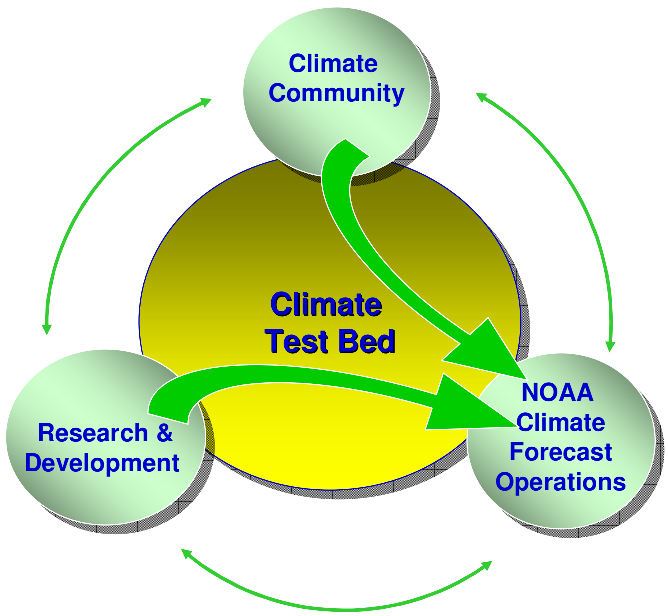
NOAA Climate Test Bed

Multi-Model Ensembles: Transition from Research to Operations and Implementation Strategy



<u>Mission:</u> to accelerate the transition of research and development into improved NOAA operational climate forecasts, products, and applications.

October 2006

Table of Contents

Page

Executive Summary **3**
1.0 Introduction **4**
2.0 CTB Activities **4**
  2.1 Science Priorities **4**
  2.2 Computing and Data Access **6**
  2.3 Current Activities **6**
    2.3.1 International MME **6**
    2.3.2 National MME **8**
    2.3.3 Applications and Extensions of MME **9**
    2.3.4 Calibration and Reforecasting **9**
    2.3.5 Staffing **10**
    2.3.6 Computing Resources **10**
  2.4 Enhancing Scientific Community Involvement **10**
  2.5 Milestones **11**
3.0 References **12**

Executive Summary

The Climate Test Bed is in close alignment with the mission and goals of NOAA. Subseasonal and seasonal climate predictions are central to the mission of NOAA, and are strongly supported by Climate Test Bed (CTB) activities. The CTB is working with a world-class coupled prediction model, the Climate Forecast System (CFS), and is accelerating improvements in routine climate forecasts and climate forecast products that are valued by the user community.

The CTB has an opportunity to increase the utility of seasonal forecasts by defining a process for introducing a Multi-Model Ensemble (MME) prediction system at NCEP. Since there is evidence that MME-based forecasts yield improvements over single model ensembles, accelerating the transition of MME-based forecasts to forecasters and MME-based products and services to end users is a high priority.

The CTB will sponsor the MME project to demonstrate the value added by composing the MME with CFS included. CTB will pursue international collaborations with EUROSIP, APCC and possibly others. These organizations make operational forecasts and have (or will have) the required hindcasts plus considerable expertise and experience. CTB will also collaborate with national centers developing 1-tier coupled models (including GFDL, NASA, and NCAR).

The CTB will pursue both experimental activities as well as a deliberate transition to operations of an MME Prediction System with all requirements necessary for producing high quality forecast products. Current science priority areas include preliminary evaluations of MME forecast skill, comparisons of simple linear combination algorithms to more complex weighted combinations, and the generation of hindcast datasets for effective calibration and application of reforecasting techniques. A future emphasis on the development of MME-based products is anticipated.

In order to pursue MME forecasts more quickly and effectively, CTB staff with the necessary expertise will be dedicated to the MME projects. The CTB will also enhance communication and collaboration between CPC, EMC and the external (research and user) community. Additional computer resources will be allocated to the MME efforts. Evaluation of the return on the allocation of these resources to the MME activities will be closely monitored by CTB management.

This White Paper recommends a MME implementation strategy, including alternatives. It also indicates some specific activities currently underway at NCEP. While MME appears to offer significant potential for improving climate forecasts, there is also the possibility that, in practice, MME will not lead to substantial improvements in operational climate outlooks. For this reason, the CTB will provide frequent updates on the progress of its MME activities for comments and recommendations.

