Responses referee #2 on the article HESS-2011-117 entitled "Improving pan-European hydrological simulation of extreme events through statistical bias correction of RCM-driven climate simulations"

by R. Rojas et al.

June 24, 2011

We would first like to thank the constructive comments on our manuscript raised by the reviewers. In the next sections, we provide a detailed list with the responses to the major remarks pointed by the reviewers hoping to clarify remaining issues. We proceed by listing the reviewers' comments (**in bold text**) and the corresponding reply.

Anonymous referee #2

Questions/Comments

C1. pg 3890, line 23: The correction seems to be either linear or asymptotically linear (the most significant differences between exponential and the linear correction showing at moderate precipitation amounts). But it makes me wonder whether it can be used to adjust the tail of the distribution (large quantiles). (pg 3892: transfer function approaches a constant slope for large amounts, implies that the influence on the tail is very limited). Would be nice if the effect on the distribution could be made explicit in a plot. Or, for instance, the correction of a certain quantile or the coefficient of variation. Piani 2010 also proposes a logaritmic fit, which however turns out to be less suitable due to the fit errors. Nevertheless, it wouldnt hurt to mention it in one sentence.

A1. Corrected (see revised manuscript). We partially agree with this comment. Using an ensemble of RCMs for Europe, Dosio and Paruolo (2011) suggested that lower-end percentiles of precipitation were more sensitive to the selection of the correction function (linear or exponential–asymptotically linear as referred to by the referee). The latter differs from what is stated by the referee. They also showed a persistent underestimation of the lower-end percentiles of precipitation after bias correction. This affects the wet day frequency and increases the chances of a given day to be considered as a dry day for a constant x_0.

Figure 1 shows the PDF for winter daily precipitation obtained from an ensemble of 11 RCMs for Europe by Dosio and Paruolo (2011). From this figure we observe that the bias corrected seasonal precipitation fits quite well the pdf obtained from the E-OBS data set. As mentioned before, a slight underestimation of lower-end percentiles is observed for all regions. At the same time, Dosio and Paruolo (2011) showed that for higher-end percentiles the bias correction had a significant impact in reducing the bias and inter-model variability. The latter is shown in Figure 2 where it can be seen that the bias correction technique corrects for both the dry bias around the median of the distribution and for the wet bias around the higher end of the distributions.

Following this we believe that it is unnecessary to repeat the work already done by Dosio and Paruolo (2011). To avoid ambiguity we have rephrased the corresponding section pointing the reader to the relevant work of Dosio and Paruolo (2011).

C2. pg 3890, line 5: The E-OBS grid values are not simply averages over the stations within gridvoxes. Haylock 2008 describes the use of thin-plate splines to interpolate the climatology and kiging to interpolate the anomalies, separately, selected from a variety of methods.

A2. Corrected (see revised manuscript). We have partially rephrased sections referring to the E-OBS data set to reflect what is suggested by the referee.

C3. pg 3892, line 4: Do modelgrid and EOBS-grid match exactly or is regridding necessary? If so, what is the effect of regridding on the smoothing? (I do not expect this to influence the hydrological simulations, however).

A3. Models from the ENSEMBLES project match exactly the E-OBS grid. However, regridding was necessary to map the input forcing data into the LIDFLOOD grid (5km). This is explained in section 2.3 of the manuscript (Hydrological simulation). We have not performed any analysis on the effect of the regridding or smoothing from the 25km to 5km resolution. The only available configuration of LISFLOOD for pan-European scale simulations is at a 5km resolution. As pointed by the reviewer, however, we do not expect this to influence the hydrological simulations.

