

Realizing the Benefits of AI across the Numerical Weather Prediction Value Chain

Adapted from “Leveraging Modern Artificial Intelligence for Remote Sensing and NWP: Benefits and Challenges,” by **Sid-Ahmed Boukabara** (NOAA/NESDIS Center for Satellite Applications and Research), **Vladimir Krasnopolsky**, **Jebb Q. Stewart**, **Eric Maddy**, **Narges Shahroudi**, and **Ross N. Hoffman**. Published in BAMS online, December 2019. For the full, citable article, see [DOI:10.1175/BAMS-D-18-0324.1](https://doi.org/10.1175/BAMS-D-18-0324.1).

Artificial Intelligence (AI) techniques have had significant recent successes in multiple fields, from financial forecasting to speech recognition. These fields and the fields of satellite remote sensing and Numerical Weather Prediction (NWP) share the same fundamental underlying needs, including signal and image processing, quality control mechanisms, pattern recognition, data fusion, forward and inverse problems, and prediction. Thus, modern AI in general and machine learning (ML) in particular, can be positively disruptive and transformational in the fields of satellite remote sensing and NWP—achieving the change needed to meet the increasing challenges of Big Data, advanced models and applications, and user demands.

Operationally, AI has only applied to “niche” applications and with limited, hard-won successes. However, forms of AI including ML have been applied successfully in remote sensing and weather prediction and have shown much promise for a wide range of geophysical problems, from aiding human interpretations to discovering new relationships in large datasets and making postforecast corrections. Clearly, AI has the potential to meet increasing requirements by and for nowcasting, forecasting, and climate projection.

AI, and specifically ML, may augment or replace components of the NWP processing chain to 1) speed up and improve the processing of satellite data; 2) facilitate data assimilation and

initialization of numerical weather and climate models; 3) speed up and improve model physics; and 4) improve postprocessing of model outputs.

Global weather forecasting uses only about 1–3% of currently available satellite data, and the processing time for traditional approaches using these data is already crippling and expected to increase. New sources of data, such as the Internet of Things providing observations from observation towers, phones, and automobiles, will also complement traditional data sources. Moreover, weather forecasting, including NWP, has unmet and increasing needs for faster and/or more accurate methods of calculation. ML can provide such methods.

Additionally, weather data users—from pilots to emergency planners to farmers—increasingly demand higher accuracy and greater resolution. As a result, weather products must combine data sources at increasing spatial and temporal resolutions, and with improved timeliness. ML will be important to enabling the new approaches needed to ingest more sources of observations into increasingly more accurate (and computationally demanding) data assimilation and forecast systems, and to process and disseminate the results.

Advantages and Challenges

Advantages of ML methods include, but are not limited to, computational efficiency, accuracy, transferability (from other fields), synergy with other approaches, and ease of use. With these advantages, ML offers new paradigms for utilizing a large fraction of available data, improving parameterizations, tailoring products for users, and meeting other NWP challenges.

An example of the advantages of ML is the use of satellite radiance or brightness temperature observations. ML accelerates computation and improves algorithms (e.g., by using training datasets derived from highly accurate physically based models that are too expensive for use in remote sensing and parameterizations). As a result, ML has the potential to improve overall accuracy and quality within tightening latency requirements. For real-time data assimilation and forecast applications, fast versions of radiative transfer are generally used. These rely on parameterizations that are based on approximations and require tuning, introducing errors. In this situation, ML can be a game-changer: orders of magnitude execution acceleration allows a larger set of observations to be exploited, with the potential of improved

