Nonstationary probabilistic downscaling of extreme precipitation

M. Kallache,[1,2] M. Vrac,[2] P. Naveau,[2] and P.-A. Michelangeli[1]

[1] Reanalysis data and general circulation model outputs typically provide information at a coarse spatial resolution, which cannot directly be used for local impact studies. Downscaling methods have been developed to overcome this problem, and to obtain local-scale information from large-scale atmospheric variables. The deduction of local-scale extremes still is a challenge. Here a probabilistic downscaling approach is presented where the cumulative distribution functions (CDFs) of large- and local-scale extremes are linked by means of a transfer function. In this way, the CDF of the local-scale extremes is obtained for a projection period, and statistical characteristics, like return levels, are inferred. The input series are assumed to be distributed according to an extreme value distribution, the Generalized Pareto distribution (GPD). The GPD parameters are linked to further explanatory variables, hence defining a nonstationary model. The methodology (XCDF-t) results in a parametric CDF, which is as well a GPD. Realizations generated from this CDF provide confidence bands. The approach is applied to downscale National Centers for Environmental Prediction reanalysis precipitation in winter. Daily local precipitation at five stations in southern France is obtained. The calibration period 1951–1985 is used to infer precipitation over the validation period 1986–1999. The applicability of the approach is verified by using observations, quantile-quantile plots, and the continuous ranked probability score. The stationary XCDF-t approach shows good results and outperforms the nonparametric CDF-t approach or quantile mapping for some stations. The inclusion of covariate information improves results only sometimes; therefore, covariates have to be chosen with care.

Citation: Kallache, M., M. Vrac, P. Naveau, and P.-A. Michelangeli (2011), Nonstationary probabilistic downscaling of extreme precipitation, *J. Geophys. Res.*, *116*, D05113, doi:10.1029/2010JD014892.

1. Introduction

[2] Currently coupled atmosphere-ocean general circulation models (GCMs) generate projections on a too coarse scale to obtain information at local-scale level [*Meehl et al.*, 2007]. Hydrological processes, for example, typically occur on finer scales than those provided by GCM outputs [*Kundzewicz et al.*, 2007]. Downscaling methods are commonly used to infer from climate simulations of models to local-scale projections. Hydrological regional models are in general driven with downscaled data, normally temperature and precipitation data. This means that precipitation is one of the most relevant variables for hydrological impact studies [cf., e.g., *Quintana Seguí et al.*, 2010].

[3] Precipitation is difficult to model and to downscale, mostly due to its high spatial and temporal variability and its nonlinear nature. A variety of downscaling approaches exist, they can be divided into dynamical downscaling approaches, where a regional model is nested into a GCM, and probabilistic downscaling approaches. No universal best method exists, and thus many comparison studies have been carried out, for example the PRUDENCE project [*Christensen et al.*, 2007], or the STARDEX project [*Goodess et al.*, 2011]. A recent overview over current downscaling approaches is given by *Maraun et al.* [2010b].

[4] Here a probabilistic downscaling approach is presented, and in the following we will concentrate on this branch of methods. Downscaling comprises techniques such as the use transfer functions (e.g., regression) [*Wigley et al.*, 1990; *Wilby et al.*, 1998], or weather types [*Vrac et al.*, 2007]. We denote point data as *local-scale* data and data having an area as support, e.g., grid cells, as *large-scale* data. Perfect prognosis methods [cf. *Rummukainen*, 1997] establish a link between local- and large-scale observations. They may apply this relationship to downscale outputs from numerical models [e.g., *Wilks*, 2006]. To do so, the model has to simulate the large-scale variable realistically. For some studies, the weather sequences of local- and large-scale variable can be related to each other event by event. Newer downscaling approaches, known as model output statistics methods, use simulated mesoscale weather [*Rummukainen*, 1997]. They establish a direct relationship between large-scale variables simulated by a model, and local-scale observations. Their aim is to correct for model errors [see, e.g., *Lenderink et al.*, 2007]. Many downscaling approaches are hybrids of perfect prognosis and model output statistics methods. An example

---

[1]CLIMPACT, Paris, France.
[2]Laboratoire des Sciences du Climat et de l'Environnement, IPSL, Gif-sur-Yvette, France.

are stochastic weather generators, which can be used for downscaling as well. They generate time series of the local-scale variable by reproducing statistical properties of the observations (commonly a weather variable) [e.g., *Wilks and Wilby*, 1999; *Vrac and Naveau*, 2007]. Until recently, most studies have been based on the downscaling of mean values [*Fowler et al.*, 2007]. One focus of climate change studies is yet the assessment of changes in extremes [*Intergovernmental Panel on Climate Change (IPCC)*, 2007], and according frameworks for the downscaling of extremes have been lately proposed by, for example, *Friederichs and Hense* [2007] and *Cannon* [2011].

[5] Extreme value theory (EVT) deals with the assessment of extreme, that is rare, events. EVT suggests the use of a Generalized Extreme Value distribution for the probabilistic modeling of maxima. According to the fundamental *three types theorem* [*Fisher and Tippett*, 1928; *Gnedenko*, 1943], there are only three distributions, which can arise as limiting distributions of rescaled sample maxima. For values above a high threshold, EVT indicates that the adequate CDF comprising these three types corresponds to the Generalized Pareto Distribution (GPD). EVT plays an important role in risk assessment studies. A more detailed presentation is provided by, for example, *Leadbetter et al.* [1983], *Embrechts et al.* [1997], *Katz* [2002], and *Khaliq et al.* [2006].

[6] Here we present a stochastic downscaling approach for extremes which is in line with *Michelangeli et al.* [2009], and which takes findings of extreme value theory into account. We infer local-scale extremes from large-scale extremes, where "extreme" is any value above a high threshold. Another approach would be to use the whole set of large-scale values to infer the local-scale extremes. But, at the present time, this objective is overly complex because it means that behavior of local extremes has to be probabilistically linked to average and extreme large-scale information. EVT is not currently adapted to such a challenging framework. Further studies are needed to assess the relationship between non-extreme large-scale events and extreme local-scale ones. Here hybrid distributions, which mix, for example, the Gamma distribution and the GPD to model the whole distribution of precipitation might be of interest [*Vrac and Naveau*, 2007; *Furrer and Katz*, 2008].

[7] In this paper we solely focus on the links from large-scale extremes and local extremes. Basically we assume that a large-scale extreme event implies a local-scale extreme event and this relationship is the subject of our study. The cumulative distribution function (CDF) of large-scale extremes is linked to the CDF of the local-scale extremes by using a transfer function. The CDFs of the predictors are assumed to be distributed according to an extreme value distribution, the GPD. By letting the parameters of this distribution depend on covariates, further influencing variables can be integrated in the approach. The resulting CDF provides statistical characteristics, such as quantiles. The exploitation of CDF characteristics is a common procedure in precipitation analysis [see, e.g., *Vidal and Wade*, 2009]. In quantile regression, for example, the local-scale CDF quantiles are derived by means of a regression equation [cf. *Friederichs and Hense*, 2007; *Cannon*, 2011]. Here, by contrast, EVT is used to determine the transfer function between the local-scale quantile and the large-scale predictor,

and the outcome is a complete CDF. The inferred local-scale CDF allows for an estimation of the potential sizes of local-scale extremes. The chronology, that is the time points of occurrence of local-scale extreme events, has to be determined separately [see, e.g., *Benestad*, 2010]. In the following we will denote the presented approach as XCDF-t, as an extension of the CDF-t approach [*Michelangeli et al.*, 2009] to extreme values. The precondition for XCDF-t is a link between the CDFs of local-scale and large-scale extremes, which stays stable between calibration and prediction period. XCDF-t does therefore not require a realistic day-to-day representation of the large-scale variable and may be classified as model output statistics downscaling method. In case the large-scale data are biased, the bias has to be the same in calibration and prediction period. However, in case covariates are used to improve the description of the local-scale extremes, XCDF-t is rather in perfect prognosis setting. In this case it has to be verified that the large-scale covariates are simulated realistically by the model.

[8] In section 2 the data are presented. The methodology is outlined in section 3, and the use of covariates is introduced in section 4. Then, in section 5, results for the downscaling of extremes of daily winter precipitation in southern France are discussed, and a comparison with other methods is given. Concluding remarks follow in section 6.

2. Data

[9] We use observations of daily precipitation (millimeters) as local-scale variable. The observations have been provided by the European Climate Assessment & Dataset (ECA&D) project, http://eca.knmi.nl. Daily NCEP (National Centers for Environmental Prediction–National Center for Atmospheric Research) reanalysis precipitation data are to be downscaled to the observations (both variables have the same units).

[10] Reanalysis data are based on observations, they are capable to reproduce certain characteristics of observed precipitation, e.g., spatial large-scale patterns. However, model physics and uncertainties lead to uncertainty in the resulting reanalysis data product. Large biases may occur, depending on type of reanalysis data, season, and region [cf. *Widmann and Bretherton*, 2000; *Bosilovich et al.*, 2008]. *Bosilovich et al.* [2008] determined a moderate average difference in precipitation of -1 mm/d between reanalysis data and observations when looking at yearly averages in the region of Southern France. This holds for all reanalysis products assessed in this study, including NCEP reanalysis data. The method proposed here does furthermore not rely on a realistic day-to-day representation of the large-scale variable. We therefore regard NCEP reanalysis as adequate large-scale data for our study purposes.

[11] Five stations in Southern France, namely Marseille, Perpignan, Mont-Aigual, Nîmes and Sète, are assessed. Winter seasons (November to Mars) for the time period 1951–1999 are analyzed. As calibration period we choose the time span 1951–1985, and as verification period 1986–1999. In Figure 1a the station locations are shown together with grid cells indicating the resolution of the large-scale precipitation and of the covariates. The Mediterranean region is regularly confronted with periods of strong wind and heavy precipitation. The average duration of the heavy precipitation events is about 29 h [*Boudevillain et al.*, 2009]. Therefore it seems
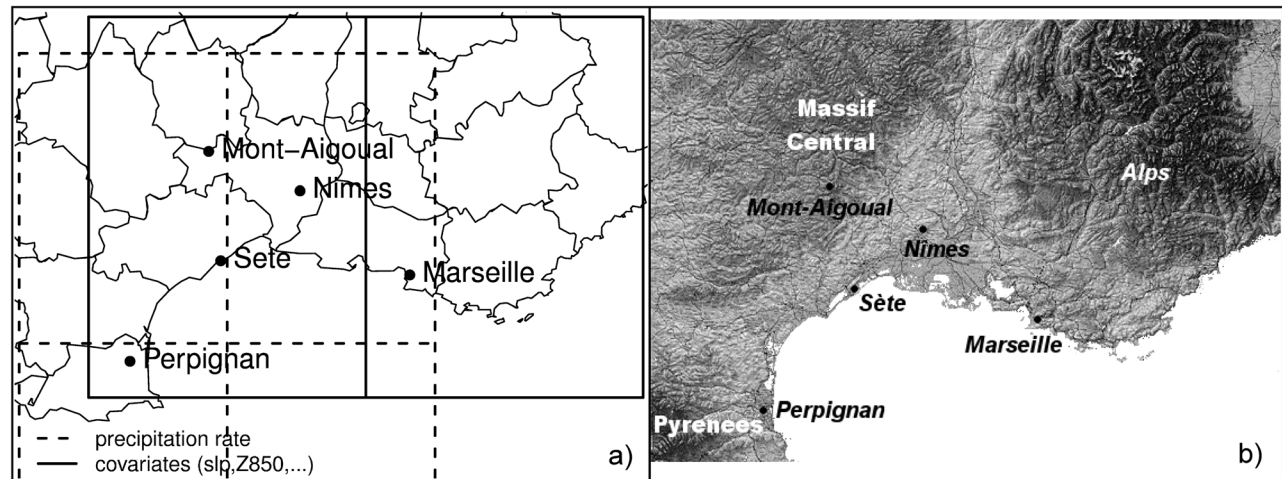
Figure 1. Stations assessed, which are located in southern France, including (a) some grid cells of NCEP reanalysis precipitation (black dashed line) and covariate variables (e.g., slp, Z850,…) and (b) the topography. Station Marseille is located at an altitude of about 54 m, Perpignan at 40 m, Mont-Aigoual at 1565 m, Nîmes at 100 m, and Sète at an altitude of about 30 m.

