Calibration Strategies
A Source of Additional Uncertainty in Climate Change Projections

BY CHUN KIT HO, DAVID B. STEPHENSON, MATTHEW COLLINS, CHRISTOPHER A. T. FERRO, AND SIMON J. BROWN

R eliable projections of weather variables from climate models are required for the assessment of future climate change impacts (e.g., flooding, drought, temperature-related mortality, and crop yield). Assessments of such impacts are made by driving impact models with relevant weather variables from climate model simulations (e.g., daily temperature for temperature-related mortality assessment). However, it is generally necessary to adjust (calibrate) the variables to correct for climate model biases rather than to drive impact models with raw climate model output. Climate models are imperfect representations of reality, and so systematic discrepancies occur between climate model simulations and observations. Model discrepancies arise from many sources, such as structural uncertainty caused by representing the atmosphere by a finite number of variables, uncertainties in physical and subgrid-scale parameterization schemes, and uncertainty in how best to choose the model parameters.

Various calibration methods have been used in climate change studies, but few of the published studies carefully investigate the sensitivity of their results to the choice of calibration method. In this article, we show how existing approaches fall into two distinct calibration pathways ("bias correction" and "change factor") and then show how the choice of pathway can produce substantially different climate projections of European daily surface air temperatures. Such differences have the potential to produce significantly different impact outcomes and hence are an important source of additional uncertainty that needs to be acknowledged. While it is common practice to adjust for biases in the mean, increasingly projections are made in terms of the distributions of variables. Account must therefore be taken of discrepancies in the ability of the model to reproduce the whole probability distribution of observed weather variables. We focus here on the simplest example of calibration of the probability distribution of variables separately at different grid-point locations, but it is noted that other studies have also developed multivariate calibration using multivariate regression (e.g., downscaling and forecast assimilation approaches) and pattern matching methods based on empirical orthogonal functions.

A SIMPLE FRAMEWORK OF TWO CALIBRATION PATHWAYS. Consider the simplest case of a single weather variable (e.g., daily temperature at a grid-point location) that is observable over a short recent present-day time period, and over a future time period, denoted by random variables X_o and X_o', respectively. Furthermore, assume that we have climate model simulations of the variable over these two periods, which will be denoted by X_m and X_m'. For future climate assessment, one needs to infer the distribution of future observations X_o' given available data samples of X_o, X_m, and X_m'. Figure 1 shows that there are two very distinct calibration pathways: mapping of future climate simulations X_m' to find X_o' ("bias correction"), and mapping of present-day observations X_o to find X_o' ("change factor").

AFFILIATIONS: HO—NCAS–Climate and the Department of Meteorology, University of Reading, Reading, United Kingdom; STEPHENSON AND FERRO—NCAS–Climate and Mathematics Research Institute, University of Exeter, Exeter, United Kingdom; COLLINS—Mathematics Research Institute, University of Exeter, and Met Office Hadley Centre, Exeter, United Kingdom; BROWN—Met Office Hadley Centre, Exeter, United Kingdom
CORRESPONDING AUTHOR: Chun Kit Ho, Department of Meteorology, University of Reading, Earley Gate, P.O. Box 243, Reading RG6 6BB, UK
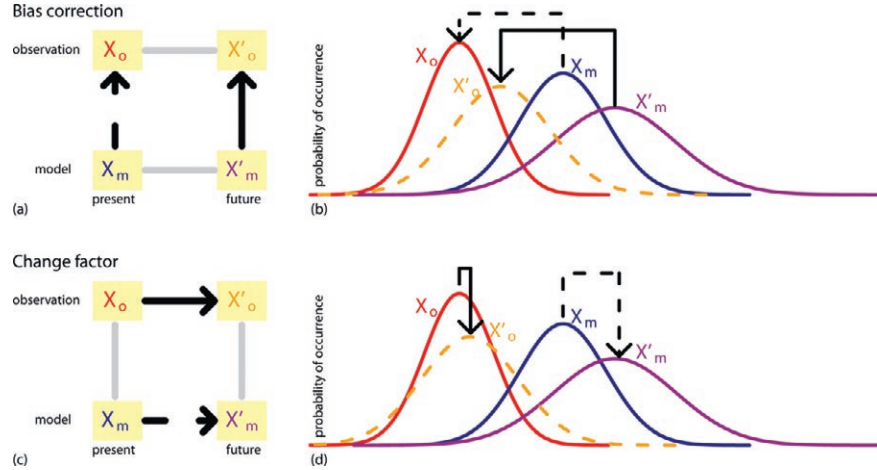E-mail: c.k.ho@reading.ac.uk

