Contents lists available at ScienceDirect

Weather and Climate Extremes

Projections of precipitation extremes based on a regional, non-stationary peaks-over-threshold approach: A case study for the Netherlands and north-western Germany

M. Roth [a,b,*], T.A. Buishand [a], G. Jongbloed [b,c], A.M.G. Klein Tank [a], J.H. van Zanten [b,d]

[a] Royal Netherlands Meteorological Institute (KNMI), Utrechtseweg 297, NL-3731 GA, De Bilt, The Netherlands
[b] EURANDOM, Eindhoven University of Technology, Den Dolech 2, 5612 AZ, Eindhoven, The Netherlands
[c] Delft Institute of Applied Mathematics, Delft University of Technology, Mekelweg 4, 2628 CD Delft, The Netherlands
[d] Korteweg-de Vries Institute for Mathematics, University of Amsterdam, 1098 XH Amsterdam, The Netherlands

ARTICLE INFO

ABSTRACT

Projections of extreme precipitation are of great importance, considering the potential severe impacts on society. In this study a recently developed regional, non-stationary peaks-over-threshold approach is applied to two transient simulations of the RACMO2 regional climate model for the period 1950–2100, driven by two different general circulation models. The regional approach reduces the estimation uncertainty compared to at-site approaches. The selection of a threshold for the peaks-over-threshold model is tackled from a new perspective, taking advantage of the regional setting. Further, a regional quantile regression model using a common relative trend in the threshold is introduced. A considerable bias in the extreme return levels is found with respect to gridded observations. This bias is corrected for by adjusting the parameters in the peaks-over-threshold model.

In summer a significant increase in extreme precipitation over the study area is found for both RACMO2 simulations, mainly as a result of an increase of the variability of the excesses over the threshold. However, the magnitude of this trend in extreme summer precipitation depends on the driving general circulation model. In winter an increase in extreme precipitation corresponding with an increase in mean precipitation is found for both simulations. This trend is due to an increase of the threshold.

© 2014 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-SA license (http://creativecommons.org/licenses/by-nc-sa/3.0/).

1. Introduction

Information on extreme precipitation is crucial for various societal sectors, e.g. for the design of sewage and drainage systems, roads and tunnels, farming, and the insurance industry. Consensus is growing that the characteristics of extreme precipitation may alter owing to climate change. In order to project the change in extreme precipitation, climate model data have been analyzed and compared to observations. Often this evaluation is done in terms of index values, such as the empirical annual 90% quantile of the precipitation amounts for each year or the 1-day or 5-day maximum precipitation amount in a year, see e.g. Klein Tank and Können (2003), and Turco and Llasat (2011). This approach shows the evolution of precipitation extremes over time. However, the indices have mostly a return period of not more than 1 year,

which is of minor importance for the planning of hydraulic infrastructure, that usually has to withstand events with much longer return periods. To estimate the changes in these rare events, extreme-value distributions have been fitted to the extremes for two subsets of the data representing current (e.g. 1980–2010) and future (e.g. 2070–2100) climate, assuming stationarity within the time slices, see e.g. Fowler et al. (2005), Ekström et al. (2005), and Kyselý and Beranová (2009). Considering only two time slices does not give a picture of the evolution of the extremes, which is e.g. necessary if one is interested in the risk of failure of a hydraulic structure during its expected lifetime. Moreover, the selection of the time slices introduces additional uncertainty. A small shift of the time slices may have large influence on the estimated change. As an alternative, extreme value distributions with time-dependent parameters, which allow the consideration of the full time period, have been used, see e.g. Coles (2001), El Adlouni et al. (2007), Sugahara et al. (2009), Kyselý et al. (2010), Beguería et al. (2011), and Tramblay et al. (2013).

The estimation of changes in rare extremes is subject to large uncertainty. A general way to reduce the estimation uncertainty is regional frequency analysis (RFA), where the similarities between

* Corresponding author at: Royal Netherlands Meteorological Institute (KNMI), Utrechtseweg 297, NL-3731 GA, De Bilt, The Netherlands. Tel.: +31 302206382.
E-mail addresses: m.roth@tue.nl (M. Roth), buishand@knmi.nl (T.A. Buishand), G.Jongbloed@tudelft.nl (G. Jongbloed), kleintan@knmi.nl (A.M.G. Klein Tank), hvzanten@uva.nl (J.H. van Zanten).

different sites in a region are exploited (Hosking and Wallis, 1997). RFA is mostly applied to (annual) block maxima (BM). An alternative to BM is to consider all peaks over a (high) threshold (POT), which is often preferable, owing to the more efficient use of the data.

A regional peaks-over-threshold model, combining the RFA approach and POT data, which can be used to analyze precipitation extremes in a changing climate, was developed by Roth et al. (2012). In this model a temporally varying threshold, which is determined by quantile regression, is used to account for changes in the frequency of precipitation extremes. The marginal distributions of the excesses are described by generalized Pareto distributions (GPD), with parameters, that may vary over time and their spatial variation is modeled by the index flood (IF) approach. For a detailed introduction to the index flood method, see Hosking and Wallis (1997).

The selection of the threshold is a crucial step in the application of the POT approach. However, there is still no standard procedure for this, and usually one relies on visual tools. Among these the plotting of the change of the estimated GPD shape parameter or the mean excess of the exceedances against the height of the threshold is popular. Unfortunately these plots rarely give clear indications of which quantile should be used for the threshold. In the present study, the individual plots are averaged over the region in order to make the desired constant or linear structure more apparent.

Daily precipitation from two simulations of a regional climate model (RCM) driven by different general circulation models (GCM) and from gridded observational data is analyzed for the Netherlands and north-western Germany. Instead of linking the POT model parameters to time, a temperature-based covariate is used to include the evolution of climate, see also Hanel et al. (2009), and Van Oldenborgh et al. (2009). Bias correction of the return levels from the regional climate model simulations is discussed. In addition to the changes in return levels, we present a risk-based design level that was recently introduced by Rootzén and Katz (2013).

Section 2 outlines the methods and Section 3 introduces the data used. Results and discussion are given in Section 4, followed by the conclusion in Section 5.

2. Methods

2.1. Introduction to the peaks-over-threshold model

To study the extremes of independent and identically distributed random variables X_i, one can consider the excesses $Y_i = X_i - u$ over a (high) threshold u. The *Balkema, De Haan, and Pickands* theorem states that the distribution of the excesses Y, conditioned on $Y \geq 0$, can be approximated by a generalized Pareto distribution (GPD), if the threshold u is sufficiently high and certain regularity conditions hold, see e.g. Reiss and Thomas (2007):

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

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 respectively. For $\xi = 0$ the GPD reduces to the exponential distribution. The independence requirement can be weakened (e.g. Leadbetter et al., 1983). In the case of short-range dependence the GPD approximation applies if one considers declustered excesses, i.e. the excesses of the local maxima (peaks) in a cluster of exceedances only. Several studies have considered the GPD also for non-stationary data, using temporally varying parameters, see for recent examples Sugahara et al. (2009), Kyselý et al. (2010), and Beguería et al. (2011).

2.2. Temporal dependence and declustering

Daily precipitation exhibits temporal dependence, also at high levels. This dependence is generally stronger in winter than in summer (due to the convective nature of most extremes in summer). As mentioned in Section 2.1 one can account for this temporal dependence by considering declustered excesses. This is usually achieved by specifying a minimum separation time t_{sep} between exceedances over the threshold, where t_{sep} is determined by the temporal dependence in the data at high levels. Here, we follow this approach but decluster all data and not the excesses only.

Let $x_s(t)$ be the rainfall at site $s \in \{1, ..., S\}$ and day $t \in \{1, ..., T\}$. To determine t_{sep} we compute first the 95% quantile for each site s and calculate the number of clusters of length $n \geq 2$. A cluster of length n is defined as n consecutive exceedances of the 95% quantile. The number of clusters decreases usually very fast with the length n. The separation time t_{sep} is set to n, if the number of clusters of length $n+2$ is sufficiently low. After this initial step we obtain the declustered data by replacing $x_s(t)$ with zero, if it is not a maximum in the subset $x_s(t - t_{sep}), ..., x_s(t), ..., x_s(t + t_{sep})$. From this it is clear that also the excesses obtained from the declustered data are separated by at least t_{sep} days.

