Spatial analysis of trends in extreme precipitation events in high-resolution climate model results and observations for Germany

L. Tomassini[1] and D. Jacob[1]

[1]   A statistical extreme value analysis is applied to very high-resolution climate model results and observations encompassing the area of Germany. Two control runs representing the current climate, as well as three scenario simulations of the regional climate model REMO, are investigated. The control runs were compared against high-resolution observations. The analysis is divided into two main parts: first trends in extreme quantiles of daily precipitation totals are estimated in a station-by-station analysis. In the second part, the spatial characteristics of the estimated trends in heavy rainfall are investigated over the area of Germany by fitting a parametric geostatistical model to these trends. The rule of thumb of estimating trends in extreme quantiles of heavy precipitation based on the Clausius-Clapeyron relation, about 6.5% per 1°C temperature increase, has been roughly confirmed for Germany by our study with respect to the observations, but the climate model computes weaker trends. In the control simulations, the climate model tends to underestimate trends in heavy rainfall compared to observations. In the scenario simulations, positive trends prevail (as in the observations). They are, however, relatively small when set in relation to the uncertainties. The trends become significantly positive to a larger spatial extent only in the A2 scenario simulation. The estimated shape of the extreme value distributions does not change significantly in the scenario simulations compared to the climate model control runs. The parameter estimates for the geostatistical model for the trends in extreme quantiles of daily precipitation sums are rather uncertain. The most striking feature of the analysis is a reduction of the spatial variance of the trends over the considered area of Germany in the scenario simulations compared to observations and, in particular, the climate model control runs.

Citation: Tomassini, L., and D. Jacob (2009), Spatial analysis of trends in extreme precipitation events in high-resolution climate model results and observations for Germany, *J. Geophys. Res.*, *114*, D12113, doi:10.1029/2008JD010652.

1.   Introduction

[2]   Many parts of the world can cope with a moderate change in mean climate or even benefit from altered climatic conditions. This is true in particular for economically strong regions like Central Europe or North America. The most damaging effects of a warming atmosphere however emanate from extreme climatic conditions that could possibly become more prominent in the future. Extreme precipitation events causing floods or landslides, heat waves or long dry periods can implicate large losses in terms of infrastructure and other goods, constitute a danger to human lives and have manifold impacts on ecosystems, agriculture and other industries and even cultural heritage.

[3]   It is therefore important to understand the relations and mechanisms that relate a general warming of the world due to increased levels of greenhouse gases in the atmosphere to extreme events that can arise on a local scale.

[4]   General theoretical considerations based on the Clausius-Clapeyron relation predict that upper quantiles in precipitation will increase more pronouncedly than mean precipitation [*Frei et al.*, 1998; *Pall et al.*, 2007]. According to this reasoning, heavy precipitation is estimated to increase by about 6.5% per degree increase in mean temperature [*Allen and Ingram*, 2002]. Such a relation has been verified for global climate model simulations mainly in the midlatitudes and regions in which the nature of the ambient flow changes little [*Kharin et al.*, 2007]. Regions with strong advection of moist maritime air by southwesterly low-level flow toward topographically structured terrain, like encountered in Western Europe, are particularly affected.

[5]   In general, dynamical studies show that regional climate models (RCMs) are able to reproduce the conditions and physical mechanisms that lead to extreme precipitation events [*Frei et al.*, 1998; *Semmler and Jacob*, 2004; *Boroneant et al.*, 2006]. Broad statistical investigations of the performance of RCMs with respect to extreme events

---

[1]Regional Climate Modeling, Max Planck Institute for Meteorology, Hamburg, Germany.

have been performed in the context of the MERCURE and PRUDENCE projects [*Frei et al.*, 2003; *Frei et al.*, 2006; *Beniston et al.*, 2007]. Despite some biases, the RCMs were found to reproduce the prominent mesoscale features of heavy precipitation, and biases in extremes were comparable or even smaller than the ones for wet day intensity and mean precipitation. It should be noted that the resolution of the models in these studies was thereby considerably coarser than in the present paper (about 50 km compared to 10 km), and the time series of observations shorter.

[6] With respect to extreme precipitation events the scale of the analysis is important, especially regarding topographically induced precipitation [*Kunz and Kottmeier*, 2006; *Zängl*, 2007]. Although even global climate models, despite their coarse resolution, show some apparent skill in reproducing extreme event characteristics on large scales [*Sillmann and Röckner*, 2008; *Randall et al.*, 2007; *Tebaldi et al.*, 2006], heavy rainfall events are often characterized by their local nature and cannot be resolved by global climate models with typical resolutions of 100 to 250 kilometers. In our study we concentrate on a relatively small geographic area, Germany, and investigate very high-resolution regional climate model results as well as quality-checked high-resolution observation data from the German Weather Service.

[7] Considering the relatively short time period covered by the observational record, the method employed in the extreme value analysis is of utter importance [*Frei and Schär*, 2001]. In our study we do not only investigate extreme event statistics on the basis of time series related to separate locations, but also compare the spatial characteristics of trends in heavy rainfall events in observations and climate model simulations, and identify changes of these spatial patterns in scenario simulations. For this purpose, techniques of extreme value statistics are combined with methods of geostatistical data analysis. Special care is dedicated to the estimation of uncertainties in the derived results. For this purpose estimates based on asymptotic theory as well as a Bayesian approach (using Markov chain Monte Carlo sampling) are considered.

[8] We do not develop spatial models of extreme events themselves [*Schlather and Tawn*, 2003; *de Haan and Pereira*, 2006; *Mendes et al.*, 2009; *Cooley et al.*, 2008; *Sang and Gelfand*, 2009], but estimate trends in a station-by-station analysis and formulate geostatistical models of these trends. This methodological procedure has been proposed by Smith (Trends in rainfall extremes, unpublished manuscript, 1999) (see also the studies by *Smith* [2003] and *Meier* [2004] for an exploratory study). It is appropriate for the focus of our analysis and has some distinct advantages. It allows for dealing with large data sets, as in our situation, considering threshold exceedances of century-long time series for more than 4000 locations.

[9] We turn to a more detailed description of the content of the present paper. Section 2 is on model and methods. In section 2.1 the analyzed observations and climate model simulations are introduced. Section 2.2 describes the statistical extreme value analysis as carried out in the present work. It proceeds in two main steps: first the stations (gridded observations as well as climate model results) are analyzed separately in a station-by-station analysis (section 2.2.1). Trends in extreme precipitation events are estimated for each such station. In a second step (section 2.2.2) geostatistical models are introduced to investigate the spatial characteristics of the estimated trends.

[10] Section 3 presents the results of the analysis. Again, there are two parts: the results of the station-by-station analysis (section 3.1) and the spatial modeling of estimated trends in heavy precipitation events (section 3.2). In the first part maximum likelihood parameter estimation is applied (section 3.1.1) as well as Bayesian parameter estimation using Markov chain Monte Carlo sampling (section 3.1.2).

[11] A discussion section, section 4, concludes the paper.

2. Data and Methods

2.1. Climate Model Simulations and Observations

[12] We analyze climate model simulations performed with the three-dimensional hydrostatic mesoscale model REMO [*Jacob*, 2001; *Jacob et al.*, 2007] at a horizontal grid spacing of 10 km. This resolution is at the high end of current regional long-term climate simulations.

[13] Figure 1a shows the whole climate model domain of the 10 km simulations (colored area), Figure 1b depicts Germany, the focus area of the present study, with selected cities.

[14] In the model, the simulation of precipitation is divided into a large-scale cloud scheme accounting for clouds developing on scales that can be described directly by the prognostic variables of the model, and in a subgrid scale scheme (also called convection scheme) for clouds on smaller scales. The stratiform cloud scheme in REMO is taken from the global climate model ECHAM4 and based on the approach of Sundqvist [*Sundqvist*, 1978], cumulus convection is parameterized by a mass flux scheme following Tiedtke [*Tiedtke*, 1989] with some modifications.

[15] For penetrative convection, unlike in the procedure recommended by Tiedtke, the mass flux at cloud base is specified following an adjustment closure proposed by Nordeng [*Nordeng*, 1994], linking the cloud base mass flux to the available convective potential energy (CAPE), which should be removed by convection in a characteristic time τ. This characteristic time depends on the horizontal resolution of the model and was set to 15 minutes for the considered simulations. For shallow and midlevel convection a moisture closure is used.

[16] The convection scheme is coupled to the large-scale cloud scheme by handing over the convective cloud liquid water detrained in the updrafts to the large-scale cloud scheme. The formulation of the convection scheme restricts convective activity to one single time step and thus includes no memory in the convection scheme, which would allow convective clouds to develop on a longer timescale.

[17] For the present study, five simulations are analyzed. In two so-called control simulations spanning the years 1961 to 2004, the regional climate model REMO at a resolution of 10 km is nested in two different REMO simulations with a resolution of 50 km. The REMO 50 km resolution simulations in turn were forced at the lateral boundaries by different runs of the global atmosphere-ocean general circulation model ECHAM5/MPI-OM. The coupled ECHAM5/MPI-OM simulations use observed anthropogenic concentrations for CO_2, CH_4, N_2O, CFCs, O_3, and sulphate, and were initialized by a preindustrial control
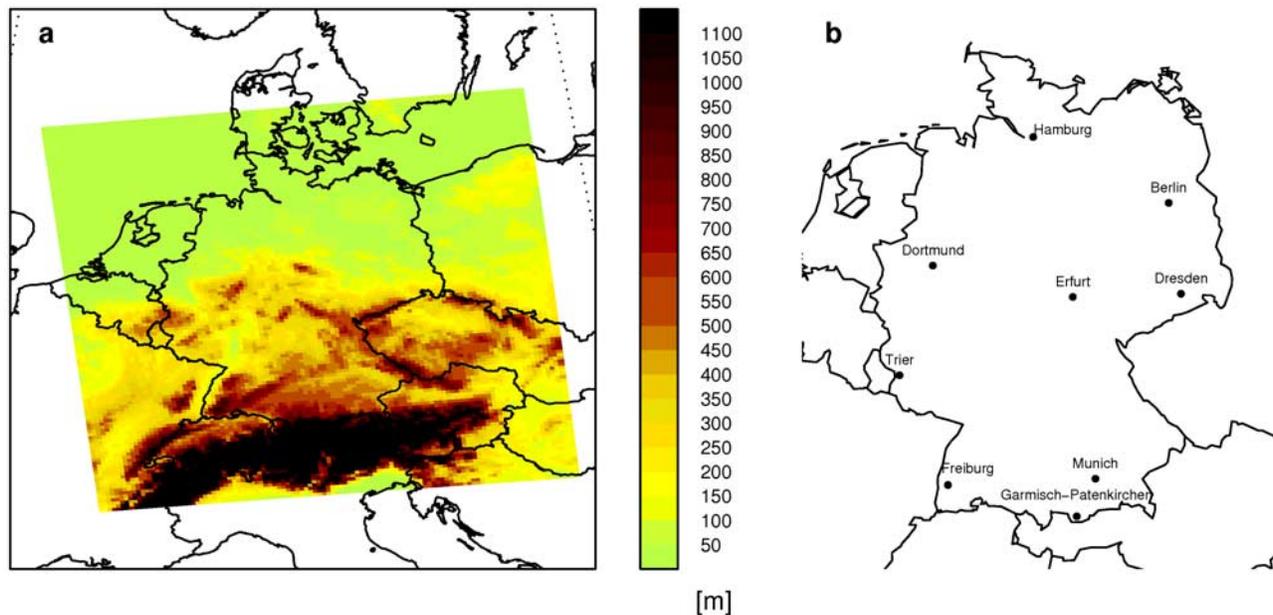
Figure 1. Colored area in Figure 1a shows the whole climate model domain. The orographic elevation as represented in the model is depicted. Figure 1b shows the focus area Germany with selected cities.

run at two different initial times (resulting in different initial conditions). These simulations were performed for the Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report [*IPCC*, 2007].

[18] Moreover, analogously produced scenario simulations for the period 2001 to 2100 according to the three SRES emission scenarios B1, A1B, and A2 are considered. All the REMO 10 km simulations were commissioned by the German Federal Environmental Agency [*Jacob et al.*, 2008].

[19] For the control period 1961 to 2004, the results of the regional climate model simulations are compared to daily rainfall observations provided by the German Weather Service (DWD). The observations consist of a gridded data set, which was constructed according to the REGNIE procedure developed by the DWD.

[20] More than 4000 rain gauge stations were used to derive a gridded precipitation data set for Germany with a resolution of 1 km by 1 km. The procedure goes in three steps: First, monthly means of station data are transferred to the 1 km grid by linear regression taking into account height, longitude, latitude, exposure (direction of the slope), and the amount of exposure (angle of the slope) as covariates. In the second step, these monthly fields are used as background fields for the construction of daily fields. Station values are attributed to the nearest grid point, while missing values are interpolated based on the deviation of station values from the monthly background fields using distance weighting. Finally, the daily time series are corrected for well-known systematic observation errors in rainfall measurements.

