Hydrol. Earth Syst. Sci. Discuss., 9, 12765–12795, 2012
www.hydrol-earth-syst-sci-discuss.net/9/12765/2012/
doi:10.5194/hessd-9-12765-2012

Hydrology and Earth System Sciences Discussions

Is bias correction of Regional Climate Model (RCM) simulations possible for non-stationary conditions?

C. Teutschbein[1] **and J. Seibert**[1,2,3]

[1]Department of Physical Geography and Quaternary Geology, Stockholm University, Svante Arrhenius Väg 8, 10691 Stockholm, Sweden
[2]Department of Geography, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland
[3]Department of Earth Sciences, Uppsala University, Villavägen 16, 75236 Uppsala, Sweden

Received: 24 October 2012 – Accepted: 27 October 2012 – Published: 12 November 2012

Correspondence to: C. Teutschbein (claudia.teutschbein@natgeo.su.se)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

In hydrological climate-change impact studies, Regional Climate Models (RCMs) are commonly used to transfer large-scale Global Climate Model (GCM) data to smaller scales and to provide more detailed regional information. However, there are often considerable biases in RCM simulations, which have led to the development of a number of bias correction approaches to provide more realistic climate simulations for impact studies. Bias correction procedures rely on the assumption that RCM biases do not change over time, because correction algorithms and their parameterizations are derived for current climate conditions and assumed to apply also for future climate conditions. This underlying assumption of bias stationarity is the main concern when using bias correction procedures. It is in principle not possible to test whether this assumption is actually fulfilled for future climate conditions. In this study, however, we demonstrate that it is possible to evaluate how well bias correction methods perform for conditions different from those used for calibration. For five Swedish catchments, several time series of RCM simulated precipitation and temperature were obtained from the ENSEMBLES data base and different commonly-used bias correction methods were applied. We then performed a differential split-sample test by dividing the data series into cold and warm respective dry and wet years. This enabled us to evaluate the performance of different bias correction procedures under systematically varying climate conditions. The differential split-sample test resulted in a large spread and a clear bias for some of the correction methods during validation years. More advanced correction methods such as distribution mapping performed relatively well even in the validation period, whereas simpler approaches resulted in the largest deviations and least reliable corrections for changed conditions. Therefore, we question the use of simple bias correction methods such as the widely used delta-change approach and linear scaling for RCM-based climate-change impact studies and recommend using higher-skill bias correction methods.

1 Introduction

In hydrological climate-change impact studies, large-scale climate variables for current and future conditions are generally provided by Global Climate Models (GCMs). To resolve processes and features relevant to hydrology at the catchment scale, Regional Climate Models (RCMs) are commonly used to transfer coarse-resolution GCM data to a higher resolution. Although this provides more detailed regional information (Fowler et al., 2007; Grotch and MacCracken, 1991; IPCC, 2007; Salathé, 2003) for hydrological simulations, there is still a mismatch of scales especially for meso- and small-scale watersheds that are often captured by only one RCM grid cell. In addition, impact modelers are also facing a risk of considerable biases (i.e. systematic model errors) in RCM simulations (Christensen et al., 2008; Teutschbein and Seibert, 2010; Varis et al., 2004). Mismatching scales in combination with these biases have led to many recently developed bias correction approaches (Johnson and Sharma, 2011; Teutschbein and Seibert, 2012) that help impact modelers to cope with the various problems linked to biased RCM output.

Bias correction approaches can be classified according to their degree of complexity and include simple-to-apply methods such as scaling factors but also more advanced methods such as probability mapping or weather generators. The correction procedures usually identify possible biases between observed and simulated climate variables, which provide the basis for correcting both control and scenario RCM runs with a transformation algorithm. Although a bias correction of RCM climate variables considerably improves hydrological simulations (Teutschbein and Seibert, 2012), there is a major drawback: all bias correction methods follow the assumption of stationarity of model errors, which means that the correction algorithm and its parameterization for current climate conditions are assumed to also be valid for a time series of a changed future climate. Whether or not this condition is actually fulfilled for future climate conditions can fundamentally not be evaluated. This motivated us to find a method to address this issue and to test how well bias correction methods perform for conditions

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

HESSD

9, 12765–12795, 2012

Bias correction for non-stationary conditions

C. Teutschbein and J. Seibert

different from those used for calibration. We applied the idea of a differential split-sample test, originally proposed by Klemeš (1986) for hydrological models, to bias correction procedures and applied it to analyze the performance of different bias correction methods for use with simulations under changed conditions.

The testing presented here was done for different commonly-used bias correction procedures (Johnson and Sharma, 2011; Teutschbein and Seibert, 2012) based on 11 RCM-simulated temperature and precipitation series for five meso-scale catchments in Sweden. The aim was to demonstrate that differential split-sample testing is a powerful tool to evaluate the performance of bias correction procedures for the most relevant conditions when it comes to impact studies, namely those conditions, which are different from those used for calibration of the procedures.

2 Methods

2.1 Study catchments

The analysis in this study was performed for five meso-scale catchments (Fig. 1) with areas ranging from 147 to 293 km^2. These catchments fall all below the standard RCM grid cell size of approximately 25×25 km and are, therefore, potentially affected by the scaling issue. The chosen catchments represent different typical Swedish climatic conditions and land-use types (Table 1) and were studied in earlier publications by Teutschbein and Seibert (2010, 2012) in terms of climate-change impacts on hydrological regimes. Continuous temperature and precipitation measurements for all five catchments were available for the standard period 1961–1990.

2.2 Data

Daily temperature and precipitation measurements for the period 1961–1990 were taken from a spatially interpolated 4×4 km national grid (Johansson, 2002) provided by the Swedish Meteorological and Hydrological Institute (SMHI). The values were

HESSD
9, 12765–12795, 2012

obtained by averaging all grid cells containing parts of the catchment. Climate simulations were obtained from the ENSEMBLES project (Van der Linden and Mitchell, 2009): we used daily precipitation and temperature series for the same period 1961–1990 (control run) simulated by 11 RCMs driven by different GCMs (Table 2). The chosen RCMs have a resolution of 25 km and, thus, the area of a single grid cell clearly exceeds the size of the study catchments. We chose to take precipitation and temperature values from the RCM grid cell with center coordinates closest to the center of the catchment, as Teutschbein and Seibert (2010) found that values of one grid cell do not differ considerably from the average over nine grid cells (i.e. over one grid cell and its eight neighboring cells) for the five Swedish catchments which are also used in this study.

2.3 Background on bias correction methods

RCM simulations are typically affected by systematic model errors: misestimated climate variables, incorrect seasonal variations of precipitation (Christensen et al., 2008; Terink et al., 2009; Teutschbein and Seibert, 2010) and the simulation of too-many drizzle (i.e. low-intensity rain) days (Ines and Hansen, 2006) are just a few examples of common biases. Therefore, climate variables simulated by individual RCMs do often not agree with observed time series (Fig. 2), which poses a problem for using simulations of a single RCM as input data for hydrological impact studies. One possible solution is to use an ensemble of RCM simulations (Déqué et al., 2007; Giorgi, 2006; Teutschbein and Seibert, 2010). As Teutschbein and Seibert (2010) concluded, multi-model approaches (i.e. ensembles) have two advantages: (1) the spread of individual ensemble members covers a more realistic range of uncertainty and (2) the ensemble mean may fit observations better, which is especially true for temperature simulations (Fig. 2, top). However, for precipitation simulations even the ensemble mean deviates considerably from observations and is not able to capture the variability in the observations (Fig. 2, bottom). This shows that it is not enough to only employ an RCM

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

HESSD

9, 12765–12795, 2012

Bias correction for non-stationary conditions

C. Teutschbein and J. Seibert

ensemble, but that additional bias correction procedures are needed (Teutschbein and Seibert, 2012).