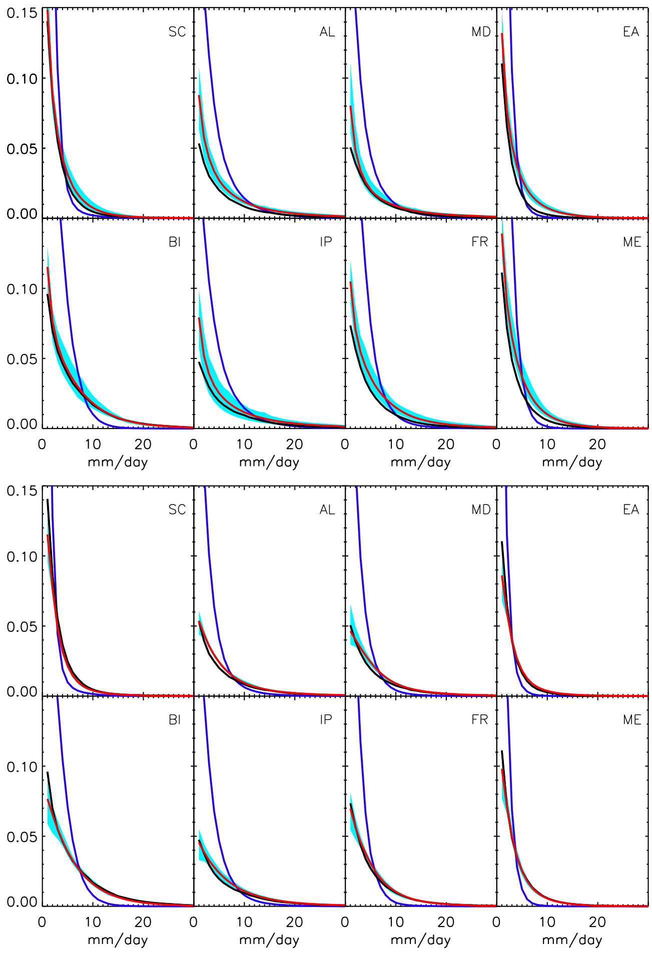

Figure 1: PDFs of winter daily precipitation averaged over geographic areas (for acronyms see Dosio and Paruolo 2011). First two rows: uncorrected RCMs; Bottom two rows: bias corrected RCMs. The light blue area represents the range of the ensemble. The black line is the PDF calculated form the E-OBS dataset. The red line is the average of the different models PDFs. The blue line is the PDF of the ensemble average of the models daily values. (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, Journal of Geophysical Research, doi:10.1029/2011JD015934, 2011. Copyright 2011 American Geophysical Union. Reproduced by permission of the American Geophysical Union).

C4. pg 3892, line 17: If this is the most deficient model, the necessity of bias-correction of its data does not imply the necessity of applying bias-corrections to RCM-data in general. Besides, If the biases are that large, and the changes due to the needed corrections is far larger than the climate-change signal, can the (corrected) future simulation still be trusted?

A4. The improvements of performing bias correction on climate data used to drive impact models have been discussed in the literature (see, e.g., Fowler and Kilsby, 2007; Christensen et al., 2008; Teutschbein and Seibert, 2010). The results of this work together with the results of Dosio and Paruolo (2011) support the idea that DMI-HIRHAM-ECHAM5 is one of the models showing strong wet and warm biases across Europe. Thus, if this bias correction method works properly for the most deficient model we could expect a similar or better performance for climate models showing weak biases.

Additionally, whether (corrected) future simulations can be trusted or not is an hypothesis that can not be tested. However, by more closely reproducing observed climate after bias correction, the confidence in future climate projection increases.

C5. pg 3893, line 13: Insert the text: "Given the availability...values derived in the range 5¡b¡0.2." (which also describes the correction of precipitation) here before discussing temperature.

