Geophysical Research Abstracts
Vol. 20, EGU2018-23, 2018
EGU General Assembly 2018

A Machine-learning Ensemble Approach to Improve Wave-condition Forecasts

Scott James (1), Fearghal O'Donncha (2), and Yushan Zhang (3)

(1) Baylor University, The Institute of Ecological, Earth & Environmental Science, Geosciences and Mechanical Engineering, Waco, United States (sc_james@baylor.edu), (2) IBM Research, Dublin, Ireland (feardonn@ie.ibm.com), (3) University of Notre Dame, Department of Chemical and Biomolecular Engineering, Notre Dame, United States (yzhang33@nd.edu)

Modern smart-grid networks use technologies to instantly relay information on supply and demand to support effective decision making. Integration of renewable-energy resources with these systems demands accurate forecasting of energy production (and demand) capacities. For wave-energy converters, this requires wave-condition forecasting to enable estimates of energy production. Accurate forecasting of wave conditions is a challenging undertaking that typically involves solving the spectral action-balance equation on a discretized grid with high spatial resolution. The nature of the computations typically demands high-performance computing infrastructure. Using a case-study site at Monterey Bay, California, a machine learning framework was trained to replicate numerically simulated wave conditions at a fraction of the typical computational cost. Specifically, the physics-based Simulating WAves Nearshore (SWAN) model, driven by measured wave conditions, nowcast ocean currents, and wind data, was used to generate training data for machine learning algorithms. The model was run between April 1st, 2013 and May 31st, 2017 generating forecasts at three-hour intervals yielding 11,078 distinct model outputs. SWAN-generated fields of 3,104 wave heights and a characteristic period could be replicated through extremely fast (1,200 times faster) simple matrix multiplications using the mapping matrices from machine learning algorithms. These machine learning models can be included in statistical learning aggregation techniques that combine forecasts from multiple, independent models into a single "best-estimate" prediction of the true state. Ensembles are developed based on multiple simulations perturbing input data (wave characteristics supplied at the model boundaries and winds). A learning-aggregation technique uses past observations and past model forecasts to calculate a weight for each model. The aggregated forecasts are compared to observation data to quantify the performance of the model ensemble and aggregation techniques. The appropriately weighted ensemble model outperforms an individual ensemble member with regard to forecasting wave conditions. The low computational cost allows for a comprehensive investigation of the inherent uncertainty of inputs (winds, ocean currents, and wave conditions) through an ensemble based approach (i.e., running the model hundreds or thousands of times with perturbed inputs). The aggregating procedure applies appropriate weighting to each ensemble element based on both past performance and real-time sensor data to improve model accuracy. By integrating ensemble forecasting with an extremely lightweight forecasting model and a statistical learning aggregation technique, this study provides a complete real-time forecasting framework that leverages all available data from observations while accounting for uncertainties in model forcing data.