A regional peaks-over-threshold model in a nonstationary climate

M. Roth[1,2], T. A. Buishand[2], G. Jongbloed[3], A. M. G. Klein Tank[2], and J. H. van Zanten[1]

M. Roth and J. H. van Zanten, EURANDOM, Eindhoven University of Technology, PO Box 513, Eindhoven, 5600 MB, Netherlands. (m.roth@tue.nl; J.H.v.Zanten@tue.nl)

T. A. Buishand and A. M. G. Klein Tank, Royal Netherlands Meteorological Institute, PO Box 201, 3730 AE De Bilt, Netherlands. (buishand@knmi.nl; albert.klein.tank@knmi.nl)

G. Jongbloed, Delft Institute of Applied Mathematics, Delft University of Technology, Mekelweg 4, 2628 CD Delft , Netherlands. (G.Jongbloed@tudelft.nl)

[1]EURANDOM, Eindhoven University of Technology, Netherlands

[2]Royal Netherlands Meteorological Institute, De Bilt, Netherlands.

[3]Delft Institute of Applied Mathematics, Delft University of Technology, Netherlands

Abstract. Regional frequency analysis (RFA) is often used to reduce the uncertainty in the estimation of distribution parameters and quantiles. In this paper a regional peaks-over-threshold (POT) model is introduced that can be used to analyze precipitation extremes in a changing climate. We use a temporally varying threshold, which is determined by quantile regression for each site separately. The marginal distributions of the excesses are described by generalized Pareto (GP) distributions. The parameters of these distributions may vary over time and their spatial variation is modeled by the index flood (IF) approach. We consider different models for the temporal dependence of the GP parameters. Parameter estimation is based on the framework of composite likelihood. Composite likelihood ratio tests that account for spatial dependence are used to test the significance of temporal trends in the model parameters and to test the IF assumption.

We apply the method to gridded, observed daily precipitation data from the Netherlands for the winter season. A general increase of the threshold is observed, especially along the west coast and northern parts of the country. This implies, that moderate extremes have increased over the observed time period. Moreover, the positive trend in the threshold induces an increase in the scale parameter of the GP distribution owing to the IF assumption. There is no additional trend in the scale parameter and the trend in the shape parameter is not significant.

1. Introduction

Design values for infrastructure are often based on characteristics of extreme precipitation. These characteristics may have changed over time owing to climate change, see e.g. *Klein Tank and Können* [2003] and *Milly et al.* [2008], which contradicts the stationarity assumption, that is usually made in hydrologic and hydraulic design. Wrongly assuming stationarity generally leads to systematic errors in design values and might have a considerable impact on the risk of failure of hydraulic structures, as shown by *Wigley* [2009]. Climate scientists have analyzed trends in moderate extremes, that occur once or several times per year, based on annual indices. Examples are the empirical annual 90% quantile of the precipitation amounts on wet days or the 1-day or 5-day maximum precipitation amount in each year, see e.g. *Klein Tank and Können* [2003] and *Turco and Llasat* [2011].

In this study we focus on rare extremes which occur less frequently than once per year. These are frequently assessed by extreme value (EV) models. To account for the temporal trend in the distribution, the parameters of the EV model are often selected to be time dependent [*Smith*, 1986; *Kharin and Zwiers*, 2005; *Brown et al.*, 2008; *Hanel et al.*, 2009; *Kyselý et al.*, 2010; *Beguería et al.*, 2011]. Because of the rarity of the extremes, the parameters in these EV models and, especially, large quantiles of the precipitation amounts have wide confidence intervals. To reduce the uncertainty in the estimates the use of data sets over a long period and/or regional frequency analysis (RFA), have been recommended e.g. by *Hosking and Wallis* [1997]. Data sets over a long period are often only available for a few stations, whereas we have multiple stations that cover a relatively short time period. The idea behind RFA is to exploit the similarities between the sites in a certain region, so that all data in the region can be used to obtain quantile estimates for a particular site. The index flood (IF) approach is a popular method in RFA. It assumes that the distributions of the extreme precipitation amounts are identical after scaling with a site-specific factor (the index flood).

The IF approach has frequently been applied to describe the distribution of block maxima (BM), i.e. the largest value in a year or season. Considering only BM discards useful data in the case of multiple extremes in a block, see e.g. *Madsen et al.* [1997a] and *Kyselý et al.* [2010]. An alternative method to analyze extremes is to consider all values that exceed a certain high threshold, which is known as peaks-over-threshold (POT) modeling. A potential advantage of POT modeling is the possibility to include more data in the analysis than in the BM approach, which may reduce the estimation variance. The use of the IF assumption together with the POT approach has been studied in *Madsen and Rosbjerg* [1997] for stationary data. Here we develop a different POT model with time-varying parameters, that satisfies the IF assumption.

In section 2 we describe the proposed model. We explain the basic methods used to deal with high quantile estimation in the case of stationary data with emphasis on the POT approach. After that, we present our model for the nonstationary climate. In section 3 we outline the estimation procedure. The choice between different models is addressed in section 4 and in section 5 the application of the model to observed daily precipitation data in the Netherlands is discussed.

2. Model description

The data we describe with our model consist of measurements at S sites over a period of T time points. The data can be represented in an $S \times T$ space-time matrix

$$\mathbf{X} := (X_s(t))_{s\in\mathcal{S},t\in\mathcal{T}},$$

where $X_s(t)$ is the random variable representing the value at site s and time t, $\mathcal{S} := \{1, ..., S\}$ and $\mathcal{T} := \{1, ..., T\}$.

In POT modeling exceedances over a high threshold $u_s(t)$ are considered, $s \in \mathcal{S}$, $t \in \mathcal{T}$. This threshold is generally site specific and may depend on time. In the case of temporal clustering of the exceedances the largest value in a cluster (peak) is considered only. These peaks will then generally be approximately independent. We assume that the $X_s(t)$ have been declustered and we define $Y_s(t)$ as the difference between the daily value at site s and time t and the corresponding value of the threshold, i.e.

$$Y_s(t) := X_s(t) - u_s(t),$$

and $\mathbf{Y}$ is defined analogously to $\mathbf{X}$. The excesses are the nonnegative part of $\mathbf{Y}$. Note, that due to the declustering $Y_s(t)$ is only non-negative if there is a peak. By $\widetilde{\mathcal{T}}$ we denote the subset of days which exhibit at least one exceedance of the threshold, i.e.

$$\widetilde{\mathcal{T}} := \{t \in \mathcal{T} | \exists s \in \mathcal{S} : Y_s(t) \geq 0\}.$$

2.1. Stationary climate

2.1.1. Site specific approach

The BM approach for a stationary climate relies on the Fisher-Tippet-Gnedenko theorem for maxima of independent and identically distributed (i.i.d.) random variables. This theorem allows, under certain regularity conditions, to approximate the distribution of the BM by an extreme value distribution, see e.g. *Embrechts et al.* [1997]. The three types of extreme value distributions can be summarized in the generalized extreme value (GEV) distribution, i.e.

$$H_{\xi^*,\sigma^*,\mu^*}(x) = \begin{cases} \exp\left\{-\left[1+\xi^*\left(\frac{x-\mu^*}{\sigma^*}\right)\right]^{-1/\xi^*}\right\}, & \xi^* \neq 0, \\ \exp\left[-\exp(-\frac{x-\mu^*}{\sigma^*})\right], & \xi^* = 0, \end{cases}$$

for $1 + \xi^*(x - \mu^*)/\sigma^* > 0$, where μ^*, σ^* and ξ^* are the location, scale and shape parameter. $\xi^* > 0$ corresponds to the Fréchet family, $\xi^* < 0$ to the Weibull family and $\xi^* = 0$ to the Gumbel family.