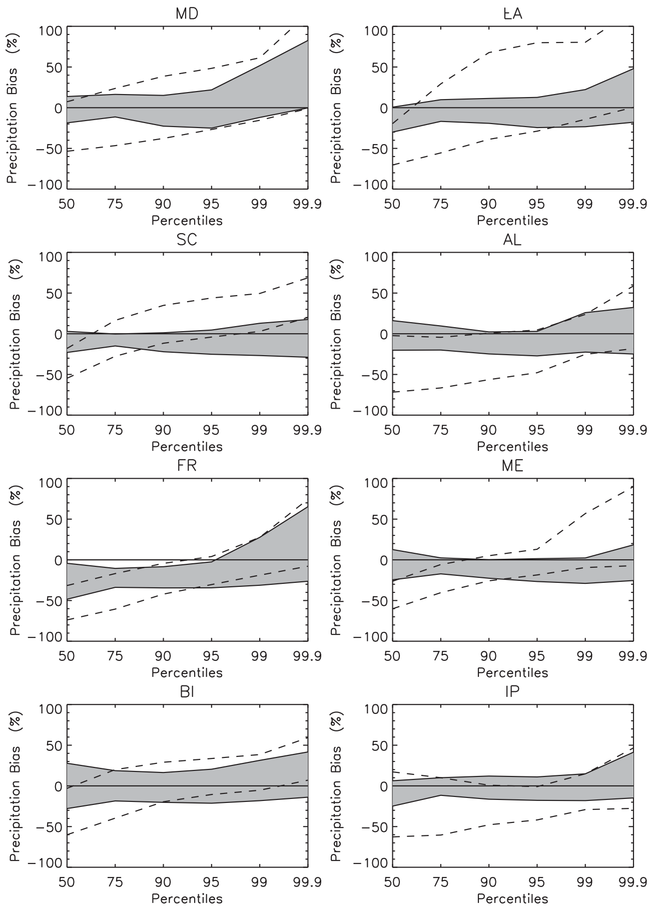

Figure 2: Bias of summer daily precipitation averaged over geographic areas shown as a function of different percentiles. The dashed lines represent the range of ensemble before the bias correction. The shadowed areas represents the ensemble range after bias correction (Dosio, A. and Paruolo, P.: Bias correction of the ENSEMBLES high resolution climate change projections for use by impact models: evaluation on the present climate, Journal of Geophysical Research, doi:10.1029/2011JD015934, 2011. Copyright 2011 American Geophysical Union. Reproduced by permission of the American Geophysical Union).

A5. Corrected (see revised manuscript).

C6. pg 3893: Does switching to an exponential fit lead to a decrease of slope b? If the low/moderate values decrease or become zero in the transfer function, one would expect the fitted slope to be even larger.

A6. Switching from linear to exponential transfer function is done under two conditions: first, if $a > 0$ (i.e. ignoring dry days completely in the corrected precipitation) and, second, when the slope of the linear fit is too extreme (with arbitrary values $b < 1/5$ and $b > 5$). For lower values of precipitation, the exponential TF is compensating for the number of dry days redistributing the anomalously frequent number of drizzle days in the simulated precipitation (Piani et al., 2010). Therefore, and as pointed by the reviewer the slope of the TF tends to increase for low value of precipitation to compensate for the definition of x_0 (dry day correction). The latter has been previously suggested by Piani et al. (2010).

C7. pg 3898, line 23: It is not really a surprise that the mean bias-corrected precipitation corresponds so well with the observations.

A7. Corrected (see revised manuscript). This text has been removed.

C8. pg 3899, line 13: moving -¿ switching, alternating Can it be explained why in particular mixing both correction methods makes the bias-correction less effective? Does it have to do with the interpolation between subsequent months?

A8. A possible explanation is discussed immediately after the sentence highlighted by the referee's comment. We observed some correspondence between grid cells showing problems to reduce the wet bias and the change in the monthly fitted function. As explained in the text, interpolating from monthly TFs to daily TFs may alter the monthly statistics (see, e.g., Piani et al., 2010). The latter may as well affect the seasonal statistics shown in the figure. At the same time, we observed for Europe that the exponential TF was fitted predominately in summer months (JJA) and September whereas the rest of the months tended towards the linear TF. This is explained by the expected poor performance of the linear TF for months with little precipitation.

We must emphasize, however, that problems are limited to few grid cells and even in these cells we observed a drastic reduction of the overestimation after bias correction.

