U.S. DEPARTMENT OF ENERGY | Office of Science

DOE/SC-ARM-25-017

ARMing the Edge: Demonstration of Edge Computing Field Campaign Report

S Collis
C Catlett
R Jackson
D Dematties
B Raut
P Beckman
N Ferrier
S Park
R Sankaran

July 2025

ARM

DISCLAIMER

ARMing the Edge: Demonstration of Edge Computing Field Campaign Report

S Collis, Argonne National Laboratory (ANL)
P Beckman, Northwestern University (NU)
C Catlett, ANL
N Ferrier, ANL and NU
R Jackson, ANL
S Park, ANL and NU
D Dematties, ANL and NU
R Sankaran, ANL and NU
B Raut, ANL

July 2025

How to cite this document:

Collis, S, P Beckman, C Catlett, N Ferrier, R Jackson, S Park, D Dematties, R Sankaran, and B Raut. ARMing the Edge: Demonstration of Edge Computing Field Campaign Report. U.S. Department of Energy, Atmospheric Radiation Measurement user facility, Richland, Washington. DOE/SC-ARM-25-017.

Work supported by the U.S. Department of Energy,
Office of Science, Office of Biological and Environmental Research

Acronyms and Abbreviations

AI	artificial intelligence
AI4ESP	Artificial Intelligence for Earth System Predictability (DOE)
ANL	Argonne National Laboratory
ARM	Atmospheric Radiation Measurement
ASCR	Advanced Scientific Computing Research (DOE)
BER	Biological and Environmental Research program (DOE)
CMV	converted meteorological visibility
CNN	convolutional neural network
DINO	Distillation No Labels
DL	Doppler lidar
DOE	U.S. Department of Energy
FFT	Fast Fourier Transform
GBT	gradient-boosted trees
ML	machine learning
MSRI	Mid-Scale Research Infrastructure (NSF)
NNSA	National Nuclear Security Administration (DOE)
NSF	National Science Foundation
NU	Northwestern University
OF	Optical Flow
PC	Phase Correlation
RMSPE	root-mean-square percentage error
SGP	Southern Great Plains
SNR	signal-to-noise ratio
WSN	Wild Sage Node

Contents

Figures

1.0 Summary

Edge computing enables "next-to-instrument" control and intelligent data volume reduction and the potential for autonomous, adaptive measurement strategies such as for automated control of scan strategies for Doppler lidar (DL). Instruments with narrow bandwidth connections (e.g., ship and remote sites) can do scene determination and save phenomenon-appropriate data. For example, Doppler spectrum can be saved when clouds are detected by the instrument or automatic moment detection can take place in camera images and only preserve spectrum when non-monomodal spectra are detected. Automated control at the edge involves changing the sampling (temporal or scanning strategy) of an instrument to suit the phenomena both present and being studied (Jackson et al. 2020). Both data processing and instrument control introduces the possibility of *a software-defined instrument*.

Figure 1. Waggle node at the U.S. Department of Energy (DOE) Atmospheric Radiation Measurement (ARM) user facility's Southern Great Plains (SGP) observatory featuring a sky-facing and a ground camera with Stevenson shield, serving as an edge computing node for real-time atmospheric data collection and analysis.

ARMing the Edge was a demonstration project showcasing the potential of edge computing for the DOE Biological and Environmental Research (BER) program. The project was part of a broader initiative, Sage, a national-scale reusable cyberinfrastructure project aimed at enabling artificial intelligence (AI)-driven analysis at the data collection point (edge). Sage is supported by multi-agency sponsors, including DOE Advanced Scientific Computing Research (ASCR), DOE National Nuclear Security Administration (NNSA), DOE BER (through Artificial Intelligence for Earth System Predictability [AI4ESP]), Argonne Laboratory-Directed Research and Development (LDRD), and is funded as a National Science Foundation (NSF) Mid-Scale Research Infrastructure (MSRI) program. The campaign deployed a Wild Sage Node (WSN) and a Sage Blade at ARM's SGP observatory. The project focused on two initial science goals: Clever reduction of scanning Doppler lidar data and detection of clouds and cloud motion via hemispheric camera imagery. Later, the WSN was relocated to assist in vehicular traffic detection and identifying contamination in aerosol measurements.

