NOAA Community Modeling & Strategic Implementation Plan (SIP) Coordination Workshop:

April 18-20, 2017 – NCWCP, College Park, MD

Introduction

NB: This file is the first of a two-part workshop summary report. For part two see file: NOAA COMMUNITY-SIP X-WG WORKSHOP Summary (Part 2-of-2).pdf

On April 18-20, 2017, approximately 229 people (177 in person; 52 virtual) from the US weather enterprise participated in a workshop at the National Center for Weather and Climate Prediction (NCWCP) in College Park, MD, to build the NOAA Weather Modeling Community and forward its development of a Strategic Implementation Plan (SIP). Part 1 focused on the community engagement and Part 2 of the workshop emphasized coordination across the 14 extant SIP working groups. This workshop advances the work of the Next Generation Global Prediction System (NGGPS).

Additional information and materials produced during and/or presented at the workshops can be found at:

- https://www.weather.gov/sti/stimodeling_nggps_presentations
- https://ral.ucar.edu/events/2017/community-modeling-workshop

Workshop - Part 1: Community Modeling

Day 1 Summary

The first day featured lively engagement and discussion with the diverse and large group of participants from across the U.S. weather modeling enterprise (public, private, academic). The workshop attendees heard programmatic and technical perspectives and status from senior NWS leadership (Ji/STI and Farrar/NCEP respectively). Then an esteemed panel of leaders from NOAA (Uccellini, McLean), DOD/Navy (McCarren), NASA (Peters-Lidard), NSF (Harr), UCAR (Busalacchi), Academia (Carr, OU), and the private sector (Glackin, IBM/The Weather Company), shared their perspectives on the goals and needs of a community modeling system. The rest of the day walked through overviews and discussions with the SIP/NGGPS Working Groups (WG) through moderated panels:

- Governance WG and System Architecture WG
- Infrastructure WG and Global Modeling Test Bed (GMTB)
- Dynamics/Nesting WG, Model Physics WG, and Convective Allowing Model (CAM) WG
- Data Assimilation WG, Ensembles WG, Verification & Validation (V&V) WG, and Post-Processing WG
- Marine Models WG, Land Surface Models (LSMs) + Hydrology WG, and Aerosols/Atmospheric Composition WG

Detailed notes and additional synopses are provided in Appendix A.

Day 2 Summary

Having taken a high-level walk through the NOAA Modeling Community as it currently stands on day 1, the morning of day 2 provided an opportunity for a deeper dive into community in eight smaller groups. Each group was asked to consider the same two questions: breakout topic 1 was, in essence, what is NOAA doing well; whereas the breakout topic 2 asked what are we missing or could do better.

Appendix B provides summaries from each of the 8 groups. At the highest level:

BREAKOUT TOPIC 1: In terms of overall direction of moving towards a community approach to modeling, what appears to be on the right track?

- Most strongly noted has been the increase in transparency and efforts to engage the community (e.g. this workshop)
- Also recognized by many are the positive changes in leadership within NOAA, the strategic planning, recognition of challenges, and the moves towards modernizing software and infrastructure
- This has led to improved processes and practices (trends in the right direction)
- Finally the role of NGGPS has been recognized as having been successful to date, and serving as a bridge moving forward

BREAKOUT TOPIC 2: What are we missing? What gaps remain?

Not surprisingly, there are a number of things NOAA should address or improve. These include:

- Getting a lot more specific with respect to governance and implement governance quickly: we're heading in the right direction and lots of work has been done, but we have a long way to go
- This leads to a related theme of process; we need to implement best practices (e.g. defining benchmarks/baselines), tools (e.g. unified verification and validation), define how we are going to handle research to operations, and more
- There is concern over resources (HPC, human etc.), and…
- A related category of management, needs attention (e.g. setting priorities, refocusing EMC staff and in the community at large)
- Programmatically we need to work on more engagement (accessibility) and sustaining the community and aligning R&D and operational resources.
- Finally, communications and outreach are needed, and we need to clarify and refine our definition and vision of community

Workshop - Part 2: Strategic Implementation Plan (SIP) Coordination

The second part of the workshop focused on advancing the development of the Strategic Implementation Plan via coordination across the fourteen SIP/SIP Working Groups. These notes are compiled in a separate document.

Appendix A

Notes from Part 1 Plenary Session, as recorded by the rapporteurs.

I. NOAA Community Modeling Workshop

&

II. Strategic Implementation Plan (SIP) Coordination Workshop

NCWCP, College Park

Part 1: The 'Community Modeling Workshop' is an opportunity for the community to tell NOAA how they would like to engage with us. A brief remark on the format of this part of the workshop: the goal is to keep prepared remarks brief (provided for context and as fodder for discussion) and to allow for maximum dialogue.

Part 2: The Strategic Implementation Plan (SIP) Coordination Workshop seeks to delve into the topical issues associated with the SIP/Next Generation Global Prediction System (NGGPS) Working Groups.

Objectives:

NOAA seeks to engage with the community to form and shape the community, and to consider how to best execute shared infrastructure, support, management, and governance. Other objectives include identifying "best practices", discussing how community-based unified modeling system will actually work, and to evolve and coordinate between the SIP/NGGPS Working Groups (WGs).

Outcomes:

These workshops are parts of an open and transparent process of ongoing engagement, and will provide an important vehicle for NOAA to listen to the community regarding how it wants to engage. Outcomes include timely sharing of vital information (e.g. how people can plug in and contribute, timelines and status, etc.), generating a greater sense of community and mutual trust, and for NOAA, tapping the wisdom of the community. Output from the workshops will help NOAA to plan and establish the NOAA Modeling Community and to refine and improve its Strategic Implementation Plan for the FV3 model.

Noteworthy:

The federal government is trying not to replicate what other organizations are already doing. And, budget may force the community to come together to recognize/identify priorities to be funded, as no one can afford unified modeling efforts on parallel tracks.

It was noted that security measures may block goals. There needs to be an appropriate balance of community engagement and security measures. Federal policy mandates stringent security measures on computers and code, and restraints are likely to increase moving forward. Therefore, the community code cannot be on the operational computer, which will lead to a strong partnership with a firewall system, and/or running parallel systems.