feasible to look at daily precipitation amounts and to neglect the duration of the events. Parts of the Massif Central, the Alps and the Pyrenees are located in the Mediterranean region. The combination of a Mediterranean influence with a mountainous area favors the occurrence of intense rain events and consequently of devastating flash floods, especially in autumn. This interplay determines the spatial structure of the precipitation patterns [*Berne et al.*, 2009; *Boudevillain et al.*, 2009]. In Figure 1b the topography of the Mediterranean region is shown, and the mountain regions are indicated. The southeastern part of the Massif Central (the Cevennes) is one of the rainiest areas of the Mediterranean region [*Nuissier et al.*, 2008]. Stations Nîmes, Mont-Aigoual, and partly Sète lie in its sphere. During the analyzed seasons the respective precipitation amounts reach up to 120 mm, 500 mm, and 140 mm per day. Perpignan is located near to the Pyrenees, which potentially trigger convection. Precipitation attains here up to 200 mm/d in the seasons analyzed. Marseille is situated on the coastline, here more moderate amounts up to 100 mm/d are recorded.

[12] Heavy precipitation events in the Mediterranean region are typically caused by the following process: The sea provides the moisture supply to the moderate to strong, southerly to easterly, low-level flow that feeds precipitation events. Orographic forcing then continuously generates new convective cells upwind of the Massif Central and thus maintains a stationary mesoscale convective system [*Ducrocq et al.*, 2008].

[13] For the analysis, all gridded data have been bilinearly interpolated to the station locations. In Figure 2, extreme precipitation data are exemplarily depicted for station Nîmes to give an impression of the local- and large-scale data used.

[14] Here excesses over a threshold constitute a set of extreme events [*Coles*, 2001]. Extremes of the large-scale data in calibration and projection period, and of the local-scale observations in the calibration period serve as input series for XCDF-t. Therefore thresholds have to be deter-

mined for those data. The thresholds have to be chosen sufficiently high to be able to assume a GPD distribution for the excesses, because the GPD distribution arises as a limiting distribution for high enough thresholds [*Coles*, 2001].

[15] XCDF-t establishes a link between local- and large-scale extremes. Therefore it would be desirable from
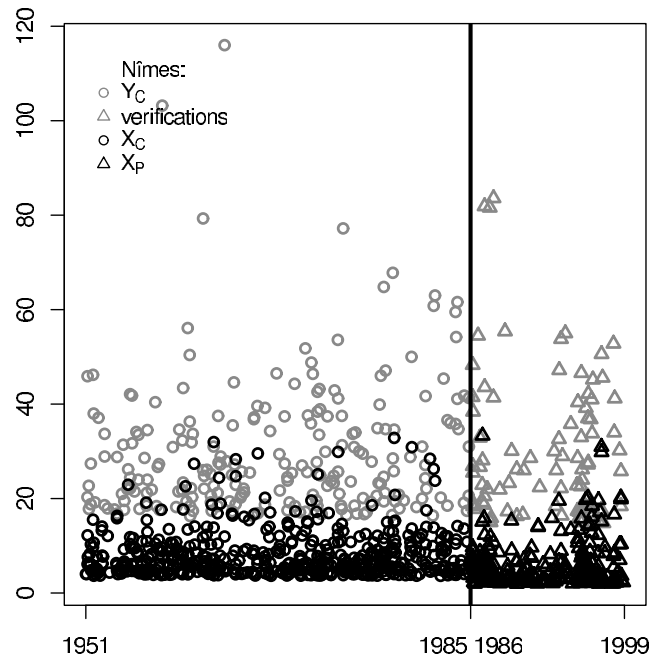
Figure 2. Extreme precipitation (mm/d) is shown for station Nîmes: observation extremes Y_C (grey circles, threshold at 7 mm) and large-scale extremes X_C (black circles, threshold at 3.63 mm) in the calibration period, and verification extremes (grey triangles, inferred threshold at 5.32 mm) and large-scale extremes X_P (black triangles, threshold at 1.94 mm) in the projection period are depicted.

Table 1. Percentage of Extreme Values (Number of Exceedances) of Days With Precipitation

Station	Y_C	X_C	X_P
Marseille	17.0% (262)	12.7% (475)	13.9% (157)
Perpignan	13.0% (189)	13.8% (475)	22.0% (231)
Mont-Aigoual	13.3% (342)	7.6% (369)	9.5% (168)
Nîmes	13.7% (222)	10.9% (528)	13.0% (231)
Sète	15.6% (209)	11.5% (396)	7.9% (83)

a technical point of view to choose the threshold for the large-scale series in the calibration time period as lowest value for which simultaneously local-scale excesses occur. However, small-scale variability is not completely determined by large-scale extreme events. Indeed, the thresholds which have been determined in this manner, were too low to be able to assume a GPD distribution for the precipitation data assessed. Therefore we chose higher thresholds for the large-scale data, which means that not all local-scale extremes occur simultaneously with a large-scale extreme. The assumed connection between local- and large-scale extremes is nevertheless strong (see below).

[16] We select a threshold separately for every station and for every series. A general rule for the threshold choice, for example, to take the topmost 5% of the data, led too often to a violation of the assumption of a GPD distribution, or the connection between large-scale and local-scale extremes showed to be too weak. A suitable threshold is determined by using the mean residual life plot, a fit over a range of thresholds and quantile-quantile plots (q-q plots) [cf. *Coles*, 2001]. We test the appropriateness of a GPD distribution for the excesses by means of the Anderson-Darling test (AD test) [*Choulakian and Stephens*, 2001]. To account for the asymmetric shape of the GPD, an AD test with modified weight function is used (here large weight is given only to departures in the right tail [see *Ahmad et al.*, 1988]). The distribution of the AD test statistics has been bootstrapped. For the thresholds chosen, the test confirms a compatibility with the GPD distribution at a 5% significance level for all series of extremes analyzed.

[17] The threshold selection results in rates of extremes from 8% to 22% of days where rain occurs (in this region we have about 30% of rainy days on average). An overview of the percentage and number of extreme values, which result from the threshold choice, is listed in Table 1. Given the data and the thresholds used in our application, on average over all stations about 59% of the local extremes are accompanied by an extreme event of the NCEP large-scale variable at the same day. Extreme event occurrences generated randomly under similar conditions coincidence for only 7.5% of the cases (results of this simulation study are not shown). About 96% of the local-scale extremes occur within the same month of at least one large-scale extreme. We consider this relation as being sufficient to link the CDF of the local-scale extremes to the CDF of all large-scale extremes.

3. XCDF-t

[18] CDF-t is an extension of the quantile mapping approach [see *Panofsky and Brier*, 1958; *Michelangeli et al.*, 2009] and belongs to the canon of statistical downscaling methods. Contrary to most statistical downscaling methods,

the goal here is to downscale a CDF, that is a statistical entity, and not directly time series of the variable of interest. XCDF-t is the adaptation of the CDF-t approach to the downscaling of extremes.

[19] Let F_{Y_C} be the CDF of observed local data at a site and for the historical calibration period. Similarly, let F_{X_C} be the CDF of the large-scale variable for the same time period and covering the site's location, i.e., GCM or, as in our application, NCEP reanalysis data. Furthermore, let F_{Y_P} and F_{X_P} be the according CDFs for the projection time period. The goal of CDF-t is to retrieve F_{Y_P} out of the other three CDFs by means of a transfer function $T(\cdot) : [0, 1] \rightarrow [0, 1]$.

$$F_{Y_P}(x) = T(F_{X_P}(x)) = F_{Y_C}\left(F_{X_C}^{-1}(F_{X_P}(x))\right), \quad (1)$$

where $T(\cdot)$ is an increasing function. For further details see *Michelangeli et al.* [2009].

[20] Here we link the CDFs of extremes, which we model as threshold excesses. Let F_{Y_C} be the CDF of a series of observed local-scale threshold excesses Y_C. Similarly, let X_C and X_P be the threshold excesses of the large-scale series. We assume that the threshold excesses are distributed according to a GPD [*Embrechts et al.*, 1997]. Thus F_{Y_C}, F_{X_C}, and F_{X_P} have the form of a GPD,

$$F(x; \sigma, \xi) = 1 - \left(1 + \frac{\xi x}{\sigma}\right)^{-(1/\xi)}, \quad (2)$$

with scale parameter σ and shape parameter ξ. The GPD is defined on $\{x : x > 0 \text{ and } (1 + \xi x/\sigma) > 0)\}$ with threshold u and excess $x = z - u$, where z is the local- or large-scale series.

[21] Let u_{Y_C} be the threshold for the local-scale observations in the calibration period, and u_{X_C} and u_{X_P}, respectively, be the thresholds for the large-scale series in calibration and projection period. These thresholds may be determined separately for each series. They have to be sufficiently high, so that the threshold excesses follow an extreme value distribution [cf., e.g., *Coles*, 2001]. When applying XCDF-t, local-scale extremes are inferred from the large-scale extremes. Therefore typically a big set of large-scale extremes is desirable for the analysis. This has to be considered as well when choosing the thresholds. The threshold u_{Y_P} of the inferred local-scale series is deduced from the other three thresholds; see Appendix A.

[22] In case F_{Y_C}, F_{X_C} and F_{X_P} have the same parameter values, F_{Y_P} trivially is a GPD with equal parameter values. Otherwise F_{Y_P} has a more complex structure. Let z be the large-scale values in the projection period and $x = z - u_{X_P}$ their threshold excesses. Then, by concatenating the distributions $F_{Y_C}(\sigma_{Y_C}, \xi_{Y_C})$, $F_{X_C}(\sigma_{X_C}, \xi_{X_C})$ and $F_{X_P}(\sigma_{X_P}, \xi_{X_P})$ as given in equation (1), we get

$$F_{Y_P}(x) = 1 - \left(1 + \frac{\xi_{Y_C}}{\xi_{X_C}}\frac{\sigma_{X_C}}{\sigma_{Y_C}}\left[\left(1 + \frac{\xi_{X_P}}{\sigma_{X_P}}(x + \kappa)\right)^{\frac{\xi_{X_C}}{\xi_{X_P}}} - 1\right]\right)^{-\left(\frac{1}{\xi_{Y_C}}\right)},$$

$$(3)$$

which is defined for $x > 0$, and where κ is a correction factor related to inflation. Indeed, to obtain the whole range of F_{Y_P}, we usually inflate and shift the data before the analysis (see

Appendix B). Therefore $u_{X_C} = u_{Y_C}$ holds, and both terms cancel each other out in equation (3).