C9. pg 3900, line 5: "This implies...amounts" Why? Even if the wet-day frequency is correct, the rest of the distribution can still be off.

A9. Corrected (see revised manuscript). This was a somewhat vague statement, which has now been removed.

C10. pg 3900, line 28: "..a tendency that....upper-end percentiles..." I doubt if the latter is common for RCMs and for different regions. I have seen many RCMs underestimating the high quantiles while overestimating the mean precipitation.

A10. Corrected (see revised manuscript). We have specified that this has been observed for climate model simulations in the framework of the ENSEMBLES project.

C11. pg 3901, line 12: Underestimation of the variability of multi-day precipitation usually indicates lack of serial correlation. Was the effect of the correction on the autocorrelation (and spatial correlation) investigated? Both can influence the simulation of discharge extremes. Serial correlation of temperature and coherence between temperature and precipitation is important for the contribution of melting snowpack to extreme discharge.

A11. We fully agree with the reviewer in this point. As highlighted by Piani et al. (2010), there is no guarantee that serial and spatial correlations of dynamically linked fields such as temperature and precipitation are maintained, which is an unwanted consequence of any bias correction technique involving univariate correction of multiple fields. We suspect, however, that the slight underestimation of the multi-day precipitation may be related to the underestimation of the wet day frequency. The latter is explained by the sensitivity of the lower-end percentiles of the bias corrected precipitation to the selected transfer function (linear vs. exponential) observed by Dosio and Paruolo (2011). At the same time, they found a systematic underestimation of the small values of bias corrected precipitation compared to the observed pdf obtained from the E-OBS data set (see Figure 14 in Dosio and Paruolo 2011). Additionally, and as explained in the manuscript, the exponential transfer function employs the dry day correction ($x_0 = -a/b$), defined as the value below which bias corrected precipitation is set to zero. x_0 is obtained directly from the optimized parameters (a, the additive correction factor and b, the multiplicative correction factor of the linear asymptote) to compensate for the number of dry days in the observed precipitation. As highlighted by Piani et al. (2010), $x_0 > 0$, since observed precipitation shows many more dry days compared to the simulated precipitation from RCMs. This is explained by the systematic bias of RCMs to simulate too many days with too little precipitation (drizzle), thus, underestimating the number of dry days. As a consequence, the underestimation of the wet days frequency after bias correction is likely explained by: first, 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, and second, by the potential overestimation of the number of dry days given that x_0 is obtained from the optimized parameters a and b solely, which may differ from the actual number of dry days in the observed precipitations.

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. The dependence between precipitation and temperature can be expressed in terms of the Copula, which poses the notable property of being invariant with respect to monotonically increasing transformations of the 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 variables have the same Copula function. They 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 that, as highlighted by Piani et al. (2010), the discussion on the nature of the relationship between precipitation and temperature and ways to express this relationship is still ongoing and is far from settled.

C12. pg 3902, line 4: Is this "maximum and minimum daily temperature" or "daily maximum and minimum temperature" mean here? Or is the maximum of daily max and minimum of daily min.

A12. Corrected (see revised manuscript). These are daily maximum and minimum temperatures.

C13. pg 3903, line 26: "Evapotranspration..." If one focusses purely on the effect of the bias correction, then it suffices to compare the evapotranspiration pattern with and without the bias-correction, without having to mention that observational data is lacking. Same for snow accumulation. If LISFLOOD calculates the actual evapotranspiration and evaporation (using Pennman-Monteith I suppose), is this still consistent with the RCM? The latter has its own water balance which should be consistent with the hydrological model (water from the soil = water into the atmosphere. Since both are models are not coupled, the total amount of moisture might not be conserved.

A13. Corrected (see revised manuscript). Notwithstanding RCMs have considerably advanced in reproducing regional an local climate, they still show deficiencies and systematic errors in reproducing local hydrology. As explained in the text, this is mainly due to resolution issues and the poor representation of land surface processes (see, e.g., Giorgi et al., 1994; Evans, 2003; Fowler and Kilsby, 2007). Therefore, and even if both RCMs and LISFLOOD should conserve the total moisture (independently from each other as they are run off line), we believe LISFLOOD will give a more reliable estimation of evaporation and evapotranspiration given the detailed information employed for its calculation (e,g. HYPRES data base and CORINE2000 land cover).