The project successfully demonstrated the use of edge computing to process Doppler lidar spectra in real time during cloudy and clear-sky periods (Jackson et al. 2023). The campaign also explored cloud motion analysis from a hemispheric sky camera, providing a new way to compute the cloud motion (Raut et al. 2023a). The potential for future research includes expanding the scope of the edge-based AI/machine learning (ML) application.

2.0 Results

The deployment at the SGP observatory yielded promising results across three main areas.

2.1 AI@Edge for Doppler Lidar

Edge computing allowed real-time selection of Doppler lidar spectra for interesting phenomenon, which are typically discarded after moments are calculated. A convolutional neural network working on the edge successfully classified clear and cloudy sky and the Doppler spectra from the lidar were saved for the cloudy cases (Raut et al. 2023b). A machine learning approach using gradient boost techniques provided more accurate classification of spectra, as shown in a confusion matrix comparing different models presented in Jackson et al. (2023).

The data were collected from Doppler lidars at the ARM SGP site and processed using range-corrected signal-to-noise ratios (SNR) and moments. Supervised learning models, namely ResNet50 and gradient-boosted trees (GBT), were used to classify sky conditions. ResNet50 outperformed the GBT model, achieving 97.6% accuracy for clear-sky detection and 94.7% for cloudy skies, while the GBT misclassified cloudy periods as clear. The convolutional neural network (CNN) model processed one hour of Doppler data in under two minutes, making it ideal for real-time use. Additionally, an unsupervised learning approach with a convolutional autoencoder and k-means clustering identified 10 atmospheric clusters, revealing distinct cloud-base heights and conditions. This method improved cloud event detection during rainy periods and outperformed conventional techniques in managing precipitation-contaminated data. The overall approach reduced storage needs, demonstrating the effectiveness of edge computing in atmospheric data management.

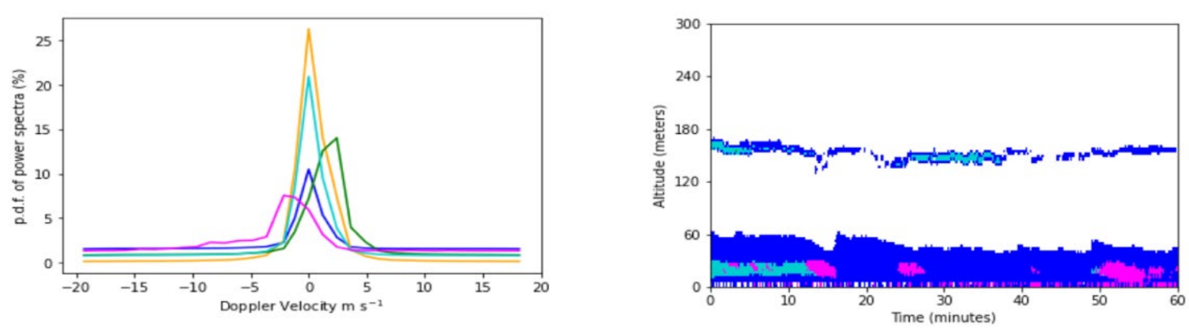

Figure 2. Automated identification at the edge: (left) Five clusters of Doppler spectra, and (right) results for one hour of lidar data.

2.2 Cloud Motion Analysis

The study compared cloud motion estimation methods using data from TSI and Sage cameras, employing both Phase Correlation (PC) and Optical Flow (OF) algorithms. PC used Fast Fourier Transform (FFT) to calculate cloud displacement in image blocks, while OF provided dense pixel-wise motion estimation. Sensitivity tests showed smaller block sizes captured finer motion but introduced noise, while larger blocks and shorter frame intervals improved stability of the motion vectors. Validation with synthetic cloud images and real-time data from the ARM SGP site also demonstrated that doubling the frame rate improved cloud motion estimation: however, higher resolution does not show significant improvement.

The OF method generated denser motion fields but required more computational resources than PC, which performed efficiently, with root-mean-square percentage error (RMSPE) increasing from 22.6% for 5-pixel displacements to 49% for 20-pixel displacements. The PC algorithm also showed significant positive correlation with wind profiles (r=0.42 for U component and r=0.59 for V component after excluding rainy periods), making it effective for tracking stable cloud layers. Raindrop contamination of the images was a significant factor in cloud motion errors with wind profilers and it was successfully flagged using a rotating TSI mirror, preventing erroneous converted meteorological visibility (CMV) calculations. Overall, PC integrated with edge computing proved suitable for real-time cloud motion estimation, outperforming OF in computational efficiency, while OF provided better quality of the motion vectors after removing spurious vectors using the Westerweel and Scarano (2005) method with pyramid scheme.