2.3. Index flood approach

Roth et al. (2012) introduced a regional approach for multi-site, non-stationary POT rainfall data:

$$y_s(t) = x_s(t) - u_s(t),$$

where $u_s(t)$ is a suitable threshold value for site s and day t. The approach is based on the index flood (IF) assumption, i.e. that the non-stationary POT data have, after scaling by a time and site dependent index variable (or index rainfall) $\eta_s(t)$, a common excess distribution. If the site-specific excess distributions are GPD with shape parameter $\xi_s(t)$ and scale parameter $\sigma_s(t)$, then we have for the scaled excesses:

$$P\left(\frac{Y_s(t)}{\eta_s(t)} \leq y \middle| Y_s(t) > 0\right) = G_{\xi(t),\gamma(t)}(y), \tag{1}$$

with $Y_s(t)$ the excess at site s and day t and $\gamma(t) = \sigma_s(t)/\eta_s(t)$ a dimensionless dispersion coefficient. The IF assumption thus implies that this coefficient and the shape parameter are constant over the region of interest. Roth et al. (2012) used the threshold $u_s(t)$ as index variable:

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

The mean number λ of the excesses over $u_s(t)$ in this approach is constant over time and space, which was achieved by using quantile regression to determine $u_s(t)$. With Eq. (1) we can compute for each site s and day t the value $r_{s,t}(\alpha)$ that is exceeded on average α times in a season:

$$r_{s,t}(\alpha) = \begin{cases} u_s(t)\left(1 - \frac{\gamma(t)}{\xi(t)}\left[1 - \left(\frac{\lambda}{\alpha}\right)^{\xi(t)}\right]\right), & \xi(t) \neq 0, \\ u_s(t)(1 + \gamma(t)\ln(\lambda/\alpha)), & \xi(t) = 0. \end{cases} \tag{2}$$

In analogy with a stationary setting, the quantity $r_{s,t}(\alpha)$ is termed the $1/\alpha$-year return level, although $1/\alpha$ no longer gives the expected waiting time between exceedances of $r_{s,t}(\alpha)$.

2.4. Determination of the threshold

The non-stationary threshold is estimated via quantile regression. However, we have to select an appropriate quantile, i.e. the value of the threshold has to be high enough to justify the GPD assumption.

2.4.1. Quantile selection

The threshold choice (TC) plot and the mean excess (ME) plot (also referred to as mean residual life plot) are widely used graphical tools for the selection of the threshold in the POT analysis. The TC plot is based on the fact that once the GPD model holds at grid point s for some threshold u_s^0 it holds for every threshold $u_s \geq u_s^0$. In particular we have that the associated shape parameter is the same for u_s and u_s^0. This property is called threshold stability. The threshold stability can be exploited by estimating the shape parameter for a range of high thresholds and plotting:

$$(u_s, \hat{\xi}_s(u_s))_{u_s \geq u_s^0},$$

where $\hat{\xi}_s(u_s)$ denotes the estimated shape parameter for threshold u_s. If the GPD model holds for u_s^0 the graph should be (approximately) constant for $u_s \geq u_s^0$. However, owing to the decreasing number of excesses above higher thresholds, the plot becomes unstable and the constant behavior is difficult to see. By combining different diagnostics one hopes for a better picture of the threshold to be used. In this study we, therefore, consider also the ME plot.

The ME plot relies on a similar consideration and utilizes the mean excess function

$$e_s(u) := E[X_s - u | X_s > u]$$

where X_s represents the daily rainfall at site s. The empirical version of the mean excess function is given by

$$\hat{e}_s(u) = \frac{\sum_{t=1}^{T}(x_s(t) - u)\mathbb{1}_{(u,\infty)}(x_s(t))}{\sum_{t=1}^{T}\mathbb{1}_{(u,\infty)}(x_s(t))},$$

where $\mathbb{1}_A$ is the indicator function for set A, i.e. $\mathbb{1}_A(x) = 1$ if $x \in A$ and otherwise zero. If the GPD model holds, $e_s(u)$ is linear in the threshold u, see e.g. Embrechts et al. (1997), and $\hat{e}_s(u)$ becomes approximately linear. However, as described in greater detail in Ghosh and Resnick (2010), there are some problems associated with the use of the ME plot. These are in particular that the ME function is only well defined for $\xi < 1$ and – as for the TC plot – the empirical ME function becomes unstable for high values of the threshold, see also the comments of Dr. Kimber in the discussion of Davison and Smith (1990). In our specific application the condition $\xi < 1$ is no restriction as even the highest estimates of ξ for daily extreme precipitation are far below 1, see e.g. Martins and Stedinger (2000) and Papalexiou and Koutsoyiannis (2012).

We want to select $\tau_0 \in [0, 1)$, such that the GPD model is valid above the τ_0 quantile at every grid point. We denote by $q_s(\tau) := F_{X_s}^{-1}(\tau)$ the τ quantile of the data at site s, by $\hat{\xi}_s(\tau)$ the corresponding estimate of the shape parameter and by $\hat{e}_s(\tau)$ the corresponding value of the empirical ME function. We assume, there exists for each s a τ_0^s, such that the GPD model holds for all excesses $X_s - q_s(\tau_0^s)$. Then we define $\tau_0 := \max_s \tau_0^s$ and it is clear that the GPD model holds for $q_s(\tau)$ for every grid point $s \in \{1, ..., S\}$ and $\tau > \tau_0$, i.e. $\hat{\xi}_s(\tau)$ should be approximately constant and $\hat{e}_s(\tau)$ should be approximately linear for $\tau > \tau_0$. Concerning the instabilities of the TC and ME plots, the constant (respectively linear) behavior might not be apparent for every single grid point s. Therefore, we propose to consider spatial averaging of the TC (respectively ME) plot to reveal the underlying structure better. The spatially averaged TC plot is given by

$$(q_{avg}(\tau), \hat{\xi}_{avg}(\tau))_{\tau \in [0,1)}, \tag{3}$$

where $q_{avg}(\tau) = S^{-1}\sum_{s=1}^{S}q_s(\tau)$ and $\hat{\xi}_{avg}(\tau) = S^{-1}\sum_{s=1}^{S}\hat{\xi}_s(\tau)$. The spatially averaged ME plot is given by

$$(q_{avg}(\tau), \hat{e}_{avg}(\tau))_{\tau \in [0,1)}, \tag{4}$$

where $\hat{e}_{avg}(\tau) = S^{-1}\sum_{s=1}^{S}\hat{e}_s(\tau)$. Note that while the spatially averaged TC plot works if the shape parameter is site-specific, the spatially averaged ME plot requires a common shape parameter. The strength

of the spatially averaged plot lies in the increased detection probability of non-constant (respectively non-linear) behavior, when the threshold is too low.

