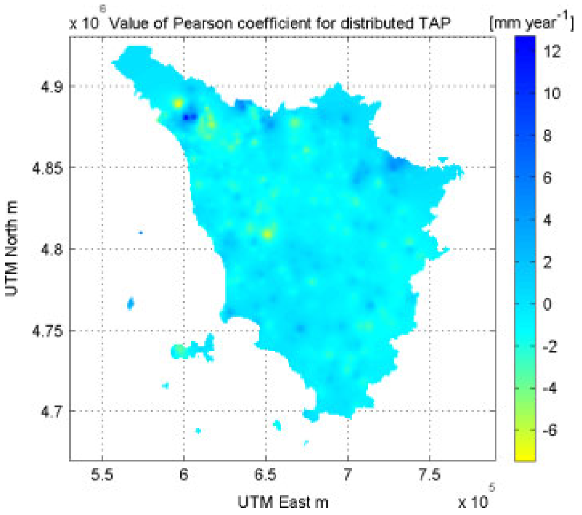

Figure 11. Spatial distribution of the Pearson coefficient for the time series of total annual precipitation (TAP) calculated with the methodology proposed by the authors. Interpolation method: IDW. This figure is available in colour online at www.interscience.wiley.com/ijoc

5. Conclusions

Providing a reliable analysis of long-term changes in hydroclimatic variables is quite a complex task. Several problems like data inhomogeneity, changes in instrumentation, and the lack of long time series contribute to the reduction of the substance of this type of analysis. These aspects should be considered in the choice of the data analysis procedure and in interpretation of the results to reduce the uncertainties and the difficulties involved in reaching a final conclusion. In this study, the comprehensive analysis of the precipitation regime carried out in Tuscany in the period 1916–2003 (although most data derive from the period 1950–2000) does not appear to offer any evidence of nonstationarity. There is practically an absence of trends in the average precipitation regime and in the intensity of extreme events of 3–6 and 12 h in almost all the stations of Tuscany. These results are reinforced once the multiple site testing issue is taken into account in the hypothesis testing. All the evidence is supported by the results of the spatial analysis, where significant changes in rainfall regime cannot be confirmed even in the sub-zones of Tuscany. The use of a great number of indexes, especially the PCI and the monthly subdivision of the total precipitation and of the number of wet days, clearly confirms that changes in the seasonal distribution of precipitation are not statistically significant.

The signs of climate change as illustrated by several authors and reports are not detected in Tuscany for the analysed dataset, even if other zones of Mediterranean area show signals of global warming impacts (Piervitali *et al*., 1997; Norrant and Douguédroit, 2006). The authors presume that the presence of numerous feedbacks could delay or delete the consequences of global warming on the precipitation regime, especially in a complex climatic system like central Italy. Future scenarios for extreme rainfall events and drought cannot be inferred from this analysis, which focuses only on the recorded data. Similar studies together with up-to-date rainfall measurements in Tuscany are recommended, as well as a continuous monitoring of the atmospheric phenomena, with the aim of checking possible anomalies such as modifications in the precipitation distribution over the year that strongly impacts on the ecosystem and the water supply activity.

Table II. Annual average precipitation depth (*mm*), on the Tuscany region, calculated with the average value of the stations with more than (1) 40 years of annual registration, (2) 20 years of annual registration, and (3) with the 1-year interpolation meshes with more than 100 stations.

	Interpolation method			
	Nearest	Linear	Cubic	IDW
TAP (1) mm	1024	997	997	1004
TAP (2) mm	1139	1025	1022	1071
TAP (3) mm	1109	1016	1018	1065

Table III. Fraction of Tuscany region with positive, negative, and no trend detected by means of the four tests for the total annual precipitation. FPR hypothesis testing method.

Average TAP 'IDW method'	No trend	Increasing trend	Decreasing trend
Mann and Kendall	0.891	0.009	0.098
Cox-Stuart	0.845	0.000	0.153
F-Fisher	0.976	0.006	0.016
Daniels	0.877	0.008	0.113
Average TAP 'nearest-neighbour method'	**No trend**	**Increasing trend**	**Decreasing trend**
Mann and Kendall	0.723	0.029	0.247
Cox-Stuart	0.822	0.006	0.171
F-Fisher	0.855	0.023	0.121
Daniels	0.718	0.030	0.251

Acknowledgements

Anonymous referees are gratefully acknowledged. Hydrologic data have been provided by the Regional Hydrologic Service of Tuscany (Italy).