The bias correction strategy assumes that the model discrepancies stay constant in time [i.e., the relationship between the distributions of X_o and X_m is the same as the relationship between the distributions of X'_o and X'_m (Figs. 1a,b)]. This allows predictions of future observables to be obtained by mapping from future model simulations, $X'_o = B(X'_m)$, with a transfer function given by $B(X'_m) = F_o^{-1}[F_m(X'_m)]$, where $F_m(\,.\,)$ is the cumulative distribution function (CDF) of X_m and $F_o^{-1}(\,.\,)$ is the inverse CDF (the "quantile function") of X_o.

The change factor strategy is based on the alternative assumption that the change from present-day to future in

FIG. 1. (a,c) Schematic diagrams showing the two main pathways for calibrating climate model projections: bias correction and change factor. (b,d) Examples of probability density functions (PDF) of the variables involved. The PDF of present-day observed (X_o), present-day modeled (X_m), and future modeled (X'_m) variables are shown as labeled. For illustration purposes, these variables have different location and scale but the same shape. Note that the dashed orange lines denoting the PDF of calibrated future projected variable (X'_o) depend upon which pathway is used.
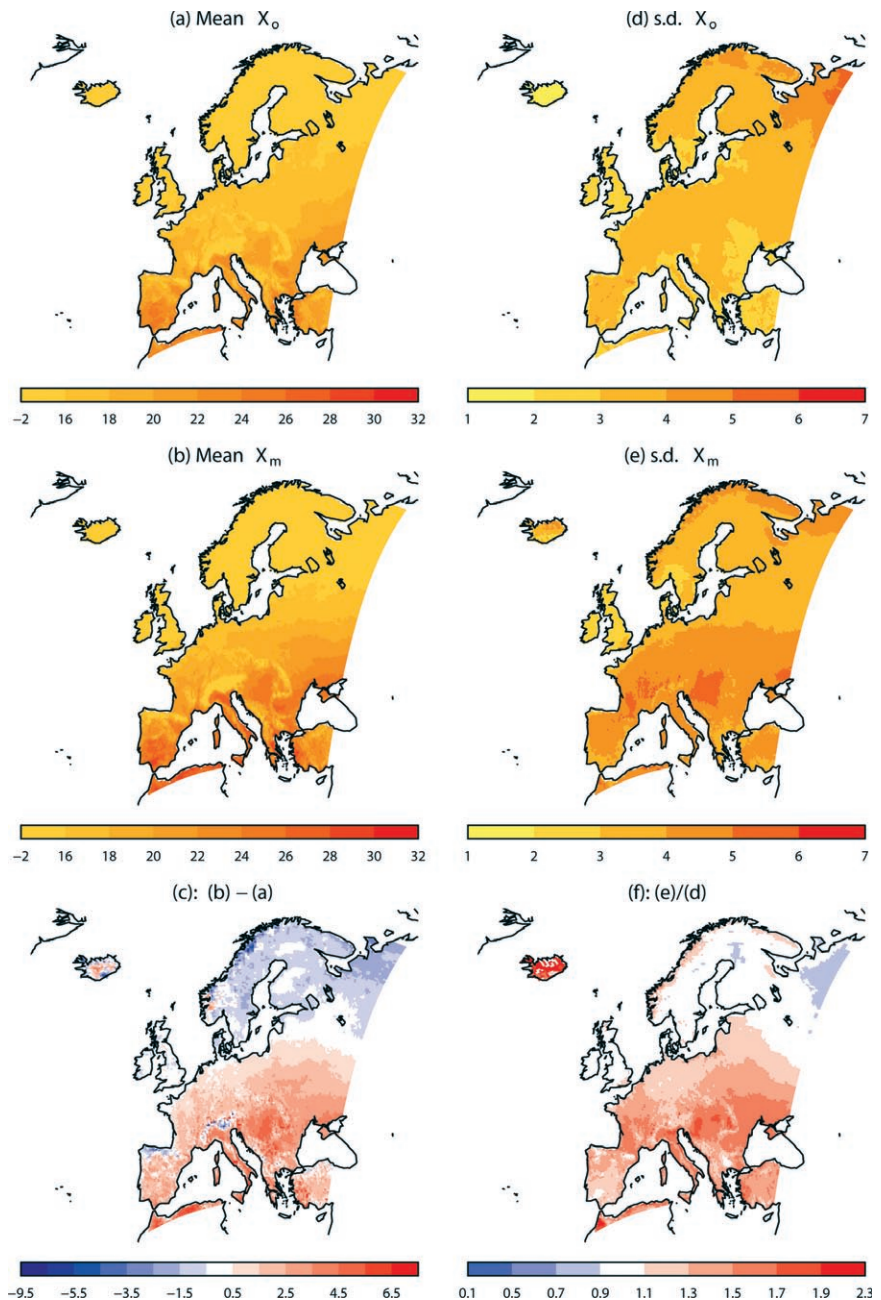
the observation distribution will be the same as the change in the model distribution [i.e., the relationship between the distributions of X_o and X'_o is the same as the relationship between the distributions of X_m and X'_m (Figs. 1c,d)]. This allows predictions of future observables to be obtained by mapping historical observations, $X'_o = C(X_o)$, with the transfer function given by $C(X_o) = F_{m'}^{-1}[F_m(X_o)]$ where $F_{m'}^{-1}(\,.\,)$ is the inverse CDF of X'_m.

