AI RFI Responses, October 26, 2018

Update to the 2016 National Artificial Intelligence Research and Development Strategic Plan RFI Responses

National Center for Atmospheric Research
P.O. Box 3000, Boulder, CO 80307 USA
www.ncar.ucar.edu

NCAR

25 October 2018

RFI Response
National Science Foundation
Update to the 2016 Artificial Intelligence Research and Development Strategic Plan

From the National Center for Atmospheric Research

Scientists and Software Engineers from the Research Applications Laboratory and the Computational Information Systems Laboratory of the National Center for Atmospheric Research (NCAR) have reviewed the RFI materials, Update to the 2016 Artificial Intelligence Research and Development Strategic Plan, and are pleased to have the opportunity to respond.

Operations and Mission of the National Center for Atmospheric Research

NCAR's mission is "to conduct research that contributes to the depth of fundamental understanding of the atmosphere and its interaction with society and the environment and develop and transfer knowledge and technology that expands the reach of atmospheric science." NCAR has a successful history of transferring technology and knowledge to U.S. government agencies, the private sector, and foreign governments. NCAR is eager to collaborate with other organizations to address the use of Artificial Intelligence (AI) in the environmental sciences.

NCAR is operated by the University Corporation for Atmospheric Research (UCAR), a non-profit organization established in 1960 to oversee a wide range of programs and facilities that support its 100+ university affiliates, as well as the national and international scientific community.

NCAR's Role in AI Research

NCAR has a long-standing research program in artificial intelligence (AI) and machine learning (ML). That research began under externally funded projects and has helped lead the meteorology and climate community to recognize the efficacy of applying these methods to weather forecasting, applications of hydrometeorology, climate simulation, and various applications including aviation meteorology, renewable energy[1], surface transportation[2], forecasting in support of agriculture and food security[3], modeling for wildland fire management, and much more. These applications have

[1] Mahoney, W.P., K. Parks, G. Wiener, Y. Liu, B. Myers, J. Sun, L. Delle Monache, D. Johnson, T. Hopson, and S.E. Haupt, 2012: A Wind Power Forecasting System to Optimize Grid Integration, special issue of *IEEE Transactions on Sustainable Energy* on Applications of Wind Energy to Power Systems, **3** (4), 670-682.

[2] Drobot, S. D, M. Chapman, P. A. Pisano, and B. B. McKeever, 2010: Using vehicles as mobile weather platforms. Data and Mobility: Transforming Information into Intelligent Traffic and Transportation Services, Springer, 203–214.

[3] Myers, W., F. Chen, J. Block, D. T. N. Meteorlogix, and M. N. Burnsville, Application of Atmospheric and land data assimilation systems to an agricultural decision support system. Conference on Agriculture and Forestry, Orlando, FL, American Meteorological Society, 2008.

become commonplace and critical to providing timely and accurate forecasts in these many realms. From the advent of NCAR's Dynamical Integrated foreCast (DICast®) engine nearly 20 years ago[4], to applications of genetic algorithms in homeland security[5], and including using methods such as random forests, fuzzy logic, neural networks, regression trees, analog ensembles, gradient boosted regression, and convolutional neural networks for various forecasting applications, NCAR has helped push the use of AI in many areas of environmental science and published widely in these fields.

More recently, NCAR has embarked on a major effort to expand the use of AI in environmental systems. For instance, the Computing and Information Systems Laboratory has teamed with other NCAR science labs including the Research Applications Laboratory, High Altitude Observatory, and Climate and Global Dynamics Laboratory to build, demonstrate, and validate ways to replace certain physical models in our various modeling systems with AI emulators. The preliminary results of these initiatives are quite encouraging and are expected to lead to some major advances in model accuracy and efficiency. This work is funded through reinvestment of NSF Base funds at NCAR.