2.4.2. Quantile regression

Quantile regression relies on the fact that a sample quantile can be viewed as the solution of an optimization problem, which can be computed efficiently using linear programming, as shown in Koenker and Bassett Jr. (1978). For a fixed site $s \in \{1, ..., S\}$, we can obtain the τ-th sample quantile of the declustered observations $x_s = (x_s(1), ..., x_s(T))$ as

$$\arg \min_{\beta \in \mathbb{R}} \sum_{t=1}^{T} \rho_\tau(x_s(t) - \beta), \tag{5}$$

where

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

This can be easily generalized by replacing β by a suitable regression model, e.g. in order to obtain for each site a linear trend in the temporal covariate $z(t)$, we determine

$$\arg \min_{\beta_s^0, \beta_s^1 \in \mathbb{R}} \sum_{s=1}^{S}\sum_{t=1}^{T} \rho_\tau(x_s(t) - \beta_s^0 - \beta_s^1 \cdot z(t)), \tag{6}$$

which can be done for each site separately. It may be useful to assume a common relative trend over the region, compare Hanel et al. (2009) who use a common relative trend in the location parameter of the generalized extreme value distribution. Then, we have to determine

$$\arg \min_{\beta_s^0, \beta^1 \in \mathbb{R}} \sum_{s=1}^{S}\sum_{t=1}^{T} \rho_\tau(x_s(t) - \beta_s^0[1 + \beta^1 \cdot z(t)]), \tag{7}$$

where β_s^0 is a site specific niveau component and β^1 a common relative trend. This common relative trend can be found by profile quantile regression, i.e. we compute

$$\sum_{s=1}^{S} \min_{\beta_s^0} \sum_{t=1}^{T} \rho_\tau(x_s(t) - \beta_s^0[1 + \beta^1 \cdot z(t)])$$

over a grid of possible values of β^1 and select the one that minimizes the sum on the right side. The β_s^0 are determined for each fixed β^1 as the solutions of the independent minimization problems:

$$\min_{\beta_s^0 \in \mathbb{R}} \sum_{t=1}^{T} \rho_\tau\left(\frac{x_s(t)}{1 + \beta^1 \cdot z(t)} - \beta_s^0\right).$$

We want to test whether β^1 differs significantly from zero. However, the distribution of β^1 under the null hypothesis, i.e. that there is no trend, is not known. Therefore, we calculate p-values by a block-wise bootstrap approach (compare e.g. Douglas et al., 2000). Seasonal blocks over the whole spatial domain are sampled with replacement, therefore the newly created data should have approximately the same spatial dependence structure as the original data but no trend component. Calculating β^1 multiple times for these bootstrap data delivers an approximate distribution of β^1 under the null hypothesis.

2.5. Estimating the excess distribution

For a specific model of the GPD parameters, e.g.

$$\gamma(t) = \gamma_0 + \gamma_1 \cdot z(t), \quad \xi(t) = \xi_0, \tag{8}$$

and fixed threshold $(u_s(t))_{s \in \{1,...,S\}, t \in \{1,...,T\}}$ we estimate the vector of parameters $\theta = (\gamma_0, \gamma_1, \xi_0)'$ by maximizing the following function:

$$\ell_I(\theta; \mathbf{y}) = -\sum_{\substack{t=1 \\ y_s(t) \geq 0}}^{T} \sum_{s=1}^{S} \left[\ln(\gamma(t) u_s(t)) + \frac{1+\xi(t)}{\xi(t)} \ln \left(1 + \frac{\xi(t) y_s(t)}{\gamma(t) u_s(t)} \right) \right]. \tag{9}$$

This is the so-called independence log likelihood, i.e. the log likelihood that would be obtained, if peaks at different sites were independent of each other, see also Moore (1987), Smith (1989), Buishand (1991), Cooley et al. (2007), and Hanel et al. (2009). We rely on this simplified likelihood function, because the estimation of the full likelihood function would be virtually impossible due to spatial dependence and large dimensionality, compare also Thibaud et al. (2013). This method provides asymptotically unbiased parameter estimates, but the spatial dependence in the data results in a (highly) increased variance of the estimates compared to the variance that would be obtained for independent data. Therefore, Smith (1990) suggested to adjust the standard errors and likelihood ratio tests in a way that is now generalized in the composite likelihood framework, see Varin et al. (2011) for an extensive overview. For applications of this approach, see e.g. Blanchet and Lehning (2010) for annual maximum snow depths over Switzerland and Van de Vyver (2012) for annual extremes of precipitation in Belgium. We refer to Roth et al. (2012) for the details of the estimation of the GPD parameters.

2.6. Bias correction

Climate models represent the current status of knowledge about the climate system but are imperfect still. Systematic differences occur between climate model data and observations, which translate into biases in derived quantities like return levels of daily precipitation. There are different ways to correct for these biases. In the discussion below, we will use the following notation. By superscript 'Mod' we denote all entities, such as parameters and return level, that are derived from the climate model data and by superscript 'Obs' those from the observations. Further, we denote by T^C the end of the overlapping or control period.

A simple way to adjust the simulated return levels, in the spirit of quantile matching, is to apply a return period specific change factor:

$$\tilde{r}_{s,t}(\alpha) = r_{s,t}^{\text{Mod}}(\alpha) \cdot \frac{\sum_{i=1}^{T^C} r_{s,i}^{\text{Obs}}(\alpha)}{\sum_{i=1}^{T^C} r_{s,i}^{\text{Mod}}(\alpha)}. \tag{10}$$

A drawback of this approach is that in general the adjusted quantiles will no longer follow a GPD. It is even possible that they do not increase monotonically with increasing return period. This can be overcome if we adjust the GPD parameters and the threshold instead, i.e.

$$\tilde{\xi}(t) = \xi^{\text{Mod}}(t) - \frac{1}{T^C} \sum_{i=1}^{T^C} (\xi^{\text{Obs}}(i) - \xi^{\text{Mod}}(i)),$$

$$\tilde{\gamma}(t) = \gamma^{\text{Mod}}(t) \cdot \frac{\sum_{i=1}^{T^C} \gamma^{\text{Obs}}(i)}{\sum_{i=1}^{T^C} \gamma^{\text{Mod}}(i)},$$

$$\tilde{u}_s(t) = u_s^{\text{Mod}}(t) \cdot \frac{\sum_{i=1}^{T^C} u_s^{\text{Obs}}(i)}{\sum_{i=1}^{T^C} u_s^{\text{Mod}}(i)}. \tag{11}$$

For the shape parameter an additive adjustment was chosen owing to the small and possibly negative values of this parameter. Using the adjusted parameters in Eq. (2), we get an estimate of $r_{s,t}(\alpha)$. The proposed correction yields that the temporal mean of the adjusted return levels over the control period approximates the temporal mean of the estimated return levels for the observational data. The approach can be generalized by taking not only the biases in the means of the threshold and the GPD parameters into account but also the biases in their trends. However, one

should be very careful doing so as it might lead to implausible effects owing to the many parameters involved.

The uncertainty in the projected return levels, due to the unknown GPD parameters, can be assessed by the following procedure. First we generate bootstrap samples of the estimated GPD parameters for both the observations and the climate model data, e.g. by the method used in Roth et al. (2012), based on exponential residuals of the excesses. From these bootstrap samples a bootstrap sample of the adjusted return level is obtained, by applying correction scheme (11) to the estimated shape and dispersion parameters in the bootstrap samples from the climate model data and observations. Note that for long return periods the uncertainty of the threshold can be neglected, but this cannot be done for short return periods. Then one must generate bootstrap samples from all days in the season of interest rather than from days with threshold exceedances only.

3. Data

3.1. Region and precipitation data

The spatial domain of the study includes the Netherlands and a part of north-western Germany, roughly northwest of the line Cologne–Hamburg. The region is relatively flat with hills up to 300 m in the extreme south and south-east. The averaged elevation is for about 60% of the grid boxes lower than 25 m and only in less than 5% of the grid boxes higher than 100 m. Fig. 1 shows the observed mean annual precipitation totals for the considered grid points, which range from about 650 mm to 850 mm, without a clear spatial pattern. Precipitation data have been provided on a $0.22° \times 0.22°$ rotated pole grid and a total of 158 grid points falls into the specified region.

For this region we consider two transient simulations from the RACMO2 RCM of the Royal Netherlands Meteorological Institute (Van Meijgaard et al., 2008) from 1950 to 2100. The first one was driven by the ECHAM5 GCM, developed at the Max Planck Institute for Meteorology in Hamburg (Roeckner et al., 2003), and the second was driven by the GCM Model for Interdisciplinary Research on Climate (MIROC) (Hasumi and Emori, 2004). In the following these two simulations are denoted as R/ECHAM5 and R/MIROC. We selected these simulations for the present case study, because their precipitation projections are quite different, although both simulations are based on the SRES A1B emission scenario (Nakićenović and Swart, 2000).

