Guidelines for Constructing Climate Scenarios

Scientists and others from academia, government, and the private sector increasingly are using climate model outputs in research and decision support. For the most recent assessment report of the Intergovernmental Panel on Climate Change, 18 global modeling centers contributed outputs from hundreds of simulations, coordinated through the Coupled Model Intercomparison Project Phase 3 (CMIP3), to the archive at the Program for Climate Model Diagnostics and Intercomparison (PCMDI; http://pcmdi3.llnl.gov) [*Meehl et al.,* 2007]. Many users of climate model outputs prefer downscaled data—i.e., data at higher spatial resolution—to direct global climate model (GCM) outputs; downscaling can be statistical [e.g., *Maurer et al.,* 2007] or dynamical [e.g., *Mearns et al.,* 2009]. More than 800 users have obtained downscaled CMIP3 results from one such Web site alone (see http://gdo-dcp.ucllnl.org/downscaled_cmip3_projections/, described by *Maurer et al.* [2007]).

A common request from those applying any of these outputs—whether to conduct impact research or to support adaptation planning—is guidance on how to select, treat, and combine the vast amount of climate model output into useful climate scenarios. A scenario is a postulated sequence of events, whether of human development, climate, etc. Specifically, two questions are often asked: (1) How best can scientists understand and characterize uncertainty? (2) What are some key considerations when selecting and combining climate model outputs to generate scenarios? Addressing these questions in the context of recent research leads to some possible guidelines for creating and applying climate scenarios [see also *Knutti et al.,* 2010]. At this juncture, with a new generation of global and regional climate projections becoming available, such guidelines may prove useful to researchers and policy makers.

Understanding and Characterizing Uncertainty

Descriptions of future climate change should include both a central estimate and some representation of uncertainty. Major contributors to uncertainty are imperfect knowledge of (1) the drivers of change, chiefly the sources and sinks of anthropogenic greenhouse gases and aerosols; (2) the response of the climate system to those drivers; and (3) how unforced variability may mask the forced response to drivers.

Quantifying uncertainty in greenhouse gas emissions and other forcings—the drivers of change—remains problematic, and although some studies have attempted to assign probabilities, many instead simply choose among the three forcing scenarios that were widely used for CMIP3. Between now and about 2050 this source of uncertainty is less important than others, because concentration scenarios diverge substantially only after that and because changes before then include a substantially delayed response to previous emissions.

The response of the climate system, the second major contributor to uncertainty, is sometimes characterized by its "climate sensitivity," defined as the change in globally averaged temperature in response to a specified radiative forcing. While this provides a simple characterization based on a single parameter, a full description of response uncertainty would also involve uncertainties in the time-evolving response, and in responses at subglobal scales and of variables other than temperature, which may be proportional to the climate sensitivity, whether on global or regional scales.

Climate sensitivity can be estimated from observations [e.g., *Hegerl et al.,* 2007], but these estimates are subject to uncertainties in both forcing and response. It is hard to rule out very high rates of warming: most studies estimate that there is at least a 5% chance that the sensitivity exceeds 7–9°C for a doubling of atmospheric carbon dioxide (CO_2). Some of these studies account for uncertainties in aerosol forcing. Model estimates of climate sensitivity, on the other hand, range only from 2.1°C to 4.4°C [*Randall et al.,* 2007]. No climate model in the CMIP3 archive represents a low-likelihood, high-sensitivity future climate.

The third important source of uncertainty—how unforced variability masks effects by known drivers of climate change—involves the fact that historical climate simulations do not, and are not intended to, reproduce the exact monthly values of climate variables. A goal of developing scenarios is to distinguish the slowly varying central tendency of change forced externally (by greenhouse gases, volcanoes, etc.) from the unforced variations, which can be important, even dominant, when trying to diagnose and interpret climate change on small time and space scales in the context of global simulations [*Hawkins and Sutton,* 2010]. Using climate projections for impact assessments depends on being able to separate forced responses from natural climate variability [e.g., *Giorgi,* 2005], which is often accomplished by analyzing the mean and range in an ensemble of simulations differing only in initial conditions.

