Toward a National Multi-Model Ensemble (NMME) System for Operational Intra-Seasonal to Interannual (ISI) Climate Forecasts

ABSTRACT
This white paper summarizes the background, the need, and proposed implementation strategy for a national multi-model ensemble (NMME) system for operational intra-seasonal to interannual (ISI) climate prediction. The proposed strategy includes implementation of an experimental NMME protocol and designing and testing a future protocol for operational NMME prediction system. The protocols will leverage existing activities and resources to achieve an NMME capability quickly, and it will be sufficiently flexible to evolve as it guides research, development, transition to operations, and evaluation. By building a network of communications and collaborations among NMME partners and stakeholders, and by identifying and applying best practices, the protocols will also seek to optimize the utility and value of ISI climate forecast products.

Background

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 necessarily discrete representation of the continuous climate system and the parameterization of sub-grid physical processes.

Users 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. These techniques involve defining "optimal" procedures for perturbing the initial conditions. For example, the singular vector approach has been used for ENSO predictability studies (Moore and Kleeman 1999; Kleeman and Moore 1997). This approach involves determining the stochastic "optimals" of non-normal systems, which can be calculated from either the true linear dynamic operator (Moore and Kleeman 1996; Kleeman and Moore 1997) or from an empirically derived propagator using observational data or coupled model output (Penland and Sardeshmukh 1995; Kleeman et al. 2003). Alternatively, the bred vector approach (Toth and Kalnay 1996; Cai et al 2003 among many others) views error growth from the perspective of non-linear chaotic dynamics, whereas the stochastic optimals result from the linear dynamics. Nevertheless, many operational intra-seasonal to seasonal prediction systems use ad hoc methods, e.g., initial conditions based on observations that are separated by a short time interval, to probe the initial conditions uncertainty problem.

Quantifying prediction uncertainty due to uncertainty in model formulation falls into two general categories that are complementary. The first category is an *a posteriori* approach where ensemble predictions from different models are combined to produce a forecast probability distribution. This approach is the basis for several international collaborative prediction research efforts (e.g., DEMETER[1], CHFP[2], APCC[3]), the operational EUROSIP system[4] and there are numerous examples of how this multi-model ensemble (MME) approach yields superior forecasts compared to any single model (Kirtman and Min 2009; Jin et al. 2008; Hagedorn et al. 2005; Doblas-Reyes et al. 2005; Palmer et al. 2004; Kirtman et al. 2002; among many others). For example, Fig. 1 (taken from Kirtman and Min 2009) shows the global distribution of the point correlation for the predicted SSTA and the observed estimates predicted using the Community Climate System Model version 3 (CCSM3) and the NOAA/NCEP Climate Forecast System version 1 (CFSv1) and the MME formed by averaging the two with equal weights. The forecasts are initialized in January and are verified for the following June (i.e., 5-month lead). For this particular lead-time and initial condition, the CCSM forecasts appear to be more skillful in the tropical Pacific, and much of this skill is captured in the MME. Conversely, the CFS has higher skill in the Atlantic and the MME appears to also capitalize on this skill, with hints of even higher skill. The correlation is comparable between CCSM and CFS in the Indian Ocean and is maintained in the MME. Figure 2 shows an example from the DEMETER project indicating that a multi-model ensemble has better skill scores (ranked probability skill score is this case) than a single model (ECMWF) of the same total ensemble size. Multi-model ensembles by virtue of using different data assimilation and initialization systems also represent initial condition uncertainty, albeit in an ad-hoc manner.

Typically, these multi-model efforts are multi-institutional. This multi-institutional aspect is important - the participating models tend to be developed and verified independently increasing the likelihood that they have complementary or independent skill. It is this complementary skill that is the scientific basis for the multi-model approach, although quantifying the complementary skill in various models remains an open research question. Moreover, the multi-institutional aspect means that the core development efforts are conducted at different laboratories and centers minimizing the risk of diluting core efforts to improve model fidelity.

