Report of the

UCACN Model Advisory Committee

05 October 2016

UMAC is a sub-committee of the

UCAR Community Advisory Committee for NCEP (UCACN)

Administered by the

University Corporation for Atmospheric Research

UMAC Members

Frederick Carr, co-chair
Richard Rood, co-chair
Alan Blumberg
Chris Bretherton
Andy Brown
Eric Chassignet
Brian Colle
James Doyle
Tom Hamill
Anke Kamrath
Jim Kinter
Ben Kirtman
Cliff Mass
Peter Neilley
Christa Peters-Lidard

Ex officio members of UMAC

Gary Lackmann
Gilbert Brunet
* Bill Kuo
Tsengdar Lee

* Josh Hacker attended for Bill Kuo

On August 9-11, 2016, the UMAC (UCACN Modeling Advisory Committee) held an in-person meeting in Silver Spring, Maryland. The invitees included the UMAC, NOAA staff, and members of the private, academic, and federal communities. The agenda and the registered attendees are attached at the end. This document includes individual narratives from all of the UMAC members. Members who were not in attendance had access to the presentation materials as well as notes from the meeting. Common themes from the individual's narrative contribute to the Summary Remarks. An acronym list is provided at the end of the narrative.

Summary Remarks

Strategic Plan: A near ubiquitous comment from the individual members of the UMAC is the need for a useful and usable strategic plan that connects together NCEP, indeed NOAA modeling, from bottom to top. Such a strategic plan not only unifies the organization towards common goals, but it links together products with balancing cost, potential scientific prize, and user need. There have been many strategic planning activities, but the plans from these efforts do not appear to be meaningful to the organization.

Throughout the individual comments, specifics are given. Noteable specifics include: the need to consider an individual's incentive structure; linking computational, scientific, workforce capacity, and cost; proper description and management of requirements. Special attention is needed to Meso-Unification / Regional modeling which remains too complex and is a cost burden and damages the scientific credibility of the organization.

There is much discussion in the individual comments of the need for NCEP to engage more actively and effectively with the community. A strategic plan, including the strategic understanding of the role of community, is necessary for NCEP and NOAA to develop a stable, substantive, and broadly beneficial community presence.

A number of committee members note that the ECMWF Strategic Plan reflects the values and goals that are needed for a focused, product-driven organizations. It serves as good example of a plan and can be found here:

http://www.ecmwf.int/sites/default/files/ECMWF_Strategy_2016-2025.pdf

Community: Many of the UMAC members emphasize the importance of NCEP to more effectively work with the community: private sector, federal, and academic. Given the increasing complexity of forecasting systems expected in the next decade, NCEP requires an expertise base that is beyond its internal workforce. Research activities outside of NCEP have the potential to improve the value to the Nation of NCEP's forecast products.

Progress on community inclusion and shared decision making is noted, especially, with the NGGPS Overarching Systems Group and the NOAA Environmental Modeling System (NEMS) meetings on coupled models. Foundational elements for community building include development of code management and documentation standards and practices. Some of these practices serve as examples that could benefit EMC, NCEP, and NOAA as a whole. We note that since the UMAC meeting the NEMS meeting on Code, Data, and Documentation Management had high levels of participation from EMC staff and management, with management requests for documents that are foundational for community development.

An item of particular concern to many members is that, now, with the choice of FV3 as the dynamical core, how does NOAA, NWS, NCEP, EMC assure that the global model that is developed is a community model? There are many specific comments from individuals worthy of attention. We note in particular that many individuals suggest the formation, in the near term, of a class of trusted super users, who are given access to the code, as well as trained under the tutelage of GFDL experts. The code used in the dycore testing should be made publically available. An item of particular note is the need to start testing FV3 at convective-allowing scales as soon as possible. This testing needs to be an open process, with community participation. This testing is strategic; it allows evaluation of the suitability of the scheme for use in a unified global-regional product suite.

Looking longer term, NOAA needs to take a knowledge-based, strategic approach to community building and participation. Members note, specifically, that community does not equate to making a model publically available, does not equate to asking for community input, is not defined, unilaterally, by individuals, organizational units, or indeed NOAA. Effective communities are defined by the participants and evolve a beneficial set of rules and definition of effective, balanced governance practices. Recommendations from individuals include a study of existing community efforts, their governance, their functioning, and their effectiveness. NCEP and EMC need to develop, internally, their reasons for community participation, the goals they need to achieve, and their parameters for participation. Though there is an urgency to develop community for FV3, the long-term, strategic importance of community participation requires planning, prototypes, and cultural change. Therefore, it makes sense to have short-term and long-term efforts on community building. For example, in the short-term, an FV3 sandbox to support dissemination of FV3 experience and exploration of community practice might make sense, while for the longer term, more strategic approaches are evolved during the design phase of the unified global system.

Global Modeling and Global Model Unification: Several members of the UMAC noted progress in global modeling and the efforts towards unifying the suite of products that use the global model. The agenda placed together, in a panel, the leads of the applications participating in the NEMS-based coupling activities. This panel includes all of the products that rely on the global atmospheric model as well as some applications that require coupling on the regional scale. It was evident that there was improved communication across the leads, and the developing of understanding of the importance of considering the impacts of decision across the system. Hence, there is the beginning of shared decision making. This progress is often associated with NEMS and the NGGPS project and needs to become part of the EMC and NCEP internal organizational culture.

NEMS is operational, working, and evolving. Given that virtually all categories of global forecasting in the next decade will require coupled models, the investments and successes with NEMS represent a strategic asset. It is elemental for global model unification, and, ultimately, integrating global and regional modeling.

High Performance Computing and WCOSS

UMAC strongly encourages the continuation of a multi-year procurement plan for High-Performance Computing, hopefully one that does not require natural disasters such as Sandy to move forward. We note that obtaining top HPC performance among national NWP centers is a moving target; e.g., the British Met Office will have 16 petaflop capacity by the end of 2016. There are two additional vital components to NOAA HPC acquisitions: First, research computing capacity has to remain commensurate with operational capabilities; otherwise new ideas and implementations cannot be properly developed and tested. Second, not all HPC resources should go just for increased CPU (super-computer) acquisition; it is vital to have "HPC investment balancing", in which disk storage, archival storage, bandwidth as well as software to support data management and workflow are obtained to optimize R&D effectiveness.

As noted throughout the UMAC narratives, NCEP is being strongly encouraged to use its available computing capacity to move boldly toward a convection-allowing model ensemble, and improved physics and data assimilation for all models. UMAC notes that there are several ways to increase computing capacity for these HPC resource-intensive goals. One is to decrease the amount of "white space" and make more effective use of the existing capacity. Second, there is no need to run the climate forecast system on the WCOSS, as the costly 99.99% reliability standard is not required for these runs. It is suggested that very fast compute clusters could be added to complete these CFS forecasts, freeing up time for convection-allowing model ensembles on WCOSS. Finally, owing to the large difference between peak and actual performance, it may be worth teaming with other large HPC centers and vendors to test methods for more efficient use of HPC capacity.

Convection-Allowing Models and Ensembles

1. There is strong agreement among UMAC that **NCEP should immediately institute a 3-km ensemble forecast system over CONUS.** Such a bold step will immediately make U.S. NWP a world leader and be greatly welcomed by NWS forecasters. The narrative immediately below outlines some crucial elements of this initiative and additional valuable guidance can be found in some of the individual comments.

While advocating for such an immediate action may seem contrary to UMAC's evidence-driven maxim, there are already ensemble systems being used by operational forecasters, namely the SSEO and the NCAR ensemble. Thus, as an initial step, it would make sense to operationalize the SSEO. Guidance for the construction of this ensemble can be found from the results of recent HMT and HWT Spring Experiments, although some additional studies may be needed.

2. While the SSEO is being used as a prototype, allowing new software tools for display and verification to be developed, **an intense research effort should proceed in parallel on the subsequent ensemble forecast system (EFS)**. Some but not all UMAC members recommend that all convection-allowing model EFS efforts should be directed to the HRRR (ARW) model, with immediate discontinuation of NMMB modeling efforts, believing that the goal of production suite simplification overrides the need to do an NGGPS-type "bake-off" between the ARW and MMB cores. Others note that we should adhere to our evidence-based theme and do the necessary evaluations. Most current evidence shows that ARW-based models outperform NMMB-based ones, but that NMMB members still provide value to multi-model ensembles (e.g., the SSEO); this needs to be tested. Most UMAC members ultimately endorse the development of a single core HRRRE system, obtaining its diversity from stochastic physics and other means. Since the SSEO involves no data assimilation, it is an efficient interim step, but the requirement of perhaps 100 or so ensembles in the data assimilation part of a convection-allowing model EFS makes a multi-core EFS much more expensive.

3. There is unanimity among UMAC members that **a strategic plan needs to be developed for convection allowing modeling (CAM) at NCEP, and its associated ensemble forecast system (EFS)** - noted as "Meso-unification" in some EMC planning documents. The most common aspiration is that atmospheric modeling at NCEP should consist of two cores (HRRR and GSM-->FV3) in the near future, with a goal of a unified system in 3-5 years if FV3 is suitable (or can be modified to do well) for convective modes. There is considerable guidance in the individual narratives about this, but common themes include using a NGGPS-like process, involve all levels of the organization (from model developers at EMC to upper-level management) as well as sister NOAA labs and the university community, connect with requirements and stakeholders, and coordinate with high-performance computing (HPC) requests/procurements. Note that the completion of this plan should not be drawn out (just 3-4 months) and that the initial SSEO step mentioned above could go forward without waiting for the plan to be completed.

4. With the desire to create a balance between keeping this summary short but still highlighting important ideas, we present the following list of crucial components to the design of convection-allowing model EFS at NCEP and identify the author(s) who address them below:

- Strategic planning (Brown, Colle, Doyle, Hamill, Kinter, Mass, Neilley)
- Having Chief Scientist for Modeling at NOAA HQ level (Brown, Brunet, Hamill, Lackmann)
- Unified modeling (Brunet, Lee)
- HPC reallocation to accommodate EFS (Carr, Neilley)
- Use of HWT to evaluate EFS (Carr, Colle, Doyle)
- Convection-allowing model and EFS verification metrics, display, postprocessing (Colle, Hamill, Lackmann)
- Importance of requirements (Hacker)
- EFS system design (Hamill, Mass)
- Data assimilation (Carr, Doyle, Peters-Lidard)

Salt and fresh water modeling issues

The coastal “salt water” modeling and associated forecasts from the NWS and NOS have been fragmented and disjointed. At a minimum, Senior Leadership from NCEP and NOS (CO-OPS and OCS) should meet on a periodic basis (minimum of twice a year) to discuss near-term and future plans for coastal and storm-surge and inundation modeling, and identify and strengthen strategic partnerships that could be supported by both. With the creation of the National Water Center (NWC) there is an opportunity to bring together the water experts, both CONUS (freshwater) and coastal (salt water) into a unified structure. This initiative should begin with a community-based workshop identifying what modeling is being done in NOAA accompanied by recommendations for evidence-based model choices, followed by the development of a strategy to integrate water modeling. An important step in this process is to ensure a rigorous review of the SLOSH model. NWS, NWC and NOS should all be using the same models to address their missions.

Narratives from UMAC Members

Below are narratives from each UMAC member in alphabetical order by last name.

Alan Blumberg (Stevens Institute of Technology):

The improvements in storm surge and estuarine circulation forecasting have been very disappointing despite detailed recommendations from UCACN and UMAC as far back as January 2014. Moreover, it was disheartening that no one in ocean leadership from NCEP was in attendance at the NCEP review.

Two types of models and used within the NHC/NOS storm surge enterprise: a simple model (SLOSH) that can be used for quick guidance and probabilistic products and a slightly more complex model (ADCIRC) that captures surge response relatively better but is computationally more demanding. Both models have serious shortcomings.