One thing to note on the uncertainty in climate projections is that on the regional to local scale, where effects are felt, studies may include extremes like cold or heat, storms, and droughts, and detection and attribution of such changes to specific causes (e.g., rising greenhouse gases) becomes more difficult. Consequently, estimating uncertainty in future changes in these local quantities has little theoretical basis.

Further, it must be emphasized that the range of available model results does not, and is not intended to, represent the true physical uncertainty of the quantity in question, although many studies implicitly assume that it does. The range of model results measures consensus, which is important but distinct from uncertainty. Some work on parameter-space exploration has explicitly attempted to quantify the physical aspects of uncertainty [see, e.g., *Stainforth et al.,* 2005].

As a final comment on sources of uncertainty in climate projection information, it is important to understand that the relevance of these sources to a given decision depends on the climate variable and scale of interest (in both space and time). Consideration of which climate aspects are most relevant to a given planning or decision-making process (variables and scales) will help steer attention toward associated aspects of climate projection information, which can lead to a more tailored and relevant discussion of these uncertainties.

Selecting and Combining Models

To distill the large number of model simulations into a small group of scenarios, it seems logical to focus on simulations that seem more credible, culling or weighting the results on the basis of some measure of skill. Weighting models may be justified when, for instance, there is a strong correlation between a physical process and a performance metric [*Knutti et al.,* 2010]. Furthermore, while many efforts have focused on ranking climate models based on how they simulate the time-averaged regional climate during a historical period [e.g., *Gleckler et al.,* 2008; *Brekke et al.,* 2008], for impact assessments, in particular, a better basis for model ranking might be their ability to simulate regional climate sensitivity to a change in global climate forcing, provided that a theoretical and observational basis for such analysis can be established.

While methods have varied, it is common to use historical model performance to weight or to choose the "best" models when constructing an ensemble. Some studies have been framed on the premise that ranking leads to better results, though it has been shown that model ranking depends on which skill metrics are considered [*Gleckler et al.,* 2008; *Brekke et al.,* 2008]. In any case, while some studies have shown that ranking models has led to a separation in future responses [e.g., *Walsh et al.,* 2008], others have shown that considering metrics of model skill has generally made little difference either to detection and attribution studies or for representing likely future change. For example, for future average temperature over the western United States, any 14 randomly selected GCMs produced results indistinguishable from those produced by a combination of the "best" models, and the ensemble skill approached the same asymptote once any 6 GCMs were included [*Pierce et al.,* 2009]. Further, using a metric of precipitation trend, 11 randomly selected GCMs produced results almost identical to those using the 11 "best" GCMs [*Knutti et al.,* 2010], and detection and attribution of changes in atmospheric water vapor were insensitive to whether the "best" or "worst" 10 GCMs were used [*Santer et al.,* 2009]. Additionally, little reduction was found in estimating regional precipitation and temperature change uncertainty over northern California [*Brekke et al.,* 2008] or the Pacific Northwest [*Mote and Salathé,* 2010] when based on different sets of "better" climate models, as illustrated in Figure 1. On the basis of these findings and focusing on CMIP3 results, it is unclear whether model culling leads robustly to a separation of future responses and is thus warranted in planning efforts. However, this topic will need to be revisited with CMIP Phase 5 when new GCM simulations will be available to establish performance metrics that may be more robust [*Knutti et al.,* 2010].

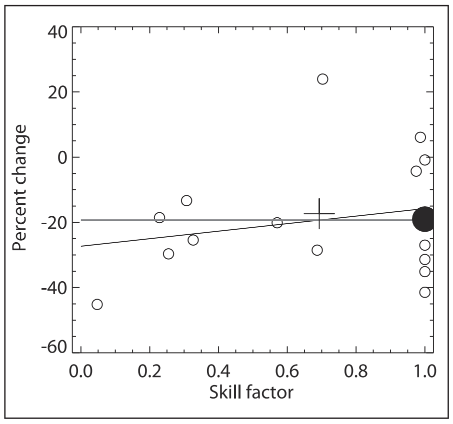

Fig. 1. Projected change (in percent) in summer precipitation for the 2080s in the U.S. Pacific Northwest from a variety of climate models (open circles, as used by Mote and Salathé *[2010]), for scenario A2 of the Intergovernmental Panel on Climate Change's Special Report on Emissions Scenarios. The* x *axis shows the bias factor of* Giorgi and Mearns *[2002]; models with simulated 1970–2000 precipitation close to the observed precipitation, within the range of natural variability, are given a skill factor of 1. Linear fit to the data is indicated (sloping line). There is little difference among changes calculated with all models unweighted (horizontal line), with only the "best" models (models with skill factor >0.9, solid circle), or with weighting the models by their skill factor (plus sign).*

