DAMP: A protocol for contextualising goodness-of-fit statistics in sediment-discharge data-driven modelling

Robert J.Abrahart[a]; Nick J.Mount[a]; NgahzaifaAb Ghani[ab]; Nicholas J.Clifford[c]; Christian W.Dawson[d]

[a]School of Geography, University of Nottingham, Nottingham NG7 2RD, UK
[b]Faculty of Civil Engineering and Earth Resources, Universiti Malaysia Pahang, 26300 Kuantan, Pahang Darul Makmur, Malaysia
[c]Department of Geography, Kings College London, Strand Campus, Strand, London WC2R 2LS, UK
[d]Department of Computer Science, Loughborough University, Loughborough LE11 3TU, UK

ABSTRACT

The decision sequence which guides the selection of a preferred data-driven modelling solution is usually based solely on statistical assessment of fit to a test dataset, and lacks the incorporation of essential contextual knowledge and understanding included in the evaluation of conventional empirical models. This paper demonstrates how hydrologic insight and knowledge of data quality issues can be better incorporated into the sediment-discharge data-driven model assessment procedure: by the plotting of datasets and modelled relationships; and from an understanding and appreciation of the hydrologic context of the catchment being modelled. DAMP: a four-point protocol for evaluating the hydrologic soundness of data-driven single-input single-output sediment rating curve solutions is presented. The approach is adopted and exemplified in an evaluation of seven explicit sediment-discharge models that are used to predict daily suspended sediment concentration values for a small tropical catchment on the island of Puerto Rico. Four neurocomputing counterparts are compared and contrasted against a set of traditional log–log linear sediment rating curve solutions and a simple linear regression model. The statistical assessment procedure provides one indication of the best model, whilst graphical and hydrologic interpretation of the depicted datasets and models challenge this overly-simplistic interpretation. Traditional log–log sediment rating curves, in terms of soundness and robustness, are found to deliver a superior overall product – irrespective of their poorer global goodness-of-fit statistics.

KEYWORDS:

Data-driven model; Modelling protocol; Rating curve; Suspended sediment; Tropical catchment; Hydrologic context

REFERENCES

1. Abrahart, R.J., Ab Ghani, N., Swan, J., 2009. DISCUSSION of An explicit neural network formulation for evapotranspiration. J. Hydrol. Sci. 54, 382–388

2. Abrahart, R.J., See, L.M., 2007. Neural network modelling of non-linear hydrological relationships. Hydrol. Earth Syst. Sci. 11, 1563–1579

3. Abrahart, R.J., See, L.M., Dawson, C.W., Shamseldin, A.Y., Wilby, R.L., 2010. Nearly two decades of neural network hydrologic modeling. In: Sivakumar, B., Berndtsson, R. (Eds.), Advances in Data-Based Approaches for Hydrologic Modeling and Forecasting. World Scientific Publishing, Hackensack, NJ, pp. 267–346.

4. Abrahart, R.J., See, L.M., Heppenstall, A.J., White, S.M., 2008. Neural network estimation of suspended sediment: potential pitfalls and future directions. In: Abrahart, R.J., See, L.M., Solomatine, D.P. (Eds.), Practical Hydroinformatics: Computational Intelligence and Technological Developments in Water Applications. Springer-Verlag, Berlin and Heidelberg, pp. 139–161.

5. Alp, M., Cigˇizogˇlu, H.K., 2007. Suspended sediment load simulation by two artificial neural network methods using hydrometeorological data. Environ. Modell. Softw. 22, 2–13.