LLNL-TR-426343

Accounting for Global Climate Model Projection Uncertainty in Modern Statistical Downscaling

G. Johannesson

March 25, 2010

Disclaimer

This document was prepared as an account of work sponsored by an agency of the United States government. Neither the United States government nor Lawrence Livermore National Security, LLC, nor any of their employees makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States government or Lawrence Livermore National Security, LLC. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or Lawrence Livermore National Security, LLC, and shall not be used for advertising or product endorsement purposes.

ACCOUNTING FOR GLOBAL CLIMATE MODEL PROJECTION UNCERTAINTY IN MODERN STATISTICAL DOWNSCALING

GARDAR JOHANNESSON
LAWRENCE LIVERMORE NATIONAL LABORATORY

1. INTRODUCTION

Future climate change has emerged as a national and a global security threat. To carry out the needed adaptation and mitigation steps, a quantification of the expected level of climate change is needed, both at the global and the regional scale; in the end, the impact of climate change is felt at the local/regional level. An important part of such climate change assessment is uncertainty quantification. Decision and policy makers are not only interested in "best guesses" of expected climate change, but rather probabilistic quantification (e.g., Rougier, 2007). For example, consider the following question:

> What is the probability that the average summer temperature will increase by at least 4°C in region R if global CO_2 emission increases by $P\%$ from current levels by time T?

It is a simple question, but one that remains very difficult to answer. It is answering these kind of questions that is the focus of this effort.

The uncertainty associated with future climate change can be attributed to three major factors:

(1) Uncertainty about future emission of green house gasses (GHG).
(2) Given a future GHG emission scenario, what is its impact on the global climate?
(3) Given a particular evolution of the global climate, what does it mean for a particular location/region?

In what follows, we assume a particular GHG emission scenario has been selected. Given the GHG emission scenario, the current batch of the state-of-the-art global climate models (GCMs) is used to simulate future climate under this scenario, yielding an ensemble of future climate projections (which reflect, to some degree our uncertainty of being able to simulate future climate give a particular GHG scenario). Due to the coarse-resolution nature of the GCM projections, they need to be spatially downscaled for regional impact assessments. To downscale a given GCM projection, two methods have emerged: *dynamical downscaling* and *statistical (empirical) downscaling* (SDS). Dynamic downscaling involves configuring and running a regional climate model (RCM) nested within a given GCM projection (i.e., the GCM provides bounder conditions for the RCM). On the other hand,
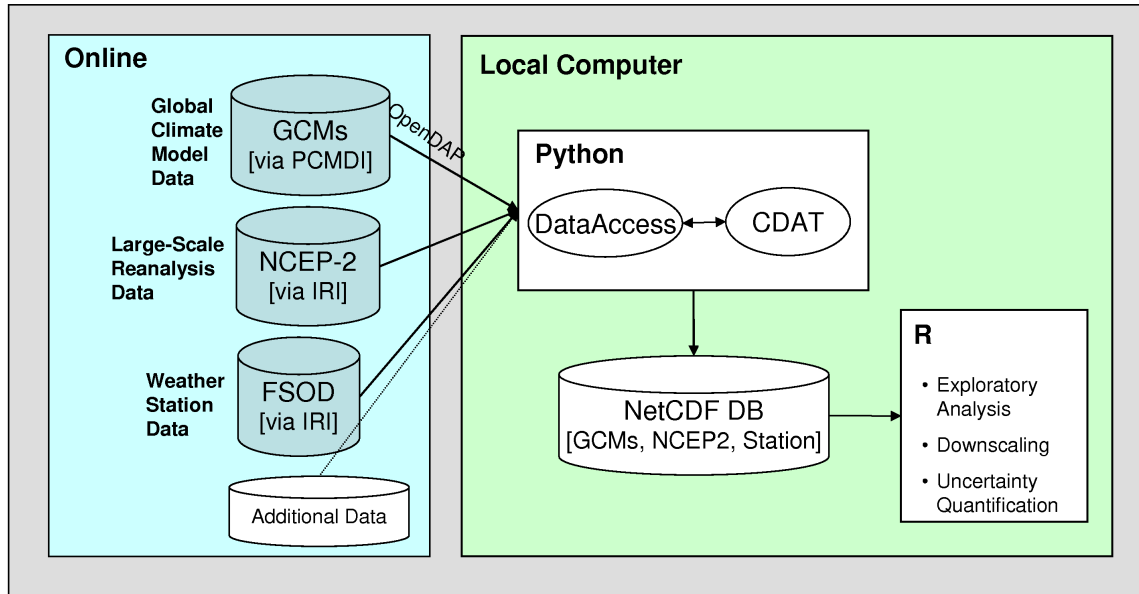
FIGURE 1.   Tools written in Python are used to extract the needed data
from various online sources for a regional assessment. The regional analysis
is then carried out using tools written in R.

statistical downscaling aims at establishing a statistical relationship between observed lo-
cal/regional climate variables of interest and synoptic (GCM-scale) climate predictors. The
resulting empirical relationship is then applied to future GCM projections. A comparison
of the pros and cons of dynamical versus statistical downscaling is outside the scope of
this effort, but has been extensively studied and the reader is referred to Wilby et al.
(1998); Murphy (1999); Wood et al. (2004); Benestad et al. (2007); Fowler et al. (2007),
and references within those.

The scope of this effort is to study methodology, a statistical framework, to propagate
and account for GCM uncertainty in regional statistical downscaling assessment. In par-
ticular, we will explore how to leverage an ensemble of GCM projections to quantify the
impact of the GCM uncertainty in such an assessment. There are three main component
to this effort: (1) gather the necessary climate-related data for a regional SDS study, in-
cluding multiple GCM projections, (2) carry out SDS, and (3) assess the uncertainty. The
first step is carried out using tools written in the Python programming language, while
analysis tools were developed in the statistical programming language R; see Figure 1