NOAA's commitment to communications needs to be near top of list transforming organizational inertia.

Josh Hacker of NCAR/DTC noted that the cultural aspect is critical beyond 2019, however, where is the motivation to change the culture? Often funding is tied to products and it is hard to separate culture and determine how to use the community. Will the "build it and they will come" be a reasonable assumption toward cultural change? Currently, lack of documentation and code access limits researcher's use of current weather service global model. If there are no barriers, researchers will work with what is in operations. In addition to resources, ease of use, architectural flexibility, documentation, and level of expertise (such as in CESM) is an enormously easier path for use.

Rusty Benson noted that the Infrastructure working group recognizes that having one place for modelers to reach into the correct repository for the code is needed; it does not necessarily need to be one repository; it just needs to be linkable to community components. This means operations needs to figure out how to work with operational mirrors. Issue tracking, software, help, etc. are also needed.

Therefore, getting access to data is critical to success. Tom Hamill noted that if there is a fire wall system, it will never be a community system. Therefore, it is important to have a community design.

Tim Schneider observed that a NOAA action is to determine if there is leeway or leverage on how NOAA treats a development system vs. an operational system.

In some cases, need to make better use of available data in modeling. In land models, for example, the data is known and available; we just need to use it.

There also needs to be prioritized recommendations, with a prioritized research agenda (across component areas and scales). Working groups may have their own prioritized agendas but these may not take into account cross-cutting priorities (is dynamic vegetation more important than aerosols for subseasonal prediction, for example?). Prioritized research on scales – where do aerosols matter and where does interaction matter? NOAA may look for research partners who have done that, for example a demonstration of aerosol interactions with physics can be strong with wildfires and dust storms.

ECMWF sees Weeks 3-4 impact with use of predictive aerosols in their model. Indications are it is important. Measuring the importance can depend on your target metric. If measuring water balance, lateral water flow is important. If measuring carbon budget, dynamic vegetation is important. NOAA needs to evaluate sensitivities per metric, how sensitive a metric is to a specific process and design architecture around this, taking into consideration the interactions across boxes, and the need for flexibility and adaptability. New considerations need to be taken into account, such as different impacts different times of year (there is more complexity and interacting depending on species).

Some other recurring themes:

1. Definition of governance: Governance of code; governance of working groups
2. Communication: Between community and NOAA; between working groups
3. Training and documentation
4. Testing hierarchy for R2O
5. Resources are needed to accomplish this

Notes

Part 1: NOAA Community Modeling Workshop

Day 1: Tuesday, April 18, 2017

Introduction

Mike Farrar, Director EMC, provided opening remarks, welcome, and administrative details.

Ming Ji, Director NWS/OSTI, provided an NGGPS programmatic perspective including the goal of everyone working together on a combined unified model per the recommendation from UMAC. He also presented what community means to NOAA, and the community model concept and challenges. Two paradigms were discussed: a generalized distributed community approach and a more centralized, mission-driven paradigm. The latter was more favored. In the context of NOAA's mission, we must also be concerned with how to attract a broader research community; IT, infrastructure and governance issues; as well as balancing research collaborations and operations; and resources to support and sustain this enterprise.

Mike Farrar provided NCEP's Strategic Implementation Plan (SIP) overview describing the three-pronged planning approach to enable change: 1) a high-level, long-range Strategic Plan/Vision, 2) including an associated detailed Roadmap, and 3) a short-term (0-3 years) Strategic Implementation Plan. There is recognition to move quickly, within a short three-year time frame while fine tuning the long-term vision. SIP's three-year community plan will describe how NCEP is moving from the GFS spectral model to a community system working across temporal and spatial scales. A multi-layer approach to community roles will contribute to model code and capability, such as trusted super-users beta testing the next version of the model, and core development partners, such as JCSDA, working on development of the next model version. SIP working groups will expand on existing Next Generation Global Prediction System (NGGPS) working groups, and include additional functional groups such as Communication and Outreach, Governance, and Convective Allowing Models (CAMs). There will be an evolution/reorganization as EMC reorganizes from separate (meso/short term/long term) modeling to combined system within EMC.

Panel Discussion: "Goals and Needs of a Community Unified Modeling System"

Panelists: Tony Busalacchi (UCAR), Fred Carr (Academia/OU), Mary Glackin (IBM/The Weather Company), Dave McCarren (DoD and Navy), Pat Harr (NSF), Craig McLean (OAR), Christa Peters-Lidard (NASA), Louis Uccellini (NWS), Moderator: Mike Farrar

Comments, questions, and answers are described below, starting with initial comments from each panel member.

Louis Uccellini began the discussion noting that the National Weather Service (NWS) is serious about implementing the changes that need to take place. He noted the Weather Bill is close to being signed, but also noted it was currently an authorization not appropriation. This SIP effort is a critical part in building a weather ready nation and will address extreme weather events and forecast. This is important because 95% of the decision-making is on the 4-8-day forecast. This advance predictive forecast will be based on the ensemble forecast system, which needs to be a seamless suite ranging from seasonal to weather day events. There are challenges, such as protective code. Recognize need to

continue improving models and unify approach/engage research community further in this effort. This SIP meeting is the front-end of this approach.

Fred Carr commented that there is no shortage of university people who want to help and contribute to the community model system. If the Weather Bill is signed into appropriation, over 30% of funding will be for extramural activities. Current gaps which NOAA may be able to leverage with the academic community include addressing physics (such as the lack of understanding in the boundary level, microphysics, etc.), data assimilation (generate ensembles), verification (which needs to be an open process of experimental runs; students can be a fresh set of eyes to view), and HPC computing. Universities are interested in working together for a common good.

Craig McLean added that the higher goal is to deliver the best weather forecast to the American people. How to get there is to change what we've been doing for past five decades. There is a changing paradigm where we need to appreciate the value of external contributions (academia, etc.). OAR is changing too. There needs to be a fair and open understanding of the process to change the paradigm, where all parties contribute information in a controlled fashion. Since OAR's inception, 50% of researchers are university people employed through cooperative institutes. Change also includes having partners in industry working alongside those in the labs. Changing practices creates level of funding/resource challenges; therefore, trying to unify development will be key.

