2019-12

Improving Tactical Environmental Support in Data Denied Areas: Applications of Machine Learning (ML)

Nuss, Wendell A.; Murphree, Tom

Monterey, California: Naval Postgraduate School

http://hdl.handle.net/10945/69975

NPS NRP Executive Summary

Improving Tactical Environmental Support in Data Denied Areas: Applications of Machine Learning

Report Date: 10/14/19 | Project Number: NPS-19-N051-A

Naval Postgraduate School, Graduate School of Engineering and Applied Science

MONTEREY, CALIFORNIA

IMPROVING TACTICAL ENVIRONMENTAL SUPPORT IN DATA DENIED AREAS: APPLICATIONS OF MACHINE LEARNING

Executive Summary Type: Final Report
Period of Performance: 10/14/2018-10/15/2019

Principal Investigator (PI): Prof. Wendell A. Nuss
Additional Researcher(s): Res. Assoc. Prof. James T. Murphree
Student Participation: LCDR Kellen Jones, USN, Meteorology, 2nd LT Jacob Lang, USAF, Meteorology, 2nd LT Darby Maier, USAF, Meteorology, Sooyeon Kim, SEAP Intern, Meteorology, Anjali Golechha SEAP Intern, Meteorology, Paulo Jauregui, Hartnell College Intern, Meteorology

Prepared for:
Topic Sponsor Lead Organization: N2/N6E
Topic Sponsor Organization (if different): N2/N6E
Topic Sponsor Name: Mr. David McCarren, CIV
Topic Sponsor Contact Information: david.h.mccarren@navy.mil, 571-232-4485

EXECUTIVE SUMMARY

Project Summary
Battlespace decisions rely on accurate environmental predictions on both long and short timescales to optimize operational success of missions. Post-processing of numerical model forecasts using a machine learning algorithm has been shown to reduce environmental prediction error. In this study, tests are done to determine the length and type of learning data needed to improve forecasts of low-level marine stratus clouds and near-surface winds. Climate analysis was also done using known climate predictors to select analog periods to define learning data. Results show improvement of short-term cloud and wind forecasts with learning periods of five days or less, making the algorithm applicable to data denied regions.

Keywords: *battlespace environment, Bayesian ensemble model output statistics, BEMOS, Bayesian regression, forecasting, machine learning, model output statistics, MOS, numerical weather prediction, NWP*

Background
Errors in forecasts of environmental conditions can have substantial negative impacts on national security operations. Forecasts of environmental conditions for long-range mission planning and short-range mission execution are generally provided by dynamically based numerical models. The magnitude of numerical model forecast error varies considerably in space and time, which makes them hard to anticipate and correct. This is especially true in data denied areas where there are little, if any, observations to assess error and correct forecasts.

Techniques to mitigate these deficiencies in numerical weather prediction (NWP) model forecasts through post-processing of model data have proven useful. The model output statistics (MOS) approach (Glahn & Lowry, 1972) uses multivariate linear regression techniques to make point forecasts of sensible weather from NWP output. This approach requires relatively large amounts of observations and model forecasts to develop the regression equations. Contemporary post-processing research by Gneiting (2014) implemented a univariate non-homogeneous regression approach to ensemble model output statistics. Richter (2012) added a Bayesian approach to ensemble model output statistics for a single variable. Wendt (2017) extended this Bayesian approach to a full multivariate approach using a hierarchical parameter structure and showed that even with limited learning data, this approach can significantly reduce error.

Critical to the application of these post-processing approaches is a sufficient set of learning data consisting of both observations and model forecasts. The performance of the Naval Postgraduate School's Bayesian ensemble model output statistics (BEMOS) post-processing showed nearly equal skill for 30-day and 1-year learning periods (Wendt, 2017), especially when similar weather patterns were used in the training. To optimize model learning to extend this application to data denied regions, longer-range, large-scale variations due to climate conditions may help identify analog periods to use as learning data in the

BEMOS post-processing. The success of this approach depends on being able to identify appropriate climate patterns in advance. The BEMOS model can also be applied to this climate prediction problem to identify particular large-scale weather patterns.