The SLOSH model forms the basis for the probabilistic tropical cyclone storm surge (P-Surge) model. The rationale for its use is that the model runs quickly and can be used to calculate the probabilities of a predicted storm surge. A probabilistic approach is essential. The methodology for how the probabilistic forecasting is being done is quite problematic regarding the model's use of atmospheric forcing and its statistical analysis framework. These problems were clearly evident during forecasting for Hurricane Hermine when at 11am September 4th 2016 the Director of the National Hurricane Center announced that all SLOSH probability based storm surge forecasts were being discontinued. The methodology must be subjected to rigorous peer review.

Now here seems to be a push in some parts of NOAA to rally around the use of ADCIRC, a vertically integrated, two dimensional hydrodynamic model. That model has been applied in different configurations by both FEMA and the USACOE in those agency's flood risk coastal modeling. The FEMA/USACOE ADCIRC setup is not optimal for storm surge forecasting. The domain size and model grid leads to problems of runtime and of inaccurate coastal flood predictions. Much work has been done by FEMA and the USACOE in getting their versions of ADCIRC to run and while tempting to leverage from that work, it is a mistake. Now the National Water Center has embraced ADCIRC too. This is a puzzle because fresh water in ADCIRC plays little role in possible flooding. ADCIRC and SLOSH for that matter are inherently incomplete with respect to bottom stress. Storm surge results from the imbalance between surface and bottom stresses, which requires a pressure gradient and hence a sea surface slope (the storm surge). To better forecast storm surge, improvements on how the bottom stresses are computed are sorely needed. Those models use the average velocity of the water column in their parameterization of bottom stress. The currents at the bottom of the water column must be used. NOS is using FVCOM and ROMS for their operational forecasting. Both of these models have more complete physics and their use will lead to more accurate forecasts of storm surges. The rationale that the physically limited ADCIRC and SLOSH model must be used because it is less demanding of computational resources than are FVCOM and ROMS is not true. The current allocation for storm surges on WCOSS resources seems surprisingly small given the significance of this issue. The allocation could easily be expanded given the new computer power so that appropriate physics based models can be run.

Almost immediately an overarching strategy must be developed that brings together NOS/CO-OPS and NOS/OCS and NCEP expertise and assets, perhaps within the National Water Center, to strengthen program missions within both of those agencies.

Chris Bretherton (University of Washington):

I was not at the UMAC meeting, so I will be brief and mainly focus on NCEP's response to date to last year's recommendations. My comments are, in part, based on my perceptions as lead external principal investigator of a climate process team on cloud and boundary layer parameterization in GFS/CFS that involves regular interaction with various model developers in EMC.

I am happy that NCEP management absorbed the message that the model suite needs to be simplified and unified so that focused resources can be brought to bear on the goal of making NCEP's forecast models the most skillful and useful in the world to U.S. end-users. I appreciate that a NWS chief scientist has recently been appointed, and I have every hope that this will help to bring reality at EMC and NCEP closer to the lofty goals. The NGGPS dycore exercise has been very productive in pointing to a promising path forward for GFS, hopefully one that will be compatible with testing/development of the model by a broad user community.

Like others on UMAC, I stress the need for a strategic plan for NCEP modeling. To be effective, this plan needs to have a broad base of rank-and-file acceptance within EMC and NCEP, so it must be formulated with the help of collective internal discussions among NCEP model developers, not just senior management and external voices - it can't just be written in isolation without taking this process seriously. It must include three elements, a vision (UMAC already gave a good start for this), a realistic but firm transition timeline include meaningful near-term milestones (EMC's UGCS dream, to be realized in the nearly indefinite future, is a good example of what happens in response to our vision when there is a perceived absence of other planning), and a management plan for how to allocate the personnel and computing resources and leverage outside collaborations to accomplish the implementation plan on a realistic budget. Moving EMC's regional modeling effort forward is one can of worms. But global modeling also needs strong strategic management.

CFSv3 is a good example of a project that is in the doldrums due to too vague a strategic plan in all three of these facets. But my personal discussions with EMC leads suggest that EMC management of global atmospheric model development generally is more reactive to whatever individuals want to do than it is proactive in forming groups that work on the most pressing model problems, and the model development process could benefit from more rigorous strategic planning and management on all time scales.

I'm really hoping that by this time next year, EMC or NCEP more broadly will have a meaningful strategic plan with all the above three elements, with buy-in from model developers and group leads as well as managers.

<u>Andy Brown</u> (The Met Office):

1. UMAC strongly advocated a plan to consolidate (as, while with unlimited resources, the best results for lots of different applications might be achieved by lots of different systems, a realistic balancing of the triangle of cost, potential scientific prize, and user need strongly points this way as a strategy).

2. On the global side, happy that significant progress is being made (and also note that we did last year advocate global unification and separate mesoscale unification as a practical first step – I think it's easy to exaggerate how quickly global will be a convection-permitting resolutions – even if able to do single shot 3-5 km global every 6 hours in a few years, it will be good while before hourly cycling, ensemble global system at those sorts of resolutions). NCEP should be congratulated. As others have said, there is of course further to go (dynamical core is only part of it), and do need credible implementation/transition plan to develop new system AND transition effort from current separate systems to implementing different configurations of the new unified system to serve different applications.

3. On the mesoscale side, it seems clear that much less evidence of progress – and a pressing need for it. I think the imperative here is a credible 3-4 year strategic plan to get to a simpler world (I'm actually less worried by the details of exactly what it proposes (which models etc) than that it should exist and that there should be a clear will to implement it).

4. I would assert that inasmuch as organizations like the Met Office and ECMWF are successful, a big factor is their strongly directed internal research and development programs + skilled staff who are bought into them (will talk about partnerships below). Development of NCEP strategic plans definitely important here (but mustn't also forget staff and cultural aspects). However, in terms of the joining up that is within NOAA's (rather than NCEP's) gift, very much agree with previous UCACN recommendation concerning chief scientist that would oversee all modelling (so e.g. OAR and NCEP plans make sense when seen as a whole, rather than being brought together after the event by some committed individuals).

5. Partnerships beyond organizations also really important. In part this is just an attitude of mind and openness to ideas from beyond the organization (I'm very happy to copy any good idea if

it's better than what we've come up with to address a problem!). I agree with the comments of others about the desirability of working with community over NGGPS (in part to cement that attitude and in part because of genuine additional value that can be pulled through from academia). However, based on my experience of where we've got most value (and I think consistent with the comments of others), I don't think just having a very large community of model users is sufficient – needs real hard thought (and investment) on how to get all that community expertise (a) focussed on some of the priority areas and (b) doing work in a way (e.g. on up to date systems; recognizing operational cost and code quality constraints upfront;.....) that maximizes the chances of it being usable by an operational system. Some of this puts an onus on NCEP to facilitate it (and possibly even commission some of it), but I think it puts a big onus on the community (and its funders) to think and act differently too i.e. it's not a one way street. So in short, absolutely yes to more working with community, but to get most value need to really review what's worked well with past efforts / what could better facilitate joint working. Some of our most productive examples have been truly jointly designed (and funded) programmes across Met Office and academia.

<u>Gilbert Brunet</u> (Environment and Climate Change Canada):

Since I was not present at the UMAC 2016 workshop, I will focus on the NOAA Action Plan for UMAC 2015 report. Overall this is good progress. Clearly the rational evidence-approach is being adopted (e.g. NGGPS). As an example the NCO 30-day parallel testing is now converted to a final test for operational environment and information technology (IT) robustness only. The implementation key decisions will be rightly based on upstream extensive numerical experiments and stakeholder evaluation and input. This is an approach used also in Environment and Climate Change Canada (ECCC), Met Office and ECMWF.

The Tiger Team and a Unified Modelling Task Force, under the leadership of the NOAA Chief Scientist, could potentially be the organizational breakthrough needed to push forward the NOAA modelling strategy. It looks like we are putting in place committees to oversee the short and long term plan of the chain of innovation (research, development, operation and services) at NOAA. It is good, but the devil is in the details.

It is very disappointing that the creation of the Chief Scientist position for NEWP was not accepted by senior NOAA leadership. It would be good to know why? This is essentially the ECCC, Met Office and ECMWF governance model and I believe it explains a great part of their successes. In my view to think otherwise is wishful thinking.

The approach to reduce complexity is relatively well defined for global NWP and I endorse the results of the NGGPS trial. My opinion is that there are lots of good dynamical cores around for NWP applications and all of them have their own strength and weakness. We should not make an obsession to find the perfect dynamical core. NWP accuracy is a fine balance between dynamical core, physics, observations and data assimilation techniques. The physics will make sure that you don't need to have a perfect dynamical core for a given horizontal and vertical resolution. I am convinced that NGGPS have found a good one to start with, so now the real challenge is to put all together the different components of the NWP system. This is the most demanding and painful effort and it is why unification is so important. You don't want to do this too often. In view of this, data assimilation techniques will be the cornerstone for the model evaluation (e.g. test new physics) and comparison to observations. It is unclear from the document what will be the data assimilation framework to do these numerical experiments and how cost-effective it will be. This is something that UMAC could ask more information about.

I find reducing complexity for convection-allowing models (CAM) modelling less convincing. I still don't understand why we are still pushing forward a major effort for hurricane modelling (HFIP) in isolation of all other convection-allowing model applications. This is not an effective way of disseminating new innovations. Hence from my perspective there is a pressing need to unify convection-allowing modelling. This is something that UMAC could be more pro-active about. In ECCC we see our long range plan (beyond ten years) for convection-allowing and urban modelling as the most critical project to push forward. Once we will have developed a decent convection-allowing model for massively parallel HPC, we will use the Ying-Yang approach to build our global NWP model. There is good evidence that this is a proper approach to keep the development of a unified model cost-effective. In fact, I am surprised that it seems that no one in the USA has proposed or is testing this approach with WRF. Anyhow this is food for thought for the NOAA long term plan for the unified modelling strategy.

<u>Fred Carr</u> (University of Oklahoma):

<u>Super-users</u>: Even before the community model structure for FV3 is established, there exists the need for training "super-users" who know the code extremely well such that they can help lead research projects or teams that address, e.g., physics , data assimilation, and postprocessing. These super-users should be at EMC, ESRL, NCAR, NSSL, universities, etc. GFDL needs to quickly develop a plan to do this training.

<u>HPC Allocation</u>: In order to accommodate the NWS field office demand for convection-allowing model ensembles as soon as possible, greater resources on WCOSS dedicated to this may be needed. Two ways to accomplish this are (1) to use all the resources allocated now to RAP, NAM, NAM Nests and Windows, and HRRR to do a large-domain (N. America) deterministic run with HRRR at 3 km that covers all the areas where nests are now [this is run in hourly RR mode but with forecasts out to 60-84 hrs every 3-6 hrs]; this should permit a HRRRE ensemble on a CONUS domain to be established; (2) move the climate model runs off WCOSS to a fast compute machine that may have lower reliability (i.e. 99% rather than 99.99%). This machine could also be used for continuous reforecast and reanalysis runs to keep up with model bias corrections, etc.

Ensemble design: [These conclusions are primarily based on summaries of the 2015 and 2016 Ensemble Design Workshops and results from the HWT Spring Forecast Experiments (SFE)] The Storm-Scale Ensemble of Opportunity (A group of 7 model runs with both ARW and NMM cores made at EMC and NSSL, and processed by SPC) still performs better than current single core ensemble systems for most verification metrics. If QPF is used, ARW-based systems perform better. There is some evidence that a properly designed single-core HRRRE (ARW) ensemble could be better than a dual core system with HRRR and NMMB, but this needs to be tested first. The HWT SFE is an excellent opportunity to test EFS design options, in concert with EMC, GSD, NCAR, etc. partners.

