EarthCube Science Requirements: Disseminating Multi-Model Seasonal Forecast Data

On Behalf of the US National Multi-Model Ensemble Intraseaonal to Interannual Prediction Group

Ben Kirtman
University of Miami - Rosenstiel School for Marine and Atmospheric Science
University of Miami - Center for Computational Science

Jin Huang
National Oceanic and Atmospheric Administration
Climate Prediction Center

The EarthCube vision is to enable data and information sharing, foster multi-discipline collaborations and catalyze knowledge dissemination. In this white paper, we provide a use case scenario that provides a "real life example" of a multi discipline, multi institutional collaboration that the EarthCube design should address.

Weather and climate predictions are necessarily uncertain. The uncertainty comes from two major sources:

(i) Initial conditions uncertainty associated either with observing system errors or the way in which observational estimates are used to initialize prediction systems (model uncertainty and errors play a significant role here);
(ii) Uncertainties in the formulation of the models used to make the predictions and to assimilate the observations. These uncertainties are associated with the discrete representation of the continuous climate system and the parameterization of sub-grid physical processes.

Users of forecast information are multidisciplinary (e.g., water resource managers, energy providers, …) with varying levels of computational sophistication, yet all require predictions with minimal uncertainty accompanied by reliable estimates of that uncertainty. There are a number of different techniques for assessing the uncertainty of intraseasonal, seasonal and interannual (ISI) predictions due to initial conditions uncertainty, and these typically involve perturbing the initial condtions. Quantifying prediction uncertainty due to uncertainty in model formulation is often estimated by making ensemble predictions from different models that are then combined to produce a forecast probability distribution. This approach is the basis for the newly emerging US National Multi-Model Ensemble (NMME) intraseasonal-to-interannual prediction system (http://www.cpc.ncep.noaa.gov/products/ctb/). This ongoing effort is multi-institutional (i.e., University of Miami, Center for Ocean-Land-Atmosphere Studies, NASA, NOAA-GFDL, NOAA-CPC, NCAR, Princeton University, University of Colorado, International Research Institute for Climate and Society) and involves six (6) major modeling efforts. Indeed, US NMME is in response to a recent US National Academies "Assessment of Intraseasonal to Interannual Climate Prediction and Predictability" (NRC 2010[1]), which was unequivocal in recommending the need for the development of a US national multi-model ensemble operational predictive capability. This national effort is required to meet the specific tailored regional prediction and decision support needs of the emerging

[1] http://www.nap.edu/catalog.php?record_id=12878

National Climate Service. The challenge is to meet this National need without diluting existing model development activities at the major centers and ensure the forecast products continue to improve and be of societal value.

The societal value of this NMME effort will be measured in terms of how forecast information enhances decision-making – can the physical science be translated into societal benefit? In terms of this NMME consortium, the value assessment will hinge on two “make or break” challenges: (i) the skill or quality of the forecasts themselves and (ii) the ease at which comprehensive forecast information can be made readily available to a disparate collection of users. Both challenges are formidable and the first is aggressively being addressed at modeling centers and in research groups world wide. The second challenge is far from being adequately addressed. This is because the user needs encompass a very wide range of space and time scales, and a heterogeneous set of variables and fields leading to an enormous data volume that cannot be served by any single institution – the centralized approach in not sustainable or scalable in the longer term. Perhaps, distributed model ideas such as those described by Mader et al. (position paper under CI) should be explored in greater detail.

The sharing or distributing both real-time forecasts and the retrospective or historical forecasts for developing decision support application tools is a daunting challenge, will necessarily require a networked or distributed approach, sophisticated sub-setting tools and on-demand processing and visualization. This need for a distributed approach is well recognized, and some leveragable software and hardware tools (e.g., THREADS servers) are available for distributed databases; nevertheless, this distributed approach to serving user application needs has not succeeded in the weather and climate community. All examples of success in sharing data for user applications have been with centralized databases. It is our assertion here that this lack of success in the distributed approach is because the distributed nature of the data is not transparent to the user - they are forced to go to multiple distribution sites, acquire inhomogeneous data that lacks sufficient standards and uniformity for easy use in application models.

The NMME project noted above, however, is perhaps the ideal environment to demonstrate success in a distributed database for application/decision support using forecast information. First, the NMME consortium has already demonstrated the ability to provide their data in a common format, the forecast protocol is uniform, and the updating of the forecast information in real-time is done in a consistent manner each month. Essentially, there is sufficient uniformity (i.e., agreed data format and protocol) so that software and server interfaces can be built such that the user does not see the distributed data, and significant inhomogeneity that this distributed approach is the only long term viable solution for making the data available to a multidisciplinary research and application community of forecast information users.