When we consider the POT approach rather than block maxima, we have to model the process of exceedance times and the distribution of the excesses separately. In a stationary climate the the threshold u is constant and the times of exceedance are usually modeled by a homogeneous Poisson process. This implies, that the mean number λ of exceedances in a block (i.e., year or a particular season) is constant over time.

The Balkema-de Haan-Pickands theorem states, that the distribution of i.i.d. excesses can be approximated by a generalized Pareto (GP) distribution, if the threshold u is sufficiently high and certain regularity conditions hold, see e.g. *Reiss and Thomas* [2007]:

$$\begin{aligned} P(Y \leq y | Y \geq 0) &= G_{\xi,\sigma}(y) \\ &= \begin{cases} 1 - \left(1 + \frac{\xi y}{\sigma}\right)^{-1/\xi}, & \xi \neq 0, \\ 1 - \exp\left(-\frac{y}{\sigma}\right), & \xi = 0, \end{cases} \end{aligned}$$

for $y \geq 0$ if $\xi \geq 0$ and $0 \leq y \leq -\sigma/\xi$ if $\xi < 0$, where σ and ξ are the scale and the shape parameter. For $\xi = 0$ the GP distribution reduces to the exponential distribution.

We are interested in the level q_α which is exceeded on average α times in a block. Since there are on average λ peaks in a block, the probability that an arbitrary peak exceeds this level equals α/λ. To obtain q_α we first determine the $(1 - \alpha/\lambda)$-quantile of the excess distribution:

$$\tilde{q}_\alpha = G^{-1}_{\xi,\sigma}(1 - \alpha/\lambda),$$

and then add the threshold u, i.e.

$$q_\alpha = u + \tilde{q}_\alpha = \begin{cases} u + \frac{\sigma}{\xi}[1 - (\frac{\alpha}{\lambda})^{-\xi}], & \xi \neq 0, \\ u + \sigma \ln(\frac{\lambda}{\alpha}), & \xi = 0. \end{cases} \tag{1}$$

We will sometimes indicate the quantile q_α as the $1/\alpha$ return level to make the comparison with studies for a stationary climate easier.

If one assumes that the exceedance times originate from a homogeneous Poisson process and the excesses are independent and follow a GP distribution, it can be shown that the subsequent relationship between the parameters of the GEV and the GP distribution holds [*Buishand*, 1989; *Wang*, 1991; *Madsen et al.*, 1997b]:

$$\begin{aligned} \mu^* &= \begin{cases} u - \frac{\sigma}{\xi}(1 - \lambda^{\xi}), & \xi \neq 0, \\ u + \sigma \ln(\lambda), & \xi = 0, \end{cases} \\ \sigma^* &= \sigma \lambda^{\xi} \\ \xi^* &= \xi \end{aligned} \tag{2}$$

Note, that the derived GEV distribution is defined only for BM greater than u.

2.1.2. Regional approach

The IF method was originally developed for annual maxima of river discharges by *Dalrymple* [1960]. It assumes that the annual maxima at different sites, after being scaled by a site specific factor, the 'index flood', have a common distribution [e.g. *Dalrymple*, 1960; *Hosking and Wallis*, 1997; *Robinson and Sivapalan*, 1997]:

$$P\left(\frac{M_s}{\eta_s} \leq x\right) = \phi(x) \quad \forall s \in \mathcal{S} \tag{3}$$

where M_s represents a typical block maximum at site s, η_s is the index flood at site s for $s \in \mathcal{S}$ and the common distribution function ϕ does not depend on the site s.

From equation (3) we see, that the site specific quantile function can be written in the following product form:

$$q_\alpha(s) := Q_{M_s}(\alpha) = \eta_s \phi^{-1}(\tau), \tag{4}$$

where Q_{M_s} is the quantile function of M_s and τ is the non-exceedance probability.

Because of using more data than those from the site of interest alone, the IF can provide quantile estimates, which are superior to at-site estimates, even if spatial homogeneity is not entirely achieved after scaling [*Cunnane*, 1988]. The IF approach was developed for river discharges but can be applied, whenever multiple samples of similar data are available, see *Hosking and Wallis* [1997]. In particular, for precipitation data the IF assumption has often been used combined with the GEV family, see e.g. *Hosking and Wallis* [1997]; *Fowler el al.* [2005] and *Hanel et al.* [2009]. To further enhance the usage of the available data, *Madsen and Rosbjerg* [1997] propose the combination of the IF assumption with the POT approach.

A natural analogue of relation (3) in the POT setting is that the site-specific exceedances, properly scaled by their index floods, have a common distribution. More formally:

$$P\left(\frac{X_s}{\eta_s} \leq x | X_s \geq u_s\right) = \psi(x) \quad \forall s \in \mathcal{S}, \tag{5}$$

where X_s represents the values at site s, η_s is the site-dependent scaling factor (index flood) and ψ does not depend on site s. Note that because of $\psi(u_s/\eta_s) = 0, \forall s \in \mathcal{S}$ and because ψ has a density with mass immediately to the right of u_s/η_s, it follows that u_s/η_s has to be the lower endpoint of the support of ψ for every $s \in \mathcal{S}$, i.e.

$$\frac{u_i}{\eta_i} = \frac{u_j}{\eta_j} \quad \forall i, j \in \mathcal{S}. \tag{6}$$

This can be only true, if the index flood is a multiple of the threshold, i.e.

$$\eta_s = c u_s \quad \forall s \in \mathcal{S},$$

for some positive constant c. Without loss of generality we can take $c = 1$. This choice of η_s also satisfies the IF equation for the excesses, i.e.

$$P\left(\frac{Y_s}{\eta_s} \le y | Y_s \ge 0\right) = \widetilde{\psi}(y) \quad \forall s \in \mathcal{S}, \tag{7}$$

where $\widetilde{\psi}(y) := \psi(y+1)$ is independent of site s.

A natural choice for a site specific threshold is a high empirical quantile of the at-site data [see also *Smith*, 1989a]. An important consequence of this choice is that the mean number of exceedances per block λ_s will be approximately constant over the region, i.e.

$$\lambda_s \equiv \lambda.$$

Under the previous assumptions the distribution of the scaled excesses has the following form:

$$P\left(\frac{Y_s}{u_s} \le y | Y_s \ge 0\right) = G_{\xi_s, \frac{\sigma_s}{u_s}}(y). \tag{8}$$

Equation (7) then implies, that we have the following restrictions on the parameters of the GP distribution

$$\frac{\sigma_s}{u_s} \equiv \gamma, \qquad \xi_s \equiv \xi \quad \forall s \in \mathcal{S}, \tag{9}$$

for a common dispersion coefficient γ and a common shape parameter ξ.