[23] We chose to set $\xi_{X_P} = \xi_{X_C}$ and to estimate only σ_{X_C} from the large-scale series of excesses in the calibration period X_P. In this way model complexity is reduced and the inflation step is not needed. The quotient ξ_{X_C}/ξ_{X_P} in equation (3) cancels out, and finally we use the following relation:

$$F_{Y_P}(x) = 1 - \left(1 + \frac{\xi_{Y_C}}{\sigma_{X_P}} \frac{\sigma_{X_C}}{\sigma_{Y_C}}(x + \kappa)\right)^{-\left(\frac{1}{\xi_{Y_C}}\right)}. \quad (4)$$

In this way, results get more robust (small estimation deviations for the shape parameters might result in large variability of this quotient). This procedure is justifiable from a physical point of view (expected changes of the shape parameter for the same variable are small), as well as from a technical point of view (small deviations from the "true" shape parameter ξ_{X_P} might be compensated by a different scale parameter estimate). Now F_{Y_P} is a GPD distribution with scale parameter $\sigma_{Y_P} = \sigma_{Y_C}(\sigma_{X_P}/\sigma_{X_C})$ and shape parameter $\xi_{Y_P} = \xi_{Y_C}$. Thus, we expect a similar speed of decay for $F_{Y_C}(\cdot)$ and $F_{Y_P}(\cdot)$.

[24] Moreover, a GPD with negative shape parameter has an upper bound [*Coles*, 2001]. To capture the whole range of the extremes of Y_P in any case, we therefore apply XCDF-t only to heavy tailed time series, that is, we constrain our analysis to series whose CDF F_{Y_C}, F_{X_C}, and F_{X_P} have a positive shape parameter. This constraint is in general fulfilled for streamflow or precipitation data [cf. *Reiss and Thomas*, 1997; *Katz*, 2002; *Katz et al.*, 2002]. However, occasionally negative shape parameters are estimated for precipitation extremes [*Maraun et al.*, 2011; *Toreti et al.*, 2010]. This may be, for example, due to particular geographic conditions, which inhibit the occurrence of very extreme precipitation events. Furthermore, a single site assessment in contrast to a multivariate, regional analysis might be the reason [*Katz et al.*, 2002], or an underestimation of heavy precipitation in case GCM outputs are assessed [*Fowler et al.*, 2010]. In practice, we first check that all series analyzed have a shape parameter estimate whose confidence band suggests a compatibility with the parameter being larger than zero. Then, for the actual analysis, we set zero as lower bound for the shape parameter estimation.

[25] The parameters of F_{Y_C}, $F_{\tilde{X}_C}$, and $F_{\tilde{X}_P}$, namely $(\sigma_{Y_C}, \xi_{Y_C})$, $(\sigma_{\tilde{X}_C}, \xi_{\tilde{X}_C})$ and $(\sigma_{\tilde{X}_P})$, are estimated by maximizing the likelihood ($\tilde{X}_C$ and $\tilde{X}_P$ denote the inflated and shifted large-scale series; see Appendix B). For estimation, a representation of the GPD distribution with orthogonal parameters with respect to the Fisher information matrix is used to get more stable results, that is, the parameters (v, ξ) with $v = \sigma(1 + \xi)$ [cf. *Chavez-Demoulin and Davison*, 2005] are estimated. Potential covariates are linked to those parameters v and ξ.

4. Adding Covariate Information

[26] When linking the parameters of F_{Y_C}, $F_{\tilde{X}_C}$, and $F_{\tilde{X}_P}$ to covariates, supplementary information can be added to the analysis. Time as covariate allows for the modeling of time dependence of the data. The use of nonstationary extreme value distributions for climate change studies gets more and more popular. *Fowler et al.* [2010], for example, assess the question when changes in extreme precipitation due to climate change will be detectable by suiting a nonstationary extreme value distribution to climate model outputs. *Wang et al.* [2004] use a nonstationary extreme value distribution with covariates to obtain projections for wave heights in the North Atlantic Ocean. *Maraun et al.* [2010a] model extreme daily precipitation in this way. They integrate seasonality by means of the covariate time, and represent the influence of atmospheric circulation by the covariates synoptic-scale airflow strength, direction and vorticity. Covariates may be linked to the extreme value distributions by means of parametric or nonparametric functions, and these link functions may be linear or nonlinear. Further details on the GPD with covariate-dependent parameters are given, for example, by *Katz* [2002], *Naveau et al.* [2005], and *Overeem et al.* [2008].

[27] Synoptic-scale, mesoscale, and local-scale ingredients influence the emergence of heavy precipitation. A covariate represents one or several of those ingredients, and it thus has a different response at each station. Convection is initiated on mountain relief shoulders, therefore precipitation amounts are for example larger in the hill or mountain regions than in the plain or foothill regions. To account for this, the covariates, which significantly improve the description of the local-scale extremes by means of the statistical model, are selected separately for each station out of a set of potentially plausible covariates.

[28] XCDF-t concatenates GPD distributions as illustrated in equation (1). The parameters of those distributions F_{Y_C}, $F_{\tilde{X}_C}$, and $F_{\tilde{X}_P}$ may be linked to different covariates or stay constant. The link functions to the covariates may differ as well for each parameter.

[29] Let $(c_1^t, \ldots, c_n^t)$ be the state of the n covariates at time point t in the projection time period. The parameters of $F_{\tilde{X}_P}$ are linked to these covariate states, so $F_{\tilde{X}_P}(\sigma_{\tilde{X}_P}(c_1^t, \ldots, c_n^t), \xi_{\tilde{X}_P}(c_1^t, \ldots, c_n^t))$ is directly retrieved. The parameters of F_{Y_C} and $F_{\tilde{X}_C}$ are estimated by linking the covariate states in the calibration time period. To apply XCDF-t for time point t in the projection period, we use quantile mapping [cf. *Haddad and Rosenfeld*, 1997] to map the state of the covariates in the projection period at time point t to the according quantiles of the covariate distributions in the calibration period. The resulting covariate states $(\tilde{c}_1^t, \ldots, \tilde{c}_n^t)$ are the states of the covariates in the calibration period under conditions of the projection period at time point t. By using the link functions, which have been estimated for the calibration period, we then obtain $F_{Y_C}(\sigma_{Y_C}(\tilde{c}_1^t, \ldots, \tilde{c}_n^t), \xi_{Y_C}(\tilde{c}_1^t, \ldots, \tilde{c}_n^t))$ and $\tilde{F}_{X_C}(\sigma_{\tilde{X}_C}(\tilde{c}_1^t, \ldots, \tilde{c}_n^t), \xi_{\tilde{X}_C}(\tilde{c}_1^t, \ldots, \tilde{c}_n^t))$. XCDF-t is now applied for every time step t in the projection period, which results in a time varying CDF $F_{Y_P}^t$.

[30] The covariate concept may be used to include small and median large-scale data in the analysis. For some stations, extreme precipitation might for example not only be linked to extreme large-scale precipitation, but as well to less extreme large-scale values. In this case, time averaged large-scale precipitation can be included as covariate.

[31] We choose a set of potentially meaningful covariates, out of which then for each station separately suiting covariates are selected. Variables related to pressure are added to this set. These are common predictors for precipitation, namely sea level pressure (slp), principal components of fields of sea level pressure, geopotential height at 850 hPa (Z850), and principal components of fields of geopotential
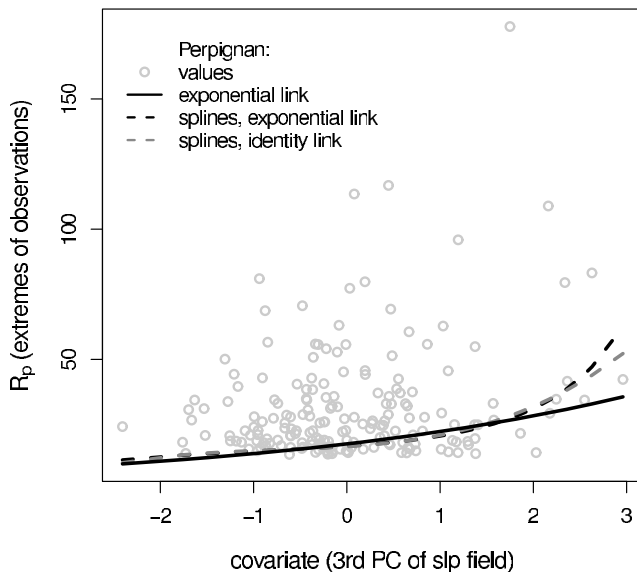
Figure 3. Relation of observation extremes Y_C to the significant covariate third PC of slp field (grey dots) for station Perpignan. The evolution of the GPD parameter v_{Y_C} with the covariate is shown as black line for an exponential link function (see equation (5)), as black dashed line for splines with exponential link function, and as grey dashed line for splines with identity link function.

height at 850 hPa. For the region under consideration, heavy precipitation typically is generated by warm and moist air from the Mediterranean Sea. This feeds in moisture and heats the lowest levels of the troposphere through air-sea sensible and latent heat fluxes. The more the low-level jets are perpendicular to the mountain range, the higher is the occurrence of heavy orographic rainfall [cf. *Ducrocq et al.*, 2008]. Therefore, maximum temperature (tmax), and wind velocity and wind direction at 10 hPa are included in the set of potential covariates. Furthermore, for most heavy precipitation events a strong synoptic-scale cyclone is present at upper levels, typically west of the threat area which generates a southerly to easterly flow that transports the warm and moist air masses from the Mediterranean Sea toward the coast and thus helps to destabilize the air column [*Nuissier et al.*, 2008]. Wind gradients potentially indicate the presence of cyclones. Therefore the difference of wind direction at 10 hPa and 850 hPa, the difference of wind velocity at 10 hPa and 850 hPa, and the difference of vertical wind intensity at 100 hPa and 850 hPa are included in the analysis. The steep and particular orography of the Mediterranean region changes wind directions, it channels the impinging moist and conditional unstable low-level flow by inducing upwind low-level convergence which contributes to the release of conditional convective instability near the Mediterranean Sea [cf. *Nuissier et al.*, 2008]. We account for this by adding further wind variables in the set of covariates, such as wind velocity, wind direction, and vertical wind intensity at 850 hPa and at 100 hPa. In addition, time and time squared are added to the set of covariates in case the potential nonstationarity of the local-scale extremes is not related to the covariates chosen. Finally, time averaged large-scale precipitation (avpr) is chosen as well as covariate. In this way it is taken into account

that local-scale heavy precipitation might be linked to average amounts of rain over the whole region.

[32] Maximum temperature is stemming from the ECA&D project, whereas all other variables are NCEP reanalysis data. The gridded data is taken from the grid cell covering the station location, and it is bilinearly interpolated to this station location. Time averaged large-scale precipitation is obtained by applying a kernel smoother with a bandwidth of 1 month. The principal components are retrieved out of data fields of nine grid points surrounding the station location. By using them, we aim to decrease the noise in the data and therefore to have a better representation of large-scale weather patters [cf. *Benestad*, 2010; *von Storch and Zwiers*, 1999]. Wind velocity and direction are calculated from the east-west and north-south wind components by a transformation into polar coordinates.

