The Decisive Role of Model Complexity in Model Selection

Marvin Höge (1,2), Thomas Wöhling (3,4), and Wolfgang Nowak (2)

(1) Center for Applied Geosciences, University of Tübingen, Tübingen, Germany (marvin.hoege@student.uni-tuebingen.de),
(2) Institute for Modelling Hydraulic and Environmental Systems (LS3)/SimTech, University of Stuttgart, Stuttgart, Germany,
(3) Department of Hydrology, Technical University of Dresden, Dresden, Germany, (4) Lincoln Environmental Research,
Lincoln Agritech, Hamilton, New Zealand

Modelers often have to select a "good" or even "best" model; this occurs when multiple model alternatives are given for an observed phenomenon and all of them seem plausible. Then, the level of plausibility of each candidate model has to be assessed, so that a subsequent model rating can indicate the preferential model choice for the desired impact study.

For a modeler, the best model fulfills a certain purpose best (e.g. flood prediction), which is typically assessed by comparing model simulations to data (e.g. stream flow). Model selection methods quantify the "best" trade-off between good fit with data and model complexity, i.e. they follow the principle of parsimony. In this context, the interpretations of model complexity implied by different model selection methods are crucial, because they represent different underlying goals of modeling.

Over the last decades, numerous model selection criteria have been proposed, but modelers who primarily want to apply a model selection criterion often face a lack of guidance for choosing the right criterion that matches their goal. We propose a classification scheme for model selection criteria that helps to find the right criterion for a certain goal.

We identify four model selection classes which seek to achieve high predictive density, low predictive error, high model probability or shortest compression of data. These goals can be achieved by following either non-consistent or consistent model selection and by either incorporating a Bayesian parameter prior or not. We show how commonly used criteria like Akaike's AIC or the BIC (Bayesian information criterion) can be classified according to our scheme and we provide guidance on choosing the right type of criteria for a specific model selection task.

This way, we want to enable modelers to distinguish between useful and useless models for their purposes and help to make sure that the objectively best model, for e.g. future predictions or process understanding, is selected.