[21] To make the data comparable to the climate model output (which does not represent single stations, but area means corresponding to the model resolution), the 1 km by 1 km observation data set was aggregated to the 10 km to 10 km climate model grid by a simple averaging procedure.

[22] Of course, the considered data are affected by these steps of interpolation and aggregation. On the other hand, as already mentioned, the results of a climate model at a specific grid location cannot be interpreted as representing an observation made at that particular geographic location. The gridded data set used in the present study is the most comprehensive available precipitation data set in Germany and is at the same time suitable for being compared to regional climate model simulations.

2.2. Statistical Methods

[23] In the following, the term "station" will refer to locations of (gridded) observations as well as model grid points, that is, we do not distinguish between observations and model simulations because the statistical analysis was applied to all time series in the same way.

[24] The statistical extreme value analysis of the data proceeds in two main steps. First the daily time series of the stations are examined independently. We estimate trends of extreme precipitation events over the different periods. For the estimation of uncertainties, estimates based on asymptotic theory as well as a Bayesian approach (using Markov chain Monte Carlo sampling) are considered. In a second step, the spatial dependence of these trends is analyzed and the adequacy of different geostatistical models investigated.

[25] The main advantages of this approach are that it is focused on trends in rainfall extremes, which are the main interest in climate change studies. Moreover, since we proceed in two steps (and do not try to make at the same time inference for the geostatistical model of trends and the extreme value model used to estimate the trends) the statistical inference becomes more tractable computationally and reliable when dealing with large data sets. These technical aspects will be touched upon in the discussion section at the end of the paper.
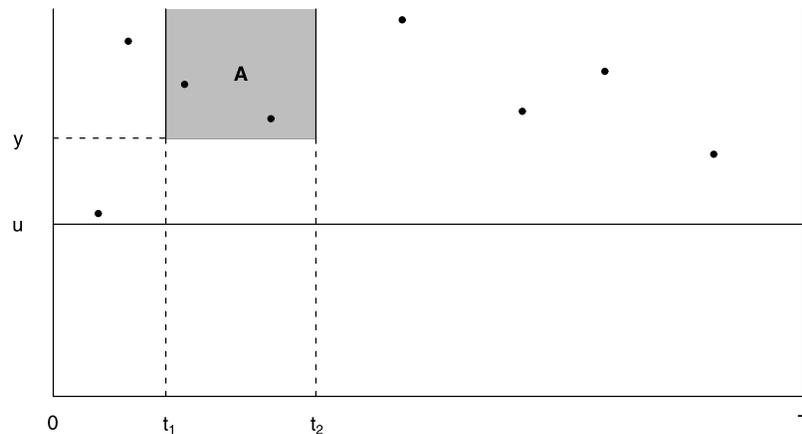
Figure 2. Illustration of the point process representation. The points on the graph represent times t and values y of exceedances over the threshold u.

[26] The statistical analysis is implemented using the statistics software R [*R Development Core Team*, 2008].

2.2.1. Station-by-Station Analysis
2.2.1.1. Statistical Model of Threshold Exceedances

[27] The daily time series of the stations are first analyzed separately. Exceedances of a (station-dependent) threshold u are modeled by a nonhomogenous Poisson point process.

[28] A Poisson point process on $A \subset \mathbb{R}$ with intensity $\Lambda(A)$ is represented by a random variable $N(A)$, the number of points in A, with the following two properties: (1) for all $B \subset A$, $N(B) \sim \text{Pois}(\Lambda(B))$, (2) the random variables $N(B)$ and $N(C)$ are independent for disjoint subsets B and C.

[29] Figure 2 illustrates the point process representation of exceedances y over the threshold u. Only exceedances of the threshold u are modeled by the point process. The distribution of the number of points $N(A)$ in A implies a probabilistic measure of intensity and frequency of threshold exceedances (i.e., heavy rainfall events in the context of the present paper). More precisely, the expectation $\mathbf{E}[N(A)] = \Lambda(A)$.

[30] The connection to traditional extreme value statistics [*Leadbetter et al.*, 1983; *Davison and Smith*, 1990; *Embrechts et al.*, 1997; *Coles*, 2001] is as follows. We first review the main theorem in this context, the Extremal Types Theorem I, following *Coles* [2001]. Let $X_1, X_2, \ldots \sim F$ identically and independently, and let $M_n = \max\{X_1, \ldots, X_n\}$. If there exist sequences of constants $\{a_n > 0\}$ and $\{b_n\}$ such that

$$\Pr\left\{\frac{M_n - b_n}{a_n} \leq z\right\} \to G(z) \qquad \text{as} \quad n \to \infty \qquad (1)$$

for a nondegenerate distribution function G, then G is a member of the generalized extreme value (GEV) family

$$G(z) = \exp\left\{-\left[1 + \xi\left(\frac{z - \mu}{\sigma}\right)\right]^{\frac{-1}{\xi}}\right\}, \qquad (2)$$

defined on $\{z \mid 1 + \xi(z - \mu)/\sigma > 0\}$, where $-\infty < \mu < \infty$, $\sigma > 0$, and $-\infty < \xi < \infty$.

[31] This means that even if the (arbitrary) distribution F of the original data (in our case, the rainfall data) is not known, according to the above theorem the maxima of long sequences (i.e., the extreme values) are distributed according to the family of distribution functions (2).

[32] By this asymptotic result, the GEV family of distribution is suitable for modeling the distribution of maxima of long sequences (e.g., yearly maxima of daily rainfall amounts). The family has three parameters: the location parameter μ, the scale parameter σ, and the shape parameter ξ, which determines the tail of the distribution. Commonly, three cases are distinguished: $\xi > 0$ (Fréchet), $\xi = 0$ (Gumbel), and $\xi < 0$ (Weibull). The Weibull distributions are short tailed, the densities of Fréchet and Gumbel distributions show polynomial and exponential decay, respectively.

[33] A similar, ''Central-Limit-Theorem''-like proposition exists for threshold exceedances: the Extremal Types Theorem II. Under the same assumptions of the Extremal Type Theorem I and for large enough threshold u, the distribution function of $(X - u)$, conditional on $X > u$, is approximately

$$H(y) = 1 - \left(1 + \frac{\xi y}{\tilde{\sigma}}\right)^{\frac{-1}{\xi}} \qquad (3)$$

defined on $\{y \mid y > 0$ and $(1 + \xi y/\tilde{\sigma}) > 0\}$, where $\tilde{\sigma} = \sigma + \xi(u - \mu)$ and X is any term in the sequence X_i. The distributions defined by (3) are called Generalized Pareto distributions (GPD).

[34] The statement can be made more precise [*Leadbetter et al.*, 1983; *Embrechts et al.*, 1997], justifying (3) as a limiting distribution as u increases. Again, the theorem states that extreme values, in the sense of exceedances over a high threshold, are distributed according to the family of distribution functions (3), independently of the (perhaps unknown) distribution function F of the original data.

[35] We now return to the nonhomogenous Poisson process representation of threshold exceedances, and its connection to the Extremal Types Theorems. Under the same assumptions of the Extremal Type Theorem I, the sequence of point processes

$$N_n = \left\{\left(\frac{i}{n+1}, \frac{X_i - b_n}{a_n}\right) \mid i = 1, \ldots, n\right\} \qquad (4)$$

converges on regions of the form $(0, 1) \times (u, \infty)$, for any $u > z_-$, to a Poisson process, with intensity measure on $A = [t_1, t_2] \times (z, z_+)$ given by

$$\Lambda(A) = (t_2 - t_1)\left[1 + \xi\left(\frac{z - \mu}{\sigma}\right)\right]^{\frac{-1}{\xi}}, \qquad (5)$$

where z_- and z_+ are the lower and upper endpoints of the distribution function G. In particular, the parameters in equation (5) are the same as the parameters of the GEV family of distributions (2).

[36] In other words, if we just consider the exceedances over the (large enough) threshold u, these exceedances can be modeled by a Poisson point process with intensity measure (5).

[37] It can be shown that the likelihood function of the described Poisson point process differs from the likelihood of the Extremal Types Theorem II only by a constant. Thus parameter estimates for the two models are the same. However, the point process representation allows for a conceptually more satisfactory formulation of nonstationary models as used in the present paper, since the parameters of the Poisson point process do not depend on the threshold u, contrary to the parameters of the Extremal Types Theorem II. Moreover, the likelihood function in the point process representation is very easily adapted to handle missing values.

2.2.1.2. Declustering

[38] The Poisson point process model introduced in section 2.2.1.1 assumes the threshold exceedances to be independent. We assume that dependent extreme precipitation events (belonging to the same precipitation front) are separated by at most one day, and that two events are independent if a day with no rainfall is in between them. Accordingly, declustering is applied as follows. Two threshold exceedances are considered to belong to the same cluster if no more than one day lies in between them, except if at this day between no precipitation has occurred. We then treat each one of these clusters as one extreme precipitation event and consider the maximum of the exceedances within each cluster as the corresponding rainfall amount [*Davison and Smith*, 1990].

[39] Inspection of rainfall observations and simulations reveals that clusters are composed rarely of more than three days.

2.2.1.3. Threshold Selection

[40] If the random variable Y is distributed according to (3) with $\xi < 1$, then

$$\mathbf{E}[Y] = \frac{\tilde{\sigma}}{1 - \xi}.$$

[41] Thus, if the conditional exceedances $Y = X - u$ were distributed exactly according to (3), then

$$\mathbf{E}[X_i - u \mid X_i > u] = \frac{\sigma + \xi(u - \mu)}{1 - \xi}.$$

[42] In particular, this expectation is linear in u. This motivates the so-called mean excess plot, defined as

$$\left\{\left(u, \frac{\sum_{i=1}^n (X_i - u)1_{\{X_i > u\}}}{\sum_{i=1}^n 1_{\{X_i > u\}}}\right)\right\}, \qquad (6)$$

depicting the averaged exceedances as a function of the threshold u. The threshold u is then selected at the value, from where on the graph (6) approximately represents a straight line.

[43] However, other considerations have to be taken into account as well. First, the common bias-variance dilemma: the higher the threshold, the better approximation (3), but the less data are available for parameter estimation. Second, in the present study we want to investigate spatial dependencies of trends in heavy precipitation events over Germany. The higher the threshold, the less spatial dependence is visible in the data. We therefore need to make a reasonable threshold selection considering these two aspects as well.

[44] For the present work we defined the threshold to be the 95% quantile of the time series of daily precipitation totals for each station. This is a reasonable choice for our purposes considering the findings of Smith (unpublished manuscript, 1999) as well as the mean excess plots (see section 3.1.1). Since in the computation of the quantile we include all days (not just days with rainfall), the 95% quantile is not an extremely high threshold. Our analysis does therefore not concentrate on the most extreme events only, but on incidents with high daily precipitation totals.

[45] Other ways of dealing with temporal dependence in the data exist like the approach to declustering by *Ferro and Segers* [2003] or other methods which avoid the need for declustering altogether [*Smith et al.*, 1997; *Fawcett and Walshaw*, 2006, 2007].

2.2.1.4. Likelihood Function and Time Dependence

[46] For each station, we now only consider the declustered exceedances and disregard the rest of the data. Let T_j be the time of the jth exceedance and $\tilde{X}_j > u$ the value of the jth exceedance of the threshold u, $j = 1, \ldots, n_u$. Denote the whole time period by the interval $[0, T]$.

[47] $(T_j, \tilde{X}_j)$ is modeled as a two-dimensional Poisson point process (section 2.2.1.1). It can be shown that the likelihood function (for each station separately) is given by

$$L = \exp\left\{-\int_0^T \left[1 + \xi\left(\frac{u - \mu_t}{\sigma_t}\right)\right]_+^{\frac{-1}{\xi}} dt\right\} \prod_{j=1}^{n_u} \frac{1}{\sigma_{T_j}}$$

$$\cdot \left[1 + \xi\left(\frac{\tilde{X}_j - \mu_{T_j}}{\sigma_{T_j}}\right)\right]_+^{\frac{-1}{\xi} - 1} \qquad (7)$$

where $a_+ = a$ if $a > 0$ and zero otherwise.

[48] As in Smith (unpublished manuscript, 1999), the temporal dependence of the parameters μ_t and σ_t is modeled as

$$\mu_t = \mu_0 e^{g(t) + s(t)}, \qquad \sigma_t = \sigma_0 e^{g(t) + s(t)}, \qquad (8)$$

where μ_0 and σ_0 are constants. Note that the same time dependence is assumed for μ_t as well as σ_t. This is due to

our desire to limit the number of parameters in the model. Estimating trend parameters for μ_t and σ_t separately would be an option, however this would increase the dimensionality of the problem. The shape parameter ξ is assumed to be constant in time.

