Big Weather

A workshop on overcoming barriers to distributed production, storage, and analysis of multi-model ensemble forecasts in support of weather prediction research and education in universities

National Center for Atmospheric Research, 9-10 January 2014

Conveners:

Joshua Hacker, National Center for Atmospheric Research, Research Applications Laboratory
Carlos Maltzahn, University of California, Santa Cruz, Department of Computer Science, Jack Baskin School of Engineering
Gretchen Mullendore, University of North Dakota, Department of Atmospheric Sciences
Russ Schumacher, Colorado State University, Department of Atmospheric Science

Funded by the National Science Foundation grant number AGS-0856145

Distribution list:

Workshop attendees/universities (see Appendix)
Steve Nelson, NSF
Michael Morgan, NSF
Eva Zanzerkia, NSF
Don Middleton, NCAR-CISL
Brant Foote, NCAR-RAL

1. Executive summary

An NSF-funded workshop held during 9-10 January, 2014, at NCAR convened to address goals at the intersection of university-based numerical weather prediction (NWP) and big data. Participants came with both research and education interests, and both NWP and computer science backgrounds. Of 25 total participants, 10 are currently M.S. or Ph.D.-level students (see Appendix).

Specific objectives of the workshop were to:

- Identify barriers limiting exchange and analysis of NWP and climate prediction output produced outside of operational centers and large laboratories
- Identify the NWP-relevant scientific progress that may be realized by removing data barriers
- Identify promising technologies emerging from computer science disciplines that could overcome those barriers
- Identify the extensibility of those emerging technologies, and expose the potential lifetime given the expected growth in local computational capacity
- Identify educational opportunities possible under a new paradigm of populating and sharing real-time NWP data

Those goals suggested the following outcomes expected from the workshop:

- A cross-disciplinary understanding of the issues slowing data-driven research in NWP, outside of operational centers
- A set of user science-driven requirements defining a community strategy for big data in short-range NWP
- A set of possibly extensible approaches to overcoming the data barriers
- A white paper addressed to NSF (post-workshop); strategic document to respond to future funding opportunities

To address physical science, computational, and educational goals, the workshop was split into three half-day sessions. Each session was split roughly between brief presentations from attendees and open discussion around the session topic.

We note that these goals and outcomes fall under a longer-term, overarching goal of accelerating scientific and educational progress in numerical weather prediction, and specifically ensemble forecasting. The national weather prediction enterprise stands to gain if we advance as a community. Operational centers and several large university-based research centers around the country have significant computational and storage capability and capacity. Those centers have less need for data sharing and analytics across the community. Conversely, small to mid-size research groups and primarily undergraduate institutions (PUIs) stand to gain much more from community-based efforts. Infrastructure to share and analyze data produced in larger programs brings intellectual capability to scientific problems, and complements many ongoing efforts, which may be NSF-funded. Any effort following this workshop is not intended to replace, or compete with, the larger programs already making significant scientific progress. Still, as we go forward, we are looking for ways that all members of the community can benefit.

The workshop elucidated nuclei around which broad-based effort can be organized. These nuclei are pieces of technology, and are intended to be easily shareable, have the ability to grow and improve over time, and ensure support (buy-in) from researchers and students. Workshop discussions clarified what nuclei would best serve our community:

1. A large ensemble consisting of contributions from multiple universities, which form a functional group representing the broader ensemble forecasting community.
2. A common storage, linking, and cataloging methodology among the participants, which allows research on data, product, and workflow provenance.
3. A virtual machine for research and education in short-range ensemble forecasting, allowing rapid spin-up, access to sharing technologies, and communication of best practices.

These nuclei provide focus for forthcoming scientific, technical, and educational efforts. In the next three sections, we discuss individual topics for research, development, and education, and point out how they are served by these nuclei.

This white paper it intended to accomplish three things: (1) summarize the key discussion points from the workshop; (2) suggest essential elements of any effort aimed at improving data sharing and analytic capability in NWP research and education; (3) offers some details about how we, as a community, might take concrete steps in that direction.

2. Science drivers for short-range ensemble forecasting

Here we enumerate science drivers that arose during six participants' presentations, followed by a 1.5 hour discussion, and list a set of requirements for a data-sharing system that could address those drivers. The list of science drivers is potentially long, but we present three that appear to cross-cut the interests of the workshop participants.