The second category of methods for quantifying prediction uncertainty can be viewed as an *a priori* technique in the sense that the model uncertainty is "modeled" as the prediction evolves. The models include a significant stochastic element in their formulation and are often referred to stochastic-dynamic models (Palmer 2001; Fleming 1971). In fact, Palmer and Williams (2010) argue that, for climate prediction, the "dynamic and thermodynamic equations have irreducible uncertainty" and that the models should include a stochastic component representing the small scales. The stochastic physics approach (Buizza et al. 1999; Palmer 2001) is perhaps the best known and is currently in use at ECMWF for generating the 41-member seasonal forecast ensemble[5]. An alternate approach has been suggested by Berner et al. (2008) who developed a semi-stochastic cellular automaton backscatter scheme to represent energy at small scales and its feedback to larger scales in a coupled model. They performed seasonal prediction experiments and showed that the inclusion of the backscatter scheme significantly improved the probabilistic skill of seasonal forecasts and reduced the systematic error. These internal (i.e., within a particular center) *a priori* techniques need to be perused further and are naturally complementary and can be easily combined into a multi-institutional multi-model approach.

---

[1] http://www.ecmwf.int/research/demeter/
[2] Climate-system Historical Forecast Project - http://www.clivar.org/organization/wgsip/chfp/chfp.php
[3] Asia-Pacific Climate Center - http://www.apcc21.net/climate/climate01.php
[4] http://www.ecmwf.int/products/forecasts/seasonal/documentation/eurosip/index.html
[5] http://www.ecmwf.int/products/forecasts/seasonal/documentation/system3/ch2.html

The Need for a US National Multi-Model ISI Prediction Capability

The recent US National Academies "Assessment of Intraseasonal to Interannual Climate Prediction and Predictability" (NRC 2010[6]) was unequivocal in recommending the need for the development of a US national MME operational predictive capability. Indeed, the national effort is required to meet the specific tailored regional prediction and decision support needs of the emerging 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.

There is little doubt that US participation in EUROSIP is beneficial to both the US and European forecasting communities and the users of the forecasts. However, as a US National Climate Service emerges and as the possible National Center for Predictions and Projections (NCPP) develops, the need for a NMME system becomes paramount for supporting continued research on MME based prediction that can transition to operations. For example, a NMME system facilitates modifications (e.g., extending the forecast to longer time-scales) to the forecast strategy, allows for better coordination of the forecast runs compared to EUROSIP (e.g., hindcast period, forecast scheduling etc.) and allows free exchange of data beyond what is supported by EUROSIP. Also, by testing various national models on weather and seasonal time-scales, the NMME system will accelerate the feedback and interaction between US ISI prediction research, US model development and the decision science that the forecast products support. For instance, the prediction systems can potentially be used to evaluate and design long-term climate observing systems, because US scientists will have open access to the prediction systems (i.e. data, data assimilation and forecast models). Our national interests require that we (1) run these ISI prediction systems operationally in the US, (2) retain the flexibility to modify the prediction systems and how they are used based on emerging national needs, and (3) ensure that there is a robust communication and collaboration network open among operational ISI forecasting, research and model development.

Critical Issues

At workshops organized by the NOAA Climate Test Bed on the February 18 and April 7-8, 2011, several key research and operational prediction groups met to discuss the current status and ways forward toward a National MME capability. Representatives of the groups presented the current status and near-term plans for their respective ISI prediction activities. In addition to the status report, the groups were asked to discuss current research gaps, which include:

- how best to combine models
- how to quantify uncertainty due to model formulation
- what are the sources of complementary skill
- how best to combine perturbed physics, perturbed initial conditions, stochastic physics and multi-model techniques
- how best to use the forecast results for coordinating model development and improvement
- how to use the forecast results for observing system design and assessment
- how best to use the NMME for decision support how to measure the success of the NMME

Developing A US Multi-Model Prediction System

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

The NOAA Climate Test Bed (CTB) has supported several "proof-of-concept" projects that have been focused on developing products and tools that *could* transition to NOAA operations. While the CTB provided the resources for developing the scientific underpinnings and technological advances, the actual transition to operations has been, in most cases, neither proposed nor supported. This white paper argues that several US National research groups and laboratories are ideally positioned to make rapid progress in developing an operational US National multi-model ISI predictive capability.