[49] The function $g(t) = \beta_1 t$ models a linear trend, while

$$s(t) = \sum_{p=1}^{P} \beta_{2p} \cos(\omega_p t) + \beta_{2p+1} \sin(\omega_p t) \qquad (9)$$

accounts for seasonal effects. In the present study we choose $P = 2$, but will use $P = 1$ as an intermediate step in parameter estimation (see section 3.1.1).

[50] The linear trend parameter β_1 will be the main focus of the present work. Primarily β_1 is the coefficient of a linear trend in the parameters of the family of distributions (2) (or, equivalently, (3)) of extreme values (i.e., heavy precipitation events in our case). Unfortunately, these parameters do not have a simple interpretation. They do not relate directly to the mean or standard deviation of the corresponding extreme value distributions, for instance. However, let $q_{T^{\text{ini}}}$ be the probability that the value y will not be exceeded during the period $[T^{\text{ini}}, T^{\text{ini}} + 1]$. An easy calculation shows that under the above assumptions $q_{T^{\text{ini}}+1}(y) = q_{T^{\text{ini}}}\left(\frac{y}{e^{\beta_1}}\right)$. In other words, if $y_{T^{\text{ini}}}$ denotes the q quantile in the period $[T^{\text{ini}}, T^{\text{ini}} + 1]$, then

$$y_{T^{\text{end}}}(q) = e^{\beta_1(T^{\text{end}} - T^{\text{ini}})} y_{T^{\text{ini}}}(q), \qquad (10)$$

that is, β_1 can be understood as an inflation factor of the quantiles of the extreme value distributions (i.e., the distributions of the threshold exceedances, or, put differently, the heavy rainfall events). Since this provides a straightforward interpretation of β_1, we will, in the presentation of our results, sometimes refer to this view on β_1.

2.2.1.5. Parameter Estimation

[51] Parameter estimation is based on the negative log likelihood

$$-\log L = \sum_{j=1}^{n_u} \left\{ \log(\sigma_{T_j}) + \left(\frac{1}{\xi} + 1\right) \log\left(\left[1 + \xi\left(\frac{\tilde{X}_j - \mu_{T_j}}{\sigma_{T_j}}\right)\right]_+\right) \right\}$$
$$+ \int_0^T \left[1 + \xi\left(\frac{u - \mu_t}{\sigma_t}\right)\right]_+^{-\frac{1}{\xi}} dt \qquad (11)$$

[52] The integral in (11) has to be approximated numerically:

$$\int_0^T \left[1 + \xi\left(\frac{u - \mu_t}{\sigma_t}\right)\right]_+^{-\frac{1}{\xi}} dt \approx h \sum_t \left[1 + \xi\left(\frac{u - \mu_t}{\sigma_t}\right)\right]_+^{-\frac{1}{\xi}} \qquad (12)$$

where $h = \frac{1}{365.25}$, the number of days per year.

[53] We mostly rely on maximum likelihood parameter estimation by minimizing the negative log likelihood (11). Uncertainty estimates are obtained by using the following result in asymptotic theory: Let $X_1, \ldots, X_n$ be iid distributed random variables with density $f(x; \theta)$. Let $\hat{\theta}_{\text{MLE}} \in \mathbb{R}^d$ denote

the maximum likelihood estimate of the parameter vector $\theta \in \mathbb{R}^d$. Then under certain regularity conditions for $f(x; \theta)$

$$\hat{\theta}_{\text{MLE}} \approx \mathcal{N}\left(\theta, n^{-1} J(\theta)^{-1}\right)$$

for large n. Here $J(\theta)$ denotes the Fisher information matrix given by

$$J(\theta)_{ij} = \mathbf{E}\left[-\frac{\partial^2}{\partial\theta_i \partial\theta_j} \log(f(X_1; \theta))\right].$$

[54] In practical applications (since $f(X_1; \theta)$ is not known), for $nJ(\theta)$ the approximation

$$I(\theta) = -\frac{\partial^2}{\partial\theta_i \partial\theta_j} \log(L(X_1; \theta))$$

is used. The estimated standard errors of the components of $\hat{\theta}_{\text{MLE}}$ are then given by

$$\widehat{\text{se}}(\hat{\theta}_{i,\text{MLE}}) = \sqrt{(I(\hat{\theta}_{\text{MLE}})^{-1})_{i,i}} \qquad (13)$$

[55] In one case of our analysis, investigating the REMO simulations according to the SRES scenario A1B, we will compare the results of the parameter estimates based on maximum likelihood theory with Bayesian parameter estimates (section 3.1.2). In the Bayesian approach, parameters are considered to be random variables. A priori assumptions on the distribution of the parameter vector θ are made. This a priori distribution $p(\theta)$ is then updated, taking the data y into account, according to Bayes' theorem, resulting in a posterior distribution $p(\theta \mid y)$:

$$p(\theta \mid y) \propto L(y; \theta) \cdot p(\theta) \qquad (14)$$

where $L(y; \theta)$ is the likelihood function (the conditional probability density of the observations y given the parameters θ).

[56] Posterior distributions for the parameters are derived using a random-walk Metropolis-Hastings Markov chain Monte Carlo algorithm [Gelfand and Smith, 1990; Smith and Roberts, 1993]. Maximum likelihood estimates are compared to the mean of the marginal posterior distributions, and the standard errors $\widehat{\text{se}}(\hat{\theta}_{i,\text{MLE}})$ are checked against the standard deviations of the marginal posterior distributions. More details are given in section 3.1.2. In the present work we will essentially assume uniform prior distributions, which implies that the posterior distribution is equal to the likelihood function. This allows for comparing the uncertainty estimates resulting form asymptotic maximum likelihood theory with the standard deviations estimated by the Markov chain algorithm. Of course also, the Bayesian approach implies an approximation (due to the finite sample size), but in principle the full likelihood function is sampled.

[57] For the maximum likelihood estimation as well as the Bayesian parameter estimation, the choice of appropriate start values for the numerical estimation procedure is crucial. In the present paper we fitted the distribution (2) to the yearly maxima of daily precipitation amounts, assuming

constant values for the parameters. The R package ismev was taken advantage of for this purpose [*Coles and Stephenson*, 2006]. The resulting estimates for μ, σ, and ξ were then used as start values for μ_0, σ_0, and ξ as defined in (7) and (8).

2.2.1.6. Model Validation

[58] Suppose the random variable Y_t is distributed according to (3), then the so-called W statistics can be calculated:

$$W_t = \frac{1}{\xi_t} \log\left(1 + \xi_t \frac{Y_t}{\tilde{\sigma}_t}\right). \tag{15}$$

[59] We have

$$P[W_t \leq x] = P\left[Y_t \leq \frac{\tilde{\sigma}}{\xi_t}(e^{\xi_t x} - 1)\right] = 1 - e^{-x}.$$

Thus $W_t \sim \text{Exp}(1)$, that is, W_t is exponentially distributed with rate parameter equal to 1. For validating the statistical model of extreme precipitation events, we can insert the parameter estimates in (15) and plot the empirical quantiles against the quantiles of the standard exponential distribution as follows:

$$\left\{\left(-\log\left(1 - \frac{i}{n_u + 1}\right), w_{(i)}\right), i = 1, \ldots, n_u\right\}, \tag{16}$$

where $w_{(i)}$ denote the ordered realizations of W_{t_k}. If the statistical model is appropriate, the points in the plot will be arranged along the diagonal.

[60] The W statistics also allows for testing the independence of the declustered threshold exceedances by investigating the autocorrelation of the transformed sample $(w_i)_{i=1,\ldots,nu}$.

2.2.2. Spatial Analysis of Trends

[61] In the second step of the present work, we investigate the spatial dependence of trends in extreme precipitation events. These linear trends $\beta_1(s)$ (see section 2.2.1.4) are first estimated in the station-by-station analysis as outlined in section 2.2.1. Now we would like to formulate a spatial statistical model to capture the spatial dependence in the trends of the different stations. This two-step procedure was proposed by Smith (unpublished manuscript, 1999).

[62] The aim of this analysis is twofold. We would like to compare the climate model simulations with observations, and to contrast the present state of the climate with the future state in order to learn about the behavior of the climate system under changing greenhouse gas concentrations. The estimated parameters of the spatial statistical model will provide information on spatial dependence both in the climate model data and the observations. Since we already work with very high-resolution data, our goal is not to predict rainfall at missing locations, the more traditional application of geostatistical modeling.

2.2.2.1. Spatial Statistical Model for the Trends

[63] We model the (true) trends $\beta_1(s)$ in extreme rainfall events as realizations of a stationary Gaussian random field. Here s indicates the spatial location. In our example, s varies over the geographic area of Germany. The covariance matrix of $\hat{\beta}_1(s)$ is denoted by $\Sigma(s_j - s'_j)$.

[64] The assumption of stationarity in space is acceptable since we did not detect any dependence of the trends on longitude, latitude, elevation, or monthly rainfall amount in our trend estimates (see section 3.1.1). Note that although we are modeling trends in extreme events, the trends themselves are not expected to show extreme value behavior.

[65] Let $\hat{\beta}_1(s)$ denote the estimate of the trend at the location s resulting from the analysis described in section 2.2.1. This estimate is assumed to be the true trend $\beta(s)$ plus an error term:

$$\hat{\beta}_1(s_j) = \beta_1(s_j) + \varepsilon(s_j), \quad j = 1, \ldots, n \tag{17}$$

[66] We assume $(\varepsilon(s_1), \ldots, \varepsilon(s_n))$ to be independently normally distributed with zero expectation and equal variance τ^2. This is a simplifying assumption, but it is difficult to establish a covariance model for the climate variability that affects the true trends due to the high dimensionality of the problem. A model formulation that takes the spatial correlation of the noise term into account would be desirable.

[67] Thus, with the assumption of independent climate noise, we have

$$\hat{\beta}(s_j)_{j=1,\ldots,n} \sim \mathcal{N}(m, \Sigma(s_j - s'_j) + \text{Diag}(\tau^2)), \tag{18}$$

where m is supposed to be a constant, and $\text{Diag}(\tau^2)$ is a $n \times n$ matrix with τ^2 in the diagonal, and zero in all other entries.

[68] For the covariance matrix Σ we consider a flexible parametric isotropic model, the Matérn family of covariances:

$$\text{Cov}(\beta_1(s_i) - \beta_1(s_j)) = \sigma^2 \frac{1}{2^{\nu-1}\Gamma(\nu)}$$
$$\cdot \left(2\sqrt{\nu}\frac{|s_i - s_j|}{\rho}\right)^\nu K_\nu\left(2\sqrt{\nu}\frac{|s_i - s_j|}{\rho}\right). \tag{19}$$

[69] Here $\Gamma(\cdot)$ is the Gamma function and $K(\cdot)$ the modified Bessel function of the third kind of order v, σ^2 (the partial sill) and ρ (the range) are form parameters, v is a shape parameter which determines the behavior of the covariance for $|s_i - s_j|$ small.

[70] For the random field of trends $\beta_1(s)$ we can define the function

$$\gamma_{\beta_1}(h) = \frac{1}{2}\text{Var}(\beta_1(s + h) - \beta_1(s)), \tag{20}$$

the so-called semivariogram. The isotropic property of the Matérn family of covariances implies $\gamma_{\beta_1}(h) = \gamma_{\beta_1}(|h|)$.

[71] Since the true trends β_1 are not known, we have to relate $\gamma_{\beta_1}(h)$ to the empirical semivariogram $\gamma_{\hat{\beta}_1}(h)$. Because of the independence of $\varepsilon(\cdot)$ and $\beta_1(\cdot)$ it follows

$$\frac{1}{2}\text{Var}(\hat{\beta}_1(s + h) - \hat{\beta}_1(s)) = \frac{1}{2}\text{Var}(\beta_1(s + h) - \beta_1(s))$$
$$+ \frac{1}{2}\text{Var}(\varepsilon(s_i)) + \frac{1}{2}\text{Var}(\varepsilon(s_j))$$
$$= \frac{1}{2}\text{Var}(\beta_1(s + h) - \beta_1(s)) + \tau^2 \tag{21}$$

[72] Or in other words,

$$\gamma_{\hat{\beta}_1}(h) = \gamma_{\beta_1}(h) + \tau^2 \qquad (22)$$

[73] In order to analyze the empirical semivariogram of the trend estimates, we can plot $\frac{1}{2}(\hat{\beta}_1(s) - \hat{\beta}_1(s'))^2$ against $|s - s'|$. However, such a diagram would contain $n(n-1)/2$ points, where n is the number of locations. It is therefore common to introduce classes $T(h)$ which contain pairs of observations (in our case, pairs of estimated trends) with similar distances in between them. The mean of $\frac{1}{2}(\hat{\beta}_1(s) - \hat{\beta}_1(s'))^2$ for $|s - s'| \in T(h)$ is then plotted against $|h|$.