Note that none of the science drivers are unique to limited-area ensembles. Global ensembles for predicting small scales are still not practical. Results from research on limited-area ensembles will remain relevant as computational power grows to allow meso-gamma scale (2-20 km) forecasts.

Convection-allowing ensembles

Deterministic mesoscale predictions have been operational at convection-allowing grid scales (100 m – 4 km) for some time now. For example, see the High-Resolution Rapid Refresh (HRRR) run at 3 km grid spacing over the contiguous United States (http://ruc.noaa.gov/hrrr/). Reasonable probabilistic treatment of these scales with ensembles is inevitable in the next 5-10 years, but many characteristics require research to ensure they are understood and effectively deployed.

While the body of evidence suggests that at tens of km, tens to 100 ensemble members saturate probabilistic skill scores, it is unknown how many ensemble members are needed to saturate probabilistic skill scores for convection-allowing scales. Large ensembles are needed to answer these basic questions.

The ability of multi-physics ensembles to handle convection-allowing scales is also unknown. Sensitivity of forecasts to microphysics and boundary-layer parameterizations is well documented. Different simulated physical structures in multiple models thwart simply statistical interpretation, because it cannot be assumed that each ensemble member is drawn from the same forecast distribution. Empirical evidence suggests that this issue does not prevent useful ensemble predictions, but whether this holds at convection-allowing scales is unknown. It is reasonable to hypothesize that different structures at these scales, which lead to different modes of convective storm formation, cannot be easily combined into a useful probabilistic forecast from an ensemble. Research is needed to find good ways to form ensembles at these scales.

***Ensemble sensitivitie*s**

Ensemble sensitivities have emerged to offer a simpler, if more computationally demanding, alternative to adjoint sensitivities in the last several years. Sensitivity methods enable physical diagnostic studies, observation network design, and observation targeting studies. In certain contexts, they can also be used to evaluate sensitivities to parameters in the models.

Optimal ensemble sensitivities are still a challenge, particularly at fine scales. Finite ensembles lead to sampling error, which leads to over-estimation of covariances. When possibly important covariances are small, the error in the covariance estimates can swamp the covariances themselves, hiding the result we seek. The obvious, but practically difficult, solution is to use very large ensembles. This solution is possible at only a few institutions, and in limited scope.

Ensemble sensitivities in a multi-physics or multi-model context are unexplored. Because each member of such an ensemble is drawn from a different distribution, interpretation may or may not be possible. At least a different interpretation may be needed. Data sets to support the necessary theoretical development are expensive to produce, and time-consuming to manage. Most single investigators, or small teams with local computing, cannot easily address this challenge.

Sensitivities to physics

A proliferation in physical closure and forcing schemes in mesoscale models such as the WRF allows a wide range of customization options, but can also drive the need for exhaustive experimentation for a new application. Besides the ensemble sensitivities discussed above, other reasons to gain access to a large multi-physics ensemble include: finding a physics suite that best reproduces observations for a case study or a number of similar events, a desire to find simulations that are not sensitive to physics choices for confidence in details of a simulation, and a desire for a range of solutions that indicate model uncertainty or inadequacy.

A common graduate student thesis project may involve reproduction of an individual event or a set of events. To ensure fidelity to the atmosphere, an initial goal is often to find simulations that verify well against observations. Because details of many simulated weather phenomena are sensitive to choices of physics, tens to hundreds of simulations may be needed to find the most promising physics suites. The process can take weeks or months, detracting from the real research questions. Easy access to an existing ensemble, with a wide range of physics combinations, and appropriate provenance, would allow an investigator to skip the actual model execution, and begin immediately on analysis to find an acceptable model configuration.

A classic use of NWP model simulations is as a proxy for the atmosphere, within which sensitivity to initial perturbations can be examined to infer predictability. Identical twin experiments are one example. In this case, the rate of divergence between a control simulation and a perturbed simulation is critical. When that rate of divergence is sensitive to details in the model configuration, confidence in the results is undermined. Given sufficient existing data, investigators can find events or time periods not sensitive to model configuration, and base predictability studies around those.

Alternately, attempts to quantify model inadequacy may benefit from a large multi-physics or multi-model ensemble. Despite nearly two decades of research showing empirical benefits of multi-model ensembles, the results have not been effectively used to produce fundamentally superior models. The community agrees that development of scale-aware and stochastic physics, for example, is a current frontier in need of attention. Objective methods to formulate them are needed. Reasoning that the empirical evidence of multi-model ensemble skill arises from something more than mere chance, it is possible that the information needed to proceed is contained in structures emerging from the variety of current model configurations. An accessible data set may enable greater attention to these frontiers.