The probability distribution of future observables depends on the choice of strategy: the CDF of predicted future observable $\hat{X}'_o$ is given by $F_o(x) = F_o\{F_m^{-1}[F_{m'}(x)]\}$ for bias correction, whereas it is given by $F_o(x) = F_{m'}\{F_m^{-1}[F_o(x)]\}$ for change factor. It is not obvious which strategy is better since both sets (or maybe neither set) of assumptions might be valid. Furthermore, it is not easy to empirically assess the assumptions (i.e., model biases are invariant with time for bias correction, and changes in modeled variable with time are consistent with changes in observations for change factor). We shall show below that this fundamental indeterminacy in the choice of calibration strategy is an additional and important source of uncertainty in climate projections.

DIFFERENT PROJECTIONS FROM THE TWO PATHWAYS. In addition to the choice of strategy, the transfer function for either strategy needs

to be estimated, and this can be done in a variety of different ways. The simplest "quantile matching" approach uses the empirical quantiles of the datasets (i.e., values sorted in ascending order) to find the transfer functions. This approach can be problematic if the data have different ranges or many tied values. The problem can be alleviated using parametric approaches where the CDFs are assumed to have known functional forms (e.g., the normal distribution). For the sake of illustration, in what follows we estimate the transfer functions using a hybrid "semiparametric" approach, whereby the predicted future distribution is estimated by the empirical distributions of the variables after adjusting for changes between the location (e.g., mean), scale (e.g., variance), and shape (e.g., skewness) parameters of such distributions. The location parameters, μ_o, μ_m, and μ'_m (for X_o, X_m, and X'_m, respectively), may be estimated by the mean or median of the data samples. The corresponding scale parameters, σ_o, σ_m, and σ'_m, may be estimated by the standard deviation or the interquartile range of the data samples. For the case where X_o and X_m have distributions with the same shape, then the bias correction strategy simply becomes $\hat{X}'_o = \mu_o + \frac{\sigma_o}{\sigma_m}(X'_m - \mu_m)$. Likewise, if X_m and X'_m have distributions with the same shape, then the change factor strategy gives the alternative transformation $\hat{X}'_o = \mu'_m + \frac{\sigma_o}{\sigma_m}(X_o - \mu_m)$. Surprisingly, these two strategies yield different ex-

pressions for the population mean of the future observations: $\mu_o + \frac{\sigma_o}{\sigma_m}(\mu'_m - \mu_m)$ for bias correction compared to $\mu'_m + \frac{\sigma'_m}{\sigma_m}(\mu_o - \mu_m)$ for change factor. The difference in predicted distribution due to the choice of strategy is fundamental and occurs with whichever method is used to estimate the transfer functions.