Typical bias correction approaches adjust RCM climate variables by employing a transformation algorithm. The concept is based on the identification of possible biases between observed and simulated climate variables, which is the starting point for correcting both control and scenario RCM runs. In addition to original (i.e. uncorrected) RCM output data, we analyzed the following 6 bias correction methods to adjust RCM simulations: (1) linear scaling, (2) local intensity scaling, (3) power transformation, (4) variance scaling, (5) a distribution transfer method and (6) the delta-change approach. In addition, a precipitation threshold was used in combination with other bias correction procedures, but not considered an appropriate "stand-alone" method. More detailed descriptions of these methods can be found in Teutschbein and Seibert (2012), Gudmundsson et al. (2012), Johnson and Sharma (2011) and the original method publications provided in Table 3.

Bias correction is often unavoidable but also a controversial subject (Ehret et al., 2012; Muerth et al., 2012). Despite their advantageous ability to reduce biases in climate model output, bias correction methods are criticized to diminish the advantages of climate models (Ehret et al., 2012) and to not have much added value in a complex modeling chain when considering other sources of uncertainty (Muerth et al., 2012). Furthermore, a common assumption of all bias correction methods (Table 3) is stationarity, or time invariance, of the biases, i.e. the empirical relationships in the correction algorithm and its parameterization for current climate conditions do not change over time and are also valid for future conditions.

2.4 Testing of bias correction methods under non-stationary conditions

The stationarity assumption is a major limitation of any bias correction procedure. The question arises as to whether it is possible to provide any confidence that the correction algorithms that are applied to today's climate are also valid for a future climate. It is a questionable assumption which merely has to be made, because we are lacking

HESSD
9, 12765–12795, 2012

appropriate methods to deal with changing climate conditions and possible changes in bias relationships. We are not able to check whether this assumption is actually true or not. However, it is possible to test how well bias correction methods can reproduce conditions different from those that they were calibrated to by using one of the operational testing methods presented by Klemeš (1986).

The hierarchical scheme for systematic testing of hydrological models outlined by Klemeš (1986) includes two approaches of interest for systematic testing of hydrological model transposibility: split-sample testing (SST) for stationary conditions and differential split-sample testing (DSST) for non-stationary conditions. SST implies the splitting of an available data record into two (preferably equally sized) segments in order to use one as calibration and one as validation period. DSST on the other hand should, according to Klemeš (1986), be used under changing conditions. The first step of this test includes the identification of two periods with the climate variable of interest having different values, for instance a warm versus a cold or a wet versus a dry period. The model is then calibrated on the period with one condition and validated on the period with the other condition, which allows analyzing the model's ability to perform under shifting conditions. SST can automatically transform into DSST, if the two segments by nature show substantial differences in their conditions (Klemeš, 1986).

We applied DSST proposed by Klemeš (1986) that was originally intended for hydrological models to test the ability of different bias correction procedures to reliably work for changed climate conditions. Both SST and DSST are seldom used to evaluate bias correction methods. We are aware of only a few other studies using such a test: Bennett et al. (2010) and Terink et al. (2010) evaluated bias correction methods using a SST with two different time periods for which observations were available. A major limitation of this approach is that the periods should be long enough to represent natural climate variability satisfactorily (Bennett et al., 2010). Furthermore, unless the two periods are different in their conditions, the methods are not evaluated for use under changed conditions. This issue motivated us to rather use DSST that is better able to evaluate performance under changing climate conditions (Li et al., 2012; Seiller et al., 2012). The available 30-yr period 1961–1990 was separated into two 15-yr periods with different climate conditions, one representing current climate and the other one future climate. Since our available 30-yr period was not long enough to show a considerable trend in precipitation or temperature data, we chose the two required segments as follows: given that climate projections indicate an increase in future precipitation and temperature for Northern Europe (IPCC, 2007), we compiled the two periods by sorting the years according to their amount of precipitation and temperature, respectively (Fig. 3). For the precipitation-bias correction assessment, we included the 15 driest years in the first subset ("calibration years") and the 15 wettest years in the second subset ("validation years"). For the temperature-bias correction evaluation, we used the 15 coldest years as "calibration years" and the 15 warmest as "validation years". This procedure was done to all 11 RCM-simulated times series and the observed times series. Thus, DSST allows the evaluation of bias correction methods under relatively challenging conditions (i.e. climate conditions considerably different from calibration) pushing them to their performance limits (Coron et al., 2012).

In this study, the differences between designed calibration and validation period were within a range of 18–36 % for precipitation (Fig. 4, left) and 0.86–1.75 °C for temperature (Fig. 4, right). These values represent a reasonable climate change signal that is likely to occur within this century (IPCC, 2007).

The assessments of precipitation and temperature were done independently from each other. Note that the years in the two periods were not consecutive and that the periods consisted of different years for the tests of precipitation and temperature-bias correction methods.

All bias corrections were first calibrated based on the first subset of years and then evaluated for the second subset of years. In this way, the performance of the bias correction methods could be tested when applied to a period with different conditions than those during calibration.

HESSD
9, 12765–12795, 2012

2.5 Evaluation of bias correction methods

The performance of each bias correction method was analyzed by using the following procedure: we were first interested in the bias correction methods' ability to reproduce annual mean values under changed conditions. Thus, we calculated the mean values for all 15 yr for each RCM simulation and each bias correction method during the calibration and the validation period for precipitation (Fig. 5) and temperature (Fig. 6). The separated observed time series served as a baseline to standardize the values. Deviations from observations are given in percentages for precipitation (Fig. 5) and degrees Celsius for temperature (Fig. 6). Thereafter, the relative errors of precipitation respectively temperature were summarized in box plots for each catchment (Figs. 7 and 8). In addition to errors of the entire 15-yr periods (Figs. 7 and 8, left), we also evaluated the relative error of the three driest (Fig. 7, center) and three wettest years (Fig. 7, right) for precipitation as well as the three coldest (Fig. 8, center) and three warmest years (Fig. 8, right) for temperature. In the last step, several statistical measures were computed. For precipitation (Fig. 9), the statistical analysis was based on the mean (μ), 90th percentiles (X_{90}), standard deviation (σ), annual maximum of consecutive 5-day precipitation ($P_{5\,\text{max}}$), probability of wet days (Pr_{wet}) and intensity of wet days (i_{wet}). For temperature (Fig. 10), we calculated the mean (μ), 10th and 90th percentiles (X_{10}, X_{90}) and standard deviation (σ).

