

Improving pan-European hydrological simulation of extreme events through statistical bias correction of RCM-driven climate simulations

R. Rojas¹, L. Feyen¹, A. Dosio², and D. Bavera¹

¹Land Management and Natural Hazards Unit, Institute for Environment and Sustainability (IES), Joint Research Centre (JRC), European Commission (EC), Ispra, Italy

²Climate Change & Air Quality Unit, Institute for Environment and Sustainability (IES), Joint Research Centre (JRC), European Commission (EC), Ispra, Italy

Received: 1 April 2011 – Published in Hydrol. Earth Syst. Sci. Discuss.: 19 April 2011

Revised: 29 June 2011 – Accepted: 12 August 2011 – Published: 24 August 2011

Abstract. In this work we assess the benefits of removing bias in climate forcing data used for hydrological climate change impact assessment at pan-European scale, with emphasis on floods. Climate simulations from the HIRHAM5-ECHAM5 model driven by the SRES-A1B emission scenario are corrected for bias using a histogram equalization method. As target for the bias correction we employ gridded interpolated observations of precipitation, average, minimum, and maximum temperature from the E-OBS data set. Bias removal transfer functions are derived for the control period 1961–1990. These are subsequently used to correct the climate simulations for the control period, and, under the assumption of a stationary error model, for the future time window 2071–2100. Validation against E-OBS climatology in the control period shows that the correction method performs successfully in removing bias in average and extreme statistics relevant for flood simulation over the majority of the European domain in all seasons. This translates into considerably improved simulations with the hydrological model of observed average and extreme river discharges at a majority of 554 validation river stations across Europe. Probabilities of extreme events derived employing extreme value techniques are also more closely reproduced. Results indicate that projections of future flood hazard in Europe based on uncorrected climate simulations, both in terms of their magnitude and recurrence interval, are likely subject to large errors. Notwithstanding the inherent limitations of the large-

scale approach used herein, this study strongly advocates the removal of bias in climate simulations prior to their use in hydrological impact assessment.

1 Introduction and scope

Europe has experienced heavy floods over the last decades, which have affected thousands of people and caused millions of Euros worth of damage. Even though it is yet impossible to link recent flooding events to global warming, climate projections indicate that in the future we can expect more extreme weather events triggering flooding. Managing risks from extreme flood events will be a crucial component of climate change adaptation. It is therefore of utmost importance to develop and implement techniques that enhance the confidence in projecting future trends in flood occurrence and intensity.

The basis for the definition of potential impacts of global warming are climate predictions. To date the most advanced tools to obtain those predictions are coupled Atmosphere-Ocean General Circulation Models (AOGCMs or in short GCMs) (Giorgi, 2005). Owing mainly to their coarse horizontal resolution (ca. 100–300 km), however, downscaling procedures are required to feed small-scale impact models and to guarantee a correct representation of the hydrologic processes at a much finer spatial resolution (Fowler and Kilsby, 2007). Possible strategies to achieve this include statistical or dynamical downscaling. The former develops statistical relationships between large-scale climate information



Correspondence to: R. Rojas
(rodrigo.rojas@jrc.ec.europa.eu)

and regional variables (Wilby et al., 1999), whereas the latter considers the application of Regional Climate Models (RCMs) driven by boundary conditions obtained from GCMs (Fowler et al., 2007). Advantages and drawbacks of both downscaling techniques have been widely discussed in literature (see, e.g. Wilby and Wigley, 1997; Murphy, 1999; Hellström et al., 2001; Haylock et al., 2006; Schmidli et al., 2007) and, hence, they will not be repeated here. For excellent discussions of downscaling techniques with focus on hydrological applications the reader is referred to Wood et al. (2004) and Fowler et al. (2007). In this paper, the dynamical downscaling approach is employed.

Despite the fact that hydrological components (e.g. surface and subsurface runoff) can be directly obtained from RCMs at lateral resolutions that agree with meso-scale catchments (e.g. 25×25 km or 50×50 km), these components are hardly ever used to assess the local hydrology. This is mainly due to the poor representation of the land surface processes, which results in a poor agreement between the surface runoff simulated by the RCMs and observations (Giorgi et al., 1994; Evans, 2003). Instead, the detailed climate information obtained from the coupled GCM-RCM is typically employed to force off-line hydrological models and, thus, to obtain more accurate representations of meso- and small-scale hydrologic processes. In recent years, several studies have appeared that implement this technique to assess the impacts of climate change on hydrological extremes, either at the local (e.g. Middelkoop et al., 2001; Etchevers et al., 2002; Prudhomme et al., 2003; Kleinn et al., 2005; Wilby et al., 2006; Steele-Dunne et al., 2008; Thorne, 2011), regional (e.g. Graham et al., 2007) or continental (e.g. Lehner et al., 2006; Dankers and Feyen, 2008, 2009) scale.

Notwithstanding RCMs have considerably advanced in reproducing regional and local climate, they are known to feature systematic errors (see, e.g. Jacob et al., 2007; Lenderink, 2010; Suklitsch et al., 2010). These biases are likely explained by model errors caused by imperfections in the climatic model conceptualization, discretization and spatial averaging within cells, and uncertainties conveyed from the GCM to the RCM (Teutschbein and Seibert, 2010). Particularly, small-scale patterns of precipitation are highly dependent on climate model resolution and parametrization. At the same time, some RCMs show systematic biases with a clear tendency to enhance these biases in more extreme cold or warm conditions (van der Linden and Mitchell, 2009). In a pan-European context, it has been noticed that climate models tend to overestimate warm summers in south-eastern Europe whereas precipitation in winter is too abundant in northern Europe (Jacob et al., 2007; Christensen et al., 2008). It has also been found that areas with a warm bias during winter generally exhibit a wet bias, whereas areas with a cold winter bias show a dry bias (Jacob et al., 2007). At the same time, Kjellström et al. (2010) found a clear tendency toward a warm bias in northern Europe for daily minimum temperature. This potentially has a significant impact on the simula-

tion of hydrological processes such as spring-flooding events driven by snow melting. A tendency to simulate too much (daily) precipitation by different RCMs, with this tendency being more pronounced for the upper-end percentiles (i.e. extreme precipitation events), has also been noticed (Kjellström et al., 2010; Lenderink, 2010). Also a persistent overestimation of the wet day frequency is generally observed (see, e.g. Leander and Buishand, 2007; Piani et al., 2010b). It should be highlighted that these biases can in part be explained by errors in the observational data set employed for comparison (Lenderink, 2010).

The presence of biases in the forcing data seriously limits its use in hydrological impact assessments (see, e.g. Wood et al., 2004) and it can result in unwanted uncertainty regarding projected climate change (van der Linden and Mitchell, 2009). More specifically, RCM outputs not corrected for biases tend to produce inaccurate probabilities for extreme events, thus rendering the extreme value analysis less reliable (Durman et al., 2001). Hence, some form of prior bias correction of the forcing data is required if a realistic description of the hydrology is sought. This procedure should aim at correcting climate simulated by the RCM during a control period to properly reflect the spatio-temporal patterns of the observed climate and, subsequently, use the “transfer function” between climate observations and simulations obtained for the control period to correct future climate simulations (Piani et al., 2010a).

In recent literature, the presence of bias in dynamically downscaled outputs has been amply recognized (see, e.g. Shabalova et al., 2003; Lenderink et al., 2007; van Pelt et al., 2009; Hurkmans et al., 2010). As a result, several techniques to correct potential bias in precipitation and temperature have been developed. Some use (a) monthly correction factors based on the ratio between present-day simulated values and observed values (Durman et al., 2001), (b) linear or non-linear transformation functions which consider changes in the mean and the variance of the observed and simulated time series (Horton et al., 2006; Leander and Buishand, 2007; Leander et al., 2008), (c) probability distribution transfer functions derived from observed and simulated cumulative distribution functions (cdfs), which is also referred to as “quantile mapping” or “histogram equalization” (Déqué, 2007; Block et al., 2009; Piani et al., 2010a,b), and (d) empirical factors to tailor the RCM outputs considering normalization, tuning of the standard deviation and calculation of residuals (Engen-Skaugen, 2007). Themeßl et al. (2011) evaluated an ensemble of seven empirical-statistical methods to correct bias in daily precipitation of a high-resolution regional climate hindcast for the Alpine region. They found that quantile mapping shows the best performance, particularly at high quantiles, which is favourable for applications related to extreme precipitation events such as flooding. For a recent review of bias correction techniques with focus on hydrological applications the reader is referred to Teutschbein and Seibert (2010).

Most of the above bias removal techniques have been applied at the local catchment scale, where dense networks of meteorological stations are available to reconstruct recent climate. With the exception of the regional study of Graham et al. (2007), all large-scale hydrological assessments to date have lacked a consistent bias correction step aimed at ensuring a correct spatio-temporal description of observed present climatic conditions. We note that there have been applications of the delta-approach at pan-European scale (see, e.g. Lehner et al., 2006), where for the control period the hydrological model has been forced by observed climate and for future climate projections the climate signal is added to current climate, hereby avoiding the need for bias correction. However, Lehner et al. (2006) used monthly time series of observed climate at 0.5° lat-long resolution (New et al., 2000) and only accounted for long-term trends and average changes in seasonal climate, neglecting a potential change in climate variability and extreme events.

The main reason advocated to skip bias correction in pan-European impact studies has been the lack of a high-quality and high-resolution meteorological data set with sufficient observation length with which to confront climate simulations of present climate. In this regard, the meteorological observations data set E-OBS (see Haylock et al., 2008), recently developed in the European project ENSEMBLES (van der Linden and Mitchell, 2009), provides the opportunity to create fully consistent bias corrected time series of precipitation and temperatures to force pan-European hydrological models. It consists of European land-only daily high-resolution gridded data for precipitation and minimum, maximum and mean surface temperature. This data set improves on previous products in its spatial resolution and extent, time period, number of contributing stations and attention to finding the most appropriate method for spatial interpolation of daily climate observations (Haylock et al., 2008). Also newly available is the global dataset of observed meteorological forcing data (Weedon et al., 2010) of the last 50 yr from the EU project WATCH (<http://www.eu-watch.org>). The data are derived from the ERA-40 reanalysis product via sequential interpolation to 0.5° lat-long resolution, elevation correction and monthly-scale adjustments based on CRU and GPCC monthly observations combined with new corrections for varying atmospheric aerosol-loading and separate precipitation gauge corrections for rainfall and snowfall.

Based on the recently established pan-European and global meteorological datasets Piani et al. (2010a,b) have proposed a “statistical bias correction” method based on quantile mapping that corrects daily values of mean, maximum, and minimum temperatures and precipitation, with the latter respecting the original precipitation-to-snow ratio. This technique works by fitting pre-defined “transfer functions” between climate observations and simulations for a given control period. A clear advantage of this technique is the flexibility of the fitting procedure by reaching a trade-off between robustness and goodness-of-fit of the alternative

“transfer functions”, and the preservation of seasonal statistics even if the bias correction is performed daily. To date, this technique has not been applied in any pan-European hydrological impact assessment.

This work presents an assessment of the benefits of correcting the bias in regional climate simulations for hydrological impact assessments at pan-European scale, with an emphasis on hydrological extreme events. To this end, we employ the bias correction method recently developed by Piani et al. (2010b). This method corrects for errors not only in the mean but also in the shape of the distribution. Dosio and Paruolo (2011) recently showed that this method is also capable of correcting the high-end percentiles of the distribution of precipitation for an ensemble of RCMs for Europe. It is therefore capable to correct for errors in the variability as well, which is crucial for extreme event analysis. As target for the bias correction we make use of the E-OBS data set (Haylock et al., 2008). We correct daily RCM fields of mean, maximum and minimum temperatures, as well as daily precipitation. The uncorrected and bias-corrected climate data are evaluated against observed climate with respect to average and extreme statistics relevant for flood simulation. The original and bias-corrected climate data are then used to force the hydrological model LISFLOOD. This spatially-distributed model has been developed for operational flood forecasting at European scale (van der Knijff et al., 2010) and has recently been applied in pan-European climate change impact assessments (Dankers and Feyen, 2008, 2009; Feyen and Dankers, 2009). Employing extreme value analysis techniques, the probability of extreme discharges is estimated and compared to results derived from long time observed discharge series at 554 stations in Europe.

The remainder of this paper is arranged as follows. Section 2 provides details on the data and methodology used in this work. This includes a description of the observed and climate data, the bias correction method, the hydrological modelling framework and the extreme value analysis for evaluating changes in hydrological extremes. We report our results in Sect. 3, provide a more in-depth discussion in Sect. 4 and offer concluding thoughts in Sect. 5.

2 Data and methods

2.1 Observed and simulated climate

Climate simulations are assessed on the basis of the high-resolution gridded E-OBS data set (v3.0) (Haylock et al., 2008) (publicly available from <http://eca.knmi.nl/>). The aim of the E-OBS data set is to represent daily areal values in alternative grid-boxes (i.e. 0.5° and 0.25° regular lon-lat grids, and 0.44° and 0.22° rotated-pole grids). The data set covers the same spatial extent as RCMs in ENSEMBLES for the time period 1950–2006. The station network used for interpolation in E-OBS comprises ca. 3000 stations for

precipitation and ca. 1900 stations for temperature, spread (unevenly) over Europe. A robust three-step process to interpolate daily observations was employed; first, interpolation of monthly precipitation totals and monthly mean temperature, second, interpolation of daily anomalies for precipitation and temperature and, third, combination of monthly and daily estimates (Haylock et al., 2008). The E-OBS data set has been specially designed to represent grid box estimates, instead of point values. This is essential to enable a direct comparison with results obtained from RCMs (see, e.g. Chen and Knutson, 2008).

Daily climate simulations are obtained from the RCM HIRHAM5 (Christensen et al., 2007) of the Danish Meteorological Institute (DMI), driven by the GCM ECHAM5 (Roeckner et al., 2003) of the Max Planck Institute for Meteorology (MPI), downloaded from the FP6 project ENSEMBLES website (<http://ensemblesrt3.dmi.dk/>). In the framework of ENSEMBLES, the HIRHAM5-ECHAM5 model was run for the period 1961–2100 with a lateral resolution of ca. 25 km (0.22° rotated-pole grids) and forced according to the SRES-A1B scenario of the Intergovernmental Panel on Climate Change (IPCC) (Nakicenovic and Swart, 2000). We selected this climate model, on the basis of a preliminary evaluation of a large number of regional climate simulations from the ENSEMBLES project. In that context, the HIRHAM5-ECHAM5 model showed to be one of the most deficient in reproducing present climate conditions in the period 1961–1990 (defined as control period in this study) when compared to the E-OBS data set.