These results suggest that for the near-term, the SSEO should be made an official operational product while additional ensemble-design experiments are performed. The overarching need for NPS simplification should be sufficient to justify a decision to concentrate on creating a well-designed data-assimilated HRRRE ensemble and to cease all development of the NMMB system.

FV3 Phase 3: EMC and its NGGPS partners should make a strong push to speed up phase 3 of NGGPS (operational implementation of FV3) as much as possible, perhaps completing the task in 2 years rather than 3. FV3 should be put on WCOSS (or an adjacent high-end compute machine) in parallel mode as soon as possible for stakeholders to examine output across all possible weather events to look for failure modes, skill dropouts, etc. The FV3 could also be adapted to run as one of the convection-allowing model systems in the HWT SFE starting in 2017 so that the community can begin to evaluate its abilities to forecast convective systems.

Community modeling effort: Need to look at pros/cons, lessons-learned of existing modeling systems (ECMWF, WRF, MPAS, CESM, HRRR, WW-III, etc.) to determine best practices for setting up a well-structured code management system that also benefits from the community. The NOAA response document to the 2015 UMAC report suggested possible use of DTG-like process for other major NPS projects (Meso, Water, etc.), as well as a “Change Review Board”. If possible, to avoid too much bureaucracy and delay, these two proposed functions could be combined into one oversight or advisory group.

Product termination: Stakeholders are important, but NPS changes should not be hostage to a few “important users”. There needs to be a well-managed sunset process in which all users are informed well in advance, “hand-held” on where to find equivalent or superior guidance, etc, and with NCEP management being supported by NWS/NOAA leaders.

Importance of Data Assimilation: Data assimilation (DA) has been insufficiently discussed in UMAC-NCEP conversations and reports. The community modeling and strategic planning exercises recommended by many in the report should remedy this situation and include a thorough discussion of the vital role of DA in all modeling systems and plans for transition to improved techniques. This discussion should include use of new observations and improved use of currents obs, improved hybrid EnKF/4DVar methods, observational and forecast error specification, use of stochastic physics in EFS, improved QC and bias removal, and DA issues associated with coupled models.

Eric Chassignet (Florida State University):

I was not able to attend this summer's UMAC meeting due to schedule conflicts My comments are short and emphasize a couple of points that have already been put forward by others.

First, I would like to reiterate the need to carefully consider and plan for community involvement outside of GFDL in the development of FV3. This is essential if FV3 is to become a national and international standard. In particular, NOAA and NCAR have a lot to gain by working closely together.

Second,, the planned "water" summit should strive to establish a clear roadmap on how the different centers and services should interact with each other and integrate their modeling strategies. A unified structure would be benefit the fresh water side (hydrology) as well as the salty side (storm surge, coastal, and deep ocean). With the establishment of the National Water Center (NWC), there is a need for coordination and clarification of the respective missions of the NWC, NWS, and NOS.

Brian Colle (Stony Brook University)

Operational Convective Allowing Ensemble (by late 2017)
Although I highlight many potential issues (opportunities) below, there needs to be a focussed initiative within the NCEP/NWS during the next year or two that has a large short-term benefit. NGGPS FV3 is the future, but this is a longer term goal. NCEP needs a shorter-term initiative that can be headlined, gets good PR (similar to the latest ECMWF strategic plan), and gets the community involved right away. That shorter-range goal should be the high resolution (~3-km) ensemble. NCEP needs an all-hands on deck type of goal to make this a top priority for 2017. One can argue that there are bigger fish to fry, but NCEP needs to walk before it can run and tackle a short-term goal that is tractable given what already exists (HRRRE, NCAR ensemble, etc.), generates excitement with the walls of NCEP, utilizes evidence that already exists about the ensemble benefit (e.g., NSSL Spring Experiment), and already has broad community support (forecasters and other users). This system will help the NWS move towards some sort of ensemble plan, which currently does not exist. No plan means that not only is the motivation for a high resolution ensemble lacking, but it hurts the NWS in many other ways: (1) there is no graphical software to display ensembles in operations (latest AWIPS build finally has something, but it is primitive), (2) there is limited validation of ensembles that connect the data to the forecasters/users, (3) most NWS offices (even those that have bandwidth) only get a handful of ensembles members and just two variables (temperature and precipitation), and

mean/spread/probabilities for a few variables, (4) there is little forecaster training on how to use ensembles in operations (outside of a COMET module or two). The lack of a plan also explains why the NCEP Ensemble workshop that I attended this past June 2016 lacked focus, with no real discussion about a high resolution ensemble.

Evidenced-based Metric Development:
Although there has been some progress this past year in using evidence-based metrics for guiding future modeling efforts, there is much more work to be done. One problem is that each group has its own set of metrics to evaluate the models, but there is no consistent set of approaches to provide the evidence to move our models forward. Some groups, such as those involved in the NSSL Spring Experiment, use more advanced approaches (feature-tracking, nearest neighbor, clustering, etc) over small regions, while others use the conventional set of metrics (RMSE, anomaly correlation, Equitable threat, etc) to make model upgrade decisions. Developing a consistent set of metrics useful for model developers will be critical to move the FV3 model forward rapidly. Recommendations:

1. Development of a list of comprehensive metrics that match the requirement, model development, and users/stakeholders. This will require a more coordinated effort between NOAA, universities, and various users of the data.
2. More metrics are needed to understand the physical processes in the models (surface fluxes, PV, moisture transport,...), which requires feature-based approaches and smaller regions where specific weather phenomena exist (East coast, central plains, etc…). Also, metrics that better relate determine whether the models have intrinsic value to the users. This can include social science or economic metrics. The need for more regular reforecasts (high-resolution version) is clear to help with model development.
3. The evidence from these metrics needs to be shared more efficiently within the community, not just buried on a web page. The web page should include outstanding issues still plaguing NWP (stable planetary boundary conditions, riming in clouds, aerosol impacts, orographic drag, etc.), so various groups can understand and contribute to the library of information, and model developers can understand the latest results.
4. We need a better way to share modeling datasets, so others can build on the current evidence. Some model data is on NOMADs, others are on the NCEP computers, which requires special access. Some of the most valuable datasets are the NSF-related field datasets and NOAA Testbeds. Millions of dollars are spent on these field efforts, but after several years this data tends to disappear or become harder to access by the community.
5. Ideally, a common piece of software for verification would be helpful, so we can easily compare results and build on existing code. The DTC's MET (Model Evaluation Tool) is heading in the right direction, but it needs more community involvement.

Community Involvement:
Community is important for model development as demonstrated by the MM5/WRF model over the years. A plan is needed to incorporate the community (especially universities and private sector) into the new FV3 modeling system. A workshop in the near future is important to lay the foundation of this involvement, but a more personal touch is needed to develop a set of of friendly superusers outside of NOAA to get the process started, so they can learn how to run and contribute to FV3. GFDL/NCEP should reach out to some other groups and ask for help rather than just a sign-up list at a workshop -- this is what worked in the early days for MM4/MM5 with NCAR. GFDL/NCEP should develop a visitor/post-doc program just for this FV3 development, ranging from model physics to data assimilation to post-processing. This community needs to be on some sort of oversight community to help guide and vet requests, and a feeling that everyone will get something out of the modeling system.

NWS forecasters are an underutilized community, which is on the front lines using NCEP models each day. A framework should be set up to get forecasters involved with providing as much feedback as possible on the models. There are other issues that involve forecasters as well, such as how to get the convective-allowing model ensemble forecasts to the field offices and what fields/products do they want. Ensemble display in the field offices is too basic, primarily mean/spread/spaghetti.

The storm surge and National Water Model (NWM) efforts should follow a similar path and at an accelerated rate. The Water Summit planned is a good start, but there needs to be much more outreach and iteration with the research groups outside of NOAA in order to incorporate the best science into these models. There are some fundamental questions on what model to use for storm surge and hydrological efforts moving forward, especially as one gets down in scale and deals with beach processes or urban-scale flooding.

Strategic Plan Efforts: Efforts to Motivate Future Computing and Development

There needs to be more forward-thinking approaches for anticipating the next big NWP advancement, which will help NCEP in many respects, from attracting visiting scientists from around the world and acquiring the next generation high performance computer. NCEP needs to think big and show how U.S. weather and seasonal climate prediction can improve if given the proper resources. The current approach is usually a reactionary process, in which a larger change is proposed when there is a large weather event (e.g., Sandy) or major forecast bust (DC/NYC snow-storm). Rather, an organized research and exploratory research effort is needed to demonstrate the potential of what the next generation guidance could do for specific weather or climate initiative. The need for a CAEM system during the next year or two (at latest) is one example, but we can think even bigger. For example, the benefits of having a global ensemble system run at explicit precipitation resolution (< 5-km grid spacing) needs to be explored for both extended and subseasonal forecast scales. One can hypothesize that this would really help the predictions as it avoids the upscale error growth from uncertainties with convective parameterization used in the critical tropics region. An NCEP team should consider how to pull this off, which would require the research community and their computer resources. If the hypothesis is proven, then NCEP has the evidence to ask for the next generation computer to make this possible. An effort like this also helps internal NCEP moral, since it provides an opportunity for those within NCEP to explore uncharted territories.

James Doyle (Naval Research Laboratory):

Simplification of the regional modeling suite is needed

The regional modeling suite appears to be overly complex and contains too many modeling systems. Simplification and consolidation of the regional modeling systems will streamline model development and allow resources to be focused on key model development areas. A detailed plan is needed to achieve a simplified and unified regional modeling suite. One way to achieve this is to focus efforts on the HRRR system and broaden the applications from this system.

Acceleration of the Transition of a Convection Allowing Ensemble System

The technology is readily available now for a high-resolution convection allowing ensemble system. These types of systems are already being run in operational or real time modes by various institutions (e.g., NCAR, UK MetOffice). At the UMAC meeting, we heard from key stakeholders underscoring the critical need for such a system and probabilistic products for forecasters and other users. There is an opportunity to build on the HRRRE effort for this system. The NWS needs to make the transition to a convection-allowing ensemble system for the U.S. a high priority and implement such a system much more quickly than in the 5-10 year time frame discussed at UMAC.

NGGPS

The process of selecting the dynamical core for the NOAA NGGPS is now complete. The process was rigorous and thorough given the short time horizon. In the next 1-2 years, FV3 should be much more rigorously evaluated at the convection allowing scales. Applying the FV3 during the NOAA Spring Experiment will be a very good first step to evaluate the characteristics of the dynamical core for convective-scale applications. Additional carefully designed tests of FV3 at the convection allowing scales need to be done and expertise in the community at large should be leveraged to assist in the evaluation. If FV3 is found to be deficient, the issues with the dynamical core will need to be addressed quickly.

The dynamical core has been chosen, but this is really a relatively small (but important) part of the overall system. NGGPS has done a nice job in the planning process for the various components for the system. More coordination within NGGPS is still needed between the various components (e.g., aerosols and physics; physics and data assimilation etc.) and within EMC and the community broadly.

FV3 Community Model

The NWS and NCEP will benefit greatly from having a vigorous and active *community of model developers*, who can contribute to innovation and model advancements using the FV3 core. EMC would benefit most from encouraging a broad community of model developers rather than just model users. Strong and well-defined governance, and active software design and management are needed to maintain a careful balance between innovation and the streamlined nature of an operational system. The community should consider maintaining and advancing a smaller number of suites of physics, where there is considerable emphasis on accurately representing the complex interactions between physics components. A strategy worth considering is for NOAA to join with the CESM system to leverage an already robust community with an established governance. The software design for the FV3 model infrastructure should be amenable to encouraging development and advancements to the system such as available within CESM. This would enhance the weather-to-climate enterprise, which should greatly benefit NOAA and the broader community in the future.