C14. pg 3904, line 26 "Depending...meters" can be omitted: only SWE is important. For the both snow accumulation and soil moisture serial correlation of temperature and precipitation are important.

A14. Corrected (see revised manuscript).

C15. pg 3905, line 8: discharge statistics at the 554... -¿ average discharge and average annual maximum discharge for each of the 554.... maybe more clear to the reader what he sees in this plot. Observed discharge is compared here with modelled discharge based on modelled climate data. If discrepancies are found (as in the extreme discharge) it is unclear where in the chain of models (RCM, biascorrection, rainfallrunoff) it came from. Since the focus is on correcting the climate data, maybe it is more useful to compare the discharge obtained from using corrected RCM data with that obtained with E-OBS data. Discrepancies described here can arise from the parameters in LISFLOOD and the discharge measurements, change in riverbasins, human intervention...etc.

A15. We agree with this comment. However, it must be emphasized that we are using exactly the same LISFLOOD setup (resolution, parameters, iput constants, etc.) for runs driven by uncorrected and bias corrected climate. Unfortunately, to run LISFLOOD forced by the E-OBS data set key information is missing. In order to perform a "clean" run to assess the impact of the bias correction on extreme discharges, we need nine fields. From the E-OBS we only have 4 (daily avg, max and min temperatures and precipitation) lacking relevant variables for the calculation of the evapotranspiration demands used to force LISFLOOD (e.g., radiative forcings, dewpoint temperatures, wind speed, albedo).

C16. pg 3906, line 8: Smoothing as mentioned by Haylock shouldnt have too much influence on discharges extremes, for riverbasins larger than the EOBS-gridspacing.

A16. There seems to be a misunderstanding on this issue. We are highlighting potential impacts of considering several stations within a single cell in the grid-box used in the E-OBS data set. Haylock et al. (2008) and Hofstra et al. (2010) refer to this problem and the impact it may have on extreme precipitation events, mainly, underestimating the actual value of maximum precipitation. Referee's comment refers to extreme discharges, where we could expect a more important impact of underestimation of precipitation, for instance for rapid onset of extreme discharges.

C17. pg 3906, line 20: Ten out of twenty seem ok, for the rest differences between simulated and observed extremes are markedly outside the confidence intervals for the observations.

A17. Corrected (see revised manuscript).

C18. pg 3908, line 3: Remove Spatially In northern Europe, the influence of dry summers on discharge extremes (taking place in winter) should be small, So the dry-summer bias is by far not so relevant as the effect on the snow.

A18. Corrected (see revised manuscript). We agree on this issue. Particularly, and on the basis of an ensemble analysis of 12 climate models at pan-European scale, for summer (JJA) season (and Spring–MAM) in Scandinavia we observe a clear effect of changes in the depth of the snowpack (mm SWE), temperature and 5-d precipitation on the seasonal extreme discharges. Even if many climate models predict an increase in the 5-days precipitation (as well as 3-days and 7-days), we observe a decrease in the seasonal maximum discharges. This decrease is clearly related to drastic reductions in the depth of the snow cover (60% or more compared to the control period 1961–1990) available in the region, which is at the same time related to a persistent increase in average and minimum temperatures. Therefore, even if precipitation is projected to increase in the future in Scandinavia, this does not compensate for the reduction in the depth of snow cover, thus, showing a decrease in the seasonal maximum discharges. A detailed explanation on this issue, however, is left for a forthcoming publication as it is beyond the scope of this article.

C19. pg 3908, line 17: simulated -¿ project, found Is there a spatial pattern in the decrease/increase of return periods, if so why?