3 Results

3.1 Relative error analysis

During the designed calibration years, all precipitation-bias correction methods resulted in good estimates of annual values (Fig. 5) and were able to improve raw RCM simulations considerably. Evaluation against observed values for the validation period, however, showed a larger spread and a clear bias for some of the methods. Considering

HESSD
9, 12765–12795, 2012

the wetter validation period, single bias-corrected RCM series had large errors up to 50 % and were generally not able to reproduce observations well. However, the observations were usually located within the projected RCM range and also acceptably close to the ensemble mean (Fig. 5). For all bias correction methods and all catchments, the relative error of annual mean values was much larger for the validation period than the calibration period (Fig. 7). Not only the deviation of the ensemble median increased for the validation period, but also the variability became much larger. Considering annual precipitation of all 15 yr (Fig. 7, left), it is difficult to state which bias correction methods had higher or lower correction skills. However, looking only at the three driest years (Fig. 7, center), the delta-change approach performed worse than the other methods for all catchments except Storbäcken. For the three wettest years (Fig. 7, right), all bias correction methods showed much larger relative errors and also an increased variability range.

All temperature-bias correction methods resulted in good estimates of annual values during the designed calibration period (Fig. 6) and brought the raw RCM ensemble mean much closer to observations. However, during the validation period, bias-corrected annual mean values showed a larger spread and a clear bias. Observations were located within the bias-corrected RCM range only for the distribution mapping, whereas all other methods showed a poor performance in reproducing observations. Considering annual mean temperature of all 15 yr (Fig. 8, left), there was no clear difference between the bias correction methods: they all tended to underestimate temperature values with median values and variability ranges being relatively similar. For the three coldest years (Fig. 8, center), the delta-change approach underestimated values more than other methods for all catchments except for the warmest and southernmost Rönne Å catchment. This pattern could also be seen for the three warmest years (Fig. 8, right): the delta-change approach resulted in the largest errors for all five catchments, with relatively little variability among the different RCMs. For both the three coldest and three warmest years of the validation years, the other methods performed fairly similar with the distribution mapping having less error.

HESSD
9, 12765–12795, 2012

Bias correction for non-stationary conditions

C. Teutschbein and J. Seibert

3.2 Statistical measures

The visual analysis of precipitation statistics (Fig. 9) reveals clear differences between all bias correction methods. In general, raw (uncorrected) RCM-simulated precipitation had a wide spread and deviated considerably from observations with 80 % of the data having a relative error of −18 to +34 % (Fig. 9a). Other methods such as Linear Scaling, LOCI and Power Transformation showed also large spreads, but the ensemble median was closer to observations than it was for uncorrected RCM simulations. Only Distribution Mapping (−9 to +9 %) and the Delta-Change approach (−10 to +16 %) showed less variability. However, the visual representation of individual statistical measures for each catchment (Fig. 9b) demonstrates that each bias correction method has certain advantages and disadvantages. For the mean and the 90th percentile, the differences between the bias correction methods were less pronounced than for other statistical measures. Considering standard deviation (σ) and maximum 5-day precipitation ($P_{5\,\mathrm{max}}$), distribution mapping clearly outperformed all other bias correction methods. The probability of wet years ($\mathrm{Pr}_{\mathrm{wet}}$) was satisfactorily reproduced by LOCI, distribution mapping and the delta-change approach. However, we suspect that the delta-change approach only performed well due to the fact that $\mathrm{Pr}_{\mathrm{wet}}$ of calibration and validation period were not considerably different which is a flaw in our analysis. The intensity of wet days (i_{wet}) was best reproduced by LOCI and distribution mapping. Overall, distribution mapping had the highest correction skills, although it is not always performing well. As an example, all bias correction methods including distribution mapping had problems to reproduce $P_{5\,\mathrm{max}}$ for the two northernmost catchments.

The star plots of temperature statistics (Fig. 10) also identify certain differences between the bias correction methods, though less pronounced than for precipitation. All methods were able to correct the mean (μ), 90th percentile (X_{90}) and standard deviation (σ) of raw RCM temperature. The most distinct differences occurred for the 10th percentile (X_{10}) where all methods tended to underestimation. However, distribution

HESSD
9, 12765–12795, 2012

mapping performed best with the least error whereas the delta-change approach deviated most from observations.

4 Discussion and conclusions

Uncorrected RCM simulations are a source of large uncertainties in climate-change impact studies (Teutschbein and Seibert, 2010, 2012). Therefore, there is a need for bias correction procedures to ensure that RCM biases do not hamper subsequent impact simulations. In this study, we do not try to answer the "main question [...], whether and when the application of bias correction methods [...] is justified or not" (Ehret et al., 2012). One needs to be aware that there are several problematic aspects related to bias correction (Ehret et al., 2012), but that there are no obvious alternatives to this post-processing of RCM data as of yet. Potential alternatives include ensemble projections and improved climate models, e.g. enhanced process descriptions and increased spatial resolutions (Ehret et al., 2012; Muerth et al., 2012; Teutschbein and Seibert, 2010, 2011, 2012), but especially the latter approach is not operational in the near future.

Teutschbein and Seibert (2012) demonstrated that most applied bias correction approaches are able to improve raw RCM data to some extent, but that there are considerable differences in the quality of adjusted RCM temperature and precipitation for current climate conditions. However, bias correction procedures are seldom evaluated for the case which is most interesting for climate-change impact studies, namely their performance for changed conditions. While it is not possible to directly test their performance for the future, we demonstrated in this paper how differential split-sample testing (DSST) can be used to analyze the transferability of bias correction approaches to different climate conditions. Using DSST allows identifying clear differences in reproducing conditions similar to and conditions different from those that the bias correction approaches were calibrated to. These differences are an indicator for improper algorithm and parameter transfers.

HESSD
9, 12765–12795, 2012

By using the coldest/driest and warmest/wettest years for separation of the periods in combination with analyzing the most extreme years, we certainly pushed the bias correction methods to their limits. This was done on purpose, because we believe that reliable simulations of the more extreme years are essential for certain impact assessments, such as drought and flood modeling under future climate conditions. To test bias correction transferability on the conditions of a less extreme climate-change signal, it is also possible to use more moderate extrapolations by applying, for instance, the generalized split-sample test (GSST) as proposed by Coron et al. (2012).

The delta-change and the linear-scaling approach are the two most common transfer methods and have been widely used (Gellens and Roulin, 1998; Graham et al., 2007a,b; Lettenmaier et al., 1999; Middelkoop et al., 2001; Moore et al., 2008; Shabalova et al., 2003), because they are straightforward and easy to implement due to their simplicity. Yet, our validation of the correction approaches based on data for conditions outside those used during calibration shows that these two methods result in particularly large deviations, small variability of the simulations based on the different RCMs (i.e. overconfident uncertainty ranges) and least reliable future projections especially for the most extreme years (driest/wettest and coldest/warmest). These findings remain to be confirmed for other catchments and other geographic regions, but based on the findings in this study we question the use of the delta-change or the linear-scaling approach to bias-correct RCM scenarios of future conditions for climate-change impact studies.

The choice of bias correction algorithm plays a large role in assessing hydrological change. For current conditions, we could easily limit this choice to the one that performed best. For simulations of future climate this is more difficult and the fundamental question is how transferable the different methods are. The differential split-sample test proved to be a suitable approach to evaluate this and was also able to confirm a better performance of high-skill methods such as distribution mapping.

Acknowledgements. This study was funded by FORMAS, the Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning (Grant No. 2007-1433). The authors thank the Swedish Meteorological and Hydrological Institute (SMHI) for providing observed meteorological data. The ENSEMBLES data used in this work was funded by the EU FP6 Integrated Project ENSEMBLES (Contract number 505539) whose support is gratefully acknowledged. We also thank Tracy Ewen for helping to clarify the text.

