8th. World Congress on Computational Mechanics (WCCM8)
5th European Congress on Computational Methods in Applied Sciences and Engineeering (ECCOMAS 2008)
June 30 –July 5, 2008
Venice, Italy

A multiscale, non-parametric, Bayesian framework for identification of model parameters.

*** P.S. Koutsourelakis**[1]

[1] School of Civil and Environmental Engineering, Cornell University, 369 Hollister Hall, Ithaca, NY 14853, E-mail:pk285@cornell.edu, URL:*http://www.cee.cornell.edu/pk285*

Key Words: *Bayesian, Sequential Monte Carlo, Trans-dimensional Markov chain Monte Carlo*

ABSTRACT

The prodigious advances in computational modeling of physical processes and the development of highly non-linear, multiscale and multiphysics models poses several challenges in parameter identification. We are frequently using large, forward models which imply a significant computational burden, in order to analyze complex phenomena. The accuracy of the results provided depends strongly on assigning proper values to the various model parameters. In solid mechanics for example laboratory testing is performed using fairly large specimens which however, do not immediately provide information about all the parameters in the material model (e.g. elastic modulus, yield stress etc) at the scale of interest, particularly in heterogeneous materials. The present paper is concerned with the problem of identifying parametric variability at various length scales from disparate measurements and experimental observations using computational models.

Traditionally, two basic approaches have been followed in addressing these problems. One one hand, deterministic optimization techniques which attempt to minimize the mean square error of model predictions and observations. Usually the objective function is augmented with regularization terms which alleviate issues with the ill-posednesss of the problem. One the other hand, statistical approaches based on the Bayesian paradigm have been employed which attempt to calculate a (posterior) probability distribution function on the parameters of interest. The latter approach offers several advantages as it provides a unified framework for dealing with the uncertainty introduced by the incomplete and noisy measurements, but most importantly it can be readily coupled with any black-box, forward solver.

In traditional Bayesian formulations, the representation of the unknown field has coincided with the forward model's. In cases where finite elements are used as forward solvers, the property of interest is assumed constant within each element and therefore the vector of unknowns is of dimension equal to the number of elements. This offers obvious implementation advantages but also poses some difficulties since the scale of resolution of the unknown field is implicitly selected by the solver and is essentially equal to the size of the elements. This is problematic in two ways. On one hand if the scale of variability is larger than the grid, a waste of resources takes place, at the solver level which has to be run at unnecessarily fine resolutions and at the level of the inference process which is impeded by the unnecessarily large dimension of the unknown vector. On the other hand, if the scale of variability is smaller than the grid, it cannot be identified even if the solver provides sufficient information for discovering this

possibility. More importantly perhaps, prior information on the scale is not prescribed by the prior, as it is done with other aspects of the model, but by the forward model.

Another significant hindrance with the application of Bayesian models is the significant computational cost. Even though the output of a Bayesian model is not just a single value for the parameters of interest but rather a distribution which compares the likelihood of all possible hypotheses, the computational effort implied by repeated calls to the forward solver can be enormous and constitute the method impractical for realistic applications. This is amplified if the posterior distribution is multi-modal i.e. several significantly different scenaria are likely. It is apparent that it would be desirable or even necessary to use forward models , e.g. models that operate at coarser resolutions, that are faster in order to make inferences. This transition of information is not straightforward if the representation is tied to the solver. Attempts have been made using parallel tempering (e.g. [2]) or hierarchical representations based on Markov trees ([4]) which however require performing inference on representations at various resolutions simultaneously,

In the present paper we adopt a nonparametric model which is independent of the grid of the forward solver and is reminiscent of non-parametric kernel regression methods. The unknown parametric field is approximated by a superposition of kernel-type functions centered at various locations. The cardinality of the representation, i.e. the number of such kernels, is treated as an unknown to be inferred in the Bayesian formulation. This gives rise to a very flexible model that is able to adapt to the problem and the data at hand and find succinct representations of the parametric field of interest. Prior information on the scale of variability can be directly introduced in the model.

Inference is performed using Sequential Monte Carlo samplers ([3]). These represent a set of flexible simulation-based methods for sampling from a sequence of probability distributions; each distribution being known up to a normalizing constant. These methods were originally introduced in the early 1950s by physicists and have become very popular over the past few years in statistics and related fields. They utilize a large set of random samples, named particles, which are propagated using simple importance sampling, resampling and updating/rejuvenation mechanisms. In our case the rejuvenation of the population is performed using Reversible Jump MCMC moves in order to explore configurations that lie in spaces of variable dimension ([1]). These moves include update moves to refine the kernel locations/amplitudes/spreads, but also birth and split moves to add new kernels and death and merge moves to eliminate erroneous or redundant kernels. The algorithm is directly parallelizable as the evolution of each particle is by-and-large independent of the rest. The sequence of distributions defined is based on using solvers that operate on different scales and which successively produce finer discretizations. This results in a hierarchical approach that makes use of the results from solvers operating at the coarser scales in order to update them based on analyses on a finer scale. Several examples on linear and nonlinear problems are presented.

REFERENCES

[1] P. Green. “Reversible jump Markov chain Monte Carlo computation and Bayesian model determination”. *Biometrika*, 82, 711-732, 1995

[2] C. H. Holloman, H. Lee, D.M. Higdon. “ Multi-resolution Genetic Algorithms and Markov Chain Monte Carlo”. *Journal of Computational and Graphical Statistics*, pp. 861-879, 2006.

[3] P. Del Moral, A. Doucet, A. Jasra. “Sequential Monte Carlo for Bayesian Computation, (with discussion)”. *Bayesian Statistics* 8, Oxford University Press, 2006.

[4] J. Wang, N. Zabaras. “Permeability estimation using a hierarchical Markov tree (HMT) model”, Seventh World Congress on Computational Mechanics, Los Angeles, California, USA, July 16-22, 2006.