A19. Corrected (see revised manuscript). Yes, there is. From an ensemble analysis including 12 climate models from the ENSEMBLE project, we have found a relationship in certain areas of Europe (Scandinavia and Northern Poland and Germany) among depth of snow cover (mm SWE), number of frost days, 5-days maximum precipitation, and seasonal maximum discharges. For Scandinavia we observe an increase in 5-days maximum precipitation, a persistent reduction in the depth of snow cover (mm SWE) and an increase in average and minimum temperatures in the future. At the same time, for northern Germany and Poland and to lesser extent in north-eastern of Carpathians, we observe a drastic decrease in the number of frost days, mild increases in precipitation and drastic reductions in snowpack depths in the future. On a seasonal basis we observe the following patterns: a) in Spring and Summer for Scandinavia there is an increase in precipitation which does not compensate for the reduction in the depth of the snowpack and, thus, maximum discharges decrease, for Winter and Autumn increase in precipitation compensates for decreases in snowpack depth and, thus, seasonal maximum discharges increase. For northern Germany and Poland, we observe an increase in the seasonal maximum discharges for Summer and Autumn associated to mild increases in precipitation, whereas we observe a decrease in the seasonal maximum discharges in Winter and Spring associated to decreases in the snowpack depth and the number of frost days in the area. There seems to be a complex (spatial-temporal) interaction among these variables which will be further explored in a forthcoming publication.

C20. pg 3909, line 16: The observed gridded data are used for fitting the correction, but also for the validation of the result, So inadequacies in the gridded data should cancel each other this way.

A20. Corrected (see revised manuscript).

C21. pg 3909, line 19: solely-¿separately,independently

A21. Corrected (see revised manuscript).

C22. pg 3909, line 20: For example...evapotranspiration The ingredient for the evapotransp should already be discussed on page 3904 (top). Even with multiple variables which are all independently corrected, their combination can still give rise to a bias. Furthermore, (actual) evapotranspiration in LISFLOOD maybe inconsistent with the climate model, since there is no feedback from the hydrological model.

A22. Corrected (see revised manuscript). We have moved the corresponding text to fit the restructuring of the new section Discussion and Conclusions. We agree that actual evapotranspiration values may differ between LISFLOOD and the RCM. However, as mentioned before, we believe LISFLOOD calculations of potential and actual evapotranspiration are more accurate than calculations from RCMs given the good representation of land surface processes (see A13), the finer resolution and the high quality land cover data set employed (CORINE2000) in LISFLOOD.

C23. pg 3910, line 5: aiming at coping with -¿ to anticipate

A23. Corrected (see revised manuscript).

C24. pg 3910, line 14: physical processes...identical time scales What is meant by this?

A24. Corrected (see revised manuscript). This statement is meant to explain that the statistical properties of a variable might be different based on different time scales (daily-monthly-yearly). Therefore, bias correction should be performed on variables observed at the same time scale.

C25. pg 3910, line 15: horizontal spatial -¿ lateral

A25. Corrected (see revised manuscript).

C26. pg 3910, line 29:-pg 3911, line 3: The errors in the gridded data due to varying station density have already been mentioned often, contrary to the errors in discharge measurements. In the last line of this paragraph, the argument of station density suffers.

A26. Corrected (see revised manuscript). We have rephrased the sentence accordingly. In addition, we have shortly discussed errors in discharge measurments.

C27. pg 3911, line 5-16: can be omitted, all issues have been mentioned already. "Other forcing...energy balance" has to do with evapotranspiration, which is not likely to dominate situations of extreme floods.

A27. Corrected (see revised manuscript).

C28. pg 3911, line 19: ..river regulation and land-use changes

A28. Corrected (see revised manuscript).

C29. Conclusion and discussion should be joined to avoid unnecessary repetition

A29. Corrected (see revised manuscript).

C30. pg 3913, line 9: to be subject to large error -¿ unreliable

A30. Corrected (see revised manuscript).

C31. pg 3913, line 17: hardly reason to become over-confident; using an ensemble of model-configurations will likely reveal a variety of problems in correcting and eventually a large spread of projected changes.