The prediction research and development activities to date have been, for the most part, ad hoc and only loosely organized. In this sense, the research and development needed to achieve a true National MME has not been sufficiently rigorous or oriented toward that mission. This process was described at the NMME workshop as an "MME of opportunity". What is needed is a "purposeful MME" process in which the requirements for operational ISI prediction are used to define the parameters of a rigorous reforecast experiment and evaluation regime.

Building on the successes of the CTB and the broader community, the CTB NMME workshops on February 18 and April 8, 2011 recommended to take a phased strategy to collaboratively develop a US MME ISI prediction system.

Phase-1 of this recommended NMME strategy to develop an experimental NMME system (i.e., MME of Opportunity), leverages existing CTB and external community ISI activities and is being implemented (view the Phase-1 NMME Implementation Plan at: http://www.cpc.ncep.noaa.gov/products/ctb/projects/NMME_FY2011WorkPlan.pdf).

Phase-2 of the recommended NMME strategy is to design and test an operational NMME protocol (i.e., a purposeful MME) that is to guide the future research, development and implementation of the NMME beyond what can be achieved in Phase-I, i.e., an "MME of opportunity". This protocol should be a "living" document that evolves as technology and understanding improve and needs change. The protocol should include specific details for the near term (i.e., FY2012) and guidance for the long-term. In other words, this protocol will evolve into a NMME seasonal-to-interannual prediction implementation plan.

a. This protocol must be designed to specifically take into account operational forecast requirements and address the transition to operations. For example, does the reforecast evaluation regime include metrics that address user requirements? Also, once the reforecasts are completed and the MME evaluation has been done, how will the operational forecasts be made available to the operational prediction entity (NOAA/CPC or the anticipated National Climate Service) on schedule? Can the operational prediction entity allow for the possibility that sometimes some of the models participating in the operational MME may not be readily available or that some members will be produced on a different schedule?

b. The protocol must have specific details regarding the number of cases, number of ensemble member and procedures for sharing all of the reforecast data among the CTB partners and the general climate research and applications community. It is essential that robust quantitative estimates of forecast quality are made and that the larger research community be engaged in the evaluation of the forecasts and the application and use of the forecasts.

c. As recommended in NRC report, the protocol must include specific details regarding sharing the data assimilation systems and the prediction models to the

CTB partners and the broader research community in order to support interactions among operations, model development and research groups. Comprehensive user support is highly desirable, although it may not be possible in the near term due to fiscal constraints.

d. The protocol must include research foci that are identified as high priorities and reviewed periodically to assess progress and changes in requirements. The protocol must also allow for flexibility and evolution in identifying and addressing high priority research needs.

e. The protocol must include details regarding the (human and computational) resource requirements to complete the transition to an operational MME, how those resource requirements will be met and who are the key points of contact in each partner organization for the partnership.

f. The protocol will recognize that the NMME must evolve as models, understanding, technology and application requirements change. The protocol must identify processes whereby the broader research and forecast use communities can inform the future of NMME development (e.g., model development, metrics, model combinations, forecast products, …). The protocol will also identify strategies for how to incorporate new versions of the participating prediction systems and/or new prediction system in the operational NMME.

g. The protocol must also include discussion of regional/user requirements, as they are the "pull" for the NMME. How these requirements can evolve based on evolving needs and target of opportunity also needs to be addressed.

The protocol should include a data dissemination/coordination strategy for hindcasts and real time forecasts.

