On the estimation of continuous 24-h precipitation maxima

H. Van de Vyver

Stochastic Environmental Research and Risk Assessment

ISSN 1436-3240
Volume 29
Number 3

Stoch Environ Res Risk Assess (2015)
29:653-663
DOI 10.1007/s00477-014-0912-5

Springer

Stoch Environ Res Risk Assess (2015) 29:653–663
DOI 10.1007/s00477-014-0912-5

ORIGINAL PAPER

On the estimation of continuous 24-h precipitation maxima

H. Van de Vyver

Published online: 24 June 2014

Abstract Extreme value analysis of precipitation is of great importance for several types of engineering studies and policy decisions. For return level estimation of extreme 24-h precipitation, practitioners often use daily measurements (usually 08:00–08:00 local time) since high-frequency measurements are scarce. Annual maxima of daily series are smaller or equal to continuous 24-h precipitation maxima such that the resulting return levels may be systematically underestimated. In this paper we use a rule, derived earlier, on the conversion of the generalized extreme value (GEV) distribution of daily to 24-h maxima. We develop an estimator for the conversion exponent by combining daily maxima and high-frequency sampled 24-h maxima in one joint log-likelihood. Once the conversion exponent has been estimated, GEV-parameters of 24-h maxima can be obtained at sites where only daily data is available. The new methodology has been extended to spatial regression models.

Keywords Generalized extreme value distribution · Extremal index · Rainfall · Sampling frequency · Spatial process

1 Introduction

Extreme hydro-meteorological events usually have a large impact on our society. For safety standards concerning life, property and for design purposes of large structures, estimation of extreme return levels are often required. A practical difficulty when analysing observational data is the effect of sampling frequency. For the calculation of return levels of 24-h precipitation extremes, hydrologist are confronted with the fact that daily data (usually 08:00–08:00 local time) only are available. To be precise, if $I(t)$ is the rainfall intensity (mm/h) at time t (expressed in h), then

$$X(t) = \int_{t-24}^{t} I(t')\,dt',$$

is the accumulated 24-h precipitation at time t. For a time window $[0, T]$, usually one year, the maximum of the continuous process $\{X(t)\}$ is

$$M(T) = \sup\{X(t)\,|\,0 \le t \le T\},$$

which in the hydrological community is often called a sliding maximum (Dwyer and Reed 1995; van Montfort 1990). Continuous measurements, if feasible, are costly to perform and difficult to record. If hourly precipitation depths $\{X(i)\} = \{X_i\}$ were recorded, a good approximation to $M(T)$ is $M_n^{(1)} = \max\{X_1, \ldots, X_n\}$, with $n = [T]$ ($[.]$ denotes the integer part). However, long-term high-frequency measurements are rather scarce in many parts of the world, especially in developing countries. On the other hand, daily observations are often longer, more reliable, and the network is geographically denser. If only daily data is available then $M_n^{(24)} = \max\{X_{24}, X_{48}, \ldots, X_{24[n/24]}\}$ will be observed, with $M_n^{(24)} \le M_n^{(1)} \approx M(T)$, such that the continuous 24-h maxima $M(T)$ may be systematically underestimated. The number of papers dealing with the discretization problem in hydrological risk estimation is limited. The earliest and most commonly cited references are Hershfield and Wilson (1958) and Weiss (1964). Hershfield and Wilson (1958) proposed an empirical

H. Van de Vyver (✉)
Royal Meteorological Institute of Belgium, Ringlaan 3, 1180 Brussels, Belgium
e-mail: hvijver@meteo.be

solution to this problem, using a multiplier (the Hershfield factor) to relate quantiles of continuous 24-h maxima and daily maxima. An historical overview and discussion is included in Dwyer and Reed (1995). More recently, Robinson and Tawn (2000) proposed a general mathematical framework that describes the effect of the sampling frequency on extreme value distributions. They point out a general relationship between the generalized extreme value (GEV) distributions of sub-sampled time series. The fundamental difference with Hershfield's empirical scaling rule is that the conversion of daily to 24-h maxima is done via the GEV-parameters. Beside the rainfall problem, this theory can be applied to other environmental extremes as well. In this paper we propose an estimation technique for the conversion exponent by combining daily maxima and high-frequency sampled 24-h maxima in one joint log-likelihood. Once the conversion exponent has been estimated, the GEV-parameters of 24-h maxima can be obtained at sites where only daily data is available.

The paper is organized as follows. In Sect. 2 we give basic facts of extremal theory for dependent and sub-sampled sequences. In Sect. 3 we propose the general assumption that there is a fixed, sufficiently small sampling interval for which the maxima are a good approximation to continuous 24-h precipitation maxima. Estimation methods for GEV-parameters of continuous 24-h precipitation maxima are introduced. A spatial extension of the new estimator is introduced in Sect. 4. Numerical experiments in Sect. 5 show the relevance of the new methodology. Finally, in Sect. 6, some conclusions are drawn.

2 Extremes of sequences: background results

2.1 Classical theory of extremes

The use of extreme value models is increasingly common in climate studies (Obeysekera et al. 2011; Siliverstovs et al. 2010; Yoon et al. 2013). These models are concerned with the statistical behavior of block maxima, i.e.

$$M_n = \max\{X_1,\ldots,X_n\}, \tag{1}$$

where $X_1,\ldots,X_n$ is a sequence of independent and identically distributed (iid) random variables. A key result in classical extreme value theory (EVT) is that if there exist sequences of constants $a_n > 0$ and b_n such that the cumulative distribution of the normalized maxima converges to a non-degenerate limit distribution $G(x)$, i.e.

$$P\left(\frac{M_n - b_n}{a_n} \le x\right) \to G(x), \quad \text{as} \quad n \to \infty,$$

then G is the Fréchet, Weibull, or Gumbel distribution (Beirlant 2004; Coles 2001; Embrechts et al. 1997; Leadbetter et al. 1983). These three classical extremal types can be rewritten in one unifying three-parameter distribution, the generalized extreme value (GEV) distribution, which takes the form

$$H(x;\mu,\sigma,\gamma) = \exp\left(-\left(1+\gamma\,\frac{x-\mu}{\sigma}\right)_+^{-1/\gamma}\right), \tag{2}$$

where $\mu \in \mathbb{R}$, $\sigma > 0$ and $\gamma \in \mathbb{R}$ are the location-, scale- and shape parameter, respectively. For notational convenience, denote by Z the maximum of the sample $X_1,\ldots,X_n$. Thus, for sufficiently large n, the probability $P(Z \le z)$ can be approximated by the GEV distribution. Inference is usually performed by maximum likelihood estimation (MLE), though moment-based estimation methods are also popular. For a given sample $\mathbf{Z} = (Z_1,\ldots,Z_k)$ of iid GEV random variables, the log-likelihood function is

$$l(\mathbf{Z}|\psi) = \sum_{i=1}^{k} \log h(Z_i;\psi),$$

with $\psi = (\mu,\sigma,\gamma)$ the vector of GEV-parameters, and $h(z;\psi)$ the probability density function. The MLE $\hat{\psi}$ is obtained by maximizing the log-likelihood function with respect to ψ.

The return level z_p associated with the return period $T = 1/p$ is obtained by inverting $H(z_p) = 1 - p$:

$$z_p = \begin{cases} \mu - \dfrac{\sigma}{\gamma}\left(1 - (-\log(1-p))^{-\gamma}\right), & \text{if } \gamma \neq 0, \\ \mu + \sigma \log(1-p), & \text{if } \gamma = 0. \end{cases} \tag{3}$$

2.2 Extremes of dependent sequences

Classical EVT can be extended to a wide class of dependent stationary sequences (Leadbetter et al. 1983). Obviously, some form of dependence restrictions is necessary to obtain extremal type results in dependent cases. The main condition to be used (termed $D(u_n)$) is defined with reference to a sequence $\{u_n\}$ of constants in terms of $F_{i_1\ldots i_r}(x_1,\ldots,x_n) = P(X_{i_1} \le x_1,\ldots,X_{i_r} \le x_r)$. For brevity, we write $F_{i_1\ldots i_r}(u_n)$ for $F_{i_1\ldots i_r}(u_n,\ldots,u_n)$.