A31. We stated this sentence as a future research goal and activity. This will serve the purpose of understanding the performance of the BC method for an ensemble of climate models and, at the same time, to verify the tendencies and relationships described in this response among snowpack depth, precipitation, number of frost days and extreme discharges in certain regions of Europe.

My recommendations to the authors would be * To separate remaining discrepancies between observed flood extremes and simulations to their source: - running LISFLOOD with gridded meteo-data (EOBS or CRU) - compare different gridded data sets to see what the largest source of error is. I suspect that discharge measurements are also part of it (trends, changes etc), in which case EOBS-driven LISFLOOD could also serve as a reference. * To inspect the effect of the correction method on spatial and temporal correlation, relevant to extreme discharge. Because of the snowmelt, also the persistence of temperature is important. * Find out why this method works so well for some river basins, while it fails for others, there must be a pattern (size and nature of the basin, geographical location). * Merge and concise Conclusions with Discussion.

A32.

- Separate remaining discrepancies between observed flood extremes and simulations using a "clean" LISFLOOD run driven by E-OBS is not possible as relevant fields to calculate the evapotranspiration (radiative forcings, dewpoint temperatures, wind speed and albedo) are missing. In addition, we believe that by combining different sources of information (e.g. E-OBS and CRU) to perform a LISFLOOD run driven by observed data makes the task of identifying the discrepancies between observed and simulated extreme discharges even harder.

- Inspect the effect of the BC method on spatial and temporal correlation of variables relevant for extreme discharges is beyond the scope of this article and it has been partially covered by Dosio and Paruolo (2011). We found an apparent relationship between switching of monthly transfer functions and cells showing a somewhat weaker performance to reduce the overestimation of precipitation. We must note here that this is not conclusive as we are only reporting results for one climate model (DMI-HIRHAM5-ECHAM5). Related to this, results from Dosio and Paruolo (2011) who applied and validated the same BC method for an ensemble of 11 models from the ENSEMBLES project provide insight on the method and the proper ways to assess the effect of the BC method on the correlations among variables after corrections. They found an excellent performance of the method for a validation period of 10 years (1991–2000). In addition, they argue that correlations are not the most suitable measure to assess relationships among variables and that to assess a proper correspondence we should resort to Copulas functions.
- Again, this goes beyond the scope of this work and it has been partially covered by Dosio and Paruolo (2011). In general, and following our results and those by Dosio and Paruolo (2011), the method works acceptably on all the domain studied. A comparison against the non-parametric version of the method showed that performance of both methods is rather similar, indicating that the parametric method (fitting of linear of exponential as used in our work) is more efficient for large-scale applications. In the work of Dosio and Paruolo (2011) insight on the spatial assessment of the method is provided.
- For the rest, we have answered and implemented all comments raised by the referee.

Bibliography

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, Geophysical Research Letters, 35, doi:10.1029/2008GL035694, 2008.

Dosio, A. and Paruolo, P.: Bias correction of the ENSEMBLES high resolution climate change projections for use by impact models: evaluation on the present cimate, Journal of Geophysical Research, in press, doi:10.1029/2011JD015934, 2011.

Evans, J.: Improving the characteristics of streamflow modeled by regional climate models, Journal of Hydrology, 284, 211–227, doi:10.1016/j.jhydrol.2003.08.003, 2003.

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.

Giorgi, F., Hostetler, S., and Shields Brodeur, C.: Analysis of the surface hydrology in a regional climate model, Quarterly Journal of the Royal Meteorological Society, 120, 161–183, doi: 10.1002/qj.49712051510, 1994.

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, Geophysical Research Letters, 113, doi:10.1029/2008JD010201, 2008.

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, Climate Dynamics, 35, 841–858, doi:10.1007/s00382-009-0698-1, 2010.

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, Journal of Hydrology, 395, 199–215, doi:10.1016/j.jhydrol.2010.10.024, 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.