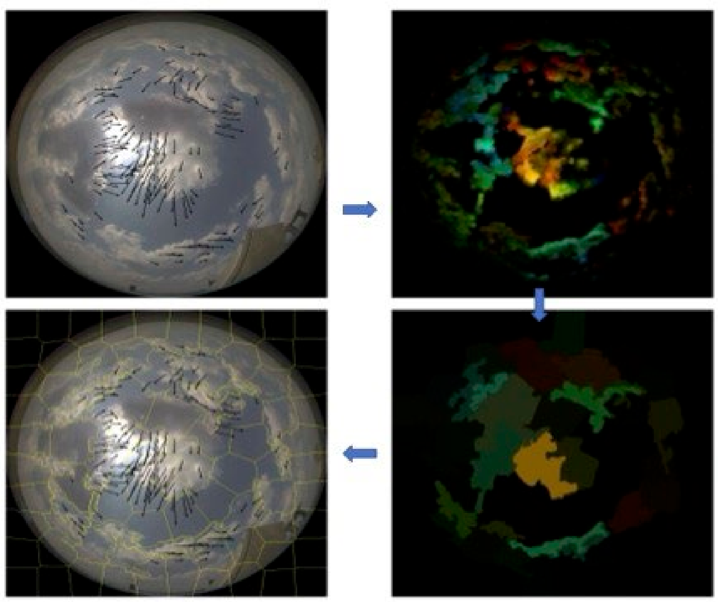

Figure 3. Motion segmentation separates potentially multi-level clouds.

2.3 Self-Supervised AI@Edge

A possibility of self-supervised learning on the edge was tested with camera data by training the Distillation No Labels model (DINO; Caron et al. 2021) for cloud classification and segmentation. We did not run DINO at the edge. However, the model processed ~85,000 full color, 2048x2048 pixel sky images archived by the Sage node at the ARM SGP site. The DINO model employs a joint embedding architecture, eliminating the need for manual labeling by autonomously extracting relevant features for cloud identification. Using both the ARM SGP data set and a 400-image labeled data set (WSISEG-Database), the method demonstrated competitive cloud segmentation and classification performance (Dematties et al. 2023). Self-organizing maps of the feature vectors effectively separated the scene by cloud coverage, diurnal variation, and cloud-base height. The results from Dematties et al. and Jackson et al. are comparable for cloud altitudes over SGP with cross-validation from multiple data sources. The self-supervised learning at the edge can be a robust and scalable solution for atmospheric monitoring and for large-scale cloud data analysis (Raut et al. 2023c).

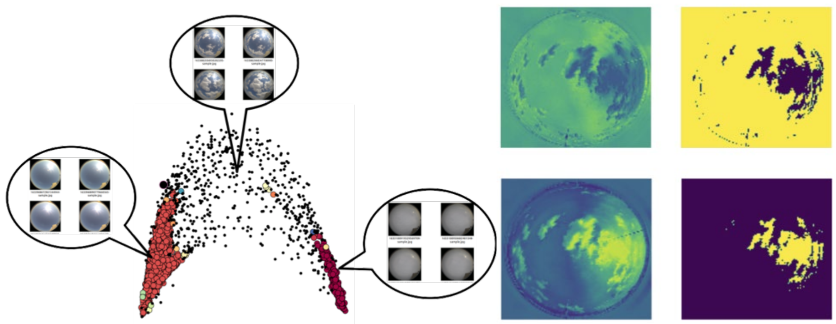

Figure 4. Cloud classification using feature vectors and segmentation using attention map from the DINO model.

2.4 Future Scope

We have tested the camera images on the waggle node with vehicle detection algorithm to tag the quality of the particle measurement at the SGP when the vehicles are detected near the site. The same has been tested offline with the ENO camera data where the planes could be detected, and the data can be tagged. The cloud cover estimation and solar energy estimation using the waggle camera at the SGP site were also used by Park et al. (2021). The project demonstrated the utility of edge computing for real-time data processing for Doppler spectra and cloud motion, suggesting future opportunities in expanding AI-based techniques for broader atmospheric applications.