2.2 Statistical bias correction

The bias correction method employed in this work falls within the category “histogram equalization” and it has been described in detail in Piani et al. (2010a,b). In this technique, the corrected variable (x_{cor}) is a function of the simulated (x_{sim}) counterpart given as $x_{\text{cor}} = f(x_{\text{sim}})$. The function f is defined such that the intensity histograms of both corrected (x_{cor}) and observed (x_{obs}) variables match. As demonstrated by Piani et al. (2010a), the function f (also referred to as “transfer function”) can be obtained by estimating the cumulative distribution functions (cdfs) of x_{obs} and x_{sim} and, subsequently, associating to each value of x_{sim} the value of x_{obs} such that $\text{cdf}_{\text{sim}}(x_{\text{sim}}) = \text{cdf}_{\text{obs}}(x_{\text{obs}})$.

Following Piani et al. (2010b), two functional forms are used to perform the bias correction of precipitation at the grid-cell level,

$$x_{\text{cor}} = a + bx_{\text{sim}} \quad (1)$$

$$x_{\text{cor}} = (a + bx_{\text{sim}}) \times (1 - e^{-(x_{\text{sim}} - x_0)/\tau}), \quad (2)$$

where a , b , x_0 , and τ are parameters of the function to be fitted. In Eq. (1) (linear case), a corresponds to an additive correction factor whereas b is a multiplicative factor. Piani

et al. (2010b) suggest that in some regions the transfer function is well approximated by a linear function at high intensities, but a systematic change of slope occurs at the lowest intensities. Based on this, they suggest Eq. (2), which represents an exponential tendency to an asymptote. Here, the asymptote is given by the linear factor ($a + bx_{\text{sim}}$), whereas τ defines the rate at which the asymptote is approached and x_0 is the “dry day correction” factor (value of precipitation below which x_{sim} is set to zero), defined here as $-a/b$. In addition to Eqs. (1) and (2), Piani et al. (2010b) also proposed a logarithmic fit, which however turned out to be less suitable due to fit errors.

From a global analysis, Piani et al. (2010b) concluded that any of these two functions may do a good job in correcting climate simulations, usually showing little improvement when moving from a two-parameter function (Eq. 1) to a four-parameter function (Eq. 2). They noted, however, that where fitting errors were high for Eq. (1), Eq. (2) performed better. As a consequence, the linear model is generally used in most cases resorting to the exponential tendency to an asymptote model (Eq. 2) when the performance of the linear model is unsatisfactory, i.e. a trade-off between robustness and goodness-of-fit.

To perform the bias correction of precipitation, series of simulated and observed daily values within a “construction period” (Y) of length n years are selected for every month m and for each grid cell. That is, x_{sim}^m (x_{obs}^m) for month m ($m = 1, \dots, 12$) is given by $\{x_{\text{sim}}^m(x_{\text{obs}}^m) : x \in Y_{i,m} ; i = 1, \dots, n\}$ (for sake of clarity we are ignoring the spatial index). Subsequently, Eq. (1) or (2) are fitted using x_{sim}^m and x_{obs}^m and monthly correction transfer functions are obtained for each grid cell. Daily precipitation values are then obtained by interpolating monthly transfer functions into daily transfer functions, using as reference points the middle-day of each month. When the estimated transfer functions for two consecutive months are both linear or both exponential-type, the daily transfer functions are obtained by a linear interpolation of the parameters of both monthly transfer functions. On the contrary, when monthly transfer functions are of a different type, an interpolation scheme is implemented that preserves the characteristics of both linear and exponential-type transfer functions. For a detailed description of this interpolation scheme the reader is referred to Piani et al. (2010b).

Following Piani et al. (2010b), only wet days (i.e. days with more than 1 mm of precipitation) are considered to perform the fitting of one of the two proposed transfer functions (Eqs. 1 or 2) through ordinary least squares (OLS) or non-linear least squares (NLS), respectively. If the number of wet days is less than 20 in the observed record, or if the mean observed precipitation value is less than 0.01 mmd^{-1} , then a simple additive correction factor equal to the difference in the means between simulated and observed series is applied. In turn, if the number of wet days is greater than 20, and the mean observed precipitation value is greater than 0.01 mmd^{-1} , the linear transfer function (Eq. 1) is fitted. The

exponential-type transfer function (Eq. 2) is selected to perform the bias correction under two conditions. First, if for the linear fit $a > 0$, which is interpreted as the corrected precipitation being always greater or equal than zero ($x_{\text{cor}} \geq 0$), i.e. ignoring dry days entirely; and, second, when the multiplicative factor (slope of Eq. 1) is too extreme, with arbitrary values defined in the range $b < 0.2$ and $b > 5$.

For temperature, series of observed and simulated daily (mean, maximum and minimum) temperatures within a “construction period” for every month are employed to fit the monthly linear transfer functions. Piani et al. (2010b) suggest that independently correcting the mean, maximum and minimum temperature results in large relative errors in the daily temperature range ($T_{\text{max}} - T_{\text{min}}$) and in the skewness ($T_{\text{mean}} - T_{\text{min}} / T_{\text{max}} - T_{\text{min}}$) of the corrected series. For that reason, the fitting of the monthly linear transfer functions is performed on the mean (T_{mean}), temperature range (T_{rg}) and temperature skewness (T_{sk}). Subsequently, daily transfer functions are obtained by a weighted linear interpolation of the parameters of the contributing monthly linear transfer functions for each variable T_{mean} , T_{rg} and T_{sk} . Using the daily transfer functions the corrected mean temperature (T_{mean}^c) is directly obtained whereas minimum and maximum temperatures are obtained from daily corrected values of T_{rg}^c and T_{sk}^c as $T_{\text{min}}^c = T_{\text{mean}}^c - T_{\text{rg}}^c \times T_{\text{sk}}^c$ and $T_{\text{max}}^c = T_{\text{mean}}^c + T_{\text{rg}}^c \times (1 - T_{\text{sk}}^c)$, respectively.

Given the availability of observed gridded data for daily precipitation, average, maximum, and minimum temperature in the E-OBS data set, the bias correction was performed for these four forcing variables. As “construction period” to build the transfer functions we defined the control period 1961–1990, i.e. 30 yr of daily data. We employed two series for each month (observed values from the E-OBS data set and simulated values from the HIRHAM5-ECHAM5 climate model) of ca. 900 values each to build the monthly transfer functions. The transfer functions obtained in this period were then applied to correct control (1961–1990) and future (2071–2100) climate simulations from the HIRHAM5-ECHAM5 model. In a parallel work, Dosio and Paruolo (2011) validated the transfer functions obtained in the control period for an independent (validation) period between 1991–2000 using an ensemble of 11 RCMs over Europe. We note here the important assumption of *stationarity*, which means that the corresponding form of the transfer function and its associated parameters are invariant over time. As a result, the transfer function estimated for present climate conditions is assumed to remain valid to correct biases in future climate simulations. As highlighted by Christensen et al. (2008), however, the stationarity assumption could be violated as biases can grow under climate change conditions and they depend on the values of the variables to be corrected.

2.3 Hydrological simulation

LISFLOOD is a GIS-based spatially-distributed hydrological rainfall-runoff model, which includes a one-dimensional hydrodynamic channel routing model (van der Knijff et al., 2010). Driven by meteorological forcing data (precipitation, temperature, potential evapotranspiration, and evaporation rates for open water and bare soil surfaces), LISFLOOD calculates a complete water balance at every (daily) time step and every grid cell defined in the modelled domain. Processes simulated for each grid cell include snowmelt, soil freezing, surface runoff, infiltration into the soil, preferential flow, redistribution of soil moisture within the soil profile, drainage of water to the groundwater system, groundwater storage, and groundwater base flow. Runoff produced for every grid cell is routed through the river network using a kinematic wave approach. Although this model has been developed aiming at operational flood forecasting at pan-European scale, recent applications demonstrate that it is well suited for assessing the effects of land-use change and climate change on hydrology (see, e.g. de Roo et al., 2001; Feyen et al., 2007; Dankers and Feyen, 2008, 2009).

The current pan-European setup of LISFLOOD uses a 5 km grid and spatially variable input parameters and variables obtained from European databases when available. Soil properties were obtained from the European Soil Geographical Database (King et al., 1994) whereas porosity, saturated hydraulic conductivity and moisture retention properties for different texture classes were obtained from the HYPRES database (Wösten et al., 1999). Vegetative properties and land use cover were obtained from the CORINE2000 data set (EEA, 2002) while elevation data and river properties were obtained from the Catchment Information System (Hiederer and de Roo, 2003). Parameters controlling snowmelt rates, overland and river flows, infiltration, and residence times in the soil and subsurface reservoirs have been calibrated against historical records of river discharge in 258 European catchments and sub-catchments. The calibration period varied for different catchments but all spanned at least 4 yr between the period 1995–2002. It may be argued that the selection of four years is too short for long-term applications, however, the selection of this period responded to a trade-off between computational time and the use of reliable and recent available information on discharges. Even if the calibration period is restricted to 4 yr, it must be stressed that LISFLOOD was calibrated having a particular interest in correctly reproducing the timing and magnitude of flooding events. The meteorological variables used to force LISFLOOD during the calibration were obtained from the Meteorological Archiving and Retrieving System (MARS) database (Rijks et al., 1998) and interpolated using an inverse weighted distance method over the 5 km grid. For catchments where discharge measurements were not available simple regionalization techniques (regional averages) were applied to obtain the parameters. The algorithm implemented

for calibration corresponded to the Shuffled Complex Evolution (SCE) (Duan et al., 1992). A more detailed description of the hydrological processes and parameters included in LISFLOOD are given by van der Knijff et al. (2010) whereas Feyen et al. (2007, 2008) discuss the calibration of LISFLOOD for different European catchments.

To drive the LISFLOOD model, the HIRHAM5-ECHAM5 daily simulations of temperature, precipitation, solar and thermal radiation, albedo, dewpoint temperature, humidity and wind speed were re-gridded to the 5 km grid used by LISFLOOD employing a nearest neighbour approach on the basis of the centre points of the 25 km grid cells of the HIRHAM5-ECHAM5 model. This resulted in forcing data on a grid fully consistent with the one employed to run LISFLOOD at pan-European scale. Two sets of forcing data were generated, one based on uncorrected regional climate simulations, and another based on the bias corrected fields of precipitation and temperature (avg, min and max). Daily information on solar and thermal radiation, albedo, dewpoint temperature, average, maximum and minimum temperatures, humidity and wind speed were used to calculate reference evapotranspiration employed by LISFLOOD using the Penman-Monteith model (Allen et al., 1998). Subsequently, fields of precipitation, average temperatures and Penman-Monteith-based evapotranspiration fields were used to force LISFLOOD in the control and future period. As a result, time series of daily discharge for each river pixel in the modelled domain (depicted in Fig. 1) were obtained for both bias corrected and uncorrected forcing data.

We assessed the benefits of correcting the bias in the climate simulations for hydrological impact assessments. For this purpose, the simulated discharges of LISFLOOD in the control period 1961–1990 were compared to observed discharges at 554 stations throughout Europe (see Fig. 1). For these stations at least 30 yr of daily discharge data were available with few exceptions of stations with 20 or 25 yr of daily discharge data in the control period. The distribution of these stations is somewhat uneven with a high concentration of stations in central Europe, whereas eastern-Europe shows the lowest station density. Despite this, these stations can be considered a good representation of different climatic conditions and hydrologic regimes with upstream areas contributing to the discharge between ca. 1000 km² and ca. 810 000 km².

2.4 Extreme value analysis and uncertainty assessment

To estimate the probability of extreme discharge levels, a Gumbel distribution was fitted to the annual maximum discharges using the maximum likelihood estimation (MLE) method (see, e.g. Beirlant et al., 2004). The Gumbel distribution is a special case of the Generalized Extreme Value (GEV) distribution with a shape parameter (ξ) explicitly set to 0 (Coles, 2001). Whereas the safest option would be to accept some degree of uncertainty about the value of ξ (Coles, 2001), by fitting a Gumbel distribution the uncertainty in the

shape parameter is explicitly neglected. Dankers and Feyen (2008), however, showed on the basis of a likelihood-ratio test that the use of the three-parameter GEV is not justified in the majority of the river cells (ca. 85 %) defined in the same European domain analysed in this work. In addition, they found no evidence that either the GEV or the Gumbel distribution produced consistently higher or lower estimates of return levels.

To obtain the 95 % confidence interval for the return levels we employed the profile-likelihood method (see, e.g. Coles, 2001; Beirlant et al., 2004). By capturing non-symmetric behaviour of confidence intervals, especially for return levels associated to long return periods, the profile-likelihood method is far more robust in assessing uncertainty compared to traditional approaches as the “Delta method” described in Coles (2001).

The profile-likelihood method works through reparametrization of the Gumbel model so that the return level z_p is one of the model parameters,

$$\mu = z_p + \sigma \log(-\log(1 - p)) \quad (3)$$

where μ and σ are the location and scale parameters of the Gumbel distribution, respectively, and $1/p$ is the return period. To obtain the profile-likelihood for a given return level, z_p is fixed to a value and the corresponding log-likelihood is maximized with respect to σ . This is repeated for a range of values of z_p . The profile-likelihood, $l_p(z_p)$, is built from the corresponding maximized values of the log-likelihood for different z_p values. Using the deviance function (see, e.g. Coles, 2001, p. 34), a $(1 - \alpha)$ confidence interval for return level z_p can be obtained as follows,

$$C_\alpha = \left\{ z_p : D(z_p) \leq \chi_1^2(1 - \alpha) \right\} \quad (4)$$

where the deviance function is defined as $D(z_p) = 2[l(\hat{z}_p) - l_p(z_p)]$, with $l(\hat{z}_p)$ being equal to the log-likelihood evaluated at the ML estimator of z_p .