2. DATA GATHERING AND PREPARATION

To carry out a regional climate change assessment, the results from multiple GCMs,
along with various relevant observational datasets need to be gathered. To carry out this
task, a Python module developed under a LLNL's Engineering Techbase LDRD effort was
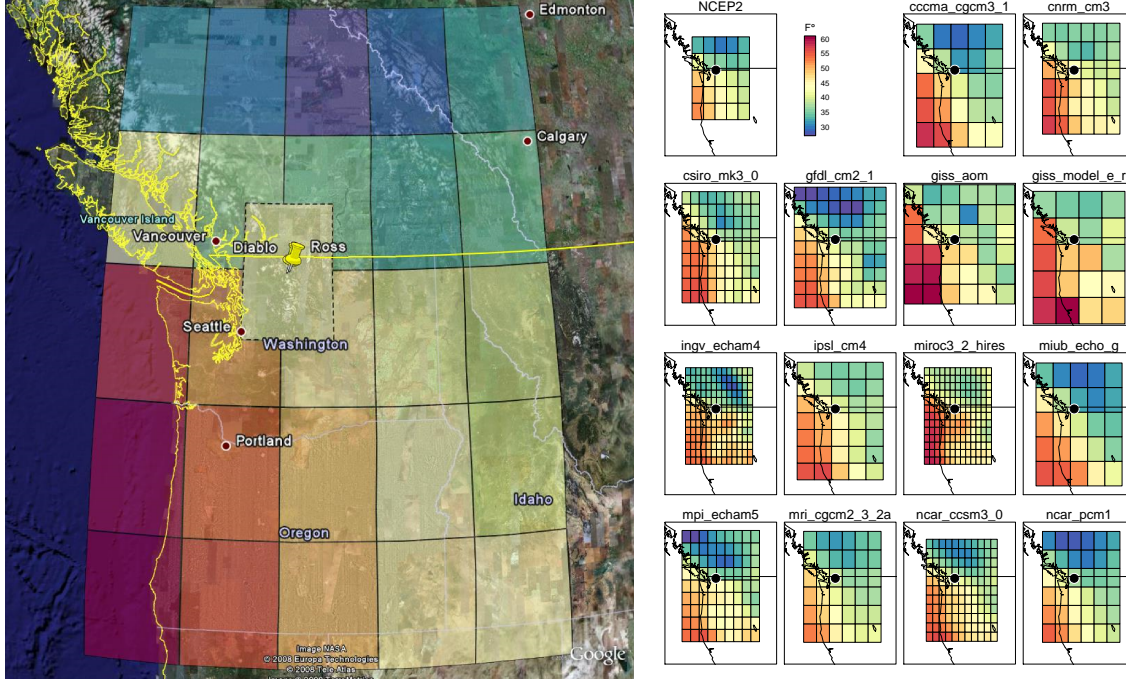
FIGURE 2.   Left, the location of the Diablo and the Ross dams in Washington State. The map is overlaid by a 5×5 grid showing the 1980–1999 average temperature at 2 meters from the NCEP/DOE Reanalysis 2 data (NCEP2). The smaller maps on the right contrast the spatial resolution of the NCEP2 data (the upper-left panel) to that of 14 different GCMs.

used and extended. This module uses the the Climate Data Analysis Tools (CDAT) module developed by the Program for Climate Model Diagnosis and Intercomparison (PCMDI) at LLNL to extract and post-process GCM data from the CMIP3 archive at PCMDI. It also fetches large-scale reanalysis data and local weather station data from the online Climate Data Library of the International Research Institution (IRI) of Climate and Society.

The regionalized data gathered this way are at various spatial and temporal resolution, and in different formats; see Figure 2(right) for a comparison of the spatial resolution of the NCEP/DOE Reanalysis 2 (NCEP2) data and 14 different GCMs. The second task of the Python module is to bring the data to a joint format (and spatial and temporal resolution) and build a local database consisting of large collection of NCDF files to be used for further analysis.

3. STATISTICAL DOWNSCALING (SDS)

The spatial resolution of the GCM projections is too course for regional assessment, as is evident from Figure 2. Statistical (empirical) downscaling (SDS) aims at establishing a

relationship between historical local/regional weather-data (say daily time series of temperature) and synoptic weather patterns—for example, NCEP2 variables. The local/regional data is provided by a network of weather stations and also fine-resolution gridded weather-data products. The synoptic weather-data is typically provided by reanalysis, which fuses together multiple sources of observational data using weather forecast-type models; an example is the NCEP2 data[1] shown in Figure 2.

Let D_t denote a given weather station observation at time $t = 1, \ldots, T^c$ (say, the recorded maximum temperature in day t) and let $\mathbf{O}_t = (O_{1t}, \ldots, O_{nt})'$ be a vector of synoptic atmospheric predictors at time t (say, NCEP2 daily temperature and sea-level pressure in a 3×3 grid centered on the region of interest; hence $n = 9 + 9 = 18$). The goal is to predict (statistically) D_t given $\mathbf{O}_t$. Various SDS models have been proposed for this task; see, for example Wilby et al. (2004) and Fowler et al. (2007) for an overview. They broadly fall in three groups: regression methods, weather generators, and weather typing (also called analogue methods or nearest-neighbor methods). In this study we focus on regression and weather generators.

Statistical regression is typically formulated as

$$D_t = F(\mathbf{O}_t; \boldsymbol{\theta}) + \varepsilon_t, \quad t = 1, \ldots, T^c,$$

where $F(\cdot; \boldsymbol{\theta})$ is the regression function, $\boldsymbol{\theta}$ a vector of unknown parameters associated with F, and ε_t a random regression error. The unknown parameter vector $\boldsymbol{\theta}$ is estimated using historical data, yielding an estimate $\hat{\boldsymbol{\theta}}$. The regression model can then be used to predict D for a given $\mathbf{O}$. The prediction is both presented as a "best guess" (expected value), $\hat{D}$, and in terms of a predictive distribution, $p(D \,|\, \mathbf{O})$.

Weather generators are very similar to the regression method above, except they are auto-regressive:

$$D_t = F(D_{t-1}, \mathbf{O}_t, \mathbf{O}_{t-1}; \boldsymbol{\theta}) + \varepsilon_t.$$

Hence, they can be used to generate a time series of downscaled values.

A key component in both methods is the regression function F. In this effort we investigate various well established regression models, including Multivariate Adaptive Regression Splines (MARS), Random Regression Forrest (RF), and Relevance Vector Machines (RVM); see Hastie et al. (2001) and Bishop (2007) for a general overview and references. These are all regression methods that are known for being able to capture non-linear and complex interaction in a large set of predictors.

