## Building Mechanistically Sound Machine Learning Models for Environmental Applications

Prof. Huichun (Judy) Zhang Frank H. Neff Professor in Civil & Environmental Engineering Case Western Reserve University, Cleveland OH

**Abstract** - Environmental-related chemical reactions have been extensively studied, but accurately modeling the kinetics of chemical reactions under different conditions remains highly challenging. Existing research often relies on traditional Quantitative Structure-Activity Relationships (QSARs), but these modeling approaches typically require complex feature engineering. In recent years, machine learning has increasingly become a highly promising tool for modeling chemical reactions because it not only achieves better results than traditional methods but also can utilize diverse chemical reference to model different chemical reactions. Here, we present several recent examples to demonstrate how machine learning models based on small datasets through careful data preprocessing and extensive feature engineering, aiming to capture changes in chemical reactivity by considering as many important features as possible, even in the case of a small dataset. Examples include aerobic biodegradation, advanced oxidation processes, as well as non-abiotic reduction and adsorption of organic and inorganic compounds. We will also discuss recent predictive models for the extent of harmful algal blooms in Lake Erie, one of the Great Lakes.