Whether or not models are culled, scenario development requires decisions on what to sample from the available ensemble. Some may focus on changes in mean climate, in which case it may be advisable to define such change based on a multimodel average rather than on any single model. However, such definitions still need to be blended with assumptions about climate variability, which may be taken from past observations. Alternatively, the ensemble of opportunity—that is, all the available model runs (as distinct from runs designed to form a meaningfully representative ensemble)—may also be used to estimate changes in both mean and variability. Further work is needed to quantify the credibility of CMIP3 and the new CMIP5 output on various space and time scales, beyond assessing relative skill and culling models, as discussed above. For estimating the central tendency or selecting a single "best" model, then, a suitable approach may be simply to take an unweighted average or median result based on as many models as possible.

In summary, and based on the evaluations cited above, it seems justifiable to forgo culling or weighting climate projections based on perceptions of credibility. This leaves a rather large ensemble of opportunity that may be sampled for climate scenario information. Such sampling may involve identifying individual climate projections that express changes that generally represent the spread of projection information, or choosing a scheme that combines projection information (e.g., ensemble median projected condition through time, or ensemble mean change in period statistics). When

Climate Scenarios cont. on page 258

By P. Mote, L. Brekke, P. B. Duffy, and E. Maurer

cont. from page 257

several simulations from the same model are available, important questions to ask involve whether differences between outputs of the same model are as large as differences between outputs of different models for various starting parameters. The answer depends on the space and time scales considered, but several studies suggest the answer is that time-averaged differences between outputs of the same model are negligible, especially for longer time horizons [e.g., *Pierce et al.*, 2009]. This implies that the formation of a large ensemble of model simulations [e.g., *Maurer et al.*, 2007] should recognize that two runs from the same model are not likely to be as different as two runs from different models, and therefore one should not simply lump all available simulations together, as this effectively gives more weight to the models contributing more simulations.

Proposed Guidelines for Model Evaluations

Results from new evaluations of models including CMIP5 (see http://cmip-pcmdi.llnl.gov/cmip5/) and the North American Regional Climate Change Assessment Program [*Mearns et al.*, 2009] are arriving, along with new downscaled data repositories. Volunteers are also contributing time on their personal computers to create a superensemble of regional climate simulations at 25-kilometer resolution for the western United States (see http://www.weatherathome.net). While these new efforts augment the options of climate scenarios available, they also complicate the development of climate scenarios.

Because modeling efforts both new and old can be difficult to navigate, the following guidelines may help scientists and managers who intend to use climate model scenarios for impact or climate diagnostic research:

1. Understand to which aspects of climate your problem or decision is most sensitive (e.g., which climate variables, which statistical measures of these variables, and at what space and time scales).
2. Determine which climate projection information is most appropriate for the problem or decision (e.g., variables, scales in space and time).
3. Understand the limitations of the method you select.
4. Obtain climate projections based on as many simulations, representing as many models and emissions scenarios, as possible.
5. It may be worth the effort to evaluate the relevant variables against observations, just to be cognizant of model biases, but recognize that most studies have found little or no difference in culling or weighting model outputs.
6. Understand that regional climate projection uncertainty stems from uncertainties about (1) the drivers of change (e.g., greenhouse gases, aerosols), (2) the response of the climate system to those drivers, and (3) the future trajectory of natural variability.
7. Use the ensemble to characterize consensus not only about the projected mean but also about the range and other aspects of variability.

These guidelines make use of several recent research efforts and may provide a better foundation for developing and applying climate scenarios to a range of research and planning questions.

References

Brekke, L. D., M. D. Dettinger, E. P. Maurer, and M. Anderson (2008), Significance of model credibility in estimating climate projection distributions for regional hydroclimatological risk assessments, *Clim. Change, 89*(3-4), 371–394, doi:10.1007/s10584-007-9388-3.