The condition $D(u_n)$ *will said to be hold if for any integers*

$$1 \le i_1 < \cdots < i_p < j_1 < \cdots < j_q \le n,$$

for which $j_1 - i_p \ge l_n$, *we have*

$$|F_{i_1\ldots i_p j_1\ldots j_q}(u_n) - F_{i_1\ldots i_p}(u_n)\,F_{j_1\ldots j_q}(u_n)| \le \alpha_{n,l_n},$$

where $\alpha_{n,l_n} \to 0$ *as* $n \to \infty$ *for some sequence* $l_n = o(n)$.

The Extremal Types Theorem for stationary sequences (Leadbetter et al. 1983) follows from the above condition:

Theorem 1 *Let $\{X_i\}$ be a stationary sequence and $a_n > 0$ and b_n given constants such that*

$$P\left(\frac{M_n - b_n}{a_n} \le x\right) \to G(x), \quad \text{as } n \to \infty,$$

for some non-degenerate distribution $G(x)$. Suppose that $D(u_n)$ is satisfied for $u_n = a_n x + b_n$ for each real x. Then G is one of the three classical extremal types, and is thus the GEV distribution.

2.3 Extremes of sub-sampled sequences

We denote by $X_1, \ldots, X_n$ the discrete time realizations of a stationary stochastic process at the finest sampling rate of interest. The maximum of the sequence sampled at d^{-1} times is

$$M_n^{(d)} = \max\{X_d, X_{2d}, \ldots, X_{d[n/d]}\}.$$

Robinson and Tawn (2000) have pointed out the relationship between the GEV-distributions of $M_n^{(d)}$ and $M_n^{(1)}$. Their main findings are presented here, but our proofs have been expressed in terms of the Extremal Types Theorem. We will use the following lemma, which is due to Chernick (Chernick 1981; Leadbetter 1983; Leadbetter et al. 1983).

Lemma 1 *Let $\{X_i\}$ be a stationary sequence (marginal d.f. F) and $\{u_n\}$ a sequence of constants such that $D(u_n)$* holds. *Let $0 < \tau < \infty$. Then*

$$P(M_n \le u_n) \to e^{-\theta\tau}, \tag{4}$$

for some $0 \le \theta \le 1$, if and only if

$$n\left(1 - F(u_n)\right) \to \tau, \quad \text{as } n \to \infty.$$

In Eq. (4), θ is called the extremal index of the sequence $\{X_i\}$.

The following result is a reformulation of a major result of Robinson and Tawn (2000).

Theorem 2 *Assume that $D(u_n)$ holds for the stationary sequence $\{X_i\}$ (and consequently also for the subsequence $\{X_{id}\}$). Assume that the conditions of Theorem 1 are satisfied for $\{X_{id}\}$, i.e. there exist sequences $a_n^{(d)} > 0$ and $b_n^{(d)}$ such that*

$$P\left(\frac{M_n^{(d)} - b_n^{(d)}}{a_n^{(d)}} \le x\right) \to H(x), \quad \text{as } n \to \infty, \tag{5}$$

with H the GEV distribution. Then

$$P\left(\frac{M_n^{(1)} - b_n^{(d)}}{a_n^{(d)}} \le x\right) \to H(x)^{d\,\theta_1/\theta_d}, \quad \text{as } n \to \infty, \tag{6}$$

where θ_1 and *θ_d are the extremal indices of the sequences $\{X_i\}$* and *$\{X_{id}\}$, respectively.*

Proof For the sequence $\{X_{id}\}$, Eq. (5) may be rewritten as $P(M_n^{(d)} \le u_n) \to e^{-\theta_d\tau}$, with $u_n = a_n^{(d)} x + b_n^{(d)}$ and $\tau = -(\log H(x))/\theta_d$. By Lemma 1, we have $[n/d]\left(1 - F(u_n)\right) \to \tau$ as $n \to \infty$, or equivalently, $n\left(1 - F(u_n)\right) \to \tau d$ as $n \to \infty$. Now, we turn to the case of the sequence $\{X_i\}$. By Lemma 1, we get

$$P(M_n^{(1)} \le u_n) \to e^{-\theta_1\tau d}, \quad \text{as } n \to \infty. \tag{7}$$

Finally, Eq. (6) follows after a little arrangement of Eq. (7). □

It is easy to verify that the GEV distributions $H(x)$ and $H(x)^{d\,\theta_1/\theta_d}$ of Eqs. (5, 6) have the same shape parameter γ, so that it does not depend on the sampling frequency. We denote by $H_d(z) := H(z; \mu_d, \sigma_d, \gamma)$ and $H_1(z) := H(z; \mu_1, \sigma_1, \gamma)$ the approximative GEV-distributions of $Z^{(d)} := M_n^{(d)}$ and $Z^{(1)} := M_n^{(1)}$. It follows from Theorem 2 that for sufficiently large n, an approximate relationship between $H_d(z)$ and $H_1(z)$ is given by

$$H_1(z) \approx H_d(z)^{d\,\theta_1/\theta_d},$$

which is the original formulation of Robinson and Tawn (2000).

3 Continuous 24-h precipitation maxima

3.1 Continuous- versus discrete-time maxima

Let $\{X(t)\}$ be a stationary process. For the continuous maximum

$$M(T) = \sup\{X(t) \mid 0 \le t \le T\},$$

we assume that there exist constants $a_T > 0$ and b_T such that

$$P\left(\frac{M(T) - b_T}{a_T} \le x\right) \to H(x), \quad \text{as } T \to \infty, \tag{8}$$

where G is one of the three extremal types. General conditions for the continuous version of the Extremal Types Theorem are rather technical, and are not summarized here. For a complete overview, see Leadbetter et al. (1983).

We further assume that the conditions of Theorem 1 are satisfied for the discrete sequence $\{X_i\}$. In order to relate the asymptotic distribution of normalized continuous time maxima $M(T)$ to discrete maxima M_n we choose a positive interval of small time-units $\delta \ll 1$, for which the maximum sampled at the associated frequency:

$$M^{(\delta)}(T) = \max\{X(i\,\delta) \mid 0 \le i \le T/\delta\}, \tag{9}$$

is a sufficiently good approximation to $M(T)$. For any two sequences $\{X(i\,\delta)\}$ and $\{X(i\,d)\}$ sampled at times δ and d respectively, the extension of Eq. (6) is

$$P\left(\frac{M^{(\delta)}(T)-b_n}{a_n}\leq x\right)\rightarrow H(x)^{\Theta_\delta},\quad \text{as } T\rightarrow\infty,\\ \text{with } \Theta_\delta=\frac{d}{\delta}\frac{\theta_\delta}{\theta_d}. \tag{10}$$

It is, however, not clear when the limit in Eq. (10) would exist if $\delta\rightarrow 0$, and one needs to understand the relation between the maxima of continuous- and discrete-time processes. A commonly used idea is selecting a grid spacing δ (which may depend on T) that converges to zero at a specific rate. The coarsest grid over which continuous and discrete maxima have the same asymptotic distribution is called the Pickands' grid (Leadbetter et al. 1983). The most complete characterization of the relation between both types of extremes for Gaussian processes is given by Piterbarg (2004). Very recently, similar results are obtained for non-Gaussian processes as well (Turkman 2011). However, the above-mentioned papers fall outside the scope of this work, because in the usual practise one selects a sufficiently large fixed T (mostly one year) for which one assumes that convergence of Eqs. (8 and 10) is satisfactory. Then we get

$$H_\delta(z)\approx H_d^{\Theta_\delta}(z), \tag{11}$$

which is equivalent to

$$\mu_\delta\approx\mu_d-\frac{\sigma_d}{\gamma}\,(1-\Theta_\delta^\gamma)\qquad\text{and}\qquad\sigma_\delta\approx\sigma_d\,\Theta_\delta^\gamma. \tag{12}$$

For any fixed T, there exists a small $\delta>0$ such that the GEV-distributions of $Z^{(\delta)}$ and $Z:=M(T)$ are sufficiently close, i.e. $H(z)\approx H_\delta(z)$ or

$$H(z)\approx H_d^{\Theta_\delta}(z). \tag{13}$$

Such a methodology was also suggested by Robinson and Tawn (2000) for approximating continuous time maxima.