Mary Glackin stated there is diversity and differences in opinion regarding community modeling in the private sector. Is modelling a community enterprise with bars already set? Will definition of success be a truly open structure or one based on super user access? If a grad student can develop in the NWS arena and have access to the code and documentation, they will be able help too. The private sector can be an active participant, including leveraging HPC, and recognizes balancing of openness of code with operational needs. She was nervous about levels of users and possible associated restrictions on outer layers of onion (different proposed user levels)? She thought her community could help in areas such as physics and with leveraging next-generation high performance computing. She commended the openness that had gotten us to this point.

Tony Busalacchi noted the candid remarks regarding the need for cultural change, and is intrigued by the potential/possibility of the unified model. He questioned whether the conversation is centered around a unified system for the research community and another unified system for operations. When talking unified system, talking one national system or separate unified systems? Having two of the same model provides universities an opportunity to contribute to the operational model. Can NCAR create a unified model for research with complimentary plans? "NOAA led" NCAR can contribute to all areas. Again, process understanding will be key. Will there be different tiers for NOAA and others? Will each layer have the same leverage? Does NOAA have funding to do this? While true environmental prediction is needed, the community may have different goals. Earth system prediction is a growth area for NOAA, however sharing decisions and transparency, similar to NUOPC, balancing O2R and R2O can be successful, similar to the water model success.

Pat Harr noted there are two frontiers, the science of predictability and prediction. One can classify research goals on how they address each of these. NSF has long history of success in the "R". The "2O" part is left to the mission agencies or private industry. There is a need for culture and framework that facilitates desired characteristics. There was also mention of how parametrizations can lead to model and prediction errors [an area where research can help].

Dave McCarren remarked that DoD, Navy and AF have different roles and requirements. For example, the Navy provides predictions of a significant fraction of storms outside of the NWS domain. The National ESPC has completed foundational work that can be leveraged, including a funded partnership, and the NUOPC interoperability layer. HFIP has grown because of national collaboration of cyclone capability. Openness of code is an issue/barrier. Development of components may be slow to integration, and it may be costly to integrate some of those changes.

Christa Peters-Lidard stated that NASA has a small research budget, so there is emphasis on shared ownership, and NASA sees the benefit from this effort. Non-NOAA organizations and NASA specifically, have product data. ESMF is funded by NASA, with a shared vision, and uses the same software framework for weather and climate models. NASA has used FV3 for 10 years, and is interested in sustaining the FV3 core going forward. NASA coordinates with NOAA LIS, including to transition to LDAS v.3, there are connections across land, hydrology, etc. The biggest challenge is governance. There is a need for an automating testing process, and there is also a need for more rapid acceptance. Embrace all potential of community – it seems the modeling system is fragmented by different organizations and different missions… everyone will need to give something up to focus on community modeling effort. Challenges: governance, streamlined test and evaluation framework needed. Partners need to have access to repository in real-time and don't want gatekeeper that slows process.

Discussion continued:

Louis Uccellini observed there are two points:

1) There is diversity in modeling vs. all right and all wrong. The modeling can be more goal directed, with some convergence needed with a percentage of basic research.

2) NOAA's made tremendous progress, and, it takes time. NWS is running more community models than 15 years ago, and thousands of users are already tuned to those models. And, although there may be support for basic research in modeling, it is not supported in congress.

Fred Carr added that there are still too many models at OU, NCEP, NCAR, and elsewhere.

Craig McLean noted that there is overlapping, and using different layers of coordination to define and align the money spent can provide insight on what everyone is spending money ON, not just spending. Although defined on the operational side, what is research looking to get out of this? The research community will need to decide what to stay with and what give up.

Mary Glackin said it was unrealistic to think we can expand research dollars in current environment.

Conversation continued noting that the federal government is trying not to replicate what other organizations are already doing. Budgets may force the community to come together to recognize and identify priorities to be funded, as no one can afford multiple unified modeling efforts on parallel tracks.

It was noted that security measures may block goals. There needs to be an appropriate balance of community engagement and security measures. NOAA has high security arrangements on computers and code, and restraints are going to be more, not less, going forward. Therefore, the community code cannot be on the operational computer, which will lead to a strong partnership with a firewall system, and/or running parallel systems. For example, the Navy has real threats, therefore an accessible code on a repository, along with non-disclosure agreements, is more realistic.

Brian Gross, NOAA's HPC NCIO, added that the community approach to HPC software may be duality. NSF and DOE have exascale community compute capability, but the code will probably need to be rewritten. Perhaps there is a need for an IT testbed (software computing and engineering), like meteorological science, with a playground for software engineers. Jim Kinter noted there is a need for NOAA-led leadership to address HPC and community collaboration, and complimented NOAA leadership for working on this. Louis Uccellini thanked the acknowledgement and responded that although there are major advancements, it is not just the modeling aspect, it is funding, defining community, and collaboration. Mary Glackin noted that each partner needs to bring something to the table. Investments will be different layers and levels between government players, industry, and academia. Louis Uccellini noted the importance of the computer and posed the question of how you open up use on software that is developed/labeled as proprietary.

Moderated Panel Discussions with SIP/NGGPS Working Groups (WG)

The WG sessions include brief remarks from each WG representative followed by group discussion.

Moderated Discussion with the Governance WG (Ricky Rood, U. Mich.) and System Architecture WG (Jim Kinter, GMU) Moderator: Mike Farrar

Governance - Decision making, roles/responsibilities, advisory boards, org. alignment.

System Architecture - NEMS evolution for coupling approach, etc.

Ricky Rood spoke representing the governing working group. The word "governance" triggers warnings. There are different types of governance, and different opinions on what governance means, and what is to be governed. Questions regarding governance include: What governance behavior is needed for a program or end-to-end system, what roles do sponsors/working groups have? There is a difference between governance and management. NOAA will try not to control the community but to build wrappers around existing communities, fill gaps, and direct research. How does NCEP settle on operational requirements? And, how do they build trust and confidence in the process? How does a program manager for delivering a product have reasonable expectations/transparency and communication paths? What are the standards for community access and resources? When is there a transition of governance to management to provide a system to meet requirements?