We would like to obtain an IF model in the BM setting, if we transfer the parameters from the IF model in the POT setting, using relationship (2). If the block maxima follow a GEV distribution, it can be shown that the IF assumption is satisfied if the dispersion coefficient $\gamma_s^* := \sigma_s^* / \mu_s^*$ and the shape parameter ξ_s^* of the GEV distribution are constant over the region, see e.g. *Hanel et al.* [2009], i.e.

$$\gamma_s^* \equiv \gamma^*, \qquad \xi_s^* \equiv \xi^*, \qquad \forall s \in \mathcal{S}. \tag{10}$$

If we transform the conditions (9) according to relationship (2) and use that λ is constant over the region, we obtain the following conditions on the GEV distribution parameters:

$$\xi_s^* \equiv \xi, \tag{11}$$

$$\gamma_s^* = \begin{cases} \frac{\lambda^\xi}{\gamma^{-1} - \frac{1}{\xi}(1-\lambda^\xi)}, & \xi \neq 0 \\ \frac{1}{\gamma^{-1} - \log(\lambda)}, & \xi = 0. \end{cases} \tag{12}$$

That is the conditions in (10) are fullfilled.

Summarizing we have developed an IF model with only one spatially varying parameter, the threshold u_s and the other parameters ξ, γ, λ constant over the region. Note, that we choose λ to be constant in the first place and therefore, obtain a site-specific threshold. This is different from the model proposed by *Madsen and Rosbjerg* [1997], where u_s is a priori fixed and only the shape parameter ξ is constant over the region, whereas σ and λ vary over the region, which violates relationship (2). Moreover, the model is only an IF model for the excesses, whereas our model is an IF model for both the excesses and the exceedances.

We get the following GP model for the excesses:

$$P\left(Y_s \leq y | Y_s \geq 0\right) = G_{\xi, \gamma u_s}(y). \tag{13}$$

Now we can rewrite equation (1) for the $1/\alpha$ return level at site s, as

$$q_\alpha(s) = \begin{cases} u_s\Big(1 + \gamma \ln(\lambda/\alpha)\Big), & \xi = 0, \\ u_s\Big(1 - \frac{\gamma}{\xi}[1 - (\frac{\alpha}{\lambda})^{-\xi}]\Big), & \xi \neq 0. \end{cases} \tag{14}$$

As in equation (4), we see the factorization in a site specific index flood and a site independent general quantile function.

2.2. Nonstationary climate

There is no general theory for the estimation of extreme quantiles of nonstationary data. Approaches to account for long term trends in extremes are mostly ad hoc *Coles* [2001]. The classical way to incorporate this nonstationarity in the POT approach, is to keep the threshold constant and model the changing exceedance frequency by an inhomogeneous Poisson process and the excessses by a GP distribution with time dependent parameters [*Smith*, 1989b; *Coles*, 2001; *Yiou et al.*, 2006; *Bengtsson and Nilsson*, 2007].

We follow a different route, which circumvents the inhomogeneous Poisson process by considering a time dependent threshold, see e.g. *Coehlho et al.* [2008] and *Kyselý et al.* [2010]. A natural way to determine this varying threshold is quantile regression, which can be described as a way to identify the temporal evolution of a given quantile in a smooth parametric way, see e.g. *Koenker* [2005]; *Friederichs* [2010] and *Kyselý et al.* [2010]. Quantile regression is further discussed in section 3.1. When we take a time dependent high quantile, given by quantile regression, instead of a constant quantile, we can assume that λ is constant over space and time. The time dependent GP distribution is used to describe the excesses of the time varying threshold.

Hanel et al. [2009] generalize the IF assumption to the nonstationary block maxima setting. Following them we generalize (5) in a similar way, which means that, after scaling by a time dependent index flood, for every time point the site specific distribution functions are constant over the region, i.e. $\forall s \in \mathcal{S}, \forall t \in \mathcal{T}$

$$P\left(\frac{X_s(t)}{\eta_s(t)} \leq x | X_s(t) \geq u_s(t)\right) = \psi_t(x), \tag{15}$$

where ψ_t is independent of the site s. As in the stationary case we take the threshold as index flood:

$$\eta_s(t) = u_s(t).$$

Now we can generalize (9) in view of (15) to

$$\xi_s(t) \equiv \xi(t), \quad \frac{\sigma_s(t)}{u_s(t)} \equiv \gamma(t), \tag{16}$$

and equation (14) can be generalized to the non-stationary setting:

$$q_\alpha(s,t) = \begin{cases} u_s(t)\Big(1 - \frac{\gamma(t)}{\xi(t)}\big[1 - (\frac{\alpha}{\lambda})^{-\xi(t)}\big]\Big), & \xi(t) \neq 0, \\ u_s(t)\Big(1 + \gamma(t)\ln(\lambda/\alpha)\Big), & \xi(t) = 0. \end{cases} \tag{17}$$

As in the stationary case, we can see the factorization into a time and site dependent index flood and a quantile function, which depends on time only.

3. Estimation of the model parameters

We have chosen the threshold as a time dependent high quantile. For the estimation of this quantile we use quantile regression, which is outlined in section 3.1. Section 3.2 illustrates the composite likelihood framework for estimating the time-dependent parameters of the excess distribution.

3.1. Threshold estimation

Quantile regression relies on the fact that a sample quantile can be viewed as a solution of an optimization problem, which can be solved efficiently using linear programming, as shown in *Koenker and Bassett* [1978]. When we fix $s \in \mathcal{S}$, we can obtain the

τ-th sample quantile of the observations $x_s = (x_{s,1}, \dots, x_{s,T})$ at site s as

$$\underset{\beta \in \mathbb{R}}{\arg\min} \sum_{t=1}^{T} \rho_\tau(x_{s,t} - \beta), \tag{18}$$

where

$$\rho_\tau(v) = \begin{cases} v(\tau - 1), & v < 0, \\ v\tau, & v \geq 0. \end{cases}$$

In linear quantile regression it is assumed, that the τ-th conditional quantile function for given covariates z has a linear structure, i.e.

$$Q_{x_s}(\tau | z) = z^T \beta(\tau), \tag{19}$$

e.g. a linear trend in time would be given by

$$Q_{x_s}(\tau | t) = \beta_0(\tau) + t \cdot \beta_1(\tau).$$

In view of (18) *Koenker and Bassett* [1978] propose

$$\underset{\beta_0, \beta_1 \in \mathbb{R}}{\arg\min} \sum_{t=1}^{T} \rho_\tau(x_{s,t} - \beta_0 - t\beta_1)$$

as estimator for $\beta(\tau)$. For details of the transformation of this optimization problem into a linear program, see *Koenker* [2005].

Note, that λ was defined as the mean number of exceedances in a block. If the linear quantile function (19) holds, we have in fact the following relationship between τ and λ,

$$(1 - \tau) \cdot T / \#N_B = \lambda,$$

where $\#N_B$ is the number of blocks.

3.2. Excess distribution estimation

Maximum likelihood (ML) estimation is a common approach to estimate the parameters in a statistical model. The ML framework has attractive asymptotic properties.

