Computational Bayesian Framework for Quantification of Predictive Uncertainty in Environmental Modeling

DOE Early Career Award DE-SC0008272
Subsurface Biogeochemical Research Program

PI: Ming Ye (mye@fsu.edu)
Department of Scientific Computing, Florida State University, Tallahassee, FL 32306

A computational Bayesian framework has been developed for quantification and reduction of uncertainty in environmental modeling. In this year of the project, the framework is used for both theoretical studies and practical applications.

The theoretical studies are focused on defining new metrics for two sensitivity analyses. The first one is to avoid biased selection of important parameters, which are important only to a single model but not the system of interest if the single model does not represent the system. This is done by defining new sensitivity indices after expanding the variance-based sensitivity analysis of a single model and a single scenario to the framework of multiple models and multiple scenarios developed in this project. To reduce the computational cost for calculating the new sensitivity indices, the sparse grid stochastic collocation method is used to evaluate the statistical moments needed for the indices. The use of the new sensitivity index is demonstrated in a numerical example of groundwater flow reactive transport modeling, which also demonstrates the dramatic reduction of computational cost. The other sensitivity analysis uses the same framework but to select important processes of an environmental model. New process sensitivity indices are defined for the situation that a process can be represented by more than one model. The new indices answer the question that how to select the important processes when it is uncertain how to model the processes. This new method of sensitivity analysis should be of particular use at the early stage of model development when the process model uncertainty is the largest.

Computational advance has been made for evaluating Bayesian evidence using the sparse grid methods. The Bayesian evidence is critical for evaluating relative plausibility of multiple models; a more plausible model has larger Bayesian evidence. Evaluating it however is computational challenging, because it requires calculating a multivariate integration. This is done in this project using the thermodynamic integration and stepping stone methods. To reduce the computational cost, the sparse grid methods are used to build cheap-to-evaluate surrogates of computationally demanding models. The Bayesian evidence is evaluated for several reactive transport models for identifying the right level of model complexity.

By collaborating with scientists at the Pacific Northwest National Laboratory, we have applied our new methods of sensitivity analysis for the Hanford 300 Area to select important parameters. Collaboration with scientists at the Oak Ridge National Laboratory is on-going for identifying important processes of climate change modeling. The collaboration of mutual interests help improve the methodology development and gain insights for practical problems.