References

Bennett, J. C., Ling, F. L. N., Graham, B., Grose, M. R., Corney, S. P., White, C. J., Holz, G. K., Post, D. A., Gaynor, S. M., and Bindoff, N. L.: Climate futures for Tasmania: water and catchments technical report, Antarctic Climate & Ecosystems Cooperative Research Centre, Hobart, Tasmania, 2010.

Block, P. J., Souza Filho, F. A., Sun, L., and Kwon, H. H.: A streamflow forecasting framework using multiple climate and hydrological models, J. Am. Water Resour. Assoc., 45, 828–843, doi:10.1111/j.1752-1688.2009.00327.x, 2009.

Boe, J., Terray, L., Habets, F., and Martin, E.: Statistical and dynamical downscaling of the Seine basin climate for hydro-meteorological studies, Int. J. Climatol., 27, 1643–1655, doi:10.1002/joc.1602, 2007.

Chen, J., Brissette, F. P., and Leconte, R.: Uncertainty of downscaling method in quantifying the impact of climate change on hydrology, J. Hydrol., 401, 190–202, doi:10.1016/j.jhydrol.2011.02.020, 2011.

Christensen, J. H., Boberg, F., Christensen, O. B., 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.

Coron, L., Andréassian, V., Perrin, C., Lerat, J., Vaze, J., Bourqui, M., and Hendrickx, F.: Crash testing hydrological models in contrasted climate conditions: an experiment on 216 Australian catchments, Water Resour. Res., 48, W05552, doi:10.1029/2011WR011721, 2012.

Déqué, M., Rowell, D. P., Lüthi, D., Giorgi, F., Christensen, J. H., Rockel, B., Jacob, D., Kjellström, E., De Castro, M., and van den Hurk, B.: An intercomparison of regional climate simulations for Europe: assessing uncertainties in model projections, Climatic Change, 81, 53–70, doi:10.1007/s10584-006-9228-x, 2007.

Ehret, U., Zehe, E., Wulfmeyer, V., Warrach-Sagi, K., and Liebert, J.: HESS Opinions “Should we apply bias correction to global and regional climate model data?”, Hydrol. Earth Syst. Sci., 16, 3391–3404, doi:10.5194/hess-16-3391-2012, 2012.

Fowler, H. J., 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.

Gellens, D. and Roulin, E.: Streamflow response of Belgian catchments to IPCC climate change scenarios, J. Hydrol., 210, 242–258, doi:10.1016/S0022-1694(98)00192-9, 1998.

Giorgi, F.: Regional climate modeling: status and perspectives, J. Phys. IV, 139, 101–118, doi:10.1051/jp4:2006139008, 2006.

Graham, L., Andréasson, J., and Carlsson, B.: Assessing climate change impacts on hydrology from an ensemble of regional climate models, model scales and linking methods – a case study on the Lule River basin, Climatic Change, 81, 293–307, doi:10.1007/s10584-006-9215-2, 2007a.

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, 2007b.

Grotch, S. L. and MacCracken, M. C.: The use of general circulation models to predict regional climatic change, J. Climate, 4, 286–303, doi:10.1175/1520-0442(1991)004<0286:TUOGCM>2.0.CO;2, 1991.

Gudmundsson, L., Bremnes, J. B., Haugen, J. E., and Engen-Skaugen, T.: Technical Note: Downscaling RCM precipitation to the station scale using statistical transformations – a comparison of methods, Hydrol. Earth Syst. Sci., 16, 3383–3390, doi:10.5194/hess-16-3383-2012, 2012.

Ines, A. V. M. and Hansen, J. W.: Bias correction of daily GCM rainfall for crop simulation studies, Agr. Forest Meteorol., 138, 44–53, doi:10.1016/j.agrformet.2006.03.009, 2006.

IPCC: Climate Change 2007: The Physical Science Basis, Cambridge University Press, Cambridge, UK, and New York, USA, 2007.

Johansson, B.: Estimation of areal precipitation for hydrological modelling in Sweden, Dissertation, Earth Science Centre, Göteborg University, Report A76, Göteborg, Sweden, 2002.

Johnson, F. and Sharma, A.: Accounting for interannual variability: a comparison of options for water resources climate change impact assessments, Water Resour. Res., 47, W04508, doi:10.1029/2010WR009272, 2011.

HESSD

9, 12765–12795, 2012

Bias correction for non-stationary conditions

C. Teutschbein and J. Seibert

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

◀| |▶

◀ ▶

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Klemeš, V.: Operational testing of hydrological simulation models/Vérification, en conditions réelles, des modèles de simulation hydrologique, Hydrolog. Sci. J., 31, 13–24, doi:10.1080/02626668609491024, 1986.

Leander, R. and Buishand, T. A.: 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. A., van den Hurk, B. J. J. M., and de Wit, M. J. 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.

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.

Lettenmaier, D. P., Wood, A. W., Palmer, R. N., Wood, E. F., and Stakhiv, E. Z.: Water resources implications of global warming: a US regional perspective, Climatic Change, 43, 537–579, doi:10.1023/A:1005448007910, 1999.

Li, C. Z., Zhang, L., Wang, H., Zhang, Y. Q., Yu, F. L., and Yan, D. H.: The transferability of hydrological models under nonstationary climatic conditions, Hydrol. Earth Syst. Sci., 16, 1239–1254, doi:10.5194/hess-16-1239-2012, 2012.

Middelkoop, H., Daamen, K., Gellens, D., Grabs, W., Kwadijk, J. C. J., Lang, H., Parmet, B. W. A. H., Schädler, 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.

Moore, K., Pierson, D., Pettersson, K., Schneiderman, E., and Samuelsson, P.: Effects of warmer world scenarios on hydrologic inputs to Lake Mälaren, Sweden and implications for nutrient loads, Hydrobiologia, 599, 191–199, doi:10.1007/s10750-007-9197-8, 2008.

Mpelasoka, F. S. and Chiew, F. H.: Influence of rainfall scenario construction methods on runoff projections, J. Hydrometeorol., 10, 1168–1183, doi:10.1175/2009JHM1045.1, 2009.

Muerth, M. J., Gauvin St-Denis, B., Ricard, S., Velázquez, J. A., Schmid, J., Minville, M., Caya, D., Chaumont, D., Ludwig, R., and Turcotte, R.: On the need for bias correction in regional climate scenarios to assess climate change impacts on river runoff, Hydrol. Earth Syst. Sci. Discuss., 9, 10205–10243, doi:10.5194/hessd-9-10205-2012, 2012.

Piani, C., Haerter, J. O., 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, 2010.

