Poster #9-16

Advancing Predictive Understanding of Terrestrial Ecosystem through Machine Learning

Dan Lu[1*] and Daniel Ricciuto[1]

[1] Oak Ridge National Laboratory, Oak Ridge, TN

Contact: lud1@ornl.gov

BER Program: TES
Project: ORNL Terrestrial Ecosystem Science Scientific Focus Area (TES SFA)
Project Website: https://tes-sfa.ornl.gov

Improving predictive understanding of terrestrial ecosystem variability and change requires data- model integration. Efficient data-model integration for complex models requires surrogate modeling to reduce model evaluation time. However, building a surrogate of a large-scale terrestrial ecosystem model (TEM) with many output variables is computationally intensive because it involves a large number of expensive TEM simulations. In this effort, we propose an efficient surrogate method capable of using a few TEM runs to build an accurate and fast-to- evaluate surrogate system of model outputs over large spatial and temporal domains. We first use singular value decomposition to reduce the output dimensions, and then use Bayesian optimization techniques to generate an accurate neural network surrogate model based on limited TEM simulation samples. Our machine learning based surrogate methods can build and evaluate a large surrogate system of many variables quickly. Thus, whenever the quantities of interest change such as a different objective function, a new site, and a longer simulation time, we can simply extract the information of interest from the surrogate system without rebuilding new surrogates, which significantly saves computational efforts. We apply the proposed method to a regional TEM to approximate the relationship between 8 model parameters and 42660 carbon flux outputs. Results indicate that using only 20 model simulations, we can build an accurate surrogate system of the 42660 variables, where the consistency between surrogate prediction and actual model simulation is 0.93 and the mean squared error is 0.02. This highly-accurate and fast-to-evaluate surrogate system will greatly enhance computational efficiency in data-model integration to improve predictions and advance our understanding of terrestrial ecosystems.

Publication

Lu, D. and Ricciuto, D.: Efficient surrogate modeling methods for large-scale Earth system models based on machine learning techniques, Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-327, in review, 2019.