Moreover, it is very flexible, e.g. it is convenient to incorporate covariates. For these reasons several authors recommend it for the estimation of extreme quantiles, especially when trends occur, see e.g. *Coles* [2001].

In order to apply the ML method, one needs the full likelihood function of the precipitation extremes, over all times and sites. Because of the spatial dependence, this requires the joint distribution of the excesses at all sites, which is difficult to describe because of the large dimensionality and estimation would be virtually impossible. One interesting way to proceed without the knowledge of the full dependence structure is to use a simplified likelihood. A class of such simplified likelihoods is summarized in the framework of composite likelihood, see e.g. *Varin et al.* [2011]. In this study we focus in particular on the independence likelihood, see also *Chandler and Bate* [2007]. The independence likelihood is the likelihood, as the name suggests, that would be obtained if the excesses at different sites were independent. We want to emphasize, that we focus on local quantiles and their spatial variation over a region, in which case the independence likelihood gives reasonable results, compare *Cooley et al.* [2007] and *Blanchet and Lehning* [2010]. This is, however, not the case if dependence parameters are of interest, as in *Padoan et al.* [2010], where a pairwise composite likelihood is used.

In the nonstationary IF model, the parameters γ and ξ of the excess distribution depend on time. We postulate a certain structure for these parameters, e.g.

$$\gamma(t) = \gamma_1 + \gamma_2 \cdot (t - \bar{t}), \qquad \xi(t) = \xi_1,$$

where $\bar{t}$ is the mean of the time points, so that γ_1 is the average of$\gamma(t)$ over t. Let $\theta = (\gamma_1, \gamma_2, \xi_1)$ be the vector of parameters, that has to be estimated. The independence likelihood is then given by:

$$\mathcal{L}_I(\theta, \mathbf{Y}) = \prod_{\substack{t\in\mathcal{T} \\ y_s(t)\geq 0}} \prod_{s\in\mathcal{S}} \frac{1}{\gamma(t)u_s(t)} \cdot \left[1 + \frac{\xi(t)y_s(t)}{\gamma(t)u_s(t)}\right]^{(-1/\xi(t)-1)},$$

where the condition on $y_s(t) \geq 0$ reflects that we only consider peaks over the threshold. Note, that by the choice of the quantile the threshold has been fixed beforehand.

The maximum independence likelihood estimator (MILE) is the parameter $\hat{\theta}_I$ which maximizes $\mathcal{L}_I(\theta, \mathbf{Y})$ or equivalently the independence log likelihood

$$\begin{aligned} \ell_I(\theta, \mathbf{Y}) = -\sum_{\substack{t\in\mathcal{T} \\ y_s(t)\geq 0}} \sum_{s\in\mathcal{S}} \Big[&\ln(\gamma(t)u_s(t)) + \\ &+ \frac{1+\xi(t)}{\xi(t)} \ln(1 + \frac{\xi(t)y_s(t)}{\gamma(t)u_s(t)}) \Big]. \end{aligned} \tag{20}$$

We have to optimize this function, with respect to the elements of θ. This can be done using the Broyden-Fletcher-Goldfarb-Shanno (BFGS) method as implemented in the `optim`-function of GNU R [*R Development Core Team*, 2011].

For testing the adequacy of the IF model, it is necessary to consider more general models with a spatially dependent dispersion coefficient, e.g. $\gamma_s(t) = \gamma_s$ and $\xi_s(t) = \xi$. The independence log likelihood for this model is obtained by replacing $\gamma(t)$ by γ_s and $\xi(t)$ by ξ in equation (20). The direct optimization of this likelihood with respect to the $(S+1)$ parameters is in the case of a large number of sites computationally very demanding. Therefore we exploit the structure of the independence likelihood by using a profile likelihood approach. In the example above we can split, for a given shape parameter, the optimization over an S-dimensional space into S optimization problems in one dimension, i.e. the maximization of the log likelihood for the excesses at site s with respect to ξ_s. This is usually much faster. If one does this on a grid of potential values for the shape parameter one can see the structure of the profile likelihood. Moreover we can construct a convergent procedure, leading to the estimator for the shape parameter. We recommend as initial value for this procedure the mean of the estimated shape parameters $\hat{\xi}_s$ of a site specific model. Another problem with the direct optimization might be the existence of local maxima in the likelihood; with the proposed approach we did not experience any problems with this issue.

The MILE $\hat{\theta}_I$ is asymptotically normal, see e.g. *Varin et al.* [2011]:

$$\sqrt{\#\widetilde{\mathcal{T}}}(\hat{\theta}_I - \theta) \xrightarrow{d} N\left(0, G^{-1}(\theta)\right),$$

where $\#\widetilde{\mathcal{T}}$ is the number of days with one or more threshold exceedances and G is the Godambe information:

$$G(\theta) = H(\theta)J^{-1}(\theta)H(\theta), \tag{21}$$

where $H(\theta)$ is minus the expected Hessian of ℓ_I at θ, also referred to as sensitivity matrix, and J is the variability matrix, i.e. the covariance matrix of the score $u(\theta, \mathbf{Y}) = \nabla_\theta \ell_I(\theta, \mathbf{Y})$. In the case of spatial independence, we have $H(\theta) = J(\theta)$ and the Godambe information reduces to the Fisher information, i.e. $G(\theta) = H(\theta)$. Here H is estimated as its observed value at $\hat{\theta}_I$, and J as

$$J = \frac{1}{\#\widetilde{\mathcal{T}}} \sum_{t \in \widetilde{\mathcal{T}}} u\left(\hat{\theta}_I, y(t)\right) u\left(\hat{\theta}_I, y(t)\right)',$$

where $y(t) = (y_1(t), \ldots, y_S(t))'$ and $u(\hat{\theta}_I, y(t))$ is the contribution of day t to $u(\hat{\theta}, \mathbf{Y})$. The latter estimate makes use of the fact that the excesses on different days are independent, see e.g. *Chandler and Bate* [2007] and *Varin et al.* [2011]. An estimate of the Godambe information $\hat{G}(\theta)$ is obtained by plugging in the estimates $\hat{H}$ and $\hat{J}$ in equation (21). This estimate $\hat{G}$ is used to assess the uncertainty of the parameters (and quantiles) of the excess distribution, see section 4

4. Model selection for the excess distribution

In this section we describe the methods used, to investigate the temporal behavior of the dispersion coefficient and the shape parameter as well as the adequacy of the IF model.