As of now, there is no real understanding of the scope of governance needed, so it needs to be defined. Is governance needed for the FV3 community or the unified global system? Will the governance support the chain to connect R2O? The roles and responsibilities on how to make decisions to ensure implementation also affect governance.

Jim Kinter described the System Architecture working group, and acknowledged co-chair Cecelia DeLuca. The working group will provide recommendations to guide development and evolution of the community within the fundamental organizational framework. Documentation is critical to support a research community. Governance – who has control of the trunk/sandbox and how things move in and out - is critical, therefore governance coordination across teams is also critical. The goal of the system architecture is to balance creativity with requirements.

Mike Farrar asked what are the top issues that need to be addressed for both research/operations and users in between (in terms of governance and system architecture)? Jim Kinter responded saying not to conflate leadership and control – there can be leaders with no control or absolute control without a leader. The key is for NOAA to exert leadership and to define how partnerships will work, perhaps not controlling the users. One approach would be a governance structure that enables everyone to participate at the level they are capable of participating, lining up modeling stakeholders and capability, and with no one fighting over turf is the goal. NOAA's first order is to find what the research community wants to use. Currently, there may be a lack of trust in both directions, with researchers focusing on only publishing papers, and operations feeling like they know as much as researchers. The optimal solution is when a grad student can take the code "off the shelf and run with it" by eliminating barriers and allowing research to progress and then having a path back into operations.

Ricky Rood noted the need to set up communication strategies and linkages to break the current culture of too many models, and transform culture of how people are working together internally and within the community. NOAA's commitment to communications needs to be near top of list transforming organizational inertia.

Josh Hacker of NCAR/DTC noted that the cultural aspect is critical beyond 2019 (end of NGGPS) - where is the motivation to change the culture beyond 2019? Often funding is tied to products and it is hard to separate culture and determine how to use the community. Will the "build it and they will come" be a reasonable assumption toward cultural change? Currently, lack of documentation, shared code expertise, and code access limits researcher's use of current weather service global model. If there are no barriers, researchers will work with what is in operation.

Tim Schneider asked if the system architecture is building new code, or is it code rewriting/refactoring? Jim Kinter noted that almost everything on his diagram was extant code, such as ocean components, waves, etc. Strawman plans are in place, however, there are documentation and interoperability issues. Additional questions include: What physics will be used? Is NEMS the right way to go forward? What tests will be used for an evidence-based decision? Sam Trahan is working on workflow and thinking about scripts level. There is a lot of code that is fundamentally sound; however there is also a lack of documentation and lot of repetitive code which can be fixed with an automated refactoring tool. There is a need for better training in writing code; perhaps using pythons scripting language to decrease the amount of code.

Cecelia DeLuca noted that coding is a focal point represented in coupling when modelers make sure the model can go into NEMS, and they can replicate strategies. Currently, gap algorithms are already underway to take care of some repetition.

Sam Trahan commented that the research culture needs to shift to contributions to an operational system as opposed to contributions to a paper. NOAA draws threads through the organization rather than outside the organization. A governance management team to address architecture and governance prioritization may be the first step. Getting planning and design in place, having a governance model, working groups, advisory panels, etc. will be beneficial to the progress.

Ricky Rood commented that one fallacy is that there will always be tension, especially with competition of ideas. People are attached to their ideas so it becomes personal. Communication as to why a decision was made will reduce and/or eliminate the tension: 1) Have a validation plan in the beginning

with merit based objectives and decisions to depersonalize the decision. 2) Have a transparent process, where those who don't win at Phase 1, know how they can participate and/or will succeed as they move along.

One of UMAC's recommendations for NOAA was an evidence-based decision process. There is broad agreement on this way of proceeding, but how does one make a decision? What is the metric of success? What does an evidence-based decision look like? This is the opportunity to engage on how to decide on the metrics and how they will be interpreted. However, how does NOAA develop a process for decision making where everyone agrees on terms of engagement?

Moderated Discussion with Infrastructure WG (Rusty Benson, GFDL) and Global Modeling Test Bed/GMTB (Bill Kuo, DTC) Moderator: Whit Anderson

Infrastructure - Standards/documentation; CM; code repository; testing; etc.

GMTB - Its role, how it will work, Common Community Physics Package (CCPP), etc.

Rusty Benson noted that the Infrastructure working group recognizes that to have one place for modelers to reach into the correct repository for the code is needed; it does not necessarily need to be one repository; it just needs to be linkable to community components. This means operations needs to figure out how to work with operational mirrors. Issue tracking, software, help, etc. are also needed. NGGPS Infrastructure working group is supplying some data but further direction on NOAA code management and governance is needed. Questions to be answered include: How to quantify rules of engagement for the different layers? Are the rules the same for every entity? There is also a need to manage with standards. If a bug fix or research trunks are getting ahead, how does NOAA bring back, merge, etc. to keep in line with slower moving operations?

Whit Anderson asked what the biggest threat is (work flow, data availability, etc.)? Bill Lapenta commented that it is not always the better or worse model; the user picks the easier supported model (help desk, documentation, etc.). Rusty Benson noted the work flow system has to be simple and easy to use. Data must be simple and easy to use as well. What if data is proprietary?

SJ Lin commented that the goal needs to be determined. What is being created: the most efficient system, the cheapest system with largest market share, or the system with the best prediction?

Bill Lapenta answered that NCEP's goal is to be second to none using the best science. The research community has a different goal, publishing a paper is a goal. Mike Farrar noted that it is not an either/or situation. Short term, NCEP is trying to make the best model with timelines to meet; this is a legitimate concern today. Long term, the next generation of scientists has an advantage for using our modeling system, and can continue to make better modeling system in the future. Must keep long-term goal in mind (making strides in improving the next generation modeling system) – beyond short-term task of delivering best system. Rusty Benson noted this is not a conflict of interest. Some layer of testbed assistance is needed to help to distill down hundreds of research ideas to be brought forth as needed, so EMC is brought the most promising advances.