1.0 Introduction

Studies by Krishnamurti et al. (1999), Palmer et al. (2004), and others have provided evidence that the forecast skill of a MME system is higher than that of the individual models. In response, the international scientific community has rallied around the MME approach to improve climate prediction. While the earlier studies suggest a way forward, it remains to determine the extent to which this approach will help CPC operational forecasts.

Several multi-model ensemble programs are underway around the world, notably those at EU (DEMETER / EURO-SIP) and at the APEC Climate Center in Korea (APCC). To maintain its status as a world leader in global Earth system modeling, NOAA should develop a systematic multi-model based prediction capability and infrastructure. The NOAA/CTB is a natural lead to develop the strategy and accelerate the transition to operations. Such a capability will allow focused research on phenomena that have been demonstrated by predictability studies to have the greatest potential to improve forecast skill on intraseasonal-to-interannual (ISI) time scales. Increased availability of improved fully coupled models, access to data sets from multi-model experiments, and experimental prediction activities will result in new capabilities for the research community at large to contribute to predictability studies and an understanding of climate variability and change based on intercomparisons of model and observational data sets. These capabilities include an improved dynamical understanding of trends that are one major component of the physical basis for operational seasonal forecasts (especially when ENSO is absent).

By FY12 NOAA will develop an Earth system modeling capability that provides an increased range of climate products for regional applications and decision support. The future vision is centered around a capability to produce a seamless suite of products that span operational climate predictions, based substantially on output from multi model ensembles with significant utilization of Earth system models, and extending to a suite of new forecast products of impacts on the environment and ecosystems with quantification of skill level at global and regional scales.

2.0 CTB Activities

2.1 Science Priorities

NOAA's Climate Prediction Center has the responsibility for the Nations official 6-10 day, 8-14 day, monthly and seasonal forecasts. Given that the MME approach offers the greatest potential for increased skill on these time-scales, the NOAA CTB has elevated the Multi-Model Ensemble (MME) effort to a high priority. The MME effort will have a two-prong approach involving both international (section 2.3.1) and national (section 2.3.2) efforts. There will be an attendant enhancement of CTB staff (section 2.3.5) and computer resources (section 2.3.6) allocated to the MME efforts.

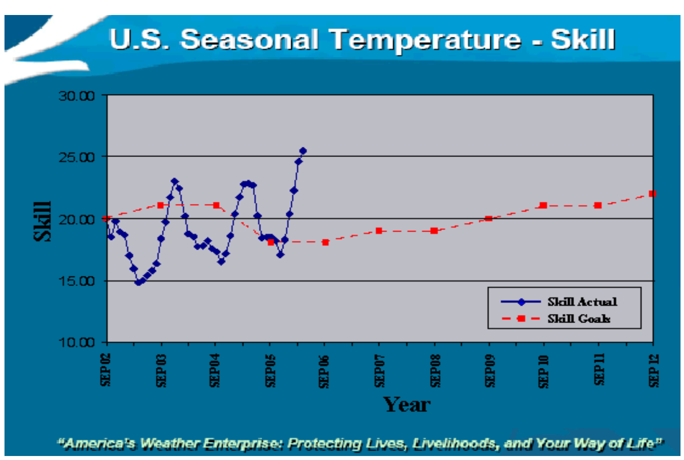
The CTB will sponsor the MME project to demonstrate the value added by composing MMEs. Both EMC and CTB will carry out the development and technical evaluation of the prototype MMEs. CPC and CTB will evaluate the prototypes and final products when they are candidates for operational implementation by EMC and especially NCEP Central Operations (NCO). CTB will also continue to emphasize CFS improvements and climate forecast products for decision support as important science priorities.

The CTB MME strategy will evolve into a more detailed and coordinated science and implementation plan, enumerating the "how", the "why", and identifying the constraints and the trade space the CTB is working within (e.g. available computer resources and FTEs). CTB Management will continue to vet the MME strategy with CTB personnel, the OB, the CST and the SAB to ensure that CTB science is not conducted in an ad-hoc manner. CTB plans will include the project management structure, specific outputs (milestones, deliverables), and a description of how they support the overall CTB goals. CTB plans for MME will be incorporated into the broader CTB Science and Implementation Plan document.