Giorgi, F. (2005), Interdecadal variability of regional climate change: Implications for the development of regional climate change scenarios, *Meteorol. Atmos. Phys., 89*(1-4), 1–15, doi:10.1007/s00703-005-0118-y.

Giorgi, F., and L. O. Mearns (2002), Calculation of average, uncertainty range, and reliability of regional climate changes from AOGCM simulations via the "reliability ensemble averaging" (REA) method, *J. Clim., 15*(10), 1141–1158, doi:10.1175/1520-0442(2002)015<1141:COAURA>2.0.CO;2.

Gleckler, P. J., K. E. Taylor, and C. Doutriaux (2008), Performance metrics for climate models, *J. Geophys. Res., 113,* D06104, doi:10.1029/2007JD008972.

Hawkins, E., and R. T. Sutton (2010), The potential to narrow uncertainty in projections of regional precipitation change, *Clim. Dyn., 37*(1-2), 407–418, doi:10.1007/s00382-010-0810-6.

Hegerl, G. C., F. W. Zwiers, P. Braconnot, N. P. Gillett, Y. Luo, J. A. Marengo Orsini, N. Nicholls, J. E. Penner, and P. A. Stott (2007), Understanding and attributing climate change, in *Climate Change 2007: The Physical Science Basis—Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by S. Solomon et al., pp. 663–745, Cambridge Univ. Press, New York.

Knutti, R., R. Furrer, C. Tebaldi, J. Cermak, and G. A. Meehl (2010), Challenges in combining projections from multiple climate models, *J. Clim., 23*(10), 2739–2758, doi:10.1175/2009JCLI3361.1.

Maurer, E. P., L. Brekke, T. Pruitt, and P. B. Duffy (2007), Fine-resolution climate projections enhance regional climate change impact studies, *Eos Trans. AGU, 88*(47), 504, doi:10.1029/2007EO470006.

Mearns, L. O., W. Gutowski, R. Jones, R. Leung, S. McGinnis, A. Nunes, and Y. Qian (2009), A regional climate change assessment program for North America, *Eos Trans. AGU, 90*(36), 311, doi:10.1029/2009EO360002.

Meehl, G. A., C. Covey, T. Delworth, M. Latif, B. McAvaney, J. F. B. Mitchell, R. J. Stouffer, and K. E. Taylor (2007), The WCRP CMIP3 multimodel dataset: A new era in climate change research, *Bull. Am. Meteorol. Soc., 88*(9), 1383–1394, doi:10.1175/BAMS-88-9-1383.

Mote, P. W., and E. P. Salathé Jr. (2010), Future climate in the Pacific Northwest, *Clim. Change, 102*(1-2), 29–50, doi:10.1007/s10584-010-9848-z.

Pierce, D. W., T. P. Barnett, B. D. Santer, and P. J. Gleckler (2009), Selecting global climate models for regional climate change studies, *Proc. Natl. Acad. Sci. U. S. A., 106*(21), 8441–8446, doi:10.1073/pnas.0900094106.

Randall, D. A., et al. (2007), Climate models and their evaluation, in *Climate Change 2007: The Physical Science Basis—Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change,* edited by S. Solomon et al., pp. 589–662, Cambridge Univ. Press, New York.

Santer, B., et al. (2009), Incorporating model quality information in climate change detection and attribution studies, *Proc. Natl. Acad. Sci. U. S. A., 106*(35), 14,778–14,783, doi:10.1073/pnas.0901736106.

Stainforth, D. A., et al. (2005), Uncertainty in predictions of the climate response to rising levels of greenhouse gases, *Nature, 433,* 403–406, doi:10.1038/nature03301.

Walsh, J. E., W. L. Chapman, V. Romanovsky, J. H. Christensen, and M. Stendel (2008), Global climate model performance over Alaska and Greenland, *J. Clim., 21*(23), 6156–6174, doi:10.1175/2008JCLI2163.1.

Author Information

Philip Mote, Oregon Climate Change Research Institute, Oregon State University, Corvallis; E-mail: pmote@coas.oregonstate.edu; Levi Brekke, U.S. Bureau of Reclamation, Denver, Colo.; Philip B. Duffy, Climate Central, Palo Alto, Calif.; and Ed Maurer, Department of Civil Engineering, University of Santa Clara, Santa Clara, Calif.