Geophysical Research Abstracts, Vol. 8, 09999, 2006
SRef-ID: 1607-7962/gra/EGU06-A-09999

Fuzzy clustering and neural networks in localized estimation of the total model uncertainty

D.Shrestha (1), **D.Solomatine** (2)

UNESCO-IHE Institute for Water Education (d.shrestha@unesco-ihe.org, sol@ihe.nl /Fax +31152122921)

Model uncertainty is often seen equivalent to its parameters' uncertainty, or to the input data uncertainty. In this paper, Uncertainty is treated as the probability distribution of the output model errors. A method for estimating model prediction uncertainty using neural networks and fuzzy logic is presented. The direct analytical estimation of the probability distribution of the model errors is often impossible; in a simplified setting it can be expressed it in the form of its quantiles, or prediction limits (interval) of the model prediction. Such expression of uncertainty comprises all types of uncertainty (model structure, model parameters and data). The idea is to partition the model input space into different zones or clusters having similar values of model errors. This is done by (a) building a data matrix (training set) by combining (some of the) historical model inputs and corresponding model error; (b) partitioning this training set using fuzzy c-means clustering, and identifying the corresponding regions of input space to the clusters.

The prediction interval is constructed for each cluster and propagated from each cluster to the examples according to their (fuzzy) membership in each cluster. Using fuzzy logic makes it possible to maintain proximity of the prediction intervals of the examples that are close in input space.

On the basis of in-sample data using computed prediction limits as targets a regression numerical prediction model (ANN) is built for, and, finally, this model is applied to estimate the prediction intervals (limits) for out-of-sample data. The method was applied to estimate uncertainty of a hydrologic forecasting model. Results show that the method is superior to other methods estimating the prediction interval.