Information criteria are used as indication of the suitability of a specific model. *Varin et al.* [2011] present composite likelihood adaptations of the Akaike information criterion (AIC) and the Bayesian information criterion (BIC), which are defined in the usual way

$$AIC = -2\ell_I(\hat{\theta}_I, Y) + 2\dim(\theta),$$

$$BIC = -2\ell_I(\hat{\theta}_I, Y) + \log(\#\widetilde{\mathcal{T}})\dim(\theta),$$

where $\dim(\theta)$ is an effective number of parameters, which can be estimated as

$$\dim(\theta) = \mathrm{tr}\left(H(\theta)G(\theta)^{-1}\right).$$

Moreover, we will test our assumptions using nested models. This means, that we consider subsets M_0 of the full model M_1 by constraining q components of the parameter vector θ. For instance we may partition $\theta = (\psi, \phi)$ such that the q-dimensional component ψ is zero under M_0. To test this hypothesis, we use the independent likelihood ratio statistic, which is a special case of a composite likelihood ratio (CLR) statistic [*Chandler and Bate*, 2007; *Varin et al.*, 2011]:

$$W = 2\Big[\ell_I(\hat{\theta}_{M_1}; y) - \ell_I(\hat{\theta}_{M_0}; y)\Big], \tag{22}$$

where $\hat{\theta}_{M_1}$ ($\hat{\theta}_{M_0}$) denotes the MILE of model M_1 (M_0). *Varin et al.* [2011] present the following asymptotic result for W under the null hypothesis

$$W \xrightarrow{d} \sum_{j=1}^{q} \lambda_j Z_j^2, \tag{23}$$

where the Z_j are independent, standard normal variates and $\lambda_1, \ldots, \lambda_q$ are the eigenvalues of

$$(G_{M_1}^{-1})_\psi \left((H_{M_1}^{-1})_\psi \right)^{-1}.$$

Here $(G_{M_1}^{-1})_\psi$ denotes the submatrix of the inverse Godambe information for the full model M_1 pertaining to the parameter vector ψ and $(H_{M_1}^{-1})_\psi$ is defined analogously.

In order to obtain the information criteria and the asymptotic distribution of W under the null hypothesis, we need to estimate the Godambe information, which is difficult when the number of parameters is large. Hence it is not feasible to examine the appropriateness of the IF assumption for regions with many sites, based on the Godambe information.

One possibility to obtain p-values for the test statistic W, without estimating the Godambe information, is to apply a bootstrap procedure, see e.g. *Varin et al.* [2011]. We follow *Hanel et al.* [2009] and use a semiparametric bootstrap approach, to take the dependence structure into account, without explicitly modeling this. The challenge is to produce bootstrap samples according to the null hypothesis, which exhibit approximately the same spatial dependence structure as the original data set. We assume that the underlying spatial dependence is not changing over time, i.e. only the marginal distributions are changing. One could think of a constant copula generating the dependence structure, that is for fixed t

$$P(Y_s(t) \leq y_{s,t} \forall s) = C\big(G_{1,t}(y_{1,t}), \ldots, G_{S,t}(y_{S,t})\big),$$

where $G_{s,t} = G_{\sigma_s(t),\xi_s(t)}$, and C a copula, for details on copula see e.g. *Nelson* [2006]. We generate the bootstrap samples in three steps. In the first step we transform the sample of the excesses $Y_s(t)$ into a sample that follows approximately the standard exponential distribution

$$Z_s(t) = \begin{cases} \frac{1}{\hat{\xi}_s^1(t)} \ln\big(1 + \frac{\hat{\xi}_s^1(t) Y_s(t)}{\hat{\sigma}_s^1(t)}\big), & \hat{\xi}_s^1(t) \neq 0, \\ \frac{Y_s(t)}{\hat{\sigma}_s^1(t)}, & \hat{\xi}_s^1(t) = 0, \end{cases} \tag{24}$$

where $\hat{\sigma}_s^1(t)$ and $\hat{\xi}_s^1(t)$ are the estimated scale and shape parameters under the full model M_1. In the second step, we sample with replacement monthly blocks of the whole spatial domain from $Z_s(t)$ to obtain a new sample $\widetilde{Z}_s(t)$ with approximately standard exponential margins and the same spatial dependence structure as that of $Z_s(t)$. In the third step we use the estimated scale and shape parameter under the null hypothesis, denoted as $\hat{\sigma}_s^0(t)$ and $\hat{\xi}_s^0(t)$, respectively, to transform the sample $\widetilde{Z}_s(t)$ to a bootstrap sample of the excesses

$$\widetilde{Y}_s(t) = \begin{cases} \frac{\hat{\sigma}_s^0(t)}{\hat{\xi}_s^0(t)} \big[\exp\big(\hat{\xi}_s^0(t) \widetilde{Z}_s(t)\big) - 1\big], & \hat{\xi}_s^0(t) \neq 0, \\ \widetilde{Z}_s(t) \hat{\sigma}_s^0(t), & \hat{\xi}_s^0(t) = 0. \end{cases} \tag{25}$$

The $\widetilde{Y}_s(t)$ follow approximately the GP model M_0 and mimic the spatial dependence structure of the original excesses.

From a number of Monte Carlo experiments, *Kyselý* [2007, 2009] concluded that the (non-parametric) bootstrap generally resulted in too narrow confidence intervals for large quantiles of the distributions, that are commonly used to describe the distribution of precipitation extremes. This has been attributed to the skewness of the estimators of the model parameters in the case of small and moderate sample sizes. This objection might be weakened, when using RFA methods, because then the estimation is based on much more data.

5. Application to precipitation data

We applied the regional peaks-over-threshold method to observed precipitation data from the Netherlands. We used the daily, gridded E-OBS data (version 5.0), which were made available by the European funded project ENSEMBLES [*Haylock et al.*, 2008]. We consider winter (DJF) precipitation for 25 km×25 km grid squares centered in the Netherlands, for the period December 1, 1950 to February 28, 2010. In total we have 69 grid boxes and 60 winter seasons of daily measurements for each grid box.

The Netherlands has a maritime climate with relatively mild and humid winters. Figure 1 shows the mean over the considered period of the largest daily precipitation value in winter (winter maximum) for each grid box. The spatial variation in Figure 1 is small, 80% of the values lie between 18.2 and 20.4 mm. Previous studies propose to view the Netherlands as a homogeneous region for which the IF assumption applies, see e.g. *Overeem el al.* [2008] and *Hanel et al.* [2009].

Daily precipitation in the winter season exhibits some temporal dependence, also at high levels. The relation between the GEV and GP distribution parameters (Equation (2)) relies on the independence assumption as does the estimation of the variability matrix J, therefore, it is necessary to select a subset of independent events. This is usually achieved by specifying a minimum separation time between exceedances over the threshold [e.g. *Kyselý et al.*, 2010].

We decluster the original data rather than the exceedances, i.e. we look at blocks of length one plus the minimum separation time and replace all but the maximum values of these blocks by zero and determine the threshold for these declustered data, as described in section 3.1. It follows that the exceedances are declustered with the same minimum separation time. The advantage of this procedure over declustering the exceedances directly, is that the expected number of exceedances per block will be approximately constant over the region, which is a basic assumption of our model. As the persistence of rain events is rather short, we specify the minimum separation time to be one day.

We choose the threshold to be the 96% linear regression quantile. Hence, we expect on average 3.61 exceedances per grid box and winter season. Figure 2 shows for each grid box the mean of this threshold for the 1950–2010 period. The trend in the threshold for the 1950–2010 period is positive over the whole domain, see Figure 3, but is relatively small in the southeastern part of the country and large (up to 40%) in the west and northern parts. *Buishand et al.* [2012] found a significant positive trend in the mean precipitation for the winter half year (October – March) in the Netherlands during the period 1951 – 2009. A clear spatial gradient was, however, not observed in the trend of the mean winter precipitation.