Given a trained SDS model, the model can be used to downscale GCM projections. In doing so, the GCM projections need to be regridded to the same spatial and temporal resolution as the synoptic data used to train the SDS model (the $\mathbf{O}_t$ time series). Let $\mathbf{M}_{it}^c$, $t = 1, \ldots, T^c$, be the time series of synoptic predictors from the i-th GCM, $i = 1, \ldots, M$, for the same time period as the historical data $(D_t, \mathbf{O}_t)$; for example, derived from GCM projections for the 20th century (e.g., the 20c3m scenario in the CMIP3 database). Similarly, let $\mathbf{M}_{it}^f$, $t = 1, \ldots, T^f$, be the time series for a given future scenario (e.g., the

---

[1]http://www.cdc.noaa.gov/cdc/data.ncep.reanalysis2.html

SREAS A1B scenario in the CMIP3 database). The SDS model can then, for example be used to generate a time series of (expected) downscaled values,

$$\hat{D}_{i1}^c, \ldots, \hat{D}_{iT^c}^c \quad \text{and} \quad \hat{D}_{i1}^f, \ldots, \hat{D}_{iT_f}^f.$$

4. Uncertainty Quantification Using an Ensemble of GCMs

Uncertainty quantification of GCM projections has emerged as an important research area; see for example Tebaldi and Knutti (2007) for an overview and recent development. The general approach is to leverage an ensemble of GCM projections to construct single "average" projection. If all the ensemble members could be considered of equal quality and independent, then a simple average of the ensemble members is a better predictor than any single member. However, this is not in general the case; GCMs vary in global and regional fidelity.

A promising approach to fuse or blend together projections from multiple GCMs is via Bayesian statistical modeling, particularly along the lines of Tebaldi et al. (2005), Lopez et al. (2006), Tebaldi and Sanso (2008), and Smith et al. (2008). Within this framework, the fidelity of each GCM to reproduce a particular attribute of past climate is used to judge its ability to project the same attribute in future climate. That is, let X_0 be the observed climate signal of interest (say the mean temperature in a given region and time period) and let X_i, $i = 1, \ldots M$, be the same climate signal as reproduced by M different GCMs. Similarly, let Y_i, $i = 1, \ldots, M$, be the future projected signal by the same GCMs. In Smith et al. (2008) the following hierarchical Bayesian model is used to formulate the relationship between X_0, X_i, Y_i, and the unknown current and future climate signals, μ and ν, respectively:

$$X_0 \sim \text{Gau}(\mu, \lambda_0^{-1}),$$
$$X_i \sim \text{Gau}(\mu, \lambda_i^{-1}), \quad i = 1, \ldots, M,$$
$$Y_i \sim \text{Gau}(\nu + \beta(X_i - \mu), (\theta\lambda_i)^{-1}), \quad i = 1, \ldots, M,$$

where $x \sim \text{Gau}(\mu, \sigma^2)$ reads "x is Gaussian (normal) distributed with mean μ and variance σ^2". In the above formulation, the GCMs are allowed to be of different precision (λ_i), the current and future climate signal from each GCM can be correlated (as captured by β), and future projections can be of less (or better) precision than past (as captured by θ). This model is completed by assigning relative non-informative prior distributions to β and θ, but $\lambda_i \sim \text{Gam}(s, r)$, where the shape and the rate, s and r, of the Gamma distribution are given a non-informative priors.

The above Bayesian model yields a joint posterior distribution over all unknown parameters, given all observed evidence;

$$p(\mu, \nu, \beta, \theta, \lambda_1, \ldots, \lambda_M, s, r \,|\, X_0, X_1, \ldots, X_M, Y_1, \ldots, Y_M).$$

This is an unattractive distribution to work with in closed form and inference is carried by generate a random sample from the target distribution using a Markov chain Monte Carlo (MCMC) sampler; see, for example Robert and Casella (2004).

The Bayesian ensemble model reviewed above has been applied to fuse together climate signal at the GCM-scale (i.e., global signals, continental, etc.). In this effort we investigate its use in fusing together multiple downscaled results. As such, let

$$X_0 = G(D_1, \ldots, D_{T^c}),$$
$$X_i = G(\hat{D}_1^c, \ldots, \hat{D}_{T^c}^c), \quad i = 1, \ldots, M,$$
$$Y_i = G(\hat{D}_1^f, \ldots, \hat{D}_{T^f}^f), \quad i = 1, \ldots, M,$$

where $G()$ is a given summary statistic operator (e.g., the average, a given quantile, etc.), and recall that the D_i's are the observed values (at the given site) and the $\hat{D}_i^c$'s and the $\hat{D}_i^f$'s are the downscaled values for the current and future time period. In this case, the GCM specific precision parameters, the λ_i's, do not only represent the fidelity of each GCM, but also the additional error introduced by post-processing the GCM output for use in the downscaling model and the additional error introduced in the downscaling process itself.

5. Application: Two Sites on the Skagit River in Washington State

We now present a small application of site-specific downscaling of daily maximum temperature using an ensemble of GCMs.

We are interested in two sites along the Skagit river in Washington State; at the Diablo and the Ross dams, which are both part of Seattle City Light's Skagit Hydroelectricity project (Figure 2). Both dams have been collecting weather information for decades, including daily maximum temperature. The question of interest here is what can we expect the daily maximum temperature to be at these two sites in 2046–65 (assuming future GHG emission according to the SRES A1B scenario) and how much higher do we expect it to be compared to 1980–99?

We shall carry out two different analysis, one at the NCEP2 pixel scale ($2.5° \times 2.5°$) and then at the two locations in question. The first analysis does not requires SDS and is therefore not impacted by the additional uncertainty that SDS introduces.

