Current State of Artificial Intelligence Exploitation in AMS Community

Eric B. Wendoloski, Timothy J. Hall, Kiley L. Yeakel, Peter J. Isaacson

The Aerospace Corporation

Abstract

The environmental community has long produced a wealth of mission specific observations, estimations, and simulations. Fusion of these sources traditionally occurs within numerical weather prediction frameworks through data-assimilation cycles that provide initial conditions to forecast models. Myriad environmental forecasting applications exist over all scales, but the environmental community as whole has been slower to adopt the application of artificial intelligence (AI) to these problem spaces than other industries (e.g., financial services, retail, etc.). However, a marked increase in AI-based applications that leverage the wealth of data available in the environmental sciences has been occurring over the last two years. This rapid increase in exploitation has been manifesting itself as a jump in AI-related presentations and publications within the AMS community and increased utilization in the operational meteorological domain. This presentation characterizes the increase in AI-based activity in AMS publications and identifies broad research areas that are reaching maturity using AI-based approaches. Additionally, this presentation discusses the catalysts responsible for this increase in activity along with research vectors that can benefit from AI-based data exploitation.

Introduction

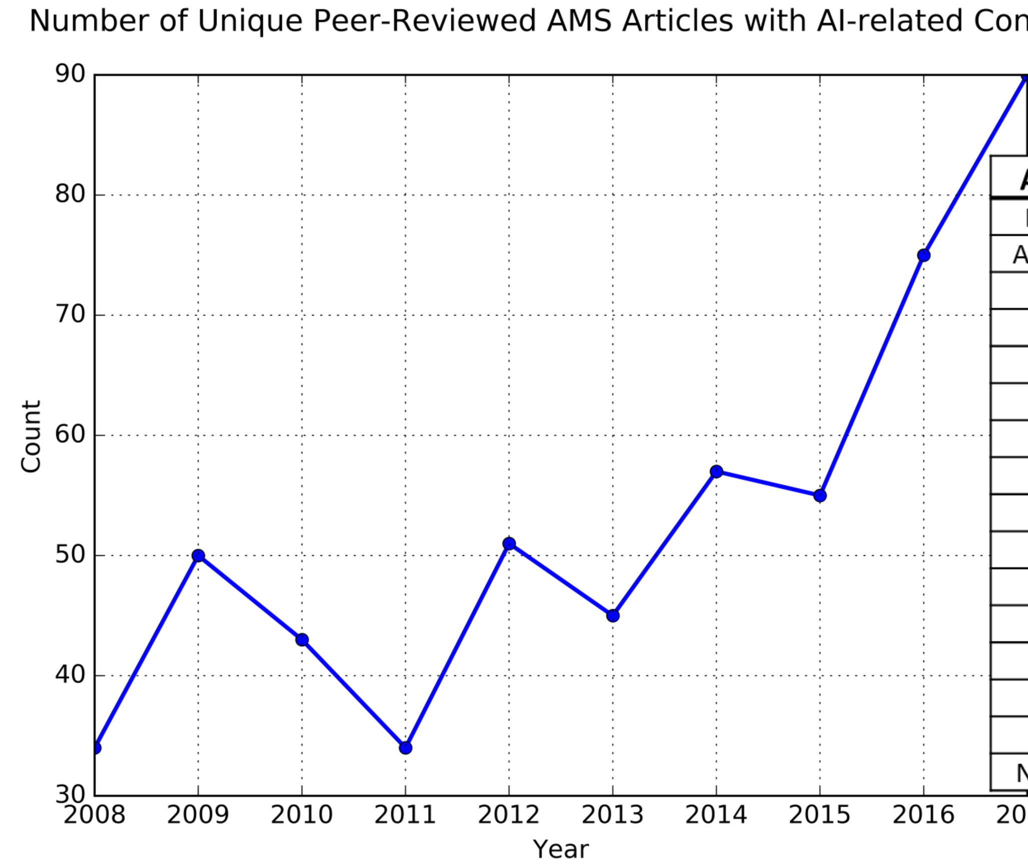
- Environmental community provides huge volume of environmental observations to the user community.
- NOAA gathers > 20 TB of data per day
  - >>20 TB collected when next-gen satellite systems, international partner data, and private industry sensor/IoT data considered
- Fusing multi-source data to leverage the combined "information" contained within is daunting
- Forecasters-in-the-loop lack time to deviate from trusted resources
- Fusion of measurements occurs in data assimilation
- Innumerable forecasting applications exist where multi-source data fusion could yield valuable information

Lag in AI Application to Weather Topics?