[33] We choose an exponential link function for parameter v (see section 3),

$$v_t = \exp\left(a_0 + a_1 c_1^t + \ldots + a_n c_n^t\right). \tag{5}$$

Here $c_1, \ldots, c_n$ are the potential covariates. Different link functions have been tested. The exponential link function allows a smooth dependence of v on the covariates. Figures and numerical results are not shown, but the exponential link function furthermore revealed to be a good compromise between an identity link function or using splines (with exponential or identity link function). The identity link function shows to be too rigid for the data assessed. Furthermore, for all stations the exponential link function achieves similar continuous ranked probability scores (see section 5.3) to the splines link function. A comparison of the log likelihood of the fits revealed that the splines link function effectuates an improvement of the fit for some series. However, this never justified the increase of model complexity in terms of the deviance statistics [see *Davison*, 2003]. As an illustration scatterplots of the extremes of the observations Y_C and the covariate are depicted for station Perpignan in Figure 3. The scatterplot (grey dots) does not show a complex structure. The evolution of the GPD parameter v with the covariate is depicted as black line for the parametric exponential relation given in equation (5), and as black and grey dashed lines for modeling this relation by using splines (actually a simple spline model with one polynomial basis of order three is used to represent the covariate, i.e., the number of knots is zero, two additional parameters per covariate have to be used). The evolution of v is similar for all three link functions depicted.

[34] We furthermore choose to leave all shape parameters constant. A covariate-dependent shape parameter ξ_t with negative and positive states at different time points would indicate a switch of the extreme value distribution characteristics from having an upper endpoint to allowing infinitely high extremes [*Embrechts et al.*, 1997]. The modeling of such a switch is beyond the scope of the paper.

[35] Before the analysis, all covariates are normalized to have mean zero and a standard deviation of one. In this way, the optimization procedure gets more robust. The normalization changes the parameter estimates a_i of equation (5), but does not affect the evolution of v_t.

[36] To select covariates, we assess separately for the local-scale series in the calibration period and for the large-scale series in calibration and projection period whether the

Table 2. Significant Covariate Variables Used for the Analysis

Station	Covariates
Marseille	time averaged large-scale precipitation, geopotential height at 850 hPa, wind direction at 850 hPa
Perpignan	third principal component of field of sea level pressure
Mont-Aigoual	maximum temperature, sea level pressure
Nîmes	sea level pressure, third principal component of field of geopotential height
Sète	third principal component of field of sea level pressure, wind velocity at 850 hPa

Table 4. Comparison of the Threshold u_{Y_P} for the Local-Scale Series in the Projection Period Obtained By Means of XCDF-t (u_{Y_P} (XCDF-t)) or By Assuming the Same Rate of Excesses for the Verifications as for the Local-Scale Series in the Calibration Period (u_{Y_P} (Verifications))

Station	u_{Y_P} (XCDF-t)	u_{Y_P} (Verifications)
Marseille	12.03	11.6
Perpignan	11.97	14.4
Mont-Aigoual	27.4	29.0
Nîmes	14.52	16.6
Sète	14.6	15.2

covariates improve the modeling of the extremes. To do so, the fits of a GPD distribution without covariates and GPD distributions with all possible combination of covariates are compared. The deviance statistics [see *Davison*, 2003] is used as model selection criterion.

[37] The principal components of the slp and Z850 fields comprise information of nine NCEP grid points. Indeed, they show to be linked more often to the large-scale extremes than to the local-scale extremes. We could not identify covariates which supply significant additional information for all three time series Y_C, X_C and X_P at a station location. We therefore chose to link only v_{Y_C} to covariates and to leave all other parameters constant. Often slp and Z850 exhibit a similar evolution. They seem to be key covariates to accompany winter precipitation extremes in Southern France, and are significant covariates for all stations assessed. Z850 and slp are related to low-pressure fields, which often accompany heavy precipitation in the Mediterranean region. Contrary to our expectations, time averaged NCEP precipitation data (avpr) does not significantly improve the GPD fit of the observations, except for station Marseille. However, this finding supports our assumption of linking the local-scale extreme CDF to the large-scale extreme CDF. Vertical wind intensity at 850 hPa showed to be a significant covariate for all stations, but this covariate worsened the verification results (see section 5.3) a lot. Hence, the link between vertical wind intensity and precipitation extremes is unstable. Therefore, we dismissed this variable. Wind is a significant covariate for stations Marseille and Sète, that is wind direction for Marseille and wind velocity for Sète. Both stations are situated on the coastline and therefore wind is not hindered or altered by the topography and may have a strong influence.

[38] All significant covariates used for the further analysis are listed in Table 2.

5. Results

5.1. Parameter Estimates and Inferred CDF

[39] In Table 3, the GPD scale and shape parameter estimates used for XCDF-t, are listed together with their 68% confidence interval (the confidence interval for the scale

parameter is deduced from the parameter v which has actually been estimated, by using the delta method [cf. *Coles*, 2001]). The decay rate of the CDF depends on sign and size of the GPD shape parameter. It determines the frequency of occurrence of the most extreme events; therefore the shape parameter of the GPD distribution is an interesting feature to look at. As is apparent in Table 3, the estimated shape parameters of all series are close to zero. This means, that the CDF of the input data decay slower than exponentially, but not at a very slow rate.

[40] To evaluate the results, we first compare the threshold u_{Y_P} determined by XCDF-t to the threshold obtained when assuming equal rates of exceedance for the verifications and for the local series in the calibration period. Here rates of exceedance for days with precipitation are used, as listed as percentages in Table 1. Both thresholds are listed in Table 4. The XCDF-t approach implies that the local-scale data experience the same change between calibration and projection time period as the large-scale variable, which determines the threshold u_{Y_P} (see Appendix A). For the data assessed it shows, that u_{Y_P} deduced with XCDF-t is close to the threshold obtained when using the verifications. The largest deviance occurs for Perpignan, where u_{Y_P} determined by XCDF-t is 17% lower. In the further comparison with observations, we use this threshold u_{Y_P} determined by XCDF-t to identify the verification extremes.

[41] In Figure 4, the resulting distribution F_{Y_P} of the local extremes for the projection period (1986–1999) is depicted as black line (here F_{Y_P} is obtained without the use of covariate information). Furthermore, the cumulative distributions of the verifications (black dots), Y_C (grey line), X_C (grey dashed line) and X_P (black dashed line) are shown. To allow for a comparison without taking our choice of distribution into account, the empirical cumulative distribution functions (ECDF) are shown and not the estimated GPD distributions. Apparently, the distribution of the large-scale variable does not undergo large changes from the calibration to the projection period. As outlined in Appendix B and illustrated in Figure 4, that is transferred to the local-scale variable, which undergoes as well small changes from the calibration to the

Table 3. Scale Parameter σ and Shape Parameter ξ Estimates and 68% Confidence Intervals for XCDF-t Without Covariates

Station	σ_{Y_C}	$\sigma_{\bar{X}_C}$	$\sigma_{\bar{X}_P}$	ξ_{Y_C}	$\xi_{\bar{X}_C}$ ($\xi_{\bar{X}_P}$)
Marseille	11.60 [10.57, 12.63]	3.82 [3.56, 4.08]	4.32 [3.97, 4.67]	0.008 [0.0, 0.072]	0.022 [0.0, 0.072]
Perpignan	15.42 [13.72, 17.12]	3.35 [3.12, 3.58]	3.22 [2.99, 3.44]	0.159 [0.075, 0.242]	0.085 [0.033, 0.138]
Mont-Aigoual	34.30 [31.74, 36.86]	3.90 [3.60, 4.20]	3.83 [3.52, 4.14]	0.091 [0.039, 0.142]	0.047 [0.0, 0.103]
Nîmes	13.23 [12.00, 14.47]	3.96 [3.69, 4.23]	3.93 [3.65, 4.22]	0.030 [0.0, 0.095]	0.111 [0.059, 0.163]
Sète	15.87 [14.23, 17.50]	3.58 [3.30, 3.85]	4.19 [3.70, 4.69]	0.144 [0.067, 0.221]	0.088 [0.030, 0.146]
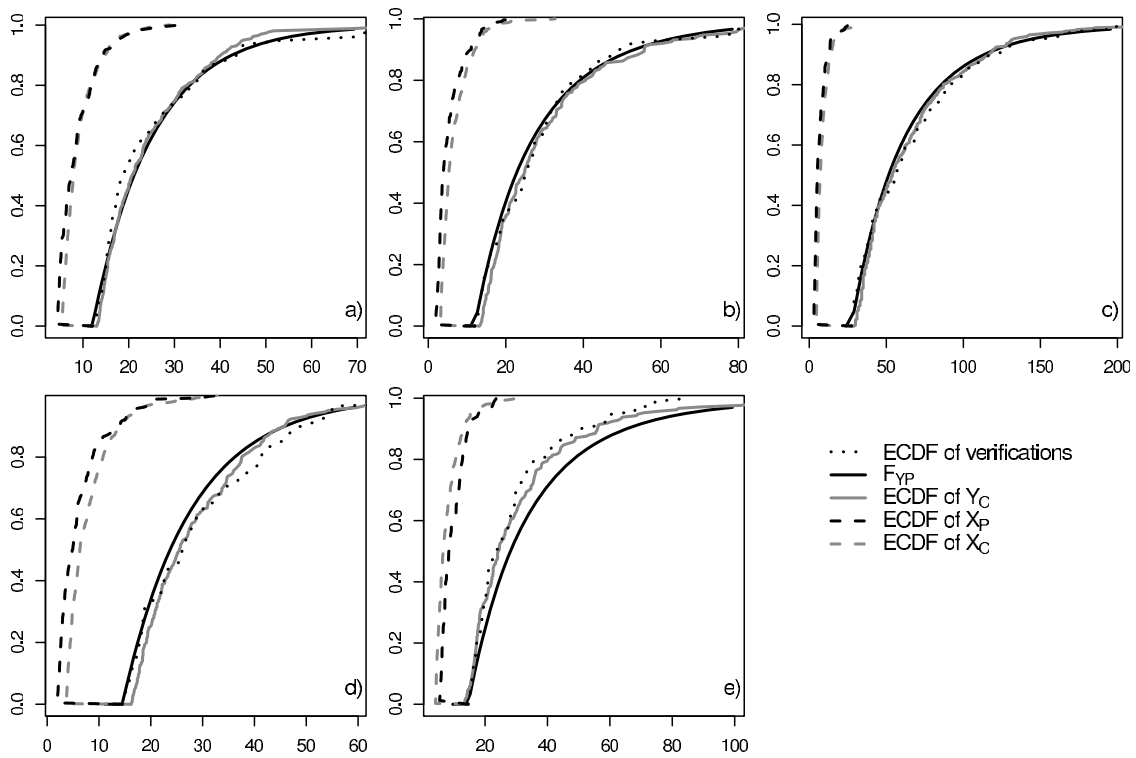
Figure 4. F_{Y_P} (black line) and the ECDF of the verifications (black dotted line), Y_C (grey line), X_C (grey dashed line), and X_P (black dashed line) are shown for stations (a) Marseille, (b) Perpignan, (c) Mont-Aigoual, (d) Nîmes, and (e) Sète. The x axis is in mm/d; the y axis is probability.
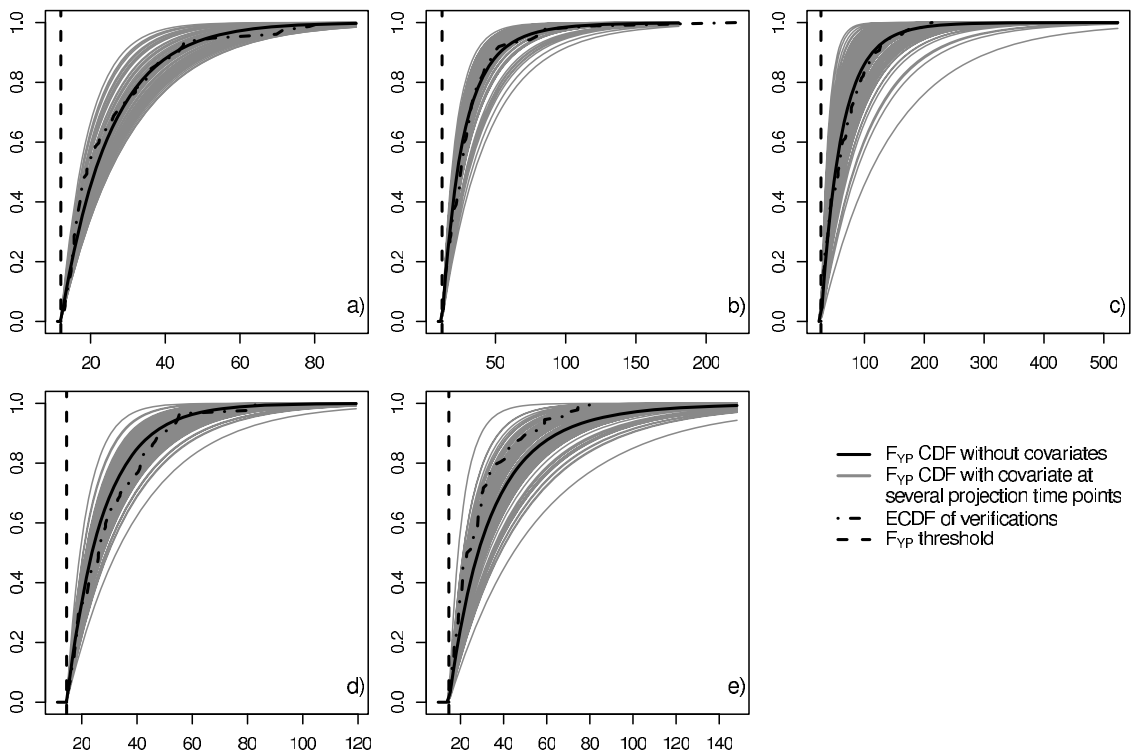
Figure 5. F_{Y_P} obtained from XCDF-t without covariates (black line), and $F_{Y_P}^t$ at different time points when including covariate information (grey lines) are shown for all five stations. For comparison reasons, the ECDF of the verifications is depicted as well (black dash-dotted line). The x axis is in mm/d; the y axis is probability.