Strategic and Implementation Plans

There is an urgent need for an overall strategic plan for the NWS modeling suite that is *actionable*, along with detailed implementation plans for the models and various components. Detailed plans and roadmaps for FV3, regional model simplification (including HRRR and HRRRE), advanced physics, data assimilation, component coupling (atmosphere, ocean, waves), etc. are all needed. More attention is needed on the data assimilation planning given how central it is to NWP. Detailed implementation plans on the shorter to medium time scales (1-3 years) are needed. Overall, I think there should be a stronger emphasis on prioritization of the various model development efforts. Identification of the model applications that will provide the most benefit to the end users should be given highest priority.

<u>Josh Hacker for Bill Kuo</u> (National Center for Atmospheric Research):

A more mature, complete, and socialized process for gathering requirements would have major implications for prioritizing production suite components, defining staff reward structures, and HPC acquisition. It appears that access to groups of stakeholders through NWS service areas and partnerships with other agencies is in place. Those groups should have a known process for submitting service requirements. Program managers are also in place, whose job it is to ensure that the service requirements enter a pipeline where they lead to science and technical requirements and priorities, and where funds are directed at the areas most critical to meet those requirements. This is necessarily a process in which information flows in both directions between EMC and the users of services. Requirements specifications are never static; they constantly evolve with changing science and technical capabilities, and changing service priorities.

Service requirements, linked to technical and scientific requirements, inform strategic planning and tactical activities. They provide NCEP/EMC with the foundation for metrics that demonstrate they are meeting the NWS mission to save lives and mitigate property loss. With some information about the impacts of weather on economic sectors, economic arguments can lead directly to service requirements. New products (and forecasting capability) deployed to address requirements provide a basis for assessing changes in economic impact. Lives saved under a successful forecast, resulting from a capability initially put in place to address a requirement, provides information to form metrics addressing the NWS ability to meet that part of its mission. Because stakeholders are intrinsically linked to requirements, they will advocate for EMC capability that meets them (thereby providing political capital).

A mature requirements process has implications for staff performance and a dynamic work force. Within EMC, the requirements met would most often be science and technical requirements that lead to the service required within the NWS service areas. Rather than staff "owning" a piece of code or a particular application, staff are part of a team where groups and individuals are rewarded based on their efficacy in meeting requirements. Contractors should work to SOWs formed to address specific efforts needed to meet requirements. As requirements evolve, so should the contracting workforce.

A mature requirements process informs grant programs. Grant programs target science and technical needs to augment EMC expertise, thereby contributing to EMC's ability to meet their requirements.

A mature requirements process is the basis for acquiring HPC resources. Rather than basing HPC capability and capacity on the whims of Congress or the voting public, or awaiting a natural disaster, the requirements drive the technical specifications for HPC resources. An ability to rigorously define the need for HPC resources to address particular service needs gives NOAA the leverage it needs whenever resources do become available.

As an example, a requirement to "provide 1-h warnings of a thunderstorm that will produce golfball-sized hail within an urban-scale area" may lead to the following: (1) a requirement for probabilistic numerical guidance with a precision that can resolve those thunderstorms, (2) a requirement that numerical guidance includes probabilities on the size of hail stones. Those lead to requirements for (3) a convection-allowing ensemble, and (4) new parameterizations for hail size. Further, those in turn lead to (5) research programs that define the ensemble size needed according to (6) the appropriate metrics not yet developed, and (7) an assessment of the literature addressing the trade-offs between single-model and multi-model ensembles. The ensemble size and computational cost of the model, combined with the service requirement for a 1-h warning of the model lead to (8) HPC requirements and specifications that inform the next acquisition. Full consideration would include even more technical and scientific requirements. . Requirements resulting activities are fully documented and provide a direct link between scientific/technical capability and a service activity to save lives and property. Once the service requirements are met, an assessment of their effectiveness is possible.

Going forward, it is important that NCEP manage the transition to coupled data assimilation in Earth-system forecasting carefully. An implementation plan for data assimilation would help control the potential for proliferation of data assimilation codes. It is recognized that various implementations of the GSI are in broad use throughout the current and planned production suite. Already, the regional and global versions of the GSI are not interchangeable; they reside in different branches of the code repository. The RAP and HRRR versions of the GSI use different code branches than the NAM. Under NGGPS, it is expected that additional Earth-system components will be coupled to the GFS and CFS. Each may come with a different set of data assimilation code. Although the success of “strongly coupled” data assimilation in the future depends on scientific advancement, the ability for NCEP to consolidate technical data assimilation components early in the process will grease research to operations transitions in the future. By coordination with the JCSDA, the DTC, and other agencies NOAA will gain access to a greater pool of talent and increase the likelihood of meaningful contributions to NCEP operations from external sources.

<u>Tom Hamill</u> (National Oceanographic and Atmospheric Administration):

I did not attend the UMAC meeting in 2016 due to schedule conflicts. I am both a NOAA/OAR employee who is collaborating with EMC and one who is reviewing EMC as part of UMAC. My work experience is largely in ensemble system design, data assimilation, statistical post-processing, and forecast verification. To try to separate my own work from my UMAC role, my notes below are separated into: (a) UMAC-specific comments and recommendations on topics not directly related to my own research, and (b) observations over the last year with regards to my own experiences trying to bring about UMAC-inspired change in my own areas of expertise.

UMAC-specific comments and recommendations:

(1) *The 2015 UMAC recommendation for a chief NWP scientist at NOAA Headquarters with responsibilities spanning NWS and OAR should be reconsidered.* The Senior Level (SL) science advisor position, which I understand to be the NWS’s way of instituting the chief NWP scientist recommended by UMAC, is not at a high-enough level to be maximally effective. The NWP problem in NOAA is not EMC’s alone; the problems have also stemmed from disconnections between OAR and NWS, and disconnections of both from the broader academic community. I have felt on several occasions over the past decade that decisions made by OAR colleagues and management were not always in NOAA’s overall best interest. In other situations I believe that EMC personnel ignored relevant OAR developments to the detriment of

the final product quality. Leaving the SL position solely in the NWS thus means that there isn't a single person responsible for planning of NWP spanning OAR and NWS, coordinating those activities, making sure we're all rowing in the same direction. The NWS SL advisor cannot currently affect OAR plans except through the chain of command, and thus with many potential people to prevent a coherent vision from being realized. Another consideration is that the NWS ownership of this position provides the appearance that NWS is in the driver's seat with model development. OAR can advise, but NWS makes the strategic decisions. Is the NWS realistically that confident in its own judgment?

(2) Related to (1) above, much of the funding in OAR/ESRL for NWP system development comes from NWS and its NGGPS program, though this funding is generally prescribed for research and development (R&D) at high technological readiness levels (TRLs). While appropriate in many circumstances (as OAR had technologies ready for advanced testing and operations), *it's also important for the long-term health of NWP system development for OAR to be addressing major scientific challenges at lower TRL's (and to have the funding stream to do so)*, challenges such as strongly coupled data assimilation. *This points back to the need for coordination at the NOAA level to make sure that the NWP strategy is coordinated between line offices and that OUR research funding is directed to longer-term NWP needs.* This would be easier with that chief NWP scientist spanning NWS and OAR.

(3) NGGPS just went through an arduous (and admirably evidence-based) process for selection of a next-generation dycore. But selection of the dycore is just the beginning. *It is extremely important for this dycore to evolve to a truly community system* - not GFDL dycore to control without community involvement, not EMC's to control in isolation. Moving to a community model is part and parcel of providing the necessary O2R to facilitate more rapid tech transfer.

Understandably, there are configuration-management concerns with community systems. How do we have a community model but simultaneously safeguard the operational code? Tom Auligne, JCSDA director, recently outlined a vision for a community data assimilation code that I thought was compelling and could be a good model for the NGGPS. He proposed a tiered repository. The outer level of the repository would permit users to submit changes largely as they wished, without strict NOAA controls. A middle tier would have more research-level code, taken from the best algorithms developed at the outer level. And finally, the current and most refined code, closely maintained and secured by NOAA, would be in the inner tier. Decisions on the migration from one tier to the next would be made in an objective fashion, such as with a change control board comprised of subject-matter experts. *Conducting a workshop in the near future to set up the structure for this community system and a governance process is strongly encouraged.*

(4) Thinking about last year's UMAC recommendations, and specifically the recommendation for "evidence-based decision making," (EBDM) I believe that UMAC left NWS without clear guidance on what that means or how to put such a recommendation into effect. Arguably, nothing that NWS has done in the last several decades has been decided in the lack of

evidence, and yet NOAA lags several competitors in model skill. UMAC could provide more specific guidance on this if desired, but perhaps a better approach would be for the NWS to define a process for EBDM and then seek constructive criticism from UMAC and others.

(5) The focus at the 2016 UMAC on a rational consolidation of the regional system is appropriate. With SREF, WRF, NAM, HRRR, and possible future hi-res ensemble systems, there is still an unhelpful diversity, both in dynamical cores and parameterization suites. *EMC needs an EPDM-based plan for the rational consolidation of regional modeling.*

Other concerns that are not independent of my own work:

(1) Having concurred with the 2015 UMAC recommendation for evidence-based decision making (EBDM), I have tried to play a constructive role in suggesting how EBDM would be formalized for ensemble system development and post-processing. I am admittedly not a neutral player in this, for our ESRL team contributes to GEFS system development, and we haven't had clear guidance on the process for our work making it into the GEFS. For lack of a process, I have sketched out previously for EMC management what I think this might look like[1]. This document is also attached as Appendix 1 of my UMAC comments. I have circulated this through EMC management in the last few months. I welcome further dialog on this and working with EMC to formalize a process.

(2) As the 2015 lead of the NGGPS post-processing team, I helped organize a workshop in early 2016 to develop recommendations on the evolution of post-processing in NOAA. The workshop was conducted, and consensus recommendations were formed. I briefed this to senior NWS and OAR management[2], and I asked them what the next steps they wished to be pursued. The one concrete request was to generate a functional requirements document for post-processing. This was completed several months ago and is available[3]. We've received no feedback yet on this document. Many of the other recommendations still await NWS management action.

If the underlying reason is a perception that post-processing isn't all that critical to system development when compared to, say, dynamical cores or parameterizations, the post-processing community can provide evidence of its substantial impact. For example, I have found that post-processing of precipitation often achieves improvements in skill comparable to several decades of model development, and I believe we are still at a relatively immature stage of post-processing algorithm development; it could do more. I would be pleased to work with NWS management to implement these recommendations.

---

[1] https://drive.google.com/open?id=1eC1zaxD0gKLvK0YnbG4uU1r3T66o8ZEd-UcG211HDtU

[2] https://www.google.com/url?q=https%3A%2F%2Fdrive.google.com%2Fopen%3Fid%3D0B9dAMg2-6U-7QkZHcU5qU28xNkk

[3] https://www.google.com/url?q=https%3A%2F%2Fdocs.google.com%2Fa%2Fnoaa.gov%2Fdocument%2Fd%2F1WA43d2C71RUny-bES16k62Bs3kM9r5JZz3SKHnycDbI%2Fedit%3Fusp%3Dsharing

Appendix 1 of Tom Hamill's comments on 2016 UMAC

Evidence-based decision making as it applies to ensemble prediction system implementations

Some thoughts by Tom Hamill, ESRL/PSD, 4 May 2016
(updated 3 August 2016)

The UCAR review committee for the NCEP production suite, the UMAC (UCAR Model Advisory Committee) recommended in 2015, among other things, that: (a) NOAA develop strategic and implementation plans for weather and climate prediction, and (b) that NOAA model development follow an evidence-based decision making process. Our interest in this document is how this applies to ensemble prediction system (EPS) development.

