

# Chapter 9

## A New Distribution Mapping Technique for Climate Model Bias Correction

Seth McGinnis, Doug Nychka, and Linda O. Mearns

**Abstract** We evaluate the performance of different distribution mapping techniques for bias correction of climate model output by operating on synthetic data and comparing the results to an “oracle” correction based on perfect knowledge of the generating distributions. We find results consistent across six different metrics of performance. Techniques based on fitting a distribution perform best on data from normal and gamma distributions, but are at a significant disadvantage when the data does not come from a known parametric distribution. The technique with the best overall performance is a novel nonparametric technique, kernel density distribution mapping (KDDM).

**Keywords** KDDM • Nonparametric distribution • Oracle evaluation • Quantile mapping • Transfer function

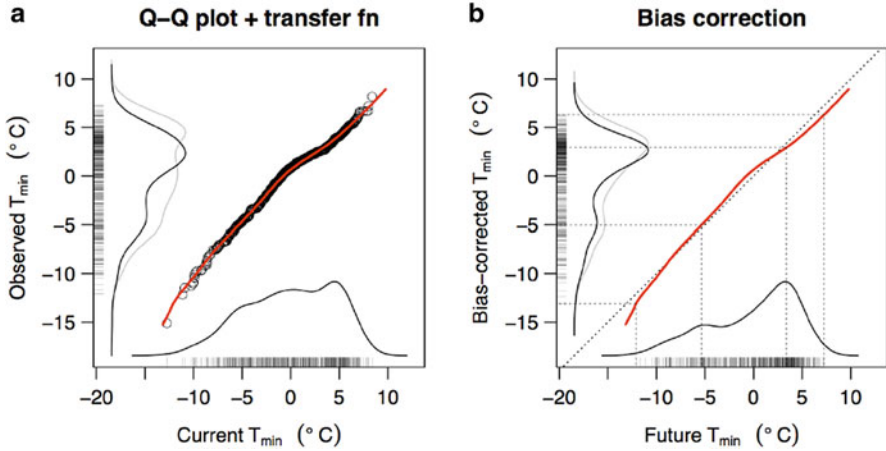
### 9.1 Introduction

Climate modeling is a valuable tool for exploring the potential future impacts of climate change whose use is often hindered by bias in the model output. Correcting this bias dramatically increases its usability, especially for impacts users. Teutschbein and Seibert (2012) tested a variety of bias-correction methods and found that the best overall performer was distribution mapping.

Distribution mapping adjusts the individual values of the model output such that their statistical distribution matches that of the observed data. This is accomplished by the method of Panofsky and Brier (1968), which constructs a transfer function

---

S. McGinnis (✉) • D. Nychka • L.O. Mearns  
National Center for Atmospheric Research, Boulder, CO, USA  
e-mail: [mcginnis@ucar.edu](mailto:mcginnis@ucar.edu)



**Fig. 9.1** Bias correction via distribution mapping. **(a)** Q-Q plot of observed versus modeled data for minimum daily temperatures with transfer function overlaid. **(b)** Plot of the transfer function showing its use in bias correction of modeled future data. *Dashed lines* illustrate how example values are bias-corrected by mapping via the transfer function. Probability density curves and rug plots of individual data values for each dataset are plotted along the edges of each figure

that transforms modeled values into probabilities via the CDF (cumulative distribution function) of the model distribution and then transforms them back into data values using the inverse CDF (or quantile function) of the observational distribution:

$$x_{\text{corrected}} = \text{transfer}(x_{\text{raw}}) = \text{CDF}_{\text{observed}}^{-1}(\text{CDF}_{\text{model}}(x_{\text{raw}})). \quad (9.1)$$

The transfer function is constructed using observed data and model output from the same current period and then applied to model output from a future period. This approach assumes that model bias is stationary and does not change significantly over time. This process is illustrated in Fig. 9.1: the first panel shows a transfer function overlaid on a quantile-quantile (Q-Q) plot of the data from which it is constructed, and the second panel shows how the future-period data is bias-corrected by mapping through the transfer function. This figure is discussed in further detail at the end of Sect. 9.2.

There are a number of different bias-correction techniques that use this distribution mapping approach; they differ primarily in how they construct the transfer function. They are referred to in the literature, often inconsistently, by a variety of different names, including among others “quantile mapping,” “probability mapping,” and “CDF matching.” In this paper, we test six such techniques, which are described in the section following, and include a novel technique based on kernel density estimates of the underlying probability distribution function (PDF). We evaluate the techniques using an “oracle” methodology of bias-correcting synthetic data for which a known correct answer exists for comparison.