The CTB is aggressively pursuing parallel efforts for an international MME as a near-term strategy (see section 2.3.1) and a national MME as a longer-term strategy (see section 2.3.2). Ultimately, a consolidated MME Prediction System that combines all known sources of independent skill is envisaged. The quality of the MME Prediction System may depend on the level of sophistication of the consolidation scheme that is used to combine these sources (see section 2.3.1 for preliminary results). The CTB has established near-term and provisional long-term milestones for the MME prediction system (see section 2.5).

A recent string of very skillful forecasts in CPC forecast operations (Fig. 1) can be attributed, at least in part, to efforts to consolidate tools during FY06 (van den Dool, 31st CDPW). Ongoing CTB efforts to consolidate tools and to add tools that bring independent skill to the climate forecast (including MME) may improve these forecasts further and help to reset the base level of skill.

Skill Goal – vs – Actual Skill



Courtesy Mike Halpert

Figure 1. Skill goal –vs- Actual skill for U.S. seasonal temperature. (0.5 month lead; 4-year running means)

2.2 Computing and Data Access

In order to transition a MME Prediction System to NCEP, the CTB will have to address major challenges in the area of computing and data access. This will require the development of a coherent strategy for the allocation, monitoring, administration and evaluation of the CTB shared computing facility. In developing its strategy, the CTB will

- Consult with the NCEP Computing Oversight Board (Lord (Chair), Laver, Toepfer, Cooley) on internal computer resource matters;
- Consult with NOAA CIO Office on management of the CTB portion of the new NOAA R&D computer;
- Request that the NCEP Computing Oversight Board consult with external experts (e.g. CSL and SCD/NCAR) on the strategy for allocation and monitoring of CTB computer resources;
- Work with the CST Co-Chairs to gather and evaluate quarterly progress reports from the Transition Project Teams on computer usage tied to CTB projects;
- Vet and document any recommended improvements in the CTB strategy for the shared computing facility.

CTB will work to ensure additional improvements to climate forecasts and related products, and that data sets are distributed. The CTB will continue to review the needs of the stakeholder and applications communities to determine if its data policy adequately addresses those needs.

2.3 Current Activities

As mentioned earlier, the CTB has adopted a two-pronged approach, involving both international and national MME activities.

2.3.1 International MME

In the near term the CTB will focus on an analysis of the EUROSIP datasets in an effort to determine if CFS adds value to a European multi-model ensemble and vice-versa. The CTB is particularly interested in the set of dynamical coupled models run by the operational centers participating in this program (UKMO, Meteo-France, and ECMWF). Because each center has produced an associated hindcast dataset, these activities are less resource intensive for CTB than those in which hindcasts are not available (see section 2.3.2). Experimental collaborative projects (using the existing hindcast datasets) followed by a deliberate transition to operations are envisaged.

CTB will add CFS to the EUROSIP models to determine whether there is benefit in terms of probabilistic skill scores. Seasonal forecast skill will be evaluated over Europe and the US for both temperature and precipitation using a multi-model ensemble consisting of the 3 European models (EURO3) and a multi-model ensemble consisting of the 3 European models plus CFS (IMME). The evaluation will be carried out using equal weighting as well as more sophisticated consolidation techniques (e.g., ridge regression). CTB will also investigate the trade space between the length of the training period (hindcast dataset) and improvements in the skill of the ensemble forecasts.

A preliminary comparison of EURO3 and IMME using equal weighting for the period 1981-2001 (van den Dool, Saha and Johansson, personal communication) showed that NCEP CFS contributes to the skill of IMME (relative to EURO3) for equal weights (especially in terms of the probabilistic Brier score and for precipitation) over Europe, in the US and in the global tropics (20°S-20°N). Also, when the skill of a particular model is low, consolidation of forecasts (based on a-priori skill estimates) will reduce the chance that the model will be included in the IMME, and thus may lead to improvements in the skill of the IMME as obtained from equal weighting. The preliminary results warrant consideration of a European-US IMME product for seasonal prediction.

The results also motivate the use of more sophisticated consolidation techniques (such as ridge regression) or best model approaches (which improve as the number of models increases) in the IMME. In comparison to ECMWF, METFR and UKMO, the CFS as an individual model does well in deterministic scoring (AC) for precipitation and very well in probability scoring (BS) for precipitation and temperature over both Europe and the US.