We compare the climate model data with gridded observations from the E-OBS data set (Haylock et al., 2008), version 6.0, which is available on the same grid. With 3.1 stations on average per grid box (in total 492) the station density in this part of the E-OBS domain is quite high. Therefore, the data are considered to be
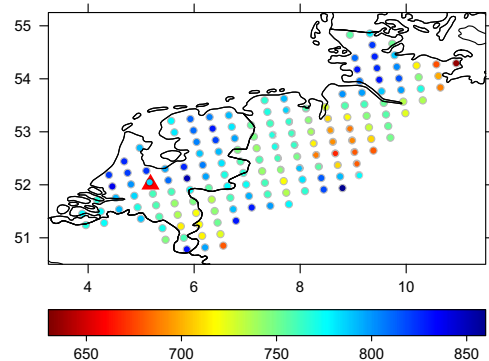

Fig. 1. Mean annual precipitation totals in mm for the considered grid cells. The red triangle indicates the position of De Bilt.

Table 1
Seasonal mean precipitation in mm/day.

| Season	Data	1950–2011	2071–2100	Change (%)
Winter	E-OBS	2.08	–	–
	R/ECHAM5	2.98	3.55	19.1
	R/MIROC	2.48	2.52	1.6
Summer	E-OBS	2.46	–	–
	R/ECHAM5	2.70	2.27	− 15.9
	R/MIROC	2.10	2.17	3.3

adequate for comparison with gridded RCM data. Table 1 contrasts the daily regional mean winter (December–February) and summer (June–August) precipitation for the climate model data with those for the observations during the control period 1950–2010. For the climate model data also the means for the future period 2071–2100 and the associated change are given. The R/ECHAM5 simulation shows a clear increase in winter and decrease in summer, which are both significant at the 1% level according to a two sample *t*-test based on the seasonal totals. For R/MIROC the changes in mean winter and summer precipitation are small. It is noteworthy that the climate simulations do not preserve the seasonal cycle: for the observations the mean precipitation is higher in summer than in winter, but for the climate model data the opposite is found.

3.2. Temperature as covariate

The enhanced greenhouse gas effect is anticipated to be small or not existent in the first decades of the simulation and increasing by the end of the 20th century. This is inconsistent with a simple linear trend over time for the threshold and GPD parameters. Instead of applying more complicated relationships with time, leading to increased estimation uncertainty, a covariate that is considered representative of the enhanced greenhouse gas effect is used.

With rising temperatures the water holding capacity of the atmosphere increases. As extreme precipitation strongly depends on the available precipitable water, compare e.g. Lenderink and van Meijgaard (2008), temperature is a natural covariate for the non-stationary POT approach. Hanel et al. (2009) used the seasonal global temperature anomaly, from the driving GCM, as covariate. However, a prominent property of the standard time covariate is that trends over time are easily comparable between different models. In order to maintain this property, and in view of the common driving scenario, we consider a common, regional temperature covariate.

First we determine for each year the seasonal mean temperature in the R/ECHAM5 and R/MIROC data, aggregated over all grid boxes in the RACMO2 domain with at least 50% land coverage. These seasonal mean temperatures were averaged for the two climate simulations and the anomalies were then computed with respect to the control period. The temperature anomalies are decomposed into a short-range dependent component and an increasing trend component using monotone regression (Wu et al., 2001). For an introduction to monotone regression, see e.g. Robertson et al. (1988). In this study the monotone median algorithm of Koenker and Ng (2005) is used, and the fitted values constitute the common covariate $z(t)$. Fig. 2 shows $z(t)$ for the winter and the summer season. Note the breaks in winter around 1960 and 2020, and in summer around 1990 and 2055. These are owing to the piecewise linear fit in the regression setting and do not have a specific physical meaning. The fitted curves are assumed to be representative of the enhanced greenhouse effect, as projected in the two climate simulations.
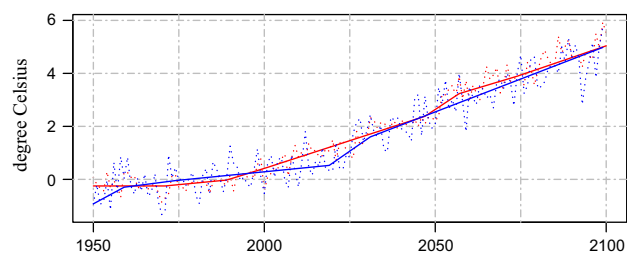
Fig. 2. Seasonal regional temperature anomaly with respect to the period 1950–2011 (dashed) and fitted monotone trend (solid) for winter (blue) and summer (red).

Over the control period the observed trend in temperature is larger than the (averaged) trend of the climate simulations in this region. Van Oldenborgh et al. (2009) explored several explanations for the stronger observed than modeled warming in western Europe, the most important being a stronger trend to westerly circulation in winter and a stronger trend towards more short-wave radiation in summer than simulated by the climate models. The pronounced trend in observed temperature is accounted for in an alternative covariate $z_{E-OBS}(t)$, based on a monotone regression fit to the observed seasonal temperature anomalies.

4. Results and discussion

4.1. Temporal dependence and declustering

Prior to the analysis of the extremes in the precipitation data, we apply the declustering scheme, outlined in Section 2.2. In winter we observe large numbers of clusters of lengths 2 and 3 of the exceedances of the 95% quantile, both in the climate model data and the observations, indicating a considerable amount of temporal dependence. As the number of clusters of length 4 is less than one per decade, we choose in winter a separation time of 2 days. Consistent with the literature, compare e.g. Kyselý and Beranová (2009), the temporal dependence in summer is much weaker, manifesting in a small number of clusters of length 3. Therefore, a separation time of 1 day is taken as sufficient for the summer season.

4.2. Threshold

4.2.1. Quantile selection

We illustrate the quantile selection approach proposed in Section 2.4 using the winter data from the R/MIROC simulation. From the spatially averaged TC and ME plots (Figs. 3 and 4), it can be seen that the 95% quantile is too low for the GPD model to hold, i.e. the spatially averaged estimates of the shape parameter to the right of the spatially averaged 95% quantile in the TC plot are neither constant nor is the ME plot linear. The 97.5% quantile seems to be high enough for the GPD model to hold. Similar pictures for the winter data from the ECHAM5 driven simulation and the observations were obtained. Therefore, we took the 97.5% quantile as threshold for all winter data. For the summer data it was necessary to reject also the 97.5% quantile and the 98.5% quantile was used as threshold, i.e. on average we consider 1.38 excesses per grid point and season. Note that this is a considerably higher quantile than in many other studies (Friederichs, 2010; Kyselý et al., 2010; Halmstad et al., 2012).

4.2.2. Quantile regression

For each grid point and season, we fitted the simple linear quantile regression model in Eq. (6) to the observed precipitation data. As

predictor for the 97.5% quantile in winter (respectively the 98.5% quantile in summer) we used the regional temperature anomaly $z(t)$. This was repeated for the climate model data both for the control period and the full period. The resulting local trends (i.e. slope parameters β_s^1) are shown in Fig. 5 for winter and Fig. 6 for summer. In winter the observed data show rather large positive slopes, which are not found for the climate model data in the control period. The

Fig. 3. Spatially averaged TC plot for R/MIROC, winter season. The dashed vertical lines mark the spatially averaged quantile for $\tau = 0.95$, 0.975 and 0.985.

Fig. 4. Same as Fig. 3 but for the spatially averaged ME plot.

large increase in precipitation extremes is consistent with the observed increase in 10-day precipitation maxima in this part of Europe. Van Haren et al. (2013a) attribute part of the latter trend to the change in circulation, which explains also partly the relatively large increase in mean winter temperature. In summer, the climate model output exhibits more negative slopes than the observations. As the considered region is bordered by the North Sea, this might be attributed to a, lower than observed, trend in sea surface temperature in this coastal area in the climate models (compare Van Haren et al., 2013b).