3.2 Rainfall extremes: estimation method for Θ_δ

In what follows, we reconsider the terminology of Sect. 1, such that $d=24$ for daily data. If only daily precipitation measurements are available, the conversion rule Eq. (13) is particularly useful for obtaining GEV-parameters of continuous 24-h precipitation maxima. It is thus of great importance to have a good estimate of Θ_δ. To illustrate the estimation, we take the 115-year time series of 10-min rainfall of the Royal Meteorological Institute of Belgium (Uccle, Brussels). The series was recorded by the same instrument at the same location since 1898, and processed with identical quality since that time (Demarée 2003). In Belgium, 08:00 h local time is the typical time of day at which daily aggregated data are recorded, so we initially construct daily data in this way by aggregating 10-min measurements. In Van de Vyver (2012) we have estimated the extremal indices θ_{24} and θ_δ with the so-called runs estimator (Smith and Weissman 1994), and substituted them in Eq. (10), resulting in $\hat{\Theta}_\delta=1.67$. However, the estimation error could not be provided. In the absence of a general theory of extremal index estimation in sub-sampled sequences, we employ here an alternative estimator. A possible method could be to estimate the GEV-parameter vectors $\psi_\delta=(\mu_\delta,\sigma_\delta,\gamma)$ and $\psi_{24}=(\mu_{24},\sigma_{24},\gamma)$ on the two sets of maxima

$$\mathbf{Z}^{(\delta)}=(Z_1^{(\delta)},\ldots,Z_k^{(\delta)}),\qquad \mathbf{Z}^{(24)}=(Z_1^{(24)},\ldots,Z_k^{(24)}),$$

respectively, and then manipulating these to estimate Θ_δ via Eq. (12). In the context of estimating the extremal index, a similar procedure was already proposed in Gomes (1993). Ancona-Navarette and Tawn (2000) advanced further on this idea and suggested that it would be better to perform a simultaneously estimation. For the estimation of Θ_δ we consider this procedure, but suitably adapted to our case. We estimate the joint distribution of $(\mathbf{Z}^{(\delta)},\mathbf{Z}^{(24)})$ based on the joint log-likelihood $l(\mathbf{Z}^{(\delta)}|\psi_\delta)+l(\mathbf{Z}^{(24)}|\psi_{24})$ $=: l(\mathbf{Z}^{(\delta)},\mathbf{Z}^{(24)}|\psi_\delta,\psi_{24})$. When introducing Eq. (12) we see that parameter vector $\phi=(\mu_{24},\sigma_{24},\gamma,\Theta_\delta)$ is a sufficiently statistic for $(\mathbf{Z}^{(\delta)},\mathbf{Z}^{(24)})$. Throughout this work, parameter vector ψ is referred to the standard GEV log-likelihood, while vector ϕ is associated with the present joint log-likelihood $l(\mathbf{Z}^{(\delta)},\mathbf{Z}^{(24)}|\phi)$. After some calculation, the joint log-likelihood can be written as

$$\begin{aligned} l(\mathbf{Z}^{(\delta)},\mathbf{Z}^{(24)}|\phi) &= -k\,(\log\sigma_{24}+\gamma\,\log\Theta_\delta)\\ &-\sum_{i=1}^{k}\left(\frac{\sigma_{24}+\gamma\,(Z_i^{(\delta)}-\mu_{24})}{\sigma_{24}\,\Theta_\delta^\gamma}\right)^{-1/\gamma}\\ &-(1/\gamma+1)\sum_{i=1}^{k}\log\left(\frac{\sigma_{24}+\gamma\,(Z_i^{(\delta)}-\mu_{24})}{\sigma_{24}\,\Theta_\delta^\gamma}\right)\\ &-k\,\log\sigma_{24}-\sum_{i=1}^{k}\left(1+\gamma\left(\frac{Z_i^{(24)}-\mu_{24}}{\sigma_{24}}\right)\right)^{-1/\gamma}\\ &-(1/\gamma+1)\sum_{i=1}^{k}\log\left(1+\gamma\left(\frac{Z_i^{(24)}-\mu_{24}}{\sigma_{24}}\right)\right). \end{aligned} \tag{14}$$

An estimation of Θ_δ can then be obtained by maximization of Eq. (14) with respect to ϕ.

Special attention has to be paid to the dependence between $\mathbf{Z}^{(\delta)}$ and $\mathbf{Z}^{(24)}$. As well known, MLE is based on the assumption of independent observations. The asymptotic properties of the independence MLE are well known, but this is not the true model, because $\mathbf{Z}^{(\delta)}$ and $\mathbf{Z}^{(24)}$ are highly correlated. A solution to account for dependence is ignoring the dependence initially, thus working with MLE under misspecification, and then making adjustments to estimates of parameter uncertainty (Chandler and Bate 2007; Davison 2003). More precisely, one has

$$\hat{\phi} \to N\left(\phi_0, I(\phi_0)^{-1}\, V(\phi_0)\, I(\phi_0)^{-1}\right), \qquad \text{as } k \to \infty, \tag{15}$$

where ϕ_0 is the vector of true parameters, $V(\phi_0) = \mathrm{Cov}\,[\nabla l(\phi_0)]$, and $I(\phi_0) = -E[\nabla^2 l(\phi_0)]$ the Fisher information matrix (∇ and ∇^2 denote the gradient and Hessian, respectively). Two points are notable here. Firstly, Eq. (15) indicates that MLE for a misspecified model is still an unbiased estimator. Secondly, if the assumed model was correct (i.e. the series were independent) then we would have $V(\phi_0) = I(\phi_0)$, and Eq. (15) reduces to the the usual asymptotic distribution for the MLE (Davison 2003). In practice, to get standard errors we need estimates of $I(\phi_0)$ and $V(\phi_0)$. The estimation of $I(\phi_0)$ is straightforward, and may be approximated by $I(\hat{\phi})$. Usually, standard optimizers are able to get finite-difference based estimates for $I(\hat{\phi})$. Because of $V(\hat{\phi}) = 0$ for the MLE, the estimation of $V(\phi_0)$ is more difficult, and we have followed the guidelines of Varin and Vidoni (2005). We finally get $\hat{\Theta}_\delta = 1.696$, with error variance $\mathrm{Var}(\hat{\Theta}_\delta) = 0.129$.

Having introduced an estimation method, we address how good model $H_{24}^{\Theta_\delta}(z)$ in Eq. (11) describes the available data. A goodness-of-fit is visually assessed by inspection of QQ-plots in Fig. 1. On the vertical axis we have plotted the observed quantiles, i.e. the ordered maxima $Z_{(1)}^{(\delta)} \leq \ldots \leq Z_{(k)}^{(\delta)}$. On the horizontal axis we have plotted the theoretical quantiles provided by $H_{24}^{\Theta_\delta}(z)$, see Eq. (11). To be more precise, the QQ-plot consists of the points

$$\left(H^{-1}\left(\frac{i}{k+1}; \bar{\mu}_\delta, \bar{\sigma}_\delta, \hat{\gamma}\right), Z_{(i)}^{(\delta)}\right), \qquad i = 1, \ldots, k, \tag{16}$$

where the GEV-coefficients are calculated by Eq. (12): $\bar{\mu}_\delta = \hat{\mu}_{24} - \hat{\sigma}_{24}\,(1 - \hat{\Theta}_\delta^{\hat{\gamma}})/\hat{\gamma}$ and $\bar{\sigma}_\delta = \hat{\sigma}_{24}\,\hat{\Theta}_\delta^{\hat{\gamma}}$, see Table 1 (first line) for these values. For comparison, the GEV-coefficients μ_{24} and σ_{24} for daily extremes are also listed, and differ strongly from those of the 24-h extremes. In addition, the GEV-parameters estimated directly from $\mathbf{Z}^{(\delta)}$ (Table 1, second line) agree very well with those obtained here. Since the points are close to the leading diagonal, we may conclude with an excellent fit of the new model.