In a similar manner, the CTB plans to evaluate hindcasts generated by the APEC Climate Center (APCC), KMA, BMRC, and BCC models towards expanding the IMME to include these models. Again CTB is most interested in the fully-coupled dynamical models that have generated appropriate hindcast datasets (in accordance with the CTB Science Plan and Implementation Strategy). Seasonal forecast skill will be evaluated over regions of interest to the participating institutions and over North America for both temperature and precipitation using ensembles with and without CFS. Collaboration between CTB staff (Dr. Jae Schemm, Dr. Wanqui Wang), KMA (Dr. Chung Kyu Park), BMRC (Dr. Oscar Alves), and BCC (Dr. Peiqun Zhang) on the exchange of hindcast datasets and the skill evaluation is underway.

Pending the outcome of the preliminary skill evaluations and our consolidation procedures, CTB will pursue more formal collaboration with participating international institutions for the exchange of operational models and data necessary to run an IMME in an operational mode. If it is useful, the EUROSIP, APCC and possibly other MMEs will also be combined into a single IMME.

Recently, NCEP in collaboration with the Korea Meteorological Administration has commenced work towards meeting operational requirements as Lead Centers for Long-Range Forecast Multi-Model Ensemble predictions. The WMO Secretariat has prepared relevant documentation and is currently considering a recommendation to encourage participating centers to provide their seasonal hindcast and forecast data. As part of its verification effort, CTB will also collaborate with CPC to submit CFS forecast verification results to the Lead Center for Verification of Long Range Forecasts, hosted by the Australian Bureau of Meteorology and the Meteorological Service of Canada. The verifications will be done globally, as well as for specific regions such as Europe and North America. The Korean Meteorological Agency (in partnership with NCEP) is also willing to host such a center, and progress is being made on this.

2.3.2 National MME

CTB is working to establish a systematic community based MME forecasting capability and infrastructure using coupled National models (NCEP-CFS, GFDL-CM2.1, NASA-GEOS, NCAR-CCSM, others). The CTB has established "Research to Operations (R2O) Guidelines" (see Appendix A of SP&IS) that specify the path to implementation of such a capability in the NCEP operational climate model suite. These guidelines establish the roles and responsibilities of the "home research institution" and NCEP (EMC, CPC, CTB) in this process. Significant resources are required both at the home research institution and at NCEP (EMC, CPC, CTB) to carry out smooth transitions.

During FY07 the CTB is working with GFDL to bring its model into the NMME framework. In particular, the CTB MME Team (see section 2.3.5) is carrying out the following experiments:

1. Complete a preliminary skill evaluation of GFDL coupled model hindcasts (carried out at GFDL; 10-member ensemble for IC's Oct-Nov and Apr-May over the period 1981-2006). Determine whether the GFDL model contributes additional skill to the CFS forecasts for these months.

2. If the preliminary skill evaluation shows that there is additional skill, then port the GFDL system to the NCEP R&D computer for reproducibility testing.

3. Repeat hindcasts (as outlined in 1 above) at NCEP using the imported GFDL system to ensure reproducibility of the climate state (ensemble mean), and that the consolidation of these new forecasts with CFS adds skill to the CFS.

While carrying out the activities above, CTB will continue to seek partnerships with the other agencies (i.e. NASA, NCAR) using the progress with GFDL as a model. In general, the national activities are more resource intensive for CTB than the international ones (section 2.3.1) because the models under consideration are not operational and the requisite hindcast datasets are not available for the skill evaluation. As a consequence, the national activities are longer-term efforts.

If it is useful, the NMME will be combined with the IMME (section 2.3.1) towards a Multi-Model Ensemble Prediction System at NCEP that takes advantage of all known sources of independent skill. Assuming that the models have been calibrated and consolidated correctly, the CTB will incorporate the MME into CPC forecast operations. The skill of CPCs operational consolidation tool will be compared to the skill of the individual tools (including the MME) and to CPC's official forecasts. As part of this evaluation, we will consolidate dynamical forecasts and statistical forecasts separately and then combine them.

2.3.3 Applications and Extensions of MME

The MME prediction system can be used to improve products from the Land Data Assimilation Systems (LDAS) being implemented at NCEP (Noah, Vic, Sac, and Mosaic). It is well known that hydrologic variables from LDAS depend on the model used and the input data. While the climatologies differ, the anomalies (defined as departures from the climatological mean) show more similarities. Thus, the MME can be used to test whether the ensemble mean eliminates uncertainties, hence improving the LDAS products.

