

## APPLICABILITY OF DATA-DRIVEN MODELLING FOR RESERVOIR WATER QUALITY MANAGEMENT

NAHM-CHUNG JUNG<sup>1</sup> and ROLAND. K. PRICE<sup>2</sup>

<sup>1</sup>Ph.D. Researcher, UNESCO-IHE, P.O. Box 3015, 2601 DA, Delft, the Netherlands  
(Tel: +82-42-860-0360, Fax: +82-42-860-0369, e-mail: chung@kowaco.or.kr)

<sup>2</sup>Professor, UNESCO-IHE, P.O. Box 3015, 2601 DA, Delft, the Netherlands  
(Tel: +31-15-215-1715, Fax: +31-215-215-1871, e-mail: r.price@unesco-ihe.org)

Data-driven models are an approach to manage uncertainty of model structure, input data and parameters. Artificial neural networks including fuzzy logic and model trees are a state-of-the-art technology in this area, because both models are more explicit than any other data mining techniques. Although they need a considerable amount of historical data, these kinds of data-driven models have been growing tendency to complement or even replace physically based models that need uncertainty analysis. This paper is supposed to analyze both data-driven models one of which will be a component of decision support system for reservoir water quality management. Until now the cases of water quality and ecological modelling using data mining techniques such as ANFIS and MTs have not nearly been reported. MTs are a little bit more explicit than ANFIS. The reason is that the splitting criterion at each node in MTs is more comprehensible than the MF (membership function) in ANFIS; the results of MT can be easily understood by decision makers. ANFIS is higher non-linear than MTs, allowing tracing of peaks (see Figure 5 and 6). Whereas MTs have shown disadvantage in tracing the peaks, because it takes the mean of parameters which lie inside two standard deviations from the mean (see Figure 7).

The choice of the better solution between ANFIS and MT for DSS (Decision Support System) should be decided after further research, for instance a time series CCF (cross correlation function) analysis between each sampling station and all sampling stations of tributaries.

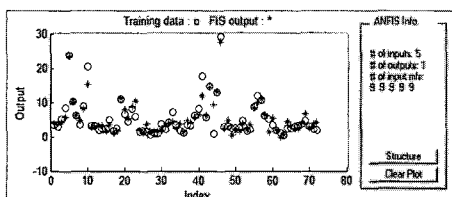


Fig. 5 Training result of ANFIS

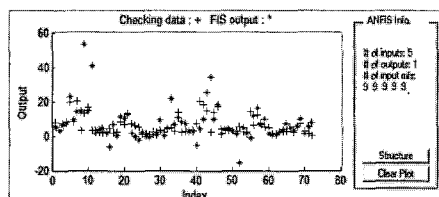


Fig. 6 Checking result of ANFIS

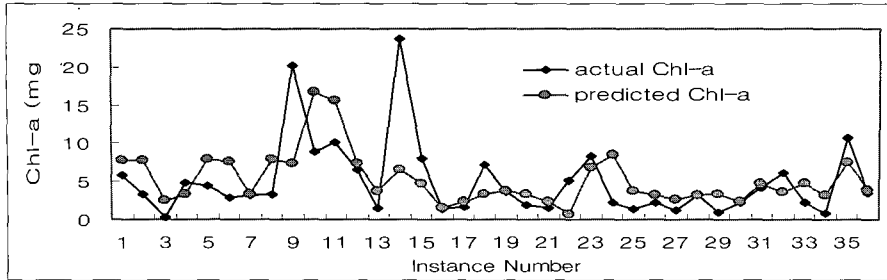


Fig. 7 Comparison of the actual and predicted Chl-a concentration at the monitoring station No.1.

### REFERECES

- B. Bhattacharya and D.P. Solomatine (2003), "Neural networks and M5 model trees in modeling water level-discharge relationship for an Indian river." ESANN'2003 proceedings-European Symposium on Artificial Neural Networks Bruges (Belgium), 23-25 April 2003, d-side public., ISBN 2-930307-03-X, pp. 407-412.
- B. Egilegor, J.P. Uribe, G. Arregi, E. Pradilla, L. Susperregi "A Fuzzy Control Adapted by a Neural Network to Maintain a Dwelling within Thermal Comfort."  
<http://www.ibpsa.org/proceedings/bs97/papers/P097.PDF>
- DIMITRI P. SOLOMATINE and KHADA N. DULAL (2003), "Model trees as an alternative to neural networks in rainfall-runoff modelling." *Hydrological Sciences-Journal-des Sciences Hydrologiques*, 48(3) June 2003.
- Donato Malerba; Annalisa Appice, Antonia Bellino, Michelangelo Ceci, Domenico Pallota; 2001), "Stepwise Induction of Model Trees." F. Esposito (Ed): *AL\*IA 2001*, LNAI 2171, pp. 20-32, 2001. © Springer-Verlag Berlin Heidelberg 2001.
- J.-S. R. JANG, C.-T. SUN & E. MIZUTANI (1997), "Neuro-Fuzzy and Soft Computing." Prentice-Hall International, Inc.
- IAN H. WITTEN & EIBE FRANK (2000), "Data Mining." A Harcourt Science and Technology Company, 525 B Street, Suite 1900, San Diego, CA 92101-4495, USA, ISBN 1-55860-552-5
- Michelangelo Ceci, Annalisa Appice, Donato Malerba, "Comapaing Simplification Methods for Model Trees with Regression ND Splitting Nodes." Dipartimento di Informatica, Università degli Studi via Orabona, 4-70126 Bari – Italy.
- Niels Landwehr, Mark Hall, and Eibe Frank, "Logistic Model Trees."  
<http://www.cs.waikato.ac.nz/~ml/publications/2003/landwehr-et-al.pdf>