One special case is worth noting. If we further assume X_o, X_m, and X'_m to have equal scale (σ_o, σ_m, and σ'_m are equal), then for bias correction, future observations of a weather variable are estimated by subtracting the difference between the mean of present-day observations and model simulations from future model projections. This is equivalent to what climate modelers refer to as "removal of mean bias." On the other hand, for change factor, future observations are estimated by adding the change in the mean of model simulations and present-day observations. This is essentially the widely used simple technique of "adding model anomalies to observations." For equal scale, the two strategies yield identical probability distributions for the future observables. However, for unequal scale, implicit calibration by use of "anomalies" can give different results depending on the pathway.

The equal scale assumption is hard to justify, since variance as well as mean are often found to differ between model simulations and observations, and to change in climate-model projections (see for example Fig. 2.32 in Folland et al. 2001). The location and scale approach is therefore more justifiable than simply location adjustment for calibration in impact

Fig. 2. Maps showing summary statistics for present-day observed (X_o) and model-simulated (X_m) summer daily mean air temperatures in Europe: (a) and (b) mean (°C); (d) and (e) standard deviation (°C). The difference in the sample mean and the ratio of standard deviation are shown in (c) and (f), respectively.

assessments. For location and scale adjustment, the two calibration strategies can give different distributions of future observations—the population mean of estimated $\hat{X}'_o$ for the two strategies are not necessarily identical (Figs. 1b,d). This implies that any impact assessments based on these climate projections are also likely to differ.

EXAMPLE: CALIBRATION OF TEMPERATURE PROJECTIONS FOR EUROPE.

We illustrate the ideas by the calibration of summer (June–August) daily mean surface air temperatures in Europe projected by the Hadley Centre regional climate model HadRM3. Daily air temperature is an important variable for many climate change impacts, such as drought, heat-related mortalities, and electricity demand. HadRM3 has a minimum horizontal grid spacing of 25 km and is forced at the boundary by the output of Hadley Centre coupled ocean–atmosphere global climate model HadCM3. Historical external forcings of greenhouse gases, aerosols, solar irradiance, volcanic eruptions, and ozone, and future greenhouse gas concentrations from the IPCC SRES A1B scenario, are used to drive HadCM3. For observed surface air temperatures, we use the gridded European surface observational data set E-OBS, which has a grid identical to that of HadRM3. In this example, we define "present day" to be the 30-yr period from 1970 to 1999, and "future" to be from 2070 to 2099. For each grid point, the location and scale parameters of X_o, X_m, and X'_m in the transfer functions are estimated by their sample mean and standard deviation, respectively. The effect of trend over these short 30-yr time slice periods is negligible compared to natural day-to-day variability, and so the values may be assumed to be almost identically distributed with constant location and scale.

We first compare spatial maps showing sample statistics of location and scale of observed and HadRM3 sim-