5.1. **Large-Scale Analysis.** To carry out our large-scale analysis, both the NCEP2 and GCM projections are interpolated (regridded) to a single $2.5° \times 2.5°$ region (a "pixel") that is centered on the two locations of interest; see Figure 2(left). Figure 3 compares the empirical distribution of average daily temperature at 2 meters in the NCEP2 dataset from 1980–99 to that of 14 GCMs for the same time period and also for the future time period 2046–65 (given the SRES A1B scenario). As expected the GCMs deviate by different amount from the NCEP2 data for the 1980–99 time period (i.e., different fidelity of reproducing the past), but all of them show an increase in temperature for the 2046–65 time period. The average temperature for these two time periods is plotted in Figure 4(top-left), along with the projected increase in average temperature by the 14 GCMs (bottom-left). The two plots on the right in Figure 4 are analogous to the two on the left, except the time series of daily average temperature from each GCM has been bias corrected such that they have the same mean and standard deviation as the NCEP2 data for the 1980–99 time
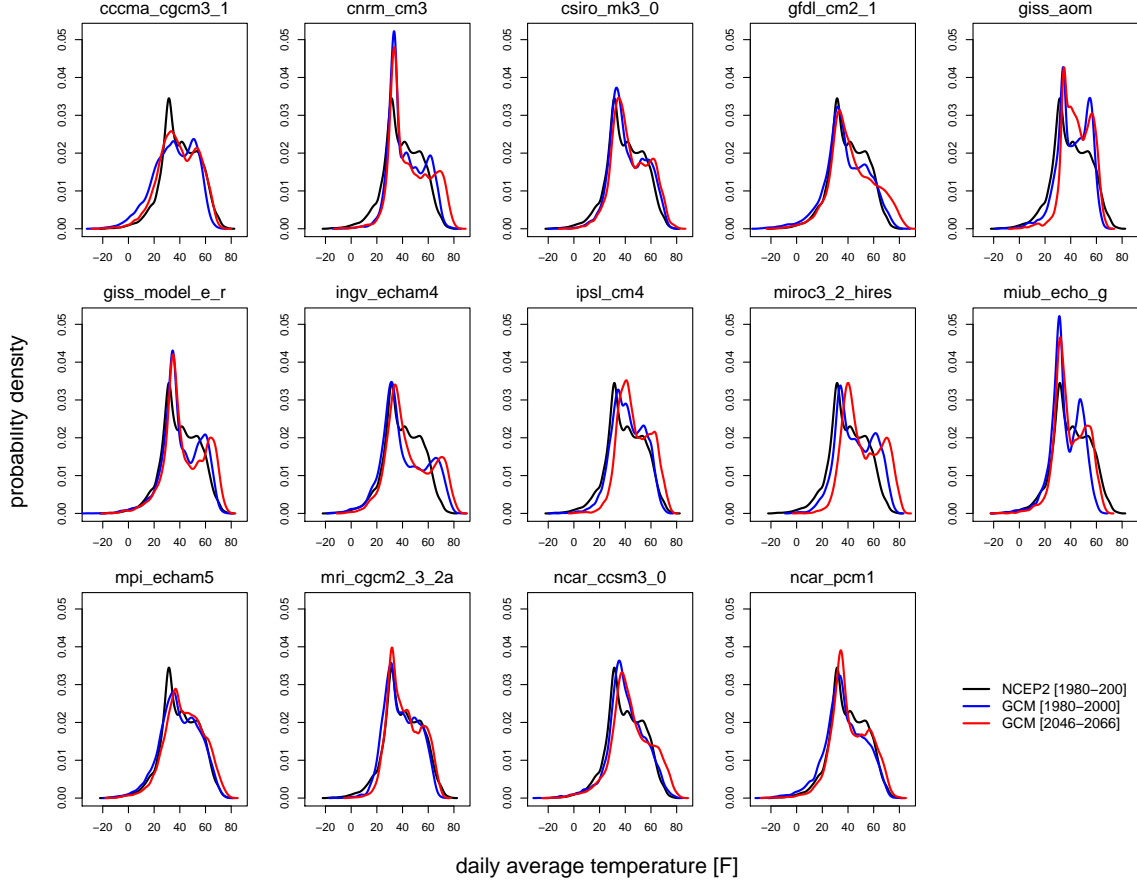
FIGURE 3.   Each panel shows the empirical density of daily average temperature at 2 meters in a $2.5° \times 2.5°$ grid-cell centered on the two locations of interest (see Figure 2). The three lines shown correspond to the NCEP2 data for 1980–99 and a given GCM for 1980–99 and for 2046–65.

period; hence, as a result, the average temperature produced by the 14 GCMs for 1980–99 is identical to that of the NCEP2 data. The result is, as expected much narrower range of values produced for the 2046–65 time period (the top-right plot). However, this has very little or no impact of the projected increase in average temperature, as is evident by the two bottom plots in Figure 4. The average temperature increase in the bottom-left plot is 3.63°F, with a standard deviation of 1.03.

So fare no attempt has been made to "weight" the 14 GCM projects differently in summarizing the expected increase in temperature. Figure 5, however presents the results of applying the hierarchical Bayesian model introduced previously. The Bayesian model yields narrower range for the expected average temperature in 2046–65 and the same can be said about the expected increase from the 1980–99 time period. The Bayesian model produces
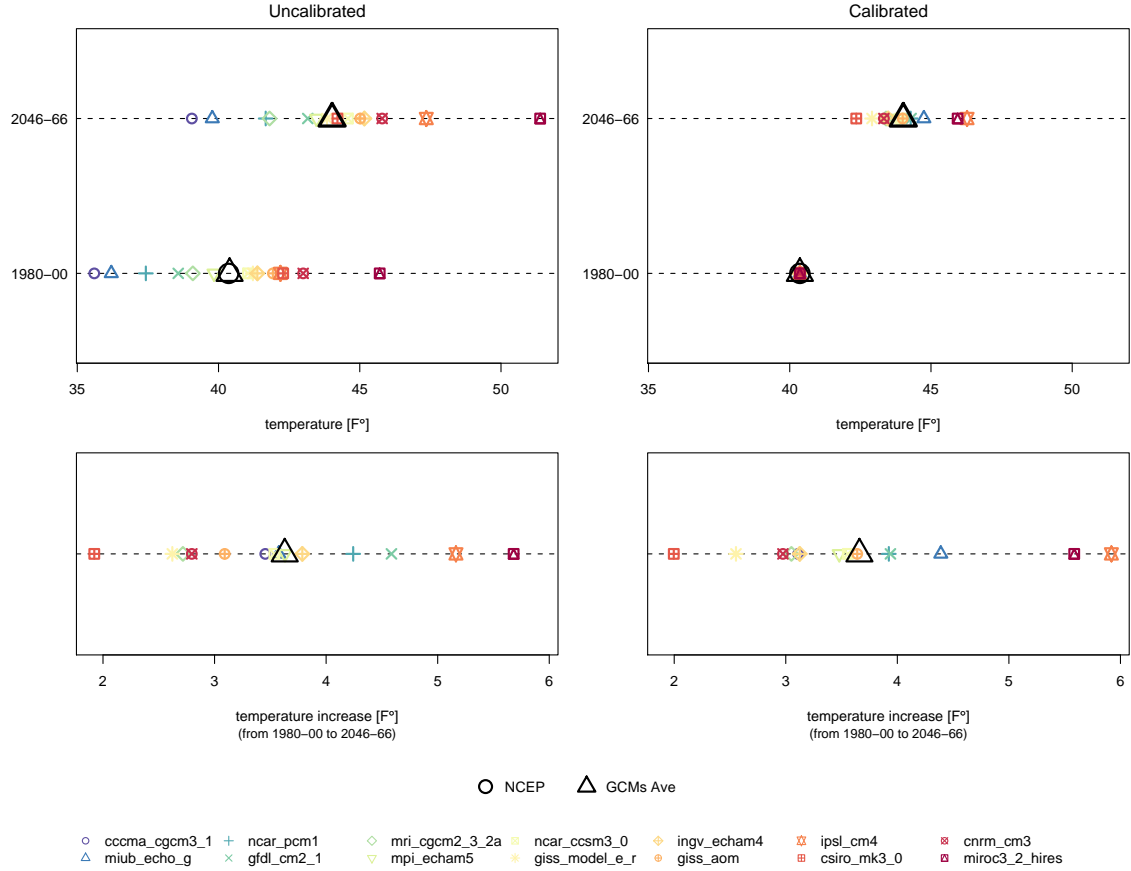
FIGURE 4.   Top-left, the average temperature at 2 meters in a $2.5° \times 2.5°$ grid-cell for 1980–99 and for 2046–65 as estimated by 14 GCMs and the NCEP2 data; this is the mean of the empirical distributions of Figure 3. Top-right shows the average temperature after bias-correcting the 14 GCMs (see text). The bottom plots show the increase in temperature as projected by the 14 GCMs.