Thoughts on Updating NSF's 2016 Artificial Intelligence Research and Development Strategic Plan

NCAR scientists are pleased to see NSF take an active interest in topics relevant to advancing the use and prominence of AI research in the scientific community. Both basic and applied research are needed to fully leverage these techniques that are revolutionizing how we interact with technology and how we model our physical environment. As scientists, we seek to understand as well as to replicate the physics and dynamics of our modeled systems. We wish to harvest the full capabilities of these techniques to advance our ability to model the natural world. However, we need to ensure that AI methods accurately represent the full spectrum of physical processes and enable interpretation of the patterns they discover. Thus, *we believe that it is critically important to include in the revised plan initiatives that include modeling the environment and applications of AI in environmental science*. There are numerous complex, nonlinear interactions among multiple parameters governing processes in the environment. These processes have resisted our efforts to develop effective simplistic mathematical models, often due to our inability to process and derive meaning from a vast amount of data. AI approaches are being utilized to develop more accurate representation of the physical processes that, in turn, can lead to a better understanding of environmental systems.

A specific recommendation is to explicitly include AI in environmental modeling and simulation as a topic that contributes to economic prosperity. NCAR scientists have quantified the U.S.

---

[4] Myers, W., G. Wiener, S. Linden, and S. E. Haupt, 2011: A consensus forecasting approach for improved turbine hub height wind speed predictions, in *Proc. WindPower 2011*, Anaheim, CA, May 24, 2011.

[5] Rodriguez, L.M., S.E. Haupt, and G.S. Young, 2011: Impact of Sensor Characteristics on Source Characterization for Dispersion Modeling, *Measurement*, **44**, 802-814.

economic sensitivity to weather variability as $485Billion/yr.[6] Thus, improvements in forecasting, including from applying AI, can provide situational awareness of the weather, which can directly impact economic costs as well as the safety and well-being of the human population. AI has proven useful in improving weather forecasts of temperature, precipitation, icing, and even the path and intensity of hurricanes. In addition, applications of AI can reduce costs of integrating renewable energy into the grid, which enables more renewable energy and the resulting decrease in CO_2 and other pollutants.[7],[8] AI methods can be used to predict which storms are most likely to produce hail.[9] These and similar applications are making substantial impacts in how the meteorology community provides necessary support to the public as well as for commercial interests.

The current state of AI is rapidly changing and that section of the strategic plan should be a living document with frequent updates. For instance, the applications of "deep neural networks" have expanded rapidly over the past couple of years. We agree with the assertion that the third wave of AI advance must focus on explanatory and general AI technologies. By understanding the AI advances in the environmental sciences, we are going to learn more about the systems they simulate.[10]

Specific Comments Regarding the Strategy of the Plan

The items listed as part of the strategy make sense, as far as they go. But there could be a few additions and modifications. Some specifics are listed below:

1. **Long-term investments in research** – We very much agree with this part of the strategy and there could be a large gain in economic benefit through government investment in research. Although industry also invests in research, their investment is more targeted toward their short-term financial gain and bottom line. On the other hand, government investment goes toward more basic research, which will advance the state-of-the-science and benefit the greater good. An example is the improvements that have come through research in applying AI to weather data. That has resulted in improved situational awareness, actionable insight, and forecasts for everyone. In addition, the application, including to surface and air transportation, has enhanced public safety, and to renewable

[6] Laxo, J.K., M. Lawson, P.H. Larsen, D.M. Waldman, 2011: U.S. Economic Sensitivity to Weather Variability, Bulletin American Meteorol, Soc., 709-720.

[7] Haupt, S.E., B. Kosovic, T. Jensen, J. Lazo, J. Lee, P. Jimenz, J. Cowie, G. Wiener, T. McCandless, M. Rogers, S. Miller, M. Sangupta, Y. Xue, L. Hinkelman, P. Kalb, J. Heiser, 2018: Building the Sun4Cast System: Improvements in Solar Power Forecasting, *Bulletin of the American Meteorological Society*, Jan. 2018, 121-135. doi: 10.1175/BAMS-D-16-0221.1

[8] Kosovic, B., S.E. Haupt, D. Adriaansen, S. Alessandrini, T. Jensen, G. Wiener, Y. Liu, L. Delle Monache, 2018: Scientific Advances in Wind Power Forecasting, submitted to *Renewable Energy*.