2.2.2.2. Parameter Estimation for the Spatial Model

[74] In order to estimate the parameters σ^2, ρ and v of the covariance model (19), as well as the "nugget effect" τ^2, the variance of the error term in (17), two different methods are applied: parameter estimation by ordinary least squares fitting of the theoretical model to the empirical semivariogram, and by maximum-likelihood estimation. In principle also a Bayesian parameter estimation would be possible at this stage. However, this proved not to be feasible computationally considering the high dimensionality of the problem. The sample size of the posterior distributions would be too small to allow for reliable parameter estimation.

3. Results

3.1. Station-by-Station Analysis

3.1.1. Maximum Likelihood Estimates

[75] As already mentioned in section 2.2.1.5, starting values for the maximum likelihood parameter estimation were obtained by fitting the distribution (2) to the yearly maxima of daily precipitation amounts, assuming constant values for the parameters. The resulting estimates for μ, σ, and ξ were then used as start values for μ_0, σ_0, and ξ as defined in (7) and (8).

[76] The parameter estimation then proceeded in two further steps. First parameter estimation was performed for the reduced model with $P = 1$ in equation (9). In a second step the resulting estimates were used as starting values for the parameter estimation of the full model with $P = 2$. The quasi-Newtonian Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm was applied for the minimization of the negative log likelihood.

[77] Figure 3 shows the estimated linear trends $\hat{\beta}_1$ in terms of increase [%] in extreme quantiles (compare section 2.2.1.4) over the considered time periods for observations, the two REMO control runs and the different scenario simulations.

[78] The observations show positive trends mainly in the southern part of Germany and at some sites in Western Germany, and no clear trends in the northeastern part. This is consistent with results of *Trömel and Schönwiese* [2007] derived for trends in monthly mean precipitation.

[79] The first REMO control simulation (Figure 3b) also shows positive trends for the South and no trends or even negative values for $\hat{\beta}_1$ in the eastern parts of Germany. The (spatial) mean of $\hat{\beta}_1$ * 1000 is -0.14, which implies a change of -0.6% over the 44 years (observations: 1.02, implying an increase of 4.6%), and the variance is 3.4 (observations: 2.6). That is, there are more negative trends

than positive ones in the control run (vice versa in the observations), and the variability is larger.

[80] It has to be stressed that the control simulation does not necessarily reflect the historical climate of the time period 1961 to 2004, but represents just one possible realization of a similar climate (with the same greenhouse gas concentration). Because of natural variability the control climate is expected to differ from the observations and such differences do not necessarily imply model deficiencies.

[81] To investigate this aspect, a second REMO control run for the same time period was analyzed (Figure 3c). Here the trends are even more negative (the mean value of $\hat{\beta}_1$ * 1000 is -0.94, implying a change of -4.1%), positive trends show up mainly in Southern Germany. This indicates that the climate model tends to underestimate the trends. Negative values dominate in both control simulations, while in the observations trends are mostly positive. The reason for this underestimation is unclear and will need further investigation.

[82] In the moderate B1 scenario positive trends start to dominate also in the climate model simulation, especially in Southern Germany. This becomes more pronounced in the A1B scenario (the mean of $\hat{\beta}_1$ * 1000 amounts to 0.51, implying a change of 5.25% over the 100 year period), and in the A2 scenario positive values prevail in all regions with strongest trends in the South and in central parts of Germany. Note that the trends are relatively small, mostly below 15%. The increase in extreme quantiles in the medium range A1B emission scenario, over the time span of 100 years, is similar in magnitude to the increase estimated for the observations (which cover a time period of 44 years). The temperature increase in the case of the A1B scenario is 2.5 to 4°C with 4°C in Southern Germany, 3 to 3.5°C in Central Germany including the eastern state of Brandenburg, and 2.5 to 3°C in the North. Thus the spatial pattern of the increase in rainfall extremes is consistent with the Clausius-Clapeyron relation, but the magnitude estimated by this thermodynamic consideration is, in the case of the scenario simulations, larger than the one predicted by the regional climate model. In the case of the observations, however, it agrees rather well with the data record.

[83] Figure 4 shows the estimated shape parameter $\hat{\xi}$ (for the GPD family of distributions, see (3)) for observations, REMO control and scenario simulations. A value $\xi < 0$ implies that the distribution has an upper bound, the case $\xi = 0$ corresponds to the exponential distribution, a value $\xi > 0$ means that the distribution has no upper bound and is long tailed.

[84] In all panels, long-tailed distributions dominate. From the observations distributions with upper bounds are also inferred at some grid locations. It is remarkable that in the REMO results there are no major differences between the control runs and the various scenario simulations.

[85] Figure 5 depicts the t statistics for observations, REMO control and scenario simulations. Red areas have a t value larger than 1.96 and the corresponding trends can be considered to be significantly positive.

[86] There are no large regions which show a significantly positive trend neither in the observations nor in the REMO control simulations. The picture changes in the scenario runs. However, only in the scenario A2, which assumes
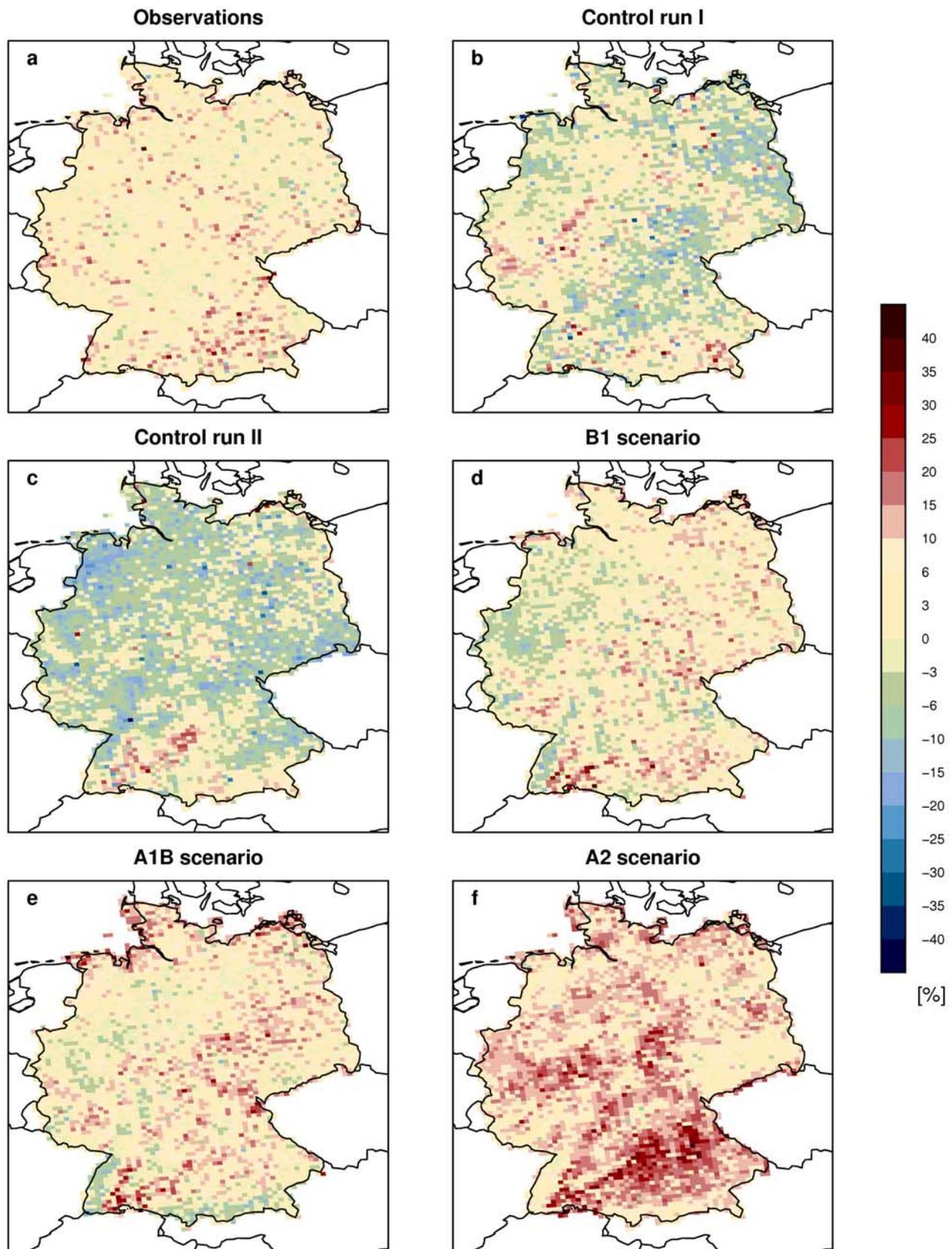
Figure 3.   Trends in heavy rainfall in terms of increase [%] in extreme quantiles over the considered time periods: (a) Observations (1961–2004); (b) REMO first control run (1961–2004); (c) REMO second control run (1961–2004); (d) REMO B1 scenario (2001–2100); (e) REMO A1B scenario (2001–2100); (f) REMO A2 scenario (2001–2100).

