Geophysical Research Abstracts
Vol. 12, EGU2010-10725, 2010
EGU General Assembly 2010

Data driven model generation based on computational intelligence

Peter Gemmar (1), Oliver Gronz (1), Christophe Faust (1), and Markus Casper (2)

(1) University of Applied Sciences (FH) Trier, i3A, Informatics, Trier, Germany (p.gemmar@fh-trier.de, +49-561-8103-454), (2) University Trier, Germany

The simulation of discharges at a local gauge or the modeling of large scale river catchments are effectively involved in estimation and decision tasks of hydrological research and practical applications like flood prediction or water resource management. However, modeling such processes using analytical or conceptual approaches is made difficult by both complexity of process relations and heterogeneity of processes. It was shown manifold that unknown or assumed process relations can principally be described by computational methods, and that system models can automatically be derived from observed behavior or measured process data. This study describes the development of hydrological process models using computational methods including Fuzzy logic and artificial neural networks (ANN) in a comprehensive and automated manner.

Methods We consider a closed concept for data driven development of hydrological models based on measured (experimental) data. The concept is centered on a Fuzzy system using rules of Takagi-Sugeno-Kang type which formulate the input-output relation in a generic structure like R_i : IF$q(t) = low$AND...THEN$q(t + \triangle t) = a_{i0} + a_{i1}q(t) + a_{i2}p(t - \triangle t_{i1}) + a_{i3}p(t + \triangle t_{i2}) + \ldots$. The rule's premise part (IF) describes process states involving available process information, e.g. actual outlet $q(t)$ is *low* where *low* is one of several Fuzzy sets defined over variable $q(t)$. The rule's conclusion (THEN) estimates expected outlet $q(t + \triangle t)$ by a linear function over selected system variables, e.g. actual outlet $q(t)$, previous and/or forecasted precipitation $p(t \mp \triangle t_{ik})$. In case of river catchment modeling we use head gauges, tributary and upriver gauges in the conclusion part as well. In addition, we consider temperature and temporal (season) information in the premise part. By creating a set of rules $\mathbf{R} = \{R_i|(i = 1, \ldots, N)\}$ the space of process states can be covered as concise as necessary. Model adaptation is achieved by finding on optimal set $\mathbf{A} = (a_{ij})$ of conclusion parameters with respect to a defined rating function and experimental data. To find $\mathbf{A}$, we use for example a linear equation solver and RMSE-function.

In practical process models, the number of Fuzzy sets and the according number of rules is fairly low. Nevertheless, creating the optimal model requires some experience. Therefore, we improved this development step by methods for automatic generation of Fuzzy sets, rules, and conclusions. Basically, the model achievement depends to a great extend on the selection of the conclusion variables. It is the aim that variables having most influence on the system reaction being considered and superfluous ones being neglected. At first, we use Kohonen maps, a specialized ANN, to identify relevant input variables from the large set of available system variables. A greedy algorithm selects a comprehensive set of dominant and uncorrelated variables. Next, the premise variables are analyzed with clustering methods (e.g. Fuzzy-C-means) and Fuzzy sets are then derived from cluster centers and outlines. The rule base is automatically constructed by permutation of the Fuzzy sets of the premise variables. Finally, the conclusion parameters are calculated and the total coverage of the input space is iteratively tested with experimental data, rarely firing rules are combined and coarse coverage of sensitive process states results in refined Fuzzy sets and rules.

Results The described methods were implemented and integrated in a development system for process models. A series of models has already been built e.g. for rainfall-runoff modeling or for flood prediction (up to 72 hours) in river catchments. The models required significantly less development effort and showed advanced simulation results compared to conventional models. The models can be used operationally and simulation takes only some minutes on a standard PC e.g. for a gauge forecast (up to 72 hours) for the whole Mosel (Germany) river catchment.