Analysing Trends in Precipitation Extremes for a Network of Stations

Agne Burauskaite-Harju
Linköping University, Linköping, Sweden

Anders Grimvall
Linköping University, Linköping, Sweden

Claudia von Brömssen
Swedish University of Agricultural Sciences, Uppsala, Sweden

Corresponding author address: Agne Burauskaite-Harju, Division of Statistics, Department of Computer and Information Sciences, Linköping University, SE-58183 Linköping, Sweden.
Phone: +46 13 282785
Fax: +46 13 142231
E-mail: agbur@ida.liu.se

ABSTRACT

Temporal trends in meteorological extremes are often examined by first reducing daily data to annual index values, such as 95th or 99th percentiles. Here, we report how this idea can be elaborated to provide an efficient test for trends at a network of stations. The initial step is to make separate estimates of tail probabilities of precipitation amounts for each combination of station and year by fitting a generalized Pareto distribution (GPD) to data above a user-defined threshold. The resulting time series of annual percentile estimates are subsequently fed into a multivariate Mann-Kendall (MK) test for monotonic trends. We performed extensive simulations using artificially generated precipitation data and noted that the power of tests for temporal trends was substantially enhanced when ordinary percentiles were substituted for GPD percentiles. Furthermore, we found that the trend detection was robust to misspecification of the extreme value distribution. An advantage of the MK test is that it can accommodate nonlinear trends, and it can also take into account the dependencies between stations in a network. To illustrate our approach, we used long time series of precipitation data from a network of stations in the Netherlands.

KEY WORDS climate extremes; precipitation; temporal trend; generalized Pareto distribution; climate indices; global warming

1. Introduction

It is generally believed that global warming causes an increase in the frequency and intensity of extreme weather events. Data representing such changes can be derived from process-based global circulation models (Kharin & Zwiers, 2000; Semenov & Bengtsson, 2002). Moreover, a number of investigators have also compiled high quality data sets and summarized indications of temporal variations in extreme precipitation and daily minimum and maximum temperatures (Klein Tank and Können, 2003; Moberg and Jones, 2005; Goswami *et al*., 2006, Griffiths *et al*., 2003, Groisman *et al*., 1999, Haylock and Nicholls, 2000, Haylock *et al*., 2005, Manton *et al*., 2001). The statistical methods that were used in the cited empirical studies are mainly descriptive in nature. In general, this means that the collected time series of daily weather data are first reduced to annual indices, such as 90th or 95th percentiles, and then the obtained sequences of index values are summarized in tables and graphs. Only a few climatologists have employed techniques in which extreme value distributions are fitted to daily weather records (IPCC, 2002; Della-Marta and Wanner, 2006).

The existing statistical literature on trends in extremes is focused considerably more on extreme value theory. Several researchers have advocated that the problem of detecting and testing for trends in environmental and meteorological extremes is best addressed by using probability models in which the parameters of some extreme value distribution are permitted to vary in a simple fashion with the year of investigation. Smith (1989, 1999, 2001, 2003) has proposed models in which the location and scale parameters of such

probability distributions vary linearly or periodically over the study period. Other statisticians have emphasized the need for less rigid models and nonparametric analysis of temporal trends when fitting parametric models to extreme values (Hall and Tajvidi, 2000; Davison & Ramesh, 2000; Pauli & Coles, 2001; Chavez-Demoulin & Davison, 2005; Yee & Stephenson, 2007), but none of those authors has suggested any formal significance test for the presence of trends in extremes at a network of stations.

The aim of the present work was to develop a test for trends in extremes that can accommodate nonlinear trends and integrate information from a network of stations. To make our test easy to comprehend, we stuck to the widespread concept of computing annual percentiles, i.e. the first step in our method is to perform separate analysis of subsets of data representing one year of daily weather records from a single station. However, in contrast to the current tradition in climatology, annual percentiles are computed by first fitting a generalized Pareto distribution (GPD) to data above a user-defined threshold and then utilizing mathematical relationships between percentiles and the shape and scale parameters of such distributions. The percentiles computed in this manner are subsequently fed into a standard Mann-Kendall test for trends in multiple time series of data (Hirsch & Slack, 1984).

It is well known that ordinary annual percentiles computed from daily precipitation records exhibit substantial interannual variation. Accordingly, we undertook extensive simulations to determine whether the variation in annual GPD-based percentiles is less irregular. Furthermore, we investigated how the performance of our trend test is influenced

by misspecification errors in the extreme value distribution. Because our test is not derived from any criteria of optimality, we also examined whether separate fitting of GPDs to subsets of data representing a single year leads to substantial loss of information. A collection of long time series of precipitation data from a network of stations in the Netherlands was used to illustrate our approach.

2. Observational data

We selected twelve high-quality time series of daily precipitation records from the European Climate Assessment (ECA) data set (Klein Tank *et al*., 2002). All stations were located in the Netherlands and had complete records for the period 1905–2004. The homogeneity of these and other ECA data series has previously been examined by Wijngaard *et al*. (2003).

Figure 1 illustrates time series of annual medians and 98th percentiles of our data set. As can be seen, there is a substantial difference in interannual variation between the medians and higher percentiles, and the graphs indicate that the trends may differ as well.