Observations

Control run I

Control run II

B1 scenario

A1B scenario
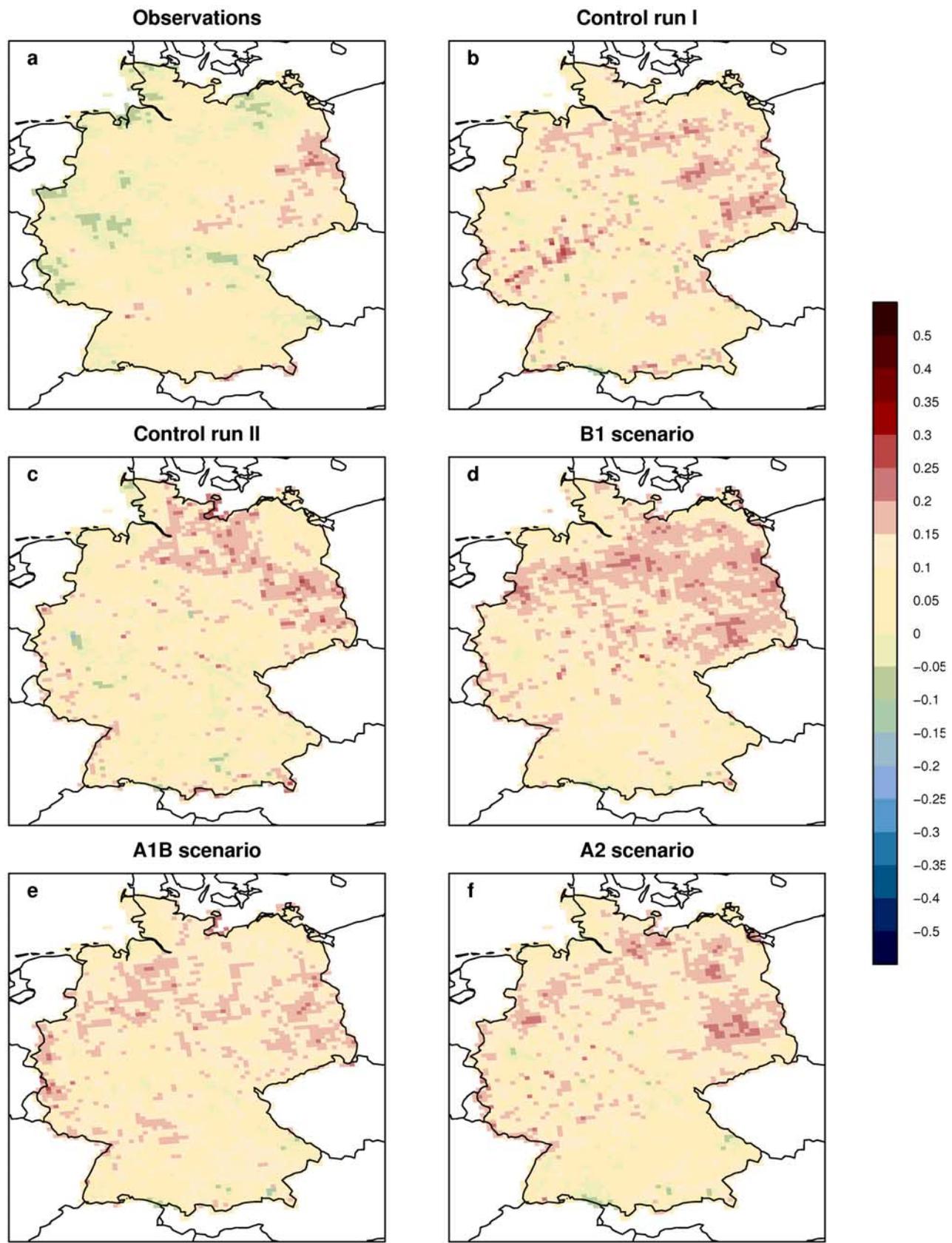
A2 scenario

Figure 4. Estimated value of the shape parameter ξ: (a) Observations (1961–2004); (b) REMO first control run (1961–2004); (c) REMO second control run (1961–2004); (d) REMO B1 scenario (2001–2100); (e) REMO A1B scenario (2001–2100); (f) REMO A2 scenario (2001–2100).
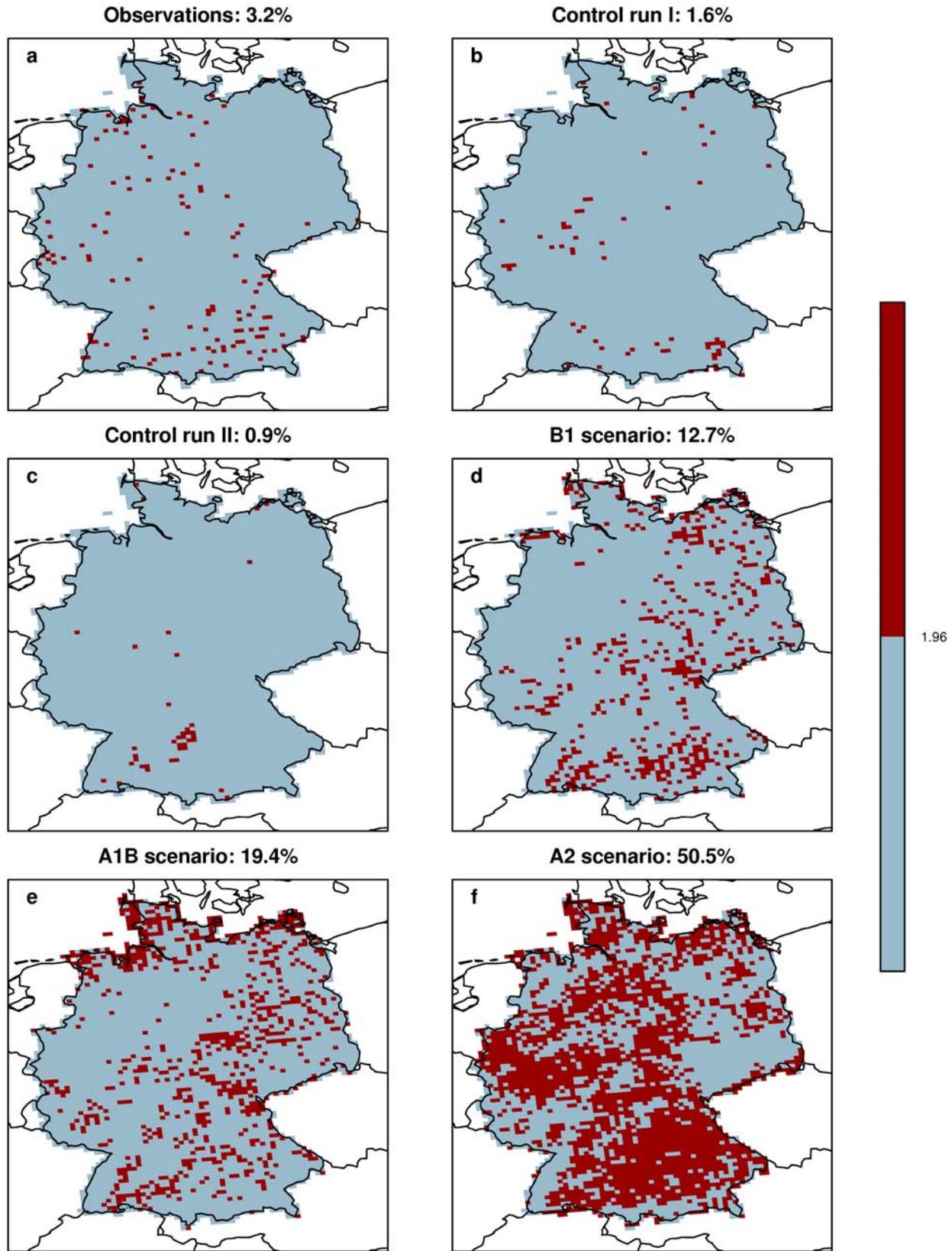
Figure 5. Values of the *t* statistics for the trend estimates are shown. Red areas have a *t* value larger than 1.96 and the corresponding trends can be considered to be significantly non zero. (a) Observations (1961–2004); (b) REMO first control run (1961–2004); (c) REMO second control run (1961–2004); (d) REMO B1 scenario (2001–2100); (e) REMO A1B scenario (2001–2100); (f) REMO A2 scenario (2001–2100). The headers indicate the percentage of significant grid boxes.
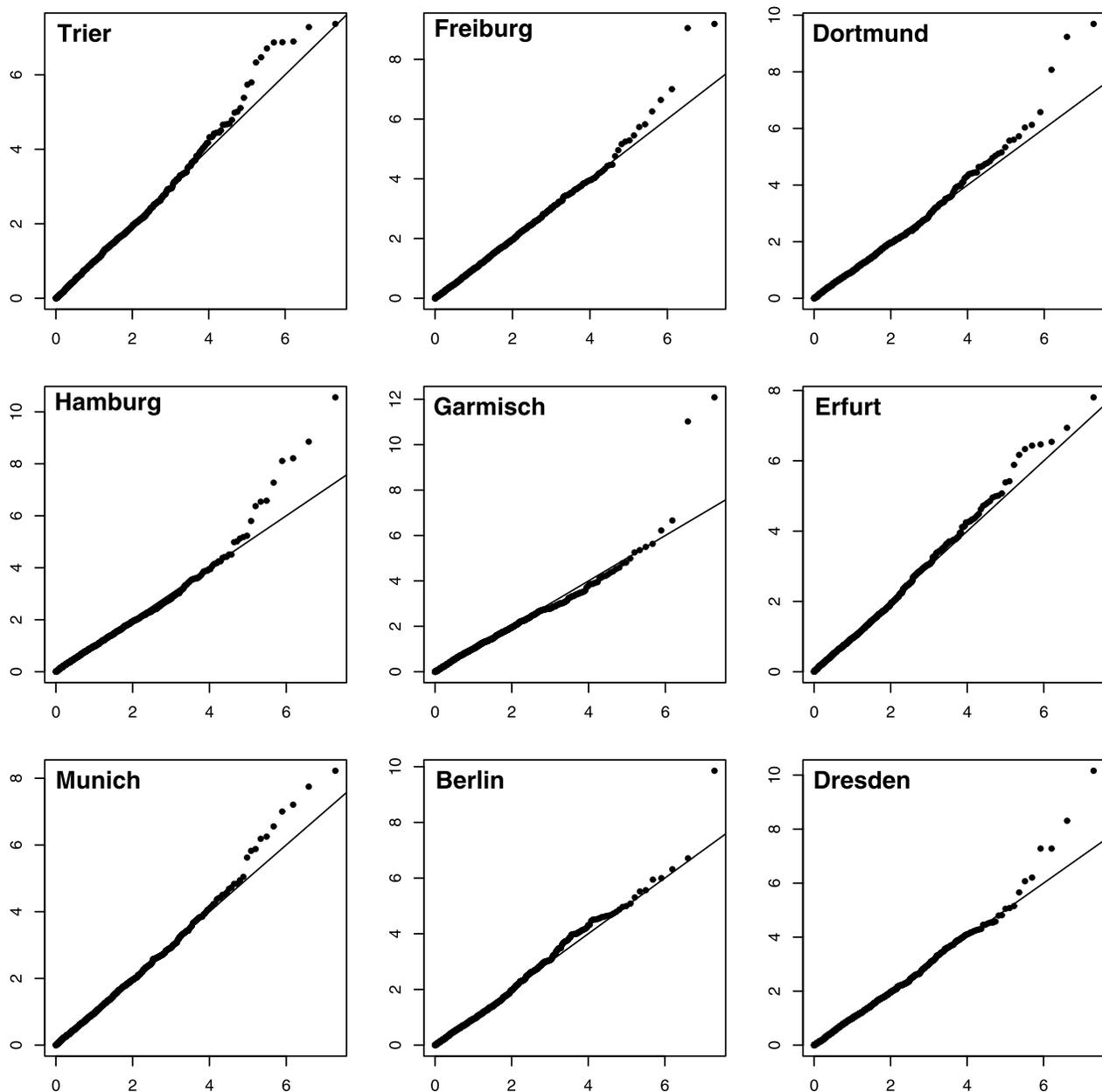
Figure 6. Quantile-quantile plots for the W statistics for the REMO A1B scenario simulation and different locations in Germany (calculated according to equation (16)).

strong increase in greenhouse gas concentrations, trends become significant in wide areas throughout Germany.

[87] Whether the significant trends at the North Sea coast in the A1B scenario is a robust feature which is reflected in reality cannot be decided with certainty. It might also be due to model deficiencies in representing characteristic features of the coastal atmospheric boundary layer and convective activity caused by the land-sea thermal contrast.

[88] For validating the statistical extreme value model we produced quantile-quantile plots (Figure 6) of the W statistics for the REMO A1B scenario simulation and different locations in Germany (see section 2.2.1.6 for explanations on the W statistics and Figure 1b for the location of the German cities).

[89] The statistical extreme value model fits the distribution of heavy precipitation events very well except for some very extreme incidents at the far end of the tails; they are generally underestimated by the statistical model. Plots for other grid locations and scenarios are similar. Also, the assumption of independence of declustered heavy precipitation events was checked on the basis of the W statistics. Plots of the empirical autocorrelations (not shown) confirm the independence assumption.

[90] Moreover, in order to check the selection of the thresholds u, mean excess plots (see section 2.2.1.3) were produced for the A1B scenario simulation and selected grid boxes (Figure 7). The plots generally look reasonable and there is no indication that the chosen thresholds (shown as dashed vertical lines in Figure 7) were too low. As discussed in section 2.2.1.3, the focus of our analysis is on estimating trends in days with large precipitation amounts and not on predicting very extreme events and their return times. Our
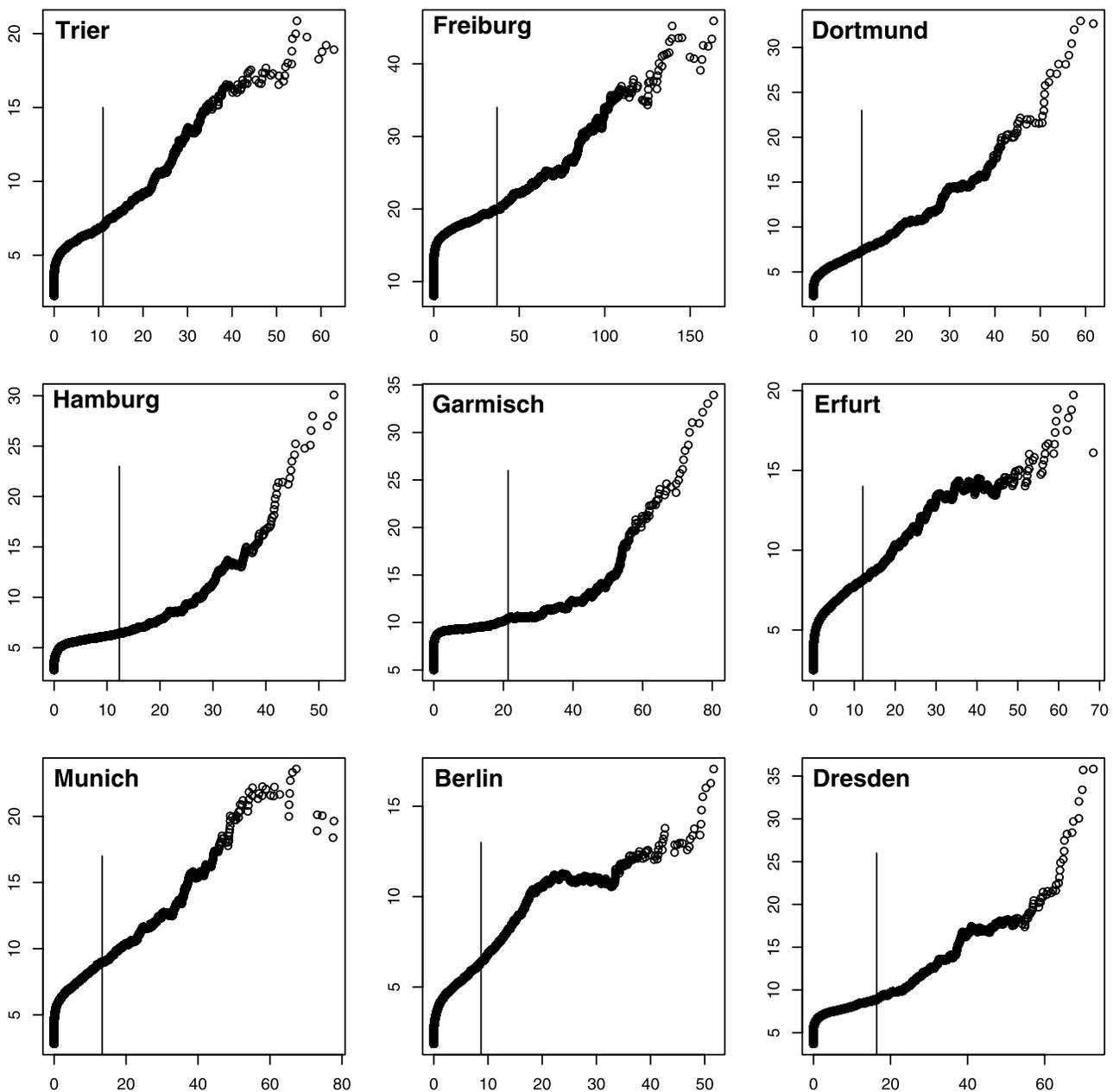
Figure 7. Mean excess plots for the REMO A1B scenario simulation and selected model grid locations.

model's underestimation of a few days with very heavy rainfall as indicated in Figure 7 is therefore not of major concern.