richer results be providing a full (joint) posterior distribution across all parameters. The Bayesian model estimates the expected increase in temperature to be 3.58°F with a standard deviation of 0.33. The expected increase is close to identical to the average of the 14 GCMs, however, the standard deviation (a measure of uncertainty) is considerable lower than the simple average of the 14 GCMs produces; 0.33 versus 1.03. The Bayesian model in effect down-weights the results from GCMs that are furthest away from the NCEP2 results for the 1980–99 time period; the GCMs that have the lowest fidelity in reproducing past climate in this grid-cell.
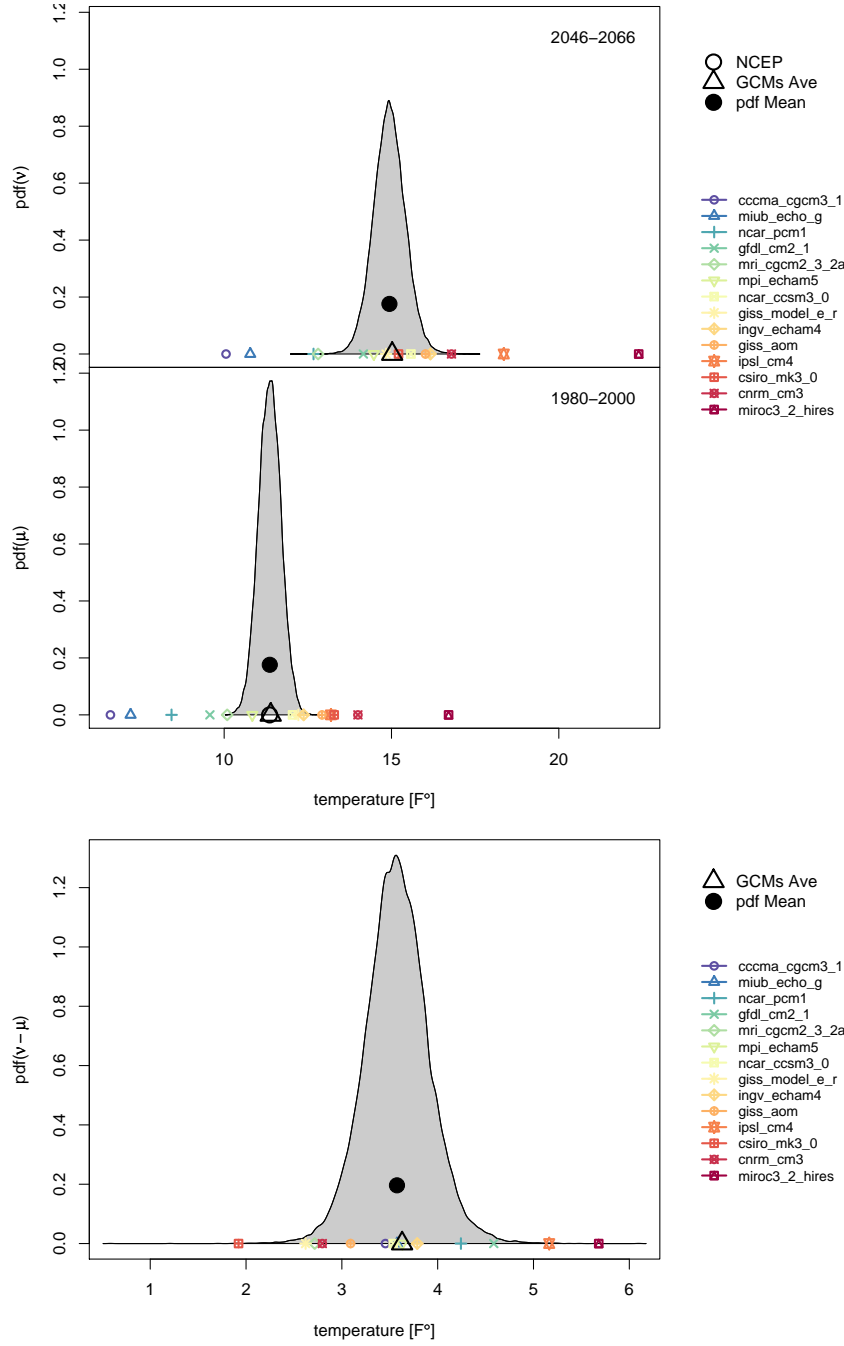
FIGURE 5. The results of the Bayesian model for the average temperature shown in Figure 4(top-left). Shown is the posterior probability density of the average temperature in 2046–65 and for 1980–99 (top). The bottom plot shows the posterior distribution for the average temperature increase from 1980–99 to 2046–65.
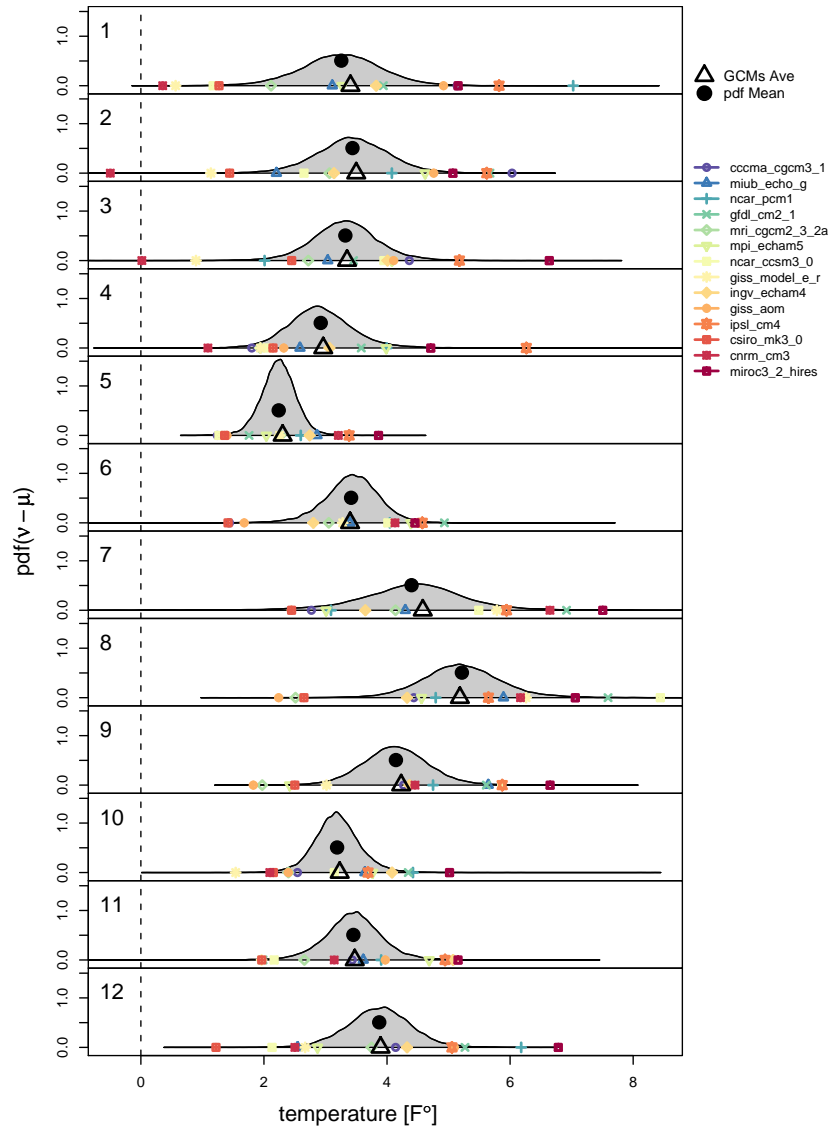
FIGURE 6.   The results of the Bayesian model in estimating the increase
in daily average temperature from 1980–99 to 2046–65 for each month sep-
arately in the $2.5° \times 2.5°$ target region of interest.