3 Results

3.1 Comparison of HIRHAM5-ECHAM5 simulations with meteorological observations

3.1.1 Precipitation

An assessment of precipitation simulated by the HIRHAM5-ECHAM5 climate model is shown in Fig. 2 (plates a, b, c, and d). These plates show the ratio between the uncorrected HIRHAM5-ECHAM5 simulations and the E-OBS observations for the average seasonal (DJF, MAM, JJA and SON) precipitation in the control period 1961–1990. Seasonal averages show a clear overestimation of precipitation in large parts of the Iberian Peninsula, Central Europe, Great Britain, northern Europe and Scandinavia, with values as high as 7–20 fold the observed precipitation. For the winter season

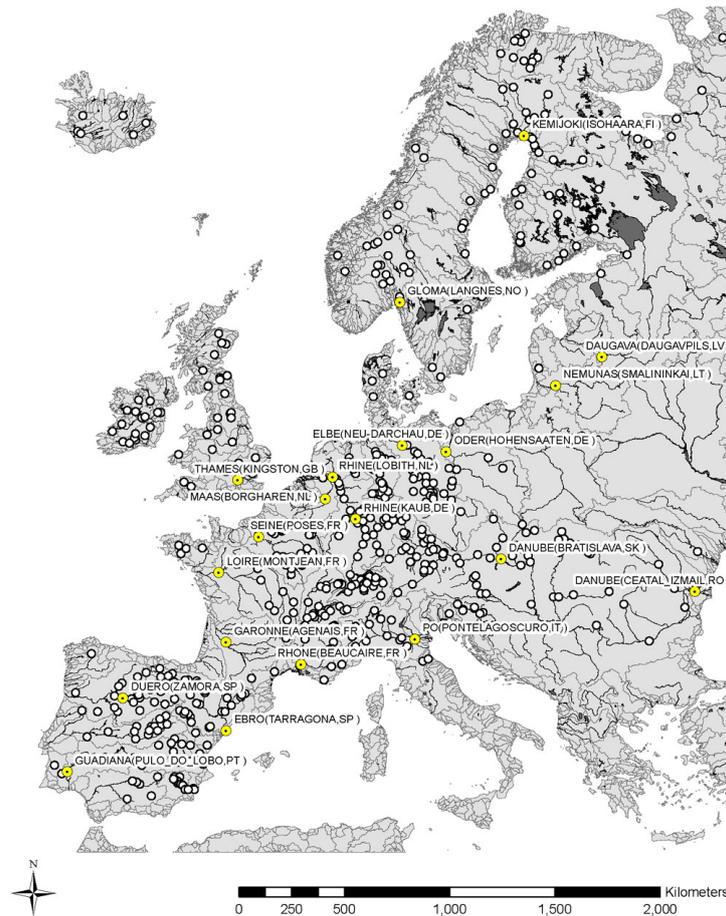


Fig. 1. Location of the 554 gauging stations used to assess simulated river discharges. These stations have at least 25 yr of daily data for the control period 1961–1990.

(Fig. 2a) the HIRHAM5-ECHAM5 model tends to overestimate precipitation all across Europe, with some weak underestimation in the Scandinavian mountains, west coast of Italy, north-west coast of Great Britain and Balkan. A similar pattern is observed for the transition seasons (MAM and SON). In summer season (Fig. 2c), on the other hand, precipitation is overestimated in northern European regions, whereas too dry conditions are simulated in south-eastern parts. The summary statistics presented in Table 1 show that, averaged over the European domain, the simulated average annual precipitation almost doubles the observed precipitation. These results are in agreement with the findings of van Meijgaard et al. (2008) and Kjellström et al. (2010), who analysed a series of RCMs driven by lateral boundary conditions obtained from ERA40 reanalysis product.

Plates e, f, g, and h of Fig. 2 show the ratio between bias corrected HIRHAM5-ECHAM5 precipitation and observations. The range of over- and underestimation is reduced, with differences in seasonal average precipitation of approximately $\pm 5\%$ of the observed precipitation. This is also reflected in the average over the European domain of the annual

and seasonal corrected precipitation, which is nearly identical to that of the observation data set (see Table 1). For summer (JJA) an overestimation of the observed precipitation is still present in the most southern areas of the Iberian Peninsula, whereas a weak underestimation of precipitation can be observed in the Balkan.

In spite of the excellent performance of the bias correction, a few grid cells associated to mountain areas (e.g. Alps, Apennines, Balkan, and Carpathians) still show a pronounced overestimation up to 3–16 fold the observed precipitation after bias correction. However, the overestimation in these areas of the uncorrected precipitation amounts to 25–30 fold the observed annual precipitation. The seasonal analysis indicates that the residual overestimation after the bias correction mainly occurs in winter season, and to a lesser extent in the transition seasons. Analysis of the type of fitting function employed in these grid cells suggests that switching between the linear and exponential type of fitting during the winter months (DJF) could significantly alter the interpolation of daily values as they are influenced by the anterior and posterior (monthly) fitting function. This could potentially

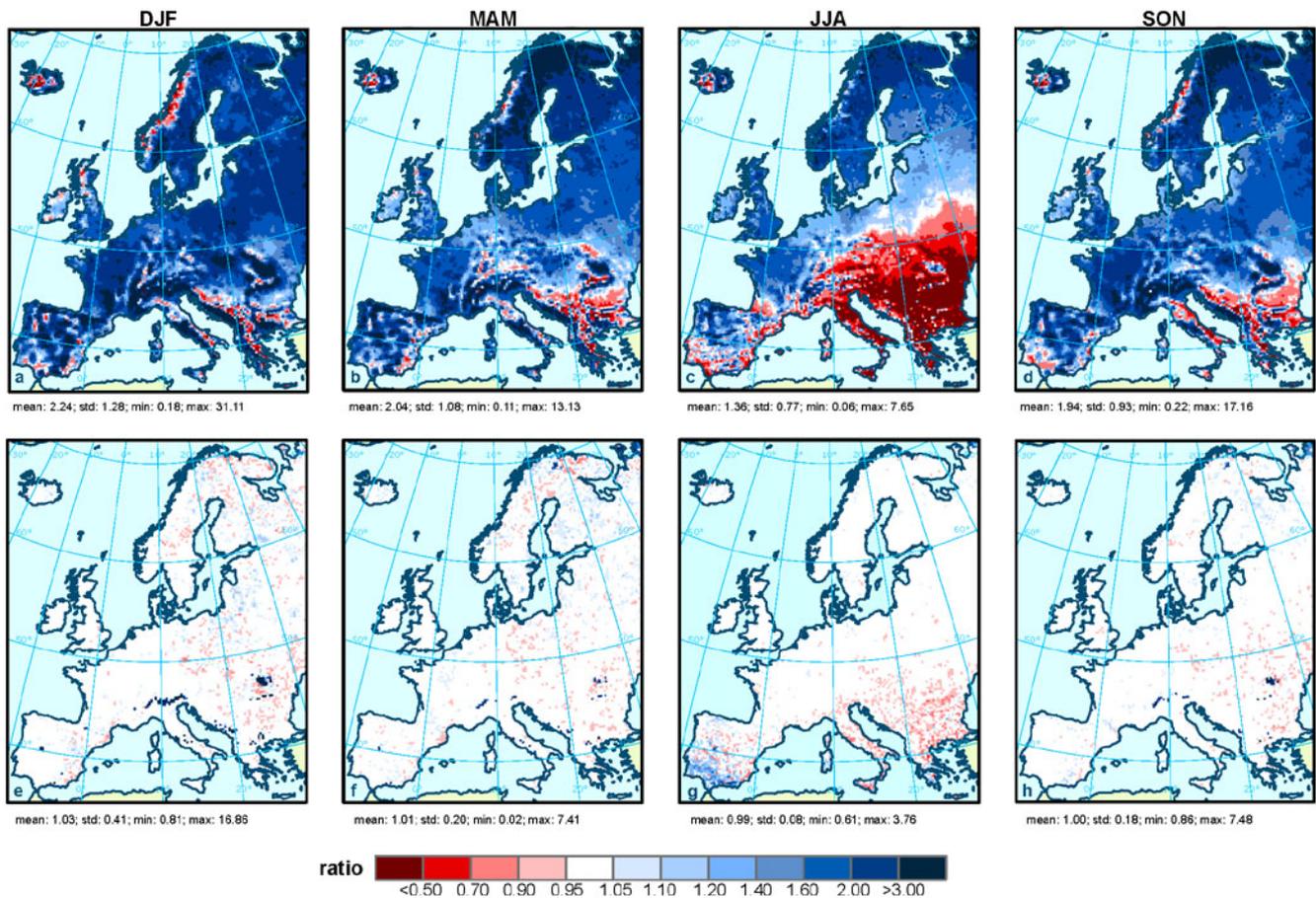


Fig. 2. Ratio between HIRHAM5-ECHAM5 simulations and observations of seasonal average precipitation for the control period 1961–1990. First row, uncorrected seasonal precipitation; second row, bias corrected seasonal precipitation.

Table 1. Summary statistics of precipitation (mmd^{-1}) for the control period 1961–1990.

	Statistics		
	Observed	Bias corrected	Uncorrected
Average annual precipitation	1.9	1.9	3.4
Average annual max precipitation	28.3	30.6	47.8
3-d annual max precipitation	47.6	48.8	78.5
5-d annual max precipitation	59.1	60.0	97.9
7-d annual max precipitation	68.5	69.5	114.4
99th percentile precipitation	23.9	25.5	37.3
Average seasonal precipitation (DJF)	1.8	1.8	3.8
Average seasonal precipitation (MAM)	1.6	1.6	3.2
Average seasonal precipitation (JJA)	2.0	2.0	2.8
Average seasonal precipitation (SON)	2.1	2.1	3.9

modify the monthly statistics. A similar problem has been recognized by Piani et al. (2010b) in their global analysis. Aside from model deficiencies, the pronounced overestimation at high altitudes can in part be explained by the fact that observed precipitation typically underestimates true precipitation, especially in winter, due to poor station network den-

sity (see, e.g. Hofstra et al., 2010; Lenderink, 2010) that does not allow to fully capture orographic effects on precipitation. Hence, the bias-corrected precipitation in these areas may more closely correspond to true precipitation amounts than the comparison with observed precipitation suggests here.

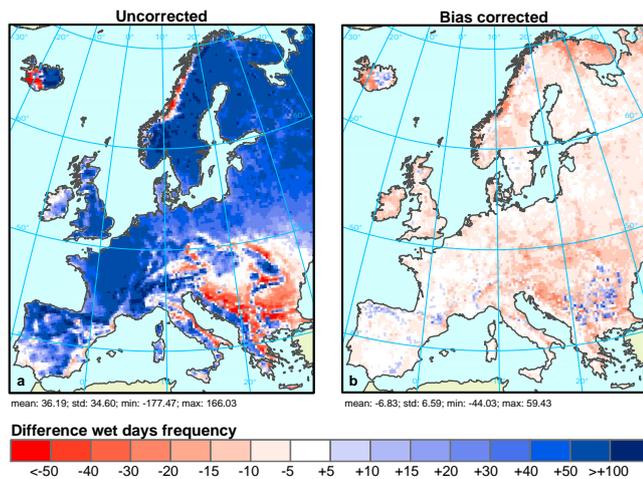


Fig. 3. Difference of average annual wet day frequency (precipitation $\geq 1 \text{ mm d}^{-1}$) between (a) uncorrected precipitation and (b) bias corrected precipitation with respect to the E-OBS data set for the control period 1961–1990.

The wet day frequency relates to the distribution of precipitation in time. The comparison presented in Fig. 3 shows a pronounced overestimation of the average annual number of wet days (daily precipitation $\geq 1 \text{ mm d}^{-1}$) in northern and western parts of Europe, as well as in mountain regions, whereas in most of southeastern Europe and parts of the Scandinavian mountains too few wet days are simulated. This is conform to what most studies report, namely, that climate models tend to simulate too many days with weak precipitation, especially in humid climate zones (see, e.g. Leander and Buishand, 2007; Piani et al., 2010b; Iizumi et al., 2011). The bias correction drastically reduces the error in the number of wet days simulated with respect to the E-OBS data set. This improvement in the distribution of wet days frequency after bias correction is explained by the advantage of fitting both “*a*” and “*b*” parameters in Eq. (1) and (2), which are related to the dry day correction factor, and by using the portion of both time series (observed and simulated) that correspond to wet days only to estimate the transfer functions.

At the same time, Fig. 3 shows a slight underestimation of the wet day frequency after bias correction (plate b). From the analysis of an ensemble of RCMs for Europe, Dosio and Paruolo (2011) suggested that low-end percentiles of bias corrected precipitation are subject to large uncertainties due mainly to the choice of the transfer function (Eq. 1 vs. Eq. 2) used to perform the bias correction. In addition, they found a systematic underestimation of the small values of bias corrected precipitation when compared to the pdf obtained from the E-OBS data set. As a consequence, the underestimation of the wet days frequency after bias correction is likely explained by the systematic underestimation of the low-end percentiles for the bias corrected precipitation observed by Dosio and Paruolo (2011), which increases the probability

for a given day to be considered as a “dry day” for a constant x_0 . In addition, and even if the fitting of “*a*” and “*b*” helped to drastically reduce the excessive number of wet days observed before bias correction, there still exists the possibility that the number of dry days is slightly overestimated given that x_0 is obtained purely from a fitting process, thus, potentially differing from the actual number of dry days in the observed precipitation.

Rain-driven flooding can be caused by short-term heavy rain events (rapid-onset flooding such as pluvial and flash floods) or prolonged periods of intense rain (slow-onset flooding such as river or fluvial floods). To evaluate the possible effect of the bias correction on both types of phenomena we evaluate different precipitation indicators, namely, the 99th percentile of daily precipitation and the maximum precipitation amount in 3, 5 and 7 consecutive days. The comparison for the seasonal 99th percentile of daily precipitation (Fig. 4 plates a, b, c, d and Table 1) shows that daily extreme precipitation is overestimated in most parts of Europe, except for southeastern Europe, Italy and southern parts of the Iberian Peninsula in summer, and some areas in the Scandinavian mountains mainly in winter. The seasonal 3, 5 and 7-day maximum precipitation (see Fig. 5 plates a, b, c and d for the 5-day maximum precipitation) show a similar pattern (which corresponds very well to the average precipitation shown in Fig. 2, plates a, b, c and d), although compared to the 99th percentile of daily precipitation the underestimation in the Balkan is more pronounced here in the transition seasons. The general overestimation of the extreme indicators is in line with observations that many RCMs contained in the ENSEMBLES project simulate too much (daily) precipitation, a tendency that is more pronounced for the upper-end percentiles (i.e. extreme precipitation events) (Kjellström et al., 2010; Lenderink, 2010). The latter, does not preclude that some RCMs forced by different conditions may underestimate the high-end percentiles while overestimating the mean precipitation.