Sam Trahan added that there are two different use cases (operations and research). There is R2O (relatively fast process) and there is the case of research papers (requiring longer-term support). He mentioned operations and R2O for HFIP to bring code into NEMS, seems to work. Researchers work

with EMC on HFIP and three months later EMC get results from research community. But researchers don't have the HPC access, therefore, they work on their HPC, in the meantime EMC's master/trunk may have changed. NOAA needs to consider where is the trunk? Will there be a frozen version to do papers? What defines a release? But all entities are doing their development, which is their goal. They decide. Rusty indicated that the Infrastructure WG hadn't reached the point of discussing release versions or the process.

Steve Smith, Director of NOAA virtual laboratory, noted that VLAB supports upwards of 5,000 users on a daily basis, from small to large, like AWIPS software development. There is a help desk, social collaboration, and work flows, and Git repository in NCEPs operations. He offered, and was welcomed, to sit on the Infrastructure working group.

Bill Kuo presented information on the Global Modeling Test Bed (GMTB). He noted that to make the model a community model, portability will be important. There is broad support for more than just the operational style of environment. Another question is how do internal libraries used at NCEP get documented and pushed out if needed?

Moderated Discussion with the Dynamics/Nesting WG (Vijay Tallapragada), Model Physics (Jim Doyle, NRL), Convective Allowing Model (CAM) WG (Stan Benjamin, ESRL) Moderator: Whit Anderson

Dynamics/Nesting - FV3 transition on global weather/S2S/climate; hurricane nests

Physics - Stochastic, scale-aware physics; relationship to the CCPP, etc.

CAM - Intermediate steps to CAM ensembles, Warn on Forecast; test/evaluation w/community

Vijay Tallapragada talked about the Dynamics and Nesting working group, and noted what everyone knew – the dynamics have been selected. The philosophy is to minimize changes to the dycore, and maintain user potential. The dycore is designed to be able to respond to changes and requires collaboration to work through what new items are needed.

Jim Doyle described the Physics working group and commented that the model needs to perform well across time scales. Testing in a DA cycling mode is needed for transition to operations because the grey zone issue will be problem and needs to be addressed, especially with variable resolution on the sphere.

Stan Benjamin discussed the Meso/Convection Allowing Models (CAM) and noted the importance of verification and validation. The model may have a convective scale specific aspect, but it would have to apply to all seasons and atmospheric coupled phenomenon.

Arun Chawla questioned the metric evaluation. Metric score cards will be complicated going across time scales, and there may be a specific set of metrics for EMC, vs. what physics gets into CCPP, and what is needed in the community. Vijay Tallapragada noted that EMC has a process and the NGGPS Verification and Validation working group led by Tara Jensen and Bonny Strong have collected metrics. Stan Benjamin noted that there are different layers of model applications, from short range hourly to 1 day, global and seasonal time scale; there could be three score cards.

Hendrik Tolman commented that EMC does not want to create verification and validation in pieces. The Verification and Validation WG is looking into what needs to go into the packages for each group and will decide on the final metrics - more to come in their presentation.

Ricky Rood asked what the committees need in governance? What are the dependencies, software, training documentation, how to manage the decision making process.

Jim Doyle observed there is quite a bit of discussion on CCPP/GMTB and how the community will participate. How to have community involvement is a question.

Sam Trahan asked how does a researcher check whether they have a good code? A researcher can run 1D physics model but will get different results with real atmospheric data, and the community researcher cannot do full size tests. How/will NOAA track who is getting resources? The CAM WG is working on it, but for those with limited resource and limited access, GMTB would allow better access.

Jim Doyle commented that coming from research community, there should be an effective way of engaging and facilitating with a hierarchal test suite. There is a set of tools and metrics (1D, then moving to 3D with some test cases). If it moves along in terms of its promise, it will move up to engage with EMC, bringing it closer to engaging in the operational environment.

Steve Goldhaber from NCAR asked that once the physics is designed, what tools will available to use in the community? Jim noted that there is an effort to develop an interoperable physics driver to be a flexible piece of software, to drop in multiple suites and other dynamic cores. The Navy, NCAR, others have the ability to share common physics suite/or a couple of suites. The focus is to engage and leverage the community to develop something similar within this community allowing this application to work across models and other dycores.

Ricky Rood also commented that the community is testing short simulations for convecting allowing models. For more robust representation for coupled global simulation (which is for a different test strategy), governance fits in. Both NOAA and the community need to be thinking ahead to subseasonal and keep looking ahead, anticipating out certain time scales. To have evaluations across time scales from convective time scale, to medium range seasonal, subseasonal, and long term, interactions among all groups is needed, including government.

Mariana Vertenstein noted the practicality of setting up collaborations and noted the challenge of having different physics on grid/concurrency in place – what if you want neighboring columns to interact? Do the interactive columns outweigh what is currently being looked at? Discussion continued noting that NOAA and the community need to learn about other ways to approach these issues. Will there be hierarchy within the driver to add layers and address layers such as regridding? What if physics runs on a different grid? Will there be concurrency? NOAA needs a strategy to build upon. Physics development needs to keep in mind the users and applications, and the need for training the community.

Bill Lapenta commented that NOAA is trying to be visionary and look 5-10 years down the road. What we are doing today, this process to supporting the evaluation groups, is to get a better system in the end.

Moderated Discussion with Data Assimilation WG (Tom Auligne, JCSDA), Ensembles WG (Tom Hamill, ESRL/PSD), Verification & Validation (V&V) WG (Tara Jensen, NCAR/RAL), and Post-Processing WG (Matt Peroutka, NWS/MDL) Moderator: Whit Anderson

DA - FV3 integration between NOAA, NASA; Joint Effort for DA Integration (JEDI); coupled DA

Ensembles **-** Strategy across scales; model uncertainty

V&V- V&V of ops forecasts vs. R&D testing/evaluation; unified/standard tools and data formats

Post-Processing - Community infrastructure (e.g., Unified Post Processor/UPP); standard formats/tools

Tom Auligné started the data assimilation discussion by editing the title from Unified Modeling System to Unified Prediction System because this is more than just modeling: there is tracking, code review, systematic testing. Open access and transparency in the process is a key component to success. Currently HPC access is too restrictive; there also needs to be a way to incentivize the community on how to get things back into the system.

There needs to be an agreement on the environment to get a community system developed, such as: "Put your books in the library vs. living room book club." Industry has already done this. It is not as much a technical problem as it is a cultural shift. Recommended a central repository with a control version and community version, to create a system that balances accessibility with control.