References

Alexander LV, Zhang X, Peterson TC, Caesar J, Gleason B, Klein Tank AMG, Haylock M, Collins D, Trewin B, Rahimzadeh F, Tagipour A, Rupa Kumar K, Revadekar J, Griffiths G, Vincent L, Stephenson DB, Burn J, Aguilar E, Brunet M, Taylor M, New M, Zhai P, Rusticucci M, Vazquez-Aguirre JL. 2006. Global observed changes in daily climate extremes of temperature and precipitation. *Journal of Geophysical Research* **111**: D05109, DOI: 10.1029/2005JD006290.

Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B* **57**: 289–300.

Benjamini Y, Yekutieli D. 2001. The control of the false discovery rate in multiple testing under dependency. *The Annals of Statistics* **29**(4): 1165–1188.

Bonaccorso B, Cancelliere A, Rossi G. 2005. Detecting trends of extreme rainfall series in Sicily. *Advances in Geosciences* **2**: 7–11.

Brunetti M, Buffoni L, Maugeri M, Nanni T. 2000. Precipitation intensity trends in Northern Italy. *International Journal of Climatology* **20**: 1017–1031.

Brunetti M, Colacino M, Maugeri M, Nanni T. 2001. Trends in the daily intensity of precipitation in Italy from 1951 to 1996. *International Journal of Climatology* **21**: 299–316.

Burlando P, Rosso R. 2002. Effects of transient climate change on basin hydrology. 1. Precipitation scenarios for the Arno River, central Italy. *Hydrological processes* **16**: 1151–1175.

Commission of the European Communities. 2007. *Limiting Global Climate Change to 2 degrees Celsius. The Way Ahead for 2020 and Beyond*. Communication from the commission to the Council, the European parliament, the European economic and social committee and the committee of the regions. Brussels, 10 January.

Conover WJ. 1999. *Practical Nonparametric Statistics* (3rd ed.). John Wiley & Sons: New York.

Cox DR, Stuart A. 1955. Some quick tests for trend in location and dispersion. *Biometrika* **42**: 80–95.

Crisci A, Gozzini B, Meneguzzo F, Pagliara S, Maracchi G. 2002. Extreme rainfall in a changing climate: regional analysis and hydrological implications in Tuscany. *Hydrological Processes* **16**: 1261–1274.

Daniels HE. 1950. Rank correlation and population models. *Journal of the Royal Statistical Society, Series B* **12**: 171–181.

De Lùis M, Raventòs J, Gonzalez-Hidalgo JC, Sanchez JR, Cortina J. 2000. Spatial analysis of rainfall trends in the region of Valencia (East Spain). *International Journal of Climatology* **20**: 1451–1469.

De Michele C, Montanari A, Rosso R. 1998. The effects of non-stationarity on the evaluation of critical design storms. *Water Science Technology* **37**(11): 187–193.

Easterling DR, Meehl GA, Permesan C, Changnon SA, Karl TR, Mearns LO. 2000. Climate extremes: observations, modelling and impacts. *Science* **289**: 2068–2074.

European Environment Agency (EEA). 2007. Climate change and water adaptation issues. Technical report No. 2/2007, EEA: 13 February, 2007.

Fowler HJ, Kilsby CG. 2003. Implications of changes in seasonal and annual extreme rainfall. *Geophysical Research Letters* **30**(13): 1720. DOI: 10.1029/2003GL017327.

Fujibe F, Yamazaki N, Kobayashi K. 2006. Long-term changes of heavy precipitation and dry weather in Japan (1901–2004). *Journal of the Meteorological Society of Japan* **84**(6): 1033–1046.

González-Hidalgo JC, De Luìs M, Raventos J, Sanchez JR. 2003. Daily rainfall trend in the Valencia Region of Spain. *Theoretical and Applied Climatology* **75**: 177–130.

Groisman PY, Knight RW, Karl TR, Easterling DR, Sun B, Lawrimore JH. 2004. Contemporary changes of the hydrological cycle over the contiguous United States: trends derived from in situ observations. *Journal of Hydrometeorology* **5**(1): 64–85.

Hopkins LC, Holland GJ. 1997. Australian heavy-rain days associated east coast cyclones, 1958-92. *Journal of Climatology* **10**: 621–635.

Hosking JRM. 1990. L-moments: analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society B* **52**(1): 105–124.