In addition to carry out analysis of the annual average temperature for the two time
periods of interest, one can also carry out seasonal analysis and analysis of other climate
variables of interest. Figure 6 shows for example the results of the Bayesian model when
applied to each month separately in assessing the expected increase in daily temperature;
it suggest larger increase in temperature in summer than winter.

5.2. **Local Analysis.** To carry out our local analysis, a MARS model was trained to downscale the NCEP2 from 1980–90 to the Diablo and the Ross sites (independently). The regression model used 45 predictors, derived from five different variables in a 3×3 local grid centered on the locations of interest; the five variables are air temperature at 2 meters, sea level pressure, north- and east-ward wind components at 500mb height, along with air temperature at 500mb. Figure 7 summarizes the downscaled results for daily maximum temperature in terms of averages across the 1980–99 and the 2046–65 time periods for the two sites. Then, as for the large-scale analysis, the Bayesian model was trained on the data to yield a probabilistic characterization of the change in daily maximum temperature at the two sites. The result of the Bayesian model for the Ross site is shown in Figure 8. The mean of the posterior probability distribution of the increase in daily maximum temperature is $3.55°$F, with a standard deviation of 0.40. A simple average of the increase projected by the 12 GCMs involved is $3.58°$F, with a standard deviation of 1.11. Again we see a dramatic decrease in the uncertainty associated with the expected increase. The results for the Diablo site are shown in Figure 9.