3.3 Sensitivity of Θ_δ when δ decreases

We examine to which extend assumption $H(z) \approx H_\delta(z)$ is valid, such that Eq. (13) might be a plausible model for practical applications. Annual maxima $Z^{(\delta)}$ are constructed

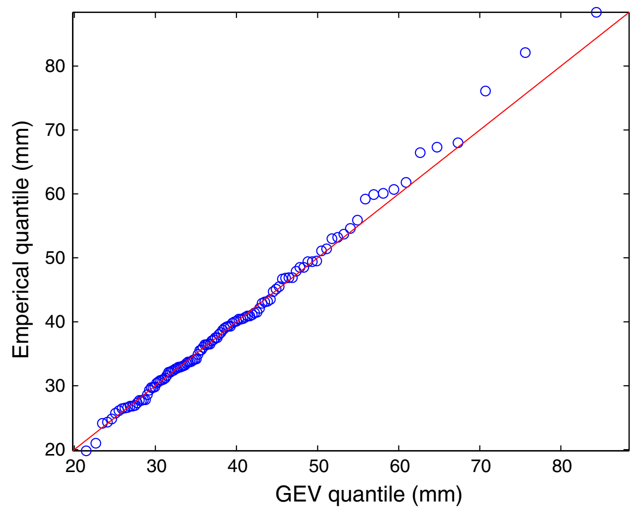


Fig. 1 QQ-plot for 24-h extremes from the Uccle series (1898–2012). *Dots*: Eq. (16). *Solid line* leading diagonal

Table 1 Estimation results for daily and 24-h extremes from the Uccle series (1898–2012)

Log-likelihood	μ_{24} (mm)	σ_{24} (mm)	γ	μ_δ (mm)	σ_δ (mm)
$l(\mathbf{Z}^{(\delta)}, \mathbf{Z}^{(24)} \mid \phi)$, Eq. (14):	29.55	8.20	0.0865	33.98	8.58
$l(\mathbf{Z}^{(\delta)} \mid \psi_\delta)$:	–	–	0.0900	33.97	8.60

for $\delta = 10$, 20, 30, 60, 120, 360, 720 and 1,440 min. Automatic stations provide 5-min measurements, but these series are shorter than 10 year, which is much too short for an extreme value analysis. Estimated Θ_δ-values, as a function of δ, are shown in Fig. 2. In addition, the estimated distribution parameters $\bar{\mu}_\delta$, $\bar{\sigma}_\delta$ and $\hat{\gamma}$ are also plotted in Fig. 2. There is a clear convergence as $\delta \to 10$ min. Furthermore, working with hourly sampled 24-h rainfall may already result in good approximations to continuous 24-h maxima.

3.4 Constancy of Θ_δ over space

Practical applications involve the following methodology. If only daily maxima are available at a certain site, and $H(z; \psi_{24})$ is fitted to this series, the distribution of continuous 24-h maxima can be approximated by $H^{\hat{\Theta}_\delta}(z; \hat{\psi}_{24})$ (cfr. Eq. (11)), where Θ_δ is estimated as before. A key point in our methodology is that Θ_δ can be modeled as a single value for the entire study region. This idea agrees with earlier empirical scaling rules (quantiles of daily vs. 24-h maxima) where the conversion factors are supposed not to vary too much in regions with the same precipitation climate (Dwyer and Reed 1995). Indeed, a storm split (at 08:00) between two observations is under-recorded by fixed interval measurements. If a typical storm arises from

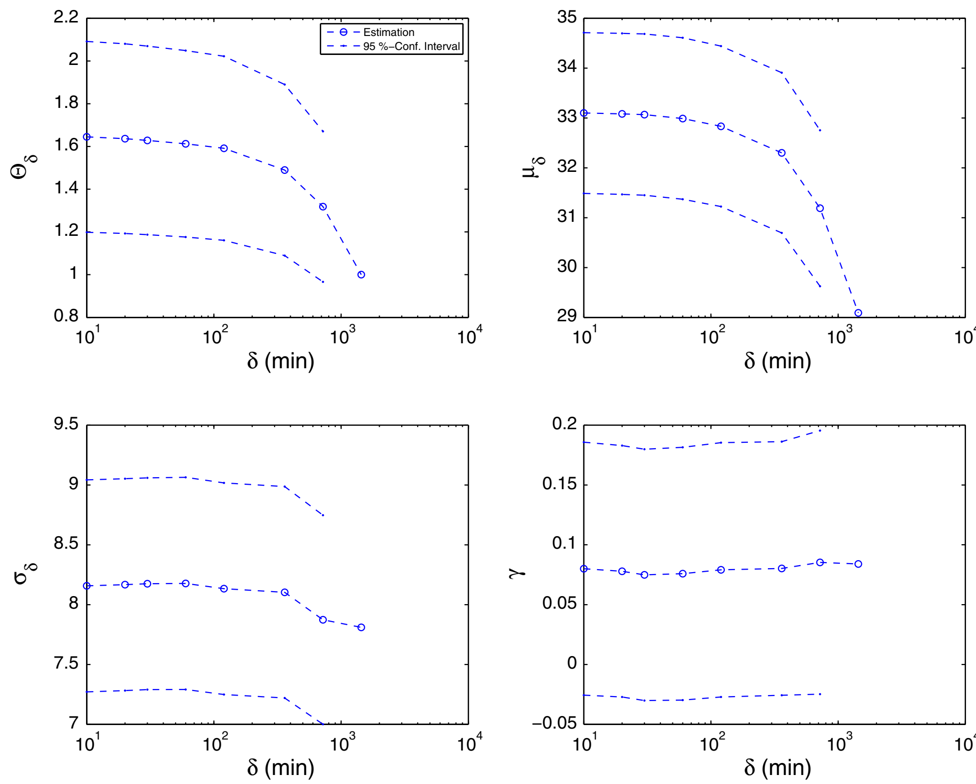


Fig. 2 Sensitivity of $\hat{\Theta}_\delta$, $\bar{\mu}_\delta$, $\bar{\sigma}_\delta$ and $\hat{\gamma}$ as $\delta \rightarrow 10$ min

relatively few hours of intense rainfall, then there is less chance of this happening than for a site experiencing longer duration events. In general, the climate in Western Europe belongs to the type Cfb in the Köppen classification (Pidwirny 2008). To illustrate the equality of Θ_δ, we have used the 115-year Uccle series, with the addition of 17 Belgian pluviograph stations which also provide 10-min measurements, but for a much shorter period (1967–2004), see Fig. 4. There are statistical tests to indicate the regional variability of parameters of a probability distribution, see for example Van de Vyver (2012). However, they are not immediately applicable here because of the spatial dependence among the series and low density of the pluviograph network. Alternatively, we can plot $\hat{\Theta}_\delta$, as estimated at each station, together with the 95 %-confidence bounds, see Fig. 3 (left). 115-year Uccle series features much shorter confidence intervals, as we would expect. Overall, no obvious inter-site differences in Θ_δ are discernible from visual examination. In Fig. 3 (right), the scatter plot of $\hat{\Theta}_\delta$ against the height H shows that fluctuations in $\hat{\Theta}_\delta$ cannot be explained by orographic effects. Taken together, the above arguments appear to support the hypothesis that fluctuations in $\hat{\Theta}_\delta$ are probably due to sample-to-sample variability, and that Θ_δ can be kept constant when modeling continuous 24-h precipitation maxima. In the following section the present MLE estimator generalizes further to a spatial estimator in which it is also assumed to have a constant Θ_δ-value. As we will see, validation results are equally good which justifies, again, our main hypothesis here.

4 Spatial estimation of Θ_δ

The foregoing analysis clearly shows the ability of the new methodology for single-site estimations. The past decade, there is a growing interest in modeling spatial extremes (Cooley et al. 2012; Davison et al. 2012; Ribatet 2011). In fact, the use of spatial data appears so often in atmospheric sciences that the construction of models for them is currently seen as a well-established area of investigation. We restrict our attention here to the modeling of marginal distributions, rather than the multivariate case based on copula models or max-stable processes. Our goal is to investigate if the new estimation method can be naturally extended to spatial precipitation data.