Rojas, R., Feyen, L., Dosio, A., and Bavera, D.: Improving pan-European hydrological simulation of extreme events through statistical bias correction of RCM-driven climate simulations, Hydrol. Earth Syst. Sci., 15, 2599–2620, doi:10.5194/hess-15-2599-2011, 2011.
Salathé Jr., E. P.: Comparison of various precipitation downscaling methods for the simulation of streamflow in a rainshadow river basin, Int. J. Climatol., 23, 887–901, doi:10.1002/joc.922, 2003.
Schmidli, J., Frei, C., and Vidale, P. L.: Downscaling from GCM precipitation: a benchmark for dynamical and statistical downscaling methods, Int. J. Climatol., 26, 679–689, doi:10.1002/joc.1287, 2006.
Seiller, G., Anctil, F., and Perrin, C.: Multimodel evaluation of twenty lumped hydrological models under contrasted climate conditions, Hydrol. Earth Syst. Sci., 16, 1171–1189, doi:10.5194/hess-16-1171-2012, 2012.
Sennikovs, J. and Bethers, U.: Statistical downscaling method of regional climate model results for hydrological modelling, in: 18th World IMACS Congress and MODSIM09 International Congress on Modelling and Simulation, edited by: Anderssen, R. S., Braddock, R. D., and Newham, L. T. H., Modelling and Simulation Society of Australia and New Zealand and International Association for Mathematics and Computers in Simulation, Date and Place of Congress: 13–17 July, 2009, Cairns, Australia, 3962–3968, 2009.
Shabalova, M. V., van Deursen, W. P., and Buishand, T. A.: 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.
Sun, F., Roderick, M. L., Lim, W. H., and Farquhar, G. D.: Hydroclimatic projections for the Murray-Darling Basin based on an ensemble derived from Intergovernmental Panel on Climate Change AR4 climate models, Water Resour. Res., 47, W00G02, doi:10.1029/2010WR009829, 2011.
Terink, W., Hurkmans, R. T. W. L., Torfs, P. J. J. F., and Uijlenhoet, R.: Bias correction of temperature and precipitation data for regional climate model application to the Rhine basin, Hydrol. Earth Syst. Sci. Discuss., 6, 5377–5413, doi:10.5194/hessd-6-5377-2009, 2009.
Terink, W., Hurkmans, R. T. W. L., Torfs, P. J. J. F., and Uijlenhoet, R.: Evaluation of a bias correction method applied to downscaled precipitation and temperature reanalysis data for the Rhine basin, Hydrol. Earth Syst. Sci., 14, 687–703, doi:10.5194/hess-14-687-2010, 2010.

Teutschbein, C. and Seibert, J.: Regional climate models for hydrological impact studies at the catchment scale: a review of recent modeling strategies, Geography Compass, 4, 834–860, doi:10.1111/j.1749-8198.2010.00357.x, 2010.

Teutschbein, C. and Seibert, J.: Bias correction of regional climate model simulations for hydrological climate-change impact studies: review and evaluation of different methods, J. Hydrol., 456–457, 12–29, doi:10.1016/j.jhydrol.2012.05.052, 2012.

Teutschbein, C., Wetterhall, F., and Seibert, J.: Evaluation of different downscaling techniques for hydrological climate-change impact studies at the catchment scale, Clim. Dynam., 37, 2087–2105, doi:10.1007/s00382-010-0979-8, 2011.

Van der Linden, P. and Mitchell, J. F. B.: ENSEMBLES: Climate Change and its Impacts: Summary of research and results from the ENSEMBLES project, Met Office Hadley Centre, FitzRoy Road, Exeter EX1 3PB, UK, 2009.

Varis, O., Kajander, T., and Lemmelä, R.: Climate and water: from climate models to water resources management and vice versa, Climatic Change, 66, 321–344, doi:10.1023/B:CLIM.0000044622.42657.d4, 2004.

HESSD

9, 12765–12795, 2012

Bias correction for non-stationary conditions

C. Teutschbein and J. Seibert

Table 1. Main characteristics of the five Swedish study sites including area, total annual precipitation (P), mean annual temperature (T), climate zones and land-use properties.

Catchments	Area (km^2)	P ($mm\,yr^{-1}$)	T (°C)	Climate zone (Köppen-Geiger)	Land-use (%) Forest	Open land	Lakes/ wetlands	Residential
1 Tännfors	227	775	−0.5	Continental subarctic towards polar tundra	32	60	8	0
2 Storbäcken	150	617	2.1	Continental subarctic	79	9	2	0
3 Vattholmaån	293	633	5.2	Warm summer continental	81	7	10	2
4 Brusaån	240	632	5.7	Warm summer continental towards maritime temperate	83	12	3	2
5 Rönne Å	147	786	7.3	Maritime temperate	23	46	27	4

HESSD
9, 12765–12795, 2012

Table 2. RCM experiments from the ENSEMBLES EU project used in this study.

Institute	Model	Resolution	Driving GCM
C4I	RCA3	25 km	HadCM3Q16
CNRM	Aladin	25 km	ARPEGE
DMI	HIRHAM	25 km	ARPEGE
ETHZ	CLM	25 km	HadCM3Q0
HC	HadRM3Q0	25 km	HadCM3Q0
ICTP	RegCM	25 km	ECHAM5-r3
KNMI	RACMO	25 km	ECHAM5-r3
met.no	HIRHAM	25 km	HadCM3Q0
MPI	REMO	25 km	ECHAM5-r3
SMHI	RCA	25 km	HadCM3Q3
UCLM	PROMES	25 km	HadCM3Q0

HESSD
9, 12765–12795, 2012

Table 3. Overview of methods used to correct RCM-simulated precipitation (P) and/or temperature (T) data, for more information on the methods see Teutschbein and Seibert (2012).

Method	Variable	Short Description	Advantages (+) and Disadvantages (–)	References
Raw RCM Output Data	P, T	– RCM-simulated time series are used directly without any bias correction	+ simplest way to use RCM data – systematic model errors are ignored – can cause substantial errors in impact studies	
Precipitation Threshold	P	– an RCM-specific threshold is calibrated such that the number of RCM-simulated days exceeding this threshold matches the number of observed days with precipitation larger than 0 mm – rarely used as a "stand-alone" method but often combined with other correction procedures	+ wet-day frequencies are corrected – the mean is not adjusted	(Schmidli et al., 2006)
Delta-Change Correction	P, T	– RCM-simulated future change signals (anomalies) are superimposed upon observational time series – usually done with a multiplicative correction for precipitation and an additive correction for temperature	+ observations are used as a basis, which makes it a robust method – potential future changes in climate dynamics are not accounted for – all events change by the same amount	(Gellens and Roulin, 1998) (Graham et al., 2007a, 2007b) (Johnson and Sharma, 2011) (Lettenmaier et al., 1999) (Mpelasoka and Chiew, 2009) (Middelkoop et al., 2001) (Moore et al., 2008) (Shabalova et al., 2003)
Linear Scaling	P, T	– adjusts RCM time series with correction values based on the relationship between long-term monthly mean observed and RCM control run values – precipitation is typically corrected with a factor and temperature with an additive term	+ allows to correct the mean + variability of corrected data is more consistent with original RCM data – wet-day frequencies and intensities are not corrected – all events are adjusted with the same correction factor	(Lenderink et al., 2007)
Local Intensity Scaling	P	– combines a precipitation threshold with linear scaling (both described above)	+ improves linear scaling + allows to correct the mean, wet-day frequencies and intensities	(Schmidli et al., 2006)
Power Transformation	P	– is a non-linear correction in an exponential form ($a \cdot P^b$) that combines the correction of the coefficient of variation (CV) with a linear scaling	+ allows to adjust mean and standard deviation (variance) ± adjusts wet-day frequencies and intensities only to some extend	(Leander and Buishand, 2007) (Leander et al., 2008)
Variance Scaling	T	– combines standard linear scaling with a scaling based on standard deviations	+ allows to adjust mean and standard deviation (variance)	(Chen et al., 2011)
Distribution mapping	P, T	– matches the distribution functions of observations and RCM-simulated climate values – a precipitation threshold can be introduced to avoid substantial distortion of the distribution caused by too many drizzle days – also known as "quantile-quantile mapping", "probability mapping", "statistical downscaling" or "histogram equalization".	+ allows to adjust mean, standard deviation (variance), wet-day frequencies and intensities	(Block et al., 2009) (Boe et al., 2007) (Déqué et al., 2007) (Ines and Hansen, 2006) (Johnson and Sharma, 2011) (Piani et al., 2010) (Rojas et al., 2011) (Sennikovs and Bethers, 2009) (Sun et al., 2011)