We test the hypothesis that the event times come from a Poisson process individually for each grid box by the dispersion index (DI) test. The DI test exploits the fact, that the variance and the mean of the Poisson distribution are the same, see *Cunnane* [1979] for details. The Poisson assumption is rejected at the 5% significance level in two of the 69 grid boxes, which is in good agreement with the expected number of rejected grid boxes under the Poisson assumption. If the exceedance times come from a homogeneous Poisson process, these should be distributed uniformly on any time interval, see e.g. *Cox and Lewis* [1966]. The Kologorov-Smirnov test does not reject uniformity in any grid box.

We exploit the fact that event times, coming from a Poisson process, are uniformly distributed over time. The resulting p-values of a Kolmogorov-Smirnov test on the uniformity are shown in Figure 4.

We consider four different models for the excess distribution, three based on the IF approach, **A** – **A″** in Table 1, and one with a spatially varying dispersion coefficient and constant shape parameter, model **B**.

In a first step we want to infer which of the IF models is the best to describe the data. For a first indication the information criteria are computed, as outlined in section 4, for each of the three IF models, see Table 2. We see from both the composite AIC and the composite BIC, that the incorporation of a trend in the dispersion coefficient γ (model **A′**) does not result in a better model. One can see on the other hand, that according to the AIC model **A″**, which has a (linear) trend in the shape parameter, is selected. That contrasts with the selection of the simplest model **A**, by means of the BIC.

The shape parameter is crucial for the estimation of very high quantiles. Model **A** estimates the shape parameter to be 0.03, i.e. just in the Fréchet domain. Model **A″** estimates a large drop in the shape parameter from 0.10 to -0.09, which would mean a change from the Fréchet family to the Weibull family. In order to gain more insight in the temporal behavior of the shape parameter, we compute the shape parameter for overlapping 20 year subsamples of the data, using model **A**, which has no trend in the model parameters. It appears that a large part of the negative trend in the shape parameter in model **A″** is due to one specific event, namely the extreme rainfall of

December 3, 1960, compare also *Buishand* [1984] and *Van den Brink and Können* [2011], resulting in a large drop of the 20 year window estimates in the year 1971, as observed in Figure 5.

The quantile estimates, obtained from model **A**, are increasing due to the positive trend in the threshold. Because of the connection of the scale parameter with the threshold, see equation (16), the positive trend in the threshold leads to a positive linear trend in the scale parameter. In contrast to the previous model, we obtain from model **A''**, quantile estimates, that exhibit a phase transition. While the 2 year return level is still increasing due to the positive trend in the threshold, we have that the 25 year return level is decreasing due to the negative trend in the common shape parameter. The 5 year return level is approximately constant, see Figure 6. An interpretation of this is much more complex, than for the quantile estimates, stemming from model **A**.

When we carry out the composite likelihood ratio test, it turns out, that neither the trend in the dispersion coefficient nor the trend in the shape parameter are significant, although the p-values are quite different for these trends, see Table 3. We can also see from Table 3, that the bootstrap procedure gives similar results as the use of the asymptotic result in equation (23).

In the second step we want to test the IF assumption. Therefore we compute the composite likelihood ratio test for the full model **B** and the nested model **A**. As earlier explained, we can not estimate the Godambe information well for model **B**. Hence, we proceed only with the bootstrap procedure. We obtain an p-value of 0.103 for 2500 bootstrap samples. This means, that the IF assumption does not have to be rejected. Note, because of the large difference in the number of parameters between model **B** and model **A**, the composite likelihood ratio test will not have much power due to the great number of alternatives. This can be considered as an intrinsic problem, when comparing regional models with site dependent parameters.

Figure 7 compares for a particular site the estimated return levels of the excess distribution based on the site specific approach with those obtained from the IF assumption. Pointwise confidence bands for the return levels based on the asymptotic normality of the MILE are also given. The quantile estimates for the two methods are quite similar, but the IF approach reduces the uncertainty in the estimation to half the uncertainty of the site specific approach.

6. Conclusions

An index flood approach for nonstationary peaks-over-threshold data has been developed. The threshold is chosen to be a large quantile that varies over time, which is also taken as the index flood. The peaks exceeding the threshold are described by Generalized Pareto distributions. The index flood assumption implies that the ratio of the scale parameter to the threshold and the shape parameter are constant over the region but may vary over time.

The approach was applied to gridded, observed daily precipitation data from the Netherlands for the winter season. A linear increase in the threshold was found, which was most pronounced in the western and northern parts of the country. This increase in the threshold leads to an increase in the scale parameter, because of the index flood assumption. No evidence was found for a change in the ratio of the scale parameter to the threshold. Though a large negative trend in the shape parameter was observed, this trend turned out to be mainly due to one exceptional event. Therefore, the extreme quantiles increase in the same way as the threshold.

Although the uncertainty in the estimation of the excess distribution was cut by half compared to a site specific estimation procedure, the remaining uncertainty is still substantial. The uncertainty could be possibly further reduced by considering longer records or by extending the region. For instance, one could think of including the neighboring part of North Germany in the analysis. The different trends in the index flood indicate, however, that one should be very careful with extending the region. Apart from analyzing more data, the estimation uncertainty might also be reduced by maximizing a pairwise likelihood that partly accounts for spatial dependence rather than the independent likelihood.

The validity of the bootstrap might be questionable and should be assessed by a Monte Carlo experiment, which includes the spatial dependence. However, this is for peaks-over-threshold data much more computational demanding than for block maxima.

Acknowledgments. The research was supported by the Dutch research program Knowledge for Climate. We acknowledge the E-OBS dataset from the EU-FP6 project ENSEMBLES (`http://ensembles-eu.metoffice.com`) and the data providers in the ECA&D project (`http://eca.knmi.nl`). All calculations were performed using the R environment (`http://www.r-project.org`).

References

Beguería, S., M. Angulo-Martínez, S. M. Vicente-Serrano, J. I. López-Moreno, and A. El-Kenawy (2011), Assessing trends in extreme precipitation events intensity and magnitude using non-stationary peaks-over-threshold analysis: a case study in North-east Spain from 1930 to 2006. *International Journal of Climatology*, 31(14): 2102–2114, doi:10.1002/joc.2218.

Bengtsson, A. and C. Nilsson (2007), Extreme value modelling of storm damage in Swedish forests. *Nat. Hazards Earth Syst. Sci.* , 7:515–521.

Blanchet, J. and M. Lehning (2010), Mapping snow depth return levels: smooth spatial modeling versus station interpolation. *Hydrology and Earth System Sciences*, 14(12):2527–2544, doi:10.5194/hess-14-2527-2010.

Brown, S. J., J. Caesar, and C. A. T. Ferro (2008), Global changes in extreme daily temperature since 1950. *J. Geophys. Res.*, 113(D5):D05115, doi:10.1029/2006JD008091.

Buishand, T. A. (1984), Zware neerslag in het winterhalfjaar. *Cultuurtechnisch Tijdschrift*, 23(5). In Dutch.

Buishand, T. A. (1989), Statistics of extremes in climatology. *Statistica Neerlandica*, 43(1):1–30, doi:10.1111/j.1467-9574.1989.tb01244.x.

Buishand, T. A., G. De Martino, H. Spreeuw, and T. Brandsma (2012), Homogeneity of precipitation series in the Netherlands and their trends in the past century. *International Journal of Climatology*, Accepted.