[9] Gagne II, D.J., A. McGovern, S.E. Haupt, R. Sobash, J.K. Williams, and M. Xue, 2017: Storm-Based Probabilistic Hail Forecasting with Machine Learning Applied to Convection-Allowing Ensembles, *Weather and Forecasting,* **32**,1819-1840.

[10] Gagne II, D.J., S.E. Haupt, and D.W. Nychka, 2018: Interpretable Deep Learning for Spatial Severe Hail Forecasting, submitted to *Monthly Weather Re*view.

energy, has resulted in higher deployment of renewables, reducing the emissions of CO_2 and other pollutants from the energy industry. Research currently being done on applying the most modern advances, such as deep learning, promise to make even bigger impacts toward providing added public safety and economic benefits; for example, better forecasting for hail allows the public to take steps to protect their safety and property. We also agree that advancing more general-purpose AI will lead to large public benefits.

2. **Effective methods for human-AI collaboration** – Once again, the meteorology community has been leading in this arena. The AI that is applied to postprocess the weather model output works well with the human forecaster. We often refer to the "human above the loop" as an approach where the forecast is generally automated, but the human forecaster oversees and has the ability to correct any obvious bad directions that may be made by the AI system. For example, AI has been applied to recommend best routes for aviation in order to avoid developing thunderstorms. But with the addition of human oversight, any perceived erroneous forecasts can be corrected via human input to assure that the computational decision agrees with the human judgement. That process combines the best of both the human and the AI.
3. **Understand and address ethical, legal, and societal implications of AI** – Application of AI technology to automating processes necessarily moves a human further away from the direct effects of decision making. However, this cannot be a license to develop and apply AI without the same deliberate process used when humans are directly affecting outcomes. AI must be applied ethically. In the meteorology community, that often includes respect for the privacy of data that are used in the machine learning and internet of things (IoT) applications. At times, those data are obtained under non-disclosure agreements, which must be respected. At other times, data may come from individuals and their privacy should be respected. For instance, when applying AI to surface transportation problems, data may be obtained from individual vehicles. Where those vehicles are publicly owned, such as by state Departments of Transportation, the data may be considered public. But when we start gathering more data from privately-owned vehicles, that data must only be used anonymously to provide better road weather forecasts for public benefit. The data must be protected from other uses that might disclose too much information about the location and the actions of the private citizen. As mentioned in the strategic plan, architectures must be designed to protect such private data.
4. **Ensure the safety and security of AI systems** – We agree that to build trust in AI systems, those systems must be reliable, informative, and interpretable. Humans must be able to understand the basic workings of the constructed technology and it must be thoroughly verified and validated, much as we as scientists are always expected to do in the weather and climate community. When we build specific weather systems using AI, we always validate against an accepted state-of-the-science baseline to assure that it does indeed provide improvement. This is where encouraging publication of the advances in the peer-reviewed literature contributes to such rigorous assessment. Before systems are deployed, they must be thoroughly vetted. The scientists must understand their shortcomings as well as their advantages and determine where they should not be applied as well as where they work well. Again, vetting in the peer-reviewed literature aids this process. As noted in the strategic plan, security against nefarious attack should also be considered for any system that is expected to operate consistently and reliably. Again, the weather community has

considered such security and regularly applies best practices in constructing reliable systems.