HESSD
9, 12765–12795, 2012

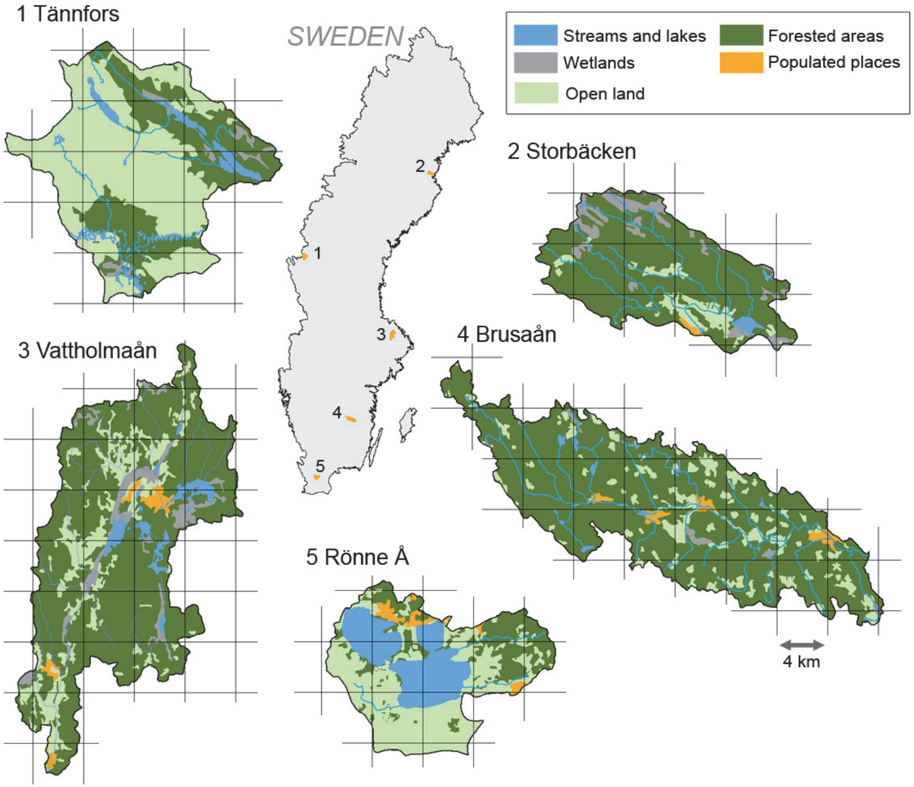

Fig. 1. Map showing locations of the Swedish study sites and the spatially interpolated 4 × 4 km national grid of observed precipitation and temperature. Catchments: (1) Tännfors, (2) Storbäcken, (3) Vattholmaån, (4) Brusaån and (5) Rönne Å.

HESSD
9, 12765–12795, 2012

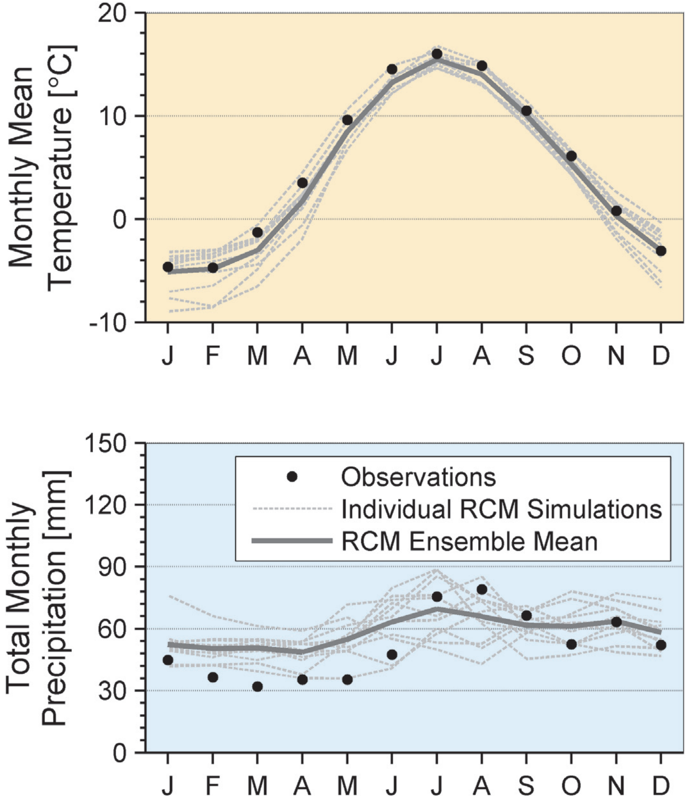

Fig. 2. Monthly mean temperature and total monthly precipitation for the period 1961–1990 as simulated by individual RCMs (gray dashed lines) for the Vattholmaån catchment in South-Eastern Sweden. Observations (black circles) and the RCM-ensemble means (gray continuous line) are displayed as well.