- Environmental community has lagged behind other industries in using machine-learning/data-analytics-enabled AI to fully exploit observations.
- Weather traditionally a problem solved with classical physics
  - Continuous innovation in physics modeling
  - Difficult to break from physical constructs
- Legacy approaches difficult to supplant
  - E.g., Legacy code persists older techniques
  - E.g., Forecaster-in-the-loop ingrained in operations
- Past difficulties in implementing innovative approaches
  - E.g., computational resources/software tools
- **Last two years demonstrate huge increase in community effort related to AI applications.**

Surge in Community AI Research/Interest

- Conference proceedings/recent publications suggest **rapid expansion** in community AI capabilities
- 2015/16 → 2017: ~**50%** increase in conference material
- 2015 → 2017: ~**65%** increase in published material.
- Expansion accompanied by notable uptick in mature work



Authors (Last, First)	# Pubs.
Lakshmanan,Valliappa	15
Anagnostou,Emmanouil	12
Sorooshian,Soroosh	12
Cassano,John	11
Zhang,Jian	9
McGovern,Amy	8
Richman,Michael	7
Hong,Yang	7
Wei,Chih-Chiang	7
Aires,Filipe	7
Howard,Kenneth	7
Smith,Travis	7
Haupt,Sue	7
Elmore,Kimberly	6
Gagne II,David	6
Nikolopoulos,Efthymios	6

Figure 1: Surge in peer-reviewed AMS articles containing AI-related content over last decade. Accompanying table lists most active authors in AMS journals in this subject area.

Maturing Applications & Major Contributors

- Increased number of mature AI-based ideas
  - Mature – products are/soon-to-be operational
- **Renewable Energy**
  - NCAR Boulder
  - Often includes application of tree-based methods, but variety employed (e.g., Haupt et al. 2017)
- **Severe Weather**
  - University of Oklahoma/CIMMS
  - Includes application of tree-based or neural-network (NN) based methods (McGovern et al. 2017)
- **Aviation**
  - MIT Lincoln Labs
  - Includes mature applications rooted in deep-learning frameworks (Veillette et al. 2017)

Figure 2: Decision-tree Structure

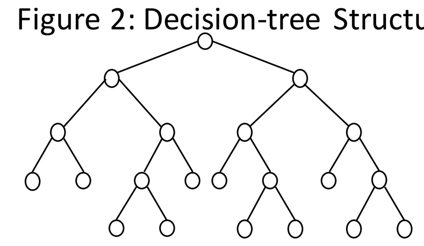


Figure 3: Neural-Network Structure

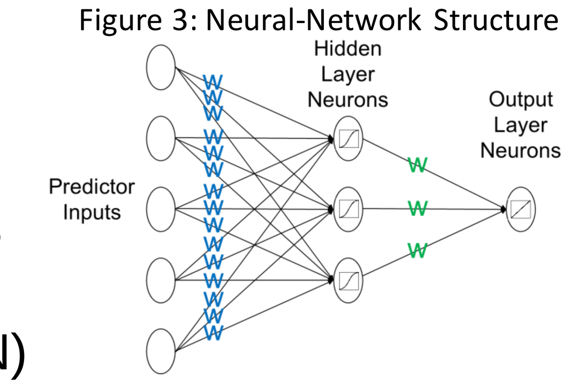
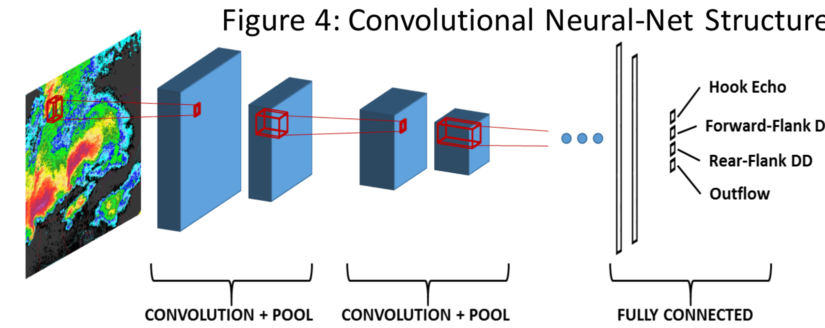