The bias correction considerably reduces the error in the seasonal 99th percentile to within the range $\pm 25\%$ of the observed value. Similar as for the average precipitation, in some grid cells associated to mountain areas (e.g. Alps, Apennines, Balkan, and Carpathians) still a clear tendency to overestimation can be observed. Also for the 3, 5 and 7-day maximum precipitation, aside from the residual strong overestimation in the sparse mountainous grid cells, the error is considerably reduced. However, notwithstanding that the European-average statistics closely reproduce the observed statistics (see Table 1), the bias correction results in a slight underestimation of the maximum precipitation amount in 3, 5 and 7 consecutive days across most of Europe. Similarly to the case of wet days frequency, this is most likely related to the underestimation of the low-end percentiles of the bias corrected precipitation (see, e.g. Dosio and Paruolo, 2011), and a potential overestimation of the number of dry days obtained from the fitting of the transfer functions. The latter, in

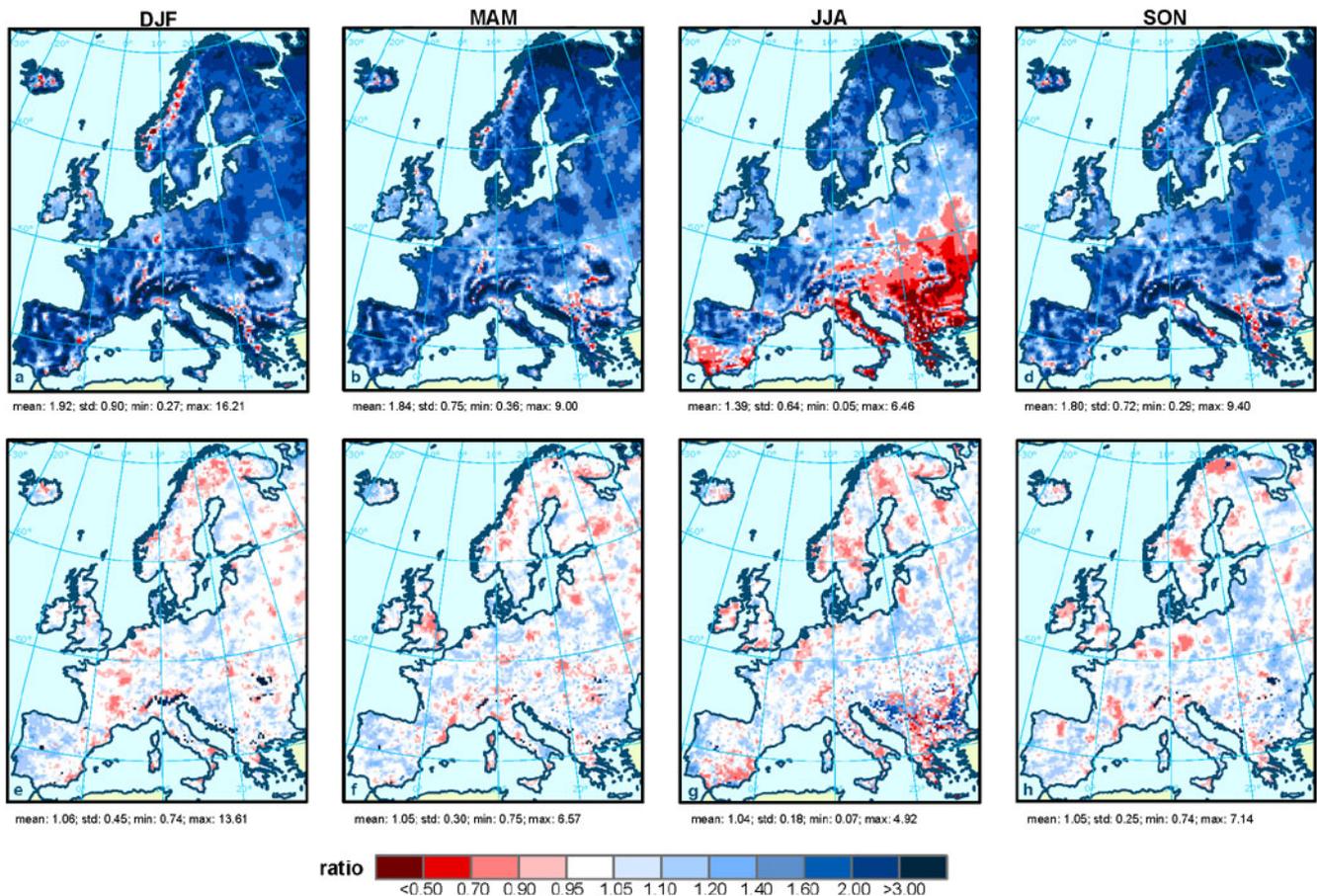


Fig. 4. Ratio between HIRHAM5-ECHAM5 simulations and observations of seasonal 99th percentile for the control period 1961–1990. First row, uncorrected seasonal precipitation; second row, bias corrected seasonal precipitation.

combination with the fact that observed extremes likely undervalue true extreme precipitation may result in an underestimation by the hydrological model of observed extreme discharges.

3.1.2 Temperature

Figure 6 (plates a, b, c and d) shows on a seasonal basis the difference in 2 m temperature (T_{2m}) between the HIRHAM5-ECHAM5 simulations and the E-OBS data set. A common feature, evident in all seasons, is the strong warm bias (up to 4°C and more) in south-eastern Europe, which extends further north and west depending on the season. In winter, T_{2m} is overestimated nearly all over Europe, except for the British Isles, along the western Scandinavian coast, in the Alps and Pyrenees, with a somewhat mixed bias-pattern in the rest of the Iberian Peninsula. In the other seasons, simulated T_{2m} is too low for most of the western half of Europe, including the Iberian Peninsula, France, Great Britain, the Alps and western parts of Scandinavia. In summer, and to a lesser

extent in spring, a cold bias up to 3°C can be observed in northern Europe. After performing the bias correction of the HIRHAM5-ECHAM5 temperature simulations, the anomalies in T_{2m} are drastically reduced (see Fig. 6 plates e, f, g and h, and Table 2). For all seasons, discrepancies with the E-OBS data set have become marginally small, with maximum differences not exceeding 0.5°C.

A comparison on annual basis of the daily maximum (T_{max}) and minimum (T_{min}) temperature is presented in Fig. 7. As maximum temperatures occur in summer season, for T_{max} a similar bias pattern can be observed as for T_{2m} in summer (see Fig. 6c). Figure 7a shows for T_{max} a warm bias in southeastern Europe with values higher than 1°C that transitions into an underestimation in the western and northern half of the continent, where the cold anomaly ranges between 1°C and 5°C compared to the E-OBS data set. Minimum temperatures, on the other hand, are overestimated all over Europe, especially in the south-east, except for the Alps, western Pyrenees and western parts of Norway (see Fig. 7b). In southeastern Europe, the overestimation of T_{min} is much stronger than that of T_{max} . Hence, the

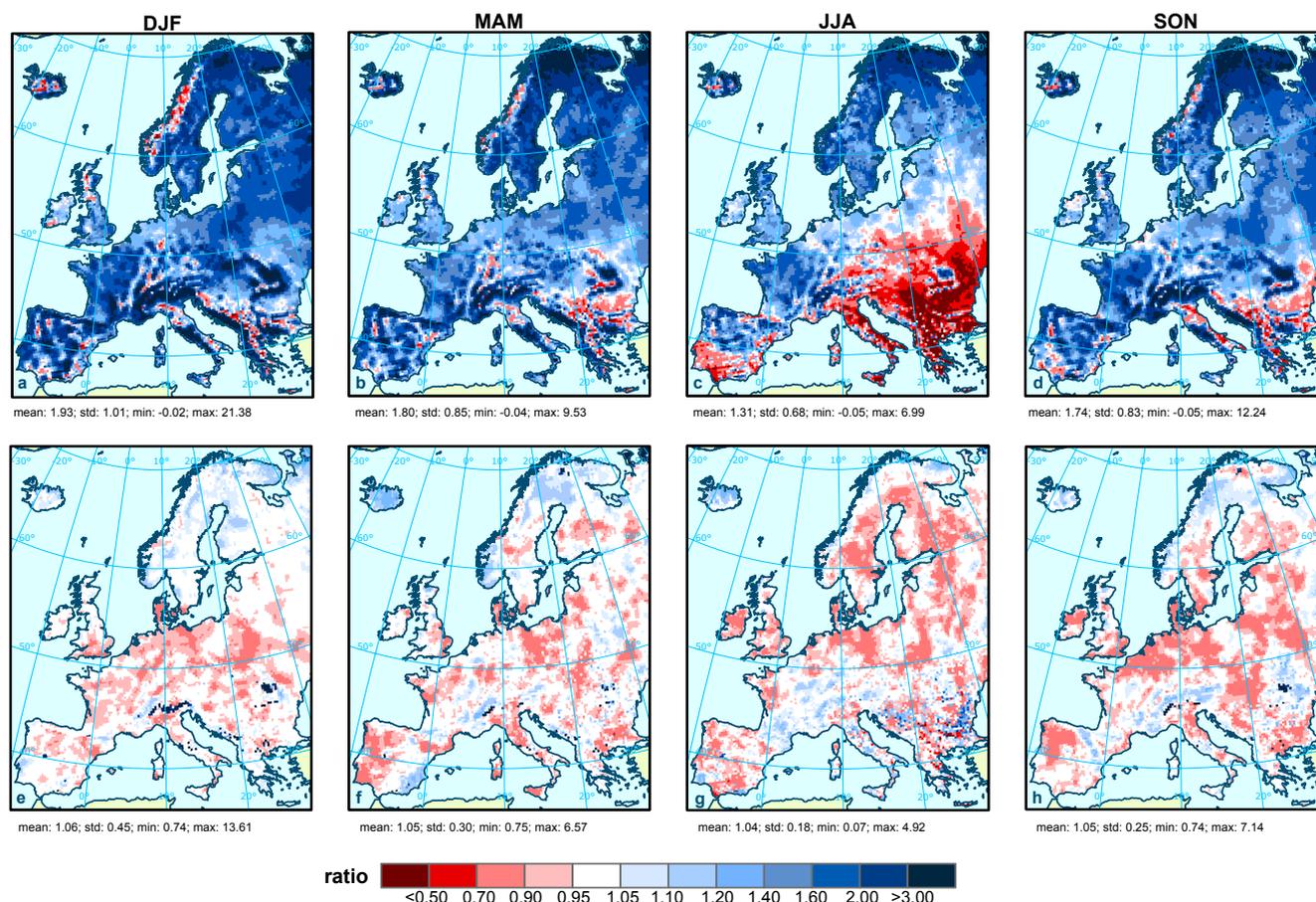


Fig. 5. Ratio between HIRHAM5-ECHAM5 simulations and observations of seasonal 5-days maximum precipitation for the control period 1961–1990. First row, uncorrected seasonal precipitation; second row, bias corrected seasonal precipitation.

Table 2. Summary statistics of (2 m) temperature ($^{\circ}\text{C}$) for the control period 1961–1990.

	Statistics		
	Observed	Bias corrected	Uncorrected
Average daily temperature	6.9	6.9	7.9
Average daily maximum temperature	11.2	11.3	10.6
Average daily minimum temperature	2.8	2.8	5.1
Winter (DJF) average daily maximum temperature	0.9	0.9	1.3
Winter (DJF) average daily minimum temperature	-5.8	-5.8	-3.1
Spring (MAM) average daily maximum temperature	10.8	10.8	9.5
Spring (MAM) average daily minimum temperature	1.6	1.6	3.6
Summer (JJA) average daily maximum temperature	21.5	21.5	20.4
Summer (JJA) average daily minimum temperature	11.2	11.2	13.5
Autumn (SON) average daily maximum temperature	11.6	11.7	11.0
Autumn (SON) average daily minimum temperature	3.9	4.0	6.4

simulated diurnal temperature range is too narrow in most regions of Europe for the uncorrected temperatures simulated by the HIRHAM5-ECHAM5 model. After performing the bias correction, T_{max} and T_{min} show maximum differences

not exceeding $\pm 0.1^{\circ}\text{C}$ across Europe. A warm anomaly of maximum 0.5°C is observed in the Scandinavian mountains, Balkan, and JJA.

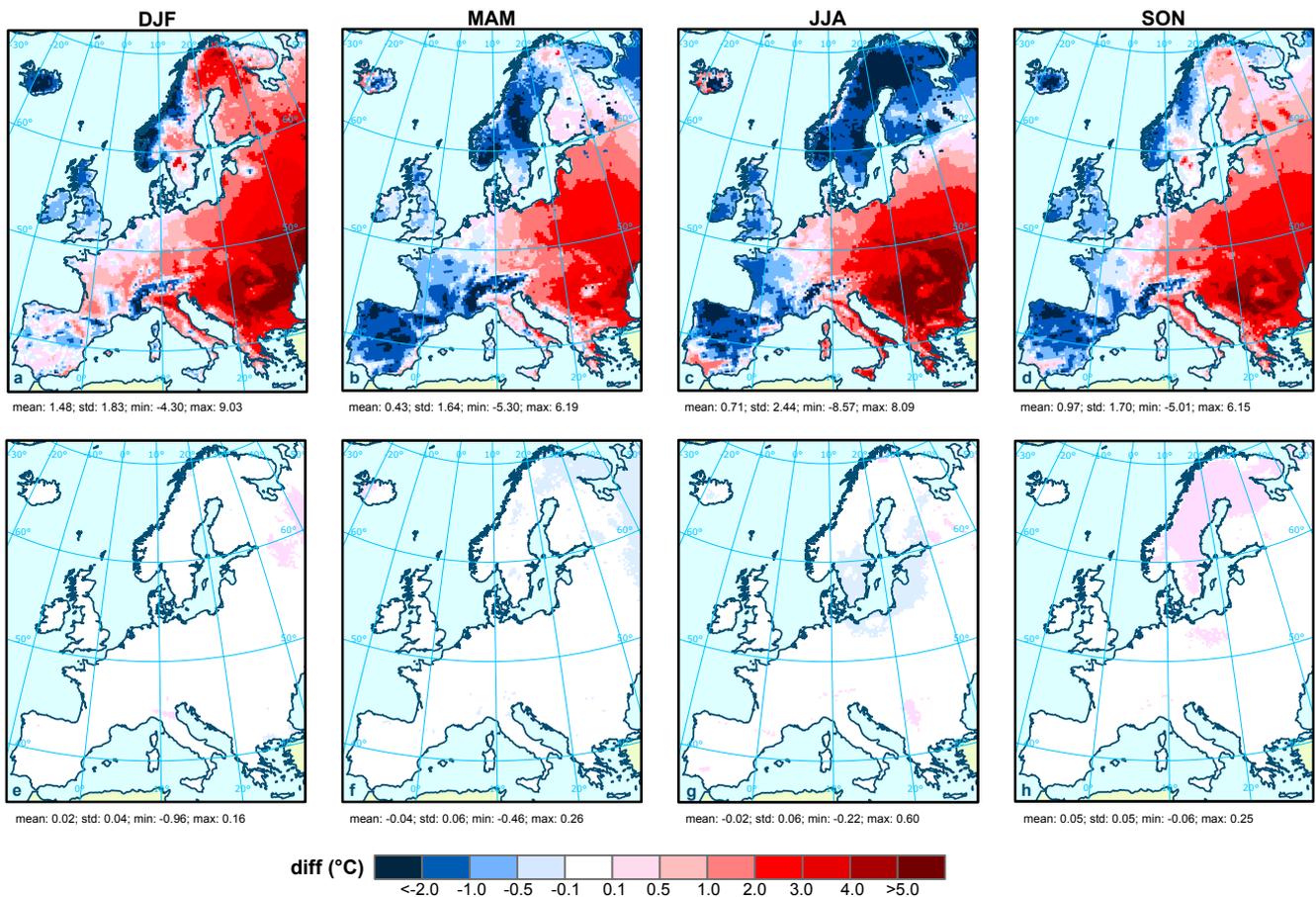


Fig. 6. Difference between HIRHAM5-ECHAM5 simulations and observations of seasonal average daily temperature for the period 1961–1990. First row, uncorrected seasonal temperature; second row: bias corrected seasonal temperature.