[91] Figure 8 shows smoothed trend estimates ($\hat{\beta}_1$ times 1000) for the first control run (upper left), the B1 scenario (upper right), the A1B scenario (lower left) and the A2 scenario (lower right).

[92] The large (spatial) variance in the estimated trends for the control run is reflected by the bumpy surface: overall aggregated trends for larger areas can hardly be inferred. For the REMO scenarios the picture is quite consistent. Positive trends in the South and along the North Sea coastline, and for the A2 scenario also a clear prevalence of positive trends in the western parts of Germany.

[93] In order to investigate the dependence of the trend estimates on other covariates, scatterplots (Figure 9) of trend estimates (again $\hat{\beta}_1$ times 1000) against elevation (left column) and monthly mean precipitation (right column) for the observations, REMO first control run, REMO B1 scenario, REMO A1B scenario, and REMO A2 scenario have been produced. In the REMO scenario runs a slight (negative) dependence of the trends on the monthly mean precipitation is visible, otherwise no correlation is present. It cannot be concluded that a negative dependence of the estimated trends on monthly mean precipitation (visible mainly in the A1B scenario simulation) is a robust feature.

3.1.2. Bayesian Estimation

[94] Results shown in this section only refer to the REMO A1B scenario simulation. In this case, a Bayesian parameter

a

b

c

d
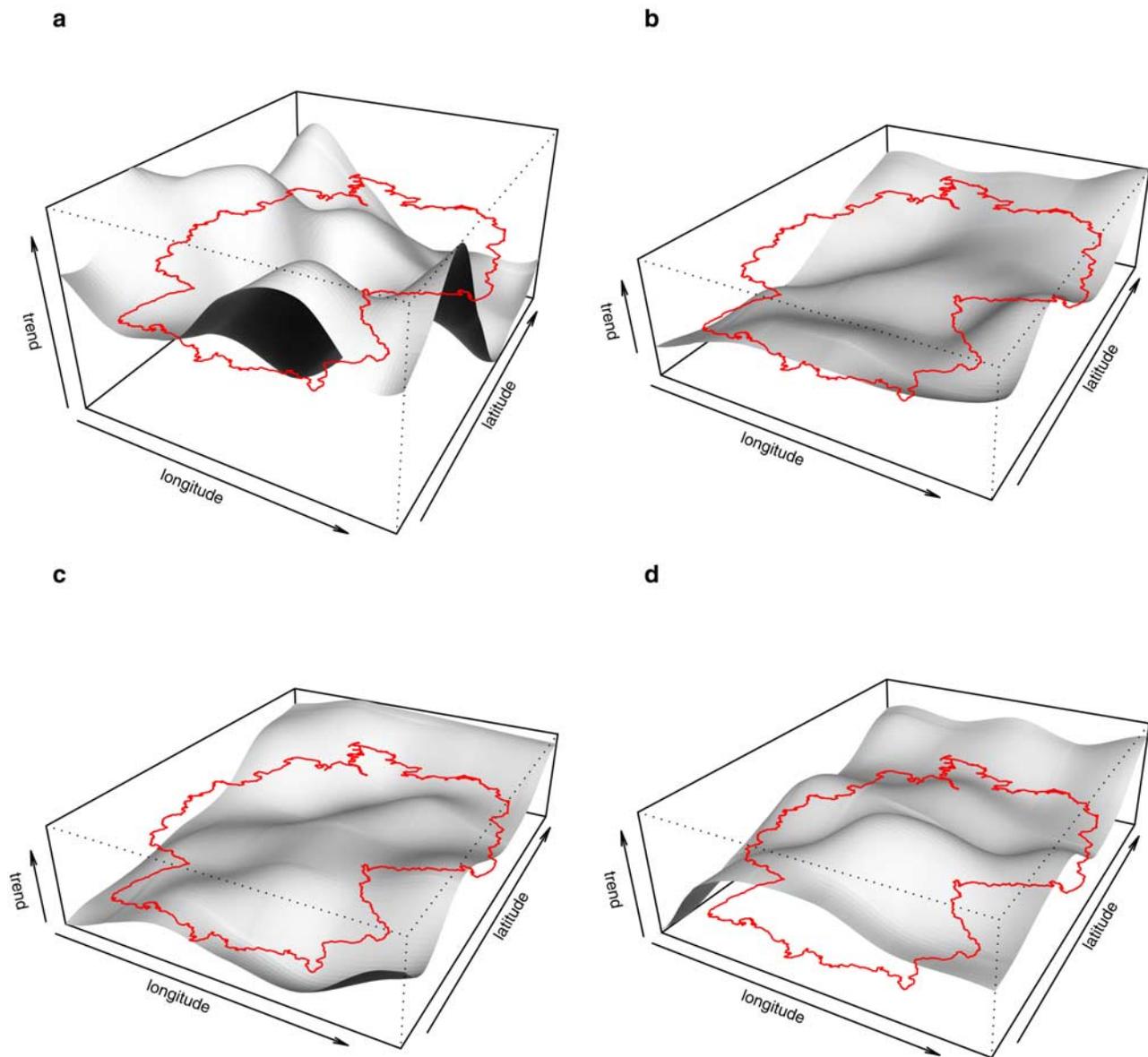
Figure 8.   Smoothed and scaled trend estimates ($\hat{\beta}_1$ times 1000) for (a) first REMO control run; (b) B1 scenario; (c) A1B scenario; and (d) A2 scenario.

estimation was carried out using Markov chain Monte Carlo sampling. Markov chain Monte Carlo in a Bayesian context refers to algorithms for sampling from the posterior distribution based on constructing a Markov chain that converges to the desired distribution. More specifically, a random-walk Metropolis-Hastings algorithm was applied to derive a posterior sample of the parameters. The sample size was chosen to be 26,000, the first 1000 points were disregarded for the parameter estimation. Uniform prior distributions were chosen for all parameters except for the scale parameter σ_0 (see (8)); here a flat lognormal prior with mean equal to 50 and standard deviation equal to 50 was applied in order to exclude negative values. In practice, however, this proved not to be a problem.

[95]   Figure 10 shows the traces of the Markov chains and the estimated probability densities for selected grid boxes of the model domain and the trend parameter β_1 (see again

Figure 1b for the location of the sites). The traces are stationary and burn-in times are short, if present at all. The probability distributions of the trend estimates are generally quite symmetric and do not show a heavy tail behavior.

[96]   Figure 11 shows the result of the Bayesian estimation using MCMC for the estimated linear trend $\hat{\beta}_1 * 1000$. The means and the standard deviations of the posterior samples were compared to the maximum likelihood estimates and the estimated standard errors (using asymptotic theory, see section 2.2.1.5), respectively, presented in section 3.1.1.

[97]   The differences in estimated changes of extreme quantiles are shown in relative terms, with the maximum likelihood estimates as references (positive values in the plot imply larger values of the MCMC estimates). The relative differences in the parameter estimates, although small in absolute values, can exceed 100%. The cause for
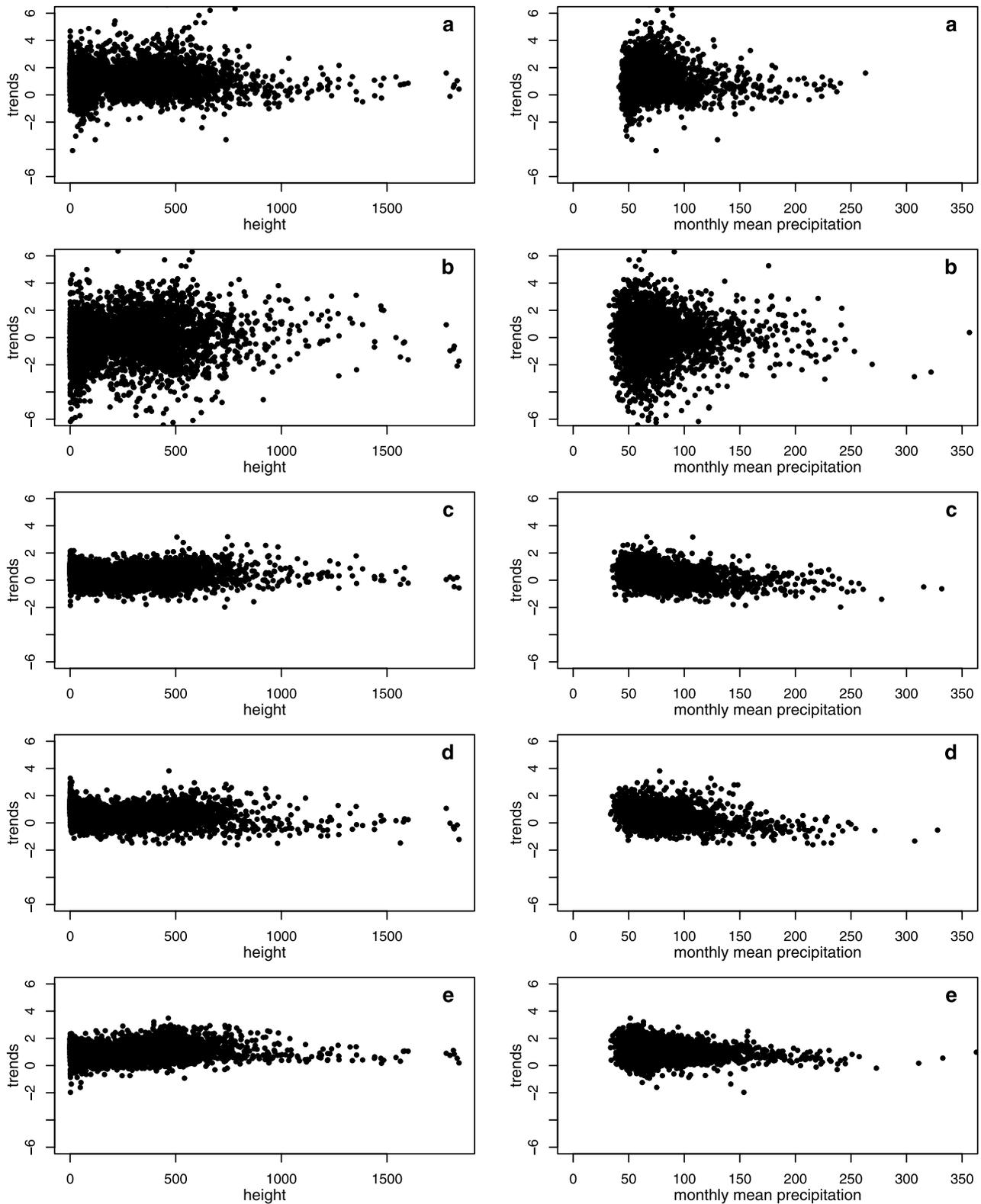
Figure 9. Scatterplots of trend estimates ($\hat{\beta}_1$ times 1000) against (left) height and (right) monthly mean precipitation. (a) Observations; (b) REMO first control run; (c) REMO B1 scenario; (d) REMO A1B scenario; and (e) REMO A2 scenario.