## 9.2 Distribution Mapping Techniques

The following techniques encompass the different approaches to distribution mapping that we found in our survey of the literature. In an effort to clear up the problem of inconsistent nomenclature, we name them here according to their distinctive methodology, rather than by the names used in the referenced papers.

**Probability Mapping (PMAP)** Probability mapping fits parametric distributions to the current and observed datasets and forms a transfer function by composing the corresponding fitted analytic CDF and quantile functions (Ines and Hansen 2006; Piani et al. 2010; Haerter et al. 2011). For example, using the normal distribution:

$$\mathbf{x}_{bc} = Q_{\text{norm}}(P_{\text{norm}}(\mathbf{x}_{\text{fut}}, \mu_{\text{cur}}, \sigma_{\text{cur}}), \mu_{\text{obs}}, \sigma_{\text{obs}}), \quad (9.2)$$

where  $Q_{\text{norm}}$  and  $P_{\text{norm}}$  are the quantile and CDF functions of the normal distribution,  $\mu$  and  $\sigma$  are its parameters, and  $x$  is a data value, each belonging to the current, future, observed, or bias-corrected dataset, as indicated by the subscript.

The family of the distribution must be specified a priori. In this paper, we use a gamma distribution to fit data bounded at zero and a normal distribution to fit unbounded data, as would be typical practice in bias-correcting climate model output *en masse*. We tested several methods of fitting distributions and found no noteworthy differences in performance, so in this analysis we use the computationally simple method of moments for fitting.

**Empirical CDF Mapping (ECDF)** ECDF mapping creates a Q-Q map by sorting the observed and current datasets and mapping them against one another. It then forms a transfer function by linearly interpolating between the points of the mapping (Wood et al. 2004; Boé et al. 2007). Note that because it relies upon the Q-Q map, this technique requires the current and observed datasets to have equal numbers of points.

**Order Statistic Difference Correction (OSDC)** This method is uncommon, but is used in a few studies, and may be confused with ECDF mapping. OSDC sorts the observed and current datasets and differences them to produce a set of corrections to be applied to the future dataset (Iizumi et al. 2011). Mathematically, the bias correction is described thus:

$$\mathbf{x}_{bc}^{(i)} = \mathbf{x}_{\text{fut}}^{(i)} - (\mathbf{x}_{\text{cur}}^{(i)} - \mathbf{x}_{\text{obs}}^{(i)}), \quad (9.3)$$

where  $\mathbf{x}_{bc}^{(i)}$  denotes the  $i$ th largest value of the bias-corrected dataset. Note that this technique requires all datasets to have equal numbers of points.

**Quantile Mapping (QMAP)** Quantile mapping estimates a set of quantiles for the observed and current datasets and then forms a transfer function by interpolation between corresponding quantile values (Ashfaq et al. 2010; Johnson and Sharma

2011; Gudmundsson et al. 2012). In this study, we employ the `qmap` package (Gudmundsson 2014) for the statistical programming language R (R Core Team 2014) to perform quantile mapping, using empirical quantiles and spline interpolation, which a separate analysis showed to be the most effective options. The number of quantiles is a free parameter that must be specified; we test three cases, using “few” (5), “some” ( $N^{1/2} = 30$ ), and “many” ( $N/5 = 180$ ) quantiles.

**Asynchronous Regional Regression Modeling (ARRM)** ARRM constructs a transfer function based on a segmented linear regression of the Q-Q map (Stoner et al. 2012). As in ECDF mapping, it begins by sorting both datasets and mapping them against one another (which requires that they have equal number of points). It then finds six breakpoints between segments by applying linear regression over a moving window of fixed width to find points where the slope of the Q-Q map changes abruptly. Finally, it constructs the transfer function as a piecewise linear statistical model using these breakpoints as knots. The implementation of ARRM used here is based on the description in Stoner et al. (2012) and has some simplifications of various checks and corner cases that are needed for dealing with real-world data but do not apply to synthetic data. We use the R function `lm()` for the linear regressions and `lm()` with `ns()` to construct the transfer function.