Tom Hamill discussed the Ensemble WG and suggested exercises in community systems, focusing on 1 day to 1 month, including graduate students. This near-term focus, such as on the Land Surface, could be farmed out to NGGPS DA teams. Not everything is ready yet, and there are issues on how to use blocks of CPU time. The current work already uses a lot of disc space, and this will need to be addressed in the future.

Tara Jensen mentioned the large membership of the Verification and Validation WG and the WG goal to identify and respond to gaps. MET is currently on GitHub, and the model evaluation tools are used for coupled systems, space weather prediction, land, waves, and atmosphere. Currently the code is C++ and wrapping in Python. The goal of V&V WG is to provide friendly interface and extensive outreach such as tutorials, in-person and online. The restrictive data issues need to be addressed, and the group needs to figure out how to support MET to the international community, and for those without access to C++. Similar to comments from other working groups, more R2O and O2R is needed. The system is capable of running with or without a work flow manager, and run on multiple processors. It is a process oriented matrix with different verification for different scales. And, although some want a standalone verification and validation, the bottom line is a goal where forecasters gain confidence in the products, which will help managers make decisions.

Matt Peroutka provided an update on the Post-Processing WG by first observing that data format has not been brought up yet. Data format is important for processing on fields. Several partners use cloud computing, and WG needs to spend time evaluating different options. As already discussed, access to computing is important, and the process to access data, as well as the portability of code and the ability to work on multiple machines, is important. Once outside the critical barrier, the data is extremely valuable. Perhaps an option is to take model output and push it to the Amazon cloud for access. DTC access to WCOSS is currently limited to five or six users. Therefore, getting access to data is critical to

success. Tom Hamill noted that if there is a fire wall system, it will never be a community system; therefore it is important to have a community design.

Tim Schneider observed that a NOAA action is to determine if there leeway or leverage on how NOAA treats a development system vs. an operational system.

Discussion continued and comments included mention of the current effort excluding foreign nationals from accessing HPCs, and the request to keep R&D machines open. It was noted that NOAA is allowed to use GitHub as long as those using it are corresponding safe inside fire wall.

Accessing the weather data is a challenge. It is hard to duplicate live data on any machine, and it is hard to get live data to an R&D machine. If NOAA is rethinking computer, raw vs. historic data, this needs to be re-evaluated because currently moving data on and off machines is limited by disc space. Sam Trahan added that the Joint Typhoon Warning Center (JTWC) has multiple, redundant paths, and that cloud computing is feasible.

Discussion continued on physics verification and the Common Community Physics Package (CCPP). It was clear that the ability to swap physics suites may not be simple. For example, if microphysics schemes will change model drive radiance, which changes post processing, additional verification and validation will be needed. Along with physics, comes systematic error, therefore analyses will have a physics error.

Christa Peters-Lidard commented on benchmarking. NASA benchmark is upper bound in terms of skill against some data (what is possible); will there be some sort of persistence skill metric that you need to get above (lower bound)? Tara Jenson commented that they are thinking of providing benchmarks but don't know how just yet.

Chidong Zhang (PMEL) asked about process oriented metrics and how to package data to make sure it is useful (for example radar data, as well as ties to IODA) for the observations database and how to address it. It was suggested he reach out to Tara directly. A workshop to engage the community on V&V was suggested.

Moderated Discussion with the Marine Models WG (Bob Hallberg), Land Surface Models (LSMs) + Hydrology WG (Christa Peters-Lidard, NASA/Goddard), and Aerosols/Atmospheric Composition WG (Ivanka Stajner, NWS/OSTI) Moderator: Whit Anderson

Marine - Current marine models, NOS coastal/bay models, future ecological models

Land/Hydro - LSMs within coupled system, connections with National Water Model, etc.

Aerosols/Atm. Comp. - Downstream vs. in-line applications; science challenges, etc.

Bob Hallberg described the goals of the Marine Models WG including bridging and unifying diverse communities. The group is working towards combining both research and operational capabilities, and adopting a common infrastructure to be sustained and supported for the long term. Governance of shared code bases is suggested, both in GitHub and in the open, with caveats that if you do no harm, code will be shared. Also, governance of specific configuration and specific runs can be used by user chosen through a vetting process.

Christa Peters-Lidard presented information on the Land Surface Models and Hydrology WG and what the land models require for development and improvement. Examples include the need for dynamic vegetation estimate to change in response to drought/fire, and basic physics is missing. The human influence – urbanization, irrigation, reservoir water diversions, fire, biogenetic – all influence on the model prediction. There is a need for different representation of land surfaces across the scales for energy and water across spatial and time scales. Scale-aware land surface models are needed. Picking the most dominant vegetation/water type will not work when looking for skill scores. The data is known and available; we just need to use it. Additional issues relevant to land include sensible weather near the surface, and metrics. ECMF papers already note this. There are cross-NOAA issues: FV3 has their own land model, GLERL has their own model, HRRR uses REC LSM. A robust evaluation is needed as the models move into an information framework, including the need for benchmarking, as we move into a new modeling approach.

Ivanka Stajner provided an update on the Aerosols and Atmospheric Composition WG and started off by describing the four aerosol models currently be used by EMC. The WG would like a unified production system. But, in order to have one system, there is a need for a unified interface to physics, dynamics and DA. Aerosols and chemistry models are different and need to be combined to make the estimate of the boundary layer easier.

Discussion continued around the challenge of the architecture for coupling, the explicit boundary layer, and column to column interaction. All have three things in common – all are related to fast and slow processes.

1) Delta T, cadence, and how all synchronized and coupled manner.

2) all have water

3) all research questions

There needs to be research on the processes – do these things matter? Is vegetation relevant to subseasonal seasonal now casting? It was noted that there are architecture implications. It was also pointed out that we have to have the capability to ask questions in the research community, which means code has to be there and available.

There also needs to be a prioritized recommendation, with a prioritized research agenda. Working groups may have a prioritized agenda, but is dynamic vegetation more important than aerosols for subseasonal prediction? Ivanka Stajner said there wasn't a weather service response to this but research partners have shown strong interactions between aerosols and microphysics in some situations. In general, the Aerosols and Atmospheric Composition WG is using community findings for justification of impact.