this probably lies in the fact that the likelihood function can have several local maxima. The numerical optimization algorithm used in the maximum likelihood estimation might not be able to find the global maxima in all cases. However,

also the fact that we use the mean of the posterior sample as parameter estimate (and not the sample point with the maximum value of the posterior density) plays a certain role. Rather surprisingly, the uncertainty estimates are much
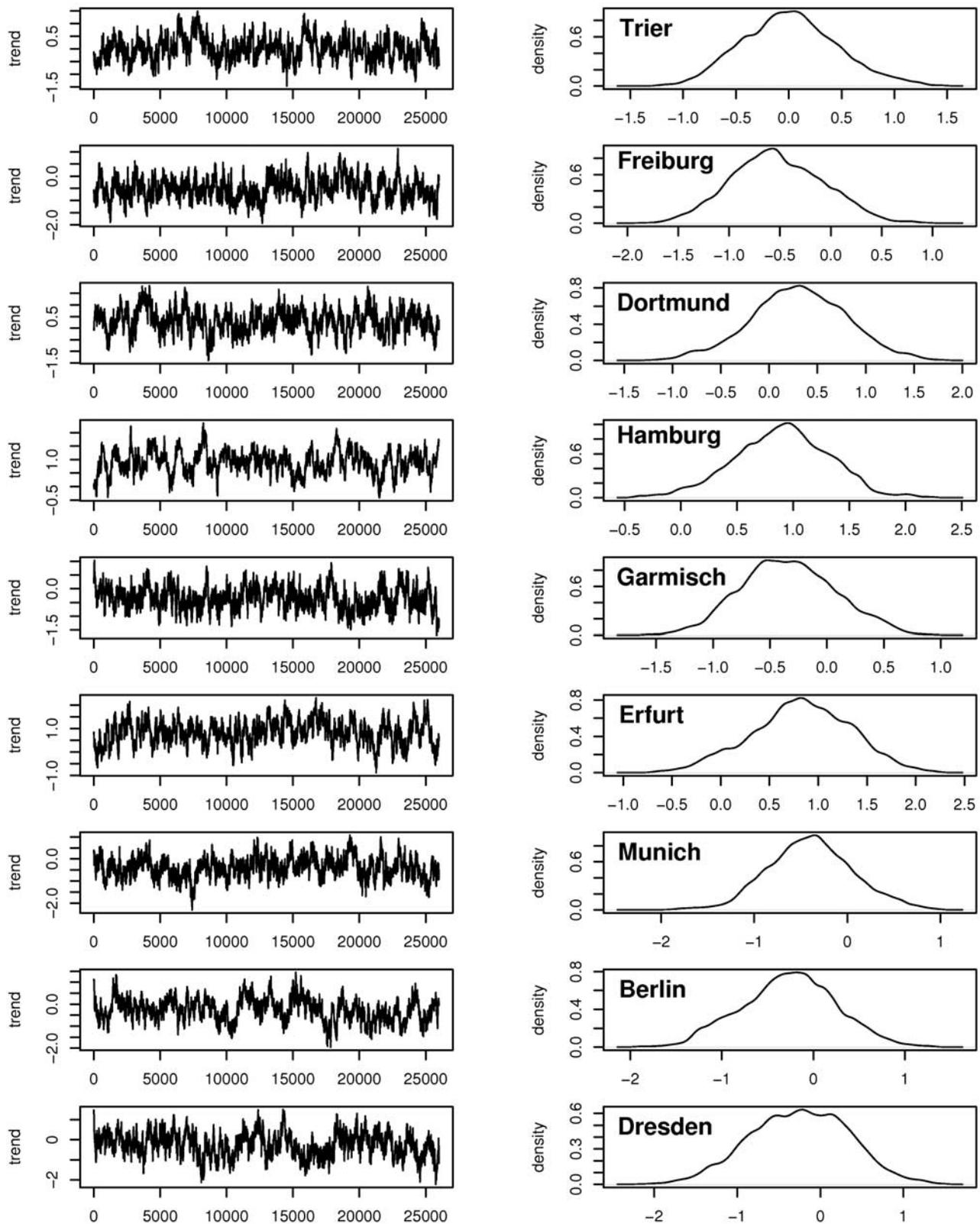
Figure 10. (left) Trace of the Markov chain samples of trends for selected grid boxes. (right) Estimated probability densities from the posterior sample.

more consistent and only differ within a range of −20% to 20%.

[98] The right panel of Figure 11 shows area of significantly positive trends (red colored) as determined by the *t* test based on the Bayesian estimates. This should be compared to Figure 5e. Although the Bayesian estimates result in slightly more extended areas of significant trends, the spatial pattern is similar overall.
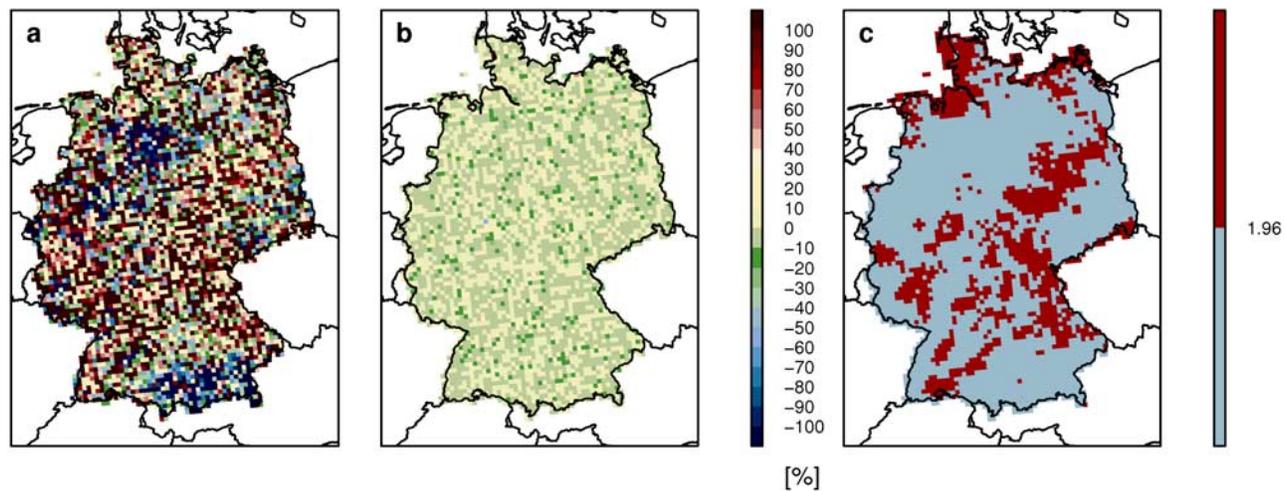
Figure 11. Results of the Bayesian estimation using MCMC. (a) Means and (b) standard deviations of the posterior samples were compared to the maximum likelihood estimates and the estimated standard errors for the REMO A1B simulation. Relative differences in estimated changes of extreme quantiles are shown, with the maximum likelihood estimates as references. (c) Area of significant trends (red colored) as determined by the t test based on the posterior sample.

3.2. Spatial Model for the Trends

[99] In the following paragraph the spatial dependence of the trends in both the observations and the climate model runs is investigated. As discussed in section 2.2.2, the goal of the analysis is to validate the climate model against observations, and to examine a possible change in spatial dependence of the trends in the scenario simulations of the regional climate model.

[100] Figure 12 shows empirical semivariograms of scaled trends ($\hat{\beta}_1$ times 1000) for the observations, second REMO control run, REMO B1 scenario, REMO A1B scenario, and the REMO A2 scenario. Pairs of stations were thereby pooled in clusters of size 10000. The results for the two REMO control simulations are similar, we therefore restrict the presentation to the analysis of one of them.

[101] The red lines show the estimated Matérn model using ordinary least squares fitting of the theoretical covariance function to the empirical semivariogram, the blue lines indicate the results of the maximum likelihood parameter estimation.

[102] For the variogram fitting the maximum considered distance between points (clusters) was set to 600 km, except for the control simulation where the maximum distance was chosen to be 500 km. In the maximum likelihood estimation no such threshold was included.

[103] The most striking feature is the difference in the semivariance. While the observations lie at around a value of 1.0, the semivariance of the REMO control run is considerably larger. The semivariance of the scenarios, including the B1 scenario, however, are consistently smaller and show a value of around 0.4. This is supported by the variance of the trend estimates (column five in Table 1). This means that the trends not only become predominantly positive in the scenario simulations, but also more consistent and less variable across Germany. However, the phenomenon of reduced semivariance can partly be explained by the fact that for the control simulation only a time span

of 44 years was considered, while the scenario simulations extend over a period of 100 years. To estimate the magnitude of this effect we selected 44 years of the A1B scenario simulation, the years 2057 to 2100, and repeated the analysis. The result (not shown) exhibits an almost tripled semivariance compared to Figure 12d, but still only about have the semivariance of the control simulation.

[104] The general shapes of the semivariograms of the observations and the REMO control run agree well and this form remains essentially unchanged in the scenario simulations. There is no clear sign that the climate model exhibits much more or less spatial correlation in the estimated trends than the observations.

[105] In Table 1 the parameter estimates related to the spatial model for the trends using the two different estimation techniques are summarized.

[106] Recall that the resolution of the climate model results as well as of the observations is 10 km. From the data we can therefore not infer the spatial covariance structure of the trends at smaller scales. The shape parameter v determines the behavior of the covariance function around zero: a shape parameter close to 0.5 implies an exponential covariance model, a large value of the shape parameter leads to a Gaussian form of the covariance function.

[107] In the case of the observations and the REMO control simulation there is a difference in the results of the parameter estimation using variogram fitting and maximum likelihood estimation. While the variogram method gives a more Gaussian covariance function, the maximum likelihood estimation leads to a exponential type of covariance function. This implies different estimates for the other parameters of the covariance model, too.

[108] For the scenario simulations the resulting covariance functions are of exponential type for both estimation methods. The uncertainty in the parameter estimates is however large. The value of the likelihood function proved to be close for very different combinations of parameters. It
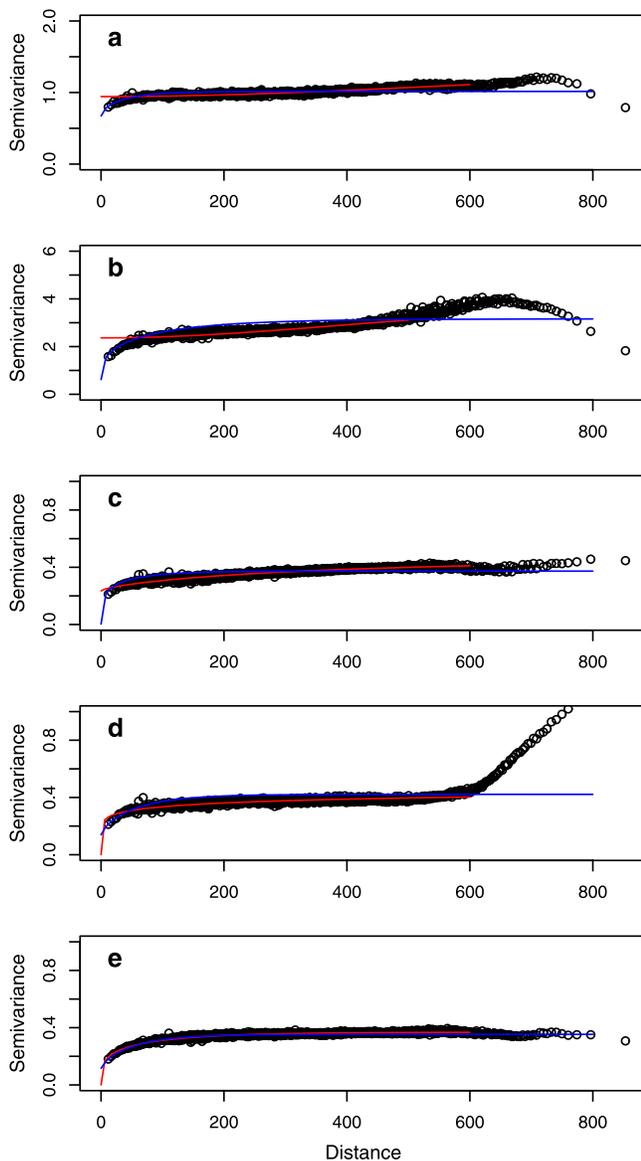
Figure 12. Empirical semivariograms of trends for (a) observations; (b) REMO second control run; (c) REMO B1 scenario; (d) REMO A1B scenario; and (e) REMO A2 scenario. The red lines show the fitted Matérn model using ordinary least squares variogram fitting, the blue lines depict the results of the maximum likelihood estimation.

seems therefore problematic to attach a direct physical meaning to the estimated parameter values. This relates in particular to the typical correlation length ρ. The uncertainty is too large to conclude that typical distances of correlation in the data are equal to the estimated value of ρ. However, also the estimates for the nugget effect depend heavily on the inferred behavior of the correlation function around zero (i.e., the value of v) and should be interpreted with care.

[109] This implies that it would be problematic to use the parameter estimates for kriging in order to statistically downscale the observations or climate model results. Kriging at scales comparable to the resolution of the data would however hardly be affected by these uncertainties in the parameter estimates and could be used, for instance, to produce a gridded data product from station data.