6. Conclusion

This report summarizes an initial attempt to account for GCM uncertainty in a SDS framework; that is, move from an ensemble of downscaled results to a joint probabilistic characterization. The initial results are promising. By using a rather simple Bayesian model that has been suggested to fuse together projections from multiple GCMs at the global scale, we have shown that a similar approach can be applied to an ensemble of SDS results.
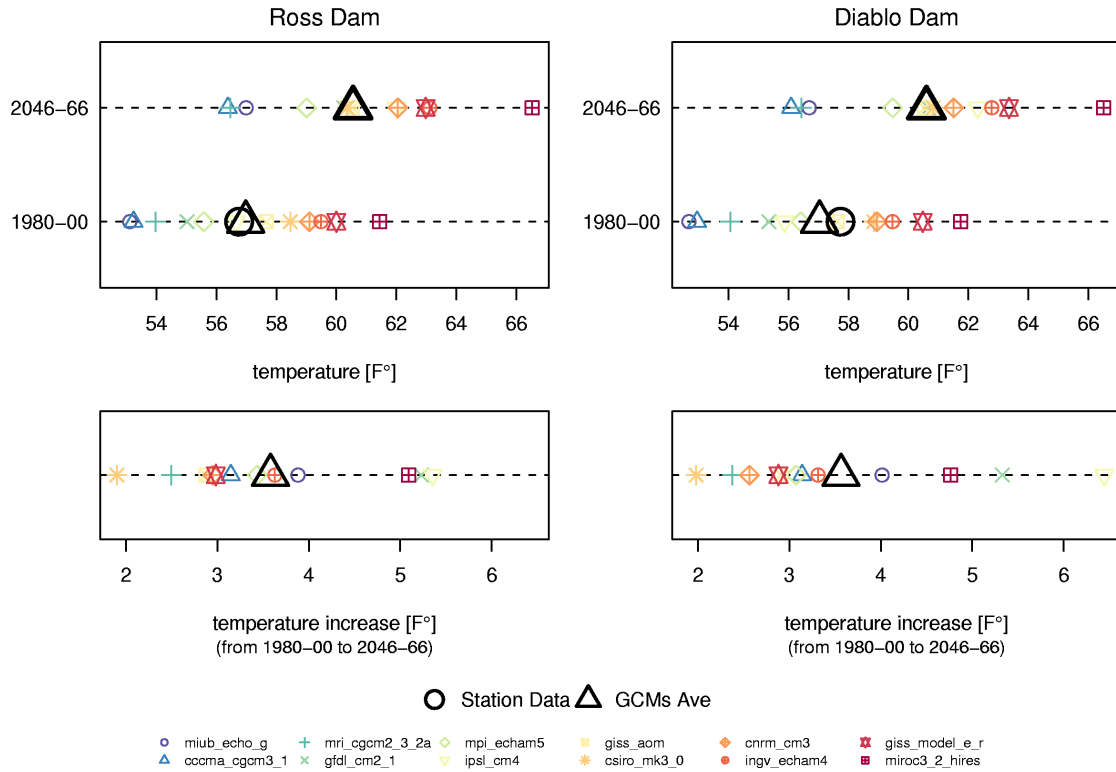
FIGURE 7. A summary of statistically downscaled daily maximum temperature at the Ross and the Diablo sites. Shown is the average of the downscaled daily maximum temperature for the 1980–99 and the 2046–65 time periods (top plots), along with the observed average at the two sites (station data). The bottom plots show the projected increase in the daily maximum temperature.
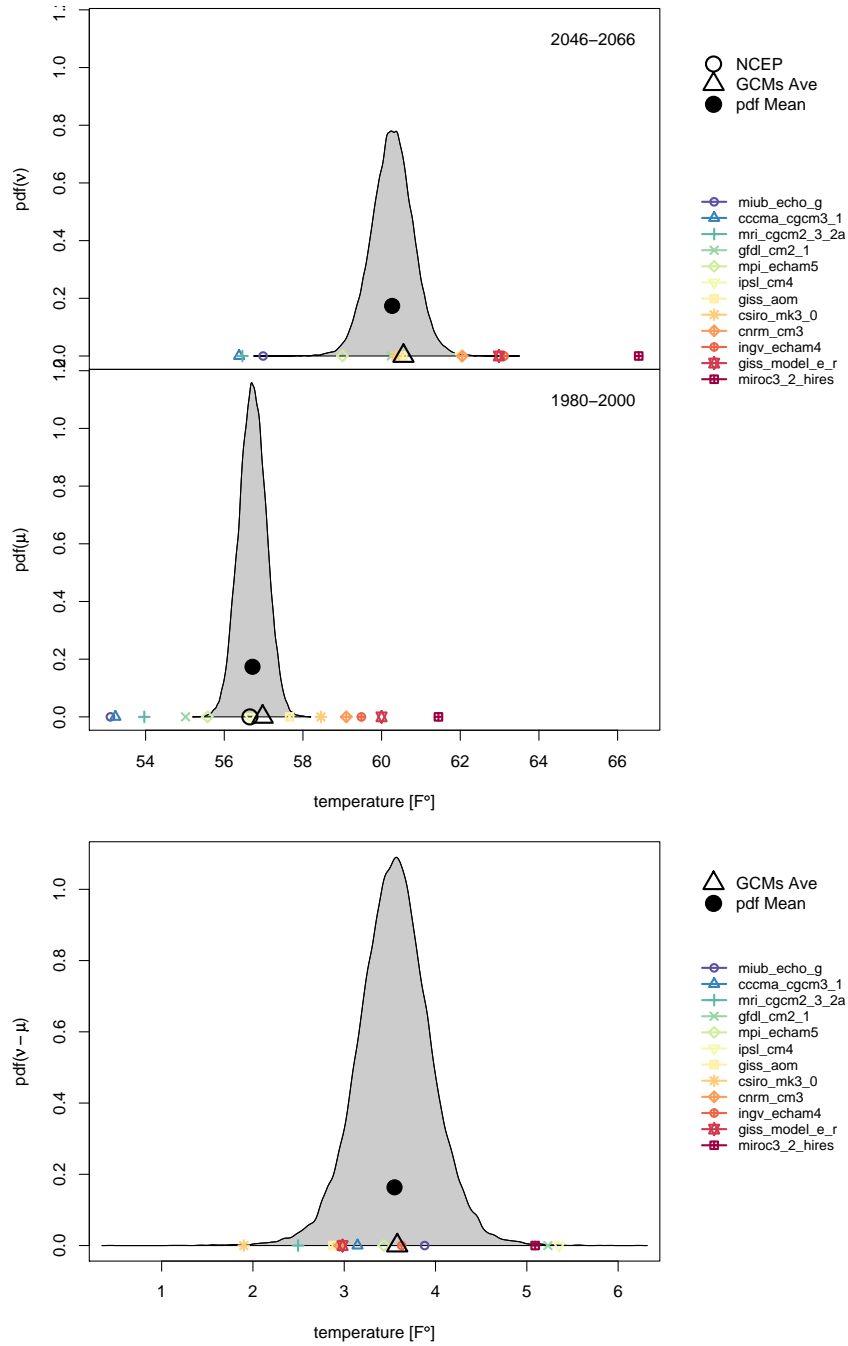
FIGURE 8. The results of the Bayesian model for the daily maximum temperature shown in Figure 7 for the Ross site. Shown is the posterior probability density of the average of the daily maximum temperature in 2046–65 and for 1980–99 (top). The bottom plot shows the posterior distribution for the average increase in daily maximum temperature from 1980–99 to 2046–65.
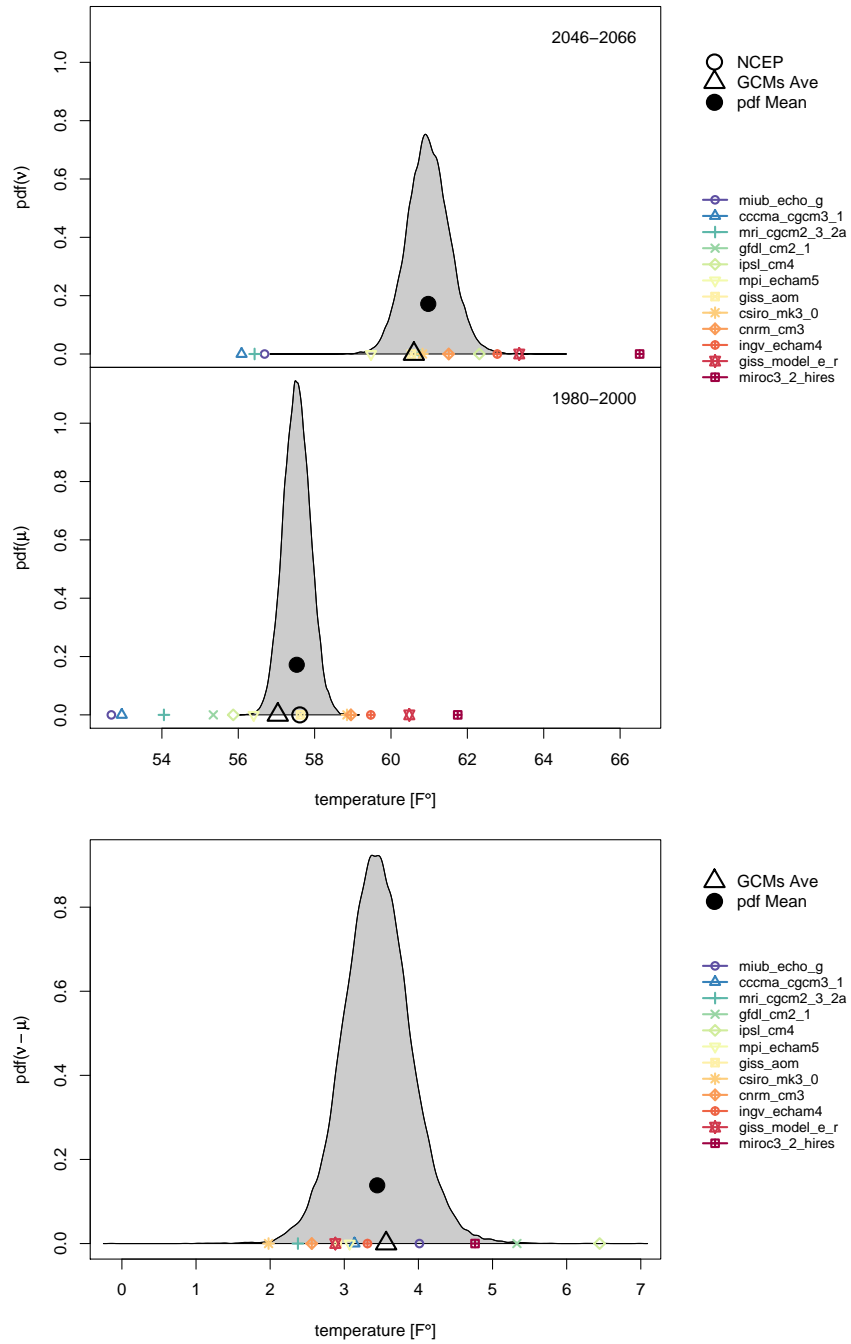
FIGURE 9.   Identical to Figure 8, except for the Diablo site.

