Model selection under computational time constraints: application to river engineering

Sergey Oladyshkin[1], Anneli Guthke[1], Farid Mohammadi[1], Rebekka Kopmann[2], Wolfgang Nowak[1]

[1] Department of Stochastic Simulation and Safety Research for Hydrosystems (IWS/SC SimTech), University of Stuttgart, Germany
[2] Federal Waterways Engineering and Research Institute (BAW), Karlsruhe, Germany

Key words: Hydro-morphodynamics; Bayesian model selection; Reduced model; Polynomial chaos expansion

Introduction

A suitable representation of hydro-morphodynamic processes is crucial for successful river management. To predict sediment transport and river bed evolution, a variety of empirical formulas can be plugged into numerical river models and modelers lack objective guidance of how to select the most appropriate one for a specific application. Such guidance can be provided by the Bayesian model selection (BMS) framework, which rests on Bayesian probability theory. Due to its flexibility and handling of uncertainty in measurements and parameters, this framework has been frequently applied in many fields of water resources research and practice. Its applicability is however limited by high computational costs. To transfer it to computationally expensive river modeling tasks, we propose to combine BMS with model reduction. Specifically, we build response surfaces via the so-called arbitrary Polynomial Chaos Expansion (aPC) to approximate several competing sediment transport models. Despite iterative refinement of the response surfaces, it needs to be acknowledged that they are only approximations of the models that we actually want to rank. We, therefore, introduce a novel correction factor in BMS, yielding a model ranking that is (more) representative of the full-complexity models. We demonstrate our proposed approach on a case study for a 10-km stretch of the lower Rhine river, with data provided by the Federal Waterways Research Institute (BAW) in Karlsruhe, Germany. We first verify the proposed framework in a synthetic benchmark experiment and then obtain a ranking of four commonly used sediment transport formulas (Meyer-Peter and Müller, Hunziker, van Rijn, and Wu) based on actual measurements. Results show that the correction factor shields us from misleading model ranking results. Equipped with this correction factor, the proposed model selection framework provides valuable guidance in choosing among computationally expensive models for river engineering or any other geoscience modeling task.

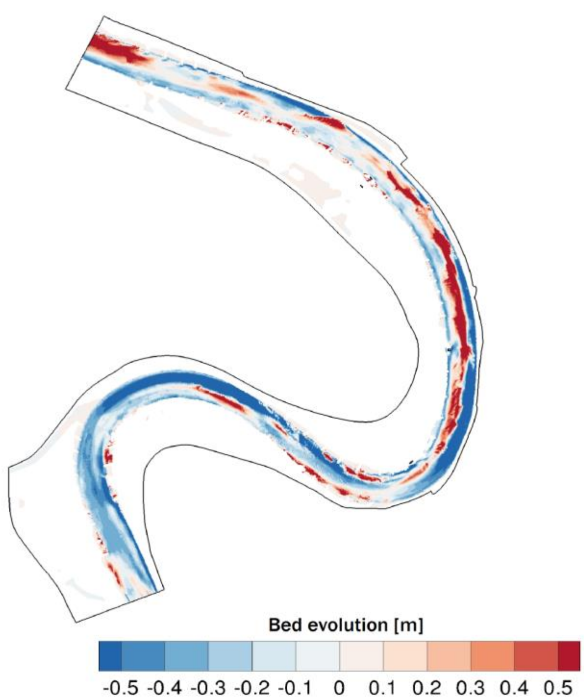

Figure 1: Bed evolution for the lower Rhine model obtained using the Meyer-Peter and Müller formula.

Bayesian model selection under computational time constraints

Bayesian model evidence (BME) is an objective metrics that can be served for model ranking. BME quantiles the likelihood of the model to have produced the observed data averaged over the complete prior parameter space. BME can be evaluated by integration over the full parameter space. This study focuses on model selection for computationally challenging simulations. In this case, the model selection analysis only becomes feasible in combination with model reduction. We establish the model reduction using so-called response surfaces iteratively constructed via arbitrary polynomial chaos expansion. It adjusts to arbitrary probability distribution shapes of input parameters and can even work with limited data that accrue during the Bayesian updating. However, reduced models are only a rough approximation of the full-complexity model. Hence, any conclusions drawn from BME values based on reduced models are only valid to the degree that we believe in the accuracy of the approximation. One may accept this limitation as a "necessary evil" of BMS via model reduction, i.e. a trade-off with efficiency. Indeed, as stated earlier, BMS only becomes feasible for applications like the one presented in this study through model reduction. Still, care must be taken when interpreting model ranking results based on model evidence for reduced models, BME_{RM}. If, using BME_{RM}, a model were penalized because its reduced variant shows a low approximation quality, model ranking via reduced models could be viewed as a conservative guess about the appropriateness of the full-complexity models. However, this is not necessarily the case: the reduced model could meet (just by chance) the observed data even better than the full-complexity model, and in that case the evidence in favor of the model would be overestimated. Hence, model ranking could turn out differently for the reduced models than for the full-complexity models, which would trigger wrong conclusions and finally possibly wrong management decisions. To avoid such misleading results, we go one step further and propose to explicitly account for the quality of approximation by the reduced model in model ranking. We thereby hope to obtain model ranking results that are more representative of the desired model ranking results for the full-complexity models. To do so, we introduce a credibility weight (denoted as $Weight_{RM}$) that reflects the approximation error by the reduced model. This yields a corrected BME value (BME_{corr}) that is more representative (although still an approximation) of the full-complexity model's BME. The resulting model ranking is listed in Table1 for the lower Rhine case study.

Model	Rank	BME_{RM}	$Weight_{RM}$	BME_{corr}
Hunziker	1	1.44E+36	1.79E-68	2.56E-32
Van Rijn	2	3.73E+36	1.94E-69	7.24E-33
Wu	3	3.79E+27	2.67E-68	1.02E-40
Meyer-Peter and Müller	4	1.84E+10	1.06E-68	1.94E-58

Table 1: BME ranking of sediment transport models accounting for the approximation quality of reduced models

Conclusions

The present paper deals with Bayesian selection of hydro-morphodynamic models under computational time constraints. We incorporate the Bayesian model evidence as a model evaluation yardstick for ranking competing models that represent the hydro-morphodynamics. To make sure that BMS framework is applicable for computationally intensive hydro-morphodynamic models, we combine it for the first time with a model reduction technique based on the arbitrary polynomial chaos expansion. Taking into consideration that the Bayesian model evidence based on the reduced model is merely an estimate of the true BME value, we newly introduce a correction factor that corrects the model ranking in view of the approximation quality. Thus, we conclude that blind use of Bayesian principles under strong computational time constraints could be misleading and an assessment of approximation quality of reduced models should be taken into consideration. Overall, the presented reduced model-based Bayesian selection framework is helpful to objectively select between different available models according to their capability to capture the observation data, even if the models are computationally too costly for standard procedures of BMS.