4. Discussion

[110] In the present paper a statistical extreme value analysis is applied to very high-resolution climate model results and observations encompassing the area of Germany. The gridded observational data set is quality checked and corrected for well-known systematic observational errors (i.e., precipitation undercatch of the measurement devices). The study is divided into two main parts: first trends in extreme quantiles of daily precipitation totals are estimated in a station-by-station analysis. In the second part, the spatial characteristics of the estimated trends in heavy rainfall are investigated over the area of Germany by fitting a parametric geostatistical model to these trends.

[111] In the control simulations the climate model tends to underestimate trends in heavy rainfall compared to observations. In the scenario simulations positive trends prevail (as in the observations). They are however relatively small when set in relation to the uncertainties. The trends become significantly positive to a larger spatial extent only in the A2 scenario simulation.

[112] This does of course not necessarily imply that increasing frequencies and intensities of extreme precipitation events do not pose a threat to Germany. In our study we only considered long-term trends in extreme quantiles in precipitation. It would be necessary to couple the climate model to hydrological discharge models in order to estimate the consequences of climate change in terms of frequency and intensity of floods. This would lead to a more appro-

Table 1. Parameter Estimates Related to the Spatial Model for the Trends Using Least Squares Variogram Fitting and Maximum Likelihood Parameter Estimation (See Equation (19) for the Definition of the Parameters)

Data	σ^2	ρ	v	τ^2	$\mathrm{Var}(\hat{\beta}_1 * 1000)$
Least squares variogram fitting					
Observations (1961–2004)	0.398	52.49	60.20	0.943	1.012
REMO Control II (1961–2004)	1.286	38.175	50.02	2.366	2.840
REMO B1 (2001–2100)	0.155	215.0	0.718	0.267	0.364
REMO A1B (2001–2100)	0.4076	351.45	0.0693	0.01	0.390
REMO A2 (2001–2100)	0.369	149.398	0.119	0.01	0.350
Maximum likelihood					
Observations (1961–2004)	0.3438	53.78	0.278	0.6714	1.012
REMO Control II (1961–2004)	2.546	155.91	0.164	0.617	2.840
REMO B1 (2001–2100)	0.3699	73.9	0.1163	0.0033	0.364
REMO A1B (2001–2100)	0.2826	71.75	0.3486	0.1392	0.390
REMO A2 (2001–2100)	0.2379	67.6	0.3479	0.1156	0.350

priate measure of potential damages to the society caused by heavy rainfall events.

[113] The Bayesian estimation essentially confirmed the uncertainty estimates resulting from the asymptotic theory. Nevertheless, the advantage of the Bayesian approach lies in the fact that it leads to probability distributions of trends, which can be used further in impact studies and risk assessments. Moreover, it allows for including prior information on the parameters in the estimation procedure [see, e.g., *Frei et al.*, 2006].

[114] The estimated shape of the extreme value distributions does not change significantly in the scenario simulations compared to the climate model control runs, that is, the distributions of the heavy rainfall events are not projected to become pronouncedly more heavy tailed. The variances of these distributions nevertheless increase (because of the trend in the scale parameter σ, see (8)).

[115] The fitting of the geostatistical Matérn model to the trend estimates proves to be a useful method to compare the spatial characteristics of the observations and the climate model results. However, the parameter estimates are uncertain. In our situation this was due mainly to the finite resolution of the gridded observations and the climate model results: data is not available on a resolution finer than the mesh size of the grid (i.e., 10 km). Therefore the model should be applied with care. The most striking result is a reduction of the spatial variance of the trends for the scenario simulations which can only partly be explained by the fact that the considered time series are shorter in the control simulations: trends become considerably more consistent and homogeneous over the considered area of Germany in the scenario runs.

[116] Bayesian hierarchical models [*Cooley et al.*, 2008; *Sang and Gelfand*, 2009] are very appropriate when analyzing spatial correlation in extreme event characteristics. They allow for borrowing strength across locations in the procedure of parameter estimation. However, in the case of very large data sets, the computational costs limit the applicability of these methods. In our situation we analyze threshold exceedances of the 95% quantiles of long daily time series for more than 4000 grid points which amounts to a substantial quantity of data. Here the two-step procedure of first estimating trends in a station-by-station analysis, and then fitting the spatial model separately, proves to be an advantageous approach. The point process model of threshold exceedances can be fitted with care, even Bayesian parameter estimation is feasible and leads to a good fit of the statistical model and reliable uncertainty estimates. In principle, a Bayesian approach would also be possible in the second step, the parameter estimation for the geostatistical Matérn model. However, in our situation we considered the dimensionality of the problem (and, consequently, the computational cost) to be too high for generating enough samples of the posterior distribution.

[117] With regard to the rule of thumb of estimating trends in extreme quantiles of heavy precipitation based on the Clausius-Clapeyron relation, about 6.5% per 1° temperature increase (see section 1), no unequivocal conclusion can be drawn from our study. The simple thermodynamic equation tends to predict stronger trends than the regional climate model, but agrees rather well with the observational record. It would be useful to conduct a similar study for other parts of the world using high-resolution data. A more thorough investigation of the reasons and physical processes causing positive trends in extreme precipitation events is planned in a future work. Apart from the increased moisture content of a warmer atmosphere, general changes in the circulation pattern and increased blocking frequencies (J. Sillmann and M. Croci-Maspoli, Atmospheric blocking and extreme events in the present and future climate, submitted to *Geophysical Research Letters*, 2009) may be responsible for more extreme heavy rainfall events in the future.

References

Allen, M. R., and W. J. Ingram (2002), Constraints on future changes in climate and the hydrological cycle, *Nature*, *419*, 224–232.

Beniston, M., et al. (2007), Future extreme events in European climate: An exploration of regional climate model projections, *Clim. Change*, *81*, 71–95.

Boroneant, C., G. Plaut, F. Giorgi, and X. Bi (2006), Extreme precipitation over the Maritime Alps and associated weather regimes simulated by a regional climate model: Present-day and future climate scenarios, *Theor. Appl. Climatol.*, *86*, 81–99.

Coles, S. (2001), *An Introduction to Statistical Modeling of Extreme Values*, 208 pp., Springer, New York.

Coles, S., and A. Stephenson (2006), *ismev: An Introduction to Statistical Modeling of Extreme Values*, R package version 1.2.

Cooley, D., D. Nychka, and R. L. Smith (2008), Bayesian spatial modeling of extreme precipitation return levels, *J. Am. Stat. Assoc.*, *102*(479), 824–840.

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

de Haan, L., and T. T. Pereira (2006), Spatial extremes: Models for the stationary case, *Ann. Stat.*, *34*(1), 146–168.

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

Fawcett, L., and D. Walshaw (2006), Markov chain models for extreme wind speeds, *Environmetrics*, *17*, 795–809.

Fawcett, L., and D. Walshaw (2007), Improved estimation for temporally clustered extremes, *Environmetrics*, *18*, 173–188.

Ferro, C. A. T., and J. Segers (2003), Inferences for clusters of extreme values, *J. R. Statist. Soc. Series B*, *65*, 545–556.

Frei, C., and C. Schär (2001), Detection probability of trends in rare events: Theory and application to heavy precipitation in the Alpine region, *J. Clim.*, *14*, 1564–1584.

Frei, C., C. Schär, D. Lüthi, and H. C. Davies (1998), Heavy precipitation processes in a warmer climate, *Geophys. Res. Lett.*, *25*(9), 1431–1434.

Frei, C., J. H. Christensen, M. Déqué, D. Jacob, R. G. Jones, and P. L. Vidale (2003), Daily precipitation statistics in regional climate models: Evaluation and intercomparison for the European Alps, *J. Geophys. Res.*, *108*(D3), 4124, doi:10.1029/2002JD002287.

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

Gelfand, A. E., and A. F. M. Smith (1990), Sampling based approaches to calculating marginal densities, *J. Am. Stat. Assoc.*, *85*, 398–409.

IPCC (2007), Climate Change 2007: The physical science basis, *Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge Univ. Press, New York.

Jacob, D. (2001), A note to the simulation of the annual and inter-annual variability of the water budget over the Baltic sea drainage basin, *Meteorol. Atmos. Phys.*, *77*(20), 66–73.

Jacob, D., et al. (2007), An inter-comparison of regional climate models for Europe: Model performance in present-day climate, *Clim. Change*, *81*, 31–52.

Jacob, D., H. Göttel, S. Kotlarski, P. Lorenz, and K. Sieck (2008), Klimaauswirkungen und Anpassung in Deutschland. Phase 1: Erstellung regionaler Klimaszenarien für Deutschland, *UBA Forschungsbericht 204 41 138*, Umweltbundesamt, Dessau.

Kharin, V. V., F. W. Zwiers, X. Zhang, and G. C. Hegerl (2007), Changes in temperature and precipitation extremes in the IPCC ensemble of global coupled model simulations, *J. Clim.*, *20*, 1419–1444.

Kunz, M., and C. Kottmeier (2006), Orographic enhancement of precipitation over low mountain ranges. Part II: Simulations of heavy precipitation events over Southwest Germany, *J. Appl. Meteorol. Climatol.*, *45*(8), 1041–1055.

Leadbetter, M. R., G. Lindgren, and H. Rootzen (1983), *Extremes and Related Properties of Random Sequences and Processes*, 336 pp., Springer, Berlin.

Meier, L. (2004), Extremwertanalyse von Starkniederschlägen, ETH Diploma thesis, Eidgenössische Technische Hochschule Zurich.

Mendes, J. M., P. de Zea Bermudez, J. Pereira, K. F. Turkman, and M. J. P. Vasconcelos (2009), Spatial extremes of wildfire sizes: Bayesian hierarchical models for extremes (online), *Environ. Ecol. Stat.*, doi:10.1007/s10651-008-0099-3.

Nordeng, T. E. (1994), Extended versions of the convective parametrization scheme at ECMWF and their impact on the mean and transient activity of the model in tropics, *ECMWF Tech. 206*, ECMWF, Reading, U. K.

Pall, P., M. R. Allen, and D. A. Stone (2007), Testing the Clausius-Clapeyron constraint on changes in extreme precipitation under CO_2 warming, *Clim. Dyn.*, *28*, 351–363.

R Development Core Team (2008), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria.

Randall, D. A., et al. (2007), Climate models and their evaluation, in *Climatic Change 2007: The Physical Science Basis*, *Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by S. Solomon et al., Cambridge Univ. Press, New York.

Sang, H., and A. Gelfand (2009), Hierarchical modeling of extreme values observed over space and time (online), *Environ. Ecol. Stat.*, doi:10.1007/s10651-007-0078-0.

Schlather, M., and J. A. Tawn (2003), A dependence measure for multivariate and spatial extreme values: Properties and inference, *Biometrika*, *90*, 139–156.

Semmler, T., and D. Jacob (2004), Modeling extreme precipitation events—a climate change simulation for Europe, *Global Planet. Change*, *44*, 119–127.

Sillmann, J., and E. Röckner (2008), Indices for extreme events in projections for anthropogenic climate change, *Clim. Change*, *86*, 83–104.

Smith, R. L. (2003), Statistics of extremes, with applications in environment, insurance and finance, in *Extreme Values in Finance, Telecommunications and the Environment*, edited by B. Finkenstadt and H. Rootzen, pp. 1–78, CRC Press, London, U. K.

Smith, A. F. M., and G. O. Roberts (1993), Bayesian computation via the Gibbs-sampler and related Markov chain Monte Carlo methods, *J. R. Stat. Soc. Series B*, *55*, 3–24.

Smith, R. L., J. A. Tawn, and S. G. Coles (1997), Markov chain models for threshold exceedances, *Biometrika*, *84*, 249–268.

Sundqvist, H. (1978), A parameterization scheme for non-convective condensation including prediction of cloud water content, *Q. J. Meteorol. Soc.*, *104*, 677–690.

Tebaldi, C., K. Hayhoe, J. M. Arblaster, and G. A. Meehl (2006), Going to the extremes. an intercomparison of model-simulated historical and future changes in extreme events, *Clim. Change*, *79*, doi:10.1007/s10584-006-9051-4.

Tiedtke, M. (1989), A comprehensive mass flux scheme for cumulus parametrisation in large-scale models, *Mon. Weather Rev.*, *117*, 1779–1800.

Trömel, S., and C.-D. Schönwiese (2007), Robust trend estimation of observed German precipitation, *Theor. Appl. Climatol.*, doi:10.1007/s00704-007-0341-1.

Zängl, G. (2007), To what extent does increased model resolution improve simulated precipitation fields? A case study of two North-alpine heavy-rainfall events, *Met. Z.*, *16*(5), 571–580.

―――――――――――――

D. Jacob and L. Tomassini, Regional Climate Modeling, Max Planck Institute for Meteorology, Bundesstrasse 53, D-20146 Hamburg, Germany. (lorenzo.tomassini@zmaw.de)