Reply

—T. N. Palmer, F. J. Doblas-Reyes,
A. Weisheimer, and M. J. Rodwell
European Centre for Medium-Range Weather Forecasts,
Reading, United Kingdom

Picture this: you are advising your regional government on infrastructure investment for climate adaptation. Increased reservoir capacity may be needed if your region is likely to become drier over the years, while better flood defenses and bigger storm sewers are necessary if the local climate will become wetter. Your regional government wants a proper risk assessment for each scenario. The notion of risk combines the economic and social consequences of some specific climate change with the probability that this climate change will occur. The economists and social scientists have estimated the consequences of not having sufficient water capacity if the climate becomes significantly drier, or of not updating flood defenses if it becomes wetter; your task is to estimate the probability that it will indeed become either drier or wetter.

DOI:10.1175/2009BAMS2916.1

An obvious source of data to obtain such probabilities is the third phase of the Coupled Model Intercomparison Project (CMIP-3) multimodel ensemble, used in the Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report (AR4; information online at www-pcmdi.llnl.gov/ipcc/about_ipcc.php). Here, for example, one can readily estimate changes in the frequency of occurrence of precipitation events for your particular region of interest resulting from various scenarios of future greenhouse gas concentrations. However, a fundamental question is whether such frequencies of occurrence can be interpreted as probabilities in the type of risk assessment outlined above.

This question raises a yet more fundamental question: what do we mean by probability in this context? A probability forecast should accurately describe uncertainties in the outcomes, given uncertainties in the various inputs. In the case of climate model prediction, key input uncertainties are associated with the way the underlying partial differential equations (PDEs) of motion are represented computationally, for example, in using traditional bulk formula parameterization to represent subgrid-scale physical processes.

Does the CMIP-3 multimodel ensemble properly sample these uncertainties? The fact that probability distributions from such an ensemble are manifestly inaccurate when verified against twentieth-century climate (see Palmer et al. 2009, their Fig 2) indicates that, at some level, such a multimodel ensemble does not properly represent underlying uncertainties in the representation of subgrid-scale processes. This is hardly surprising; the CMIP-3 ensemble has not in any way been designed to sample the key uncertainties arising from the truncation of the PDEs of climate. Moreover, it is not at all clear how one would design such a multimodel ensemble as a matter of principle.

What to do? The CMIP-3 frequencies of occurrence somehow have to be discounted to take account of the fact that the multimodel ensemble is not a complete representation of model uncertainty, but we need some objective way of doing this discounting. Your regional government is pressing you for your input!

Palmer et al. (2008, hereafter P08) have suggested that estimates of seasonal forecast reliability provide a direct way to calibrate multimodel climate change frequencies of occurrence, thereby making these frequencies of occurrence more appropriate to the type of risk analysis discussed above. Scaife et al. (2009, hereafter S09) disagree, feeling that seasonal variability is irrelevant to the longer-term climate change issue. Unfortunately, S09 have misunderstood both the procedure and the underlying scientific basis for the discounting technique in P08. We would like to take this opportunity to explain these procedures again.

S09 begin by commenting that P08 "suggested that the same chain of physical processes is responsible for both seasonal variations and long-term climate change." We are not claiming that the two chains are the same, but rather that they have links in common.

Take the following simple example: Models used for long-term climate change prediction and shorter-term seasonal climate forecasting incorporate parameterizations of deep convection. Deep convection itself is a rather fast time-scale physical process whose effects are crucial to both accurate seasonal forecasts and accurate predictions of climate change.

On the other hand, predictions of long-term climate change are also sensitive to cryospheric and biogeochemical processes, which are not important for seasonal prediction. However, just because these biogeochemical and cryospheric processes occur on relatively long time scales, it would be wrong to conclude that they are therefore independent of the faster time-scale processes that are central to seasonal forecasting. For example, the extent to which the rainforests may become a carbon source for the atmosphere in the future will depend on changes in the patterns of seasonal-mean rainfall, and simulations of this will therefore depend on the representation of convective rainfall processes in climate models. Hence, if the representations of the fast time-scale processes in seasonal forecasts are found to not be fully reliable (in certain regions and at certain time of year), is not this information relevant to the problem of converting raw frequencies of occurrence into useful probabilities? Of course, the long-term climate models' representations of the biogeochemical and cryospheric processes might also be unreliable, but this is not a matter that seasonal prediction studies can address. P08's suggestion of using seasonal forecast information should be considered necessary but not sufficient for the provision of reliable probabilities.

Hence, when S09 claim that "there is growing evidence that seasonal and centennial climate changes can operate by different mechanisms," are they really saying the role of deep convection in transporting heat, moisture, and momentum in the atmosphere is somehow only relevant on short time scales and not on long time scales?

It hardly seems possible that S09 really do believe this, so it is more likely that they have misunderstood the discounting procedure proposed in P08. A clue that this is the case lies in their comment:

> Both initial conditions and omission of greenhouse gas increases lead to errors in the seasonal hindcasts used by P08. These errors inevitably reduce the reliability of seasonal forecasts, but they do not affect centennial climate projections where greenhouse gases are specified to increase and initial condition information is long forgotten.

There are two issues raised by this quote. We deal with the less important one first.

In principle we agree that omission of greenhouse gas increases could lead to seasonal forecast unreliability, but the study by Doblas-Reyes et al. (2006) showed that this would not lead to unreliability in either circulation or precipitation forecasts. Precipitation forecasts, not temperature forecasts, are the focus of attention in PO8.

S09's comment about initial conditions is by far the more important. The word "reliability" has already been used several times in this response. Many people may use the word "reliable" as an alternative for "skilful." However, in probability forecasting, these two words are not at all synonymous and it is crucially important in understanding the argument in P08 to note that we are not using seasonal forecast "skill" as the metric that determines the strength of calibration. Through their comment about the role of initial conditions, we believe S09 may have fallen into this trap.