Table 5. Shape Parameter Estimates ξ and 68% Confidence Intervals for Series Y_P Obtained With and Without Covariate Information[a]

Station	Y_P (no cov)	Y_P (cov)	Verifications	Y_C
Marseille	0.014 [0.0, 0.066]	0.009 [0.0, 0.050]	0.151 [0.009, 0.292]	0.008 [0.0, 0.072]
Perpignan	0.145 [0.057, 0.239]	0.058 [0.001, 0.146]	0.196 [0.084, 0.308]	0.159 [0.075, 0.242]
Mont-Aigoual	0.083 [0.021, 0.148]	0.026 [0.0, 0.086]	0.0 [−0.002, 0.002]	0.091 [0.039, 0.142]
Nîmes	0.025 [0.0, 0.094]	0.009 [0.0, 0.055]	0.0 [−0.001, 0.001]	0.030 [0.0, 0.095]
Sète	0.132 [0.045, 0.220]	0.046 [0.0, 0.127]	0.068 [−0.063, 0.199]	0.144 [0.067, 0.221]

[a]Also given are the shape parameter estimates for a GPD fit to the verification extremes and the local-scale data in the calibration period Y_C.

projection period. It gets also apparent, that the forms of the local-scale CDF and the large-scale CDF differ a lot, and that XCDF-t well preserves the CDF form of the local-scale variable. This demonstrates that XCDF-t does neither compress nor extend the local-scale series unintentionally. All in all we find good agreement between the ECDF of the verifications and F_{Y_P}, except for station Sète. At Sète, the observed precipitation amounts decrease from the calibration to the projection period. The large-scale data experience a directly opposed change, which leads to the discrepancy.

[42] The examination of heavy summer precipitation did not show as good results (results not presented). In winter, extremes are in general caused by frontal systems, thus local- and large-scale precipitation are commonly strongly related. In summer, however, local extreme events are often generated by processes on very small scales. Then it is not so obvious to find a large-scale predictor, whose extremes explain the local-scale extremes. The downscaling of extreme summer precipitation is still an open research challenge [*Maraun et al.*, 2010b]. The Mediterranean summer is anyway characterized by high temperatures, long periods of droughts and rather a lack of precipitation than extreme precipitation.

[43] To assess the surplus value of covariate information, we compare the F_{Y_P}'s with the ECDF of the verifications. In Figure 5, the F_{Y_P} obtained from XCDF-t without covariate information (black line) and the $F_{Y_P}^t$'s obtained with XCDF-t using covariate information at each time point t where a verification exceeds the threshold u_{Y_P} (grey lines) are shown. The ECDF of the verifications is added as black dash-dotted line. The differences between the F_{Y_P} obtained without covariate information and the ECDF of the verifications may be due to a nonstationary evolution of the observations, which can be captured by use of covariate information. Indeed, for all stations the area spanned by the grey lines covers the ECDF of the verifications completely. This hints to an improvement of the results when using covariate information.

## 5.2. Statistical Characteristics of the Inferred CDF F_{Y_P}

[44] We assess the decay of the CDF F_{Y_P} by fitting a GPD to realizations of F_{Y_P} and by interpreting the shape parameter

Table 6. Percentage of Verifications Which Exceed the Lower 95% Body of F_{Y_P}[a]

Station	F_{Y_P} (no cov)	F_{Y_P} (cov)
Marseille	**6.5**	7.8
Perpignan	6.9	**4.6**
Mont-Aigoual	**5.4**	7.8
Nîmes	**3.4**	6.9
Sète	1.1	**4.3**

[a]Values closest to 5% are boldface.

estimate (cf. equation (4)). In Table 5, the median and the 68% confidence interval of the distributions of the shape parameter ξ_{Y_P} are listed. These distributions are obtained from a GPD fitted to Y_P series of length 2000, which are generated from bootstrapped F_{Y_P} to capture the whole bandwidth of ξ_{Y_P} (for details, see Appendix C1). In case covariate information is incorporated, we use the CDFs $F_{Y_P}^t$ of the time points t when a verification exceeds the threshold u_{Y_P}. For each time point we generate 30 series, which results in about 2400 to 5000 shape parameter values. In Table 5 the shape parameter estimates of the verification extremes and the Y_C series are listed as well for comparison reasons. The shape parameters of F_{Y_P} and the verification extremes should lie close by. This is the case for stations Perpignan, Nîmes and Sète, where the shape parameter of the verification extremes is included in the confidence interval of the F_{Y_P} obtained without covariate information. The difference of the shape parameter estimates is with 0.137 the largest for station Marseille, for the other stations the difference is much smaller and does not exceed 0.083. When including covariate information, the inferred shape parameter ξ_{Y_P} is closer to the shape parameter estimate of the verifications for stations Mont-Aigual, Nîmes and Sète.
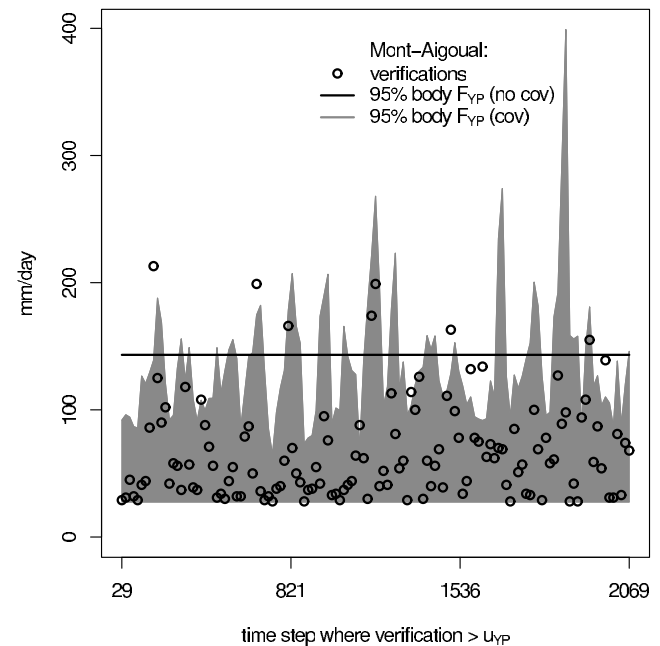
Figure 6. Lower 95% body of F_{Y_P} obtained without covariates (below black line) and with covariate information at each time point where a verification exceeds the threshold u_{Y_P} (grey area) for station Mont-Aigoual. The verification extremes are depicted as black dots.
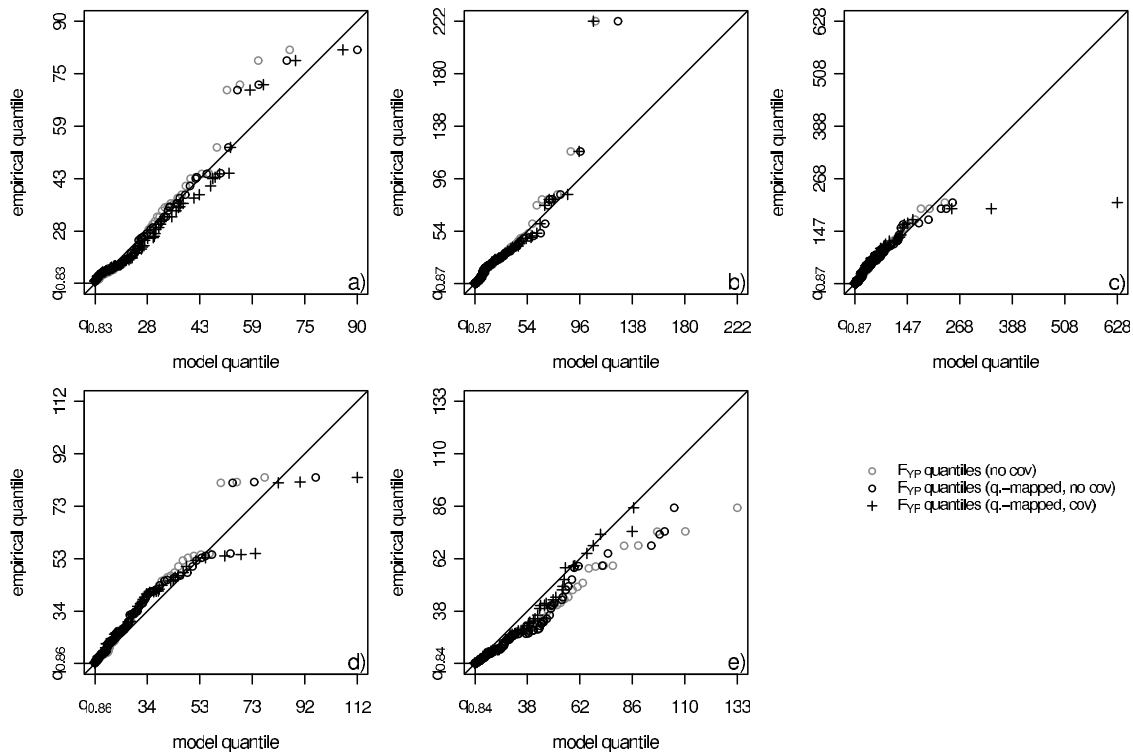
Figure 7. The q-q plots for F_{Y_P} without covariates (x axis) and ECDF of verifications (y axis) are shown (grey dots) for the stations (a) Marseille, (b) Perpignan, (c) Mont-Aigoual, (d) Nîmes, and (e) Sète. In addition, q-q plots for the ECDF of Y_P (obtained by quantile mapping X_P between F_{Y_C} and the inferred F_{Y_P}) are depicted for XCDF-t without covariate information (black dots) and with covariate information (black crosses).