4.1 Estimation of Θ_δ by using spatial regression models

A spatial regression model for annual maxima (daily, or continuous 24-h precipitation) is defined as $Z(\mathbf{r}) \sim$ GEV $[\mu(\mathbf{r}), \sigma(\mathbf{r}), \gamma(\mathbf{r})]$, where $\mathbf{r}$ is the location (expressed in longitude/latitude, or other geographic coordinates). The parameters characterize the regional variability in extreme precipitation, and are possibly related to climatological and orographic effects. In Van de Vyver (2012, 2013), the following spatial regression models were investigated:

$$
\begin{aligned}
&\mathrm{GEV}_{10}^{(alt)}: \quad \mu(\mathbf{r}) = \mu^{(0)} + \mu^{(1)} H(\mathbf{r}), \quad \sigma(\mathbf{r}) = \sigma^{(0)},\\
&\mathrm{GEV}_{11}^{(alt)}: \quad \mu(\mathbf{r}) = \mu^{(0)} + \mu^{(1)} H(\mathbf{r}), \quad \sigma(\mathbf{r}) = \sigma^{(0)} + \sigma^{(1)} H(\mathbf{r}),\\
&\mathrm{GEV}_{10}^{(mar)}: \quad \mu(\mathbf{r}) = \mu^{(0)} + \mu^{(1)} mar(\mathbf{r}), \quad \sigma(\mathbf{r}) = \sigma^{(0)},\\
&\mathrm{GEV}_{11}^{(mar)}: \quad \mu(\mathbf{r}) = \mu^{(0)} + \mu^{(1)} mar(\mathbf{r}), \quad \sigma(\mathbf{r}) = \sigma^{(0)} + \sigma^{(1)} mar(\mathbf{r}),
\end{aligned}
$$

and $\gamma(\mathbf{r}) = \gamma$ for all models. Here, covariates $H(\mathbf{r})$ and $mar(\mathbf{r})$ respectively refer to height and mean annual rainfall at location $\mathbf{r}$. It was found that *mar* is by far the strongest covariate for extreme precipitation, a result which was also confirmed in Blanchet and Lehning (2010), Coles and Tawn (1996), Cooley et al. (2007), Smith (1990). In particular, numerous validation tests in Van de Vyver (2012, 2013) indicate an excellent performance of $\mathrm{GEV}_{11}^{(mar)}$. Next, a natural extension of Eq. (11) for spatial regression models is

$$
H\big(z; \mu_\delta(\mathbf{r}), \sigma_\delta(\mathbf{r}), \gamma\big) \approx H^{\Theta_\delta}\big(z; \mu_{24}(\mathbf{r}), \sigma_{24}(\mathbf{r}), \gamma\big), \tag{17}
$$

and our aim is now to estimate Θ_δ by using Eq. (17). Of course, the single-site estimation of Θ_δ (Sect. 3.2) can be used in practise, but combining data from several sites lead to a significant reduction in the error $\mathrm{Var}(\hat{\Theta}_\delta)$. For spatial data we adopt the following notation: $\mathbf{Z}^{(\delta)} = (Z_{ij}^{(\delta)})$ and $\mathbf{Z}^{(24)} = (Z_{ij}^{(24)})$ where subscripts $i = 1, \ldots, k$ and $j = 1, \ldots, n_s$ refer to the year and station number, respectively. Putting everything together, a spatial extension of log-likelihood Eq. (14) becomes

$$
\begin{aligned}
l(\mathbf{Z}^{(\delta)}, \mathbf{Z}^{(24)} | \phi) = & -k n_s \gamma \log \Theta_\delta - k \sum_{j=1}^{n_s} \log \sigma_{24}(\mathbf{r}_j)\\
& - \sum_{i=1}^{k} \sum_{j=1}^{n_s} \left(\frac{\sigma_{24}(\mathbf{r}_j) + \gamma\,(Z_{ij}^{(\delta)} - \mu_{24}(\mathbf{r}_j))}{\sigma_{24}(j)\, \Theta_\delta^{\gamma}} \right)^{-1/\gamma}\\
& - (1/\gamma + 1) \sum_{i=1}^{k} \sum_{j=1}^{n_s} \log \left(\frac{\sigma_{24}(\mathbf{r}_j) + \gamma\,(Z_{ij}^{(\delta)} - \mu_{24}(\mathbf{r}_j))}{\sigma_{24}(\mathbf{r}_j)\, \Theta_\delta^{\gamma}} \right)\\
& - k \sum_{j=1}^{n_s} \log \sigma_{24}(\mathbf{r}_j) - \sum_{i=1}^{k} \sum_{j=1}^{n_s} \left(1 + \gamma \left(\frac{Z_{ij}^{(24)} - \mu_{24}(\mathbf{r}_j)}{\sigma_{24}(j)} \right) \right)^{-1/\gamma}\\
& - (1/\gamma + 1) \sum_{i=1}^{k} \sum_{j=1}^{n_s} \log \left(1 + \gamma \left(\frac{Z_{ij}^{(24)} - \mu_{24}(\mathbf{r}_j)}{\sigma_{24}(\mathbf{r}_j)} \right) \right).
\end{aligned} \tag{18}
$$

The parameter vector ϕ depends on the selected spatial model. For example, if we consider $\mathrm{GEV}_{11}^{(mar)}$, one has $\phi = (\mu_{24}^{(0)}, \mu_{24}^{(1)}, \sigma_{24}^{(0)}, \sigma_{24}^{(1)}, \gamma, \Theta_\delta)$.

Beside the dependence between $\mathbf{Z}^{(\delta)}$ and $\mathbf{Z}^{(24)}$, one has additionally to take into account the spatial dependency of the data. Analogously to the methodology of Sect. 3.2, the likelihood function Eq. (18) was constructed as if the data was independent, but Eq. (15) is proposed for obtaining standard errors which allows for dependence. In the context of model selection we use the Takeuchi Information Criterion (TIC), defined as

$$
\mathrm{TIC} = -2\, l(\hat{\phi}) + 2\ \mathrm{Tr}\ \Big(I(\hat{\phi})^{-1}\, V(\hat{\phi}) \Big),
$$

where $l(\hat{\phi})$ is the maximized log-likelihood Eq. (18). The best model will be that having the lowest value of TIC.

4.2 Implementation and validation

We have used the 18 Belgian pluviograph stations which provide 10-min measurements (period: 1967–2004), see Fig. 4. Estimation results have been listed in Table 2. Compared to the single-site estimation in Sect. 3.2, there is a marked reduction of $\mathrm{Var}(\hat{\Theta}_\delta)$. Model $\mathrm{GEV}_{11}^{(mar)}$ has the lowest TIC-value, and is thus the version to choose, as we would expect.

Again, as in Sect. 3.2 for single-site estimations, we should evaluate how good model $H^{\Theta_\delta}\big(z; \mu_{24}(\mathbf{r}), \sigma_{24}(\mathbf{r}), \gamma\big)$ in Eq. (17) describes the available data. In the present context, classical QQ-plots are not useful as the spatial data is not identically distributed. A possible extension of classical QQ-plots consists in transforming the data to variables that satisfy the iid property. Assume $Z^{(\delta)}(\mathbf{r}_j) \sim$ GEV $[\mu^{(\delta)}(\mathbf{r}_j), \sigma^{(\delta)}(\mathbf{r}_j), \gamma]$, $j = 1, \ldots, n_s$. The transformation

$$
\tilde{Z}^{(\delta)}(\mathbf{r}_j) = \frac{1}{\gamma}\ \log\ \left(1 + \gamma\, \frac{Z^{(\delta)}(\mathbf{r}_j) - \mu^{(\delta)}(\mathbf{r}_j)}{\sigma^{(\delta)}(\mathbf{r}_j)} \right), \tag{19}
$$

results in a Gumbel distributed random variable Beirlant (2004), Coles (2001), i.e. $P\big(\tilde{Z}^{(\delta)}(\mathbf{r}_j) \le z\big) = \exp\big(-\exp(-z)\big)$. The Gumbel distribution obtained does not any longer depend on the covariates, and hence the random variable $\tilde{Z}^{(\delta)}(\mathbf{r}_j) =: \tilde{Z}^{(\delta)}$ is identically distributed. Analogously to Sect. 3.2 we can calculate $\bar{\mu}^{(\delta)}(\mathbf{r}_j)$, $\bar{\sigma}^{(\delta)}(\mathbf{r}_j)$ and $\hat{\gamma}$ via Eq. (12), and then substitute these values in Eq. (19). Next, the spatial data $(Z_{ij}^{(\delta)})$ play the role of $Z^{(\delta)}(\mathbf{r}_j)$ in Eq. (19). The quantile function associated with the Gumbel distribution is given by

$$
Q(p) = -\log(-\log p), \qquad 0 < p < 1,
$$

yielding the Gumbel QQ-plot coordinates

$$
\left(-\log\left(-\log \frac{i}{m+1} \right), \tilde{Z}_{(i)}^{(\delta)} \right), \qquad i = 1, \ldots, m, \tag{20}
$$

where $\tilde{Z}_{(1)}^{(\delta)} \le \cdots \le \tilde{Z}_{(m)}^{(\delta)}$ are the corresponding order statistics of $\tilde{Z}^{(\delta)}$, and $m = k\, n_s$ the total number of spatial data. If the Gumbel model provides accurate description of the