3.0 Publications and References

Caron, M, H Touvron, I Misra, H Jégou, J Mairal, P Bojanowski, and A Joulin. 2021. "Emerging properties in self-supervised vision transformers." *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)* 9650–9660, https://doi.org/10.48550/arXiv.2104.14294

Dematties, D, BA Raut, S Park, RC Jackson, S Shahkarami, Y Kim, R Sankaran, P Beckman, SM Collis, and N Ferrier. 2023. "Let's Unleash the Network Judgement: A Self-Supervised Approach for Cloud Image Analysis." *Artificial Intelligence for the Earth Systems* 2(2), https://doi.org/10.1175/AIES-D-22-0063.1

Jackson, RC, BA Raut, D Dematties, S Collis, N Ferrier, P Beckman, R Sankaran, Y Kim, S Park, S Shahkarami, and R Newsom. 2023. "ARMing the Edge: Designing Edge Computing-Capable Machine Learning Algorithms for ARM Doppler Lidar Processing." *Artificial Intelligence for the Earth Systems* 2(4): 1–38, https://doi.org/10.1175/AIES-D-22-0062.1

Jackson, RC, S Collis, N Ferrier, P Beckman, and R Sankaran. 2020. "ARMing the Edge: Using Machine Learning at the Edge to Improve Cloud Microphysical Observations." Virtual presentation at the ASR/ARM Topical Workshop on Machine Learning and Statistical Methods for Observations, https://youtu.be/mbYHVn6rK_Y?feature=shared

Park, S, Y Kim, NJ Ferrier, SM Collis, R Sankaran, and PH Beckman. 2021. "Prediction of Solar Irradiance and Photovoltaic Solar Energy Product Based on Cloud Coverage Estimation Using Machine Learning Methods." *Atmosphere* 12(3): 395, https://doi.org/10.3390/atmos12030395

Raut, BA, R Jackson, D Dematties, S Park, J Swantek, S Shahkarami, N Conrad, Y Kim, S Shemyakin, R Sankaran, P Beckman, N Ferrier, and SM Collis. 2022. "Waggle Edge Computing for the Atmospheric Radiation Measurement (ARM) User Facilities: ARMing the Edge." Presentation at the ARM/ASR Principal Investigators Meeting, Rockville, Maryland.

Raut, BA, P Muradyan, R Sankaran, RC Jackson, S Park, SA Shahkarami, D Dematties, Y Kim, J Swantek, N Conrad, W Gerlach, S Shemyakin, P Beckman, NJ Ferrier, and SM Collis. 2023a. "Optimizing Cloud Motion Estimation on the Edge with Phase Correlation and Optical Flow." *Atmospheric Measurement Techniques* 16(5): 1195–1209, https://doi.org/10.5194/amt-16-1195-2023

Raut, B, S Park, D Dematties, S Collis, N Ferrier, Y Kim, S Shahkarami, P Beckman, and R Sankaran. 2023b. Doppler spectra from ARM Doppler Lidar during ARMing the Edge. U.S. Department of Energy, Atmospheric Radiation Measurement. https://doi.org/10.5439/2204486

Raut, B, D Dematties, RC Jackson, S Park, S Shahkarami, Y Kim, R Sankaran, P Beckman, SM Collis, and N Ferrier. 2023c. "A Self-supervised Approach for Cloud Image Analysis." Presentation at the 103rd American Meteorological Society Annual Meeting, 22nd Conference on Artificial Intelligence for Environmental Science.

Westerweel, J, and F Scarano. 2005. "Universal outlier detection for PIV data." *Experiments in Fluids* 39: 1096–1100, https://doi.org/10.1007/s00348-005-0016-6

4.0 Lessons Learned

The key challenge in this campaign was the integration of edge computing with existing ARM instrumentation (Doppler lidar) without disturbing the operational scans. Technical challenges, such as power management and software optimization for field deployment, were notable but SGP team's support made it easy. Several suggestions were made to the waggle team for better data management for accessibility and ML-ready data, paving the way for future field campaigns leveraging edge infrastructure.

ARM

www.arm.gov

U.S. DEPARTMENT OF ENERGY

Office of Science