MMM/ RAL/ CISL SEMINAR

Mesoscale & Microscale Meteorology MMM LABORATORY

Coupling Data Science and Numerical Weather Prediction for High-Impact Weather Forecasting

David John Gagne

NCAR/ASP

Meteorologists have access to more model guidance and observations than ever before, but this additional information does guarantee better forecasts. New tools are needed to reduce the cognitive load on forecasters and to provide them with accurate, reliable consensus guidance. Techniques from the data science community, such as machine learning and image processing, can summarize and calibrate numerical weather prediction model output and to generate deterministic and probabilistic forecasts of high-impact weather. I developed data-science-based approaches to improve the predictions of two high-impact weather domains: hail and solar irradiance. Both hail and solar irradiance produce large economic impacts, have non-Gaussian distributions of occurrence, are poorly observed, and are partially driven by processes too small to be resolved by numerical weather prediction models. Hail forecasts were produced with convection-allowing model output from the Center for Analysis and Prediction of Storms and National Center for Atmospheric Research ensembles. The machine learning hail forecasts were compared against storm surrogate variables and physics-based diagnostic models of hail size. By coupling the machine learning model to predicting hail size distributions and estimating the distribution parameters jointly, the machine learning methods were able to show skill and reliability in predicting both severe and significant hail. Machine learning model and data configurations for gridded solar irradiance forecasting were evaluated on two numerical modeling systems. The evaluation determined how machine learning model choice, closeness of fit to training data, training data aggregation, and interpolation method affected forecasts of clearness index at unseen Oklahoma Mesonet sites. The choice of machine learning model, interpolation scheme, and loss function had the biggest impacts on performance. Errors tended to be lower at testing sites with sunnier weather and those that were closer to training sites. All of the machine learning methods produced reliable predictions but underestimated the frequency of cloudiness compared to observations.

This seminar will be webcast live at:
http://www.fin.ucar.edu/it/mms/fl-live.htm

Recorded seminar link can be viewed here:
https://www.mmm.ucar.edu/events/seminars

Thursday, 13 October 2016, 3:30 PM
Refreshments 3:15 PM
NCAR-Foothills Laboratory
3450 Mitchell Lane
Bldg 2, Main Auditorium, Room 1022

NCAR

NSF