2.1 Requirements

Science drivers for a system that allows easy sharing and analysis of a multi-model ensemble system lead to a set of requirements on that system. Here we list those requirements, link them back to the science drivers discussed above, and note some design implications:

- Resolution sufficient to permit explicit convection.
    - Allows studies relevant to future ensemble systems, whether global or limited area.
    - Most universities cannot run an ensemble at these scales, but may be able to run deterministic simulations in near-real time.

- An ensemble of order 100 members.
    - Allows analysis to find weak, but important, signals.
    - Most universities cannot run an ensemble this large, but may be able to run smaller ensembles in near-real time.

- A wide variety of model configurations.
    - Allows a thorough exploration of variability or consistency among members, to inform a variety of studies.
    - Universities can be free to choose one or more configurations that meet their local needs.

Contributors to any system addressing these requirements should be able to extract value equal to or greater than the effort put in. Discussions during the workshop touched on several design characteristics that could maintain value for contributors while meeting the requirements underpinning the science. In particular, finding the appropriate level of coordination is essential. A balance between contributing some part of what is already running in an NWP system at a university, and enforcing enough common ground to make the data tractable, can be struck to ensure that the most scientific value is extracted.

2.2 Approaches

During the workshop, the group felt that some coordination would be helpful. One approach would be to require that each group contribute one or more ensemble members on a common, pre-defined grid at approximately 10-km grid spacing covering the CONUS. As a group, we can then agree on 3-5 subdomains at around 3.3-km grid spacing, which is convection-permitting. Those sub-domains would be defined based on the regional interests of contributing groups. In all domains, individual groups would have the freedom to configure other aspects of the model such as physics, diffusion, time step, initial conditions, and boundary conditions. The CONUS ensemble meets the need for a large and variable ensemble, while minimizing the need for regridding prior to analysis. The convection-allowing ensemble would be smaller, but contain some variability.

Such a university-based ensemble raises some ancillary service needs. An example is a re-gridding service. Re-gridding can be used to put disparate forecasts on a common grids for plotting or analysis, and more generically to match forecasts with verifying observations. In the spirit of a new paradigm, the capability would be deployed as a service that executes on a machine close to the raw data so that large data transfers are unnecessary. Basic regridding, or observation-forecast pairing, services enable immediate research opportunities. One example is post-processing to calibrate large and diverse ensembles, toward optimizing reliability and resolution of the forecasts. This immediately leads to work identifying how skill varies with ensemble size, both with and without calibration.

A university-based ensemble, where each participant has the ability to analyze member forecasts from many other universities, has added educational benefits. Details are outlined later, but we note it here to make obvious the link between research and educational needs served under the paradigm we envision.

Hardware or software infrastructure that might result from such a system should be extensible and adaptable to other science areas of scientific inquiry. We next discuss key components of a system to meet our scientific needs, but that we expect would have broader utility.

3. Infrastructure/software for data investigation

Throughout this workshop the need for technological solutions for managing large amounts of data distributed across the nation and across the world was apparent. During the afternoon of the first day of the workshop we started out with 7 speakers talking about existing efforts to build such an infrastructure spanning from practical systems available today to current research trends in scalable scientific data management. We then followed this session by a 2-hour lively discussion.

The discussion revealed a number of important findings:

1. A key barrier for this community is the lack of what we call here "nuclei", shared digital artifacts that allow the community to share and curate data and to continually encode and improve best-practices, tools, data products, workflows, and analyses of that data. These nuclei will only be adopted if the individual cost of maintaining them are significantly lower than the individual benefit of using them. Many successful examples of these nuclei exist, the most famous examples being Linux and Wikipedia. Immediate candidates specific for the regional NWP community might be:

    a. A virtual machine that has a standard environment with all tools necessary to conduct successful weather modeling.
    b. A library of workflows and recipes on how to accomplish particular data products, charts, or visualizations.
    c. A library of data types with access methods that allow the efficient access and processing of remote data.
    d. An world-wide web of catalogue servers that allows access, reference, and provenance tracking of distributed data sets.

Of these candidate nuclei, (a) is a restatement of the nucleus (3) identified in Section 1. Candidates (b)-(d) are elements of the broader nucleus (2).