These observations on simulated temperature are in agreement with other studies that report a warm bias in winter and summer season, especially in southeastern parts of Europe (Jacob et al., 2007; Christensen et al., 2008), and a consistent cold bias in the Alpine region (Suklitsch et al., 2010). A strong underestimation of T_{\max} over northern Europe has been observed by Nikulin et al. (2011) when examining an ensemble of RCA3 simulations driven by six different GCMs (including ECHAM5), as well as by Kjellström et al. (2007) for 10 different RCMs driven by HadAM3H. Possible causes for temperature bias include missing processes such as a realistic description of dust aerosols or regional feedback processes. Especially in the summer season, such regional feedback processes include soil moisture exhaustion or decreasing cloud cover, which can impact the temperature developments. Also, the capability of RCMs to simulate the regional climate depends on the realism of the large-scale circulation that is provided as lateral boundary conditions, in particular in regions where the large-scale dynamic forcing plays an important role compared to local forcing. This is indeed the case over large parts of Europe. van

Ulden and van Oldenborgh (2006) for example have shown that a warm bias in Central Europe of the order of degrees was induced by a bias in the large-scale circulation boundary conditions. Also in central Europe, Plavcová and Kyselý (2011) report a substantial underestimation of the diurnal temperature range throughout the year in ENSEMBLES RCM experiments. Besides the deficiencies in the simulation of atmospheric circulation, particularly too strong advection and overestimation of westerly flow at the expense of easterly flow in most RCMs, they suggest that biases in simulating anticyclonic, cyclonic and straight flow also contribute to the underestimated diurnal temperature range. Also, although to a lesser extent than for precipitation, inaccuracies in the observational datasets and interpolation of station data may explain part of the discrepancies between simulated and observed temperature (see, e.g. Lorenz and Jacob, 2010).

3.2 Impact of bias correction on the simulation of hydrological components

Flood generation is a highly non-linear process that depends on factors such as the intensity, volume and timing

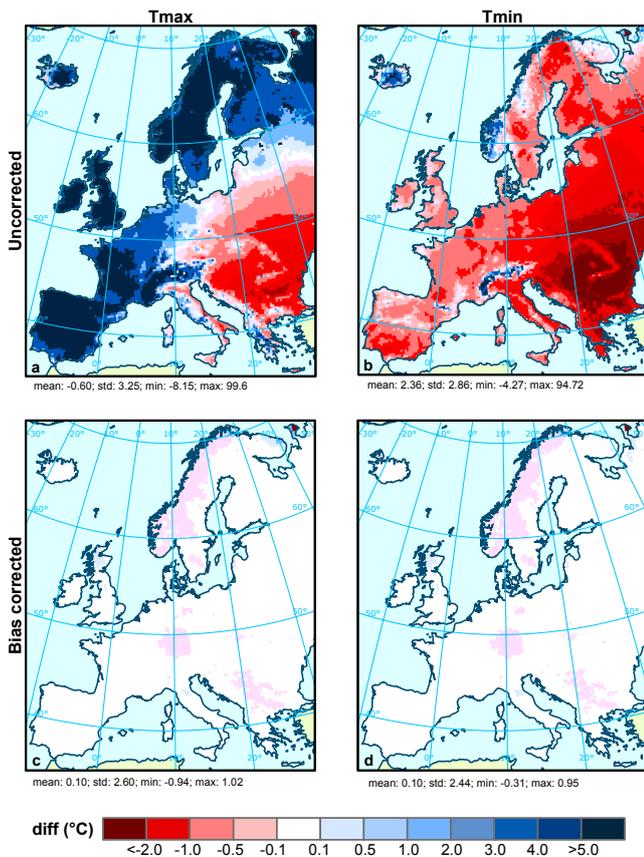


Fig. 7. Differences for daily maximum (T_{\max}) and minimum (T_{\min}) temperature simulated by HIRHAM5-ECHAM5 and the E-OBS data set for the control period 1961–1990. First row, uncorrected annual temperature; second row: bias corrected annual temperature.

of precipitation as well as on antecedent conditions of the river basin (e.g. soil wetness, snow or ice cover). Because of the small to meso-scale character of these factors, the correct representation of temperature and precipitation, both spatially and temporally, are key to enhance the predictive capacity of the LISFLOOD model. Whereas the link between extreme (long or intense) precipitation and flooding is obvious, in this section we detail the effect of the bias correction on processes that indirectly affect flooding.

Evapotranspiration (ET) regulates the flow of moisture back into the atmosphere and as such affects the storage capacity of soils. An overestimation of ET results in lower stocks of moisture in the soil, hence a larger amount of water that can be stored during subsequent wet periods. The effect of the bias correction on evapotranspiration is shown in Fig. 8. The run with uncorrected climate data results in higher ET amounts in northern-central Europe due to higher temperatures and precipitation during most of the year. In southeastern Europe, notwithstanding the strong warm bias, lower ET amounts are simulated, as the underestimation of precipitation in summer limits the water available for ET

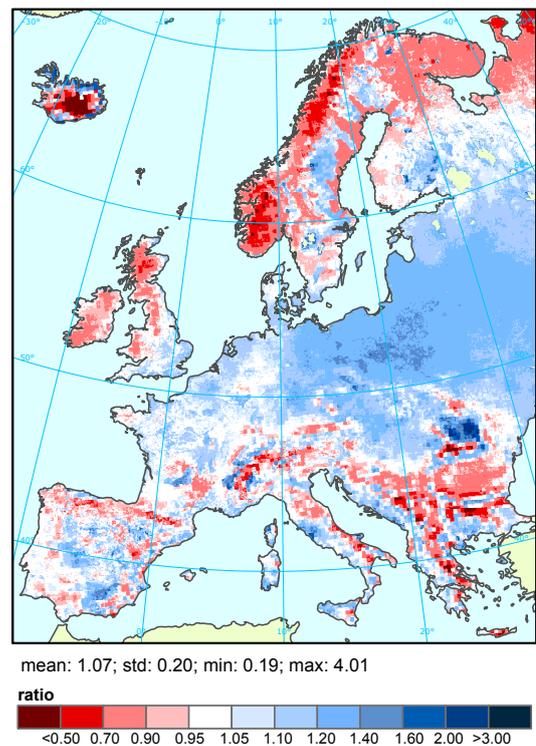


Fig. 8. Ratio for LISFLOOD evapotranspiration (mmd^{-1}) simulations between uncorrected and bias corrected climate data for the control period 1961–1990.

(see, e.g. Fig. 2c). In Scandinavia, on the other hand, lower ET amounts are simulated due to the cold bias in the uncorrected summer temperature (see, e.g. Fig. 6c).

It must be noted that variables such as dewpoint temperature, solar and thermal radiation that are employed together with the bias-corrected temperature fields to calculate the evapotranspiration terms driving LISFLOOD (see, e.g. van der Knijff et al., 2010) are not corrected for potential bias. This may violate the energy balance and potentially introduce bias in the subsequent hydrological simulations.

Many snow-dominated regions in Europe are subject to snowmelt-induced floods that result from rapid melting of the snowpack, sometimes amplified by rainfall. Ice-jam floods, related to freeze-up and break-up periods, also frequently cause winter and spring floods. The accumulation of snow during cold periods depends on the amount of precipitation, as well as on the temperature, which determines whether precipitation falls as rain or snow. A correct simulation of the (variability in) temperature also affects the possible onset of snowmelt or ice-jam floods. We show for the actual snowpack depths and the days with snow cover the difference between the simulation driven by uncorrected and bias corrected climate data (Fig. 9). From Fig. 9 we observe a thicker snowpack and more days with snow cover mainly in northern Europe and mountain areas when not correcting for bias in

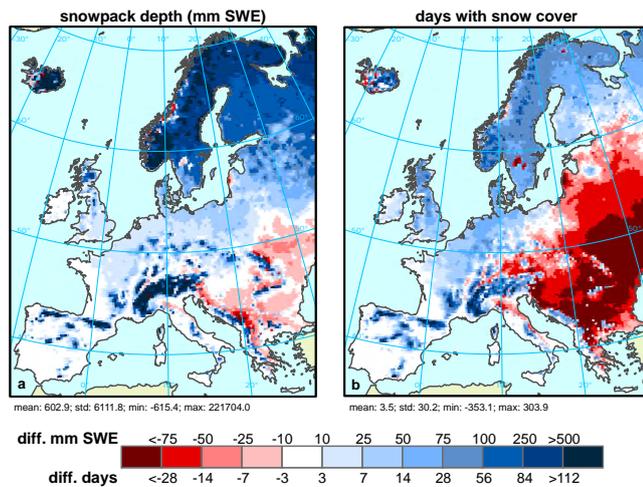


Fig. 9. Difference for LISFLOOD hydrological simulations driven by uncorrected and bias corrected climate data for the control period 1961–1990 for (a) snowpack depth (mm SWE) and (b) days with snow cover.

the climate simulations. The difference in snowpack in these areas ranges between 100 and 1000 mm snow water equivalent (SWE), and locally at high altitudes even up to 5000 mm and more. The more pronounced accumulation of snow and higher number of days with snow cover in northern Europe is due mainly to the strong overestimation of precipitation (see Fig. 2), which outweighs the effect of the warm bias in cold periods (which potentially reduces the ratio snow to rain). In central, central-eastern and south-eastern regions, the warm bias in winter reduces the number of days with snow cover (see Fig. 9b), but the wet bias in winter still results in a higher maximum snowpack. The considerable increase in snowpack in mountain ranges such as the Alps and Pyrenees follows from the wet and cold bias in cold periods.

3.3 Impact of bias correction on the simulation of hydrological extremes

Figure 10 shows observed versus simulated average discharge and average annual maximum discharge for each of the 554 validation stations (see Fig. 1). In general, the hydrological model driven by the uncorrected climate data largely overestimates average and extreme river flows (Fig. 10 plates a and c). This follows from the strong overestimation of the average and extreme precipitation in most regions of Europe. In regions dominated by (spring) snowmelt, the pronounced over-accumulation of snow in winter with the uncorrected climate data also likely contributes to the overestimation of extreme discharges. LISFLOOD simulations forced by the bias corrected climate data (Fig. 10 plates b and d) show a strong amelioration in reproducing the observed discharge statistics. Visual inspection and the values for the coefficient of determination (r^2) and model efficiency (EF) show that

the observed flow statistics are reasonably well reproduced after implementing the bias correction method, with a general tendency of better performance for average flows and with increasing catchment size.

Notwithstanding the overall good agreement between observed and simulated discharge statistics when employing bias-corrected forcing data, large discrepancies do occur at a small number of stations (see plates b and d of Fig. 10), where the relative errors can be 1 or 2 orders of magnitude (note the logarithmic scale). Deviations from the observation-based statistics can be attributed to errors in the hydrological model, its static input and in the calibration and regionalization of its parameters. Part of the disagreements can also be linked with man-made modifications of flow regimes in many catchments in Europe (see, e.g. Dynesius and Nilsson, 1994) that are not accounted for in the hydrological model. Also, albeit that the E-OBS data set is currently the best available for Europe, it is known to feature errors and uncertainties, which are translated to the corrected forcing fields during the bias correction step. More specifically, the E-OBS data set compares better to the mean of the variables than to the extremes, with larger differences for precipitation than for temperature (Haylock et al., 2008; Hofstra et al., 2010). This, in combination with the slight underestimation of the 3, 5 and 7-day maximum precipitation after bias correction (see Fig. 5) may also explain why for the annual maxima the greater part of the stations falls below the one-to-one line.

Figure 11 shows the results of fitting by MLE a Gumbel distribution to the annual maximum discharges for the period 1961–1990 in 20 selected stations. These stations are a representative subsample of the 554 validation stations (see Fig. 1) covering different hydro-morphologic and climatologic conditions and catchment size (ranging from 9948 to 80 700 km²). In general, there is a strong overestimation of the empirical return levels for the annual maxima (black crosses) when the hydrological model is driven by uncorrected forcing data. The latter is more pronounced for higher return levels. A clear improvement in reproducing the empirical return levels is observed when LISFLOOD is run using bias corrected forcing data (see blue and red dashed-lines in Fig. 11). Although for some stations still a marked disagreement exists between the return levels obtained from empirical plotting positions and the return levels obtained from bias corrected forcing data, for 50 % of stations the confidence intervals more closely envelop the empirical return levels. This is also reflected in Table 3, which summarizes the 95% confidence intervals obtained for a return period of 100-yr for the stations shown in Fig. 11. The reduction in the confidence intervals reaches up to 70 % compared to the estimations based on uncorrected forcing data. It is worth mentioning that we employed a series of 30 annual maxima (1961–1990) to perform the Gumbel fitting, thus, any extrapolation beyond 30 yr could potentially be dominated by large uncertainties. Also, the confidence intervals only reflect the uncertainty in the

Table 3. 95 % confidence intervals obtained using the profile-likelihood method for a return period of 100-yr for the gauging stations depicted in Fig. 1 for the control period 1961–1990. Observed column corresponds to Q_{100} obtained from a Gumbel distribution fitted from the observations. All values in $\text{m}^3 \text{s}^{-1}$ (values in parentheses show the percentage reduction).