It is striking that for both climate simulations the trends obtained for the control period do not resemble those obtained for the full period. In particular, the spread of the estimated trends over the region is much smaller for the full period than for the control period, consistent with the decreased estimation uncertainty for the full period. Moreover, there is no clear spatial pattern in the trends for the full period nor for the control period. Therefore, we assume in the following a common relative trend over the region, i.e. we fit the common relative trend model, given in Eq. (7), to the daily precipitation data. Table 2 shows that the climate simulations give quite different trends for the control period in summer even with opposite sign, but for the full period the trends are much closer. This is consistent with the results for the local trends. The values for the observations are quite large. This is partly caused by the fact that the temperature covariate $z(t)$ underestimates the temperature trend in the observations (Section 3.2). If we replace the covariate $z(t)$ by $z_{E-OBS}(t)$ we obtain an increase of 3.7% per degree Celsius change for the observed winter data and 2.8% in summer, resembling the estimates from the climate simulations for the full period.

We computed p-values for the common trend by the blockwise bootstrap, outlined in Section 2.4.2. It turns out that the trends are significant only for the full period in winter for both
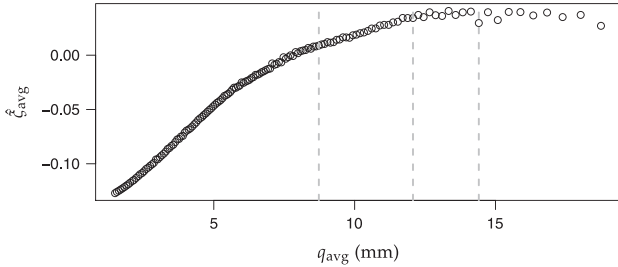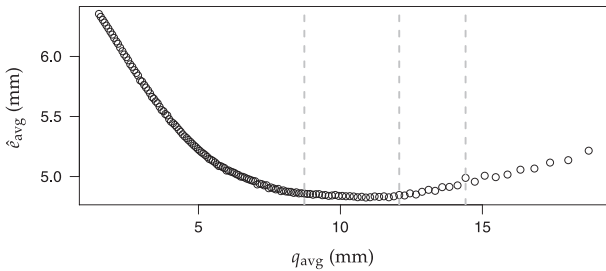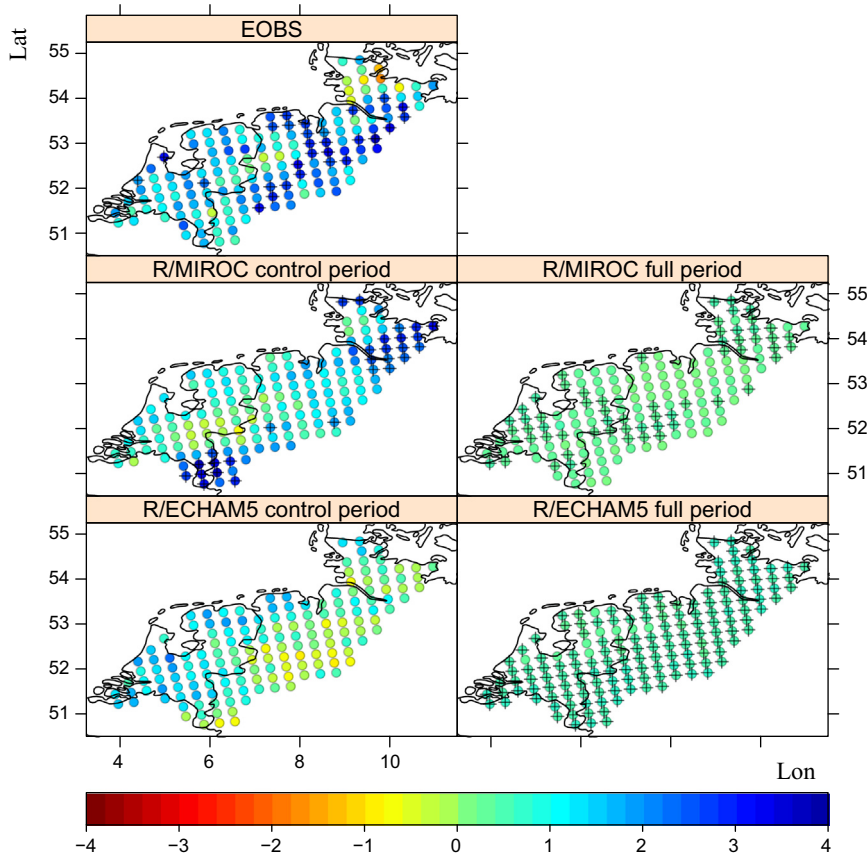

Fig. 5. Change in the threshold (mm per degree warming; pluses indicate significance at the 5% level) for the winter season, i.e. β_s^1 in Eq. (6).
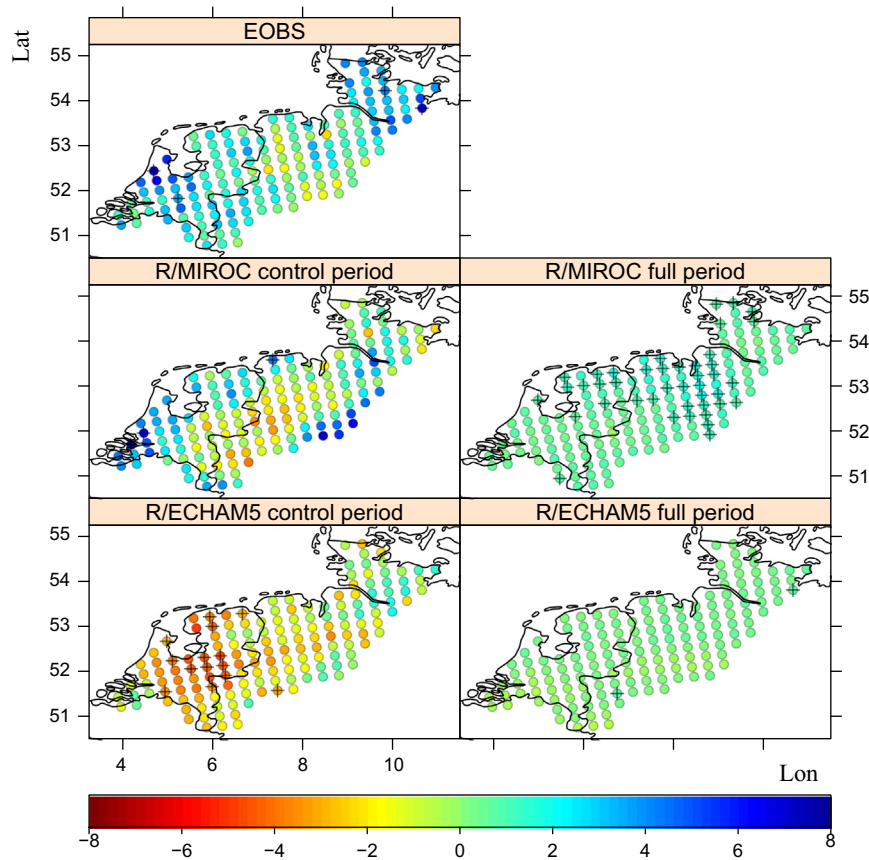
Fig. 6. Same as Fig. 5 but for the summer season.

Table 2
Change in the threshold (% per degree warming; bold script indicates significance at the 5% level), i.e. 100 β^1 in the model with a common relative trend, Eq. (7).

Season	Data	1950–2011	1950–2100
Winter	E-OBS	10.6	–
	R/ECHAM5	3.8	**3.6**
	R/MIROC	9.4	**2.2**
Summer	E-OBS	7.4	–
	R/ECHAM5	−6.3	1.5
	R/MIROC	8.8	**4.6**

Table 3
Slope parameter of the linear regression model for the GPD dispersion coefficient, i.e. γ_1 in Eq. (8), for each season (bold – significant at 5% level).

Season	Data	1950–2011	1950–2100
Winter	E-OBS	0.0905	–
	R/ECHAM5	−0.0093	0.0010
	R/MIROC	0.0324	0.0086
Summer	E-OBS	0.0085	–
	R/ECHAM5	−0.0470	**0.0258**
	R/MIROC	0.0247	**0.0400**

models and in summer only for R/MIROC, which is consistent with the significance of the local trends, compare the number of pluses in Figs. 5 and 6. Therefore, we are considering the following temporally varying thresholds only for those data sets, for the other data sets the thresholds are taken to be constant.