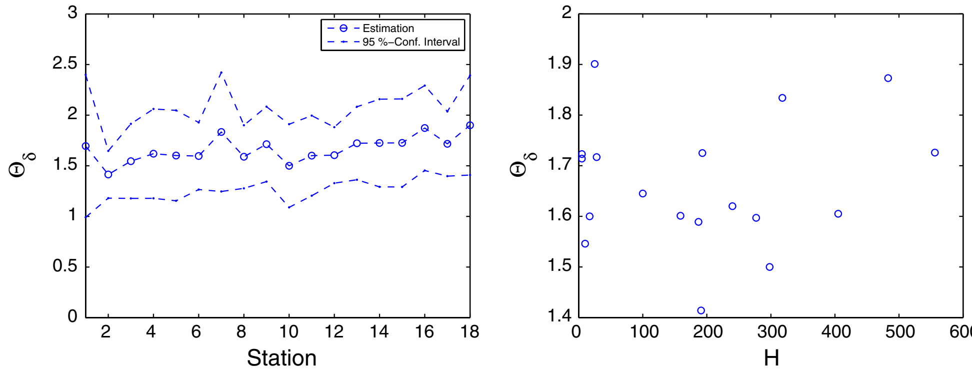


Fig. 3 *Left*: $\hat{\Theta}_\delta$ against station number, together with 95 %-confidence intervals. *Right*: scatterplot of $\hat{\Theta}_\delta$ against height *H*

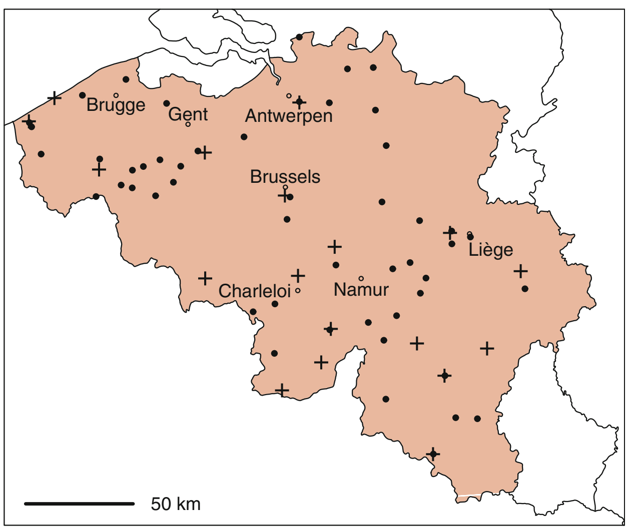


Fig. 4 Location of daily precipitation stations (*dots*), and 10-min precipitation stations (*crosses*)

data, one expects the points on the Gumbel QQ-plot to be close to the leading diagonal. Observing Fig. 5, one can conclude that model $H^{\Theta_\delta}(z;\mu_{24}(\mathbf{r}),\sigma_{24}(\mathbf{r}),\gamma)$ in Eq. (17) reasonably describes the 24-h extremes.

5 Application: return level estimation

Let us finally recall that the ultimate goal of extreme value analysis is to provide return levels Eq. (3) of extreme events. For practical applications we assume that only daily measurements were recorded at a certain station, and the GEV-parameters are $\psi_{24}=(\mu_{24},\sigma_{24},\gamma)$. The GEV-parameters $\psi_\delta=(\mu_\delta,\sigma_\delta,\gamma)$ of 24-h maxima are approximated by using the conversion rule Eq. (12), here shortly denoted by $\psi_\delta=f(\psi_{24},\Theta_\delta)$. The MLE is $\bar{\psi}_\delta=f(\hat{\psi}_{24},\hat{\Theta}_\delta)$, where we use the conversion exponent provided by $\mathrm{GEV}_{11}^{(mar)}$, i.e. $\hat{\Theta}_\delta=1.655$ with $\mathrm{Var}(\hat{\Theta}_\delta)=0.0019$ (see Table 2). By substitution of $\bar{\psi}_\delta$ in Eq. (3), the MLE $\bar{z}_p$ is obtained. By the delta-method, the estimation error of $\bar{z}_p$ is

$$\mathrm{Var}\,(\bar{z}_p)=\nabla z_p^T\,V_{\psi_{24}}\,\nabla z_p+\left(\frac{\partial(z_p\circ f)}{\partial\Theta_\delta}\right)^2\,\mathrm{Var}\,(\hat{\Theta}_\delta),\tag{21}$$

evaluated at $(\hat{\psi}_{24},\hat{\Theta}_\delta)$, where $V_{\psi_{24}}$ is the variance-covariance of $\hat{\psi}_{24}$, and

$$\begin{aligned}\nabla z_p^T&=\left[\frac{\partial(z_p\circ f)}{\partial\,\mu_{24}},\frac{\partial(z_p\circ f)}{\partial\,\sigma_{24}},\frac{\partial(z_p\circ f)}{\partial\,\gamma}\right]\\&=\Big[1,-\gamma^{-1}\left(1-y_p^{-\gamma}\,\Theta_\delta^\gamma\right),\sigma_{24}\,\gamma^{-2}\left(1-y_p^{-\gamma}\,\Theta_\delta^\gamma\right)\\&\quad-\sigma_{24}\,\gamma^{-1}\,y_p^{-\gamma}\,\Theta_\delta^\gamma\,\log y_p+\gamma^{-1}\,\sigma_{24}\,y_p^{-\gamma}\,\Theta_\delta^\gamma\,\log\Theta_\delta\Big],\end{aligned}$$

and

$$\frac{\partial(z_p\circ f)}{\partial\Theta_\delta}=\sigma_{24}\,\Theta_\delta^{\gamma-1}\,y_p^{-\gamma},\qquad\text{where }\;y_p=-\log(1-p).$$

To validate the new return level estimator $\bar{z}_p$, we use as reference the direct MLE, which is based on the 10-min sequence of 24-h aggregations, here denoted by $\hat{z}_p:=z_p(\hat{\psi}_\delta)$.

Return level plots, together with the standard error, are plotted in Fig. 6 for two sites having 10-min

Table 2 Spatial estimation of Θ_δ

Method	$\hat{\Theta}_\delta$	$\mathrm{Var}(\hat{\Theta}_\delta)$	TIC	Method	$\hat{\Theta}_\delta$	$\mathrm{Var}(\hat{\Theta}_\delta)$	TIC
$\mathrm{GEV}_{10}^{(alt)}$	1.607	0.0019	9,667.68	$\mathrm{GEV}_{11}^{(alt)}$	1.612	0.0018	9,653.00
$\mathrm{GEV}_{10}^{(mar)}$	1.656	0.0020	9,597.22	$\mathrm{GEV}_{11}^{(mar)}$	1.655	0.0019	9,582.71