Stations	Country	Observed (Gumbel fitting)	Forcing input data	
			Uncorrected	Bias corrected
Guadiana, Pulo do Lobo	PT	4521	[9598–14 239]	[4027–6008] (57.3 %)
Danube, Bratislava	SK	7435	[11 356–15 957]	[8951–12 973] (12.6 %)
Elbe, Neu-Darchau	DE	3038	[7742–11 343]	[3373–5109] (51.8 %)
Oder, Hohensaaten-Finow	DE	1933	[5058–7394]	[2730–4055] (43.3 %)
Maas, Borgharen	NL	2105	[5610–7938]	[2260–3352] (53.1 %)
Rhine, Lobith	NL	9336	[23 190–31 827]	[10 105–13 985] (55.1 %)
Thames, Kingston	GB	500	[866–1172]	[341–490] (51.3 %)
Danube, Ceatal Izmail	RO	13 825	[31 519–41 268]	[22 430–31 932] (2.5 %)
Nemunas, Smalininkai	LT	2963	[7290–10 901]	[5678–9260] (0.8 %)
Daugava, Daugavapils	LV	3887	[6033–8782]	[3156–4842] (38.7 %)
Seine, Poses	FR	2296	[8940–12 041]	[2598–3798] (61.3 %)
Loire, Montjean	FR	5715	[15 116–20 099]	[4664–6724] (58.7 %)
Garonne, Agenais	FR	5143	[11 612–15 898]	[3490–4804] (69.3 %)
Rhone, Beaucaire	FR	7142	[15 366–19 898]	[4753–6269] (66.6 %)
Ebro, Tarragona	SP	3245	[6839–9199]	[1687–2400] (69.8 %)
Po, Pontelagoscuro	IT	8280	[31 623–45 357]	[10 338–14 659] (68.5 %)
Gloma, Langnes	NO	2982	[9172–11 838]	[3090–4092] (62.4 %)
Kemijoki, Isohaara	FI	4222	[11 396–15 591]	[3193–4418] (70.8 %)
Duero, Zamora	SP	1340	[4323–5778]	[841–1275] (70.2 %)
Rhine, Kaub	DE	5975	[13 781–18 590]	[6782–9358] (46.4 %)

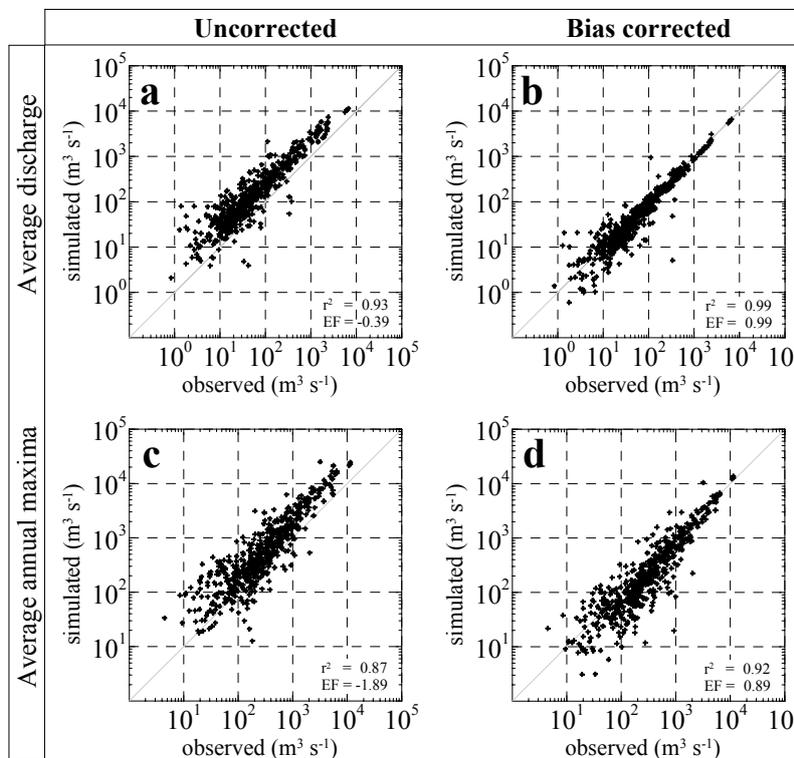


Fig. 10. Observed versus simulated average discharge (a, b) and average annual maximum discharge (c, d) for each of the 554 stations depicted in Fig. 1 for the control period 1961–1990. Left and right columns show hydrological simulations driven by uncorrected and bias corrected climate data, respectively.

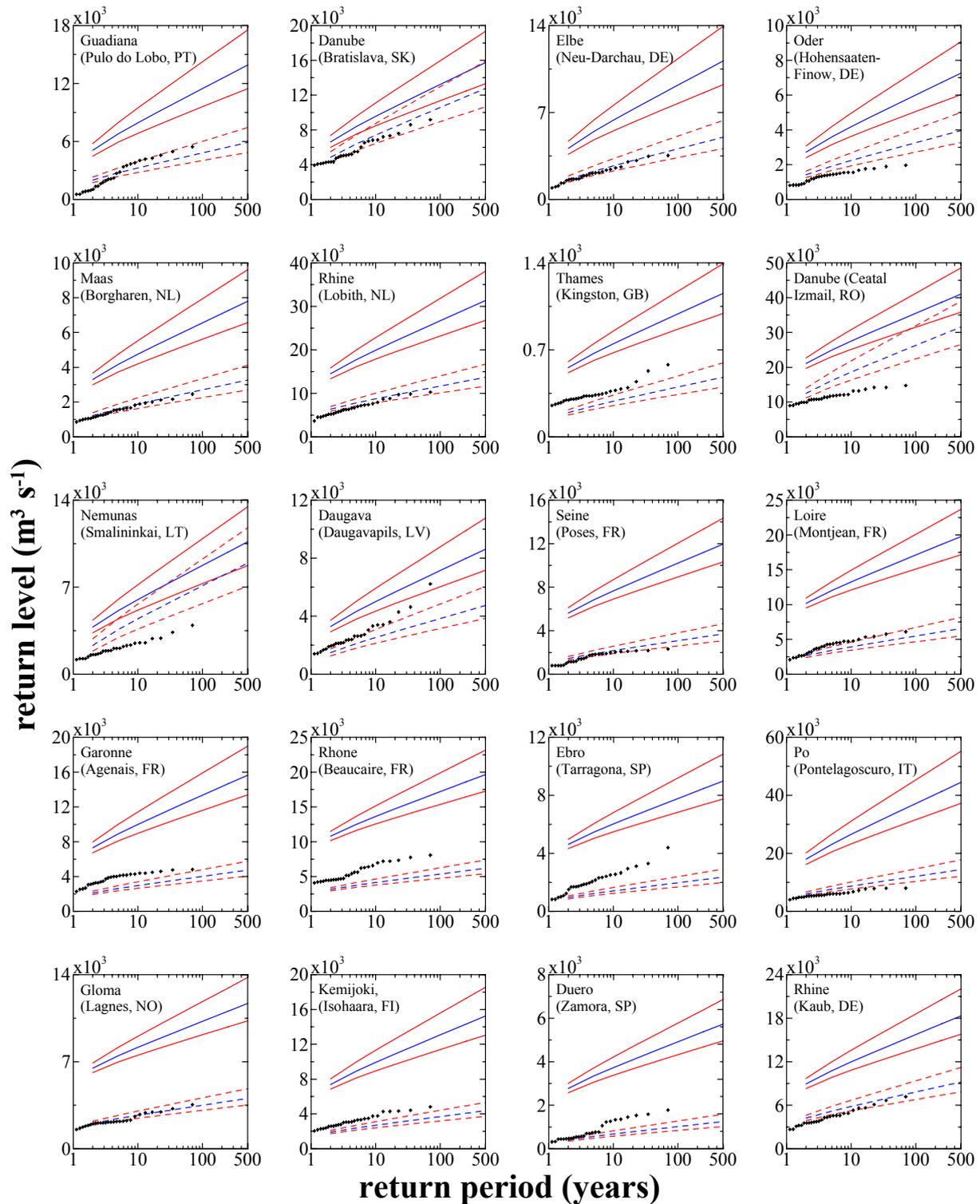


Fig. 11. Return level plots of simulated discharge levels in the 20 selected gauging stations showed in Fig. 1, based on a Gumbel distribution fit to the annual maxima for the control period 1961–1990. Full lines represent hydrological simulations driven by uncorrected forcing data whereas dashed lines represent the bias corrected counterpart. Also included in the plates are the 95 % confidence intervals (red dashed- and full-lines) derived using the profile-likelihood method. Black crosses represent return levels obtained from empirical plotting positions of observed annual maximum discharges at the selected stations.



Fig. 12. Ratio of discharges for a return period of 100 yr between hydrologic simulations driven by uncorrected and bias corrected forcing data for the control period 1961–1990.

extreme value fitting, and other sources of uncertainty from the hydrological modelling exercise are not accounted for.

In most regions of Europe the empirical return levels are overestimated when LISFLOOD is driven by uncorrected forcing data. This can be deduced from the analysis of the 20 stations above in combination with Fig. 12, which shows the ratio of 100-yr return level discharges between runs driven by the uncorrected and bias corrected forcing data for the control period. There is, at the same time, a well-defined area (north of the Carpathian mountains) where the 100-yr flood discharges based on bias corrected forcing data are slightly higher (ca. 25 %) than those obtained when using uncorrected forcing data. In the latter, this region is subject to a large overestimation of the temperature in all seasons, as well as to an underestimation of precipitation in summer and relatively moderate overestimations of precipitation in the other seasons. In the hydrological model, compared to the run with the bias corrected forcing data, this translates in higher evapotranspiration losses during growing season (see Fig. 8), less days with frost and snow cover, reduced snow accumulation in winter (see Fig. 9), and drier soils and depleted groundwater stocks after summer.

3.4 Future recurrence intervals

The previous sections have shown the benefit of bias removal for reproducing with the LISFLOOD model observed average and extreme discharge statistics in the control period 1961–1990. Based on the assumption of a stationary error model, future climate is corrected using the transfer functions derived from the control climate. As such, in accordance to the observations in the control period (Fig. 12) but not shown here, future flood magnitudes based on the uncorrected climate simulations largely overshoot those based on corrected climate in most regions of Europe. Hence, although not verifiable, they very likely provide a more biased estimate of future flood magnitudes.

Another interesting aspect, however, is the future recurrence interval of a given flood level observed in the control period, which does not depend on the absolute magnitude, but on the change thereof between the control and future period. Here we therefore show what could be the extent of the errors induced by analysing recurrence intervals based on hydrological simulations driven by uncorrected forcing data compared to its bias corrected counterpart.

We first calculated the discharges associated to a 100-yr flood event for the hydrological simulations driven by uncorrected and bias corrected climate data. This was done employing the Gumbel distribution fitted in the control period (1961–1990). Future recurrence intervals of a 100-yr event observed in the control period were then obtained by translating the respective control period 100-yr discharges using the Gumbel distribution fitted in the future time slice (2071–2100). This procedure was repeated for both simulations driven by uncorrected and bias corrected climate data. Deviations from the control period recurrence interval indicate whether the control period 100-yr event will be more or less frequent in a future climate. Figure 13a shows the future recurrence intervals of a control period 100-yr event, where red indicates that a control 100-yr flood will become more frequent (top scale of legend). In most rivers across Europe the return interval of what is currently a 100-yr flood may in the future decrease to 50 yr or less. A notable exception is the considerable increase in flood recurrence interval (or decrease in frequency) in the northeast, where warmer winters and a shorter snow season will likely reduce the magnitude of the spring snowmelt peak. Also in some other rivers in central and southern Europe an increase in recurrence interval is found.

Figure 13b shows the difference in recurrence interval between the runs driven by corrected and uncorrected climate data. River stretches in white show a good agreement in the change in recurrence interval, red implies that for the simulations driven by bias-corrected forcing data floods are projected to become more frequent compared to the uncorrected driven run, and blue the opposite (bottom scale of legend). Notable is that the simulations driven by uncorrected climate

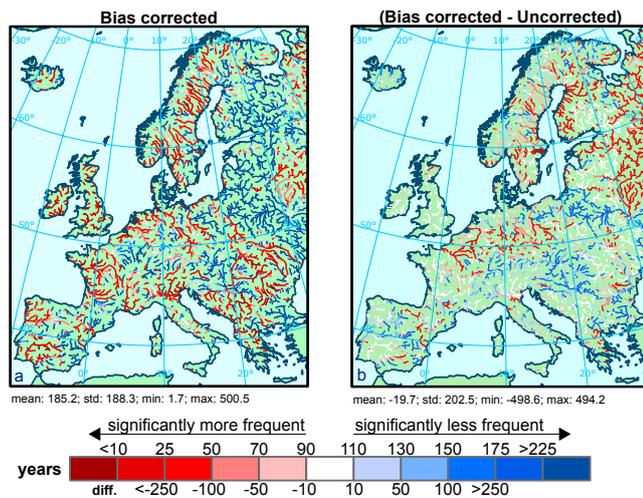


Fig. 13. (a) Recurrence interval of a 100-yr event observed in the control period (1961–1990) for the long-term future (2070–2099), and (b) difference (yr) for the recurrence interval of a 100-yr event between hydrological simulations driven by bias corrected and uncorrected climate data.

data tend to simulate less frequent floods compared to the simulations driven by corrected data in most of northern Europe. Analysis of the control and future snowpack maps reveals that this is caused by a stronger relative reduction in snowpack depth between control and future climate when the climate simulations are not corrected for bias. In the rest of Europe, except for the British Isles where both runs project similar changes in recurrence interval, the recurrence intervals for the uncorrected driven run deviate in both directions from those derived from the bias-corrected driven run. This suggests that estimated recurrence levels based on uncorrected forcing data might wrongly indicate future recurrence intervals of a 100-yr event observed in the control period.