2. Important technical, political, cultural barriers exist that make the

formation of nuclei difficult:

a. Lack of available mass storage with long-term durability: local storage capacity is relatively affordable but making it available to the community is costly, and its long-term durability is uncertain. Cloud storage is available but is still expensive and its availability depends on continued funding.
b. Cost and lack of incentives for contributing data products: the scientific community gets rewarded for paper publications but not for publishing the data products that underlie the publication. This lack of cultural and political incentive is exacerbated by the cost of maintaining the availability of data products.
c. Insufficient opportunity and trust to rely on remote availability of data products: due to a history of technical challenges of sharing data sets, members of the community feel more comfortable if the data sets they rely on is stored locally. Scratch spaces at large-scale high-performance computing sites enforce policies that only store data for very limited time periods (necessitating frequent and large amounts of data movements to preserve data).

3. Another key barrier (not only for this community) is the lack of infrastructure to manage the provenance of data products. A key requirement for effective provenance management is a shared environment with a global name space for data sets and other artifacts of scientific work. Provenance management is also a key enabler for sharing scientific work, not only for qualitative properties such as attribution but also quantitative metrics such as uncertainty introduced by common operations such as regridding and multi-resolution encoding. Provenance information reveals workflows and recipes on how to achieve a particular kind of data product and is thus essential for the ability to reproduce results and provides a rich basis for tutorials and education.

3.1 Requirements

The main conclusion from these findings is that the community needs to find a way to build a distributed infrastructure that fosters nuclei and supports provenance using existing community resources. The infrastructure has to consist of open source software so community members can freely contribute and continually improve it. NCAR and UCAR/Unidata are in a position to help develop tools and fund equipment that allows members of the community to install such an infrastructure. However, a concerted effort is needed, for example in form of a Software Institute, to coordinate these efforts, to support open source communities, and to encourage and monitor adoption.

3.2 Approaches

One possible way forward is to start out from the most well-catalogued data items with the most well-managed provenance information, such as publications and technical reports, and expand that management to the provenance of figures and tables. We envision a scenario where one reads a pdf paper online and is able to click on a chart that then leads the reader to a page that not only includes all available access methods to the underlying data product, but also the recipe that created the chart using these access methods.

We think a distributed catalogue service such as RAMADDA with an extensible document type system can provide an affordable infrastructure for serving such pages by leveraging Web service infrastructure that already exists in most organizations. Not all organizations will be able to locally store the whole provenance of data products that lead to a particular chart. Catalogue servers will have to be able to provide at least provenance information of data products and, if possible, reference remotely stored data sets (including data sets stored in the cloud).

Our immediate challenges can be addressed with existing capability, and here we are advocating the Unidata-supported RAMADDA to facilitate the nuclei identified in Section 1. As RAMADDA servers become more popular, their resource demands also become greater; underlying local or network file systems will need to be replaced with higher performance parallel file systems.

Assuming our community is able to coalesce around RAMADDA, then further consideration of the data storage will be needed as data volume, cataloging, and provenance grows. Behind RAMADDA, big data volumes can be abstracted on open-source, parallel storage, systems such as Ceph and Galileo. Then even small organizations can afford to run them on small clusters; distributed big-data enterprise on regional NWP also becomes a reality.

The following are elements of a road map on how to get there:

1. Encourage members of the weather prediction community to apply for Unidata equipment grants.
2. With these funds install hardware necessary for a deploying a sharing infrastructure.
3. Use RAMADDA as a key component of this sharing infrasture. RAMADDA is a REpresentative State Transfer (RESTful) catalogue server with an extensible and powerful document type system. This type system, for example, allows the storage of pdf files and enhances them so their figures link to data sets and chart recipes, or allows the representation of data sets with all of its access methods.
4. Develop new RAMADDA document types or specialize existing ones to allow the online creation of figures based on online data sets such that these figures can be embedded into pdf files.
5. Develop an enhanced paper authoring environment that automates these tasks while blending into existing work patterns with as little extra effort as possible.
6. Get Google Scholar and other attribution tracking sites to track these new kinds of citations of data sets (possibly via figures or tables).
7. Extend RAMADDA with the ability to link to remote data sets managed by other RAMADDA servers using advanced data access methods such as regridding, filtering, and other server-side data transformations (using the OpenDAP protocol and possibly extending it).
8. Extend RAMADDA to keep track of the provenance of each data product it stores to the extent that the provenance information is executable and reproduces the given data set, and can provide workflows for similar data sets.

