Geophysical Research Abstracts, Vol. 9, 08953, 2007
SRef-ID: 1607-7962/gra/EGU2007-A-08953

Let's accentuate the negative: using feedback loops to examine and compare four different neural network river discharge forecasters

R.J. Abrahart (1) and L.M. See (2)

(1) School of Geography, University of Nottingham, UK; (2) School of Geography, University of Leeds, UK (Email: bob.abrahart@nottingham.ac.uk)

Following the advent and virtual explosion of computational data-driven solutions in the field of hydrological forecasting and prediction there remains a pressing need to reconsider the manner in which such tools are tested. It is argued that traditional models and traditional assessment procedures are in most instances designed to assess the magnitude and extent of correctness i.e. performance under prescribed conditions. The question of practical operational matters such as model robustness are seldom if ever tested or reported. For example, in terms of pollution thresholds, it would be important to determine the 'worst case scenario'. The longer term challenge must be to develop more powerful or more comprehensive assessment metrics; it is, however, for the moment perhaps easier to consider potential improvements to the manner in which existing statistical measures are applied or can be extended. This paper argues that a number of benefits can be gained from the application of traditional statistics based on the industrial and engineering principles of destructive testing and failure analysis. 'Destructive Testing' involves pushing the model to its breaking point. Subsequent inspection and assessment of the shattered product is used to determine appropriate conformance in respect of its design and process requirements. 'Failure Analysis' comprises a methodical and well-organised examination of related failures. It is used to acquire information about failures, to determine the root cause of failures and to recommend corrective action.

Four neural network discharge forecasting models were evaluated with regard to the detrimental effect of incorporating model output predictions as inputs in a repeated manner based on the use of feedback loops i.e. 'marching forward'. Each model was developed to predict one-hour-ahead discharge for the Upper River Wye in Central Wales. Eight global evaluation metrics were used to assess the progressive degradation of each model at each step. The introduction of accumulated error was observed to generate important changes in the predicted model outputs; producing effects that varied [1] from model to model and [2] over different parts of the hydrographic record. The magnitude and timing of forecasts was disrupted. Significant flood events appeared to be the worst affected casualties. However, on a relative basis, considerable damage was also inflicted upon the lower water level outputs. The manner and pattern of model destruction will be described and used to provide important insights into the process of developing neural network rainfall-runoff solutions.