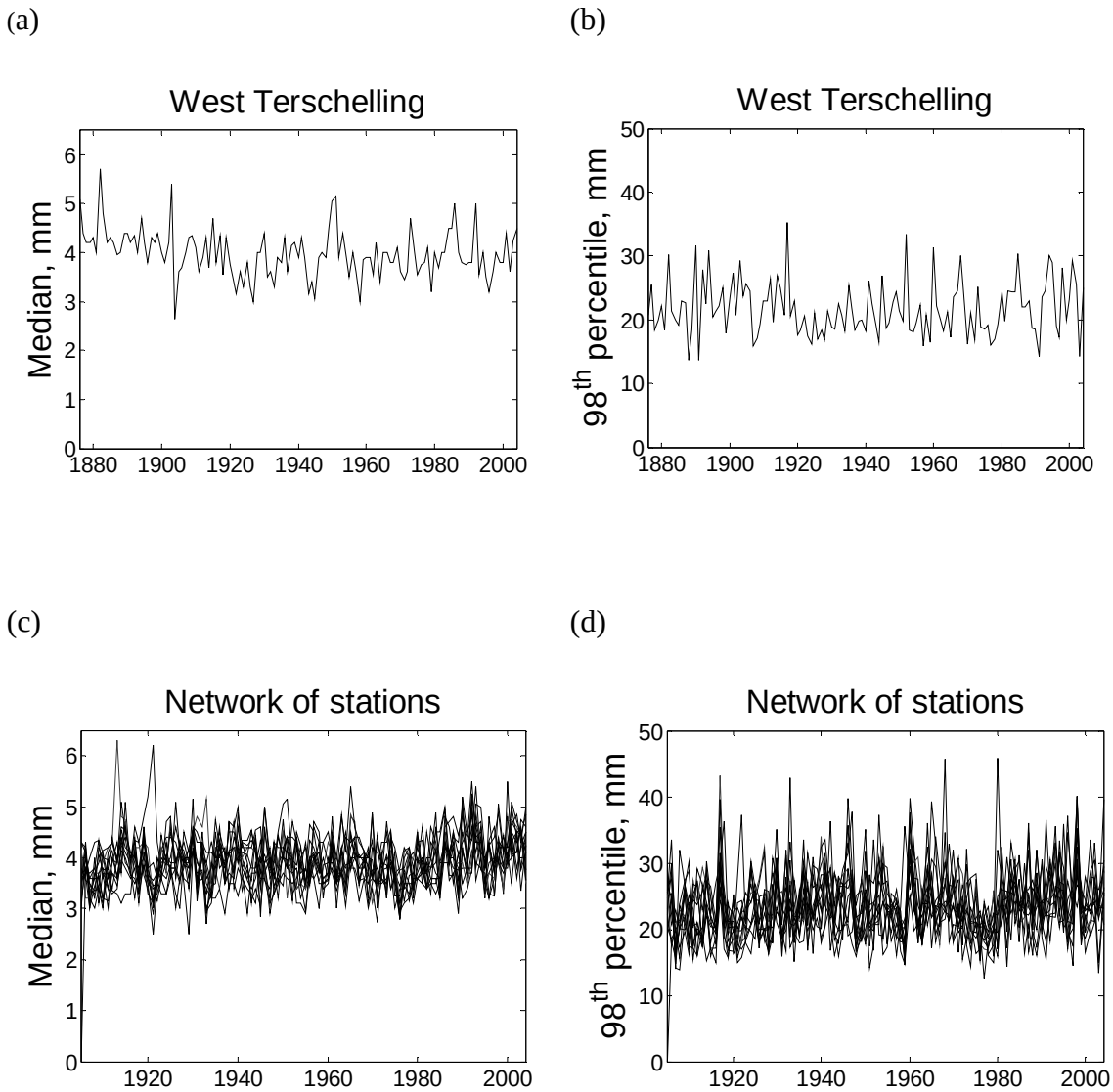

FIGURE 1. Annual medians (a and c) and 98th percentiles (b and d) of daily precipitation on rainy days (precipitation intensity ≥ 1 mm) in the Netherlands. Data from West Terschelling meteorological station (a and b) and a network of twelve meteorological stations (c and d). Each curve represents one station.

3. Methodology

The procedure we propose for detection of trends in extremes in multiple time series of meteorological data comprises the following steps:

(i) Partitioning the given data into subsets, normally one subset for each year and site.

(ii) Fitting GPDs to the exceedances over a user-defined high threshold in each subset of data.

(iii) Calculating percentiles from the estimated shape and scale parameters of the fitted GPDs.

(iv) Analysing temporal trends in the computed sequences of percentiles.

More detailed information about the extreme value distributions, percentiles, and trend tests we used is given below.

3.1. Extreme value distributions

Given a threshold u, the excess distribution F_u of a random variable X with distribution F is the distribution of $Y = X - u$ conditioned on $X > u$. Accordingly, we can write

$$F_u(y) = \Pr\{Y \le y\} = \Pr\{X \le u + y | X > u\} = \frac{F(u+y) - F(u)}{1 - F(u)}. \qquad (1)$$

For a sufficiently high u value, the GPD provides a good approximation of the excess distribution F_u (Balkema & de Haan, 1974; Pickands, 1975). The cumulative distribution function of a GPD is usually written

$$G(y;\sigma,\xi) = 1 - \left(1 + \xi \frac{y}{\sigma}\right)^{-1/\xi}. \tag{2}$$

where $\sigma > 0$ is a scale parameter and $\xi \neq 0$ a shape parameter.

The case $\xi = 0$ is a result of elementary calculus

$$\lim_{\xi \to 0} G(y;\sigma,\xi) = 1 - \exp\left(-\frac{y}{\sigma}\right). \tag{3}$$

In other words, the GPD is then identical to an exponential distribution with mean σ. Positive values of ξ produce probability distributions with heavy tails, whereas negative values produce distributions that are constrained to the interval $(0, -\sigma/\xi)$. Some techniques for threshold u selection have been summarized by Smith (2001).

3.2. Ordinary and GPD-based percentiles

Let $y_1, y_2, \ldots, y_N$ be an ordered sample from a probability distribution with cumulative distribution function (cdf) F, and define an empirical cdf by setting

$$\hat{F}(y_k) = k/(N+1),\ k = 1, ..., N \tag{4}$$

and connecting the points $(y_k, \hat{F}(y_k))$ k = 1, …, N with straight lines. Then we can estimate the 100pth percentile of F by setting

$$\begin{aligned} &y(p) = \hat{F}^{-1}(p), \quad if\ 1/(N+1) \le p \le N/(N+1) \\ &y(p) = y_1, \quad if\ \ p < 1/(N+1) \\ &y(p) = y_N, \quad if\ \ p > N/(N+1) \end{aligned} \tag{5}$$

where $\hat{F}^{-1}$ denotes the inverse of $\hat{F}$. In the following, this empiric percentile is referred to as the ordinary percentile of our sample.