Also, should there be prioritized research questions. Prioritized research on scales – where do aerosols matter and where does interaction matter? NOAA may look for research partners who have done that, for example a demonstration of aerosol interactions with physics can be strong with wildfires and dust storms.

ECMWF sees Weeks 3-4 impact with use of predictive aerosols in their model. Indications are it is important. Measuring the importance can depend on your target metric. If measuring water balance,

lateral water flow is important. If measuring carbon budget, dynamic vegetation is important. NOAA needs to evaluate sensitivities per metric (how sensitive a metric is to a specific process and design architecture around this) taking into consideration the interactions across boxes and the need for flexibility and adaptability. New considerations need to be taken into account such as different impacts different times of year. There is more complexity and interacting depending on species.

Mariana Vertenstein noted that if you wanted to bring different parameterizations into MOM, you should be able to do so. However, if you put something in and don't test it, it's broken. There needs to be a way to scientifically validate what is put in and tested at different resolutions, etc. There needs to be feedback what kind of testing and test cases are needed.

Discussion continued regarding data contribution, reproducibility, testing. There was a comment that nothing is ever deleted in GitHub, and past versions can be recovered. A question was asked if there is a way to date or track the past versions, and/or for example, replace at timescales when ocean starts to matter? Other examples include: there are four different models in operation for aerosols so the system should have options, in certain configuration, to allow for more complexity; land models converging will have a single model with physics options and choose metrics, there can be a choice between dynamic vegetation for three day for cast or for seasonal.

Use as foundation to build community earth system model across time scales, unified.

Day 2: Wednesday, April 19, 2017

Motivations & Instructions for Breakout session, Moderator: Tim Schneider

Key questions for breakout groups to address:

Based on what you've seen thus far (what looks right and what does not?), please identify:

1. **Best practices**: What are the major things that we're getting right?
2. **Gaps**: What are the major things that we're missing, or that we're heading down the wrong tracks on?

Breakout group discussion/feedback:

Do we need to build incentives? From research perspective, many researchers want to have something that is documented/has impact. This drives/satisfies some researchers. Can help to have low bar for entry and path to delivery (including easy access) to get increased participation by community. In Europe, your career is valued on both what you published and "value" of what you published.

Are there any successful models worldwide for distributed community systems? Can we build paradigm for this? Matt Peroutka commented that federal government model for leading this (weather?) enterprise/model is changing (more privatized). Reason why businesses contribute intellectual property: they are reducing their maintenance costs. Jim Doyle said that at end of day, need core group of developers committed full time to the system but seems one group won't be able to accomplish core/component work on their own (complex problem). Private companies can't develop complete weather system on their own, so advantage exists in collaborating as a win-win.

Sam Trahan – NOAA is generating more data than they can store. Business is one example of where to look to for technological solutions. IBM may have some useful suggestion, as they already support NOAA (funded). Commercial enterprise software companies have been working these issues for years (Hacker). (Hendrik) Various groups in NOAA are working these issues (such as with AWIPS) but there is a lot of inertia in the system. (Peroutka)Which working group most cares about big data issue? Tim said thought infrastructure group.

Stan Benjamin – Hierarchical repository issue – could we have/use CESM repository. Mike Farrar said this was considered.

Has anyone given anything of value to the federal government? Dave McCarren commented there is concern over legal restrictions in place for the government that may hinder hand-over.

Mike Farrar comments: Recognized based on structure of WG, we have already recognized the existing gaps. Have the right groups established to address issues. Heard consensus on gaps, but also heard varied priorities on infrastructure issues. Probably need to focus more on issues like access/repositories. Also, not clear that vision has been clearly articulated (what is goal in terms of building system, how is community defined…….). Need to do a better job of communicating.

Proposal for FV3GFS code (V0) release shared – plan to share via GitHub by mid-May.

Other items listed in Break-out Group feedback:

Groups seem to have similar good and bad points/gaps.

Community could be so big that it may be difficult to manage – how far does community extend? Users?

Interaction between groups in repository – collaboration can accelerate products into operations

How to measure success? Metrics?

Cultural change is starting.

How to put recommendations into actions – what's next?

What to do with data?? Sharing i/o cloud computing?

Procedure for R2O? O2R? Transitions

Motivators/incentive for researchers

Infuse researchers

Who making decisions?

Need governance defined.

How to build incentive for community participation? If we build a good tool then that is incentive to use. Mission requirement vs. resource requirement. – impact on operational, publications, both sides of the coin.

Crowd sourced – operating system

Post-processing – if software and datasets are useful and readily accessible, build it and they would come.

Love to set grad students on it If code from NOAA, it would be great. Business' contribution. Maintenance cost still coming up……

Need a core group of developers committed to system.

Grey area to incl. community in ways never done before. The complicated system / components are huge.

How to interface? Other core and super users developing system because of complexity of problem.

Private sector taking advantage of center – and pushing the collaboration to take advantage of.

Working alongside, HPC, operational forecasting together

NOAA mass store and data not sufficient for NOAA's purpose. Generating more data than store. Reduce what is stored.

Some problems can be solved by technological solutions while others governance, etc. are long term government

Key is to have a clear pathway from R2O with metrics and easy to use testing suites. Have a vendor to use computation testing.

Stan Benjamin – Have CESM /DTC as a way to have repository with governance - can we use it as an existing framework?

Josh Hacker group recognized what gaps are, WG already working on. Gaps are not created equal, some need to work on sooner. Infrastructure – repository, code availability, test cases,

Barriers and support systems to enable to do science. Need to be tackled sooner.

Not clear and ideas are well articulated and understood by our partners. Fundamental what mean by community, what goal to build system, what want out of it? Competing with other modeling system for researcher, building a system for researchers as well?

New work for communications group to articulate the vision.

Private sectors, academia, intend to do this mid-May will release internally within NOAA virtual lab with Git repository.

Mirror and everyone will be able to get to it. Not sure of support upfront. Version 0 of code.

ADJOURN – Part 1 Complete

Part 2 SIP/NGGPS Working Group coordination breakouts Day 2: Wednesday, April 19, 2017

Each working group breakout session will produce summary.