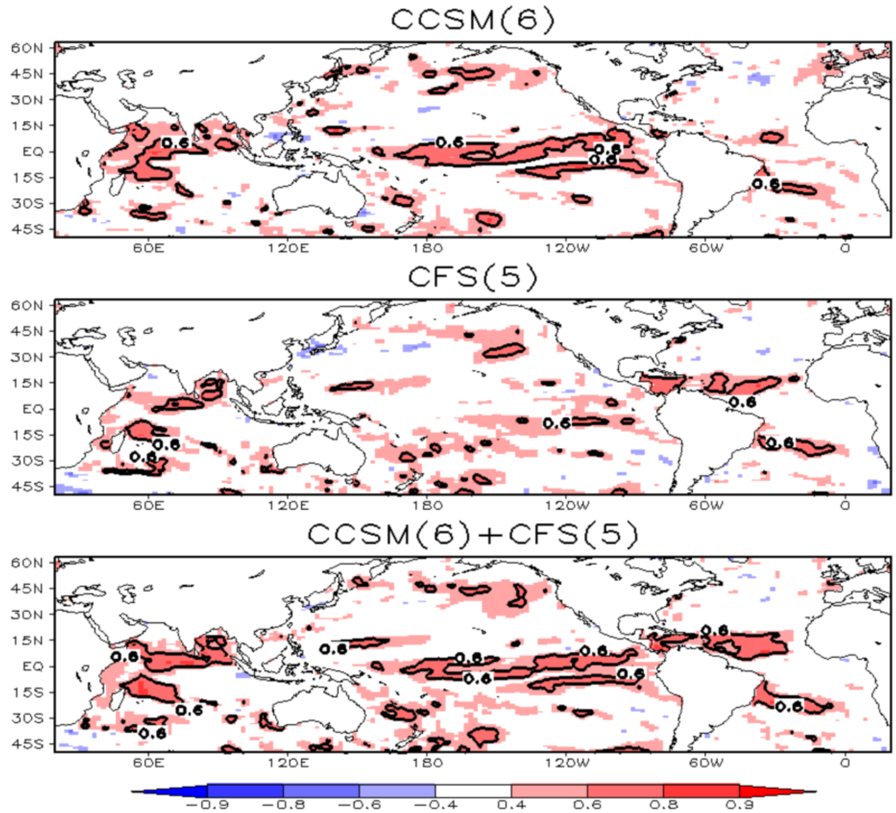

Figure 1: SSTA local point correlation for the (a) CCSM ensemble (6-members), (b) the CFS ensemble (5-members) and (c) the multi-model ensemble (11-members). The plots correspond to a lead-time of five months and January initial conditions (taken from Kirtman and Min 2009).

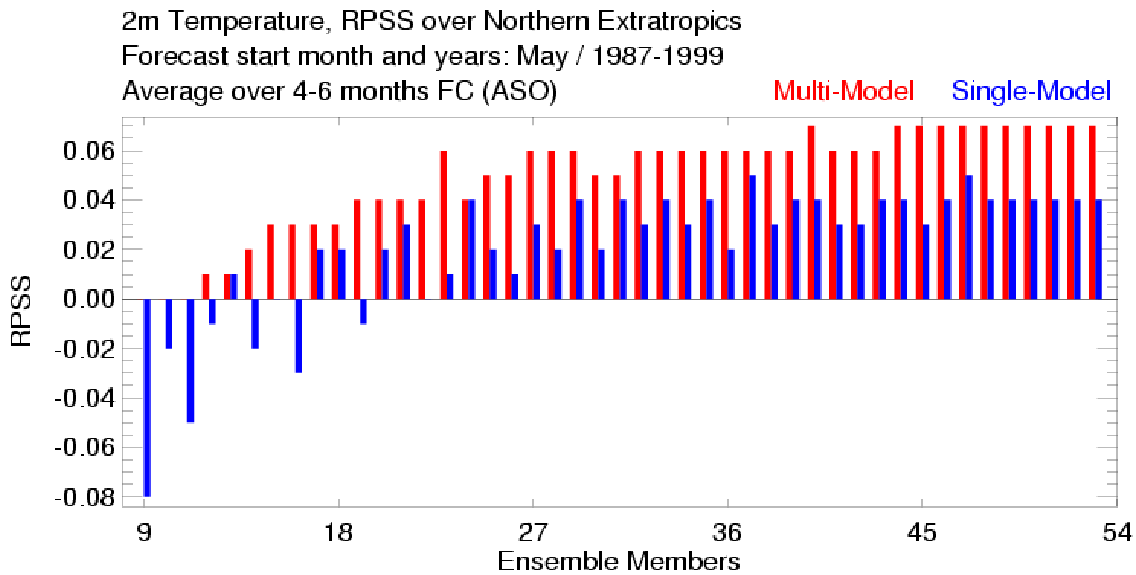

Figure 2: Comparison of DEMETER multi-model versus Single Model (ECMWF) Ranked Probability Skill Scores of 2m-Temperature over the Northern Hemisphere Extratropics (30-90N) for reforecasts initialized in May 1987-1999 verifying in August-October (i.e., 4-6 month lead).