[45] By means of a simulation study, we furthermore test whether applying XCDF-t results in biased shape parameters. For this study we set the projection period equal to the calibration period 1951–1985 and apply the XCDF-t. Then the shape parameter of the verifications, Y_C in this case, is compared to the distribution of shape parameters of Y_P series. For all stations but for Perpignan it shows here, that the shape parameter of the verifications does well lie within the distribution of the inferred shape parameter ξ_{Y_P}. We thus conclude that XCDF-t does not bias the decay behavior of the CDF. Results of this test are not shown.

[46] The 95% quantile is also an interesting statistical characteristics of a CDF. In Table 6, the percentages of verification extremes, which exceed this quantile of F_{Y_P} (or CDFs $F_{Y_P}^t$ in case covariate information is used), are listed. The percentages should lie close to 5%, which indeed is the case for most stations. In Figure 6, the 95% quantile of F_{Y_P}, obtained without covariate information, is exemplarily shown for station Mont-Aigual (black line). The 95% body of the $F_{Y_P}^t$ with covariate information (grey area) is depicted for all time points t when a verification (black dots) exceeds the threshold u_{Y_P}. The covariate information apparently helps to capture some of the most extreme verifications. This improvement, however, is canceled out by the large fluctuation of the 95% quantile of the $F_{Y_P}^t$, which leads to the omission of other verification extremes and therefore to the similar results listed in Table 6. All in all it shows that covariate information has to be introduced with care.
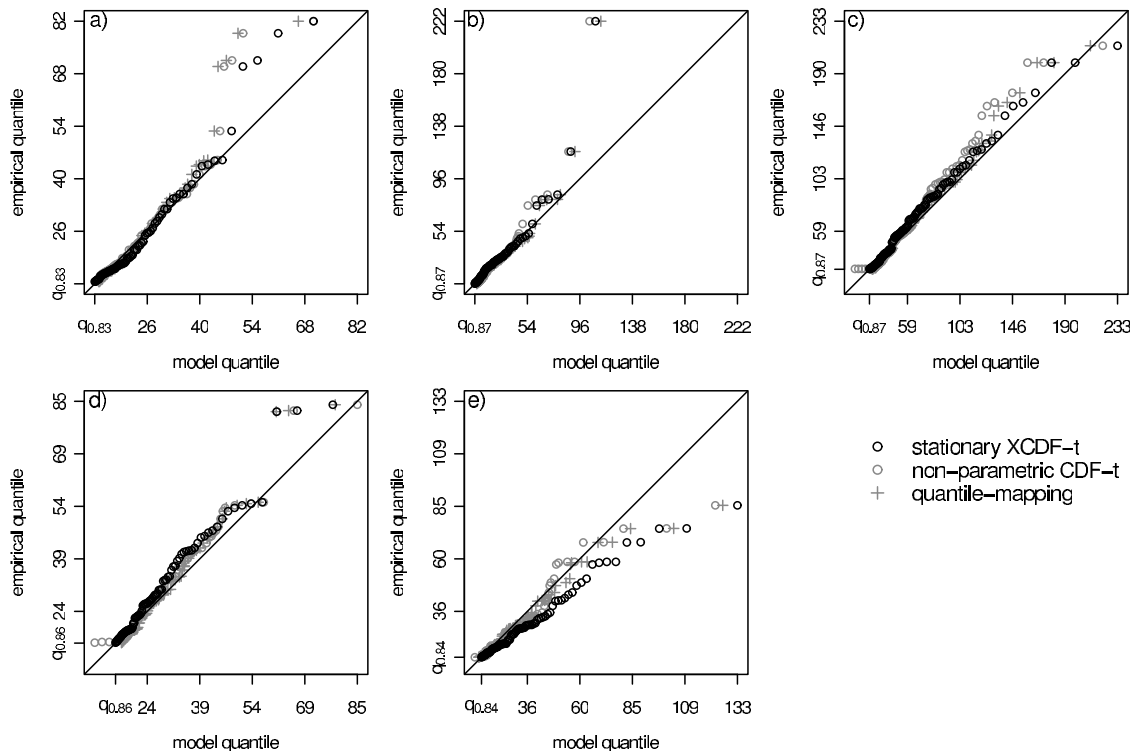
5.3. Verification

[47] The q-q plots are useful to compare the complete distributions. In Figure 7, q-q plots for the ECDF of the verifications and F_{Y_P} obtained without covariate information are depicted (grey dots). The labels of the axes include the quantile order of the thresholds (with respect to days with precipitation) to illustrate the extremity of the threshold excess series. Local-scale series Y_P in the projection period can be determined by quantile mapping X_P between the large-scale CDF F_{X_P} and the inferred local-scale CDF F_{Y_P} [cf. *Michelangeli et al.*, 2009]. In Figure 7, q-q plots between the ECDF of the in this manner determined local series Y_P and the verifications are shown as black dots for XCDF-t without covariate information and as black crosses for XCDF-t with covariate information (here the local-scale CDF $F_{Y_P}^t$ is used for quantile mapping at each time point where a large-scale

Table 7. CRPS Values, Averaged Over All Time Points Where Verification Excesses are Available[a]

Station	F_{Y_P} (no cov)	F_{Y_P} (cov)
Marseille	7.26	**7.07**
Perpignan	10.29	**10.02**
Mont-Aigoual	20.78	**19.66**
Nîmes	**7.95**	8.02
Sète	**8.22**	9.11

[a]Boldfaced values indicate a better score result.

Figure 8. The q-q plots for verification extremes and XCDF-t without covariate information (black dots, corresponding to the grey dots in Figure 7), the classic nonparametric CDF-t approach (grey dots) and quantile mapping (grey crosses) for stations (a) Marseille, (b) Perpignan, (c) Mont-Aigoual, (d) Nîmes, and (e) Sète.

extreme occurs). The standard q-q plots (grey dots) show a good agreement between the verifications and our model for stations Marseille (Figure 7a), Mont-Aigual (Figure 7c), and Nîmes (Figure 7d). For the other two stations, the probability of occurrence of the most extreme extremes is either underestimated or overestimated. We furthermore remark a proximity of the results of the standard q-q plots, where F_{Y_P} itself is used, and those of the ECDF of the Y_P series, which has been obtained by quantile mapping F_{Y_P}, see grey and black dots. For most of the stations, the ECDF of the series Y_P is even closer to the ECDF of the extreme verifications. When comparing the results obtained with and without covariate information (black dots and crosses), we find a clear improvement of the results for stations Marseille (Figure 7a) and Sète (e) when covariate information is included. For all other stations the inclusion of covariate information leads to a degradation of results. This is probably due to the temporal variability of the covariates, which do not entirely suit the local-scale data. The results might be improved by considering further covariates such as vorticity or humidity. Furthermore, transformed covariates may be of use, where lagged covariate information is taken into account, or covariate information is aggregated over several days. However, these extensions are beyond the scope of this paper.

[48] As further verification criterion, the continuous ranked probability score (crps) is employed. The crps integrates the quadratic distance of F_{Y_P} to the verification over all quantiles and corresponds therefore to the integral of the Brier score over all thresholds. The better (i.e., the closer to zero) the crps,

the closer the verification is to the center of F_{Y_P}. It is best in case of a perfect deterministic forecast, where an infinitely narrow distribution around the verifying observation has been achieved [cf. *Gneiting and Raftery*, 2007]. We integrate piecewise to approximate the crps, as, for example, proposed by *Hersbach* [2000]. In Table 7 the score results, averaged over the verification period, are listed. The inclusion of covariate information slightly improves the crps for stations Marseille, Perpignan, and Mont-Aigoual. This finding differs from the q-q plot results, because the crps aggregates information from the whole distribution.

5.4. Comparison With Other Approaches

[49] In Figure 8, q-q plots for XCDF-t (black dots, no covariate information used), the classical nonparametric CDF-t approach [cf. *Michelangeli et al.*, 2009] (grey dots), and quantile mapping [*Panofsky and Brier*, 1958] (grey crosses) are depicted. The two latter approaches are applied as well to the series of threshold excesses. The changes between calibration and verification period are small in our verification study (see Figure 4). This is advantageous for quantile mapping, because here the changes between F_{X_C} and F_{X_P} are not considered. In Figure 8 an improvement gets visible for stations Marseille (Figure 8a), and Mont-Aigual (Figure 8c) when applying XCDF-t, that is when modeling extreme values with a suitable GPD distribution. All methods have difficulties with the data of station Sète (Figure 8e), but results worsen when using XCDF-t (this would not be the case if

covariate information was used (see Figure 7)). For the other two stations, results are on a par.

6. Summary and Conclusion

[50] In this paper we present a downscaling approach, which links the CDF of large-scale extremes to the CDF of local-scale extremes by means of a transfer function. This link is assumed to stay stable between calibration and prediction period. XCDF-t does not require a realistic day-to-day representation of the large-scale variable and may be classified as model output statistics downscaling method. Only in the case that covariates are used for the modeling of the local-scale extremes, the realistic simulation of these covariates has to be verified. As result XCDF-t supplies the CDF of the local-scale extreme values in the projection time period F_{Y_P}. F_{Y_P} is obtained in parametric form, thus series distributed according to F_{Y_P} can be generated. However, the chronology of the inferred local-scale extremes in the projection period has to be derived separately, as presented, for example, by *Benestad* [2010].

[51] We assess daily precipitation in winter in Southern France for the calibration period 1951 to 1985 and the verification period 1986 to 1999. NCEP precipitation is used as large-scale variable. The results are verified by means of the 95% quantile, q-q plots, and the continuous ranked probability score. The verification extremes represent on average over all stations the upper 15% of the data at days with precipitation. It shows that an acceptable representation of extreme precipitation of Marseille, Nî, and Mont-Aigoual is obtained. Those stations are subject to very different geographic conditions: Marseille is situated on the coastline, here the moisture supply of the sea is important. Nîmes is located on the foothills of the Massif Central, and Mont-Aigoual on one of its hilltops. For the latter two stations, the interplay of specific synoptic-scale conditions with the local topography is relevant for the creation of heavy precipitation events [cf. *Berne et al.*, 2009; *Boudevillain et al.*, 2009]. Thus it appears to be feasible, even under very different geographic conditions, to deduce the CDF of local-scale precipitation extremes from the CDF of large-scale extremes.