HESSD
9, 12765–12795, 2012

Bias correction for non-stationary conditions

C. Teutschbein and J. Seibert

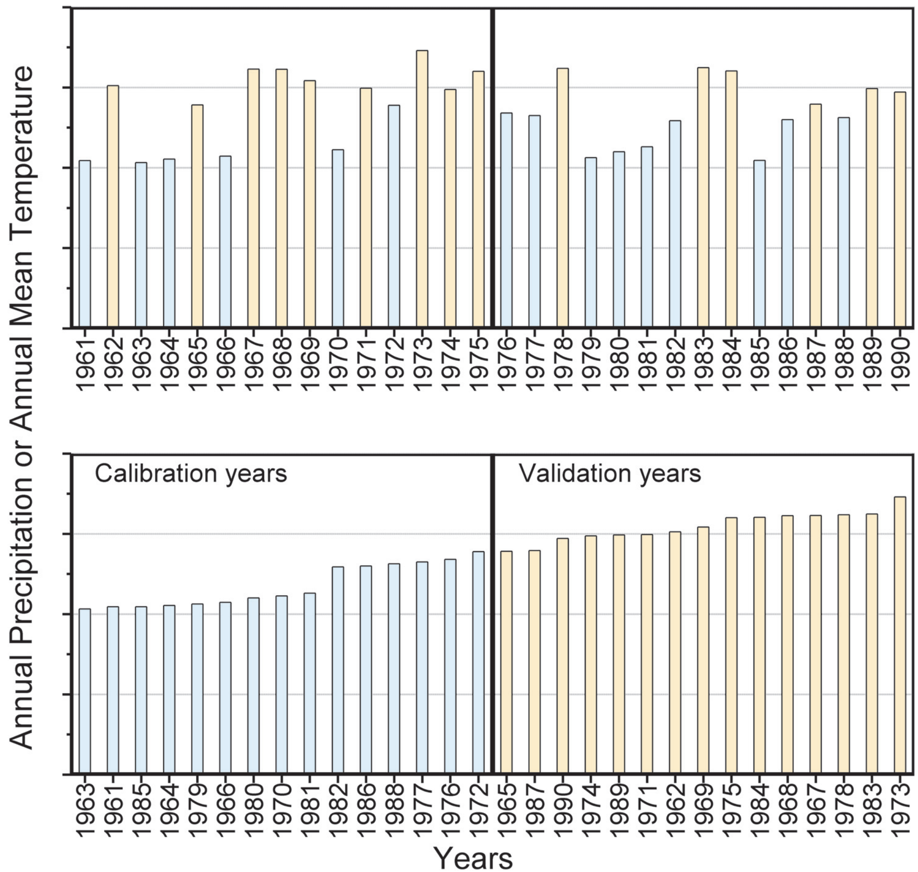

Fig. 3. Exemplary procedure of the differential split-sample test. First, the natural order of annual values (top) is sorted ascending (bottom). The lower-value years are then used for calibration, the higher-value years for validation. This test was done independently for precipitation and temperature.

HESSD
9, 12765–12795, 2012

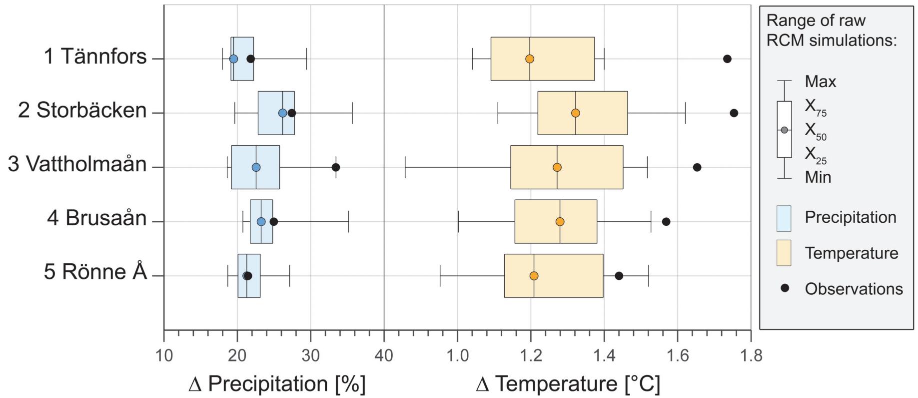

Fig. 4. Differences between mean values of designed calibration and validation period for precipitation (left) and temperature (right) shown for raw RCM simulations (colored boxes) and observations (black circles).

HESSD
9, 12765–12795, 2012

C. Teutschbein and J. Seibert

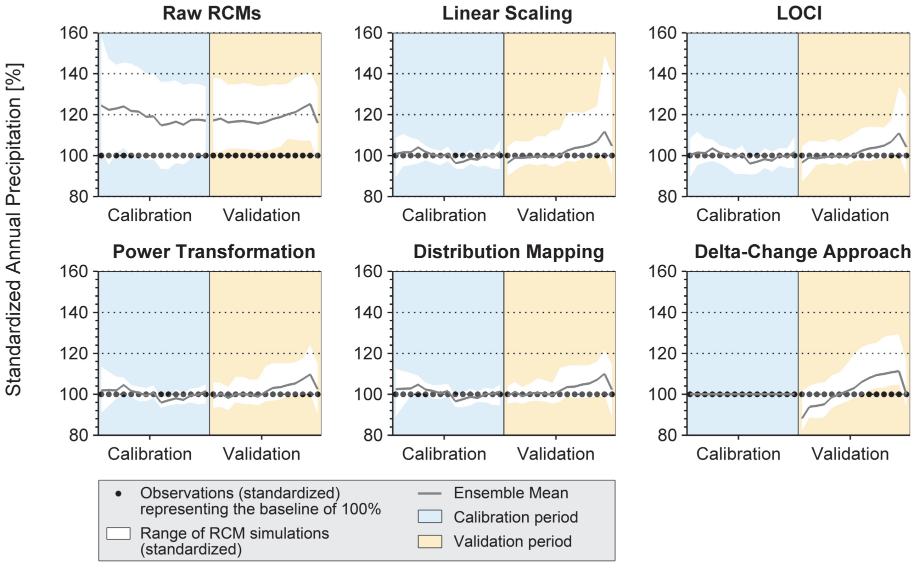

Fig. 5. Performance of different precipitation-bias corrections (white shaded area) and their ensemble mean (dark gray curve) for designed calibration (dry years, blue) and validation period (wet years, orange) compared to observations (black circles) on an annual basis in the Brusafors river basin (#4) in Southern Sweden. Precipitation in ascending order was standardized based on observations.

HESSD
9, 12765–12795, 2012

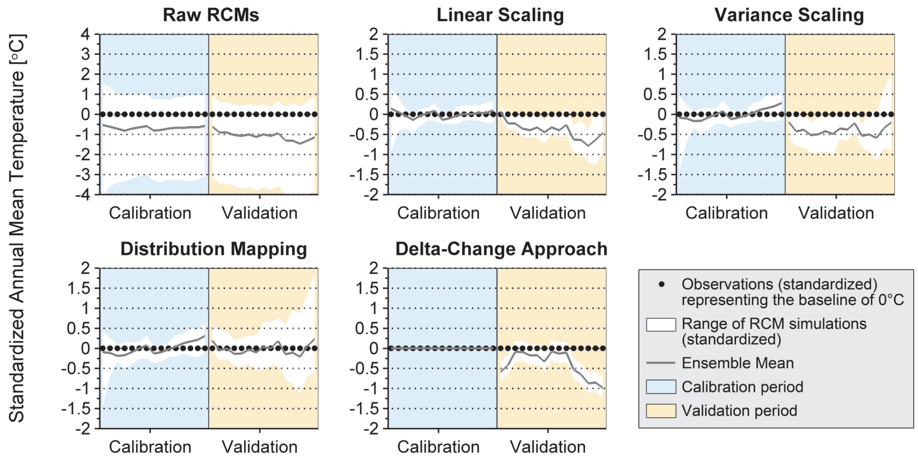

Fig. 6. Performance of different temperature-bias corrections (white shaded area) and their ensemble mean (dark gray curve) for designed calibration (cold years, blue) and validation period (warm years, orange) compared to observations (black circles) in the Storbäcken river basin (#2) in Northern Sweden. Temperature in ascending order was standardized based on observations. Please note the different scale of the upper left subplot (raw RCM simulations).