4 Discussion

Despite the strong improvement after implementing the bias correction, some discrepancies between observations and corrected climate simulations of precipitation and temperature remain. This could be related to inherent limitations of the bias correction method employed, limitations of the RCM to properly simulate/capture local orographic effects given the lateral resolution employed in this work (25×25 km), or inadequacies in the observed gridded data set that was used as target to perform the bias correction of the climate simulations.

Within the first group of limitations we should note that by only correcting average, maximum and minimum temperatures, without considering radiative forcings and dewpoint temperature, the energy balance may no longer be preserved. At present, there seems to be no workaround to this prob-

lem as there is no high-quality high-resolution gridded observed data set for these variables to perform the bias correction. Also, results shown herein suggest that any potential bias introduced due to the not-closure of the energy balance does not outweigh the improvement gained by using bias-corrected temperature and precipitation fields.

As noted by Piani et al. (2010b), interpolation of monthly transfer functions to obtain daily values of precipitation or temperature may cause the monthly statistics to not match the observed statistics as daily values in one month are influenced by the values of preceding and following months. In our case, with the exception of few grid cells, temperature and precipitation statistics for the control period 1961–1990 for average conditions are in full agreement with observed statistics, and even if the bias correction of these variables is done on a daily basis, seasonal statistics show a good agreement as well. This, however, does not guarantee that explicit spatio-temporal correlations between different variables, e.g. temperature and precipitation, will be preserved.

Dosio and Paruolo (2011) argue that correlations are appropriate measures of dependence only for variables following multivariate Gaussian or Elliptical distributions and that a more appropriate measure is given by Copula functions (see, e.g. Salvadori and De Michele, 2004). The dependence between precipitation and temperature can be expressed via copulas, which have the notable property of being invariant with respect to monotonically increasing transformations of random variables. In our case, the transfer functions used for bias correction (linear and exponential) are monotonically increasing transformations of precipitation and temperature, implying that uncorrected and bias corrected values have the same Copula function. Dosio and Paruolo (2011) suggest that based on the Copula function as a proper measure of dependence among fields, the univariate bias correction employed in this work preserves the joint prediction of precipitation and temperature and hence the bias correction does not alter the dependence structure between precipitation and temperature. We note here, however, that the discussion on the nature of the relationship between precipitation and temperature and ways of expressing this relationship is still ongoing and is far from settled as highlighted by Piani et al. (2010b).

Clear limitations of the RCMs relate to the poor description of land surface processes, the poor ability to capture the modulation of precipitation in areas of complex orography (Herrera et al., 2010), as well as unresolved/unexplained processes conveyed from the GCMs employed to drive the corresponding RCMs (Giorgi, 2005).

At the same time, it can be argued that the observed gridded data set (E-OBS) employed to perform the bias correction may not always properly match true temperature and/or precipitation due to over-smoothing effects introduced by insufficient station density for the cell-based interpolation (see, e.g. Hofstra et al., 2010; Boberg et al., 2010). Thus, part of the mismatch between the RCM simulations and

observational dataset might actually not be related to model errors, but instead to errors in the observed gridded data set (Lenderink, 2010). In the bias correction, this can potentially corrupt properly simulated temperature or precipitation in areas with low station densities. Also, using an alternative observed gridded data set (e.g. CRU TS1.2 see Mitchell et al., 2004) can lead to different transfer functions. The latter, however, is expected to have limited influence as alternative observed data sets often present similar climatology despite differences in horizontal extent or station density (see, e.g. Rauscher et al., 2010).

Notwithstanding the considerable improvement in reproducing observed discharge statistics after implementing the bias correction, some notable discrepancies between simulated and observed average annual maxima are still present at a number of river stations. Several factors may contribute to this mismatch. Part of the disagreement can be attributed to the remaining error in the meteorological forcing fields after the bias correction.

The disagreement is in part also attributable to not accounting for river regulation in the current LISFLOOD setup. The main reason is a deplorable lack of relevant data at European scale, as well as the large uncertainty regarding future river regulation and land-use changes. As a consequence, results are likely to underestimate the human influence on high flows, and, for future climate, rather reflect the impact of climate change on natural flows.

Errors in the conceptualization and parametrization of the hydrological model further affect the simulated discharges. We do not make an attempt to account for hydrological uncertainty, as it is outside the scope of this study. Several other studies (e.g. Wilby, 2005) showed, however, that this layer of uncertainty is generally much lower than the uncertainty of the climate input to the hydrological model.

The relative contribution of these factors is difficult to quantify because observed events cannot be compared on an individual basis with simulations for the control climate, as it does not reproduce the historical weather.

5 Conclusions

In this work we assessed the benefits of removing bias in climate forcing data for pan-European hydrological impact assessment, with emphasis on extreme events. Results show that the bias correction method employed performed very well in removing bias in average, maximum and minimum temperatures. Even though daily maximum and minimum temperatures were corrected for indirectly, and the daily transfer functions were obtained on a monthly basis, observed annual and seasonal statistics were fully preserved.

For precipitation, the bias correction method was able to drastically reduce the strong overestimation generally observed across Europe. Only in certain mountain areas, e.g. Alps, Apennines, and Carpathian, a persistent overestimation

remained, mainly in winter. This is likely due to the alternation between linear and exponential fitting-type functions during winter months in these areas, which can significantly alter the interpolation of daily values based on the anterior and posterior (monthly) fitting functions.

Validation of simulated discharge statistics at 554 gauging stations showed that LISFLOOD simulations driven by uncorrected forcing data strongly overestimated average and extreme discharge statistics at the majority of stations. Simulations with corrected climate simulations were more consistent with historical discharge records. A strong improvement was observed not only in average discharges but also in the annual maxima and the probability of flood levels derived by extreme value analysis. For the extremes, however, we observed a slight tendency to underestimate observed flow statistics. This can in part be explained by errors in the E-OBS data set used as target in the bias correction. In the latter, gridded extreme precipitation may undervalue true extremes due to under-catch of precipitation, especially in mountain stations, as well as to the over-smoothing effect in the grid-based interpolation in areas with low station density. Despite this, the E-OBS data set is currently the best available at pan-European scale. In small-scale studies such problems may be alleviated by using more dense data sets to reconstruct historical climate.

At a number of stations, the bias-correction could not (fully) remove the (sometimes considerable) discrepancies between observed and simulated discharge statistics. This suggests that they relate to other limitations of the hydrological modelling exercise, such as not accounting for river regulation, or errors in the conceptualization and parametrization of the hydrological model.

In accordance to the results for the control period, future flood magnitudes for the hydrological simulations driven by uncorrected climate data largely exceed those based on corrected climate. Although strictly not verifiable, this suggests that projections of future flood magnitude based on uncorrected climate simulations are likely unreliable. Moreover, results show that the recurrence interval of a 100-yr flood event observed in the control period, which depends on the change in flood magnitude rather than on the absolute value thereof, can substantially vary in both directions from its corrected-based counterpart.

This research has shown the benefits of bias removal in climate simulations for hydrological impact assessment at pan-European scale. Despite the potential limitations in the approach employed, the considerable improvement in the simulation of extreme events and their probability of occurrence in the control period 1961–1990 increases the confidence in the projections of future flood hazard. Our next steps involve the implementation of this technique to the full ensemble of climate simulations available in the FP6 ENSEMBLES project, and the hydrological impact assessment using fully consistent and bias corrected climate simulations for an ensemble of climate models.

Acknowledgements. We thank past and present collaborators at the JRC who have for several years assisted in developing and calibrating LISFLOOD. We would like to thank Claudio Piani (ICTP, Trieste) and Jan Haerter (MPI, Hamburg) for providing the bias-correction algorithms and assisting in implementing the codes. This work received funding from the EU-FP7 project ClimateCost (Full Costs of Climate Change, Grant Agreement 212774, <http://www.climatecost.cc/>). Data used in this work was funded by the EU-FP6 project ENSEMBLES (Contract number 505539). We thank Alessandra Bianchi for managing the data set in netCDF format. We acknowledge the data providers in the ECA&D project (<http://eca.knmi.nl>), the Global Runoff Data Centre (GRDC) (Germany), CEDEX-IEH (Spain), Banque HYDRO (France), and ARPA Emilia Romagna (Italia).

Edited by: T. Elliot

References

- Allen, R., Pereira, L., Raes, D., and Smith, M.: Crop evapotranspiration – Guidelines for computing crop water requirements, vol. FAO irrigation and drainage paper 56, FAO, 1998.
- Beirlant, J., Goegebeur, Y., Segers, J., Teugels, J., De Waal, D., and Ferro, C.: Statistics of extremes – Theory and applications, Wiley series in probability and statistics, Wiley, Chichester, 1 Edn., 2004.
- Block, P., Filho, F., Sun, L., and Kwon, H.: A streamflow forecasting framework using multiple climate and hydrological models, *Journal of the American Water Association*, 45, 828–843, doi:10.1111/j.1752-1688.2009.00327.x, 2009.
- Boberg, F., Berg, P., Thejil, P., Gutowski, J., and Christensen, J.: Improved confidence in climate change projections of precipitation further evaluated using daily statistics from ENSEMBLES models, *Clim. Dynam.*, 35(7–8), 1509–1520, doi:10.1007/s00382-009-0683-8, 2010.
- Chen, C. and Knutson, T.: On the verification and comparison of extreme rainfall indices from climate models, *J. Climate*, 21, 1605–1621, doi:10.1175/2007JCLI1494.1, 2008.
- Christensen, J., Boberg, F., Christensen, O., and Lucas-Picher, P.: On the need for bias correction of regional climate change projections of temperature and precipitation, *Geophys. Res. Lett.*, 35, L20709, doi:10.1029/2008GL035694, 2008.
- Christensen, O., Drews, M., Christensen, J., Dethloff, K., Ketelsen, K., Hebestadt, I., and Rinke, A.: The HIRHAM Regional Climate Model Version 5 (beta), Tech. Rep. 06-17, Danish Meteorological Institute, 2007.
- Coles, S.: An introduction to statistical modeling of extreme values, Springer series in statistics, Springer, London, 1 Edn., 2001.
- Dankers, R. and Feyen, L.: Climate change impact on flood hazard in Europe: An assessment based on high-resolution climate simulations, *J. Geophys. Res.*, 113, D19105, doi:10.1029/2007JD009719, 2008.
- Dankers, R. and Feyen, L.: Flood hazard in Europe in an ensemble of regional climate scenarios, *J. Geophys. Res.*, 114, D16108, doi:10.1029/2008JD011523, 2009.
- de Roo, A., Odiijk, M., Shcmuck, G., Koster, E., and Lucieer, A.: Assessing the effects of land use changes on floods in the Meuse and Oder catchments, *Phys. Chem. Earth, Part B*, 26, 593–599, doi:10.1016/S1464-1909(01)00054-5, 2001.
- Déqué, M.: 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, doi:10.1016/j.gloplacha.2006.11.030, 2007.
- Dosio, A. and Paruolo, P.: Bias correction of the ENSEMBLES high resolution climate change projections for use by impact models: evaluation on the present climate, *J. Geophys. Res.*, 116, D16106, doi:10.1029/2011JD015934, 2011.
- Duan, Q., Sorooshian, S., and Gupta, H.: Effective and efficient global optimization for conceptual rainfall – runoff models, *Water Resour. Res.*, 28, 1015–1031, doi:10.1029/91WR02985, 1992.
- Durman, C., Gregory, J., Kassel, D., Jones, R., and Murphy, J.: A comparison of extreme European daily precipitation simulated by a global and a regional climate model for present and future climates, *Q. J. Roy. Meteorol. Soc.*, 127, 1005–1015, doi:10.1002/qj.49712757316, 2001.
- Dynesius, M. and Nilsson, C.: Fragmentation and flow regulation of river systems in the northern third of the world, *Science*, 266, 753–762, doi:10.1126/science.266.5186.753, 1994.
- EEA: Corine Land Cover 2000, Tech. Rep. 89, European Commission, European Environment Agency (EEA), available at: <http://dataservice.eea.europa.eu/>, 2002.
- Engen-Skaugen, T.: Refinement of dynamically downscaled precipitation and temperature scenarios, *Climatic Change*, 84, 365–382, doi:10.1007/s10584-007-9251-6, 2007.
- Etchevers, P., Golaz, C., Habets, F., and Noilhan, J.: Impact of a climate change on the Rhone river catchment hydrology, *J. Geophys. Res.*, 107, 4293, doi:10.1029/2001JD000490, 2002.
- Evans, J.: Improving the characteristics of streamflow modeled by regional climate models, *J. Hydrol.*, 284, 211–227, doi:10.1016/j.jhydrol.2003.08.003, 2003.
- Feyen, J., Kalas, M., and Vrugt, J.: The value of semi-distributed parameters for large-scale streamflow simulation using global optimization, *Hydrolog. Sci. J.*, 53, 293–308, doi:10.1623/hysj.53.2.293, 2008.
- Feyen, L. and Dankers, R.: Impact of global warming on streamflow drought in Europe, *J. Geophys. Res.*, 114, D17116, doi:10.1029/2008JD011438, 2009.
- Feyen, L., Vrugt, J., Ó Nualláin, B., van der Knijff, J., and de Roo, A.: Parameter optimization and uncertainty assessment for large-scale streamflow simulation with the LISFLOOD model, *J. Hydrol.*, 332, 276–289, doi:10.1016/j.jhydrol.2006.07.004, 2007.
- Fowler, H. and Kilsby, C.: Using regional climate model data to simulate historical and future river flows in northwest England, *Climatic Change*, 80, 337–367, doi:10.1007/s10584-006-9117-3, 2007.
- Fowler, H., Blenkinsop, S., and Tebaldi, C.: Linking climate change modelling to impacts studies: Recent advances in downscaling techniques for hydrological modelling, *Int. J. Climatol.*, 27, 1547–1578, doi:10.1002/joc.1556, 2007.
- Giorgi, F.: Climate change prediction, *Climatic Change*, 73, 239–265, doi:10.1007/s10584-005-6857-4, 2005.
- Giorgi, F., Hostetler, S., and Shields Brodeur, C.: Analysis of the surface hydrology in a regional climate model, *Q. J. Roy. Meteorol. Soc.*, 120, 161–183, doi:10.1002/qj.49712051510, 1994.
- Graham, L., Hagemann, S., Jaun, S., and Beniston, M.: On interpreting hydrological change from regional climate models, *Climatic Change*, 81, 97–122, doi:10.1007/s10584-006-9217-0, 2007.