The products from the MME will be extended to hydrologic variables and will be used to monitor and predict hydrologic events (droughts and floods) over the United States. Forecasts of drought indices and / or variables such as the Standardized Precipitation Index (SPI) will be used for drought prediction. The SPI has been widely used to monitor drought; the only input required is precipitation, which is readily available from the MME. Because drought extends to seasonal and longer time scales, statistical tools will be combined with dynamical tools, including the MME for drought prediction. If 6-month, 12-month, and 24-month forecasts of the SPI have skill, then we will use this approach to predict other variables (e.g. soil moisture or evaporation).

An important issue to resolve is that daily / weekly output from Demeter / APCC / CFS are needed for the hydrologic forecasts. Also, forecasts of the SPI (and other drought indicators) may require use of high-resolution Global Forecast System output for weeks 1 and 2, and then appended CFS or MME output thereafter.

2.3.4 Calibration and Reforecasting

In order to maximize the skill of probabilistic forecasts, they should be calibrated by correcting errors determined from prior forecasts and observations. For many forecast problems (e.g. seasonal precipitation forecasts) an effective calibration is much more difficult without a long time series of past forecasts from the same model that is run operationally.

Recent work at NOAA/OAR/ESRL has demonstrated that calibration based on reforecasts can improve probabilistic forecast skill. This applies to probabilistic forecast products produced using ensemble forecasts.

Since reforecast-based statistical corrections of ensemble forecasts have been demonstrated to dramatically improve the skill and reliability of probabilistic forecasts generated from ensembles, CTB will consider a reforecast-based approach.. This approach will leverage the talents and capabilities extant at NCEP, MDL, and OAR.

This effort should help NOAA prepare to regularly produce and utilize reforecast-based products. The result will be state-of the art, calibrated, skillful probabilistic forecasts. These efforts will also help to put MME in a larger Weather-Climate probabilistic perspective.

2.3.5 Staffing

In order to carry out the activities described above, the CTB has restructured its Transition Project Teams (TPT) to include a new Multi-Model Ensemble TPT. This TPT includes subgroups focused on the MME development, consolidation and verification. The MME development group will deal with all aspects of MME (national, international, empirical, statistical, dynamical models, etc.), while the consolidation (verification) group will deal with all aspects of the consolidation (verification) effort.

In order to carry out the specific MME activities described above, the CTB has reallocated several CTB staff onto the MME team. CTB management will continue to increase support for these activities as they mature, commensurate with success.

Routine climate model development and assessment meetings involving EMC, CPC and CTB personnel as well as external collaborators are used to report and evaluate progress on MME and to optimize working relationships on MME activities.

2.3.6 Computing Resources

During FY07 the collaboration with GFDL will require a significant fraction of the available CTB computer time. CTB Management considers this to be quite reasonable, but will continue to evaluate progress and consider the trade space with other activities to ensure a proper balance. The CTB Management will work with the CST, OB, SAB, other NCEP personnel and the broader climate community to ensure that there is a consensus on this balance.

2.4 Enhancing Scientific Community Involvement

A close working relationship between the CTB and the scientific community outside NCEP is essential for achieving the CTB goal of accelerating the development of an MME Prediction System. This will allow the CTB to more effectively leverage outside resources to support NCEPs forecasting mission.

In order to address the basic question about how to better engage other major modeling groups, the CTB agrees to take leadership in establishing an interagency initiative for a Multi-Model Experimentation Facility (MMEF) at NCEP. Strong interagency support might help establish this facility, especially in garnering additional computing resources. The MMEF would promote parallel activities leading to the development of an operational MME Prediction System at NCEP that incorporates fully coupled national models (e.g. NCEP, GFDL, NASA, and NCAR/COLA), international models (e.g. EURO-SIP, APCC), and statistical models that add independent skill (value) to the ensemble. Current CTB international and national MME activities are summarized in sections 2.3.1 and 2.3.2. The MMEF would also help the community grapple with a number of difficult logistical issues (e.g., How would the various forecast systems be run? Does each system need to include the full associated data assimilation system?).

As an important step towards addressing these issues, and to promote stronger collaboration between NCEP and the national and international MME community, the CTB proposes an MME Workshop to further address the science of MME methods, and to get more information about the lessons learned from existing national and international MME efforts. This Workshop will likely be held late in FY07 or in FY08 as MME collaborations are established and science priority areas are defined .