[52] The q-q plots reveal a good estimation of the whole distribution of local-scale extremes for station Mont-Aigoual. The emergence of local extreme precipitation at this station is clearly linked to extreme precipitation within the region. For station Séte the occurrence of heavy precipitation is overestimated. This may be caused by the fact that Séte is located in the middle of several geographical spheres. The link between local-scale and large-scale extreme precipitation may differ according to the respectively dominating sphere. This is not captured by our model. For stations Perpignan, Marseille and Nîmes, the q-q plots reveal an underestimation of the most extreme events Here apparently further influences besides large-scale extreme precipitation are important for the emergence of local-scale extreme precipitation.

[53] Additional information may be added to XCDF-t in form of covariates of the parameters of the ingoing extreme value distributions. The covariates and their link functions are selected by means of the deviance statistics out of a set of appropriate covariates. A covariate typically represents one or more influences at synoptic scale or mesoscale, and thus has

different local responses. Therefore distinct significant covariates may be chosen for each station. In our analysis, sea level pressure showed to be the most useful covariate for the downscaling of winter precipitation in Southern France, whereas time averaged large-scale precipitation was only significant for station Marseille. The results show, that covariates might improve results. For most stations, XCDF-t with covariate information improved the representation of the most extreme verifications. The careful selection of covariates and their links to the GPD parameters are yet crucial.

[54] The inclusion of covariate information improves the results for Marseille, but not for Nîmes. Wind direction is a significant covariate for Marseille. Marseille is situated on the coastline, so here wind is indeed another strong influence for the emergence of extreme precipitation. For Nîmes, variables related to a synoptic depression are significant, that is slp and Z850. For this region, heavy precipitation is often accompanied by a cyclone, but its intensity does not play a major role [*Nuissier et al.*, 2008]. This may be the reason why the occurrence of heavy precipitation is overestimated when including those pressure related variables in the assessment of Nîmes. Furthermore, the local interaction between topography and synoptic flow forces and focuses the convection in the areas of Nîmes, Mont-Aigual and (to a lesser extend) Perpignan. Here it is not straightforward to establish a link between large-scale precursors and local-scale heavy precipitation [*Nuissier et al.*, 2008], which reduces the applicability of covariates in this context.

[55] An improvement of the covariates choice may be achieved by selecting them by means of score results for a verification period. To do so, the verification period must have a representative length. Furthermore, the construction of more complex covariates might be meaningful in order to account for further influencing factors on the emergence of heavy precipitation. These could be, for example, the persistence of a synoptic depression or the duration of vertical wind intensity. In the end, the determination of adequate covariates and according link functions gets very close to the modeling of physical processes.

[56] In comparison with the classical nonparametric CDF-t approach and quantile mapping, XCDF-t has worse results for station Sète, is on a par for stations Perpignan and Nîmes, and improves results for stations Marseille and Mont-Aigual.

[57] Altogether XCDF-t shows good verification results, also in comparison with the other approaches. The combination of EVT with a downscaling framework is a promising path and XCDF-t can be a useful tool in local impact studies. The R software package CDFt will be extended to contain the routines presented in this paper.

[58] Other than many classical statistical downscaling methods, the CDF-t and XCDF-t methods can be directly calibrated on large-scale GCM outputs, when no covariates are used. GCM simulations may be biased even at large scales. The GCM simulations for the 20th and 21th century for example do not represent the real temporal evolution of the large-scale weather states in the past [*Maraun et al.*, 2010b]. However, if the bias of the GCM outputs stays the same in calibration and prediction period, then GCM outputs may replace the reanalysis data as large-scale variable in both periods when applying XCDF-t without covariates. In this way IPCC SRES scenario simulations [*IPCC*, 2000] may be

used to generate local projections by means of XCDF-t without covariates.

[59] In case covariate information is used on the local-scale CDF, XCDF-t has to be calibrated on reanalysis data. The appropriateness of the in this manner calibrated framework has to be verified when being applied to GCM data. This implies that the covariates have to be realistically simulated by the GCM, also the chronology of the covariates. Then the GCM outputs can be downscaled with the inclusion of covariate information. In most of the cases observations or reanalysis data are needed to achieve realistically simulated GCM covariates, the GCM covariates may for example be nudged toward observations. Therefore XCDF-t with covariate information may rather be applied to downscale GCM outputs for feasibility studies, which is along the lines of, for example, work by *Widmann and Bretherton* [2000].

[60] For settings where the GPD distribution is not appropriate, CDF-t with and without covariates may be applied by using other distributions.

[61] Another potential application of the methodology is projection in space: Information at several sites may be combined to generate "regional" CDFs F_{Y_C}, F_{X_C}, F_{X_P} in order to infer observations at a neighboring site, where no observations have been recorded.

Appendix A: Derivation of Threshold u_{Y_P} of the Local-Scale Variable in the Projection Period

[62] $F_{\tilde{X}_C}^{-1}(F_{\tilde{X}_P}(u_{\tilde{X}_P})) = u_{\tilde{X}_C}$ holds. For $\lambda = 1$, $\tilde{X}_P$ and $\tilde{X}_C$ are shifted but not inflated. Then κ is zero (cf. equation (B4)), and equation (B5) becomes for $x = u_{\tilde{X}_P}$: $F_{Y_P}(u_{\tilde{X}_P}) = F_{Y_C}(u_{\tilde{X}_C})$. Because of the shifting $u_{\tilde{X}_C} = u_{Y_C}$ holds, therefore $F_{Y_P}(u_{\tilde{X}_P}) = 0$. Thus, u_{Y_P} is determined as $u_{Y_P} = u_{\tilde{X}_P} = u_{Y_C} + (u_{X_P} - u_{X_C})$. Hence the change Δ between calibration and projection time series is the same for local- and large-scale series: $\Delta = (u_{Y_P} - u_{Y_C}) = (u_{\tilde{X}_P} - u_{\tilde{X}_C})$. When the large-scale series have been inflated, the subtraction of the inflation correction factor κ leads to

$$u_{Y_P} = u_{Y_C} + \left(u_{\tilde{X}_P} - u_{\tilde{X}_C}\right) - \kappa$$
$$= u_{Y_C} + (u_{X_P} - u_{X_C}), \tag{A1}$$

and thus the threshold of the not inflated series is restored.

Appendix B: Data Preprocessing: Inflation and Shifting

[63] We inflate and shift the large-scale series, so that Y_C and X_C have approximately the same range before applying XCDF-t. Otherwise, the edges of the CDF F_{Y_P} may not be obtained. An alternative way to tackle this problem is, for example, to use shifted edges of the CDF F_{Y_C} to complete F_{Y_P} [*Déqué*, 2007]. However, by the here proposed inflation and shifting, the fully parametric form of F_{Y_P} is maintained, which can be used to generate new series (see Appendix C). In addition, this preprocessing helps to avoid unwanted compression or extension of F_{Y_P} with respect to F_{Y_C} when applying XCDF-t, which might otherwise occur due to the unboundedness of the upper tail of a GPD with positive shape parameter. Inflation is not necessary in case we constrain the procedure so that the inferred CDF F_{Y_P} is also a GPD (see equation (4)).

[64] Let the inflation factor be λ with

$$\lambda = \frac{\max(Y_C) - u_{Y_C}}{\max(X_C - u_{X_C} + u_{Y_C}) - u_{Y_C}}. \tag{B1}$$

The inflated series are furthermore shifted, so that local- and large-scale series have the same threshold in the calibration period. Thus we obtain

$$
\begin{aligned}
\tilde{X}_C &= \lambda \times X_C + (u_{Y_C} - \lambda \times u_{X_C}), \\
\tilde{X}_P &= \lambda \times X_P + (u_{Y_C} - \lambda \times u_{X_C}), \\
u_{\tilde{X}_C} &= \lambda \times u_{X_C} + (u_{Y_C} - \lambda \times u_{X_C}) = u_{Y_C}, \\
u_{\tilde{X}_P} &= \lambda \times u_{X_P} + (u_{Y_C} - \lambda \times u_{X_C}).
\end{aligned} \tag{B2}
$$

Now XCDF-t is given by

$$F_{Y_P}(x) = F_{Y_C}\left(F_{\tilde{X}_C}^{-1}\left(F_{\tilde{X}_P}(x)\right)\right). \tag{B3}$$

Inflating and shifting change the XCDF-t transfer function $T(\cdot)$. However, the effect of inflation on the resulting CDF F_{Y_P} is very small, this has been tested by means of simulation studies (results not shown). In case local- and large-scale variables have approximately the same range, the effect of inflation gets negligible.

[65] When applying XCDF-t, the change between calibration and projection time period for the local-scale CDF is deduced from the change Δ the large-scale CDF has experienced. We therefore use variables with the same units, so even the quantity of the transferred change is interpretable. To maintain this feature in case of inflation, we readjust by using the thresholds: Without inflation, local- and large-scale variable thresholds would experience the same change between calibration and projection period, that is $\Delta = u_{X_P} - u_{X_C} = u_{Y_P} - u_{Y_C}$ (see Appendix A). Inflation causes a change $\tilde{\Delta} = \lambda \times \Delta$. To correct for this, we subtract the correction factor

$$\kappa = \frac{(\lambda - 1)}{\lambda}\left(u_{\tilde{X}_P} - u_{\tilde{X}_C}\right) = (\lambda - 1)\Delta \tag{B4}$$

from u_{Y_P} and all quantiles of F_{Y_P}, and equation (B3) becomes

$$F_{Y_P}(x) = F_{Y_C}\left(F_{\tilde{X}_C}^{-1}\left(F_{\tilde{X}_P}(x + \kappa)\right)\right). \tag{B5}$$

This equation is finally used to deduce equation (3).

Appendix C: Generation of Realizations of F_{Y_P}

[66] Let x be distributed according to F_{Y_P}, then $F_{Y_P}(x) = u$ (with u being uniformly distributed). Thus realizations of F_{Y_P} are generated by transforming equation (B5) to

$$x = F_{\tilde{X}_P}^{-1}\left(F_{\tilde{X}_C}\left(F_{Y_C}^{-1}(u)\right)\right) - \kappa. \tag{C1}$$

[67] To account for parameter estimation uncertainty, the following bootstrap procedure is applied to generate series Y_P:

[68] 1. Estimate parameters for F_{Y_C}, $F_{\tilde{X}_C}$, and $F_{\tilde{X}_P}$ from the data and obtain F_{Y_P} by means of XCDF-t. Then simulate a series Y_P from F_{Y_P}.

[69] 2. Generate new series Y_C^*, $\tilde{X}_C^*$, and $\tilde{X}_P^*$ from F_{Y_C}, $F_{\tilde{X}_C}$, and $F_{\tilde{X}_P}$. From these series, estimate parameters for $F_{Y_C}^*$,

$F_{\tilde{X}_C}^*$, and $F_{\tilde{X}_P}^*$, and obtain $F_{\tilde{Y}_P}^*$ by applying XCDF-t. Then simulate a series Y_P^* from $F_{\tilde{Y}_P}^*$. Repeat this step R times, e.g., $R = 1000$.

[70] The Y_P^* are used to create confidence bands for Y_P itself or for characteristics of Y_P.