Chandler, R. E. and S. Bate (2007), Inference for clustered data using the independence loglikelihood. *Biometrika*, 94(1):167–183, doi:10.1093/biomet/asm015.

Coelho, C. A. S., C. A. T. Ferro, D. B. Stephenson, and D. J. Steinskog (2008), Methods for exploring spatial and temporal variability of extreme events in climate data. *Journal of Climate*, 21(10):2072–2092, doi:10.1175/2007JCLI1781.1.

Coles, S. (2001), *An Introduction to Statistical Modeling of Extreme Values*. Springer, London.

Cooley, D., D. Nychka, and P. Naveau (2007), Bayesian spatial modeling of extreme precipitation return levels. *Journal of the American Statistical Association*, 102(479):824–840, doi:10.1198/016214506000000780.

Cox, D. R. and P. A. W. Lewis (1966), *The Statistical Analysis of Series of Events*. Methuen, London.

Cunnane, C. (1979), A note on the Poisson assumption in partial duration series models. *Water Resour. Res.*, 15(2):489–494, doi:10.1029/WR015i002p00489.

Cunnane, C. (1988), Methods and merits of regional flood frequency analysis. *Journal of Hydrology*, 100(1-3):269 – 290, doi:10.1016/0022-1694(88)90188-6.

Dalrymple, T. (1960), Flood-frequency analyses, manual of hydrology: Part 3. *USGS Water Supply Paper*, 1543-A.

Embrechts, P., C. Klüppelberg, and T. Mikosch (1997), *Modelling Extremal Events.* Springer, Berlin.

Fowler, H. J., M. Ekström, C.G. Kilsby, and P.D. Jones (2005), New estimates of future changes in extreme rainfall across the UK using regional climate model integrations. 1. assessment of control climate. *Journal of Hydrology*, 300(1-4):212 – 233, doi:10.1016/j.jhydrol.2004.06.017.

Friederichs, P. (2010), Statistical downscaling of extreme precipitation events using extreme value theory. *Extremes*, 13:109–132, doi:10.1007/s10687-010-0107-5.

Hanel, M., T. A. Buishand, and C. A. T. Ferro (2009), A nonstationary index flood model for precipitation extremes in transient regional climate model simulations. *J. Geophys. Res.*, 114(D15):D15107, doi:10.1029/2009JD011712.

Haylock, M. R., N. Hofstra, A. M. G. Klein Tank, E. J. Klok, P. D. Jones, and M. New (2008), A European daily high-resolution gridded data set of surface temperature and precipitation for 1950-2006. *J. Geophys. Res.*, 113(D20):D20119, doi:10.1029/2008JD010201.

Hosking, J. R. M., and J. R. Wallis (1997), *Regional Frequency Analysis*. Cambridge University Press, Cambridge, UK.

Kharin, V. V. and F.W. Zwiers (2005), Estimating extremes in transient climate change simulations. *Journal of Climate*, 18(8):1156–1173, doi:10.1175/JCLI3320.1.

Klein Tank, A. M. G., and G. P. Können (2003), Trends in indices of daily temperature and precipitation extremes in Europe, 1946-99. *Journal of Climate*, 16(22):3665–3680, doi:10.1175/1520-0442(2003)016<3665:TIIODT>2.0.CO;2.

Koenker, R. (2005), *Quantile Regression*. Cambridge University Press, Cambridge, UK.

Koenker, R. and G. Bassett Jr. (1978), Regression quantiles. *Econometrica*, 46(1):33–50.

Kyselý, J. (2007), A cautionary note on the use of nonparametric bootstrap for estimating uncertainties in extreme-value models. *Journal of Applied Meteorology*, 47:3236–3251, doi:10.1175/2008JAMC1763.1.

Kyselý, J. (2009), Coverage probability of bootstrap confidence intervals in heavy-tailed frequency models, with application to precipitation data. *Theor Appl Climatol*, 101:345–361, doi:10.1007/s00704-009-0190-1.

Kyselý, J., J. Picek, and R. Beranová (2010) Estimating extremes in climate change simulations using the peaks-over-threshold method with a non-stationary threshold. *Global and Planetary Change*, 72(1-2):55 – 68, doi:10.1002/joc.1874.

Madsen, H., P. F. Rasmussen, and D. Rosbjerg (1997a), Comparison of annual maximum series and partial duration series methods for modeling extreme hydrologic events: 1. at-site modeling. *Water Resour. Res.*, 33(4):747–757, doi:10.1029/96WR03848.

Madsen, H., C. P. Pearson, and D. Rosbjerg (1997b), Comparison of annual maximum series and partial duration series methods for modeling extreme hydrologic events: 2. regional modeling. *Water Resour. Res.*, 33(4):759–769, doi:10.1029/96WR03849.

Madsen, H. and D. Rosbjerg (1997), The partial duration series method in regional index-flood modeling. *Water Resour. Res.*, 33(4):737–746, doi:10.1029/96WR03847.

Milly, P. C. D., J. Betancourt, M. Falkenmark, R.M. Hirsch, Z.W. Kundzewicz, D.P. Lettenmaier, and R.J. Stouffer (2008), Stationarity is dead: whither water management? *Science*, 319(5863):573–574, doi:10.1126/science.1151915.

Nelsen, R. B. (2006), *An Introduction to Copulas.* Springer, New York.

Overeem, A., T. A. Buishand,and I. Holleman (2008), Rainfall depth-duration-frequency curves and their uncertainties *Journal of Hydrology*, 348:124–134, doi:10.1016/j.jhydrol.2007.09.044

Padoan, S. A., M. Ribatet, and S. A. Sisson (2010), Likelihood-based inference for max-stable processes. *Journal of the American Statistical Association*, 105(489):263–277, doi:10.1198/jasa.2009.tm08577.

R Development Core Team (2011), R: A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing*, Vienna, Austria.

Reiss, R. D. and M. Thomas (2007), *Statistical Analysis of Extreme Values: with Applications to Insurance, Finance, Hydrology and Other Fields*. Birkhäuser, Basel, 3rd edition.

Robinson, J. S. and M. Sivapalan (1997), An investigation into the physical causes of scaling and heterogeneity of regional flood frequency. *Water Resour. Res.*, 33(5):1045–1059, doi:10.1029/97WR00044.

Rosbjerg, D., H. Madsen, and P. F. Rasmussen (1992), Prediction in partial duration series with generalized Pareto-distributed exceedances. *Water Resour. Res.*, 28(11):3001–3010, doi:10.1029/92WR01750.

Smith, J. A. (1989a), Regional flood frequency analysis using extreme order statistics of the annual peak record *Water Resour. Res.*, 25(2):311–317, doi:10.1029/WR025i002p00311.

Smith, R. L. (1986), Extreme value theory based on the r largest annual events. *Journal of Hydrology*, 86(1-2):27 – 43, doi:10.1016/0022-1694(86)90004-1.

Smith, R. L. (1989b), Extreme value analysis of environmental time series: An application to trend detection in ground-level ozone. *Statistical Science*, 4(4):367–377.