Let us now assume that the exceedances of a threshold u can be described by a GPD with parameters ξ and σ. The inverse of such a cdf can then be written

$$G^{-1}(p, \sigma, \xi) = \frac{\sigma}{\xi}\left[(1-p)^{-\xi} - 1\right], \quad 0 < p < 1 \tag{6}$$

and GPD-based percentiles can be computed according to

$$y_{GPD}(p) = u + \hat{G}^{-1}(N(1-p)/N_u) = u + \frac{\hat{\sigma}}{\hat{\xi}}\left(\left(N(1-p)/N_u\right)^{-\hat{\xi}} - 1\right). \tag{7}$$

where N_u denotes the number of observations above the threshold u, and $\hat{\sigma}$ and $\hat{\xi}$ denote maximum-likelihood estimates of the model parameters.

3.3. Tests for temporal trends

The presence of monotonic trends in sequences of percentiles from a network of stations was assessed by a Mann-Kendall (MK) test.

In the univariate case, the MK test for an upward or downward trend in a data series $\{Z_k, k = 1,\ 2,\ \ldots,\ n\}$ is based on the test statistic

$$T = \sum_{j<i} \operatorname{sgn}\left(Z_i - Z_j\right) \tag{8}$$

Achieved significance levels can be derived from a normal approximation of the test statistic. Provided that the null hypothesis is true (i.e. that all permutations of the data are equally likely) and that n is large, T is approximately normal with mean 0 and variance $n(n-1)(2n+5)/18$.

In the multivariate case, the presence of an overall monotonic trend can be assessed by computing the sum of the MK statistics for all coordinates. Provided the null hypothesis is true and n is large, this sum will also be normally distributed with mean zero. However, its variance will be strongly influenced by the interdependence of the coordinates. Therefore, we used a technique proposed by Hirsch and Slack (1984) and Loftis and co-workers (1991) to estimate the covariance of two MK statistics and compute achieved significance levels.

4. Simulation studies

We conducted a set of simulation studies to examine the performance of ordinary and GPD-based percentiles and to determine how such percentiles can be incorporated into univariate tests for temporal trends in extremes.

4.1. Precision and accuracy of ordinary and GPD-based percentiles

The precision and accuracy of ordinary and GPD-based percentiles were assessed both for samples from known probability distributions and for outputs from a stochastic weather generator.

The first type of data comprised 3000 samples from each of the following: (i) a heavy-tailed distribution (GPD(40, 0.2)); (ii) a finite-tail distribution (GPD(40, –0.2)); (iii) a mixture of two distributions (GPD(40, 0.2) and GPD(40, –0.2)). The sample length was

normally distributed with a mean and standard deviation (μ=155, σ=13) corresponding to the number of rainy days per year at Heathrow Airport in the UK.

Synthetic weather data were created using the EARWIG weather generator (Kilsby *et al*., 2007), which is based on the Neyman-Scott rectangular pulses rainfall model. More specifically, we generated 3000 samples of precipitation data, each representing one year of daily rainfall records (precipitation $\geq$ 1 mm) in a 10 x 10 km grid around Heathrow Airport.

For each time series and year, ordinary percentiles were computed according to formula (5). Moreover, GPD percentiles were computed according to formula (7) after GPD distributions had been fitted to the 20% largest observations each year. Figures 2 and 3 present the obtained means and standard deviations for 90th to 99th percentiles.