The science drivers from Section 2, and the data investigation capability outlined in Section 3, provide a foundation for educational opportunities at both big research universities and primarily undergraduate institution. The next section documents the workshop discussions on those.

4. Educational Opportunities

On the second day of the workshop, the focus was educational needs and opportunities associated with the three nuclei. Education is a high priority because of a strong need in the community for better training about ensembles in operational (weather forecasting) and research (numerical weather prediction) centers. In addition, the universities are built to educate, hence providing a real opportunity to add value beyond what the operational centers can provide.

Two primary audiences were identified:

- atmospheric sciences majors
- non-majors and the broader community

Although non-majors and the community are an audience of particular interest, we did not spend much time discussing possible solutions, as it was agreed that the *first step* would involve atmospheric sciences majors, including both operations and research foci. Once tools and procedures have been developed for working with students in the field, it will be a next step to provide access to non-majors and the community. The rest of the educational discussion is therefore focused on atmospheric sciences students.

The primary barriers to introducing ensemble prediction into the classroom are:

- foundational knowledge: to fully introduce ensemble prediction theory, a student needs a background that would include computer science, statistics, calculus, and physics/dynamics
- computing facilities and knowledge: processor speeds, storage, software tools, and support personnel may all be lacking, particularly at smaller universities; additionally, there is often slow knowledge transfer between computer scientists and earth scientists, so the earth scientists stick to what they know: “find, download, convert the data (90% of time spent) and analyze, interpret the data (10%)”
- transparency: systems continue to become more distributed, with

input data, the models themselves, and even the post-processing, all happening remotely; there is a strong need to make the process less abstract, particularly for students just being introduced to the concepts
- forecasting exercises deterministic: there is some development to make forecast exercises more probabilistic, but change is slow, e.g. the most popular national forecasting exercise, WxChallenge, is deterministic

4.1 Requirements

In order to facilitate cooperation and maximize learning, several tools were identified. The first is covered by nucleus (2) and the second by nucleus (3):

1) A data catalog service, such as RAMADDA.

This data catalog would allow sharing across working groups, ranging in size from within a department to between universities and agencies. A data catalog such as RAMADDA also allows sharing of tools (e.g. IDV scripts) and analyses.

2) A virtual machine, either run in a virtual environment or run natively.

The virtual machine would provide the environment and tools for streamlined access to and processing of model output, thereby negating one of the major hurdles to use in the classroom at many universities, caused by insufficient time and/or support staff to get the models and visualization tools up and running. This virtual machine could be developed and maintained by NCAR and/or UCAR.

It was clear from discussion that these tools need to be part of the ensemble development *from the beginning,* whether or not an educational component is immediately tested. Such tools need to be integrated into the ensemble framework, not tacked on later on.

In addition to the above tools, sharing course development progress and facilitating dissemination was also discussed. As this is by its very nature a distributed project, with a significant online presence, it may be a great fit for an online course (e.g. Coursera, COMET). It is likely that initial modules would at least serve as the building blocks for a later online course.

4.2 Approaches

As listed above, significant foundational knowledge is needed to cover full ensemble theory, meaning that most undergraduates do not discuss NWP and ensemble theory until their senior year (if at all). However, students are using forecast model output from freshman year, and often have been using it as a “weather enthusiast” well before entering university. So, instead of waiting until senior year, when many forecast model misconceptions have already been cemented, ensemble prediction topics should be integrated throughout the atmospheric sciences curriculum, from freshman to senior year. (Note that development of modules for underclassmen also help in translating to non-majors and the community in later stages of the project.)

To achieve this goal, it was discussed that workshop participants from different universities could pick a certain class and develop a module for that class that would introduce an important aspect of ensemble prediction. For example, “What are Ensembles and Why Do We Use Them” in Intro to Met., “Intro. to Data Assimilation” in Instrumentation, “Interpreting Ensemble Output” in Synoptic Met., “Parameterizations” in Physical Met., etc. A key component would be hands-on activities, however the Educational Requirements discussed above would be crucial in making hands-on activities possible. Additionally, there is value in integration of these large ensemble prediction data sets into computer sciences classrooms to demonstrate real-life big data needs. Even among atmospheric science graduate students early in their programs, there is still often a lack of deeper understanding of NWP models beyond as a “black box,” so a large university ensemble could provide opportunities for innovative education at that level as well.

4.3 Graduate student breakout session

During the workshop on Day 2, a graduate-student breakout session emphasized the potential value of a coordinated effort around the nuclei identified above. The students were given broad instructions to come up with a method to incorporate the multi-university ensemble into the classroom.