We take as given that in the near future NOAA and EMC will expand upon the Next Generation Global Prediction System (NAAPS) Strategic and Implementation Plans to develop holistic Weather and Climate Modeling Strategic and Implementation Plans. These documents will outline the high-level thrusts of system development and how to make rapid progress consistent with Weather Ready Nation and NOAA Strategic Plans. Presumably ensemble experts will be consulted in the development of such a plan, and NOAA's ensemble prediction system (EPS) development will follow the guidance from such plans.

The other UMAC recommendation, an evidence-based process for decision making, is not yet in place for ensemble prediction systems. Definition of and use of such a process should help the NWS achieve better EPS products more quickly.

This document is thus offered for consideration by NWS and EMC management, a prospectus for how we could set up an evidence-based process for ensemble system implementations. What does "evidence-based decision making" mean in the context of EPS? Where does this decision making come into play in the development process? Below, we propose that there are typically four key decision points in the R2O process, points where a group of qualified individuals could be particularly helpful in making recommendations. Accordingly, this document thus suggests: (a) the potential composition of a review board; (b) the general criteria the review board should be considering when making recommendations, and (c) the decision points where their advice is needed.

1. Composition of a review board and their terms of operation.

The pre-eminent consideration for membership on the review board is a scientific understanding of ensemble prediction and its issues. While understanding of EMC's systems is desirable and at least one board member should be from EMC, we suggest casting a wider net to entrain the best subject-matter experts from other organizations, including NOAA/OAR, NASA, the US Navy, various universities, and perhaps experts from operational centers in other countries. We suggest a panel of roughly five people, not so big as to be unwieldy but big enough to provide a diversity of experience. Following World Meteorological Organization practice, members would have 3-year terms which are renewable. Nominations for membership might be reviewed at a higher level, be it by the EMC Director, the NOAA Science Advisory Board, or other. This review panel would meet in person or via videoconference as needed, with their evaluation criteria and suggested decision points described below. The panel would provide written recommendations to the EMC Director and NOAA and NWS leaders.

2. Suggested criteria to evaluate a potential ensemble prediction system development.

a. Physically based. Is the potential change to the EPS one that can be defended on scientific principles? Has the potential change been examined by relevant scientists, and have they agreed that the change makes the system more realistic?
b. Improvement. Does the system with the potential change beat a previously agreed-upon baseline across a previously agreed upon number of metrics, in statistically significant amounts?
c. Code simplicity. Does the change make the code simpler, or at least does it add complexity only in proportion to the increased physical resemblance to the real prediction system? Does the potential change make future (physically based) modifications easier? Is the code written in such a way that it can potentially be reused for other applications (regional models, climate models)? Does the code facilitate a potential reduction in the number of modeling systems that NCEP must run and maintain, allowing the model with the potential improvement to take up the product development from another modeling system?
d. Code performance. Does the potential change increase the CPU expense and/or disk storage, and if so, is the improvement in skill roughly concomitant with the increase in CPU and disk space?
e. Documented. Is the methodology sufficiently documented so that it can be maintained?

3. Proposed stages and gates for the development process.

Here we assume that the development process is split into “stages” where work is performed, and “gates” where there is a critical review and a decision about whether to proceed on to the next stage. For a given system implementation, stages/gates 1-2 may be proceeding in parallel with multiple strands of research, and stages 3-4 are more typically with an integrated system that bundles improvements that have successfully passed through the first two stage/gates. Early stages may have a higher percentage of the work performed by OUR lab scientists or university investigators, later stages nearer to implementation are likely to have greater involvement from EMC.

Stage 1: Ideation. Based on many potential sources (development at other centers, conference results, research breakthroughs in academia, and/or the recent personal work of the scientist or team), an idea is formed about how to improve the system. Presumably this idea addresses an agreed-upon need, such as the priorities outlined in the NGGPS implementation plan. The time allotted to this stage may vary significantly. The product generated at this stage is typically a research proposal, outlining the background research, the hypothesis, the proposed test plan with milestones and resources needed. The test plan is formulated with the first gate in mind, and the evaluation criteria discussed above. As the research is early on, the focus at this stage in on whether the proposed method is physically based and whether the literature provides supporting evidence of its potential.

> Gate 1: Defense of proposal. Whether this defense takes the form of peer review of a written proposal (e.g., with soft money) and/or an oral defense (e.g., for base funding) in front of a qualified panel, the expectation is that there is a decision point where higher management informed by a panel of experts either approves the project for the next stage, or not. Approval may be contingent; the researcher(s) may be asked to provide more evidence, to submit a revised proposal that trims the scope of the project, or that modifies the research plan according to guidance from the panel.

Stage 2: Preliminary experimentation. The researchers now execute their proposed test plan, developing the new improvement, testing it against an agreed-upon baseline. Likely the researchers will engage in a substantial iterative process, where they learn about deficiencies, debug code and/or modify the hypothesis and the technical approach. During this phase, the expectation is that the scientist will be regularly consulting with experts and peers, subjecting their intermediate results and code to scrutiny, accepting and acting on relevant feedback. The product generated from this stage will be a set of preliminary results and a report/presentation that addresses the evaluation criteria above.

> Gate 2: Decision on whether preliminary experimentation warrants pre-operational development. The scientist or team involved in the research presents their preliminary findings to a review panel. This panel, with strong representation from the operational center (EMC) and other relevant parties (STI, NCO, CPC, other relevant NCEP centers), will evaluate the results and make the decision as to whether the results presented show enough promise to proceed with operational development. The review panel may make suggestions for this operational development phase; for example, they may recommend testing of the system at a higher resolution, or testing in conjunction with other related developments. The primary focus at this stage is whether the preliminary experimentation provided an improvement, with some attention to code performance and code simplicity.

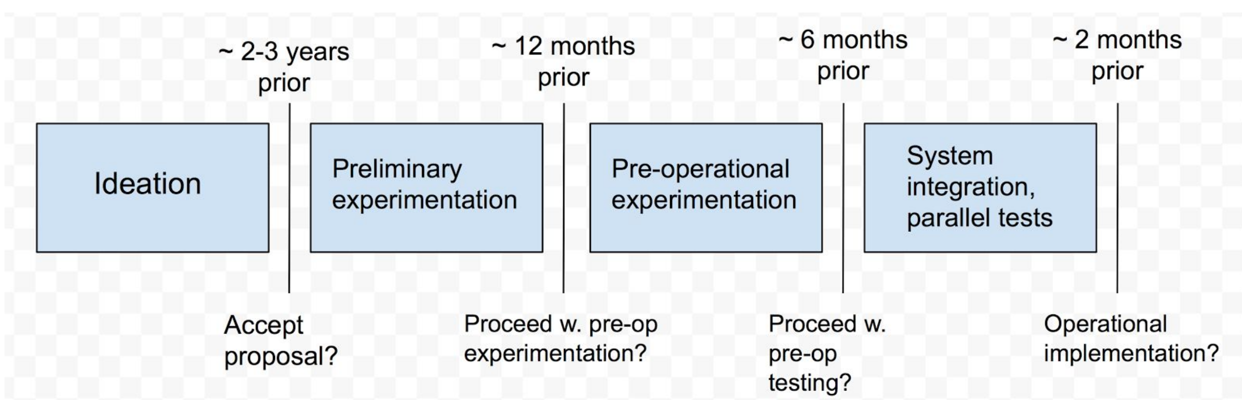
Stage 3: Pre-operational development. Presuming a positive recommendation at gate 2, in this stage the code developed in stage 2 is adapted to the operational environment and tested more rigorously. This testing might include experiments at the anticipated operational resolution, testing over a broader set of cases, testing over a broader set of metrics, and/or against more stringent reference standards (perhaps against the anticipated next version of the system, with other improvements). The scientists involved will prepare documentation for the next formal review, addressing the evaluation criteria above.

> Gate 3: Decision on whether to proceed with pre-operational testing. NWS management and chosen outside experts will evaluate the review material, making a recommendation as to whether to proceed. Other options may be discussed (return to the previous stage for more development, delay one implementation cycle, etc.). The focus is on performance, code simplicity and performance, and adequacy of documentation.

Stage 4: System integration and parallel testing. EMC staff will now take the software and documentation previously prepared in stage 3, adapting it and merging it with other proposed software enhancements that have also proceeded through gate 3. They will develop an integrated new version of the system, presumably test the integrated version to make sure all components are working properly, and validate the system performance over a moderate number of cases relative to an established reference standard. Presuming these results are satisfactory, the upgraded software is now sent to NCO for their final adaptation and is made into a formal "parallel" model run, and is compared side-by-side with the current operational model over the chosen test period. Results are synthesized into a report.

> Gate 4: Implementation decision. NWS Senior management and chosen advisors make a decision as to whether to proceed with the operational implementation. If they do, responsibility for the final steps becomes more of a responsibility of NCO.

A rough GANTT-type chart for stages and gates is provided below.



Suggested next step:

We suggest consideration of this by EMC, NCEP, and STI leadership. Assuming that NCEP is in agreement with the fundamental details here, we suggest that EMC, ESRL, and other interested parties work together to: (a) more concretely define the concept of operations; (b) draft a terms of operation for the review panel; (c) define evaluation criteria, and (d) set the initial composition of review board.

End of Appendix 1 of Tom Hamill's comments on 2016 UMAC

Anke Kamrath (National Center for Atmospheric Research):

The overall research computing investments and available allocable cycles for research are not linked to strategic goals and overall model development efforts underway. There is risk that important development efforts underway and planned can not be achieved (or only partially achieved) due to a lack of adequate compute resources. The multi-year planning required for compute resources should be done in concert with modeling efforts. It is notable that for the first time compute requirements are being gathered as part of the project management structure which will at a minimum allow the identification of "the gap" in needed and planned resources. It is also notable that efforts are made to use unused production cycles for development purposes - but this does not mitigate the overall need for additional resources and for a coupled compute and modeling investment plan.

Community involvement has been successful in many atmospheric sciences modeling efforts (e.g., WRF, CCSM) and also within NOAA - the new National Water Model (WRF-Hydro) effort seems off to a good start. Community involvement is key to the success of next generation models at NOAA/NWS if the goal is indeed to have the best forecasts in the world. There seemed to be skepticism by a few staffers involved in this UMAC about the value of community involvement - and only focused on the downsides. An inventory of successful models for community involvement should be compiled. From there a careful assessment of the appropriate governance and collaboration model should be selected that indeed achieves the key goals of the modeling enterprise. The planned Fall workshop could be a good time to do this.

The staff "reward" model within parts of NOAA/NCEP (e.g., mesoscale efforts) seems counter to the goal of having the best forecast in the world. Based on the sheer number of models and components, the "reward" model seems to be to have "your own model" and "make your own advancements". The overall staff reward model within NOAA/NCEP should be reviewed to and better understood to ensure that advancing the broader forecasting enterprise is the goal - not just personal achievement or other unintended outcomes. This will be vital for long-term success and appropriate engagement in organizational goals by staff.

<u>Jim Kinter</u> (George Mason University):

[Because I did not attend the UMAC meeting, my comments are more general, based on reviewing the materials from the meeting, conversations with UMAC members and many others, and email traffic subsequent to the meeting. The focus of my remarks is on global modeling (including the plans for unified global model development), since I have little experience in mesoscale, information technology issues, the community modeling paradigm, and strategic planning. – JK]

Global Modeling:

There has been a lot of progress on global modeling. The development and execution of a designed, evidence-based, rigorous process for selecting the dynamical core of the atmospheric component was a welcome new paradigm. The fact that it was possible to deliver strong guidance for the decision in a timely manner is a proof of concept for this paradigm, indicating that this process should be repeated for many other aspects of the global model development cycle (data assimilation, ensembles, sub-grid-scale physical parameterizations, coupled system architecture, ocean component, sea ice component, etc.). It is important to note that the selection process was designed for the purpose of advising an important decision. It is not sufficient to simply declare that there is evidence supporting a given position so that the decision is “evidence-based”. The key elements of a good design - process, requirements, evaluation metrics and timeline – were defined and employed, and that is why the dycore selection process is already viewed as a success. It is important now that NCEP embrace the FV3 dynamical core and aggressively act to integrate it into the future operational global prediction suite.