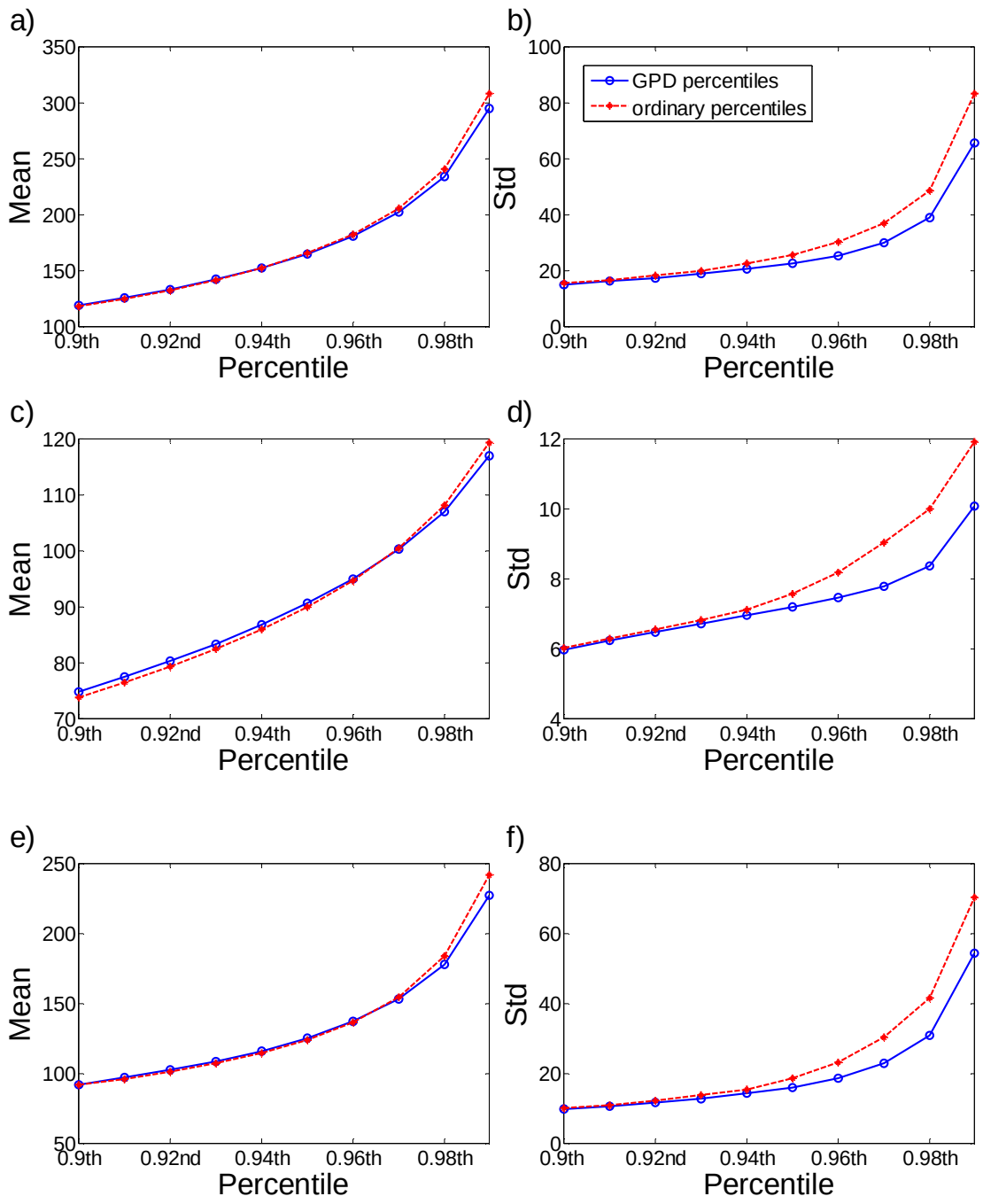

FIGURE 2. Estimated mean values (a, c, e) and standard deviations (b, d, f) of GPD-based (solid line) and ordinary (dotted line) percentiles. The percentiles were computed from simulated data with a heavy-tailed (a, b) or a light-tailed (c, d) distribution, or a mixture of the two distributions (e, f).

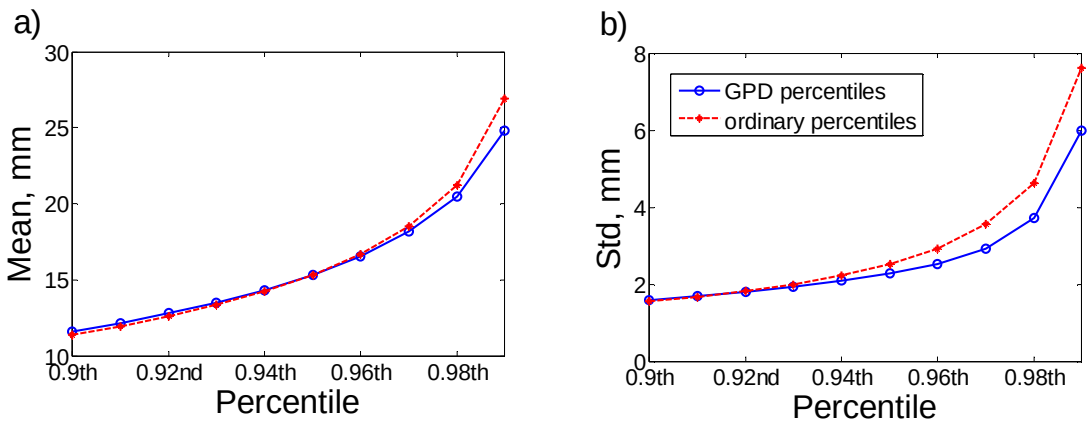

FIGURE 3. Estimated mean values (a) and standard deviations (b) of GPD-based (solid line) and ordinary (dotted line) percentiles. The percentiles were computed from rainfall data obtained by running the weather generator EARWIG for a 10x10 km area around Heathrow Airport in the UK.

As illustrated, the means of the ordinary and GPD-based percentiles were almost identical in all simulations. Inasmuch as the former percentiles were unbiased, we concluded that the GPD-based percentiles were also practically unbiased. The differences in standard deviations were more pronounced, at least for the higher percentiles. Regardless of the probability distribution of the simulated precipitation records, we found that the GPD-based estimators of such percentiles had higher precision than the ordinary percentiles. Closer examination of the probability distribution of the two types of percentiles also indicated that the GPD-based estimators were less skewed. Together, these results showed that GPD-based percentiles performed better than ordinary percentiles for a wide range of underlying distributions of daily precipitation records.

4.2. Performance of univariate trend estimators involving ordinary and GPD-based percentiles

Regressing annual percentiles on time represents a simple method for assessing temporal trends in the magnitude of extreme events. Here we examined the performance of such trend estimators based on ordinary and GPD percentiles, respectively.

The underlying data were generated using a simple rainfall model in which the number of rainy days each year had a normal distribution and the amount of precipitation on rainy days was exponentially distributed (Zhang *et al*., 2003). Furthermore, the parameters of the normal and exponential distributions were selected to mimic precipitation data from West Terschelling meteorological station (53.22ºN, 5.13ºE; 7 m above sea level) in the Netherlands. Accordingly, the number of rainy days per year was assumed to be normal with mean 133 and standard deviation 17, and the average precipitation on rainy days was set to 5.6 mm for the first year of the simulated time period.

Precipitation records representing 100-year-long time periods were generated by concatenating statistically independent one-year datasets. Moreover, trends were introduced by letting the expected amount of precipitation

$$\mu = \beta_1 + \beta_2 t \tag{9}$$

and, hence, also all percentiles values

$$y = \beta_1 \log\left(\frac{1}{1 - p/100}\right) + \beta_2 \log\left(\frac{1}{1 - p/100}\right) t \,. \tag{10}$$

change linearly with years.

We have already noted that GPD-based percentiles performed better than ordinary percentiles for a wide range of underlying probability distributions. In the present simulation experiments, we observed an analogous difference between trend estimators based on the two types of percentiles. The results given in Table I were obtained by regressing annual 98th percentiles on time, and they show that, regardless of the true slope of the trend line, the standard deviation was lower for the GPD-based estimators than for the estimators based on ordinary percentiles. In addition, all the investigated trend estimators were practically unbiased, and the mean square error was dominated by random errors. This provided additional support for GPD-based approaches.