The students were able to outline a graduate course in ensemble prediction that included topics such as model uncertainty. It also included significant use of the proposed tools to get hands-on work with a modeling system. Comments from the students confirmed that a comprehensive and accessible toolset for quickly executing and modifying ensembles with real models would be valuable. The computer science students commented that easier access to ensemble systems, and especially the system outputs, is valuable for providing hands-on examples for their discipline as well.

Involving graduate students in the workshop process was valuable for both students and investigators. Students commented that participation provided a broader view of challenges and research directions than they are typically exposed to in their own graduate research niche. The student involvement was valuable for investigators because the students have an "in the trenches" view of current big-weather challenges, and tend to be more savvy than more senior researchers on the latest techniques and tools (e.g. cloud storage and social networking).

5. Broader Impacts

The broader impacts of this endeavor are many, but here we highlight three specific areas: student involvement, inclusion of non-majors in ensemble prediction, and value to the EarthCube (http://www.earthcube.org) initiative. Significant student involvement has already been initiated, as ten graduate students, representing several different disciplines, participated in the workshop. Of particular note is the graduate student breakout session discussed above.

As mentioned above, this paper focuses on the steps necessary for research and education primarily in the atmospheric and computer sciences, but the tools being developed will be the first steps towards facilitating collaboration and communication with both researchers in a wide array of other disciplines and also with the general public. To state it simply, facilitating access for ourselves facilitates access for everyone. Other disciplines that would be involved include any discipline that uses atmospheric data (e.g. biology, chemistry, emergency management). Communication with the general public, e.g. communicating scientific uncertainty and elucidating the scientific process, would be greatly improved with the increases in transparency.

Ensemble prediction requires large amounts of data, and, by necessity, researchers have already developed the capabilities to centralize data from a myriad of observational platforms, push that data to institutes across the globe, and process terabytes of data at multiple operational steps. The large value to the EarthCube initiative of a multi-university ensemble prediction system is *active testing* of cutting edge tools and workflow schemes. A multi-university ensemble represents a community already using the most advanced data processing systems in earth science. In addition, such an ensemble is small enough to be flexible and allow quick implementation times, but also large enough to involve multiple types of users (varied in size and technical expertise) and several disciplines.

Appendix: Attendees (* denotes student)

Last	**First**	**Affiliation**
Ancell	Brian	Texas Tech University, Department of Geosciences
Arms	Sean	Unidata
Baker	Allison	National Center for Atmospheric Research, Computational and Information Systems Laboratory
Berner	Judith	National Center for Atmospheric Research
Brammer*	Alan	SUNY Albany, Department of Atmospheric and Environmental Sciences
Buck*	Joe	University of California, Santa Cruz, Department of Computer Science, Jack Baskin School of Engineering
Burghardt*	Brock	Texas Tech University, Department of Geosciences

Cao*	Yang	University of California Los Angeles, Department of Atmospheric and Oceanic Sciences
Capehart	William	South Dakota School of Mines and Technology, Department of Atmospheric and Environmental Sciences
Crume*	Adam	University of California, Santa Cruz, Department of Computer Science, Jack Baskin School of Engineering
Exby	John	National Center for Atmospheric Research, Research Applications Laboratory
Hacker	Josh	National Center for Atmospheric Research, Research Applications Laboratory
Jewett	Brian	University of Illinois, Department of Atmospheric Sciences
Lawson*	John	Iowa State University, Department of Geological and Atmospheric Sciences
Maltzahn	Carlos	University of California, Santa Cruz, Department of Computer Science, Jack Baskin School of Engineering
McWhirter	Jeff	RAMADDA
Mullendore	Gretchen	University of North Dakota, Department of Atmospheric Sciences
Pallickara	Sangmi	Colorado State University, Computer Science Department
Peterson*	Brittany	Univ of North Dakota, Department of Atmospheric Sciences
Potterjoy*	Jon	Penn State University, Department of Meteorology
Ramamurthy	Mohan	Unidata

Reinhart*	Tony	Texas Tech University, Department of Geosciences
Schumacher	Russ	Colorado State University, Department of Atmospheric Science
Tyle	Kevin	SUNY Albany, Department of Atmospheric and Environmental Sciences
Watkins*	Noah	University of California, Santa Cruz, Department of Computer Science, Jack Baskin School of Engineering