4.3. Excess distribution

For the GPD parameters it was assumed that the dispersion coefficient varies linearly with the temperature covariate and that the shape parameter is constant, see Eq. (8). The significance of the trend in the dispersion coefficient was determined by the composite likelihood ratio test, which is an extension of the classical likelihood ratio test, that takes the spatial dependence into account (compare Varin et al., 2011; Roth et al., 2012). The results are given in Table 3. Just as for the threshold, the control period does not provide a clear picture and the strong positive trend in the dispersion coefficient of the winter observation data vanishes, if the observed regional temperature anomalies z_{E-OBS} are used as

covariate. However, for the whole period both climate simulations agree reasonably well in each season. In winter, where a significant trend in the threshold was found for both models, the trend in the dispersion coefficient is negligible. In summer the dispersion coefficient is significantly increasing for R/ECHAM5 as well as for R/MIROC. Hanel and Buishand (2011) reported a similar trend in the dispersion coefficient of the 1-day summer precipitation maxima in this region for an ensemble of 15 transient regional climate model simulations.

The need to include a trend in the dispersion coefficient can be demonstrated with the standard exponential residuals (Roth et al., 2012) of the model fit. Fig. 7 shows these transformed residuals for the R/MIROC summer season obtained using the model without and with trend in the dispersion coefficient. Whereas the former exhibit a highly significant linear trend, the linear trend in the latter is close to zero and there is no indication for other trends.

Table 4 shows for both climate simulations and the observations the shape parameter ξ, the temporally averaged dispersion coefficient γ, and the temporally and regionally averaged threshold u.
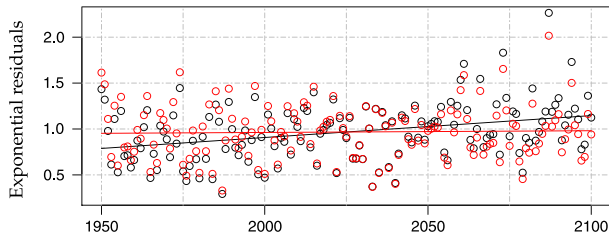
Fig. 7. Standard exponential residuals for R/MIROC summer averaged over season and the whole domain (black circles – model without trend in the dispersion coefficient, black line – corresponding linear regression, red – same for the model with trend in the dispersion).

Fig. 8. 50-year summer return level, $r(0.02)$, at the grid point closest to De Bilt for R/MIROC with 95% confidence bands. The solid black line (respectively gray band) is based on regional 30-year time slices and the solid red line (respectively red band) is based on the non-stationary approach. The blue shadow indicates the 95% confidence band based on the non-stationary at-site model.

Table 4
Mean threshold u (mm), dispersion coefficient γ, and shape parameter ξ for both seasons.

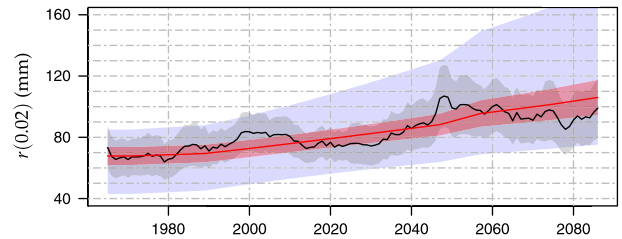| Season	Data	1950–2011			1950–2100		
		u	γ	ξ	u	γ	ξ
Winter	E-OBS	11.8	0.417	−0.001	–	–	–
	R/ECHAM5	14.0	0.303	0.050	14.8	0.339	−0.024
	R/MIROC	11.5	0.348	0.061	12.1	0.371	0.056
Summer	E-OBS	17.9	0.321	0.161	–	–	–
	R/ECHAM5	17.5	0.413	0.102	17.8	0.418	0.198
	R/MIROC	15.9	0.506	0.222	17.6	0.581	0.207

Fig. 9. Bias-corrected 50-year summer return level, $r(0.02)$, at the grid box closest to De Bilt for R/MIROC (red) and R/ECHAM5 (blue) with 95% confidence bands obtained by the bootstrap procedure, outlined in Section 2.6. The constant 50-year return level of the observations is given as reference (dashed line).

While the shape parameters and the threshold values from the climate simulations are comparable to the observed ones, the dispersion coefficients from the climate model data are too low in winter and too high in summer. Note that the shape parameter for the E-OBS summer season almost equals the constant 0.15, found by Koutsoyiannis (2004) for annual block maxima of daily precipitation in different parts of the world.
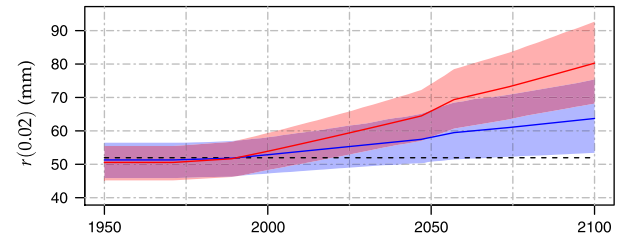
4.4. Return levels

For the selected threshold and the estimated parameters we can compute a time-dependent return level using Eq. (2). As an alternative to the non-stationary POT approach, we consider a 'time slice' approach. Therefore, we estimate for each 30-year window of the data (i.e. 1950–1979, 1951–1980, etc.) a common shape parameter and common dispersion coefficient as if the data were stationary. Then the return levels are computed based on these estimates. For the MIROC driven simulation Fig. 8 shows, for both approaches, the 50-year return level of daily precipitation in the summer season at the grid point closest to De Bilt. The confidence bands are obtained using the asymptotic normality of the maximum independence likelihood estimator of the GPD parameters (Davison, 2003; Varin et al., 2011). This ignores the uncertainty in the threshold, which is small for the 50-year return level compared to the uncertainty due to the GPD parameters.

Overall, the figure shows a good agreement between both methods. However, the 95% confidence band for the non-stationary POT approach is considerably narrower than that for the time-slice approach, owing to the increased number of data points used for the estimation. In fact, the relative standard error reduces from about 8.7% for the moving window approach to about 4.5% for the non-stationary approach. Moreover, a monotone trend in the return level is more plausible than the irregular pattern of the trend for the time-slice approach, where slightly different selections of the windows can produce quite different estimates of the change, e.g. it matters a lot if the period 2063–

2092 ($r(0.02) = 85$ mm) or the period 2071–2100 ($r(0.02) = 99$ mm) is taken as future period. This is even more delicate because the control period can be chosen in different ways too. Fig. 8 also shows the confidence band if only at-site data are used to estimate the parameters of the non-stationary model. These are four times wider than those obtained by the regional estimation approach.

The 50-year return level of the 1-day summer maximum precipitation near De Bilt from the MIROC driven simulation in Fig. 8 is significantly larger than the estimate of 52 mm from the observations. Therefore, a bias correction is needed. Bias correction can be very sensitive to trends, e.g. when they are close to zero but of different signs. Thus, we take only significant trends in the threshold and GPD parameters into account. Note that the proposed correction scheme, see Section 2.6, assumes that the bias is constant over time. Maraun (2012) finds this justified for seasonal precipitation sums over most of Europe and in particular over the here considered study area. Although, one cannot directly deduce from this that the bias in extremes is constant as well, a small experiment, splitting the historical period in two subperiods, justifies the use of the proposed bias correction.

Fig. 9 shows the bias-corrected 50-year return level for the summer season for both the R/ECHAM5 and R/MIROC data, based on the adjusted threshold and GPD parameters. The results for winter are shown in Fig. 10. As the trend in the control period is negligible for both models and seasons, the bias correction removes almost the whole bias, which would not be the case if the trend in the observations was significant. We see that in summer R/MIROC projects a much stronger increase (45%) than R/ECHAM5 (15%), while the situation in winter is the opposite. Then, R/MIROC projects a 11% increase in extreme precipitation, and R/ECHAM5 a 22% increase. The latter is about the same as the trend in the mean, see Table 1. In winter similar changes in mean and extreme precipitation have been reported for other RCM simulations in parts of Europe (Frei et al., 2006; Kyselý et al., 2011; Hanel and Buishand, 2012). However, in summer the trend
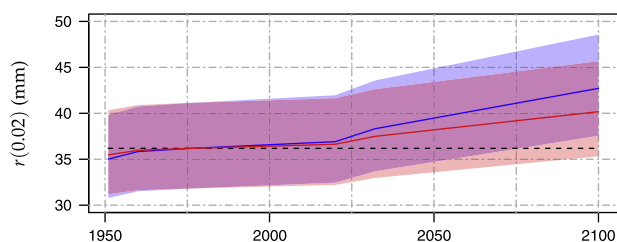
Fig. 10. Same as Fig. 9 but for the winter season.

