THE EFFECT OF TRAINING METHODOLOGY ON MACHINE LEARNING MODELS FOR ESTIMATING NUTRIENT CONCENTRATIONS

***J. Barrett Carter**, Eban Bean, and Aditya Singh*
Department of Agricultural and Biological Engineering, University of Florida, Gainesville, FL, USA

Machine learning (ML) techniques have the potential to drastically improve the study and management of water resources. However, the development of an effective and reliable ML model requires adequate training and evaluation steps, and while the use of machine learning techniques for multivariate regression has increased drastically in recent years, there is no standard procedure for training and evaluating ML models. This places a limitation on the advancement and application of ML techniques in practice. Given that there are multiple steps in the training and evaluation of a ML model, which can include a combination of randomized resampling methods such as K-fold cross validation, bootstrap, and jackknife, a slight modification to the organization of these steps could have significant implications on the final assessment of the model and its comparison to other models. This hypothesis was tested by implementing three training and evaluation architectures to produce ML models for estimating stream nitrate and phosphate concentrations from UV-Visible absorbance spectra collected from 171 stream samples. Model performances resulting from different architectures were then compared statistically and graphically to assess the impacts of training methodology on model performance. The results from this study may be used to guide researchers towards a path of more transparent and standard procedures for the development of ML models, which could ultimately help to improve the reproducibility and reliability of such techniques when used in practice such as in water resource management applications.

PRESENTER BIO: Barrett Carter is a third-year Ph.D. student specializing in estimating water quality parameters using UV-Visible spectroscopy. Barrett earned Bachelor's and Master's degrees in Biological Engineering from the University of Arkansas and has a total of six years of experience working as a research assistant at the undergraduate and graduate levels.