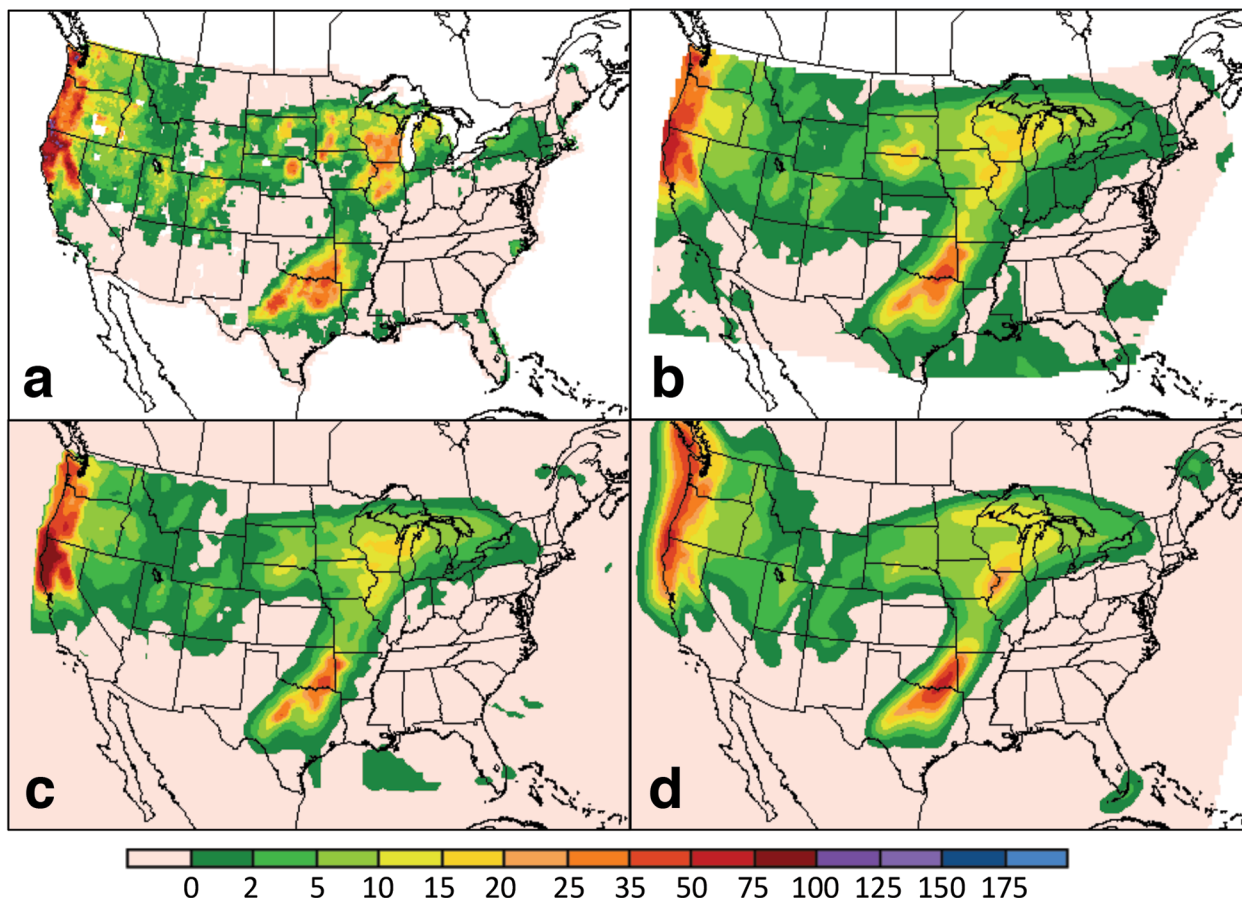
accuracy/quality. This improved accuracy/quality is obtained by training the ML algorithms, emulations, approximations, and models using data simulated by the most accurate radiative transfer calculations. Once tuned, ML will be exceedingly fast without loss of accuracy.

All of this comes with several caveats. For example, the ML models must be accurate for all situations, including very rare cases, therefore putting emphasis on a well-balanced, representative training set. It is worth noting that AI is not a replacement for physical models, but rather a complement. Continued research is needed to create highly accurate and detailed model physics, which can aid understanding of physical processes and which can produce simulated data to improve fast parameterizations based on ML emulations.

Significant experience has been accumulated in developing geophysical ML applications; however, a particularly powerful advantage of ML in geophysics is to leverage existing knowledge from other fields. Current AI systems help predict outcomes in, for example, cancer detection, electrocardiogram (EKG) analysis, algorithmic trading, stock market analysis and prediction, portfolio management, signal extraction in noisy environments, and automatic composition in any desired style of music. There are strong connections between these allied fields and problems in both satellite remote sensing and NWP—through shared fundamental procedures such as forward and inverse problems, morphing, mapping, and pattern recognition. For example, the approaches useful in facial recognition should also be useful in identifying meteorological features such as hurricanes. In addition to reuse of AI methods, “transfer learning” allows for taking an existing AI network trained for one task and retraining only the final stages of that network for a new task.