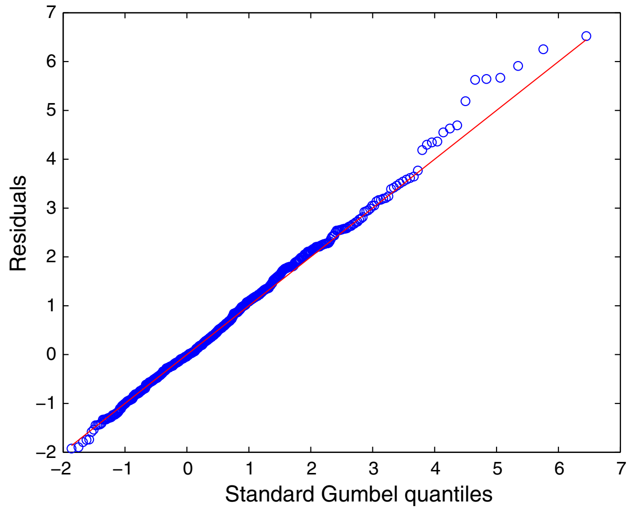
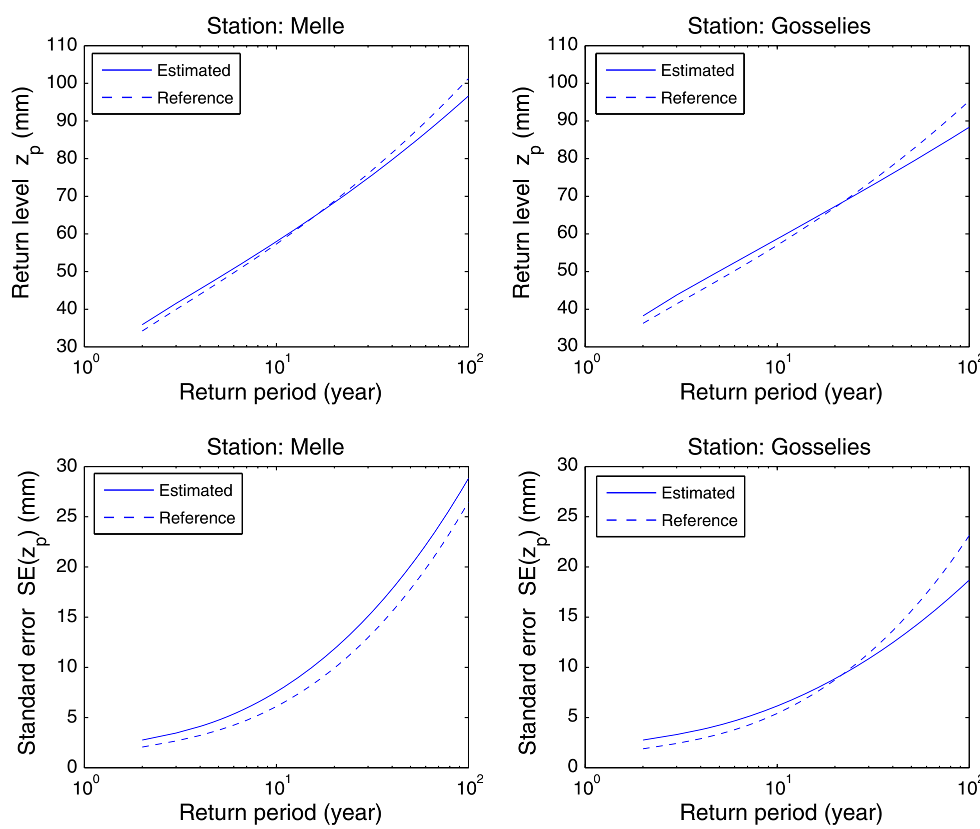


Fig. 5 Gumbel QQ-plot for 24-h extremes from the Belgian pluviograph series (Fig. 4). *Dots*: Eq. (20). *Solid line* leading diagonal

measurements. For comparison, the curve agrees well with the reference. The error Eq. (21) differs to the reference error, but there is no clear relation found between both errors. For example, in some stations the error is seen to be smaller than the reference, in other stations not.

Analogously to single-site estimation, we consider spatial data of daily measurements. We have collected rainfall extremes of 68 daily series (Fig. 4), each of them containing 60 years of data (1951–2010) (Van de Vyver 2012). Next, we fit the spatial model $\mathrm{GEV}_{11}^{(mar)}$ to this data using R-package *SpatialExtremes* (Ribatet 2011), resulting in parameter values $\hat{\psi}_{24} = (\hat{\mu}_{24}^{(0)}, \hat{\mu}_{24}^{(1)}, \hat{\sigma}_{24}^{(0)}, \hat{\sigma}_{24}^{(1)}, \hat{\gamma})$. A spatial regression model for continuous 24-h precipitation maxima is then obtained by Eq. (17), giving $\bar{\psi}_{\delta} = (\bar{\mu}_{\delta}^{(0)}, \bar{\mu}_{\delta}^{(1)}, \bar{\sigma}_{\delta}^{(0)}, \bar{\sigma}_{\delta}^{(1)}, \hat{\gamma})$. Finally, return level computation and spatial extension of the error variance Eq. (21) is straightforward. The 20- and 100-year return levels as a function of *mar* (typically ranging from 700 to 1,400 mm), together with the standard error are displayed in Fig. 7. As reference we used a spatial model that is fitted directly to 24-h maximum series from 18 pluviograph stations, see Sect. 4.2 and Fig. 4. There is an acceptable agreement between the return levels obtained by both approaches. We can see that for *mar*-values up to 900 mm, the error estimated by the new methodology is just a little larger than the reference error, which is obtained from a much sparser network. It is easy to verify that the errors in the return levels are quadratic in the covariate *mar*, so that the difference between both errors considerably grows for larger *mar*-values. Here, we can see that this is the case in the higher-altitude parts of Belgium, where *mar*-values are larger than 1,000 mm.

Fig. 6 *Top* return level plots for 24-h precipitation maxima. Estimation from daily maximum series, and then conversion by Eq. (12). Reference: direct estimation from the individual 24-h maximum series. *Bottom* corresponding standard error



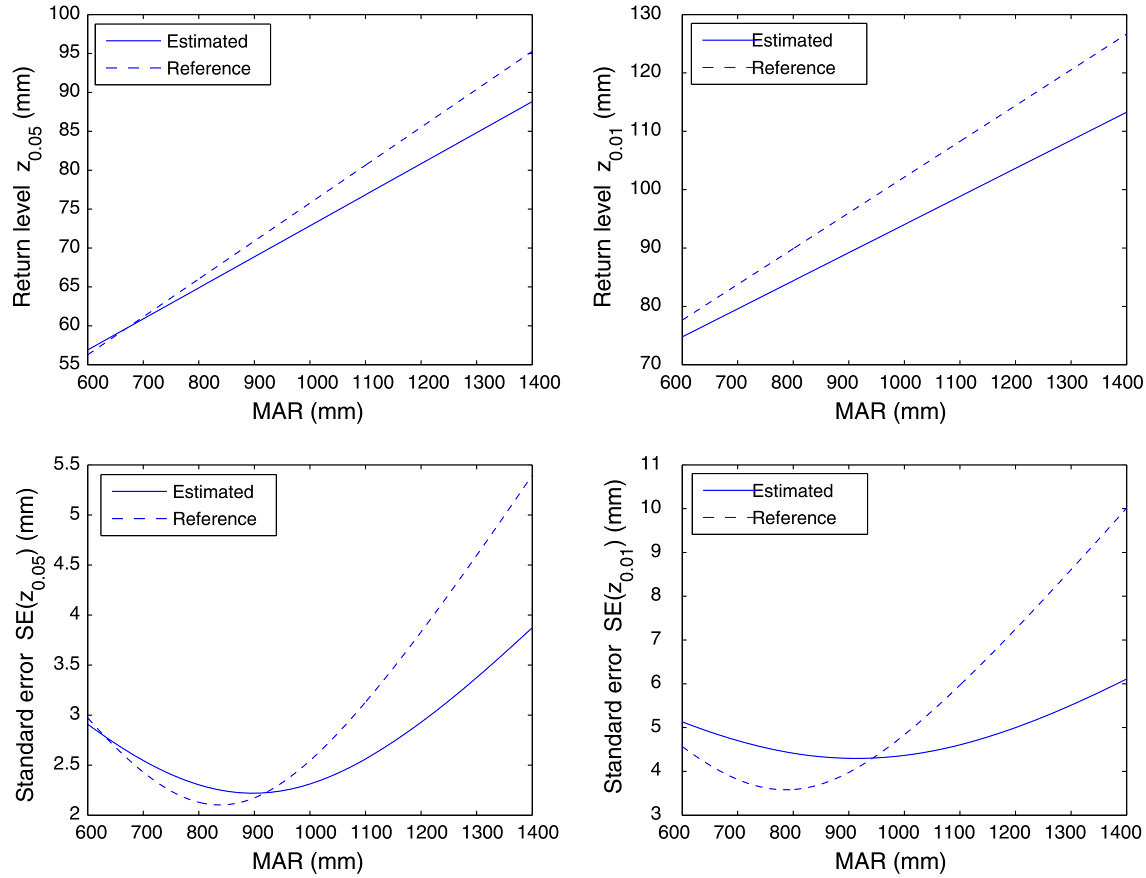


Fig. 7 *Top* return level plots for 24-h precipitation maxima provided by $\mathrm{GEV}_{11}^{(mar)}$. Estimation from daily maximum series, and then conversion by Eq. (17). Reference: $\mathrm{GEV}_{11}^{(mar)}$ which was directly estimated from 24-h maximum series, 18 in total. *Bottom* corresponding standard error. Return periods are $T = 20$ year (*left*) and $T = 100$ year (*right*)