See documents on Workshop Webpage:

https://ral.ucar.edu/events/2017/community-modeling-workshop

Appendix B – Breakout Session Summaries

Breakout summaries as provided by each group.

GROUP A

BREAKOUT TOPIC 1: In terms of overall direction of moving towards a community approach to modeling, what appears to be on the right track?

1. New openness and leadership for engaging the community.
    a. Selection of a dycore for the unified model from outside NCEP
    b. General strategy for community of developers (e.g., tiered code repository, etc.)
    c. Consideration of full Earth System modeling (atm, ocn, land, ice, aerosol, BGC)
2. Use of standards and modularity for the unified framework (e.g. ESMF/NUOPC, IDP etc.) with cross agency leveraged efforts
3. Evidence-based decision making in all aspects of planning and development

BREAKOUT TOPIC 2: What are we missing? What gaps remain?

1. Resource challenges
    a) Interagency collaboration and resources.
    b) Consolidation of modeling systems.
    c) Difficulties in providing access to visitors, collaborators.
2. Communications & Outreach WG issues:
    a) How is community defined? Clarify the motivation and scope of this effort?
    b) How do other "community" Earth System models differ from NOAA's?
3. What is the success criterion for NOAA's unified modeling system?
    a) Definition of unified metrics, V & V tools
    b) Definition of success at the end of i) 3 years; ii) 5 years ; iii) 10 years.

GROUP B

BREAKOUT TOPIC 1: In terms of overall direction of moving towards a community approach to modeling, what appears to be on the right track?

1. Community has engaged, EMC devoted considerable resources
2. Paradigm of JEDI for repository structure
3. Discussion of governance, breadth of topic

BREAKOUT TOPIC 2: What are we missing? What gaps remain?

- High level framework for project

- Overall priorities and goals, metrics, benchmarks,
- Guidelines for entering repository, release of code, documentation, training, testing and evaluation
- EMC becomes customer and contributor, role change
- Connection to observations community

GROUP C

BREAKOUT TOPIC 1: In terms of overall direction of moving towards a community approach to modeling, what appears to be on the right track?

1. Encouraging evidence of a cultural change as demonstrated by holding the meeting at all and engaging community.

2. Many of the right priorities are being highlighted as evidenced by the choices of working groups--both topic and membership. Challenges being recognized.

3. First steps to building community.

4. First steps to building relationships within a community.

BREAKOUT TOPIC 2: What are we missing? What gaps remain?

1. How working group recommendations and SIP will be put into action is not clear.

2. Apparent disconnect between research priorities and funding, both internal and extramural.

3. Visibility into process for how innovations make it to operations is lacking.

4. Explicit recognition of the importance of HPC (optimization, next-gen, etc).

GROUP D

BREAKOUT TOPIC 1: In terms of overall direction of moving towards a community approach to modeling, what appears to be on the right track?

- Choosing a 3-year strategy before beginning the SIP process correct
- Attempted to gather community
- Identified time for X-C WG interactions

BREAKOUT TOPIC 2: What are we missing? What gaps remain?

- Development expertise limited to a few scientists. How to expand the knowledge base ?
- Dynamic core, at least potentially others
- Missing a clear, coherent definition of community development, role and scope.
- Who owns the central repository or decision making process (governance) ?
- Central repository shall be accessible outside firewalls, universal visibility and limited commit privileges (e.g.: GSI approach)
- Ease of access for community (roadblocks), inputs, outputs, computing

- Systematic validation of coding standards, compute and forecast performance; regression testing

GROUP E

BREAKOUT TOPIC 1: In terms of overall direction of moving towards a community approach to modeling, what appears to be on the right track?

1. Cooperation and communication among many different groups and agencies is increasing.

2. It is encouraging to see EMC developing a strategic implementation plan, and including the community in the development of that plan.

BREAKOUT TOPIC 2: What are we missing? What gaps remain?

1. We need a procedure for the transition of research to operations. Anything RL5 and above needs to have a transition path outlined with input from EMC. There should be set procedures (decided by EMC and SMEs) for the testbeds for bringing anything into operations.

2. Incentives need to be provided to other groups/agencies to transition model development efforts toward FV3.

3. Annual science meetings and subject-specific working group meetings between EMC and the community are needed.

GROUP F

BREAKOUT TOPIC 1: In terms of overall direction of moving towards a community approach to modeling, what appears to be on the right track?

1. Initial community Interactions are good, e.g. Mike Farrar's community diagram.

2. Building operational configurations from community-based codes.

3. Good leadership, transparent decision making processes

4. NGGPS is successful so far. Good to follow. Bridge to beyond 3 years. HFIP may be a good example to follow.

BREAKOUT TOPIC 2: What are we missing? What gaps remain?

1. Articulating vision for community modeling

2. Access (security and amount) to resources constraints our approach greatly.

3. Decision making and governance

GROUP G

BREAKOUT TOPIC 1: In terms of overall direction of moving towards a community approach to modeling, what appears to be on the right track?

1. Culture shift earth system enterprise toward community engagement

2. Long term planning and vision, establishment of SIP working groups

3. Positive response from community

BREAKOUT TOPIC 2: What are we missing? What gaps remain?

1. Defined requirements and goals through governance (authority?)

2. Identification/allocation of resource needs (hpc, funding, staff, helpdesk, etc.)

3. User friendly, developer friendly, and attractive.

4. Sufficient incentives for development and support toward completing transitioning to operations.

5. Science steering committee (evidence based, identify priorities)

GROUP H

BREAKOUT TOPIC 1: In terms of overall direction of moving towards a community approach to modeling, what appears to be on the right track?

1. Transparency is essential and workshop is a good start.

2. Increased use of cross- and multi-discipline metrics to drive decision making.

3. Move toward modernization of software infrastructure and standards (refactoring, modularity, repositories).

BREAKOUT TOPIC 2: What are we missing? What gaps remain?

1. Establish process for continual input and engagement.

2. Governance needs to be detailed and well-defined for prioritization, resources, schedule, and model performance. Governance seems to have different working definition depending on perspective.

3. User support: Who is in charge and how will it be resourced?