As the options for the CTB MME Implementation Strategy become clearer, the CTB will work closely with NOAA Climate Program Office to develop priorities for future AO-driven support consistent with near-term and long-term efforts to build MME forecasting capabilities.

CTB is aggressively pursuing outreach activities with RISA's, RCC's and other intermediaries as part of its effort to develop a clear plan for decision support. CTB personnel are participating in RISA PI meetings and workshops organized by NWS/Climate Services Division in an effort to gather information on user needs for climate forecast information. Recently, the CTB provided quick turnaround to a RISA "wish list" of user-required climate forecast products. This feedback outlined specific actions that the CTB would take were deemed relatively painless, moderately difficult, and requiring careful thought. Similar interactions are anticipated with other intermediaries.

2.5 Milestones

CTB has established a timeline and milestones for accelerated implementation of the MME prediction system at NCEP. These milestones are provisional, and subject to change. They are based on the assumption that each model will bring independent skill to a MME with the CFS. The current set of provisional milestones is as follows:

FY07: Submit CFS verification results to the Lead Center for Long Lead Verification (verifications are to be done globally, but can also examine regions such as Europe and North Africa);

FY07: Evaluate GFDL CM 2.1 hindcasts (1979-2006) for their skill for ISI prediction;

FY07: If additional skill to the CFS is found, then port GFDL code to NCEP for timing and reproducibility testing (joint with EMC).

FY07: Preliminary skill evaluation of an IMME consisting of the EURO-SIP models plus CFS;

FY08: Preliminary skill evaluation of an IMME consisting of the APCC models plus CFS;

FY08: If additional skill to the CFS is found, then establish formal collaboration with EUROSIP and APCC on MMEs (Joint with OD);

FY08: If additional skill to the CFS is found, then consolidate the EURO-SIP and APCC IMMEs into a single IMME;

FY08: If additional skill to the CFS is found, then add the GFDL Model to the NMME (NCEP, GFDL) for subseasonal, seasonal and experimental forecast products;

FY08: If additional skill to the CFS is found, then add the NASA Model to the NMME (NCEP, GFDL, & NASA) for subseasonal, seasonal and experimental forecast products;

FY09: If additional skill to the CFS is found, then add the NCAR Model to the NMME (NCEP, GFDL, NASA, &NCAR) for subseasonal, seasonal and experimental forecast products;

FY10: If additional skill to the CFS is found, consolidate the NMME and IMME into a single MME Prediction System;

FY10: If additional skill to the CFS is found, incorporate the MME into the CPC consolidation scheme together with operational statistical tools.

FY10: Transition the MME Prediction System to NCEP Central Operations.

3.0 References

Krishnamurti, T. N., C. M. Kishtawal, T. E. LaRow, D. R. Bachiochi, Z. Zhang, C. E. Williford, S. Gadgil, and S. Surendran, 1999: Improved weather and seasonal climate forecasts from multi- model superensemble. Science, 285, 1548–1550.

Palmer, T. N., A. Alessandri, U. Andersen, P. Cantelaube, M. Davey, M. Déqué, E. Díaz, F. J. Doblas-Reyes, H. Feddersen, R. Graham, S. Gualdi, J.-F. Guérémy, R. Hagedorn, M. Hoshen, N. Keenlyside, M. Latif, A. Lazar, E. Maisonnave, V. Marletto, A. P. Morse, B. Orfila, P. Rogel, J.-M. Terres and M. C. Thomson, 2004. Development of a European Multi-Model Ensemble System for Seasonal to InterAnnual Prediction (DEMETER). *Bull. Amer. Meteor. Soc.*, **85**, 853-872.

Saha, S, S. Nadiga, C. Thiaw, J. Wang, W. Wang, Q. Zhang, H. M. van den Dool, H.-L. Pan, S. Moorthi, D. Behringer, D. Stokes, M. Peña, S. Lord, G. White, W. Ebisuzaki, P. Peng, P. Xie , 2006 : The NCEP Climate Forecast System. *J. Climate* (in press)

USCCSP, 2003: *Strategic Plan for the U.S. Climate Change Science Program*. Climate Change Science Program and Subcommittee on Global Change Research. Climate Change Science Program Office, 1717 Pennsylvania Avenue, NW, Suite 250, Washington, DC 20006