TABLE I. Bias, root-mean-square error (RMSE), and standard deviation of temporal trends computed by regressing GPD and ordinary 98th percentiles on time.

Trend slope, %	Trend slope	Trend slopes derived from GPD percentiles			Trend slopes derived from ordinary percentiles		
		Bias	RMSE	Std. dev.	Bias	RMSE	Std. dev.
0	0.0000	0.00018	0.00969	0.00969	0.00000	0.01158	0.01159
5	0.0110	–0.00084	0.00989	0.00986	–0.00114	0.01192	0.01187
10	0.0219	–0.00019	0.01003	0.01004	–0.00024	0.01231	0.01231
15	0.0329	–0.00052	0.01019	0.01018	–0.00019	0.01266	0.01267
20	0.0438	–0.00092	0.01050	0.01046	–0.00049	0.01277	0.01276
25	0.0548	–0.00030	0.01104	0.01104	0.00033	0.01362	0.01363
30	0.0657	0.00004	0.01056	0.01057	0.00086	0.01355	0.01353
35	0.0767	–0.00100	0.01081	0.01077	–0.00018	0.01336	0.01337
40	0.0876	–0.00143	0.01134	0.01126	–0.00075	0.01375	0.01374
45	0.0986	–0.00148	0.01184	0.01175	–0.00040	0.01426	0.01426
50	0.1095	–0.00085	0.01222	0.01220	0.00058	0.01519	0.01519

4.3. Performance of univariate trend tests relying on models with time-dependent GPD parameters

The time dependence of GPD parameters and percentiles can be modelled in a parametric or nonparametric fashion. The simulation experiments described here aimed to investigate how this modelling can influence the precision and accuracy of univariate trend estimators. In particular, we compared the mean and standard deviation of trend estimators

derived from models in which precipitation percentiles were either (i) linearly increasing or (ii) block-wise constant.

To simplify the simulations, we omitted the dry days and assumed that each year consisted of 133 rainy days. Furthermore, we assumed that the daily rainfall amounts were exponentially distributed, implying that also the exceedances over any given threshold were exponentially distributed with the same mean. The mean amount of precipitation was set to increase linearly on a daily basis.

The entire simulation experiment comprised 1000 data series, each representing 100 years of daily precipitation records, and the block size was varied from 1 to 50 years. For each series and block size, we estimated the GPD parameters for all blocks and calculated percentiles (7), and then estimated the trend slope by regressing derived percentiles on time. In addition, we computed maximum-likelihood estimates for a GPD model in which the scale parameter varied linearly with time and the shape parameter was constant (Smith 1989, 1999, 2001, 2003). Estimated parameters were used to derive percentiles (7) and to determine trend slope in percentile series.

When analysing data containing a linear trend, it should be optimal to use trend estimators that are specifically designed to detect such trends. This was partly confirmed by our simulations (Table II). Estimators derived from models in which the scale parameter of a GPD varied linearly with time had higher precision than estimators derived from models in which the distribution of the precipitation amount was assumed to be stepwise constant.

However, it should be mentioned that percentile estimators based on formula (7) are biased, because they do not take into account the possibility of a trend in the probability of exceeding the user-defined threshold. The nonparametric trend estimators offered considerable accuracy and also relatively good precision, regardless of the block size, and thus they represent a viable option when it is neither feasible nor desirable to set up a parametric model of the trend function of the scale parameter.

TABLE II. Slope estimates of linear trends in 98th percentiles of rainfall data.

	True trend slope	ML estimator of linear trend	Trend estimators relying on stepwise constant precipitation distribution					
			Block size (years)					
			1	3	5	10	20	50
Mean	0.0218	0.0128	0.0215	0.0217	0.0218	0.0218	0.0219	0.0220
Std. dev.	0	0.0092	0.0103	0.0103	0.0102	0.0102	0.0104	0.0113

5. Assessment of observational data from a network of stations

We used our proposed two-step procedure to analyse the data set presented in Figure 1 for temporal trends in extremes. This was done by first deriving GPD percentiles for each year and station, and then feeding those percentiles into a multivariate MK test.

Figure 4 illustrates ordinary and GPD-based 98th percentiles for one of the investigated stations (West Terschelling) in the Netherlands. As seen in the simulation

experiments aimed at elucidating the precision and accuracy of percentile estimators, we found less variation in the GPD percentiles than in the ordinary percentiles. In particular, it can be noted that the GPD-based series had fewer outliers. Similar patterns were observed at several of the other investigated stations.

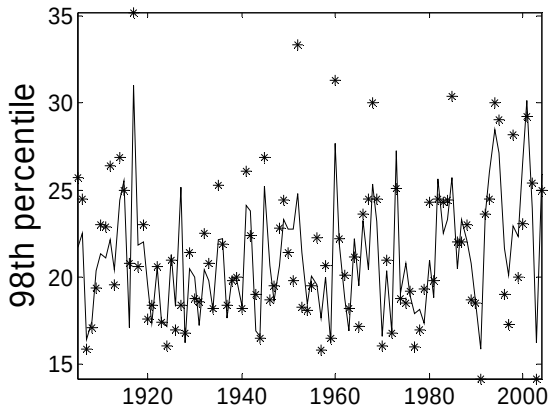

FIGURE 4. Annual estimates of 98th percentiles of the amount of daily precipitation at West Terschelling station (the Netherlands). The solid line represents GPD percentiles, and the stars indicate ordinary percentiles.

The difference between the ordinary and GPD-based percentiles was further demonstrated by the results of univariate MK tests for monotonic trends occurring at the examined network of stations. Some trends emerged more clearly in the GPD percentiles, because the interannual variation was generally smaller in those values. For example, there were five stations with a significant upward trend in the GPD-based 99th percentiles but no significant trends at all in the ordinary percentiles (Figure 5). The advantage of using GPD-

based estimators was less apparent in the lower percentiles. Figure 6 illustrates the spatial pattern in the detected trends.