5. **Develop shared public datasets and environments for AI training and testing** – This is certainly happening in the weather community as well as outside it. Several organizations interested in advancing the use of AI in the weather community are actively working to identify and deploy high-quality datasets for AI training and comparison for advances in the use of AI. The American Meteorological Society (AMS) has fostered a Committee on the Application of Artificial Intelligence in the Environmental Sciences since at least the late 1990s and began teaching short courses in the topic in 2001, eventually archiving the lectures in a book.[11] In 2008, that committee began holding AI forecasting contests, providing a common dataset that researchers could use to test their algorithms. More recently, the AMS AI Committee has been banding with other organizations to provide complex series of datasets that researchers can explore and compare their applications with others. When the committee turned to the kaggle competition website in 2014 to host their contest, they were surprised to find that some non-meteorology AI experts applied techniques that beat those currently being applied in the weather community. Those techniques were quickly adopted by many others doing parallel research. Thus, the weather community has evidence that such common datasets are effective in helping to advance the state-of-the-science.
6. **Measure and evaluate AI technologies through standards and benchmarks** - As described under #4 above, to build trust in systems requires standard rigorous assessment, which is always expected when AI is applied in a scientific environment. For instance, an analog ensemble approach has been recently applied to improve the prediction of hurricane intensification and track. To validate that method requires comparing it to the current best available systems, which has been accomplished. Another example is when building modeling and AI systems to improve solar power forecasting under Department of Energy funding, part of the project was to determine and build the rigorous systems to advance this forecasting methodology beyond the current difficult baseline.[12],[13]
7. **Better understand the national AI R&D workforce needs** – We agree that facility in AI is becoming critical to functioning well in future workplaces. We need people capable of building and maintaining the computational infrastructure necessary to continue to push the state-of-the-science. These people must know how to deal with "Big Data" at all levels, including how to manage datasets. When AI is used in production involving the types of big datasets that we use in meteorological applications (blending in situ and remote

[11] Haupt, S.E., A. Pasini, and C. Marzban, Eds. 2009: *Artificial Intelligence Methods in the Environmental Sciences*, Springer, 424 pp.

[12] Jensen, T.L, T.L. Fowler, B.G. Brown, J. Lazo, S.E. Haupt. 2016: Metrics for evaluation of solar energy forecasts. NCAR Technical Report TN-527+STR, 67 pp, doi:10.5065/D6RX99GG.

[13] Haupt, S.E., B. Kosovic, T. Jensen, J. Lee, P. Jimenez, J. Lazo, J. Cowie, T. McCandless, J. Pearson, G. Weiner, S. Alessandrini, L. Delle Monache, D. Yu, Z. Peng, D. Huang, J. Heiser, S. Yoo, P. Kalb, S. Miller, M. Rogers, and L. Hinkleman, 2016: The SunCast Solar Power Forecasting System: The Results of the Public-Private-Academic Partnership to Advance Solar Power Forecasting. NCAR Technical Report TN-526+STR, 307 pp, doi:10.5065/D6N58JR2.

observations with output from many models smartly to improve forecasts[14]), this aspect becomes as important as producing the new AI algorithms. These personnel trained in AI and Big Data must be prevalent in public positions as well as in private enterprise. In the meteorology community, several universities are providing training in AI at the undergraduate as well as graduate level. National laboratories must emphasize applications of AI. At NCAR, we have re-invigorated a cross-laboratory effort in AI, both in basic research as well as in the applications that we have fostered over the years. We firmly believe that applying these principles more broadly across a range of topics is needed to assure continued advances in our technology.

8. **Continued development of core AI based applications that push the science, serve the public and benefit the nation -** The development of innovations in AI have stemmed from the research community, academia, national laboratories, and the private sector. Some of these advances have sought to resolve and improve the predictability and understanding of environmental problems. *Using AI to address these environmental science problems successfully has contributed to the enhanced safety and well-being of the population and economic prosperity of the nation. The strategic plan should support and promote the continued development of weather, climate, and environmental applications.*

Summary:

We are pleased that NCAR has long been part of the AI community and are proud that the weather community has been a leader in these areas. We look forward to continuing to work with NSF and other U.S. government agencies to advance the research that will provide next generation models and predictions for environmental processes.

[14] Haupt, S.E. and B. Kosovic, 2017: Variable Generation Power Forecasting as a Big Data Problem, *IEEE Transactions on Sustainable Energy,* **8** (2), pp. 725-732. DOI: 10.1109/TSTE.2016.2604679.