**Kernel Density Distribution Mapping (KDDM)** is a novel technique described here for the first time. Conceptually, it is very similar to probability mapping, but instead of using fitted parametric distributions, it uses nonparametric estimates of the underlying probability density function (PDF). These estimates are created using kernel density estimation, a well-developed statistical technique that can be thought of as the smooth, non-discrete analog of a histogram. A kernel density estimate is constructed by summing copies of the kernel function (any symmetric, usually unimodal function that integrates to one) centered on each point in the dataset. Mathematically, the kernel density estimator  $\hat{f}(x)$  is

$$\hat{f}(x) = \frac{1}{n} \sum_{i=1}^n K_h(x - x_i), \quad (9.4)$$

where  $K_h$  is the kernel function scaled to bandwidth  $h$ . In this analysis, we use the default kernel (Gaussian) and bandwidth selection rule (Silverman’s rule of thumb) for R’s `density()` function (R Core Team 2014).

KDDM begins by estimating the PDFs for the current and observed datasets using kernel density estimation. The resulting nonparametric PDF estimates are then numerically integrated to approximate CDFs by evaluating them on a suitably fine grid, applying the trapezoidal rule, and linearly interpolating the results to produce a function. KDDM then forms a transfer function by composing the forward CDF for the current dataset and the inverse CDF for the observed dataset. Mathematically, defining  $\tilde{P}(x)$  as the approximate CDF,

$$\tilde{P}(x) = \int \hat{f}(x) dx, \quad (9.5)$$

and the KDDM bias correction is

$$x_{bc} = \tilde{P}_{obs}^{-1}(\tilde{P}_{cur}(x_{fut})). \quad (9.6)$$

This algorithm can be implemented very compactly in R, requiring only a dozen lines of code. It is also quite fast, requiring only twice as much computation time as the fastest methods and running 100 times faster than the slowest method.

Figure 9.1 demonstrates the application of the KDDM technique to bias-correct output from the North American Regional Climate Change Assessment Program (Mearns et al. 2007, 2009) using observations from the Maurer et al. (2002) dataset for a 2-week window in mid-October near Pineville, Missouri. The first panel shows a Q-Q plot, where the observations and current-period model output have been sorted and plotted against one another (small circles). The KDDM transfer function is overlaid, as are rug plots and PDF curves for each dataset. The second panel shows the bias correction of future-period model data by mapping through the transfer function. In both panels, the model PDF curve is mirrored in light gray on the y-axis to show the resulting change in the distribution. Before bias correction, we aggregated all three datasets across three decades (1970–2000 for the current and observed, 2040–2070 for the future) and removed the means.

### 9.3 Oracle Evaluation Methodology

To evaluate the techniques, we compare them to an ideal correction called the “oracle.” To create the oracle, we generate three sets of synthetic data to represent observed, modeled current, and modeled future data, using different parameters for each case. The differences between the synthetic current and future datasets correspond to climate change, and the differences between the synthetic observed and current datasets to model bias. Because we know the generating distribution and the exact parameter values used to generate these datasets, we can then construct a perfect transfer function using probability mapping. Applying this transfer function to the current dataset makes it statistically indistinguishable from the observed dataset; applying it to the future dataset generates the “oracle” dataset.

We then evaluate each technique by applying it to the future dataset and measuring the technique’s performance in terms of how far the bias-corrected result deviates from the perfect correction of the oracle. We perform this procedure using three different distributions, iterating over 1,000 realizations of the datasets each time. Each dataset contains 900 data points, which is the size of the dataset we would use when bias-correcting daily data month-by-month across a 30-year period, a common use case for working with regional climate model output.

The three distributions we use are the normal distribution, the gamma distribution, and a bimodal mixture of two normal distributions. We use the normal distribution to establish a baseline; its ideal transfer function is a straight line. We use the gamma distribution because precipitation has a gamma-like distribution. We

use a mixture distribution because similar distributions can be observed in real-world datasets that are often corrected under an assumption of normality, even though the actual distribution is more complex and may be impossible to fit. The observed data in Fig. 9.1 exhibits this kind of non-normal distribution.

For variables with an unbounded distribution, like temperature, it is necessary to remove the mean before bias correction, adjust it independently for climate change, and add it back in afterward, or else the transfer function will mix the climate change signal into the bias, producing an error component. For variables that are bounded at zero, like precipitation, the mean should not be removed, but it may be necessary to stabilize the variance by applying a power transform. We use a fourth-root transformation for the gamma dataset, following Wilby et al. (2014).