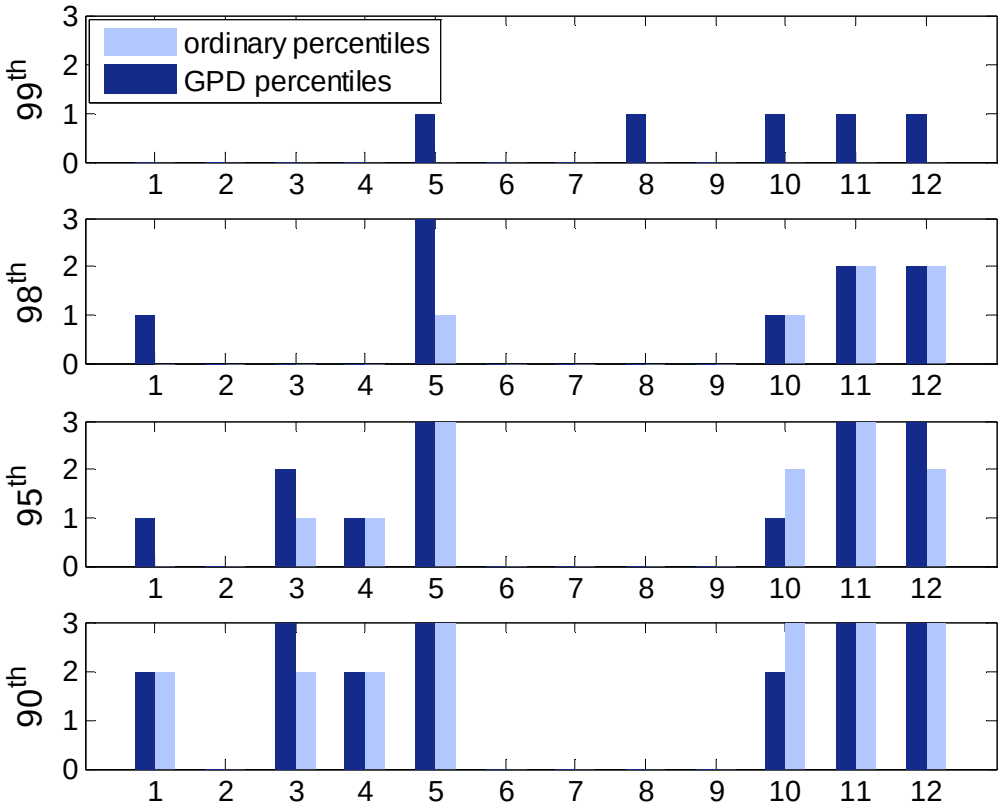

FIGURE 5. Achieved significance levels (p-values) of Mann-Kendall tests for monotonic increasing trends in time series of 90th, 95th, 98th, and 99th percentiles. The data came from 12 meteorological stations in the Netherlands.

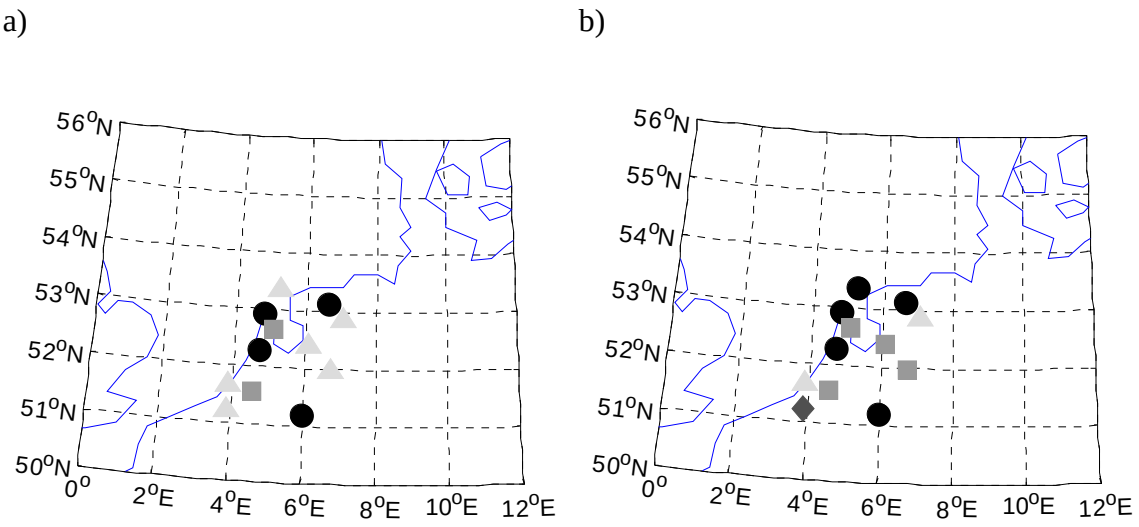

FIGURE 6. Achieved significance levels (p-values) of Mann-Kendall tests for monotonic increasing trends in time series of ordinary (a) and GPD-based (b) 98th percentiles. The data originated from 12 meteorological stations in the Netherlands. P-value: ● 0–0.05, ♦ 0.05–0.1, ■ 0.1–0.25, ▲ > 0.25.

The difference between ordinary and GPD-based percentiles emerged even more clearly in the multivariate MK tests for overall monotonic trends at the investigated stations. As shown in Table III, there was a dramatic difference in the achieved significance levels, especially for the higher annual percentiles of daily precipitation records.

TABLE III. P-values of multivariate Mann-Kendall tests for monotonic trends in time series of annual 90th to 99th percentiles of daily precipitation records from 12 meteorological stations in the Netherlands.

Percentile	*p*-value based on ordinary percentiles	*p*-value based on GPD percentiles
99th	0.07860	0.01488
98th	0.01774	0.00291
95th	0.00129	0.00041
90th	0.00009	0.00007

6. Conclusions and discussion

Two aspects of the present study demonstrated that there is a potential to substantially improve the methods that are currently preferred for analysing trends in meteorological extremes at a network of stations. First, we used a parametric extreme value model to achieve the greatest possible precision in our annual estimates of the probabilities of extreme meteorological events. Second, we fed such annual estimates into a multivariate trend test that was able to accommodate statistically dependent data from a network of stations. It is also worth noticing that the test we developed is easy to comprehend, because it adheres to the climatological tradition of computing annual index values (Klein Tank and Können, 2003; Moberg and Jones, 2005; Goswami *et al*., 2006, Griffiths *et al*., 2003, Groisman *et al*., 1999, Haylock and Nicholls, 2000, Haylock *et al*., 2005, Manton *et al*., 2001).