Information Technology:

A concern has been expressed about the capability and capacity of research computing in support of NCEP’s modeling mission. UCACN has recommended consistently since 2009 that, in addition to enhancing high-performance computing (HPC), it is critically important to strike a balance at NCEP among HPC, disk storage and archival storage (as well as software to support data management and workflow). The inadequacy of disk capacity at NCEP has been an issue for much longer than that, such that the few data sets generated in R&D mode that can be saved have often gone unanalyzed, stored on tape (effectively, write only memory). Unfortunately, it is harder to make “the business case” for adequately provisioning data management and workflow for research, which UCACN and UMAC members clearly feel is critical.

It should be noted that about 8 years ago, NCAR’s HPC advisory committee recommended at that NCAR should take into account the balance of compute capacity and data management capability in their next procurement. NCAR management heeded that advice and used a large portion of their money to deploy data analysis facilities (servers called Geyser and Caldera and a data system called Glade with an order of magnitude more disk storage and a hierarchical storage management system, all connected at high bandwidth to the Yellowstone supercomputer). This bold decision required that some of the funds allocated by NSF for NCAR computing be used to obtain something other than HPC capacity. The fact is that Yellowstone is a smaller system than NCAR could have afforded had they spent the full NSF award on HPC. The computer group at NCAR took some heat for that decision, because consequently NCAR didn't have as big a supercomputer as some of its peers. However, the result is greatly enhanced productivity and researchers who are much better equipped to handle data management issues.

Community Modeling:

NCEP's long-standing reputation for being an insular shop has been eroding over the past several years, and the development of NCEP and NWS testbeds, and, importantly, the involvement of the UCAC, the UMAC, the external participants in the dynamical core evaluation team, the external principal investigators in the NGGPS program and the external principal investigators in OAR's labs, grant programs and task forces, are all signs of NCEP's commitment to seek and use the advice and collaborative contributions of the outside community. This is a very positive development that is most welcome and should be lauded and encouraged.

It is time to take the next step and create an organizational structure for community modeling that more robustly involves the relevant and willing community members to realize the vision: *Operational prediction employs the best possible numerical forecast system, embodying all relevant and evidence-based knowledge at any given time.* There is a widely accepted view in the community outside NCEP that accelerating progress in weather and climate prediction requires a broader community perspective (viz. several NRC reports, UCAC and UMAC reports of past years, etc.). On the other hand, those in the research community and those responsible for operational prediction have very different constraints in resources, timeframe, and skill sets. The input coming from the research community has been largely constructive, but there is a potential for broader discussions of strategic planning and governance to appear to be a threat to the operational prerogative and control. Furthermore the perceived definition of "community modeling" seems to vary considerably: some people consider the status quo (e.g., involving some external PIs in NGGPS) as sufficient, and others consider large-scale efforts like CESM as the appropriate organizational model. There are also intermediate structures that are being discussed.

Therefore it is critical that the plan for transforming model development at NCEP into a community organizational model should have the following characteristics:

- inclusive and encouraging;
- well-designed (established and agreed process for interaction and communication, governance model, requirements, evaluation metrics and timeline)
- adequately resourced

- appropriately equipped with infrastructure, documentation and support.

Specifically on the last point, it is critical that the development tools, which include model components, coupling infrastructure, data assimilation capability, test harness, validation data sets, run scripts and other ancillary software and data resources – must be maintained and supported in a common repository that is open, accessible and available to both the operational developers and relevant, willing researchers. This will enable researchers to conduct their research with the same tools that are currently used in operations, automatically making their work more relevant to NCEP and smoothing the transition to operations. It will also permit the operational developers access to the latest research results, shortening the R2O timeline.

Importantly, the requirements for the community-modeling partnership structure should be drawn from both the research and operational development sides. In some cases, the requirements that come from the research side will conflict with operational requirements, e.g. architectural flexibility vs. optimal performance. As a result, the strategic planning process and governance model must take such potential conflicts into account (e.g., how are they discussed, resolved?).

As an aside, right-sizing a community modeling structure is important – should the structure support only a small, select team of developers, or everybody including a very broad army of graduate students, or something in between? The remarks above suggest that the community modeling organizational structure should be open and inclusive, but it need not necessarily be supported at the level of, say, the CESM, for which there is an abundance of support, including help desk staffing, for relative novices with only a modicum of training in Earth system science. The specific focus on prediction is likely to be self-limiting to those willing researchers whose work is relevant, but some attention should be paid to the question of scope in the design.

As another aside, there is a tension between research-driven development and implementation-driven development. On the one hand, one would want to implement changes when research has shown that there is a demonstrably superior way. That pace of development is not regular nor easy to predict. On the other hand, there is an expectation that developments are continuously implemented so that things are constantly getting better and users perceive the improvement. The latter drives incremental changes, making radical changes harder to implement, despite any demonstration of superiority. Also, operations requires decisions ("this is out, that is in"), which is a somewhat alien concept to researchers ("we could tweak it some more and get much better"). The strategy for establishing a community modeling organizational structure must account for this tension.

In the spirit of evidence-based decision-making, a survey should be made of the relevant community efforts (CESM, WRF, WaveWatch-3, ECMWF, Met Office, etc.) to determine the pros and cons, lessons learned and best practices. This survey should advise the creation and design of the community modeling organizational structure.

Specific Comments on UGCS:

The presentation by Suranjana Saha laid out a plan for producing predictions at all time scales from data assimilation to “weather” (up to 10 days; aside – why not 2 weeks?) to sub-seasonal (up to 6 weeks) to seasonal (up to 9 months), in which all predictions are ensemble-based, a continuous cycle of coupled reanalysis and reforecasting is employed, and a 7-component coupled model (atmosphere, ocean, sea ice, land surface, ocean waves, atmospheric chemistry/aerosols and ionosphere) is the forward integrator. This follows the outline that was provided in response to UMAC recommendations, which focused on products needed at different lead times. This plan is consistent with the scientific consensus that probabilistic prediction, employing ensembles, and full representation of processes and feedbacks, employing a coupled model, is the way to go at all time scales. That is, single runs and using an uncoupled atmosphere-only model are no longer adequate. What is not provided in the presentation is a clear strategy for involving the community (this is labeled a “challenge”), nor is there a timeline for development and implementation. Importantly, the coexistence of the current seasonal prediction system (CFSv2), development work on the next implementation of that system (NFSv3) and the development work on the Unified Global Coupled System (UGCS) is not mentioned in the presentation, but is, in some ways, the elephant in the room. There needs to be a plan with a timeline indicating when CFSv2 transitions to CFSv3 and then to UGCS.

Strategic Plan:

As others have commented, there is an urgent need for a forward-thinking strategic modeling plan. The plan should be actionable, i.e., not overly broad or lofty, and it should include definitions of governance, processes for gathering requirements, processes for interaction between the research and operational development communities, and metrics of success. An important part of the plan should be the process for defining the architecture of the NCEP Production Suite. It should be accompanied by detailed multi-year implementation plans that are consistent with the strategic plan and link to annual operating plans. This process is well-established and its effectiveness has been demonstrated, e.g. at ECMWF where a 10-year strategic plan is revised every 5 years and is linked to multi-year implementation plans. The strategic plan should also be accompanied by a timeline or roadmap that clearly discriminates between projects and initiatives that are strictly short-term (“patch”) developments and long-term (“strategic”) developments, including a clear definition of the time ranges for short- and long-term. As UMAC has recommended in the past, the strategic plan should employ a two-time scale approach for phase 1 – separate strategies for unified global modeling and unified mesoscale modeling – and phase 2 – a unified modeling strategy across time and space scales. The plan should include components for quality control, data assimilation, model development, testing and diagnostics, and linkages to products.

Bill Kuo (National Center for Atmospheric Research):
Bill Kuo was represented by Josh Hacker at this meeting. Josh Hacker's comments are above in this document.

Gary Lackmann (North Carolina State University):

I am an ex-officio UMAC member, by virtue of my status as U CAN co-chair. My comments therefore may be considered in a slightly different category relative to those from the original UMAC members. Nevertheless, I am very familiar with the UMAC report and recommendations.

Simplification of production suite, leadership, and strategic planning:

Especially in the area of regional modeling, there remains an urgent need to reduce the number of modeling systems. This was one of the key recommendations provided by the 2009 UCACN review of EMC, and it was repeated in the UMAC report. The reasons for lack of progress on this recommendation are complex, but a NOAA-HQ-level NWP chief scientist position of the type recommended by the UMAC appears to be essential for progress. Some difficult decisions must be made, and buy-in is required at all levels within the EMC organization regarding the need to reduce complexity.

The proliferation of modeling systems is also symptomatic of a lack of strategic planning. Developing a strategic plan was also a recommendation of the 2009 review, and going even farther back, was also a strong recommendation from the 1997 NCEP review. The reasons for the lack of progress in this area are likely numerous and complex, but again the idea that personnel throughout the organization must understand the need for planning, and that the implementation of such plans is essential to ensuring a world-class outcome.

High-resolution mesoscale ensemble

The cornerstone of the regional modeling system should be a convection-allowing ensemble. There is much evidence supporting the value of such an ensemble system, and there should be no delay in moving towards testing and implementation. The UMAC has seen a proposal to extend the HRRR system to construct a HRRR Ensemble (HRRRE), by the ESRL-led team. In contrast, at our recent UMAC meeting, we also heard a proposal to implement a 3-km ensemble as part of a 5-10 year plan. This is too slow, because by that point (the years 2021-2026), global models will be moving to such mesoscale resolutions. At our meeting, we heard from a "super stakeholder", representing many NWS forecasters. He indicated that forecasters are ready to utilize this kind of ensemble immediately (even while lacking adequate software to display and analyze the ensemble). The HRRRE proposal from the ESRL-led team should be evaluated and discussed, with testing beginning as soon as possible. Other modeling systems should be critically evaluated, and several must be phased out.

A state-of-the-art high-resolution ensemble will not in of itself produce high-quality public forecasts. Thus, the move to a high-resolution ensemble system must be accompanied by aggressive development of ensemble-friendly display and analysis software in the Weather Forecast Office (WFO) field offices, and a plan to provide rapid and efficient data access at high resolution. The convection-allowing ensemble should also be accompanied by a reforecast effort (perhaps for select cases or time slices) and be designed to facilitate innovative product development (e.g., MOS-like ensemble guidance and graphics). These ensemble data should be made easily accessible to the community and real-time post-processing (at a basic level) should be performed in order to facilitate ensemble utilization and related innovations.

Strengthened and modernized verification efforts, moving beyond traditional metrics and skill scores, should also accompany the ensemble implementation. The use of object-oriented verification methods (e.g., MODE-time domain and other MET-package capabilities from the DTC) should be aggressively pursued. Identifying a set of meaningful metrics should also involve stakeholders, especially including field forecasters.

NCEP must continue to work towards optimizing engagement with the research community in ensemble development efforts, including collaborations with NCAR, NOAA labs, academic institutions, private sector, etc. These collaborations should be supported by funding to help engage the research community on related science problems. A focused funding call aimed at specific operational research problems (O2R), similar in structure and management to the model used by the CSTAR program, would go a long way towards strengthened community engagement and accelerated R2O.