Turco, M. and M. C. Llasat (2011), Trends in indices of daily precipitation extremes in Catalonia (NE Spain), 1951 - 2003. *Natural Hazards and Earth System Science*, 11(12):3213–3226, doi:10.5194/nhess-11-3213-2011.

Van den Brink, H. W. and G. P. Können (2011), Estimating 10000-year return values from short time series. *International Journal of Climatology*, 31(1):115–126, doi:10.1002/joc.2047.

Varin, C., N. Reid, and D. Firth (2011), An overview of composite likelihood methods. *Statistica Sinica*, 21:5–42.

Wang, Q.J. (1991), The POT model described by the generalized Pareto distribution with Poisson arrival rate. *Journal of Hydrology*, 129(1–4):263 – 280, doi:10.1016/0022-1694(91)90054-L.

Wigley, T. (2009), The effect of changing climate on the frequency of absolute extreme events. *Climatic Change*, 97:67–76, doi:10.1007/s10584-009-9654-7.

Yiou, P., P. Ribereau, P. Naveau, M. Nogaj, and R. Brázdil (2006), Statistical analysis of floods in Bohemia (Czech Republic) since 1825. *Hydrological Sciences Journal* 51(5):930–945, doi:10.1623/hysj.51.5.930

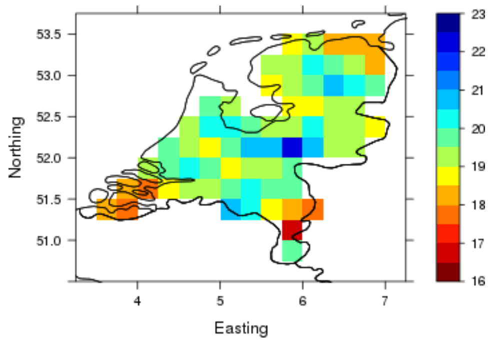

Figure 1. Mean of the winter maxima in mm

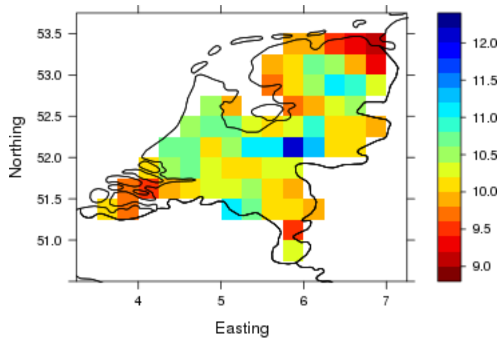

Figure 2. Mean of the threshold for the 1950–2010 period in mm

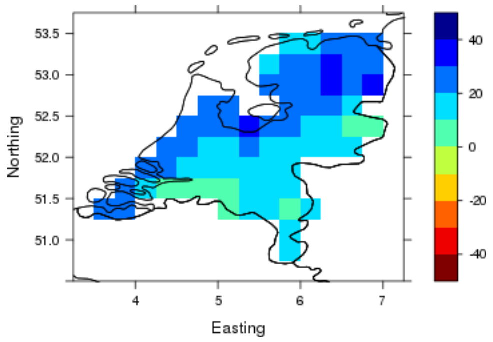

Figure 3. Trend in the threshold for the 1950–2010 period in %. The trend was defined as the difference between the last and the first value of the threshold divided by the mean value of the threshold.

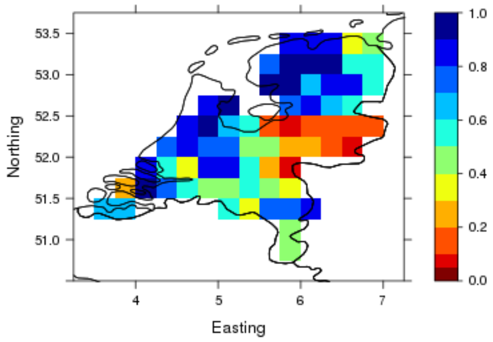

Figure 4. p-values of the uniformity for the event times

Table 1. Overview of models used

Model	**dispersion**	**shape**
A	γ	ξ
A′	$\gamma_1 + \gamma_2 * (t - \bar{t})$	ξ
A″	γ	$\xi_1 + \xi_2 * (t - \bar{t})$
B	$\gamma_1, \ldots, \gamma_S$	ξ

Table 2. Information criteria for the IF models[a].

Model	AIC	BIC
A	78387.28	**78715.59**
A′	78435.60	78880.41
A″	**78333.28**	78748.95

[a]The lowest AIC and BIC values are printed in bold.

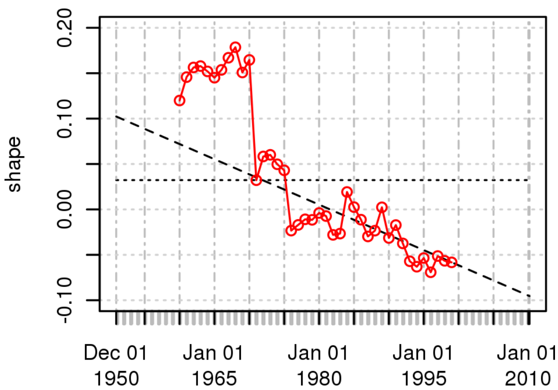

Figure 5. Evolution of the shape parameter over time (dotted - model **A**, dashed - model **A″**, solid red line with points - 20 year window estimates for model **A**

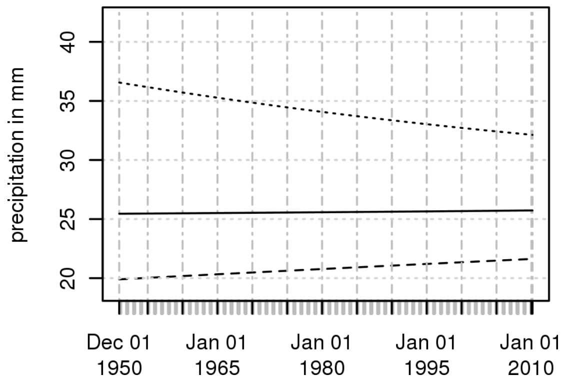

Figure 6. Trends of different return levels of daily precipitation for model **A″** (dashed – 2 year, solid – 5 year, dotted – 25 year)

Table 3. p-values of the CLR-test against model **A** (2500 samples)

Model	asymptotic	bootstrap
A′	82.9%	81.3%
A″	26.7%	12.2%

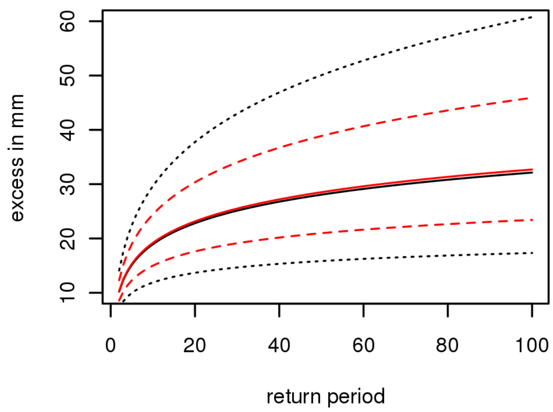

Figure 7. Estimated return levels of the excesses with 95% pointwise confidence bands for the year 1980 at the grid box around De Bilt (black – site specific, red – IF)