EGU21-8381
https://doi.org/10.5194/egusphere-egu21-8381
EGU General Assembly 2021

The Method of Forced Probabilities: a Computation Trick for Bayesian Model Evidence

Peter Walter[1], Ishani Banerjee[2], Anneli Guthke[2], Kevin Mumford[3], and Wolfgang Nowak[2]
[1]Simtech/LS3, University of Stuttgart, Stuttgart, Pfaffenwaldring 5a, 70569, Germany
[2]Institute for Modelling Hydraulic and Environmental Systems (IWS)/LS3, University of Stuttgart, Stuttgart, Pfaffenwaldring 5a, 70569, Germany
[3]Department of Civil Engineering, Queen's University, Kingston, Ellis Hall-206, ON K7L3N6, Canada

Bayesian model selection (BMS) can be used to objectively rank competing models of different structure and with different parameters upon comparison with validation data sets. This technique requires the evaluation of Bayesian Model Evidence (BME). BME is the likelihood of the data to occur under the assumed models, where the likelihood is averaged over the probability distribution of the model and its parameters.

Exact and fast analytical solutions for BME exist only with strong assumptions. For that reason, other techniques and approximations for BMS/BME have been developed. While mathematical approximations via information criteria may suffer from strong biases in real-world applications, numerical methods do not rely on any assumptions but require high computational effort. This becomes prohibitive if the data set is very large, e.g. highly resolved in space and time.

To still enable the use of BME as a probabilistic and rigorous model performance metric, we have developed the “Method of Forced Probabilities”: this method is a fast way to numerically compute BME for models that predict time series and fulfill the Markov Chain property in time. The core idea is to swap the direction of evaluation: instead of comparing thousands of forward runs of the model with the observed data (many model runs on random parameter realizations), we force the model to follow the data during each time step and record the individual probabilities of the model performing these exact transitions (single evaluation).

As a test case for demonstration, we use invasion percolation (IP) models to simulate multiphase flow in porous media. The underlying, highly resolved data set was obtained from an experiment of a slow gas injection into water-saturated, homogeneous sand in a 25cmx25cm acrylic glass cell. Images were obtained at a rate of 30 images per second using the light transmission technique. Since IP models fulfill the Markov chain property, the Method of Forced Probabilities can be applied to evaluate their BME. Results confirm that the proposed method presents a scalable, inexpensive alternative to standard Monte Carlo methods for analyzing the model-data mismatch.