Work to build a community around FV3 model core:

The operational global modeling community that could potentially work in collaboration with NCEP to develop and improve a NWP system built around the FV3 dycore has not yet been clearly identified. NCEP must facilitate discussions designed to identify and motivate the desired research community, and then engage with that community to identify additional requirements and desirable attributes for the community modeling system. This could initially take the form of a town hall or other forum to collect feedback and gauge the interest of the community in learning about, accessing, utilizing, and eventually enhancing the NCEP global modeling system.

A related recommendation is to identify the most important areas, critical to operational forecasting, where the research community could contribute to model development (e.g., scale-aware physics); this can be viewed as “O2R”. Then establish a funding mechanism to accelerate this work, including a funded visitor program, and proposal calls highlighting opportunities for research problems of operational-relevance. A clearly defined R2O strategy, along the lines suggested by others, is essential for the effectiveness of this process. Testbeds such as the Developmental Tested Center (DTC) should be actively engaged in all of these tasks, beginning with efforts to document and train users with the global modeling system in preparation for making it available to the broader community.

Tsengdar Lee (National Aeronautics and Space Administration):

I only participated in the first day of the UMAC face to face meeting on the first day. I would stay at the high-level and provide my observations and comments based on what I had seen and heard. First of all, I am very pleased to see very significant progress in the planning and organizing the global modeling activity at NCEP. The NGGPS has come a long way. The decision to use FV3 dynamical core is critical and it is solidly based on the evidence from a rigorous testing. I recommend that the NOAA instigate a formal process to make this evaluation process a community process. Making a community process is not just building a testbed. It is more about providing a transparent evaluation suite with machine executable end-to-end workflow in a commoditized computing environment. This way, members of the community may choose to participate by building and expanding on the baseline, testing and evaluating of the results, publishing and advocating for the changes in a well controlled manner. The workflows and test suite does not have to be static. But there should be a transparent processes in updating the workflows and test suites that involves the community. The use of open Request For Information process should be seriously considered.

Since the FV3 core will be the dynamic core for the next generation NCEP global model, strategic collaboration with NASA Global Modeling and Assimilation (GMAO) should be enhanced. NASA GMAO has used FV3 for more than 10 years. Various model products (reanalysis and nature run) and evaluations (papers and detail tech reports) have been available for some time. GMAO also shares the same data assimilation system which originated from and co-developed with NCEP. In addition, the Land-Information System (LIS) developed at NASA Goddard Space Flight Center has been integrated into the current GMAO model. The Land-Information System is becoming a core component of the National Water Model which is an added benefit. Already, NASA GMAO has tight relationship with GFDL SJ Lin and basically can be an early indicator for many of the new developments by NGGPS. NASA is also testing additional physics and model features. More specifically many aerosol and radiatively sensitive atmospheric chemical species are added to the model. There are plans to produce high resolution long-term simulated products for community consumption. All of these efforts can be leveraged by the NGGPS implementation team.

I am shocked by the unwillingness to change by the mesoscale groups. I fully appreciate that there are so many end users depending on the routine output of model products to feed their applications. However, the status quo approach is not going to be sustainable for long. Strategic thinking and scientific and technical leadership will be required. While the global model side has found a direction and starting to implement a plan (yes, I know the plan is not well documented and communicated) it is time for the mesoscale groups to think about a strategy to adopt a similar organization structure and strategic direction to develop the mesoscale modeling plan. I would argue to a seamless model - FV3 dynamical core with some kind of multi-scale model physics used for both global and mesoscale. NASA GMAO has experimental development (with GFDL) on this front also. The attached figure shows the current direction NASA is moving. In this case, the mesoscale model is simply a locally refined inner grid within the global model.

NCEP, by building a stronger collaboration with GFDL and NASA GMAO, may be able to reach a critical mass and expedite the seamless modeling approach. The collaborations may form a core effort and become a nuclei for new developments.

Cliff Mass (University of Washington)

Specific recommendations include:

1. **Global Modeling**

Although considerable effort was made to design the NGGPS process to be evidence-driven, there were substantial deficiencies in the process, such as the lack of emphasis and testing for convective-allowing resolutions and the heavy use of aged GFS physics for key tests. Essentially, the testing was not forward looking. Considering that the decision has been made to adopt the FV-3 dynamical core, FV-3 should immediately be tested at convective-allowing resolutions and compared to MPAS or other platforms. If the grid structure or other aspect of FV-3 is found deficient, an effort should be undertaken to remedy the problem (e.g. regridding or replacement).

The dynamical core selection is only the first step in building the next generation global operational NWP system for the U.S. In order to build the best global prediction system in the world, NGGPS requires the contributions of the entire U.S. modeling and research enterprise. Thus, the new global system MUST be a community modeling system. To do so requires that the NWS develop the community governance and cooperative structures that will foster such a community system. During the next few months, the framework for a national cooperative effort on global prediction should be established, with strong inputs from academic and private sector communities. This includes community governance structures, code sharing, maintenance and version control, support structures, verification and evaluation tools, workshops, and financial support/organization of model improvement. DTC is in unique position to aid in many of these areas.

Since the FV-3 system will not be available immediately, GFS physics should be improved immediately (e.g., microphysics and shallow convection) and the GEFS ensemble should be enlarged, with resolution dropped to approximately 15-20 km grid spacing, thus allowing the removal of the similar resolution SREF.

The CFS model is probably the most clear-cut example of NCEP **not** following an evidence-based approach. It has little skill beyond 2-3 weeks and statistical approaches are superior beyond that time. Yet, large amounts of computer resources are used for CFS, whose improvement pace is slow at best. This effort should be carefully evaluated and restructured.

2. Convection-Allowing Model Ensemble System

A wide variety of NWS stakeholders, internal NWS requirements, and a long list of advisory committees, workshops, and NRC reports have called for an operational convective-allowing ensemble system over the U.S. One group after another supported the development of a such a system during the UMAC meeting. Progress has been slow for a number of reasons, including the lack of HPC resources. The situation is now amenable for rapid progress in creating an operational convection-allowing model system by the NWS. Massive new computer resources are in place, extensive testing of prototype systems has been quite successful (by SPC, NCAR, ESRL, and others). Substantial new research has shown that stochastic physics can provide an effective approach for producing physics diversity. The NWS initiate build a prototype operational convection-allowing model ensemble during next year that will go operational by the end of 2017.

3. Simplification of the Regional and National Scale Modeling at EMC

A major recommendation of the previous UMAC report is that the regional/national scale forecasting suite be substantially reduced and simplified. Little or no progress has been made to address this recommendation. Today, EMC runs a complex, inefficient, redundant, and difficult to support collection of models, domains, and modeling approaches over the U.S. This collection of runs needs to be reduced substantially during the next year to release personnel and computing resources for more useful applications, such as the convection-allowing ensemble system noted above. The complex collection of regional nests should be replaced by national domains, the NAM-RR effort should be terminated, and SREF replaced by a larger size/higher resolution GEFS with stochastic physics. Serious consideration should be given to going to a single dynamical core, with WRF-ARW being the preferred candidate.

4. Strategic and Implementation Planning

NOAA should develop a 5-10 year strategic plan for environmental prediction and should do so within the next 6 months. This plan should lay out a vision with specific major goals (such as convection-allowing global ensembles by a certain date), as well as major research needs for that period. In addition, NOAA needs to develop highly detailed implementation plans for the 1-3 year time horizon, giving detailed guidance regard tasks, budgets, and personnel. Finally, NOAA needs to develop a list of explicit requirements that US NWP must fulfill to successfully meet national needs.

Peter Neilley (The Weather Company/IBM):

A. Major Conclusions and Recommendations

A.1 **Global Modelling**: NCEP has made wonderful progress in the NGGPS process and has largely followed the guidelines for modernization outlined in the previous UMAC report, particular in following a scientific, evidence-based decision process. While much is still left to be accomplished before successful implementation of a FV3-based model, there is reason for considerable optimism. However, there are also some reasons for concern, particularly in successful engagement of the community and ability for the FV3 to support convective allowing resolutions with skill. Recommendations:

1. NOAA should actively seek to catalyze a broad community of academic, private-sector, government and international partners in forming a vibrant and active community of users and developers of the FV3-based global model. NOAA should not dictate specific outcomes of this community. Rather, it should trust that a broadly engaged community will contribute directly and indirectly to the long-term success of the NGGPS across all scales. It is important that this community be created and run by the community and not by NOAA in order to ensure that it achieves the broadest possible adoption. However, NOAA should take a leadership role is sponsoring and facilitating the community. Proposals to organize such a community under the new global model testbed housed at UCAR seem promising.
2. NOAA should undertake a scientific program to access, understand and perhaps improve the ability of a FV3-based model to successfully forecast convective systems. The dycore selection process, while considering some convective-allowing circumstances, did not adequately assess the ability of the FV3 to handle convection. If such an assessment fundamentally undercovers limitations with the FV3 to handle convection skillfully, then a program to improve the FV3 or identify an alternative dycore must be undertaken to prepare the nation for convective allowing global ensembles which we expect to be performing operationally within the next 5-7 years.

A.2 Regional Modelling: In contrast to the NGGPS program for global modelling, NCEP has made no material progress towards the urgent need to simplify its regional modelling portfolio. The complexity of the regional modelling suite is staggering, and it is the single most significant barrier to our nation making progress towards the weather ready nation objectives. It is difficult to understand why NCEP has not make progress towards this despite numerous and continuous reports from UCAC/UMAC and others to simplify the portfolio for nearly a decade. Recommendations:

1. Immediately cease all R&D activity on any regional modelling not connected to WRF-ARW based models, and focus those resources on developing a ARE-based portfolio of regional modelling capabilities. While this recommendation is inconsistent with the “evidence-based decision making” approach strongly advocated by UMAC, the urgency and critical nature of the current situation requires such a drastic action based on an overwhelming majority opinions of the community of users and model practitioners

based on real-world on the ground experience in using the models in addition to a body of scientific evidence support it.

2. Quickly develop a 5-10 year strategic vision and plan for regional modelling based on a single regional modelling core. This plan should accommodate the critical need for a convective allowing ensemble in the first few years. The plan should also allow for eventual unification of the regional modelling with the global FV3 dycore provided that sufficient evidence is found that the FV3 convective forecasts are competitive or superior to ARE-based convective forecasts.

B. Other comments and recommendations.

B.1 HPC architecture. In order to ensure optimal use of computing resources, NOAA is encouraged to invest in optimizing the production suite to take advantage of modern HPC architectures more. NOAA should consider an assessment of specific models and which type of computing architectures are needed and optimal for that specific model. In particular, some of the 2-D models used in hydrological prediction may be satisfactorily run on architectures (e.g. GPUs) that are different than that optimal for full 3-D models where memory bandwidth, and interprocessor communications are critical. NOAA should explore hybrid computing solutions that enable different modeling suites to be run on different computing architectures that are economically optimized for that model, and move away from a “one-size fits all” approach to HPC.

B.2 Making Tradeoffs based on value to the Nation: Inevitably there are tradeoffs in computing and human resource allocations towards the spectrum of possible advances in the production suite, such as the balance of resources spent between long-term and short-term forecasting initiatives. Optimally, these decisions should be made based on an objective assessment of the value these initiatives will return to the nation and not on pure meteorological assessments of skill. NCEP (and NOAA in general) could benefit from a more quantitative processes involved in assessing the payoff of its strategic decisions to the Nation for example by incorporating the economic sciences systematically in those assessments.

B.3 Water Program. The Water Center program has made impressive progress in the past year in creating a pipeline of material improvements in our total water related forecasting needs including the upcoming release of the high-resolution streamflow prediction product. Computing needs to meet The Water Center’s planned portfolio in the next five years remain uncertain and as such there is concern that these initiatives may encroach on resources for other high priority programs (such as convection allowing ensembles). NOAA should develop a 5-yr computing roadmap for the Water Center so that resource conflicts across the entire portfolio can be identified and strategically resolved in a manner consistent with the previous (B.2) recommendation.