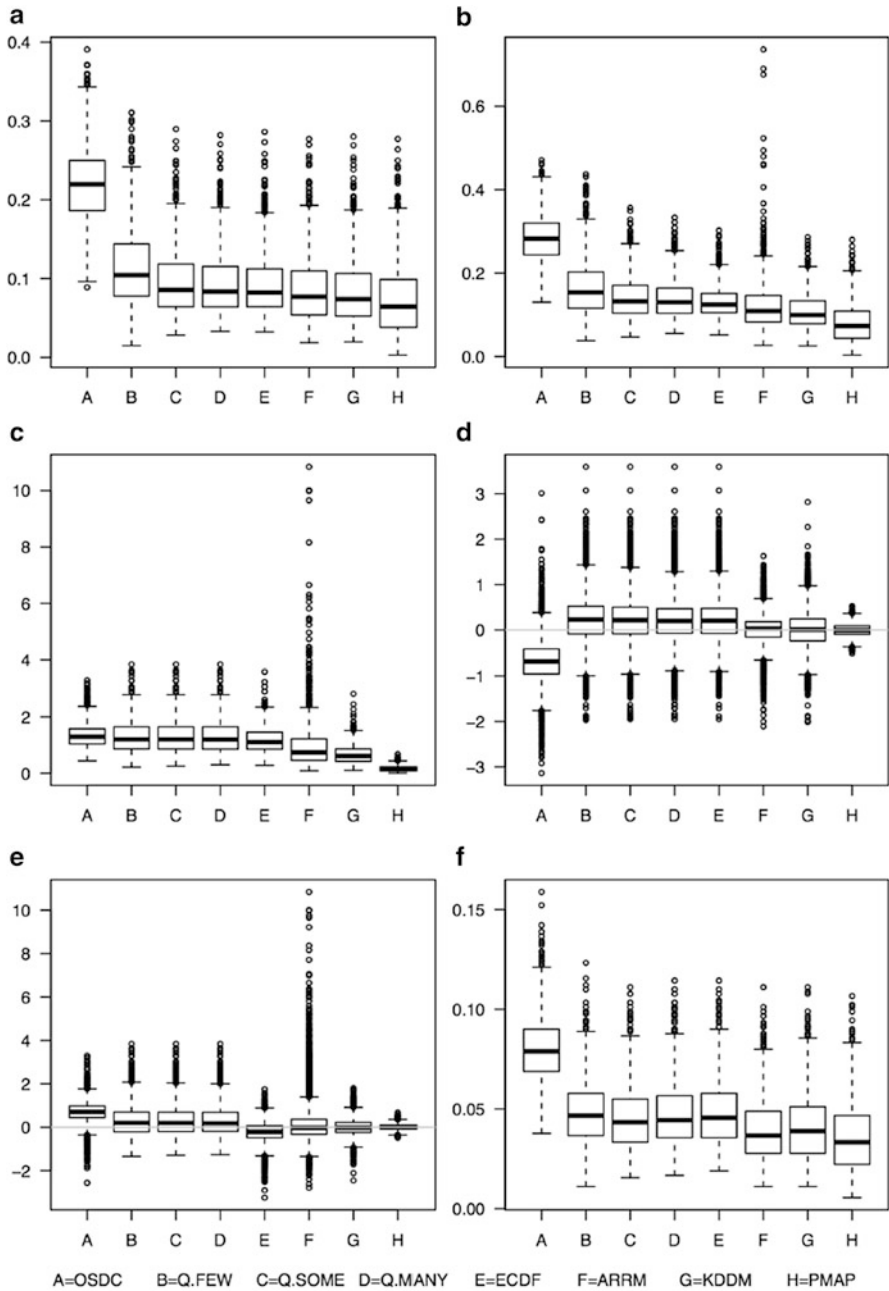
## 9.4 Evaluation Results

We evaluate each technique using six metrics. Mean absolute error (MAE) and root-mean-square error (RMSE) measure the average difference from the oracle, weighted toward larger errors in the case of RMSE. Maximum error measures the absolute value of the single largest difference from the oracle. Left and right tail errors are the difference from the oracle of the upper and lower 1 % of values in each dataset. Finally, the Kolmogorov-Smirnov (K-S) statistic measures the maximum distance between the CDFs of the two datasets.