HESSD
9, 12765–12795, 2012

Bias correction for non-stationary conditions

C. Teutschbein and J. Seibert

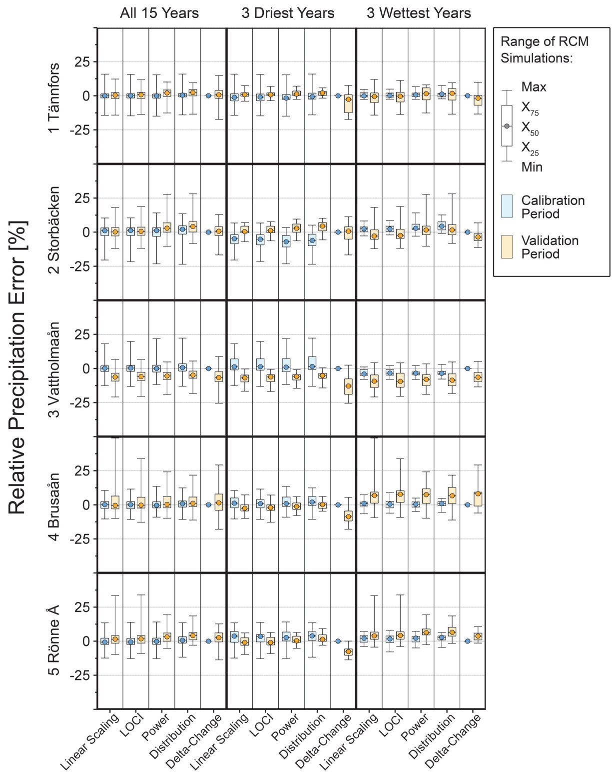

Fig. 7. Relative deviation of differently bias-corrected RCM-simulated precipitation from observations during calibration (blue boxes) and validation period (orange boxes). The boxes comprise uncertainty of different RCM simulations and the respective 15 yr of each period. Relative error for all 5 catchments is shown for the entire period of 15 yr (left panel) and the extremes, i.e. the three driest years (central panel) and the three wettest years (right panel).

HESSD
9, 12765–12795, 2012

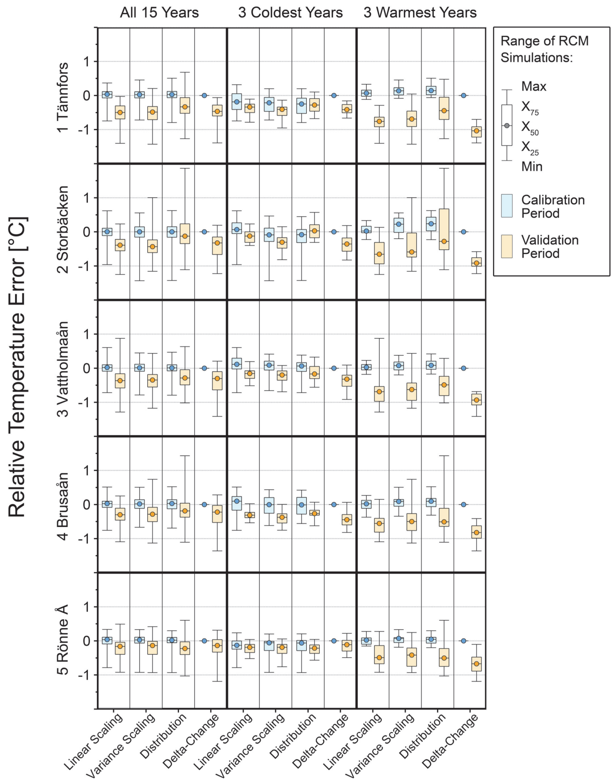

Fig. 8. Relative deviation of differently bias-corrected RCM-simulated temperature from observations during calibration (blue boxes) and validation period (orange boxes). The boxes comprise uncertainty of different RCM simulations and the respective 15 yr of each period. Relative error for all 5 catchments is shown for the entire period of 15 yr (left panel) and the extremes, i.e. the three coldest years (central panel) and the three warmest years (right panel).

HESSD
9, 12765–12795, 2012

Bias correction for non-stationary conditions

C. Teutschbein and J. Seibert

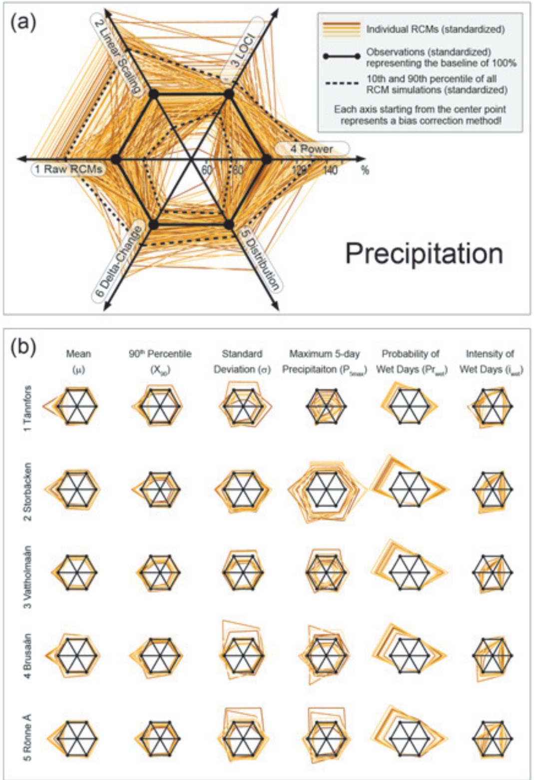

Fig. 9. Statistical performance comparison of different bias correction methods (different axes of the star plots) for precipitation of the validation period (wetter years) as simulated by different RCMs (orange lines). The bias correction methods were parameterized based on the calibration period (drier years). The results were standardized based on observations (black circles) for the validation period. Subplot **(a)** summarizes all graphics of subplot **(b)** and gives an explanation on how to interpret the graphics. Subplot **(b)** breaks down subplot **(a)** into several catchment locations and statistical measures.

HESSD
9, 12765–12795, 2012

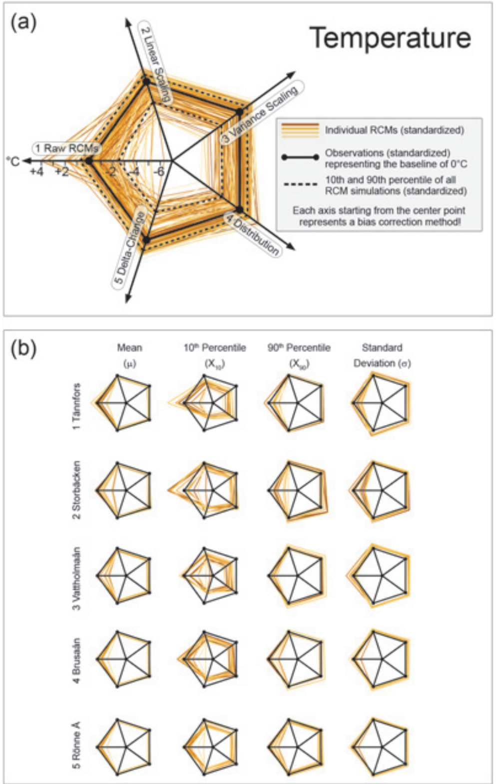

Fig. 10. Statistical performance comparison of different bias correction methods (different axes of the star plots) for temperature of the validation period (warmer years) as simulated by different RCMs (orange lines). The bias corrections methods were parameterized based on the calibration period (colder years). The results were standardized based on observations (black circles) for the validation period. Subplot **(a)** summarizes all graphics of subplot **(b)** and gives an explanation on how to interpret the graphics. Subplot **(b)** breaks down subplot **(a)** into several catchment locations and statistical measures.

HESSD
9, 12765–12795, 2012

Bias correction for non-stationary conditions

C. Teutschbein and J. Seibert