

Experiments in Auto-Calibration of Large Scale Environmental Models using a Data Mining Approach

David A. Swayne, Markiyana Sloboda, Simon Hood, Vimal Sharma, William Booty, David Lam
Computing Research Laboratory for the Environment
*Department of Computing and Information Science, University of Guelph,
Guelph, Ontario, Canada*

Abstract

Many artificial intelligence paradigms are good candidates for the very difficult problem of model autocalibration. Quasi-Newton methods, optimization techniques and genetic algorithms are typical candidates for current algorithm research for this problem. A less well-travelled road to discovery consists of techniques for knowledge acquisition and data mining techniques for autocalibration. This problem is largely intractable except by “expert” guessing, and the results are usually unsatisfactory.

Our approach has been, from the environmental side, to use well-known and understood models, rather than develop our own. Many of these models involve large amounts of computer time for a single instance. We have partially overcome the computer time issue through the use of virtual files and running the codes on a clustered supercomputer architecture.

These two preliminary steps have left us with a very typical knowledge discovery problem: several thousand model runs from which we hope to abstract the following: “good” solutions, knowledge of the trends toward these good solutions, and possibly a transformation of the model into (for instance) a Bayesian network approximation of the original model. This would allow us to “throw away” the computational model, instead selecting an output set which is a good approximation to what the model would produce under ideal circumstances.

This research was not well-timed for a paper presentation at iEMSs in Barcelona, but our project has produced some preliminary successes, from which we hope to provoke discussion at this workshop.