To meet challenges in geophysics, the key concerns and caveats when applying ML are:

- Will the ML model be reliable, or will it be prone to be less accurate or even to fail for rare or unusual cases? ML models can only learn what is in their training datasets.
- Can ML models satisfy constraints based on physical principles (e.g., conservation of mass)?
- How can AI overcome the trust barrier (i.e., the reluctance of some to accept ML model output if they can’t understand what it is doing in its hidden layers)?



- Can ML models be extended to produce uncertainty estimates (i.e., error bars)?
- Can ML development be disciplined enough to produce reproducible results? Can ML models be easily integrated into operational procedures?

Concerns about reliability (i.e., the ability to handle unusual situations) can be mitigated by representative training. Since ML models should not be used to extrapolate far beyond the domain covered by the training set, the training set should be representative (i.e., large and diverse). Without a representative training set, one could accidentally introduce a bias where a feature (or signature) limited or absent in the training set would become associated by the trained model with other sufficiently similar features. Thus, the choice of training dataset and specific input variables can be critical. When simulations are available (e.g., from a realistic nature run), the quality of the simulations is critical because the ML model trained on simulated data will be applied in reality.

With regard to trustworthiness, nonlinear statistical models of any type are difficult to understand and interpret. However, in response to the concern that ML

Comparison of (a) the NOAA Climate Prediction Center analysis of precipitation for the 24-h period ending at 1200 UTC 24 Oct 2010 to 3 forecasts, including (b) the arithmetic mean of the precipitation forecasts provided by 8 models, (c) the nonlinear artificial neural network average of the same models, and (d) the prediction by a human analyst who used satellite images and ground observations in addition to model forecast. Color bars on the left side of the figures show the color codes for mm day⁻¹ of precipitation. (After Krasnopolsky and Lin, 2012, Fig. 6.)

models are black boxes that cannot be interpreted or understood, efforts are underway to develop explainable AI and physics-guided neural networks. These developments hold the promise of explaining the connections in Big Data extracted by ML and uncovering new physical phenomena. Ghahramani¹ has argued that ML must be able to represent and manipulate uncertainty about models and predictions.

With regard to reproducibility, AI techniques eliminate some forms of human error inherent in many traditional data assimilation techniques. These traditional techniques require manual coding of human estimates based on intuition and approximation to solve what is fundamentally an ML problem—that is, to numerically minimize an objective metric of errors. ML requires experimentation and tuning, which likewise—at least currently—requires some degree of human intuition and intervention. ML development should adopt modern software practices (version control, containerization, etc.) to ensure that interested researchers can reproduce published results.

Concerns about ease of integration place additional demands on the ML software system used for development. For example, to develop an ML parameterization of physics, a very specific normalization of outputs is sometimes required. In some operational cases, the trained tool must be converted to another programming language to be consistent with the parent application.

Applications

AI can be used comprehensively along every point in the value chain of NWP, from inputs to outputs and usage. Some examples:

Satellite observation gap filling is often required due to incomplete coverage or sensor problems. AI can be of value in these situations.

Remote sensing retrievals already have a long history of using AI. This practice is considered mature. It can, however, benefit further from modern ML tools developed for other fields.

Quality control procedures have been built up over time and adjusted to account for multiple sensors. Machine learning can be effective in detecting and correcting observation problems.

Data assimilation and *fusion* can benefit from modern ML techniques for the analysis of observations.

Machine learning is well equipped to detect, evaluate, and therefore correct errors made by physics-based models.

Current data assimilation systems typically use only a small fraction of the available observations. Data fusion combines and/or converts observations and/or imagery into new information.

Model physics calculations are a computational bottleneck in numerical models. For example, accurate radiative transfer calculations are time consuming for retrievals, data assimilation, and model parameterizations. ML models can significantly speed up these calculations. New, fast, and more accurate physical parameterizations can be developed using data simulated by high-resolution and very time-consuming models.

Nowcasting extrapolates current conditions, often imagery, into the near future (minutes to hours). For example, combining data from lightning sensors, satellite imagery, and NWP model output in a convolutional neural network framework can create seamless weather radar mosaics (data fusion) and forecasts out to 12 hours (nowcasting).

Postforecast processing and *correction* are important methods of improving forecasts, either on the original forecast grid or by adapting to local conditions (as in the model output statistics method). ML is well equipped to detect, evaluate, and therefore correct errors made by physics-based models; to adapt forecasts from large-scale models to local conditions; and to nonlinearly average ensembles of forecasts.