Figure 4: Convolutional Neural-Net Structure



Catalysts for Recent Surge

- Increased demand for hyperlocal forecasts
- Renewable energy forecasting

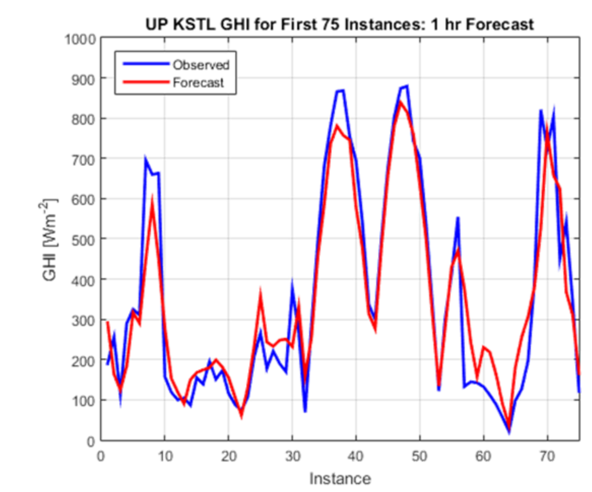


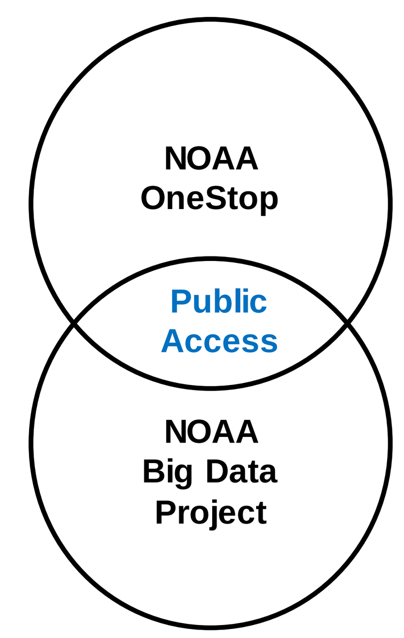
Figure 5: Machine-learning-based fusion of NWP and in-situ obs. better forecast solar irradiance during intermittent cloud cover (Isaacson et al. 2016).

- Smart-phone age
- Enhanced trust in machine-learning
- Higher data volume (e.g., GOES-16, JPSS-1)
- Ease of development

(R-Project 2018)

Catalysts for Continuing Growth

- Enhanced Ease of Data Access



NOAA OneStop Interface

Amazon Web Services

Google Cloud Platform

IBM

Microsoft

Open Commons Consortium

Key Datasets Hosted By Cloud Providers

Earth on AWS
GOES
NEXRAD
UKMET MOGREPS NWP
NASA Earth Exchange
Landsat, MODIS, Sentinel
DigitalGlobe Events (pre/post)
Agriculture Imagery
Land Use (variety of sources)

Google EarthEngine/BigQuery
NOAA GHCN
NOAA GSOD
NOAA ICOADS – surface marine
DMSP
Landsat, MODIS, Sentinel
GFS, CFSv2, OpenAQ
Agriculture Imagery
Land Use (variety of sources)