Table 5
Design life level (mm) for a 50-year life time at De Bilt. The values for R/ECHAM5 and R/MIROC are based on the bias-corrected GPD parameters for the summer season. The values in parentheses denote the 95% confidence interval for the percentage change as obtained by the bootstrap procedure, outlined in Section 2.6.

p	Data	1950–1999	2016–2065	Change (%)
0.10	E-OBS	82	–	–
	R/ECHAM5	82	92	12.9 (4.3, 22.6)
	R/MIROC	81	104	27.9 (19.8, 36.7)
0.05	E-OBS	95	–	–
	R/ECHAM5	94	106	13.4 (4.4, 23.5)
	R/MIROC	93	120	28.7 (20.0, 37.6)

in mean precipitation does not give a good indication of the trend in extremes.

In Fig. 10 one can see that the width of the confidence band for the return level is the same for both climate model simulations and that it does not vary over time. This is owing to the bias-correction scheme. As there is no trend in the GPD parameters of the climate model data in the winter season, the bias correction always yields the values of the (bootstrapped) GPD parameters from the observations. Thus, the width of the confidence band is determined by the variance of these parameter estimates. This implies that only longer observation records can reduce the uncertainty. However, in summer the trend in the dispersion coefficient is significant for the climate model data. Including this trend in the POT model increases the width of the confidence band towards the end of the 21st century, as shown in Fig. 9.

Return levels of precipitation are often the basis of hydrologic design. In a stationary climate the risk that a system fails within its expected life time is directly related to the return period, e.g. the probability that a 50-year return level is exceeded at least once during 50 years is 64%. This holds no longer for a non-stationary climate. Rootzén and Katz (2013) recently proposed a risk-based design level that can be used both for a stationary and a non-stationary climate. The p design life level for a specified period gives the value that is exceeded with probability p during the design life period. Table 5 shows the 10% and 5% design life level for De Bilt, computed for two different 50-year design life periods, using the bias-corrected GPD parameters for the summer season. For both climate simulations a significant increase is found for the future period 2016–2065. The 95% confidence interval for the change in design life level is based on the same bootstrap samples as those used to generate the confidence band for the 50-year return level in Fig. 9. The relative changes for $p=0.10$ and $p=0.05$ are almost identical. The change in the design life level for the ECHAM5 driven simulation is half the size of the change in the MIROC driven simulation, similar to the difference in the change in the summer 50-year return level, see Fig. 9. However, the change in the design life level can be compared only roughly with that in the return level, as it takes the whole design life period into account rather than only two points in time.

A caveat is that the 5% design life level corresponds to a very rare rainfall amount. In a stationary climate the associated

return period to this design life level is 1000 (exactly 975) years, which is on the edge of reasonable extrapolation. In general, for a stationary climate, the return period corresponding to the p design life level for a design life period of n years, follows from

$$R = \frac{1}{1 - \sqrt[n]{1-p}} \approx \frac{1}{1 - \left(1 - \frac{p}{n}\right)} = \frac{n}{p}. \tag{12}$$

5. Conclusion

In this study we applied the regional non-stationary POT model of Roth et al. (2012) to precipitation extremes of two transient climate simulations for the period 1950–2100, conducted with the regional climate model RACMO2, driven by the general circulation models ECHAM5 and MIROC. The simulated 1-day summer and winter precipitation extremes in the Netherlands and north-western Germany were compared with those in the gridded observation data set E-OBS for the period 1950–2011.

Visual inspection of the spatially averaged TC plot and the spatially averaged ME plot leads to a rejection of the 95% quantile as threshold, which is often used in the literature for precipitation data. It may be useful to develop regional goodness-of-fit tests to make the selection of the threshold more objectively. A regional quantile regression model using a common relative trend for the threshold was used to smooth the large spatial scattering of the local trends in the threshold for the control period. The absence of a spatial pattern in the trends for both the control and full period justifies this model. The non-stationary approach leads to return levels, that are consistent with those obtained by a 30-year moving window approach, but exhibit less uncertainty. For the considered 50-year return level the uncertainty is reduced by a factor of two, compared to the moving window approach. The simulated return levels exhibit a considerable positive bias, which was corrected for by adjusting the parameters of the peaks-over-threshold model. The uncertainty in the adjusted return levels is then strongly governed by the variance of the estimated GPD parameters from the observations. Therefore, accurate estimates of the GPD parameters, based on high quality observed precipitation records, are needed to project future extremes.

For the winter season, the two climate model simulations project a significant trend in the threshold. This is in line with a positive trend in the location parameter of the generalized extreme value distribution fitted to the 5-day winter precipitation maxima in this region for an ensemble of 15 transient regional climate model simulations (Hanel and Buishand, 2011). The difference between the R/ECHAM5 and R/MIROC simulations does not give the full range of possible future projections. In summer the differences between the climate model data are larger. The ECHAM5 driven simulation projects a 15% increase of the 50-year return level, based on a significant increase of the dispersion coefficient, which corresponds well to the increase in the dispersion coefficient of the generalized extreme value distribution found in Hanel and Buishand (2011). The MIROC driven simulation projects additionally a significant increase of the threshold, resulting in a 45% increase of the 50-year return level.

References

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

Blanchet, J., Lehning, M., 2010. Mapping snow depth return levels: smooth spatial modeling versus station interpolation. Hydrol. Earth Syst. Sci. 14 (12), 2527–2544, http://dx.doi.org/10.5194/hess-14-2527-2010.

Buishand, T.A., 1991. Extreme rainfall estimation by combining data from several sites. Hydrol. Sci. J. 36 (4), 345–365, http://dx.doi.org/10.1080/02626669109492519.

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

Cooley, D., Nychka, D., Naveau, P., 2007. Bayesian spatial modeling of extreme precipitation return levels. J. Am. Stat. Assoc. 102 (479), 824–840, http://dx.doi.org/10.1198/016214506000000780.

Davison, A.C., 2003. Statistical Models. Cambridge University Press, Cambridge, UK.

Davison, A.C., Smith, R.L., 1990. Models for exceedances over high thresholds J. R. Stat. Soc. B 52 (3), 393–442.

Douglas, E.M., Vogel, R.M., Kroll, C.N., 2000. Trends in floods and low flows in the United States: impact of spatial correlation. J. Hydrol. 240, 90–105, http://dx.doi.org/10.1016/S0022-1694(00)00336-X.

Ekström, M., Fowler, H.J., Kilsby, C.G., Jones, P.D., 2005. New estimates of future changes in extreme rainfall across the UK using regional climate model integrations. 2. Future estimates and use in impact studies. J. Hydrol. 300 (1–4), 234–251, http://dx.doi.org/10.1016/j.jhydrol.2004.06.019.

El Adlouni, S., Ouarda, T.B.M.J., Zhang, X., Roy, R., Bobée, B., 2007. Generalized maximum likelihood estimators for the nonstationary generalized extreme value model. Water Resour. Res. 43 (3), W03410, http://dx.doi.org/10.1029/2005WR004545.

Embrechts, P., Klüppelberg, C., Mikosch, T., 1997. Modelling Extremal Events. Springer, Berlin.

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

Frei, C., Schöll, R., Fukutome, S., Schmidli, J., Vidale, P.L., 2006. Future change of precipitation extremes in Europe: intercomparison of scenarios from regional climate models. J. Geophys. Res. 111, D06105, http://dx.doi.org/10.1029/2005JD005965.