ulated present-day temperatures (Fig. 2). Although the spatial patterns of observed mean temperatures are generally well simulated by HadRM3 (Figs. 2a, b), the regional climate model has a warm bias of around 2°–4°C over southern Europe and has a small cold bias over parts of Scandinavia (Fig. 2c). The model
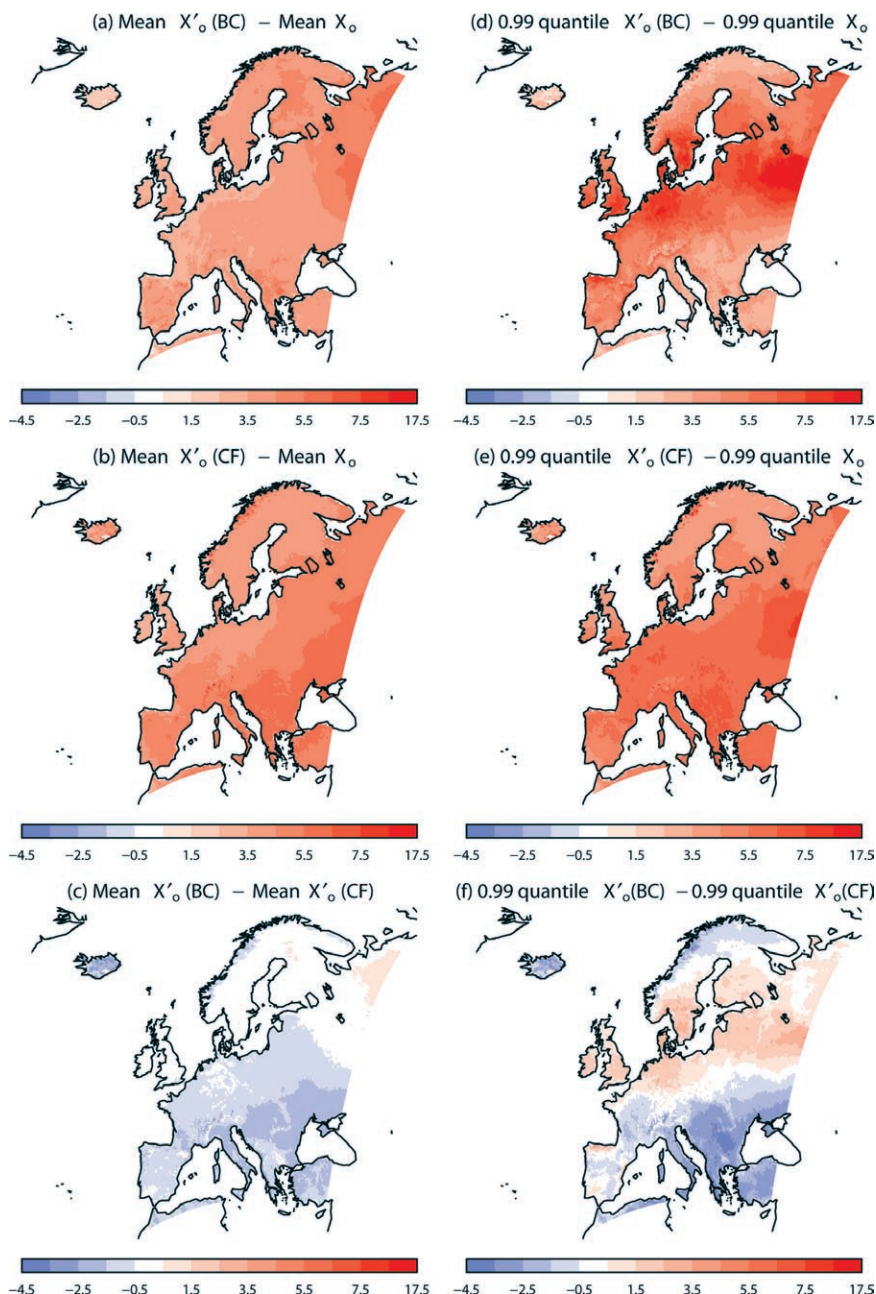
FIG. 3. (a,b) HadRM3 projected changes in the mean of summer daily mean air temperatures in 2070–2099 calibrated by bias correction (BC) and change factor (CF) (X'_o) relative to present-day observations (X_o) (in °C). (c) Difference between the changes calibrated by the two strategies [(a) minus (b) in °C]. (d)–(f) As in (a)–(c), respectively, but for the 0.99 quantile.

also overestimates the present-day variance over most parts of continental Europe, especially in the south, where the standard deviation of modeled daily mean temperatures is more than 50% greater than that of observed temperatures (Fig. 2f). This comparison confirms the need for calibration of both location and scale before HadRM3 temperature projections are used for any impact assessments.