6 Conclusions

We conclude with some remarks about the usefulness of the new methodology. The main aim was to estimate GEV-distribution parameters of 24-h precipitation maxima, given the fact that only daily measurements are available. It was already pointed out in Robinson and Tawn (2000) that for a known GEV-distribution $H_{24}(z)$ of daily maxima, the distribution of high-frequency sampled 24-h maxima is of the form $H_{24}^{\Theta_\delta}(z)$. Throughout this work, it was assumed that a sampling time δ exists for which the associated maxima are a sufficiently good approximation to continuous 24-h precipitation maxima. Although we have included 10-min observations in our analysis, it turned out that hourly-sampled time observations may already give good approximations to continuous 24-h maxima. Assuming a constant Θ_δ-value, a condition which is certainly met in a region with the same precipitation climate, the conversion rule $H_{24}^{\Theta_\delta}(z)$ can be applied to obtain GEV-parameters of 24-h maxima.

An important amount of work was devoted to estimating the conversion exponent Θ_δ. Here we have introduced a new MLE-scheme which enables us to assess the estimation error of $\hat{\Theta}_\delta$. Firstly, the estimation of Θ_δ was demonstrated with the 115-year time series of 10-min rainfall recorded in Uccle. Secondly, the estimation of Θ_δ was further refined by combining 10-min data from several sites, resulting in a marked reduction in $\mathrm{Var}(\hat{\Theta}_\delta)$. The results of practical return level estimation for single series are promising, showing that it is possible to overcome sampling issues. For spatial modeling of 24-h maxima, it is likely that the new scheme is more useful than using the 10-min series directly because this network is less dense than for daily series. The question arises whether or not the number of 10-min series is too low for modeling spatial differences in extreme 24-h precipitation. In particular, in higher-altitude parts of Belgium where $mar > 1{,}000$ mm, the estimated errors in the return levels provided by the new methodology are superior compared to those of the direct estimation. However, such a comparison is not free of subjective or empirical considerations since the standard error is an estimation too, and is only derived from the model itself. Taking account that the above conclusion is made on the assumption of having known *mar*-values. In practical return level mapping, gridded *mar*-values are obtained by spatial interpolation (kriging), which introduces additional uncertainties (Van de Vyver 2013). Complete inference of spatial regression models which includes all these types of errors is reserved for a future paper.

This analysis can be immediately extended in various ways. Firstly, the estimation of 48-h precipitation extremes is an equally important issue in hydrological sciences (Dwyer and Reed 1995; van Montfort 1990), although, the difference with two-daily maxima is much smaller than in the present case. Likewise, the effect of sampling frequency is also present on other climatological variables such as air temperature and wind speed (Dwyer and Reed 1995), such that the new methodology could be applied there as well. Secondly, seasonal effects are not distinguished, possibly due to different climate patterns in different months. It might be interesting to examine whether there are significant differences between the conversion exponents of different seasons. Finally, an interesting further development is the introduction of Bayesian techniques in our methodology in order to provide a more complete inference.

Acknowledgments This work is supported, in part, by the Belgian Science Policy Office (BELSPO) under Contract No. SD/RI/03A.

References

Ancona-Navarette MA, Tawn JA (2000) A comparison of methods for estimating the extremal index. Extremes 3:5–38

Beirlant J, Goegebeur Y, Segers J, Teugels J (2004) Statistics of extremes. Wiley, Sussex

Blanchet J, Lehning M (2010) Mapping snow depth return levels: smooth spatial modeling versus station interpolation. Hydrol Earth Syst Sci 14:2527–2544

Chandler RE, Bate S (2007) Inference for clustered data using the independence loglikelihood. Biometrika 94:167–183

Chernick MR (1981) A limit theorem for the maximum of autoregressive processes with uniform marginal distributions. Ann Probab 9:145–149

Coles S (2001) An introduction to statistical modelling of extreme values. Springer, Berlin

Coles S, Tawn JA (1996) Modeling extremes of the areal rainfall process. J R Stat Soc B 58:329–347

Cooley D, Nychka D, Naveau P (2007) Bayesian spatial modeling of extreme precipitation return levels. J Am Stat Assoc 102: 824–840

Cooley D, Cisewski J, Erhardt R, Jeon S, Mannshardt E, Omolo B, Sun Y (2012) A survey of spatial extremes: measuring spatial dependence and modeling spatial effects. REVSTAT 10: 135–165

Davison A (2003) Statistical models, Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge

Davison A, Padoan S, Ribatet M (2012) Statistical modelling of spatial extremes. Stat Sci 27:161–186

Demarée GR (2003) Le pluviographe centenaire du plateau d'Uccle: son histoire, ses données et ses applications. La Houille Blanche 4:1–8

Dwyer IJ, Reed DW (1995) Allowance for discretization in hydrological and environmental risk estimation. Report 123, Institute of Hydrology, Wallingford

Embrechts P, Klüppelberg C, Mikosch T (1997) Modelling extremal events for insurance and finance. Springer, Berlin

Gomes MI (1993) On the estimation of parameters of rare events in environmental time series. In: Barnett V, Turkman KF (eds.) Statistics for the environment 2: water related issues: pp 225–241

Hershfield DM, Wilson WT (1958) Generalizing of rainfall-intensity-frequency data. IUGG/IAHS 43:499–506

Leadbetter MR (1983) Extremes and local dependence in stationary sequences. Z Wahrscheinlichkeit 65:291–306

Leadbetter MR, Lindgren G, Rootzén T (1983) Extremes and related properties of random sequences and processes. Springer, New York

Obeysekera J, Irizarry M, Park J, Barnes J, Dessalegne T (2011) Climate change and its implications for water resources management in south Florida. Stoch Environ Res Risk Assess 25:495–516

Piterbarg VI (2004) Discrete and continuous time extremes of Gaussian processes. Extremes 7:161–177

Pidwirny M (2008) Fundamentals of physical geography, 2nd Ed. PhysicalGeography.net

Ribatet M (2011) SpatialExtremes: an R package for modelling spatial extremes, http://cran.r-project.org/web/packages/SpatialExtremes/

Robinson ME, Tawn JA (2000) Extremal analysis of processes sampled at different frequencies. J R Stat Soc B 62:117–135

Siliverstovs B, Otsch R, Kemfert C, Jaeger CC, Haas A, Kremers H (2010) Climate change and modelling of extreme temperatures in Switzerland. Stoch Environ Res Risk Assess 24:311–326

Smith RL (1990) Regional estimation from spatially dependent data. Unpublished manuscript, University of Chapel Hill

Smith RL, Weissman I (1994) Estimating the extremal index. J R Stat Soc B 56:515–528

Turkman KF (2011) Discrete and continuous time extremes of stationary processes. In: Handbook of statistics. Time series analysis: methods and applications 30:565–582

Van de Vyver H (2012) Spatial regression models for extreme precipitation in Belgium. Water Resour Res 48:W09549

Van de Vyver H (2013) Practical return level mapping for extreme precipitation in Belgium. Publication scientific et technique 62, L'Institut Royal Météorologique de Belgique

van Montfort MAJ (1990) Sliding maxima. J Hydrol 118:77–85

Varin C, Vidoni P (2005) A note on composite likelihood inference and model selection. Biometrika 92:519–528

Weiss LL (1964) Ratio of true to fixed-interval maximum rainfall. J Hydraul Div ASCE 90:77–82

Yoon P, Kim T-W, Yoo C (2013) Rainfall frequency analysis using a mixed GEV distribution: a case study for annual maximum rainfalls in South Korea. Stoch Environ Res Risk Assess 27:1143–1153