- Haylock, M., Cawley, G., Harpham, C., Wilby, R., and Goodess, C.: Downscaling heavy precipitation over the United Kingdom: a comparison of dynamical and statistical methods and their future scenarios, *Int. J. Climatol.*, 26, 1397–1415, doi:10.1002/joc.1318, 2006.
- Haylock, M., Hofstra, N., Klein, A., Klok, E., Jones, P., and New, M.: A European daily high – resolution gridded data set of surface temperature and precipitation for 1950–2006, *Geophys. Res. Lett.*, 113, doi:10.1029/2008JD010201, 2008.
- Hellström, C., Chen, D., Achberger, C., and Räisänen, J.: Comparison of climate change scenarios for Sweden based on statistical and dynamical downscaling of monthly precipitation, *Clim. Res.*, 19, 44–55, doi:10.3354/cr019045, 2001.
- Herrera, S., Fita, L., Fernandez, J., and Gutierrez, J.: Evaluation of the mean and extreme precipitation regimes from the ENSEMBLES regional climate multimodel simulations over Spain, *J. Geophys. Res.*, 115, D21117, doi:10.1029/2010JD013936, 2010.
- Hiederer, R. and de Roo, A.: A European flow network and catchment data set, Tech. rep., Joint Research Centre, European Commission, 2003.
- Hofstra, N., New, M., and McSweeney, C.: The influence of interpolation and station network density on the distributions and trends of climate variables in gridded daily data, *Clim. Dynam.*, 35, 841–858, doi:10.1007/s00382-009-0698-1, 2010.
- Horton, P., Schaeffli, B., Mezghani, A., Benoît, H., and Musy, A.: Assessment of climate-change impacts on alpine discharge regimes with climate model uncertainty, *Hydrol. Process.*, 20, 2091–2109, doi:10.1002/hyp.6197, 2006.
- Hurkmans, R., Terink, W., Uijlenhoet, R., Torfs, P., Jacob, D., and Troch, P.: Changes in streamflow dynamics in the Rhine basin under three high-resolution regional climate scenarios, *J. Climate*, 23, 679–699, doi:10.1175/2009JCLI3066.1, 2010.
- Iizumi, T., Nishimori, M., Dairaku, K., Adachi, S., and Yokozawa, M.: Evaluation and intercomparison of downscaled daily precipitation indices over Japan in present-day climate: Strengths and weaknesses of dynamical and bias correction-type statistical downscaling methods, *J. Geophys. Res.*, 116, D01111, doi:10.1029/2010JD014513, 2011.
- Jacob, D., Bärring, L., Christensen, O., Christensen, J., de Castro, M., Déqué, M., Giorgi, F., Hagemann, S., Hirschi, M., Jones, R., Kjellström, E., Lenderink, G., Rockel, B., Sánchez, E., Schär, C., Seneviratne, S., Somot, S., Ulden, A., and Hurk, B.: An inter-comparison of regional climate models for Europe: Model performance in present – day climate, *Climatic Change*, 81, 31–52, doi:10.1007/s10584-006-9213-4, 2007.
- King, D., Daroussin, J., and Tavernier, R.: Development of a soil geographic database from the Soil Map of the European Communities, *Catena*, 21, 37–56, doi:10.1016/0341-8162(94)90030-2, 1994.
- Kjellström, E., Bärring, L., Jacob, D., Jones, R., Lenderink, G., and Schär, C.: Modelling daily temperature extremes: recent climate and future changes over Europe, *Climatic Change*, 81, 249–265, doi:10.1007/s10584-006-9220-5, 2007.
- Kjellström, E., Boberg, F., Castro, M., Christensen, J., Nikulin, G., and Sánchez, E.: Daily and monthly temperature and precipitation statistics as performance indicators for regional climate models, *Clim. Res.*, 44, 135–150, doi:10.3354/cr00932, 2010.
- Kleinn, J., Frei, C., Gurtz, J., Lüthi, D., Vidale, P., and Schär, C.: Hydrologic simulations in the Rhine basin driven by a regional climate model, *J. Geophys. Res.*, 110, D04102, doi:10.1029/2004JD005143, 2005.
- Leander, R. and Buishand, T.: Resampling of regional climate model output for the simulation of extreme river flows, *J. Hydrol.*, 332, 487–496, doi:10.1016/j.jhydrol.2006.08.006, 2007.
- Leander, R., Buishand, T., van den Hurk, B., and de Wit, M.: Estimated changes in flood quantiles of the river Meuse from resampling of regional climate model output, *J. Hydrol.*, 351, 331–343, doi:10.1016/j.jhydrol.2007.12.020, 2008.
- Lehner, B., Döll, P., Alcamo, J., Henrichs, T., and Kaspar, F.: Estimating the impact of global change on flood and drought risks in Europe: A continental integrated analysis, *Climatic Change*, 75, 273–299, doi:10.1007/s10584-006-6338-4, 2006.
- Lenderink, G.: Exploring metrics of extreme daily precipitation in a large ensemble of regional climate model simulations, *Clim. Res.*, 44, 151–166, doi:10.3354/cr00946, 2010.
- Lenderink, G., Buishand, A., and van Deursen, W.: Estimates of future discharges of the river Rhine using two scenario methodologies: direct versus delta approach, *Hydrol. Earth Syst. Sci.*, 11, 1145–1159, doi:10.5194/hess-11-1145-2007, 2007.
- Lorenz, P. and Jacob, D.: Validation of temperature trends in the ENSEMBLES regional climate model runs driven by ERA40, *Climatic Change*, 44, 167–177, doi:10.3354/cr00973, 2010.
- Middelkoop, H., Daamen, K., Gellens, D., Grabs, W., Kwadijk, J., Lang, H., Parmet, B., Schadler, B., Schulla, J., and Wilke, K.: Impact of climate change on hydrological regimes and water resources management in the Rhine basin, *Climatic Change*, 49, 105–128, doi:10.1023/A:1010784727448, 2001.
- Mitchell, T., Carter, T., Jones, P., Hulme, M., and New, M.: A comprehensive set of high-resolution grids of monthly climate for Europe and the globe: the observed record (1901–2000) and 16 scenarios (2001–2100), Tech. rep., Tyndall Centre for Climate Research, Norwich, UK, 2004.
- Murphy, J.: An evaluation of statistical and dynamical techniques for downscaling local climate, *J. Climate*, 12, 2256–2284, doi:10.1175/1520-0442(1999)012<2256:AEOSAD>2.0.CO;2, 1999.
- Nakicenovic, N. and Swart, R. (Eds.): IPCC Special Report on Emissions Scenarios, Cambridge University Press, Cambridge, UK, 2000.
- New, M., Hulme, M., and Jones, P.: Representing twentieth-century space-time climate variability. Part II: Development of 1901–96 monthly grids of terrestrial surface climate, *J. Climate*, 13(13), 2217–2238, doi:10.1175/1520-0442(2000)013<2217:RTCSTC>2.0.CO;2, 2000.
- Nikulin, G., Kjellström, E., Hansson, U., Strandberg, G., and Ullerstig, A.: Evaluation and future projections of temperature, precipitation and wind extremes over Europe in an ensemble of regional climate simulations, *Tellus A*, 63, 41–55, doi:10.1111/j.1600-0870.2010.00466.x, 2011.
- Piani, C., Haerter, J., and Coppola, E.: Statistical bias correction for daily precipitation in regional climate models over Europe, *Theor. Appl. Climatol.*, 99, 187–192, doi:10.1007/s00704-009-0134-9, 2010a.
- Piani, C., Weedon, G., Best, M., Gomes, S., Viterbo, P., Hagemann, S., and Haerter, J.: Statistical bias correction of global simulated daily precipitation and temperature for the application of hydrological models, *J. Hydrol.*, 395, 199–215, doi:10.1016/j.jhydrol.2010.10.024, 2010b.

- Plavcová, E. and Kyselý, J.: Evaluation of daily temperatures in Central Europe and their links to large-scale circulation in an ensemble of regional climate models, *Tellus A*, 63(4), 763–781, doi:10.1111/j.1600-0870.2011.00514.x, 2011.
- Prudhomme, C., Jakob, D., and Svensson, C.: Uncertainty and climate change impact on the flood regime of small UK catchments, *J. Hydrol.*, 277, 1–23, doi:10.1016/S0022-1694(03)00065-9, 2003.
- Rauscher, S., Coppola, E., Piani, C., and Girogi, F.: Resolution effects on regional climate model simulations of seasonal precipitation over Europe, *Clim. Dynam.*, 35, 685–711, doi:10.1007/s00382-009-0607-7, 2010.
- Rijks, D., Terres, J., and Vossen, P.: Agrometeorological applications for regional crop monitoring and production assessment, Tech. Rep. EUR 17735, European Commission, Joint Research Centre, Ispra, Italy, 1998.
- Roeckner, E., Bäuml, G., Bonaventura, L., Brokopf, R., Esch, M., Giorgetta, M., Hagemann, S., Kirchner, I., Kornblueh, L., Manzini, E., Rhodin, A., Sclese, U., Schulzweida, U., and Tompkins, A.: The atmospheric general circulation model ECHAM5 Part I: Model description, Tech. Rep. 349, Max Planck Institute for Meteorology, 2003.
- Salvadori, G. and De Michele, C.: Frequency analysis via copulas: Theoretical aspects and applications to hydrological events, *Water Resour. Res.*, 40, W12511, doi:10.1029/2004WR003133, 2004.
- Schmidli, J., Goodess, C., Frei, C., Haylock, M., Hundscha, Y., Ribalaygua, J., and Schmith, T.: Statistical and dynamical downscaling of precipitation: An evaluation and comparison of scenarios for the European Alps, *J. Geophys. Res.*, 112, D04105, doi:10.1029/2005JD007026, 2007.
- Shabalova, M., van Deursen, W., and Buishand, T.: Assessing future discharge of the river Rhine using regional climate model integrations and a hydrological model, *Clim. Res.*, 23, 233–246, doi:10.3354/cr023233, 2003.
- Steele-Dunne, S., Lynch, P., McGrath, R., Semmler, T., Wang, S., Hanafin, J., and Nolan, P.: The impacts of climate change on hydrology in Ireland, *J. Hydrol.*, 356, 28–45, doi:10.1016/j.jhydrol.2008.03.025, 2008.
- Suklitsch, M., Gobiet, A., Truhetz, H., Awan, N., Göttel, H., and Jacob, D.: Error characteristics of high resolution regional climate models over the Alpine area, *Clim. Dynam.*, 37(1), 377–390, doi:10.1007/s00382-010-0848-5, 2010.
- Teutschbein, C. and Seibert, J.: Regional climate models for hydrological impact studies at the catchment scale: A review of recent model strategies, *Geography Compass*, 4, 834–860, doi:10.1111/j.1749-8198.2010.00357.x, 2010.
- Thiemeßl, M., Gobiet, A., and Leuprecht, A.: Empirical-statistical downscaling and error correction of daily precipitation from regional climate models, *International J. Climatol.*, 31(10), 1530–1544, doi:10.1002/joc.2168, 2011.
- Thorne, R.: Uncertainty in the impacts of projected climate change on the hydrology of a subarctic environment: Liard River Basin, *Hydrol. Earth Syst. Sci.*, 15, 1483–1492, doi:10.5194/hess-15-1483-2011, 2011.
- van der Knijff, J., Younis, J., and de Roo, A.: LISFLOOD: a GIS-based distributed model for river basin scale water balance and flood simulation, *Int. J. Geogr. Inf. Sci.*, 24, 189–212, doi:10.1080/13658810802549154, 2010.
- van der Linden, P. and Mitchell, J.: ENSEMBLES: Climate change and its impacts: Summary of research and results from the ENSEMBLES project, Tech. rep., Met Office Hadley Centre, available at: http://ensembles-eu.metoffice.com/docs/Ensembles_final_report_Nov09.pdf, 2009.
- van Meijgaard, E., van Ulft, L., van de Berg, W., Bosveld, F., van den Hurk, B., Lenderink, G., and Siebesma, A.: The KNMI regional atmospheric climate model RACMO, version 2.1, Tech. Rep. TR-302, KNMI, De Bilt, The Netherlands, 2008.
- van Pelt, S. C., Kabat, P., ter Maat, H. W., van den Hurk, B. J. J. M., and Weerts, A. H.: Discharge simulations performed with a hydrological model using bias corrected regional climate model input, *Hydrol. Earth Syst. Sci.*, 13, 2387–2397, doi:10.5194/hess-13-2387-2009, 2009.
- van Ulden, A. P. and van Oldenborgh, G. J.: Large-scale atmospheric circulation biases and changes in global climate model simulations and their importance for climate change in Central Europe, *Atmos. Chem. Phys.*, 6, 863–881, doi:10.5194/acp-6-863-2006, 2006.
- Weedon, G., Gomes, S., Viterbo, P., Osterle, H., Adam, J., Bellouin, N., Boucher, O., and Best: The WATCH forcing data 19582001: a meteorological forcing dataset land-surface and hydrological-models, Tech. rep., WATCH project, Technical Report Nr. 22, 2010.
- Wilby, R.: Uncertainty in water resource model parameters used for climate change impact assessment, *Hydrol. Process.*, 19, 3201–3219, doi:10.1002/hyp.5819, 2005.
- Wilby, R. and Wigley, T.: Downscaling general circulation model output: a review of methods and limitations, *Progr. Phys. Geog.*, 21, 530–548, doi:10.1177/030913339702100403, 1997.
- Wilby, R., Hay, L., and Leavesley, G.: A comparison of downscaled and raw GCM output: implications for climate change scenarios in the San Juan River basin, Colorado, *J. Hydrol.*, 225, 67–91, doi:10.1016/S0022-1694(99)00136-5, 1999.
- Wilby, R., Whitehead, P., Wade, A., Butterfield, D., Davis, R., and Watts, G.: Integrated modelling of climate change impacts on water resources and quality in a lowland catchment: River Kennet, UK, *J. Hydrol.*, 330, 204–220, doi:10.1016/j.jhydrol.2006.04.033, 2006.
- Wood, A., Leung, L., Sridhar, V., and Lettenmaier, D.: Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs, *Climatic Change*, 62, 189–216, doi:10.1023/B:CLIM.0000013685.99609.9e, 2004.
- Wösten, J., Lilly, A., and Le Bas, C.: Development and use of a database of hydraulic properties of European soils, *Geoderma*, 90, 169–185, doi:10.1016/s0016-7061(98)00132-3, 1999.