The effect of initial condition uncertainty on the forecast can be found by running an ensemble with representations of initial error and studying the consequent ensemble spread. If the spread is sufficiently large that forecast probabilities are no different from climatological probabilities (for a particular variable and a particular region), then the corresponding forecasts will be reliable but have no skill.

However, in P08 it is certainly not proposed that one should discount climate change frequencies of occurrence just because initial data errors have driven the seasonal forecast ensemble to model climatology; that would make no sense at all! Rather, we only propose to discount when the ensemble spread is sharper than a climatological distribution and where verification suggests that the ensemble is overconfident in predicting such a sharp probability distribution. In a single-model ensemble, the cause of such overconfidence does not lie in the representation of errors in the initial data, but in the total neglect of model error. For a multimodel ensemble, the use of different climate models introduces an ad hoc measure of model uncertainty, but it is an incomplete estimate of the true model uncertainty (e.g., in representing the process of deep convection) associated with the models used. Hence, for example, we can see that overconfident ensemble predictions of seasonal precipitation in parts of Europe are due in part to the failure of models to simulate anticyclonic blocking adequately; this in turn seems to be related to a range of model shortcomings, not the least of which is representation of deep convection. (In passing, we note that S09 are simply wrong to state that "many seasonal hindcasts used in P08 do not even initialize year-to-year sea ice changes at all in their forecasts.")

Hence, even though seasonal prediction is partly an initial value problem and climate change prediction is (to first order) not, by focusing on the reliability of the seasonal forecasts rather than their skill we can extract useful information from the seasonal forecasts that is relevant for the climate change problem.

Short of waiting for climate change to happen, how can we move this discussion forward? Here is one idea: We know that forecasts of precipitation climate change can be sensitive to the resolution of the climate model used (both horizontal and vertical resolution). Hence, consider the following: let us label the precipitation climate change signal from a high-resolution model as "truth" and the climate change signal from a low-resolution model as "model". In some regions of the world truth and model will agree, and in other regions they will disagree. Now, let us perform some seasonal "forecasts" using the low-resolution model, which is initialized and verified using the high-resolution model. Our claim is that variations in the regional reliability of these seasonal "forecasts" will be a predictor of the difference between the climate change signal's truth and model. Using simple cost–loss models of potential economic value, one could even verify, in this idealized context, the claim that calibrated probabilities would make for better decisions than uncalibrated frequencies of occurrence.

The basic claim in P08 is that the CMIP-3 frequencies of occurrence, calibrated by the unreliability from seasonal forecasts, will be better for the type of (multibillion dollar) decision making described above than are the raw uncalibrated frequencies of occurrence. However, we do not claim that our proposed calibration is optimal. Studies of the type suggested in the paragraph above may indicate more optimal methodologies for calibration than that suggested in P08. The key issue that P08 wish to raise, however, is that there is a resource, provided by seasonal forecasting, that is not currently being used sufficiently by the climate change prediction community.

One final comment on S09: by looking at the coldest and driest winters over the United Kingdom, S09 claim that the projected climate change signal is too large to be discounted by any moderate increase in blocking. However, blocking is not only manifested in terms of the very long-lived extreme events. An increase of the Tibaldi and Molteni (1990) blocking index by a factor of 3 would shield all Atlantic storms from the United Kingdom and drive many of them south to the Mediterranean regions, which are currently predicted to get drier in almost all of the CMIP-3 models. By contrast, if blocking became less commonplace in the future, then Atlantic storms might be more commonplace over the United Kingdom in summer, which is also currently predicted to become drier under climate change.

Climate change is one of the most serious long-term threats facing humanity in this and coming centuries. The current generation of comprehensive climate models is our best tool for estimating the nature of this threat, and they show a risk of very severe climate consequences unless we can significantly cut emissions of greenhouse gases. The need to continue to warn government and society of these risks has never been greater.

On the other hand, climate prediction is among the most computationally demanding problems in science. Because of the consequent need to approximate certain key processes in the computation representation of climate, the current generation of climate models is manifestly imperfect, for example, in simulating modes of intraseasonal and interannual variability. This points to deficiencies in representations of fast time-scale processes that are crucial in determining long-term climate.

We need better computational infrastructures to improve the models, but this may take time; regional governments need information to assess how to adapt to inevitable climate change now! We believe it would be irresponsible not to take account of all possible information in providing such advice. In P08 we make one quantitative suggestion in this direction.

REFERENCES

Doblas-Reyes, F. J., R. Hagedorn, T. N. Palmer, and J.-J. Morcrette, 2006: Impact of increasing greenhouse gas concentrations in seasonal ensemble forecasts. *Geophys. Res. Lett.,* **33,** L07708, doi:10.1029/2005GL025061.

Palmer, T. N., F. J. Doblas-Reyes, A. Weisheimer, and M. J. Rodwell, 2008: Toward seamless prediction: Calibration of climate change projections using seasonal forecasts. *Bull. Amer. Meteor. Soc.,* **89,** 459–470.

——, ——, ——, G. J. Shutts, J. Berner, and J. M. Murphy, cited 2009: Towards the Probabilistic Earth-System Model. [Available online at http://aps.arxiv.org/abs/0812.1074.]

Scaife, A., C. Buontempo, M. Ringer, M. Sanderson, C. Gordon, and J. Mitchell, 2009: Comment on "Toward seamless prediction: Calibration of climate change projections using seasonal forecasts." *Bull. Amer. Meteor. Soc.,* **90,** 1549–1551.

Tibaldi, S., and F. Molteni, 1990: On the operational predictability of blocking. *Tellus,* **42A,** 343–365.