References

Ahmad, M. I., C. D. Sinclair, and B. D. Spurr (1988), Assessment of flood frequency models using empirical distribution function statistics, *Water Resour. Res.*, 24(8), 1323–1328.

Benestad, R. E. (2010), Downscaling precipitation extremes: Correction of analog models through PDF predictions, *Theor. Appl. Climatol.*, 100(1–2), 1–21, doi:10.1007/s00704-009-0158-1.

Berne, A., G. Delrieu, and B. Boudevillain (2009), Variability of the spatial structure of intense Mediterranean precipitation, *Adv. Water Resour.*, 32, 1031–1042.

Bosilovich, M., J. Chen, F. R. Robertson, and R. F. Adler (2008), Evaluation of global precipitation in reanalyses, *J. Appl. Meteorol. Climatol.*, 47, 2279–2299, doi:10.1175/2008JAMC1921.1.

Boudevillain, B., et al. (2009), Projet Cyprim, partie I. cyclogenèses et précipitations intenses en région méditerranéenne: Origines et caractéristiques, *Meteorologie*, 66, 18–28.

Cannon, A. J. (2011), Quantile regression neural networks: Implementation in R and application to precipitation downscaling, *Comput. Geosci.*, doi:10.1016/j.cageo.2010.07.005, in press.

Chavez-Demoulin, V., and A. C. Davison (2005), Generalized additive modelling of sample extremes, *J. R. Stat. Soc., Ser. C*, 54(1), 207–222.

Choulakian, V., and M. A. Stephens (2001), Goodness-of-fit tests for the Generalized Pareto Distribution, *Technometrics*, 43(4), 478–484.

Christensen, J. H., T. R. Carter, M. Rummukainen, and G. Amanatidis (2007), Evaluating the performance and utility of regional climate models: The PRUDENCE project, *Clim. Change*, 81, suppl., 1–6, 2007.

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

Davison, A. C. (2003), *Statistical Models*, Cambridge Univ. Press, Cambridge, U. K.

Déqué, M. (2007), Frequency of precipitation and temperature extremes over France in an anthropogenic scenario: Model results and statistical correction according to observed values, *Global Planet. Change*, 57, 16–26.

Ducrocq, V., O. Nuissier, D. Ricard, C. Lebeaupin, and T. Thouvenin (2008), A numerical study of three catastrophic precipitating events over southern France. II: Mesoscale triggering and stationarity factors, *Q. J. R. Meteorol. Soc.*, 134, 131–145, doi:10.1002/qj.199.

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

Fisher, R. A., and L. H. C. Tippett (1928), Limiting forms of the frequency distribution of the largest or smallest member of a sample, *Proc. Cambridge Philos. Soc.*, 24, 180–190.

Fowler, H. J., S. Blenkinsop, and C. Tebaldi (2007), Linking climate change modelling to impact studies: Recent advances in downscaling techniques for hydrological modelling, *Int. J. Climatol.*, 27, 1547–1578, doi:10.1002/joc.1556.

Fowler, H. J., D. Cooley, S. R. Sain, and M. Thurston (2010), Detecting change in UK extreme precipitation using results from the climatepredic-tion.net BBC climate change experiment, *Extremes*, 13, 241–267, doi:10.1007/s10687-010-0101-y.

Friederichs, P., and A. Hense (2007), Statistical downscaling of extreme precipitation events using censored quantile regression, *Mon. Weather Rev.*, 135, 2365–2378, doi:10.1175/MWR-3403.1.

Furrer, E. M., and R. W. Katz (2008), Improving the simulation of extreme precipitation events by stochastic weather generators, *Water Resour. Res.*, 44, W12439, doi:10.1029/2008WR007316.

Gnedenko, B. V. (1943), Sur la distribution limite du terme maximum d'une série aléatoire, *Ann. Math.*, 44, 423–453.

Gneiting, T., and A. E. Raftery (2007), Strictly proper scoring rules, prediction, and estimation, *J. Am. Stat. Assoc.*, 102, 359–378.

Goodess, C. M., et al. (2011), An intercomparison of statistical downscaling methods for Europe and European regions—Assessing their performance with respect to extreme weather events and the implications for climate change applications, technical report, Univ. of East Anglia, Norwich, U. K., in press.

Haddad, Z. S., and D. Rosenfeld (1997), Optimality of empirical Z-R relations, *Q. J. R. Meteorol. Soc.*, 123, 1283–1293.

Hersbach, H. (2000), Decomposition of the continuous ranked probability score for ensemble prediction systems, *Weather Forecasting*, 15, 559–570.

Intergovernmental Panel on Climate Change (IPCC) (2000), *Special Report on Emissions Scenarios: A Special Report of Working Group III of the Intergovernmental Panel on Climate Change*, edited by M. Nakicenovic et al., Cambridge Univ. Press, Cambridge, U. K.

Intergovernmental Panel on Climate Change (IPCC) (2007), *Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by M. L. Parry et al., Cambridge Univ. Press, Cambridge, U. K.

Katz, R. W. (2002), Techniques for estimating uncertainty in climate change scenarios and impact studies, *Clim. Res.*, 20, 167–185.

Katz, R. W., M. B. Parlange, and P. Naveau (2002), Statistics of extremes in hydrology, *Adv. Water Resour.*, 25, 1287–1304.

Khaliq, M. N., T. B. M. J. Ouardab, J.-C. Ondob, P. Gachona, and B. Bobéeb (2006), Frequency analysis of a sequence of dependent and/or nonstationary hydro-meteorological observations: A review, *J. Hydrol.*, 329,534–552, doi:10.1016/j.jhydrol.2006.03.004.

Kundzewicz, Z. W., et al. (2007), Freshwater resources and their management, in *Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by M. L. Parry et al., chap. 3, pp. 173–210, Cambridge Univ. Press, Cambridge, U. K.

Leadbetter, M. R. G., G. Lindgren, and H. Rootzen (1983), *Extremes and Related Properties of Random Sequences and Series*, Springer, New York.

Lenderink, G., A. Buishand, and W. van Deursen (2007), Estimates of future discharges of the river Rhine using two scenario methodologies: Direct versus delta approach, *Hydrol. Earth Syst. Sci.*, 11(3), 1145–1159, 2007.

Maraun, D., H. W. Rust, and T. J. Osborn (2010a), Synoptic airflow and UK daily precipitation extremes: Development and validation of a vector generalised linear model, *Extremes*, 13(2), 133–153.

Maraun, D., et al. (2010b), Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user, *Rev. Geophys.*, 48, RG3003, doi:10.1029/2009RG000314.

Maraun, D., T. J. Osborn, and H. W. Rust (2011), The influence of synoptic airflow on UK daily precipitation extremes. Part I: Observed spatiotemporal relationships, *Clim. Dyn.*, 36(1–2), 261–275, doi:10.1007/s00382-009-0710-9.

Meehl, G. A., et al. (2007), Global climate projections, in *Climate 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., chap. 10, pp. 10-2–10-8, Cambridge Univ. Press, Cambridge, U. K.

Michelangeli, P.-A., M. Vrac, and H. Loukos (2009), Probabilistic downscaling approaches: Application to wind cumulative distribution functions, *Geophys. Res. Lett.*, 36, L11708, doi:10.1029/2009GL038401.

Naveau, P., M. Nogaj, C. Ammann, P. Yiou, D. Cooley, and V. Jomelli (2005), Statistical methods for the analysis of climate extremes, *C. R. Geosci.*, 337, 1013–1022.

Nuissier, O., V. Ducrocq, D. Ricard, C. Lebeaupin, and S. Anquetin (2008), A numerial study of three catastrophic precipitating events over southern France. I: Numerical framework and synoptic ingredients, *Q. J. R. Meteorol. Soc.*, 134, 111–130.

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

Panofsky, H. A., and G. W. Brier (1958), *Some Applications of Statistics to Meteorology*, Pa. State Univ. Press, Univ. Park.

Quintana Seguí, P., A. Ribes, E. Martin, F. Habets, and J. Boé (2010), Comparison of three downscaling methods in simulating the impact of climate change on the hydrology of Mediterranean basins, *J. Hydrol.*, *383*, 111–124.

Reiss, R. D., and M. Thomas (1997), *Statistical Analysis of Extreme Values with Applications to Insurance, Finance, Hydrology and Other Fields*, Birkhauser, Basel, Switzerland.

Rummukainen, M. (1997), Methods of statistical downscaling of GCM simulations, *Tech. Rep. 80*, Swed. Meteorol. and Hydrol. Inst., Norrköping, Sweden.

Toreti, A., E. Xoplaki, D. Maraun, F. G. Kuglitsch, H. Wanner, and J. Luterbacher (2010), Characterisation of extreme winter precipitation in Mediterranean coastal sites and associated anomalous atmospheric circulation patterns, *Nat. Hazards Earth Syst. Sci.*, *10*, 1037–1050, doi:10.5194/nhess-10-1037-2010.

Vidal, J.-P., and S. Wade (2009), A multimodel assessment of future climatological droughts in the United Kingdom, *Int. J. Climatol.*, *29*, 2056–2071, doi:10.1002/joc.1843.

von Storch, H., and F. Zwiers (1999), *Statistical Analysis in Climate Research*, Cambridge Univ. Press, Cambridge, U. K.

Vrac, M., and P. Naveau (2007), Stochastic downscaling of precipitation: From dry events to heavy rainfalls, *Water Resour. Res.*, *43*, W07402, doi:10.1029/2006WR005308.

Vrac, M., M. Stein, and K. Hayhoe (2007), Statistical downscaling of precipitation through nonhomogeneous stochastic weather typing, *Clim. Res.*, *34*, 169–184, doi:10.3354/cr00696.

Wang, X. L., F. W. Zwiers, and V. R. Swail (2004), North Atlantic Ocean wave climate change scenarios for the twenty-first century, *J. Clim.*, *17*, 2368–2383.

Widmann, M., and C. S. Bretherton (2000), Validation of mesoscale precipitation in the NCEP reanalysis using a new gridcell dataset for the northwestern United States, *J. Clim.*, *13*, 1936–1950.

Wigley, T. M. L., P. D. Jones, K. R. Briffa, and G. Smith (1990), Obtaining sub-grid scale information from coarse resolution general circulation model output, *J. Geophys. Res.*, *95*, 1943–1953.

Wilby, R. L., T. M. L. Wigley, D. Conway, P. D. Jones, B. C. Hewitson, J. Main, and D. S. Wilks (1998), Statistical downscaling of general circulation model output: A comparison of methods, *Water Resour. Res.*, *34*(11), 2995–3008.

Wilks, D. S. (2006), *Statistical Methods in the Atmospheric Sciences*, volume 91 of *International Geophysics Series*, Academic, Amsterdam.

Wilks, D. S., and R. L. Wilby (1999), The weather generation game: A review of stochastic weather models, *Prog. Phys. Geogr.*, *23*(3), 329–357, doi:10.1177/030913339902300302.

M. Kallache and P.-A. Michelangeli, CLIMPACT, 79 Rue du Faubourg Poissonnière, F-75009 Paris, France. (mk@climpact.com)

P. Naveau and M. Vrac, Laboratoire des Sciences du Climat et de l'Environnement, IPSL, CEA Saclay Batiment 701, Orme des Merisiers, F-91191 Gif-sur-Yvette, France.