Geophysical Research Abstracts
Vol. 12, EGU2010-10288, 2010
EGU General Assembly 2010

On the significance of the Nash-Sutcliffe efficiency measure for event-based flood models

Roger Moussa

Institut National de la Recherche Agronomique, UMR LISAH, Montpellier Cedex 1, France (moussa@supagro.inra.fr, 0033-4-67-63-26-14)

When modelling flood events, the important challenge that awaits the modeller is first to choose a rainfall-runoff model, then to calibrate a set of parameters that can accurately simulate a number of flood events and related hydrograph shapes, and finally to evaluate the model performance separately on each event using multi-criteria functions. This study analyses the significance of the Nash-Sutcliffe efficiency (NSE) and proposes a new method to assess the performance of flood event models (see Moussa, 2010, "When monstrosity can be beautiful while normality can be ugly : assessing the performance of event-based-flood-models", Hydrological Science Journal, in press). We focus on the specific cases of events difficult to model and characterized by low NSE values, which we call "monsters". The properties of the NSE were analysed as a function of the calculated hydrograph shape and of the benchmark reference model. As application case, a multi-criteria analysis method to assess the model performance on each event is proposed and applied on the Gardon d'Anduze catchment.

This paper discusses first the significance of the well-known Nash-Sutcliffe efficiency (NSE) criteria function when calculated separately on flood events. The NSE is a convenient and normalized measure of model performance, but does not provide a reliable basis for comparing the results of different case studies. We show that simulated hydrographs with low or negative values of NSE, called "monsters", can be due solely to a simple lag translation or a homothetic ratio of the observed hydrograph which reproduces the dynamic of the hydrograph, with acceptable errors on other criteria. In the opposite, results show that simulations with a NSE close to 1 can become "monsters" and give very low values (even negative) of the criteria function G, if the average observed discharged used as a benchmark reference model in the NSE is modified. This paper argues that the definition of an appropriate benchmark for model performance, and in particular for measures such as the NSE values should become part of the "best practices" in hydrologic modelling. Every modelling study should justify the choice of an appropriate benchmark which will be necessarily different for different types of case studies depending on the main hydrological processes and the driving variables. The benchmark should be simple such as that every hydrologist can understand its explanatory power and, therefore, appreciate how much better the actual hydrologic model is.

Then, on the basis of the previous results, we propose a new simple quick-to-use method to analyse the model performance using multi-criteria functions. In addition to the classical criteria functions, we define new complementary criteria functions, derived from the NSE on the basis of a lag translation (NSET) or a homothetic ratio (NSEH) of the calculated hydrograph. Using the NSET and NSEH, the modeller can see whether the shape of an event is well simulated but not the timing or the volume. We call "normal" the events for which all NSE, NSET and NSEH are high. We call "pseudo-monster" the events for which NSET and NSEH are high while NSE is low; hence, modelling these events can be improved when modifying the production or/and the transfer function. The remaining events for which NSE, NSET, NSEH and the remaining classical criteria functions are low are called "real monsters".

This approach enables a correct diagnostic of the model failures and will guide the modeller in searching a solution to improve the simulations of the "pseudo monsters" : review the quality of the input data, modify the concepts of the model, modify the performance criteria functions, etc. However, modelling the "real monsters" remains unresolved, and model parameterisation, calibration, validation and evaluation is still a difficult issue for

hydrological modelling.