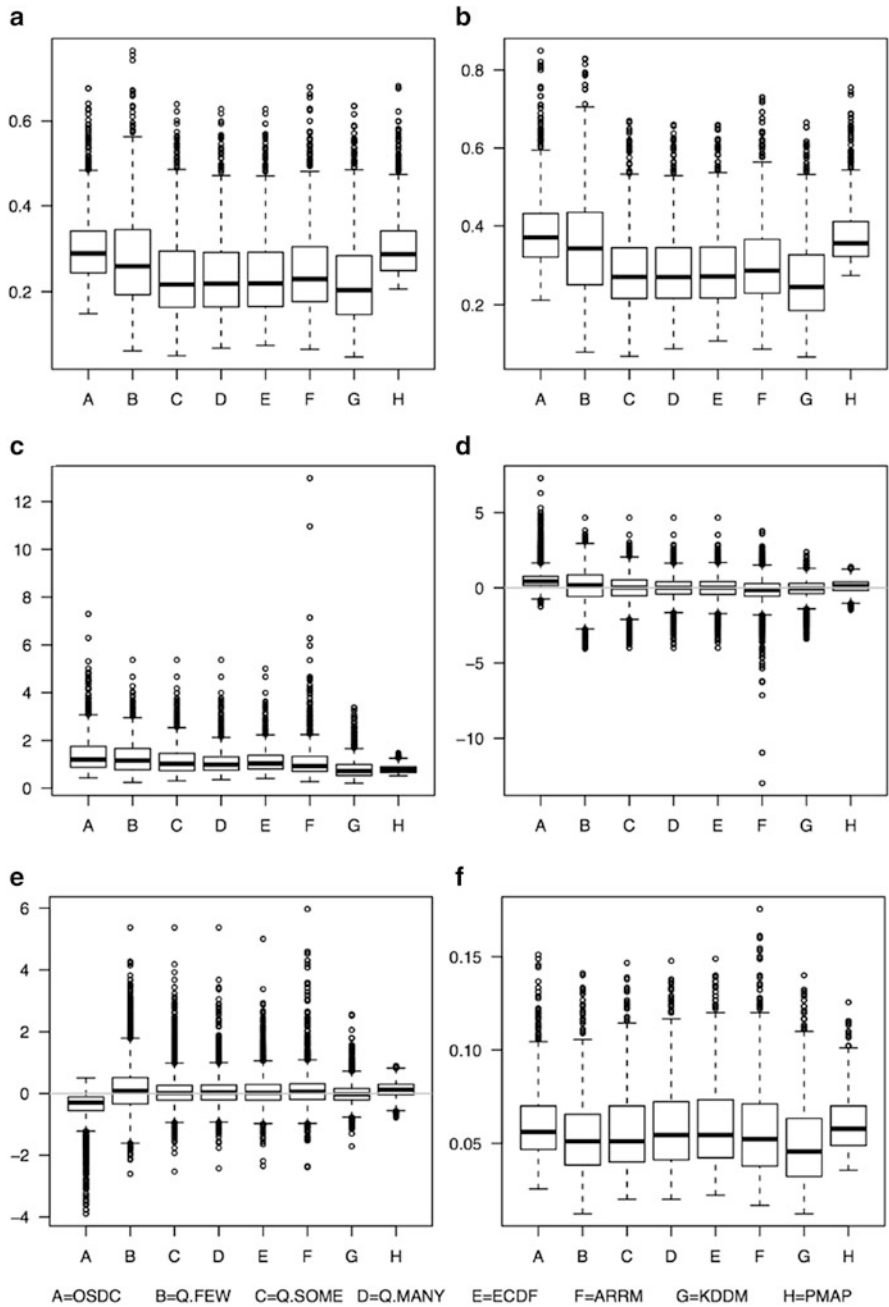
Boxplots of the six metrics show similar patterns for both the normal (Fig. 9.2) and gamma distributions (not shown): OSDC generally performs worst, followed in order of improving performance by QMAP, ECDF, ARRM, KDDM, and PMAP. For the mixture distribution (Fig. 9.3), the same overall pattern holds among the nonparametric techniques, but PMAP's performance is now worse than most of the other techniques on the MAE, RMSE, and K-S metrics. This illustrates a particular hazard of distribution-fitting techniques: when real-world data doesn't follow a fittable distribution, performance may be much worse than expected.

We conclude that although probability mapping is the best performer if the data comes from a known parametric distribution, because that assumption does not hold generally (even though it is common practice to pretend otherwise), the technique is not the best choice for general purpose or automated bias correction of large datasets.