IPCC, Intergovernmental Panel on Climate Change. 2007. Fourth Assessment Report of the IPCC. November. http://www.ipcc.ch (April 26th 2008).

Jenkinson AF. 1955. The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quarterly Journal Royal Meteorological Society* **81**: 158–171.

Katz RW, Brown BG. 1992. Extreme events in a changing climate: variability is more important than averages. *Climate Change* **21**: 289–302.

Katz RW, Brush G, Parlange M. 2005. Statistics of extremes: Modeling ecological disturbances. *Ecology* **86**: 1124–1134.

Katz RW, Parlange MB, Naveau P. 2002. Statistics of extremes in hydrology. *Advances in Water Resources* **25**(8–12): 1287–1304.

Kendall MG. 1975. *Rank Correlation Methods*. Griffin: London.

Livezey R, Chen W. 1983. Statistical field significance and its determination by Monte Carlo techniques. *Monthly Weather Review* **111**: 46–59.

Manabe S, Wetherald RT. 1987. Large scale changes of soil wetness induced by an increase in atmosphere carbon dioxide. *Journal of Atmospheric Science* **44**: 1211–1235.

Mann HB. 1945. Nonparametric tests against trend. *Econometrica* **13**: 245–259.

Mason SJ, Mimmack GM. 1995. Changes in flood frequencies in South Africa. *Proceedings of the 6th International meeting on Statistical Climatology*. Galway, Ireland. 299.

Mason SJ, Waylen PR, Mimmack GM, Rajaratnam B, Harrison JM. 1999. Changes in extreme rainfall events in South Africa. *Climate Change* **41**: 249–257.

Meehl GA, Karl T, Easterling D, Changnon S, Pielke R Jr., Changnon D, Evans J, Groisman P, Knutson TR, Kunkel KE, Mearns LO, Parmesan C, Pulwarty R, Root T, Sylves RT, Whetton P, Zwiers F. 2000. An introduction to trends in extreme weather and climate events: observations, socioeconomic impacts, terrestrial ecological impacts, and model projections. *Bulletin of the American Meteorological Society* **81**(3): 413–416.

Naidu CV, Rao BRS, Rao DVB. 1999. Climatic trends and periodicities of annual rainfall over India. *Meteorological Applications* **6**: 395–404.

Norrant C, Douguédroit A. 2006. Monthly and daily precipitations trends in the Mediterranean (1950–2000). *Theoretical and Applied Climatology* **83**: 89–106.

Oliver JE. 1980. Monthly precipitation distribution: a comparative index. *Professional Geographer* **32**: 300–309.

Piani F, Crisci A, De Chiara G, Maracchi G, Meneguzzo F. 2005. Recent trends and climatic perspectives of hailstorms frequency and intensity in Tuscany and Central Italy. *Natural Hazard and Earth System Sciences* **5**: 217–224.

Piervitali E, Colacino M, Conte M. 1997. Signals of climatic change in the Central-Western Mediterranean Basin. *Theoretical and Applied Climatology* **58**: 211–219.

Ramos MC, Martínez-Casanovas JA. 2006. Trends in precipitation concentration and extremes in the Mediterranean Penedès-Anoia region, NE Spain. *Climatic Change* 457–474.

Storch Von H, Zwiers FW. 1988. Recurrence analysis of climate sensitivity experiments. *Journal of Climate* **1**: 157–171.

Stuart A. 1956. The efficiencies of test of randomness against normal regression. *Journal of the American Statistical Association* **51**: 285–287.

Tebaldi C, Hayhoe K, Arblaster JM, Meehl GA. 2006. Going to the extremes, an intercomparison of model-simulated historical and future changes in extreme events. *Climatic Change* 185–211.

Toreti A, Desiato F. 2007. Temperature trend over Italy from 1961 to 2004. *Theoretical and Applied Climatology* **91**: 51–58, DOI: 10.1007/s00704-006-0289-6.

Ventura V, Paciorek CJ, Risbey JS. 2004. Controlling the proportion of falsely rejected hypothesis when conducting multiple tests with climatological data. *Journal of Climate* **17**: 4343–4356.

Wagner D. 1996. Scenarios of extreme temperature events. *Climatic Change* **33**: 385–407.

Yue S, Pilon P, Cavadias G. 2002. Power of the Mann-Kendall and Spearman's rho tests for detecting monotonic trends in hydrological series. *Journal of Hydrology* **259**: 254–271.