Because of the warm biases in the climate model simulations, the increases in the mean of daily mean temperatures in the period 2070–2099 relative to present-day observations estimated by the two calibration strategies are both lower than the raw model projections, especially in southern Europe (not shown). The warming in mean temperatures estimated by bias correction (Fig. 3a) is generally lower than that by change factor (Fig. 3b), by around 2°C over southern Europe (Fig. 3c). To put this 2°C difference in the mean of temperatures in context, this is comparable to the difference in the projected globally averaged temperature increase for different emissions scenarios by the models in IPCC AR4, for example between SRES scenarios A2 and B1 (see Fig. 10.4 in Meehl et al. 2007).

For many climate change impacts, changes in the tails of distributions of weather variables are often more important than changes in the mean. In our case of daily temperatures, an example is that the daily number of heat-related deaths increases substantially at higher air temperatures. Here we examine the changes in the 0.99 empirical quantile of the daily mean temperature, which corresponds to the temperature with a return period of approximately one summer (i.e., hottest day of the year). For most places, both calibration strategies suggest that the increase in the 0.99 quantile between present-day and future is larger than the corresponding increase in the mean, but they give noticeably and even qualitatively different spatial patterns of warming. For bias correction (Fig. 3d), the largest increase in the 0.99 quantile occurs over northern Europe, including southern Scandinavia, northern Germany, the United Kingdom, western Russia, and also the northern coast of Spain. In contrast, the increase is more spatially uniform for temperatures calibrated by change factor (Fig. 3e). The warming predicted using bias correction is more than 3°C lower than that estimated by change factor in southern Europe (Fig. 3f). This is mainly related to the difference in the mean of temperatures calibrated by the two strategies (refer to Fig. 3c). In the north, however, the bias correction strategy estimates more pronounced warming of the 0.99 quantile of temperatures than change factor, by up to 4°C in some places. Examination of spatial

maps of sample skewness statistics (not shown) reveals that there are two reasons for such a difference. First, HadRM3 simulated present-day temperatures are more positively skewed compared to observations in parts of this region, a feature observed in simulations of a number of other regional climate models. Second, HadRM3 projects temperatures to become even more positively skewed with time over northern continental Europe. The transfer functions used here have not accounted for such differences in shape, and how such differences can be calibrated is beyond the scope of this article.

CONCLUDING REMARKS. This article has identified and discussed the assumptions behind two distinct strategies used in calibrating climate projections: bias correction and change factor. In an example of European temperature projections by a regional climate model, we have shown that the two strategies give substantially different spatial patterns of warming. It is, however, not obvious which strategy gives more plausible results for use in impact assessments. This gives rise to an additional source of uncertainty due to calibration indeterminacy. Previous impact assessments have identified various sources of uncertainties in impact projections: uncertainty in future greenhouse gas emissions, uncertainty related to the choices of parameters and structure of climate models, uncertainty in impact modeling, and uncertainty in adaptation actions. In addition to these known uncertainties, impact modelers should also be aware of the possible uncertainty associated with the choice of calibration strategy and should explore the sensitivity of impact projections to different strategies as part of their impact assessment. When adopting a particular calibration strategy, impact modelers should also carefully assess whether the underlying assumptions in the calibration method are valid.

Given the importance of calibration in impact assessments, further research in this area is clearly required. For example, it can be argued that the simple assumptions underlying both bias correction and change-factor strategies are rather unrealistic, and so more rigorous statistical frameworks need to be developed and tested (e.g., Bayesian models that are capable of predicting true climate by properly accounting for model discrepancy and observational errors). Furthermore, the bias correction strategy presented here assumes no changes in climate model biases with time. As this assumption may not be appropriate for certain physical processes—such as the relationship between temperature and soil moisture—

further work is needed to explore the possibility of performing calibration taking nonstationarity in biases into account.