As emphasized in the introduction, the use of parametric extreme value models to estimate tail probabilities is strongly supported in the statistical literature (Gumbel, 1958; Leadbetter *et al*., 1983; Beirlant *et al*., 2004). It is an undisputable fact that estimators tailored to specific parametric models can have smaller variance than distribution-free estimators. In addition, this difference in variance is often pronounced when data are scarce, and extreme events are, by definition, infrequent. Our simulation experiments produced the anticipated results. The results: the sample variance of the GPD-based percentile estimates was considerably smaller than that of ordinary purely empirical percentiles (Table I).

A potential drawback of parametric modelling is that the probabilities of extreme events may be systematically over- or underestimated, if the extreme value model is not fully correct. However, our study showed that, from the standpoint of trend detection, bias is not a major problem. Figure 1 clearly shows that the large interannual variation in the intensity of extreme meteorological events is the main obstacle to successful assessment of temporal trends in such events. This conclusion was confirmed by the simulation experiments summarized in Table I. When the mean squared error (MSE) of the GPD-based percentile estimates was decomposed into variance and squared bias, it was apparent that the MSE value was dominated by random errors, whereas the (squared) bias played a minor role.

If temporal changes in the parameters of an extreme value distribution can be described by a simple mathematical function, and a single station is taken into consideration, it is feasible to construct statistical tests for trends in extremes (Smith, 1989; Davison & Smith, 1990; Coles & Tawn, 1990; Rootzen & Tajvidi, 1997). In the present study, we used a nonparametric trend test because we wanted to avoid making uncertain assumptions about the shape of the trend curve. Some investigators have previously emphasized that nonparametric techniques are particularly suitable for exploratory analyses of trends in extremes (Davison & Ramesh, 2000; Hall & Tajvidi, 2000; Pauli & Coles, 2001; Chavez-Demoulin & Davison, 2005, Yee & Stephenson, 2007). Furthermore, resampling has been utilized to assess the uncertainty of the estimated trend curves (Davison & Ramesh, 2000). Our method goes one step further by providing a formal trend test. In addition, it is designed to detect any monotonic change over time, which makes it well suited for studies of climate change, because the atmospheric concentration of greenhouse gases has long been increasing.

The multivariate MK test that we selected for the final trend assessment is especially convenient for evaluating trends at a network of stations, because it can accommodate statistically dependent time series of data. In particular, it can be noted that our procedure does not require any explicit modelling of the multivariate distribution of the meteorological observations at the investigated sites. The only thing that matters in such an MK test is the pattern of plus and minus signs that is observed when estimates of tail probabilities are compared for all possible pairs of years at each of the stations. Furthermore, our approach makes it possible to determine the statistical significance of an

overall trend in meteorological extremes by employing a fully automated procedure developed by Hirsch and Slack (1984).

A potential weakness of our procedure, as well as all other methods presently used to assess trends in climate extremes, is related to the presence of long-term cyclic patterns or autocorrelations in the analysed data. The multivariate MK test assumes that the vectors used as inputs are temporally independent. By computing tail probabilities separately for each year, it is assumed that there is no correlation between years. The robustness of our test to serial correlation can be enhanced by computing tail probabilities over periods of two or more years, or by reorganizing annual records into a new matrix with larger time steps (Wahlin & Grimvall, 2008). However, there is a trade-off between the length of the mentioned period and the power of the trend tests.

Another potential limitation of our technique is associated with the fact that it is not derived from any optimality criterion. In particular, there may be loss of power because there is no parametric modelling of the temporal dependence. However, our simulations of such effects showed that the loss of power is moderate (Table II), and therefore we conclude that this weakness does not outweigh the advantages of our procedure.

Finally, it can be noted that our case study provided further evidence of the strength of our procedure. The results presented in Table III demonstrate that our test was able to detect trends in precipitation extremes that would have been overlooked if we had based our inference on ordinary percentiles that did not rely on any extreme value model.

Acknowledgements

The authors are grateful for financial support from the Swedish Environmental Protection Agency.

References

[1] Balkema AA, de Haan L. 1974. Residual lifetime at great age. Annals of Probability 2: 792–804.

[2] Beirlant J, Goegebeur Y, Segers J, Teugels J. 2004. Statistics of Extremes: Theory and Applications. Wiley Series in Probability and Statistics. Wiley: Chichester.

[3] Coles SG, Tawn JA. 1990. Statistics of coastal Flood prevention, Philosophical Transactions of the Royal Society of London, Series A 332: 457- 476.

[4] Chavez-Demoulin V, Davison AC. 2005. Generalized additive modeling of sample extremes. Applied Statistics 54(1): 207-222.

[5] Davison AC, Smith RL. 1990. Models for exceedances over high thresholds. Journal of the Royal Statistical Society, Series B 52: 393–442.

[6] Davison AC, Ramesh NI. 2000. Local likelihood smoothing of sample extremes. Journal of Royal Statistical Society, Series B 62(1): 191-208.

[7] Dietz EJ, Killeen TJ. 1981. A nonparametric multivariate test for monotone trend with pharmaceutical applications. Journal of the American Statistical Association 76: 169–174.

[8] Goswami BN, et al. 2006. Increasing trend of extreme rain events over India in a warming environment. Science 314: 1442. DOI: 10.1126/science.1132027

[9] Griffiths GM, Salinger MJ, Leleu I. 2003. Trends in extreme daily rainfall across the South Pacific and relationship to the South Pacific convergence zone. International Journal of Climatology 23: 847–869. DOI: 10.1002/joc.923

[10] Groisman P, Karl T, Easterling D, Knight R, Jamason P, Hennessy K, Suppiah R, Page C, Wibig J, Fortuniak K, Razuvaev V, Douglas A, Førland E, Zhai P. 1999. Changes in the probability of extreme precipitation: important indicators of climate change. Climatic Change 42: 243–283.

