

Contributors

Robert Pincus, *NOAA-CIRES Climate Diagnostics Center*; Crispian Batstone, *NOAA-CIRES Climate Diagnostics Center*; Robert J. P. Hofmann, *University of Colorado and NOAA/ESRL/PSD*; Karl E. Taylor, *Lawrence Livermore National Laboratory*; Peter J. Gleckler, *Lawrence Livermore National Laboratory*

Research Highlight

In the thirty years since the advent of short-term weather forecasts with numerical models, the accuracy of those forecasts has improved steadily. Accuracy is measured by waiting until the forecast time arrives, then computing the degree to which the forecast matches the observations using a set of agreed-upon “skill scores.” Projections of future climate change have been made for nearly as long as weather forecasts, but the models used for these predictions have not been evaluated routinely, in part because the prediction times are so far in the future. It is possible, though, to test the degree to which a climate model reproduces the statistics of the present-day climate, and to expect that models that do a better job in today’s world will also show skill in projecting coming changes.

A recent ARM-supported paper defines skill scores appropriate for climate models, focusing on objective measures of skill in reproducing the present-day distribution of clouds, precipitation, and the effect of clouds on the earth’s energy balance - precisely those aspects of the climate that the ARM Program seeks to measure and improve. The scores are computed by comparing simulations of the last two decades of the twentieth century with multi-year or multi-decade sets of global satellite observations.

Skill scores are computed for all models that participated in the Fourth Assessment of the Intergovernmental Panel on Climate Change (IPCC), and several results emerge. First, although some models lead the pack in many measures, every model has its weak spots. The model that does best overall is, in fact, not a single code but rather the “IPCC mean model,” constructed by averaging the fields produced by all models in the sample. The mean model does so well mostly because errors in individual models are distributed on both sides of the observations. In addition, the degree to which models disagree with the observations is much larger than the amount by which independent sets of observations disagree, which means that there is room for models to improve before observational accuracy becomes a limit.

The climate model skill scores reported in this paper allow for model improvements to be tracked over time. In the present-day set of models no measure of skill is connected to the climate sensitivity, so the ability to measure model skill doesn’t immediately narrow the range of uncertainty in climate projections. On the other hand, it’s difficult to imagine accurate projections of future change coming from a model that does a poor job in simulating the present climate - and now there’s a way to measure the success of a model at doing the latter job.

Reference(s)



Measuring Climate Model Skill in Producing Present-Day Clouds

Pincus, R, CP Batstone, RJP Hofmann, KE Taylor, and PJ Glecker. 2008. "Evaluating the present-day simulation of clouds, precipitation, and radiation in climate models." *Journal of Geophysical Research* 113, D14209, doi:10.1029/2007JD009334.

Working Group(s)
Cloud Modeling