IBM NESDP
NOAA RAP NWP
NOAA NMFS Datasets

Microsoft Azure
NASA EARTHDATA
WorldcLIM (gridded climate)

OCC Environmental
GOES-16
NEXRAD

Catalysts for Continuing Growth Continued…

- Scalable Cloud Computing; no need for local cluster
  - **Amazon Web Services (AWS) Deep Learning AMI**
    - Generic build of GPU-enabled deep-learning tools (TensorFlow, Theano, Keras, Caffe, etc.)
- Best of both worlds programming languages
  - Julia –Python/R interactive style, speed of C++
  - Python packages already ported (e.g., Scikit-Learn)

Physics is a Key Hurdle for Future Adoption

- Understanding the learned physics of AI-based models
- Tree-based methods demonstrate variable importance
- NN based applications tend to bury physics
- Marzban and Viswanathan (2017) – intuitive example of NN-learned logic gates.

Vectors Prime for Further AI Exploitation

- Multi-scale automatic threat-area recognition
  - Implications for forecaster-on-the-loop transition
- Greater exploitation of lightning data including GOES GLM
- Nonlinear forecast calibration (enhanced MOS)
- Satellite Environmental Data Record (EDR) retrievals
- Data assimilation with machine-learning basis
- Automated data source selection for problem domain
- Data extension to denied regions (i.e., marine environments)
- Extreme heat waves/Long-term droughts & "flash" droughts
- Transportation forecasting with IoT backbone
- Water quality (harmful algal blooms, oil slick detection)
- Seasonal land-falling Atlantic hurricanes
- Sea ice extent/navigability

Useful Links

- AWS Deep Learning AMI: https://aws.amazon.com/amazon-ai/amis/
- Earth on AWS: https://aws.amazon.com/earth/
- Google Big Query: https://cloud.google.com/bigquery/public-data/
- Google Earth Engine: https://earthengine.google.com/datasets/
- IBM NOAA Earth Systems Data Portal (NESDP): https://noaa-crada.mybluemix.net/
- Julia Programming Language: https://julialang.org/
- NOAA OneStop: https://www.ncdc.noaa.gov/onestop/ ; Kearns, E., 2017: Facilitating new opportunities for data users via NOAA's Big Data Project. Accessed 14 November 2017, http://www.nsc2017.org/wp-content/uploads/presentations/NSC2017_Session_4.0_Ed_Kearns.pdf
- OCC Environmental Commons: http://edc.occ-data.org/

References

- Haupt, S., and coauthors, 2017: Building the Sun4Cast system: improvements in solar power forecasting. *BAMS*, Early Release, doi: *10.1175/BAMS-D-16-0221.1*.
- Isaacson, P., G. Bloy, E. Wendoloski, and C. Mutchler, 2016: Short Term Forecasting of Solar Irradiance using Data Analytics and Probabilistic Cloud Cover Forecasts, **96th Annual American Meteorological Society Conference**, New Orleans, LA.
- Marzban, C., and R. Viswanathan, 2017 : On the complexity of neural-network learned functions. *15th Conf. on Artificial and Computational Intelligence and its Applications to Environmental Sciences,* Seattle, WA, Amer. Meteor. Soc., 2.1, https://ams.confex.com/ams/97Annual/webprogram/Paper316115.html .
- McGovern A., and coauthors, 2017: Using Artificial Intelligence to Improve Real-Time Decision Making for High-Impact Weather. *BAMS*, **98**, 2073-2090, doi: *10.1175/BAMS-D-16-0123.1*
- R-Project, 2018: GNU general public license [Available online at https://www.r-project.org/Licenses/GPL-2]
- Veillette, M., E.P. Hassey, C.J. Mattioli, H. Iskenderian, and P.M. Lamey: Using convolutional neural networks to create radar-like precipitation analyses for aviation. *18th Conf. on Aviation, Range, and Aerospace Meteorology,* Seattle, WA, Amer. Meteor. Soc., J2.2, https://ams.confex.com/ams/97Annual/webprogram/Paper304433.html.