



Quantifying Model Similarity and Adequacy for Model Development

Aline Schäfer-Rodrigues-Silva, Timo Seitz, Anneli Guthke, and Wolfgang Nowak

University of Stuttgart, Institute for Modelling Hydraulic and Environmental Systems, Department of Stochastic Simulation and Safety Research, Stuttgart, Germany (aline.schaefer@iws.uni-stuttgart.de)

Hydrological processes can be represented by various numerical models differing in the considered system components and interactions in the conceptual model and their mathematical and numerical implementation. This leads to uncertainty in model development and model choice, the so-called conceptual uncertainty. Multi-model approaches are statistical methods that address this uncertainty or enable us to select the (sub-)model that suits our modelling goals best. For any of these methods it is crucial to know how similar the models under consideration are and how well they agree with different types of measurement data. However, quantifying the similarity is challenging if we want to consider different prediction variables over the entire spatial domain and/ or at each time step of the discretized system. Usually, this leads to high-dimensional problems that make an intuitive visualization impossible. To address this issue, we use a Monte-Carlo approach and sample the output probability density functions of each candidate model for all quantities of interest and calculate the distances between these functions. In doing so, we reduce the dimensionality of the problem and, by using multidimensional scaling, we are finally able to visualize the models and the data on a two-dimensional plane. In this representation, we can see intuitively how similar models are and if they surround the observed data or if they cluster in certain areas and thus can be considered as (partly) redundant. Quantifying the similarity between models or between a model and the observed data helps us to find the best multi-model strategy for a certain research question, to assess a model's adequacy given measurement data, and to identify promising strategies for model improvement.