REFERENCES

Benestad, R. E., I. Hanssen-Bauer, and E. J. Forland (2007). An evaluation of statistical models for downscaling precipitatio and their ability ot caputre long-term trends. *International Journal of Climatology 27*, 649–665.

Bishop, C. M. (2007). *Pattern Recognition and Machine Learning*. New York: Springer.

Fowler, H. J., S. Blenkinsop, and C. Tebaldi (2007). Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. *International Journal of Climatology 27*, 1547–1578.

Hastie, T., R. Tibshirani, and J. Friedman (2001). *The Elements of Statistical Learning*. New York: Springer. Data mining, inference, and prediction.

Lopez, A., C. Tebaldi, M. New, D. Stainforth, M. Allen, and J. Kettleborough (2006). Two approaches to quantifying uncertainty in global temperature changes. *Journal of Climate 19*, 4785–4796.

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

Robert, C. P. and G. Casella (2004). *Monte Carlo Statistical Methods (Second Edition)*. New York: Springer.

Rougier, J. (2007). Probabilistic inerence for future climate using an ensemble of climate model evaluations. *Climate Change 81*, 247–264.

Smith, R. L., C. Tebaldi, D. Nychka, and L. Mearns (2008). Bayesian modeling of uncertainty in ensembles of climate models. *Journal of the American Statistical Associate X*, X–X.

Tebaldi, C. and R. Knutti (2007). The use of the multimodel ensemble in probabilistic climate projections. *Philosophical Transactions of the Royal Society A X*, X–X.

Tebaldi, C. and B. Sanso (2008). Joint projections of temperature and precipitation change from multiple climate models: A hierarchical Bayes approach. *Journal of the Royal Statistical Society A X*, X–X.

Tebaldi, C., R. L. Smith, D. Nychka, and L. Mearns (2005). Quantifying uncertainty in projections o regional climate change: A Bayesian approach to the analysis of multimodel ensembles. *Journal of Climate 18*, 1523–1540.

Wilby, R. L., S. P. Charles, E. Zorita, B. Timbal, P. Whetton, and L. O. Mearns (2004). Guidlines for use of climate scenarios developed from statistical downscaling methods. *Supporting material to the IPCC*.

Wilby, R. L., T. M. L. Wigley, D. Conway, P. D. Jones, B. C. Hewitson, J. Main, and D. S. Wilks (1998). Statistical downscaling of general circulation model output: A comparison of methods. *Water Resources Research 34*, 2995–3008.

Wood, A. W., L. R. Leung, V. Sridhar, and D. P. Lettenmaier (2004). Hydrologic implications of dynamical and statistical approaches to downscaing cimate model outputs. *Climate Change 63*, 189–216.