FOR FURTHER READING

Arnell, N., D. Hudson, and R. Jones, 2003: Climate change scenarios from a regional climate model: Estimating change in runoff in southern Africa. *J. Geophys. Res.—Atmos.*, **108**, doi:10.1029/2002JD002782.

Benestad, R. E., 2001: A comparison between two empirical downscaling strategies. *Int. J. Climatol.*, **21**, 1645–1668, doi:10.1002/joc.703.

Buser, C. M., H. R. Künsch, D. Lüthi, M. Wild, and C. Schär, 2009: Bayesian multi-model projection of climate: Bias assumptions and interannual variability. *Climate Dyn.*, **33**, 849–868, doi:10.1007/s00382-009-0588-6.

Christensen, J. H., F. Boberg, O. B. Christensen, and P. Lucas-Picher, 2008: On the need for bias correction of regional climate change projections of temperature and precipitation. *Geophys. Res. Lett.*, **35**, L20709, doi:10.1029/2008GL035694.

Diaz-Nieto, J., and R. L. Wilby, 2005: A comparison of statistical downscaling and climate change factor methods: Impacts on low flows in the River Thames, United Kingdom. *Clim. Change*, **69**, 245–268, doi:10.1007/s10584-005-1157-6.

Folland, C. K., and Coauthors, 2001: Observed climate variability and change. *Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change*, J. Houghton et al., Eds., Cambridge University Press, 99–181.

Giorgi, F., and N. Diffenbaugh, 2008: Developing regional climate change scenarios for use in assessment of effects on human health and disease. *Climate Res.*, **36**, 141–151, doi:10.3354/cr00728.

Haerter, J. O., S. Hagemann, C. Moseley, and C. Piani, 2011: Climate model bias correction and the role of timescales. *Hydrol. Earth Syst. Sci.*, **15**, 1065–1079, doi:10.5194/hess-15-1065-2011.

Hawkins, E., T. M. Osborne, C. K. Ho, and A. J. Challinor, 2011: Calibration and bias correction of climate projections for crop modeling: An idealized case study over Europe. Submitted to *Agric. For. Meteorol.*

Haylock, M. R., N. Hofstra, A. M. G. Klein Tank, E. J. Klok, P. D. Jones, and M. New, 2008: A European daily high-resolution gridded dataset of surface temperature and precipitation for 1950–2006. *J. Geophys. Res.*, **113**, D20119, doi:10.1029/2008JD010201.

Ho, C. K., 2010: Projecting extreme heat-related mortality in Europe under climate change. Ph.D. thesis, University of Exeter, 196 pp. [Available online at http://hdl.handle.net/10036/107739.]

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

Jones, R., 2001: An environmental risk assessment/management framework for climate change impact assessments. *Nat. Hazards*, **23**, 197–230.

Meehl, G. A., and Coauthors, 2007: Global climate projections. *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, S. Solomon et al., Eds., Cambridge University Press, 747–845.

Moberg, A., and P. D. Jones, 2004: Regional climate model simulations of daily maximum and minimum near-surface temperatures across Europe compared with observed station data 1961–1990. *Climate Dyn.*, **23**, 695–715, doi:10.1007/s00382-004-0464-3.

Murphy, J., 1999: An evaluation of statistical and dynamical techniques for downscaling local climate. *J. Climate*, **12**, 2256–2284.

Piani, C., J. Haerter, and E. Coppola, 2010a: 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.

——, G. P. Weedon, M. Best, S. M. Gomes, P. Viterbo, S. Hagemann, and J. O. Haerter, 2010b: Statistical bias correction of global simulated daily precipitation and temperature for the application of hydrological models. *J. Hydrol.*, **395**, 199–215, doi:10.1016/j.jhydrol.2010.10.024.

Rougier, J. C. 2007: Probabilistic inference for future climate using an ensemble of climate model evaluations. *Clim. Change*, **81**, 247–264, doi:10.1007/s10584-006-9156-9.

Wood, A. W., L. R. Leung, V. Sridhar, and D. P. Lettenmaier, 2004: Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. *Clim. Change*, **62**, 189–216, doi:10.1023/B:CLIM.0000013685.99609.9e.