Geophysical Research Abstracts
Vol. 16, EGU2014-261, 2014
EGU General Assembly 2014

Iterative regularization for ensemble-based Bayesian data assimilation in the geosciences

Marco Iglesias
University of Nottingham, Nottingham, United Kingdom (marco.iglesias@nottingham.ac.uk)

The solution to a Bayesian inverse problem is the conditional probability (posterior) distribution of uncertain properties/parameters of the mathematical model given observational data. The Bayesian posterior that results from large-scale models typical in geophysical applications is defined on a high-dimensional space (e.g. 10^5-10^7 dimensions). Moreover, the posterior does not have a standard form that can be characterized with a few parameters. In general, sampling can be used to accurately approximate the posterior. However, millions of samples are often needed to fully resolve the posterior. With a Markov Chain Monte Carlo (MCMC) approach, for example, the computation of each sample of the posterior involves one forward model solve. Therefore, the computation of millions of MCMC steps to fully resolve the posterior is prohibitive for large-scale applications.

Due to the limitation to fully resolve the posterior, practical data assimilation often uses ad hoc approximations. These approximations of the posterior are exact when the prior is Gaussian and the model is linear (i.e. when the posterior is Gaussian). Examples of these approximations are most of the ensemble methods (e.g. EnKF) that have been developed in the last decades. While these approaches can be proven to sample the posterior in the aforementioned linear-Gaussian case, in general, their approximation properties are still open problems. Nevertheless, due to the urgency of developing fast and efficient data assimilation techniques, ensemble methods remain as one of the main focus of data assimilation community.

In this talk I discuss the computational aspects of two novel iterative ensemble methods inspired by ideas from iterative regularization techniques. The aim of these methods is to capture/approximate the Bayesian posterior. The first method is a variational ensemble approach that uses the regularizing LM scheme for the minimization of a set of randomized least-squares functionals. The second method is an iterative ensemble Kalman scheme. Both methods use the discrepancy principle for (i) the adaptive selection of regularization parameters and (ii) the early termination of the scheme. The main focus of our work is to show the capability of these ensemble methods for capturing/approximating the posterior that arises from Bayesian inverse problems in reservoir modeling.

One of the novel aspects of the present work is the construction of moderate size data assimilation benchmarks for which the posterior distribution is fully resolved with a state-of-the-art MCMC method. While similar benchmarks have been previously used for assessing ad hoc methods, these are based on 1D small problems. In contrast, our benchmarks are developed on 2D models with dimensions of two of order of magnitudes larger than the former. Larger size benchmarks reveal the limitations of standard methods for which the lack of proper regularization is detrimental to the ability to capture the posterior. We display numerical experiments to show that the proposed regularizing schemes outperform the accuracy of the standard methods.