[11] Gumbel EJ, 1958. Statistics of Extremes. Columbia University Press: New York.

[12] Hall P, Tajvidi N. 2000. Nonparametric analysis of temporal trend when fitting parametric models to extreme-value data. Statistical Science 15(2): 153-167.

[13] Haylock M, Nicholls N. 2000. Trends in extreme rainfall indices for an updated high quality data set for Australia, 1910–1998. International Journal of Climatology 20: 1533–1541.

[14] Haylock MR, Peterson TC, Alves LM, Ambrizzi T, Anunciação YMT, Baez J, Barros VR, Berlato MA, Bidegain M, Coronel G, Corradi V, Garcia VJ, Grimm AM, Karoly D, Marengo JA, Marino MB, Moncunill DF, Nechet D, Quintana J, Rebello E, Rusticucci M, Santos JL, Trebejo I, Vincent LA. 2005. Trends in total and extreme South American rainfall in 1960–2000 and links with sea surface temperature. Journal of Climate 19: 1490-1512.

[15] Hirsch RM, Slack JR. 1984. A nonparametric trend test for seasonal data with serial dependence. Water Resources Research 20: 727-732.

[16] IPCC. 2002. IPCC Workshop on Changes in Extreme Weather and Climate Events. Beijung, China 11-13 June. Workshop Report.

[17] Kilsby CG, Jones PD, Burton A, Ford AC, Fowler HJ, Harpham C, James P,Smith A, Wilby RL. 2007. A daily weather generator for use in climate change studies. Environmental Modelling & Software 22: 1705-1719.

[18] Kharin VV, Zwiers FW. 2000. Changes in the extremes in an ensemble of transient climate simulations with a coupled atmosphere–ocean GCM. Journal of Climate 13: 3760–3788.

[19] Klein Tank AMG, Wijngaard JB, Können GP, Böhm R, Demarée G, Gocheva A, Mileta M, Paschiardis S, Hejkrlik L, Kern-Hansen C, Heino R, Bessemoulin P, Müller-Westermeier G, Tzanakou M, Szalai S, Pálsdóttir T, Fitzgerald D, Rubin S, Capaldo M, Maugeri M, Leitass A, Bukantis A, Aberfeld R,van Engelen AFV, Førland E, Mietus M, Coelho F, Mares C, Razuvaev V, Nieplova E, Cegnar T, López AJ, Dahlström B, Moberg A, Kirchhofer W, Ceylan A, Pachaliuk O, Alexander LV, Petrovic P. 2002. Daily dataset of 20th-century surface air temperature and precipitation observations for European Climate Assessment. International Journal of Climatology 22: 1441–1453.

[20] Klein Tank AMG, Können GP. 2003. Trends in indices of daily temperature and precipitaion extremes in Europe, 1946-99. Journal of Climate 16: 3665-3680.

[21] Leadbetter R, Lindgren G, Rootzen H. 1983. Extremes and Related Properties of Random Sequences and Processes. Springer-Verlag: New York.

[22] Manton MJ, Della-Marta PM, Haylock MR, Hennessy KJ, Nicholls N, Chambers LE, Collins DA, Daw G, Finet A, Gunawan D, Inape K, Isobe H, Kestin TS, Lefale P,

Leyu CH, Lwin T, Maitrepierre L, Ouprasitwong N, Page CM, Pahalad J, Plummer N, Salinger MJ, Suppiah R, Tran VL, Trewin B, Tibig I, Yee D. 2001. Trends in extreme daily rainfall and temperature in Southeast Asia and the south Pacific: 1961–1998. International Journal of Climatology 21: 269–284.

[23] Moberg A, Jones PD. 2005. Trends in indices for extremes in daily temperature and precipitation in central and western Europe, 1901-99. International Journal of Climatology 25: 1149-1171.

[24] Pickands J. 1975. Statistical Inference Using Extreme Order Statistics. Annals of Statistics 3: 119–131.

[25] Pauli F, Coles S. 2001. Penalized likelihood inference in extreme value analysis. Journal of Applied Statistics 28:547-560.

[26] Rootzen H, Tajvidi N. 1997. Extreme value statistics and wind storm losses: a case study. Scandinavia Actuarial Journal: 70-94.

[27] Semenov VA, Bengtsson L. 2002. Secular trends in daily precipitation characteristics: greenhouse gas simulation with a coupled AOGCM. Climate Dynamics 19: 123–140.

[28] Smith RL. 1989. Extreme value analysis of environmental time series: An application to trend detection in ground-level ozone (with discussion). Statistical Science 4: 367-393.

[29] Smith RL. 1999. Trends in rainfall extremes. Preprint, University of North Carolina.

[30] Smith RL. 2001. Environmental statistics. CBMS course in environmental statistics, University of Washington, http://www.stat.unc.edu/postscript/rs/envstat/env.html.

[31] Smith RL. 2003. Statistics of extremes, with applications in environment, insurance and finance. Extreme Values in Finance, Telecommunications and the Environment: 1-78. Chapman and Hall/CRC Press: London.

[32] Wahlin K, Grimvall A. 2008. Roadmap for assessing regional trends in groundwater quality. Submitted to Environmental Monitoring and Assessment.

[33] Wijngaard JB, Klein Tank AMG, Können GP. 2003. Homogeneity of 20th century European daily temperature and precipitation series. International Journal of Climatology 23: 679-692.

[34] Yee TW, Stephenson AG. 2007. Vector generalized linear and additive extreme value models. Extremes 10: 1-19. DOI 10.1007/s10687-007-0032-4

[35] Zhang X, Zwiers FW, Li G. 2004. Monte Carlo experiments on the detection of trends in extreme values. Journal of Climate 17: 1945-1952.