B.4 Situational Use of Computing Resources. Consistent with the objectives of the Weather Ready Nation strategy, it is imperative that we are positioned to perform our best during high-impact weather threatening the nation. As such, strategies need to be developed to allow more situationally-dependent operations of our production suite to focus forecasting assets on daily weather problems of highest priority. While some of this is ongoing today, it is largely accommodated for by taking advantage of unused computing resources sitting idle in the WCOSS. However, sufficient unused resources will not always be available for optimized forecasting of certain weather situations. Therefore, strategies should be developed to shed some lower priority operations on the WCOSS to ensure the optimal set of forecast products needed for high-impact events are made. This may involve changing the SLAs NOAA currently has with its stakeholders, and as such will inevitably be contentious. NOAA has had a tendency to avoid changes in its operations that lead to such contentious issues, even with a single stakeholder, but that needs to change. NOAA should implement variable stakeholder SLAs in order to allow it the flexibility to optimize its services to the nation during high-impact weather. There is precedent for this paradigm with GOES rapid-scan satellite operations during severe weather that supplant other hemispheric imaging schedules.

B.5 Requirements. NCEP/NOAA uses the term "requirements" to direct and defend most of its activities. However, the process to collect, prioritize, and translate requirements into activities remains confusing and elusive to the community. Simplicity and transparency in the requirements process is needed in order for the community to have trust and respect in the requirements NOAA purports.

Christa Peters-Lidard (National Aeronautics and Space Administration):

I appreciate the efforts NWS has put into responding to our previous report, particularly the comprehensive response document and traceability matrix. My general impression is that the global efforts are moving in a good direction, albeit with significant work ahead, while progress on the mesoscale efforts has been lacking. The National Water Model has also progressed rapidly, and there is a strong need to integrate with the "salty" side of NOAA.

Below are my specific recommendations:

1: Support evidence-based evaluation of National Water Model:
The deployment of the National Water Model has occurred at a rapid, yet necessary, pace to demonstrate the feasibility of the effort. It is essential that an open evaluation and benchmarking service, with clearly defined requirements and metrics, be deployed as a highest priority in FY17 to provide the evidence that this approach provides skill comparable to the current operational "forecaster in the loop" RFC forecasts.

2: NOAA Water Roadmap and Summit:
The development of a NOAA Water Roadmap (expected October 2016) will provide an opportunity to convene a NOAA Water Summit including NWS (EMC, OWP), NOS and relevant stakeholders and partners to present and discuss this roadmap with the community. The report

should be shared with UMAC and/or UCACN should augment its membership in advance of this summit and take a leading role in providing feedback on the NOAA Water Roadmap.

3: NGGPS architecture, project management, community engagement:
The developing NGGPS architecture and physics teams should look to other related efforts such as the NASA/GSFC GMAO GEOS-5 and NCAR CESM for examples of robust Earth system model architecture design and community engagement. I recommended looking to CESM as a model for “community” since unlike WRF, that community is forced to choose a unified physics suite for CMIP and related assessment reports. The global unification effort has hard work ahead to develop a true earth system model based on the “best” architectural design (e.g., GMAO) and strong leadership in each of the subsystems, such as land, sea ice, ocean, atmosphere and data assimilation. Right now, it appears that the NEMS/NUOPC Technical Manager is performing effectively as the Project Scientist for the NEMS/NUOPC effort, but without the authority to make decisions. Borrowing from the NASA Mission management model, any complex project has both a project scientist and a project manager. The project scientist and the project manager must have the authority to make decisions that balance schedule, budget, and scope, independent of the scientists'/model developers' individual interests. Branch heads or other management should NOT be in these roles.

4: Simplify mesoscale suite:
All mesoscale systems should be “frozen” immediately, so that all new development can focus first on developing an ensemble HRRR/ARW and eventually on a mesoscale FV3.

5: EMC Organization:
EMC needs robust teams for atmospheric physics, land, ocean, sea ice, including data assimilation for each component. For example, the land team needs a new federal lead, since the current lead has been promoted to EMC Deputy Director. With new EMC leadership, the opportunity exists to completely reorganize EMC to be more forward looking, and get away from the separate global/meso structure.

6: Data Assimilation:
The current vision for JEDI is focused on developing unified forward models across the EM spectrum. While this makes sense for most satellite sensors (make sure you include L-band SMAP/SMOS), a comprehensive approach to data assimilation must also support other data types such as GRACE (gravity), ground-based rainfall radar data, and point data such as rawinsondes, surface met stations and stream gauges. Further, work at NASA/GMAO has shown that these forward models require extensive calibration, due to large uncertainties in the parameters.

<u>Richard Rood</u> (University of Michigan):

There has been substantial progress in a systems approach to the applications in the suite of global applications. Notably, there is improved communication across the EMC applications leads, including regular interactions with the broader community. A formalized requirements process and project management practice have been implemented.

This progress is anchored around the NGGPS Overarching Systems Team and NOAA Environmental Modeling System (NEMS) meetings on coupled models. The Overarching Systems Team is, formally, associated with NGGPS; however, its reach is farther than NGGPS. Some regional applications have been working within this Project; however, the reach does not extend to entire portfolio of regional models. At this point, the reach should not include all of the models associated with the Meso-Unification / Regional efforts of NOAA. The Meso-Unification / Regional models need some fundamental strategic planning, including culling of systems, and development of project management practices before they can participate productively in a NOAA-wide strategy.

The NOAA Environmental Modeling System (NEMS) has advanced to a state where it is operational with some of NOAA's system. There are benefits realized in software management, agility in building coupled systems, computational efficiency, inclusion of scientific evaluation into software design and development. and new paradigms within EMC of shared decision making. These initial successes need to be recognized and built upon, and used as a model to inform improved management practices across the institution. I note that in ECMWF's strategic plan for the next decade, they state, "An integrated modelling framework for representing the Earth system including the atmosphere, oceans, land surfaces and the cryosphere is paramount to achieve a seamless analysis and prediction capability." NEMS is such a framework and offers NCEP a strategic asset that if well exploited will contribute to closing the forecast gaps.

NGGPS stands as a managed project within NOAA, with several successful elements. With the choice of the dynamical core, the project is moving into an implementation phase for the global atmosphere model. Decisions in this implementation phase have profound consequences; they are complex, but they are also urgent. Crucial decisions need to be made about community (code management, shared decision making), design and architecture of the global system, and the roadmaps for assimilation and physics.

I note that the progress within NGGPS, often, does not appear wanted by those in EMC. EMC needs to internalize NGGPS and NEMS as its future. There really are no alternatives, and continuation of the model research, development, and implementation practices of previous decades will not yield a world-class, global forecast capability.

With regard to the many regional models, the UMAC meeting placed the leads to many of these modeling activities together. This was very revealing. The development of a strategic approach to reduce the complexity should include these leads, with very specific questions on identifying redundancy and how to reduce complexity. It is clear that there is a culture that supports, if not rewards, individuals and small groups each believing that some idiosyncrasy of their approach offers advantage. This excess complexity and fragmentation needs urgent, remedial attention. It is not organizationally or programmatically justifiable. The UCACN and the UMAC have been giving consistent message on this issue for years, with no apparent progress. If the status quo is maintained, it is perhaps right to assume that this message is unwanted, and the committees should discontinue their service.

Similar to the issues with regional modeling, the UCACN and UMAC have provided consistent messages on storm surge, fresh, and salt water modeling. There is evidence that organizations (e.g. U.S. Geological Survey) and individual researchers are starting to provide customers with more effective products. Again, the UMAC and the UCACN have been giving consistent message on this issue for years, with no apparent progress. If the status quo is maintained, it is perhaps right to assume that this message is unwanted, and the committees should discontinue their service.

Acronyms

The UMAC model glossary can be found at this link:
https://www.earthsystemcog.org/projects/umac_model_advisory/Model_Glossary

ARW: Advanced Weather Research and Forecasting Model
CESM: Community Earth System Model
CFS: Climate Forecast System
CFSv2: Climate Forecast System version 2
CFSv3: Climate Forecast System version 3
CONUS: Continental United States
CPC: Climate Prediction Center
CPU: Central Processing Unit
CSTAR: Collaborative Science, Technology and Applied Research
DTC: Developmental Testbed Center
DTG: Dynamical-core Test Group
ECCC: Environment and Climate Change Canada
ECMWF: European Center for Medium-range Weather Forecasts
EFS: Ensemble Forecast System
EMC: Environmental Modeling Center
ESRL: Earth System Research Laboratory
FEMA: Federal Emergency Management Agency
FV3: Finite Volume 3 (Dynamical core from GFDL)
FVCOM: Finite Volume Community Ocean Model
GEFS: Global Ensemble Forecast System
GFDL: Geophysical Fluid Dynamics Laboratory
GFS: Global Forecast System
GMAO: Global Modeling and Assimilation Office
GRACE: Gravity Recovery and Climate Experiment
GSI: Grid-point Statistical Interpolation
GSM: Global Spectral Model
NHC: National Hurricane Center
HFIP: Hurricane Forecast Improvement Program
HPC: High-Performance Computing
HRRR: High Resolution Rapid Refresh
HRRRE: High Resolution Rapid Refresh Ensemble
HWRF: Hurricane Weather Research and Forecast system
HWT: Hazardous Weather Testbed
JCSDA: Joint Center for Satellite Data Assimilation
JEDI: Joint Effort for Data assimilation Integration
LIS: Land Information System

MET: Model Evaluation Tools
MM5: Fifth-Generation Mesoscale Modeling System
MODE: Method for Object-based Diagnostic Evaluation
MOS: Model Output Statistics
MPAS: Model for Prediction Across Scales
NAM: North American Mesoscale modeling system
NCAR: National Center for Atmospheric Research
NCEP: National Centers for Environmental Prediction
NCO: NCEP Central Operations
NEMS: NOAA Environmental Modeling System
NEWP: Numerical Environmental and Weather Prediction
NGGPS: Next Generation Global Forecast System
NHC: National Hurricane Center
NMM: Nonhydrostatic Multiscale Model (on the B grid)
NOAA: National Oceanographic and Atmospheric Administration
NOMADS: National Operational Model Archive and Distribution System
NOS: National Ocean Service
NOS/CO-OPS: NOS/Center for Operational Oceanographic Products and Services
NOS/OCS: NOS/Office of Coast Survey
NPS: NCEP Production Suite
NSF: National Science Foundation
NSSL: National Severe Storms Laboratory
NWC: National Water Center
NWP: Numerical Weather Prediction
NWS: National Weather Service
NUOPC: National Unified Operational Prediction Capability.
OAR: Office of Oceanic and Atmospheric Research
OWP: Office of Water Prediction
R2O: Research to Operations
RAP: Rapid Refresh Model
ROMS: Regional Ocean Modeling System
SFE: Spring Forecasting Experiment
SMAP: Soil Moisture Active Passive
SMOS: Soil Moisture Ocean Salinity
SREF: Short Range Ensemble Forecast
SSEO: Storm-Scale Ensemble of Opportunity
SLOSH: Sea, Lake, and Overland Surges from Hurricanes
STI: Office of Science and Technology Integration
UCAR: University Corporation for Atmospheric Research
UCACN: UCAR Community Advisory Committee for NCEP
UGCS: Unified Global Coupled System
UMAC: UCACN Modeling Advisory Committee
USACE: U. S. Army Corps of Engineers

WCOSS: Weather & Climate Operational Supercomputing System
WRF: Weather Research and Forecasting Model
WW-III: Wavewatch - III

Below, the agenda, its modification, and attendees are attached.