For general use, KDDM emerges at the best overall performer. In addition to scoring best out of all the nonparametric methods, it does not require that the data be easily fittable, performs nearly as well as PMAP when the data is fittable, can accommodate differently sized input and output datasets, and is nearly as fast as the fastest methods. KDDM is also very simple to implement and therefore less vulnerable to coding errors than more complicated methods. Finally, because kernel density estimation is a well-developed topic in statistical analysis, there is an established body of knowledge that can be leveraged to generalize KDDM to new applications and optimize its performance in special cases.



**Fig. 9.2** Comparative performance of different distribution mapping techniques on normal data. (a) Mean absolute error (b) Root mean square error (c) Maximum error (d) Left tail error (e) Right tail error (f) K-S Statistic



**Fig. 9.3** Comparative performance of different distribution mapping techniques on data coming from a mixture distribution. (a) Mean absolute error (b) Root mean square error (c) Maximum error (d) Left tail error (e) Right tail error (f) K-S Statistic



To further expand the usefulness of this technique, we plan to write a paper evaluating distribution mapping techniques applied to reanalysis-driven RCM output. We also plan to develop an R package for bias correction and a multivariate bias-correction technique based on KDDM.

**Acknowledgments** Thanks to Anne Stoner for providing a very helpful review of our implementation of the ARRM algorithm that corrected some errors. Thanks also to Dorit Hammerling and Ian Scott-Fleming for helpful comments on drafts of the paper. This research was supported by the NSF Earth Systems Models (EaSM) Program award number 1049208, and by the DoD Strategic Environmental Research and Development Program (SERDP) under project RC-2204.

## References

- Ashfaq M et al (2010) Influence of climate model biases and daily-scale temperature and precipitation events on hydrological impacts assessment. *JGR* 115:D14116
- Boé J et al (2007) Statistical and dynamical downscaling of the Seine basin climate for hydro-meteorological studies. *Int J Climatol* 27:1643–1655
- Gudmundsson L (2014) qmap: statistical transformations for post-processing climate model output. R package version 1.0-2
- Gudmundsson L et al (2012) Technical note: downscaling RCM precipitation to the station scale using statistical transformations – a comparison of methods. *HESS* 16:3383–3390. doi:[10.5194/hess-16-3383-2012](https://doi.org/10.5194/hess-16-3383-2012)
- Haerter JO et al (2011) Climate model bias correction and the role of timescales. *HESS* 15:1065–1079. doi:[10.5194/hess-15-1065-2011](https://doi.org/10.5194/hess-15-1065-2011)
- Iizumi T et al (2011) Evaluation and intercomparison of downscaled daily precipitation indices over Japan in present day climate. *JGR* 116:D01111
- Ines AVM, Hansen JW (2006) Bias correction of daily GCM rainfall for crop simulation studies. *Agr Forest Meteorol* 138:44–53
- Johnson F, Sharma A (2011) Accounting for interannual variability: a comparison of options for water resources climate change impacts assessments. *WRR* 47:W045508
- Maurer EP et al (2002) A long-term hydrologically-based data set of land surface fluxes and states for the conterminous United States. *J Climate* 15(22):3237–3251
- Mearns LO et al (2007, updated 2013) The North American Regional Climate Change Assessment Program dataset. National Center for Atmospheric Research Earth System Grid data portal, Boulder, CO. Data downloaded 2012-03-23. doi:[10.5065/D6RN35ST](https://doi.org/10.5065/D6RN35ST)
- Mearns LO et al (2009) A regional climate change assessment program for North America. *Eos Trans AGU* 90(36):311–312
- Panofsky HA, Brier GW (1968) Some applications of statistics to meteorology. Pennsylvania State University Press, University Park, pp 40–45
- Piani C et al (2010) Statistical bias correction for daily precipitation in regional climate models over Europe. *Theor Appl Climatol* 99:187–192
- R Core Team (2014) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>
- Stoner A et al (2012) An asynchronous regional regression model for statistical downscaling of daily climate variables. *Int J Climatol* 33(11):2473–2494
- Teutschbein C, Seibert J (2012) Bias correction of regional climate model simulations for hydrological climate-change impact studies. *J Hydrol* 456–457:11–29
- Wilby RL et al (2014) The Statistical DownScaling Model – Decision Centric (SDSM-DC): conceptual basis and applications. *Clim Res* 61:251–268. doi:[10.3354/cr01254](https://doi.org/10.3354/cr01254)
- Wood AW et al (2004) Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. *Clim Change* 62:189–216