A Future Synergy with AI

ML approaches have been shown to be useful in many aspects of environmental prediction. Further, the use of ML by the satellite remote sensing and NWP community is ripe for more rapid advances due to recent progress in applications in the allied fields.

Without AI or some other transformation approach, current difficulties of large amounts of data are expected to be exacerbated by the expected increase in volume and diversity of environmental data from small satellites, IoT, etc.; by the commoditization of technology leading to more Earth-observing satellites; and by sensors with higher spatial and spectral density and resolution. The efficiency of ML will allow assimilating high-resolution data that are now either aggressively thinned or not considered at all. The spatiotemporal resolution and accuracy of ML-based data processing algorithms will thereby increase substantially.

¹ Ghahramani, Z., 2015: Probabilistic machine learning and artificial intelligence. *Nature*, **521**, 452–459, DOI:10.1038/nature14541.

While AI sometimes provides novel capabilities, in many cases AI replaces or enhances existing methods. Therefore, in the future, the distinction between AI and physical science may become blurred. A cohesive and synergistic coexistence of physical models with ML enhancements should emerge, with, for example, ML trained on previous forecast errors to make postforecast corrections for human-induced uncertainties—when modeling, when implementing systems, and even when issuing advisories. Such an AI system supports human practitioners by providing evidence for reconsideration of the forecast (in weather as in medical diagnosis).

Can ML improve on NWP models that are based on the laws of physics and data assimilation systems based on optimal Bayesian estimation? In fact, they can, and both physics-based models and optimal estimation techniques in operational systems have a lot in common with ML techniques. Already, many data assimilation techniques overlap with ML neural network techniques. Many data assimilation techniques may be thought of as “hand-crafted” or “artisanal” ML, prone to errors of implementation, misunderstandings of physical processes, errors of approximation, etc. For example, 3D- and 4D-variational data assimilation are based on clear Bayesian principles, but their implementation is rife with approximations and assumptions. Modern ML provides the opportunity to build on existing techniques, uncover their limitations, and possibly correct their implementation flaws.

ML still requires human expertise to succeed. New model physics parameterizations, for example, are not a standard ML problem. The successful development of a geophysical ML model usually requires domain knowledge about the Earth system. Close collaborations between computer scientists, geophysicists, data scientists, remote sensing experts, and modelers will be essential.

Finally, for AI to be adopted by geophysicists, weather forecasters, and emergency responders, the ML tools must be trustworthy, reliable, and accurate. It should be possible to accelerate ML adoption using lessons from how humans adopt new computer processes. For example, a study of extreme convective event nowcasting concluded that human–ML systems should provide clear presentations and provide explanations of how the ML results are obtained. The human–ML interface means allowing people the flexibility to choose and combine different tools and to correct the ML results. ●●

≡ METADATA

BAMS: Why did you write this article?

Sid-Ahmed Boukabara: *Artificial Intelligence (AI) carries a significant potential for all aspects related to the exploitation of the satellite data and for NWP, from calibration, remote sensing to data fusion, assimilation, and postforecast processing. AI brings significant increases in efficiency, performance, and skills, but some challenges will likely need to be addressed for AI to be fully embraced in this field.*

BAMS: Given the swift rise of AI in our sciences, these challenges don’t seem exclusively technical.

SAB: What was really surprising was the easiness by which machine learning algorithms could be implemented and optimized and how much leveraging could be done—to use architectures and setups used in totally different fields, and transfer them into the atmospheric sciences, and in particular NWP.

BAMS: So what is the major challenge?

SAB: *The biggest challenge is not technical, but instead is mainly a cultural resistance to change. In particular, when working on transitioning this work into operations, our team experienced several forms of resistance to the adoption of AI approaches. It is expected that this challenge will likely get reduced in intensity in the future as the community continues to make further progress toward the AI adoption route.*

BAMS: Given the long history in atmospheric sciences of grappling with large datasets and their problems, this seems somewhat surprising.

SAB: *It was a surprise to some of us as well. The situation is not uniform obviously, but there is a clear sort of “snub” toward AI by some, alluding to the fact that it is a black box, it is not real science, etc. It could indicate a lack of understanding and a cautionary position. Perhaps it is not the case in academic circles but it is definitely the case, as you would expect, in operational circles.*