The objectives of this study are to test the application of the BEMOS post-processing model to various short-range forecast problems as well as to determine its applicability to identify climate patterns in past data. Specifically, the impact of time and spatially limited learning data on forecast error is examined. Furthermore, the ability to identify and predict high-impact climate patterns using the BEMOS model is evaluated.

Findings and Conclusions

This study examines the utility of using a machine learning technique to post-process numerical predictions of environmental variables to reduce error as well as identify impactful climate regimes to improve environmental support. We have conducted a series of tests to determine how well this approach can be applied in data denied areas that are of operational interest. Specifically, short-range low-level cloud and surface wind forecasts as well as long-range climate pattern forecasts have been examined.

To test the application of the BEMOS machine learning to long-range, large-scale prediction of high-impact weather, the problem of Santa Ana offshore winds was used. A characteristic long-wave pattern that is associated with Santa Ana wind events was identified using climate analyses. Results using the Madden-Julian oscillation phase as a predictor showed strong predictive skill for the Santa Ana long-wave pattern that extended to as much as 21–28 days of lead time. These results indicate that the machine learning algorithm can potentially select climate patterns to use in training short-term forecasts within the particular climate pattern.

Direct model forecasts of both clouds and surface winds exhibit considerable errors that have large operational impacts when incorrectly forecast. Numerical model forecasts of marine stratus cloud forecasts often fail to depict actual cloud cover even for time periods as short as 3–6 hours. The machine learning algorithm was applied to 1 km resolution model forecasts using cloud water, relative humidity, surface elevation, and vertical motion as predictors. These predictors were then trained against 1 km visible imagery as verification for training periods of one to three days. Results showed that forecast details about the location and timing of the clearing of existing clouds were generally improved through the machine learning. However, the machine learning failed to adequately generate clouds when conditions were already clear and the predictor variables contained no clouds.

Numerical model wind forecast error is often rather large in coastal regions where complex topography and coastline geometry impact the underlying dynamics. To correct these forecasts, the statistical model was run using three predictors (wind speed, direction, and surface pressure) for the Monterey Bay region and calibrated using surface observations. Learning periods ranging from three hours to seven days were examined. In addition, data from the same time each day were compared to data from multiple times per day to use in the learning. Results showed that five days using the same time of day produced the best forecast improvement. Wind speed errors generally decreased by 0.5–1.5 m/s with the largest reduction occurring for the higher wind times of day, which represents a 15–25% improvement in the absolute wind speed forecast.

Recommendations for Further Research

The results from the study indicate that improvements in cloud and wind speed forecasts can be achieved by applying a machine learning approach to correct for typical forecast errors. This study was limited to tests in the Monterey Bay region for a small sample of forecasts. Determining the forecast improvement over a broad range of weather forecast events should be examined. In addition, coupling the climate prediction of analogous weather patterns to define similar learning data periods needs to be tested and compared to other learning periods such as most recent 5–7 days. Finally, the probability distribution produced by the BEMOS algorithm should be extracted to help decision makers utilize the forecasts most effectively.

References

Glahn, H.R., & Lowry, D.A. (1972). The use of model output statistics (MOS) in objective weather forecasting. *Journal of Applied Meteorology, **11**,* 1203–1211.

Gneiting, T. (2014). Calibration of medium-range weather forecasts, *European Centre for Medium-Range Weather Forecasts (ECMWF) Technical Memorandum 719*. Retrieved from https://www.ecmwf.int/sites/default/files/elibrary/2014/9607-calibration-medium-range-weather-forecasts.pdf.

Richter, D. (2012). *Bayesian ensemble model output statistics for temperature* (Diploma thesis). Heidelberg University, Heidelberg, Germany.

Wendt, R.D.T. (2017). *A hierarchical mulitvariate Bayesian approach to ensemble model output statistics in atmospheric prediction* (Doctoral dissertation). Naval Postgraduate School, Monterey, CA. Retrieved from http://hdl.handle.net/10945/56188

Acronyms

Bayesian ensemble model output statistics	BEMOS
Model output statistics	MOS
Numerical weather prediction	NWP