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

Ghosh, S., Resnick, S., 2010. A discussion on mean excess plots. Stoch. Proc. Appl. 120 (8), 1492–1517, http://dx.doi.org/10.1016/j.spa.2010.04.002.

Halmstad, A., Najafi, M.R., Moradkhani, H., 2012. Analysis of precipitation extremes with the assessment of regional climate models over the Willamette River Basin, USA. Hydrol. Process. 27 (18), 2579–2590 ,http://dx.doi.org/10.1002/hyp.9376.

Hanel, M., Buishand, T.A., 2011. Analysis of precipitation extremes in an ensemble of transient regional climate model simulations for the Rhine basin. Clim. Dyn. 36, 1135–1153, http://dx.doi.org/10.1007/s00382-010-0822-2.

Hanel, M., Buishand, T.A., 2012. Multi-model analysis of RCM simulated 1-day to 30-day seasonal precipitation extremes in the Czech Republic. J. Hydrol. 412–413, 141–150, http://dx.doi.org/10.1016/j.jhydrol.2011.02.007.

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

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

Hosking, J.R.M., Wallis, J.R., 1997. Regional Frequency Analysis: An Approach Based on L-Moments. Cambridge University Press, Cambridge, UK.

Nakićenović, N., Swart, R., 2000. Special Report on Emissions Scenarios: A Special Report of Working Group III of the Intergovernmental Panel on Climate Change. Cambridge Univ. Press, New York.

Hasumi, H., Emori, S., 2004. K-1 Coupled Model (MIROC) Description, Center for Climate System Research, K-1 Technical Report. University of Tokyo, Tokyo.

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

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

Koenker, R., Ng, P., 2005. Inequality constrained quantile regression. Sankhya 67 (2), 418–440.

Koutsoyiannis, D., 2004. Statistics of extremes and estimation of extreme rainfall: II. Empirical investigation of long rainfall records. Hydrol. Sci. J. 49 (4), 591–610.

Kyselý, J., Beranová, R., 2009. Climate-change effects on extreme precipitation in Central Europe: uncertainties of scenarios based on regional climate models. Theor. Appl. Climatol. 95, 361–374, http://dx.doi.org/10.1007/s00704-008-0014-8.

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

Kyselý, J., Gaál, L., Beranová, R., Plavcová, E., 2011. Climate change scenarios of precipitation extremes in Central Europe from ENSEMBLES regional climate models. Theor. Appl. Climatol. 104, 529–542, http://dx.doi.org/10.1007/s00704-010-0362-z.

Leadbetter, M.R., Lindgren, G., Rootzén, H., 1983. Extremes and Related Properties of Random Sequences and Processes. Springer, Berlin.

Lenderink, G., van Meijgaard, E., 2008. Increase in hourly precipitation extremes beyond expectation from temperature changes. Nat. Geosci. 1, 511–514, http://dx.doi.org/10.1038/ngeo262.

Maraun, D., 2012. Nonstationarities of regional climate model biases in European seasonal mean temperature and precipitation sums. Geophys. Res. Lett. 39, L06706, http://dx.doi.org/10.1029/2012GL051210.

Martins, E.S., Stedinger, J.R., 2000. Generalized maximum-likelihood generalized extreme-value quantile estimators for hydrologic data. Water Resour. Res. 36 (3), 737–744, http://dx.doi.org/10.1029/1999WR900330.

Moore, R.J., 1987. Combined regional flood frequency analysis and regression on catchment characteristics by maximum likelihood estimation. In: Singh, V.P. (Ed.), Regional Flood Frequency Analysis. Reidel, Dordrecht, pp. 119–131.

Papalexiou, S.M., Koutsoyiannis, D., 2012. Battle of extreme value distributions: a global survey on extreme daily rainfall. Water Resour. Res. 49 (1), 187–201, http://dx.doi.org/10.1029/2012WR012557.

Reiss, R.D., Thomas, M., 2007. Statistical Analysis of Extreme Values: With Applications to Insurance, Finance, Hydrology and Other Fields, 3rd ed. Birkhäuser, Basel.

Robertson, T., Wright, F.T., Dykstra, R.L., 1988. Order Restricted Statistical Inference. Wiley, Chichester.

Roeckner, E., Bäuml, G., Bonaventura, L., Brokopf, R., Esch, M., Giorgetta, M., Hagemann, S., Kirchner, I., Kornblueh, L., Manzini, E., Rhodin, A., Schlese, U., Schulzweida, U., Tompkins, A., 2003. The Atmospheric General Circulation Model ECHAM 5. Part I: Model Description. Technical Report No. 349 Max-Planck-Institute for Meteorology, Hamburg.

Rootzén, H., Katz, R.W., 2013. Design life level: quantifying risk in a changing climate. Water Resour. Res. 49 (9), 5964–5972, http://dx.doi.org/10.1002/wrcr.20425.

Roth, M., Buishand, T.A., Jongbloed, G., Klein Tank, A.M.G., van Zanten, J.H., 2012. A regional peaks-over-threshold model in a nonstationary climate. Water Resour. Res. 48, W11533, http://dx.doi.org/10.1029/2012WR012214.

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

Smith, R.L., 1990. Regional Estimation from Spatially Dependent Data. Unpublished Manuscript, University of North Carolina, Chapel Hill.

Sugahara, S., da Rocha, R.P., Silveira, R., 2009. Non-stationary frequency analysis of extreme daily rainfall in Sao Paulo Brazil. Int. J. Climatol. 29 (9), 1339–1349, http://dx.doi.org/10.1002/joc.1760.

Thibaud, E., Mutzner, R., Davison, A.C., 2013. Threshold modeling of extreme rainfall. Water Resour. Res. 49 (8), 4633–4644, http://dx.doi.org/10.1002/wrcr.20329.

Tramblay, Y., Neppel, L., Carreau, J., Najib, K., 2013. Non-stationary frequency analysis of heavy rainfall in southern France. Hydrol. Sci. J. 58 (2), 280–294, http://dx.doi.org/10.1080/02626667.2012.754988.

Turco, M., Llasat, M.C., 2011. Trends in indices of daily precipitation extremes in Catalonia (NE Spain) 1951–2003. Nat. Hazard Earth Syst. 11 (12), 3213–3226, http://dx.doi.org/10.5194/nhess-11-3213-2011.

Van de Vyver, H., 2012. Spatial regression models for extreme precipitation in Belgium. Water Resour. Res. 48, W09549, http://dx.doi.org/10.1029/2011WR011707.

Van Haren, R., van Oldenborgh, G.J., Lenderink, G., Hazeleger, W., 2013a. Evaluation of modeled changes in extreme precipitation in Europe and the Rhine basin. Environ. Res. Lett. 8, 014053, http://dx.doi.org/10.1088/1748-9326/8/1/014053.

Van Haren, R., van Oldenborgh, G.J., Lenderink, G., Collins, M., Hazeleger, W., 2013b. SST and circulation trend biases cause an underestimation of European precipitation trends. Clim. Dyn. 40, 1–20, http://dx.doi.org/10.1007/s00382-012-1401-5.

Van Meijgaard, E., van Ulft, L.H., van de Berg, W.J., Bosveld, F.C., van den Hurk, B.J.J.M., Lenderink, G., Siebesma, A.P., 2008. The KNMI Regional Atmospheric Climate Model RACMO Version 2.1. Technical Report 302 Royal Netherlands Meteorological Institute, De Bilt.

Van Oldenborgh, G.J., Drijfhout, S., van Ulden, A., Haarsma, R., Sterl, A., Severijns, C., Hazeleger, W., Dijkstra, H., 2009. Western Europe is warming faster than expected. Clim. Past 5 (1), 1–12, http://dx.doi.org/10.5194/cp-5-1-2009.

Varin, C., Reid, N., Firth, D., 2011. An overview of composite likelihood methods. Stat. Sin. 21, 5–42.

Wu, W.B., Woodroofe, M., Mentz, G., 2001. Isotonic regression: another look at the changepoint problem. Biometrika 88 (3), 793–804, http://dx.doi.org/10.1093/biomet/88.3.793.