Multi-Level Dual-Annealing Stochastic Approximation Monte Carlo Algorithm to Optimal Parameter and Uncertainty Estimation for Climate Model Prediction

Funded Project: Stochastic nonlinear data-reduction methods with detection & prediction of critical rare events

PI: Guang Lin[1]; Joint Work with Yichen Cheng[2], Ben Yang[1], Yun Qian[1], Ruby Leung[1], Faming Liang[2]
[1]Pacific Northwest National Laboratory, Richland, WA; [2]Texas A&M University, College Station, TX

Uncertainty quantification has become an important topic for climate modeling, for which we may need to choose the best set of parameters based on several sets of observations. In this study, we apply Bayesian approach and change the problem of getting estimation of parameters to the problem of maximizing the posterior probability density (PPD) function. To find the modes, a series of methods have been proposed, however, they all require a relatively long series of samples. This may cause some practical problem for computational expensive climate model, because usually for one evaluation of PPD for climate problem, it will take many hours in parallel computer clusters. This computational challenge requires us to come up with a new method which can help us to find reasonably good results within a small number of iterations.

In this study, we proposed a new method named *Multi-Level Dual-Annealing Stochastic Approximation Monte Carlo* (ML-DA-SAMC) Algorithm for global optimization which is based on *very fast simulated annealing* (VFSA) with the element of *stochastic approximation Monte Carlo* (SAMC) blended in. This new algorithm employs both space annealing and temperature annealing techniques to speed up the optimization process. To further improve the computation efficiency, we first perform DA-SAMC on a coarse spacing grid. Then applying the optimal parameter PPD calculated from the coarse spacing grid as prior and perform DA-SAMC on a fine spacing grid to speed up the convergence and obtain the optimal parameter PPD on fine spacing grid with minimum computational cost. We demonstrate that using our new ML-DA-SAMC method, we will get the optimized solution more efficiently as compared to VFSA. Also, a new method to calculate the posterior distribution for the parameters is developed. And we show that it is asymptotically unbiased and performed much better than *multiple very fast simulated annealing method* (MVFSA) even when the number of samples is very small. We demonstrate the capability and efficiency of the new developed method by comparing the results obtained from the ML-DA-SAMC algorithm with MVFSA and VFSA methods through a heat convection problem, and an optimal parameter estimation problem in the Kain-Frisch Convective Parameterization Scheme in the Weather Research and Forecasting model.