Hydrol. Earth Syst. Sci. Discuss., 10, C149–C154, 2013
www.hydrol-earth-syst-sci-discuss.net/10/C149/2013/

Hydrology and Earth System Sciences Discussions

Interactive comment on "Legitimising neural network river forecasting models: a new data-driven mechanistic modelling framework" *by* N. J. Mount et al.

H. Maier (Referee)

holger.maier@adelaide.edu.au

Received and published: 27 February 2013

The authors address a very important issue in relation to the application of ANNs and data-driven models in hydrology, which is the degree to which such models are able to capture underlying physical processes from the data, which has a large impact on the credibility of such models. The paper is well written and organised and the proposed approach and illustrative case study are clear and useful. However, I think the quality and contribution of the paper could be improved if the authors considered the following points:

1. Different levels of generality (e.g. data-driven models, ANNs, MLPs) apply / are

referred to in different sections of the paper, which can be confusing at times. It would be useful to articulate clearly which aspects of the proposed approach are applicable at the different levels of generality. For example, it is my understanding that the proposed framework is generally applicable to all data-driven models, including ANNs, whereas the actual sensitivity analysis approach is only applicable to MLPs. Is this correct? If so, it would be helpful and assist readers if the distinction between the generic conceptual approach and the specific method, as well as their areas of applicability, were made clearer in the paper.

2. I am also not clear what is mean by the term "neural network river forecasting". Maybe I am being pedantic here, but you are not really forecasting "the river", but certain hydrological variables in rivers. It is also not clear to me why the framework is restricted to the forecasting of hydrological variables in rivers. I understand the difficulties associated with defining a research area in this field, as although methods are more generally applicable, one only has experience in a particular application area... Maybe a good way forward would be to say that the discussion in the paper is restricted to the application of the approach to the forecasting of hydrological variables in rivers, but that the approach is more widely applicable.

3. I think there is also some confusion about the purpose of the proposed framework. It is clear that you are trying to assess how well underlying physical processes have been captured by a calibrated ANN model. However, what is less clear is how this information should be used. There is some discussion about model validation, yet the way the methods are used are to decide which model structure (e.g. in terms of the number of inputs and the number of hidden nodes) is best. I think it would be useful to distinguish clearly between the different steps in the model development process. To my way of thinking, at least, validation occurs after the "optimal" model has been developed (i.e. after the best inputs and model structure have been determined) and is only done on "the" selected model structure and compared with some criteria that determine whether the developed model is valid or not based on the outcome of the

validation procedure. Consequently, the way the approach is used in the case studies is to determine the best model, not to validate the model. Again, this might be somewhat pedantic, but I think this is an important distinction and the quality of the paper could be improved by tightening the language surrounding some of these issues.

4. Following on from 3 above, it would be useful to refer to and discuss some of the methods that have been developed to improve input selection, model structure selection etc., not just literature on knowledge extraction. For example, many methods have been developed to select appropriate inputs to ANNs that take into account nonlinearity and correlation between potential inputs (e.g. partial mutual information) that are designed to overcome the issues discussed in the paper and case study. A distinction needs to be made between filter (model-free) and wrapper (model-based) input selection algorithms. The approach introduced in this paper is applicable to wrapper based input-selection algorithms, but not if filters are used. However, it could be used to validate a model developed using inputs obtained using a filter method. Similarly, there are methods for determining the optimal number of hidden nodes. One example is using Bayesian methods (Kingston et al., 2008) and information criteria (e.g. AIC, BIC) have also been used in order to strike a balance between predictive performance in terms of the selected error measures and model complexity. Again, this is done as part of the model development process (i.e. pre-validation) in an attempt to obtain the most parsimonious model that results in adequate predictive performance and one would think that such models also represent the underlying physical processes better. There is still a need to validate such models, which is why it is important to distinguish between methods used for model development and those for validation.

Kingston G.B., Maier H.R. and Lambert M.F. (2008) Bayesian model selection applied to artificial neural networks used for water resources modeling, Water Resources Research, 44, W04419, doi:10.1029/2007WR006155.

5. It would also be useful to include additional literature in relation to elucidating the internal workings of ANN models and approaches that have been developed to ensure

calibrated ANNs are physically plausible. For example, the overall connection weight approach developed by Olden and Jackson (2002) is a seminal paper in this field that has been cited over 130 times since its publication. The approach has been applied a number of times for the forecasting of hydrological variables in rivers by Kingston et al., including ensuring that ANN weights are obtained during model calibration so that the physical processes captured by the ANN make physical sense, even if this is at the expense of prediction accuracy in terms of error metrics (Kingston et al., 2005). In addition, the use of sensitivity analysis for determining the relative contributions of inputs for trained ANN models used for the forecasting of hydrological variables has been used previously by Maier and Dandy before the application by Sudheer (2005). Another example of extracting knowledge from ANNs is the use of neurofuzzy models, which are a class of ANN model. I would encourage the authors to place their work in the context of this and related other work to clarify the contribution of this paper and how it fits into this wider research field. The references below are just a (rather biased) sample of relevant papers in this field, but illustrate that a significant amount of work has been done in this area.

Maier H.R. and Dandy G.C. (1997) Determining inputs for neural network models of multivariate time series. Microcomputers in Civil Engineering - Journal of Computer-Aided Civil and Infrastructure Engineering, 12(5), 353-368.

Maier H.R., Dandy G.C. and Burch M.D. (1998) Use of artificial neural networks for modelling the incidence of cyanobacteria Anabaena spp. in River Murray, South Australia. Ecological Modelling, 105(2/3), 257-272.

Maier H.R. and Dandy G.C. (2000) Application of artificial neural networks to forecasting of surface water quality variables: issues, applications and challenges, in Artificial Neural Networks in Hydrology, edited by R.S. Govindaraju and A.R. Rao, Kluwer, Dordrecht, The Netherlands, 287-309.

Maier H.R. and Dandy G.C. (2001) Neural network based modelling of environmental variables: a systematic approach. Mathematical and Computer Modelling, 33(6-7), 669-682.

Maier H.R., Sayed T. and Lence B.J. (2001) Forecasting cyanobacterium Anabaena spp. using B-spline neurofuzzy models. Ecological Modelling, 146(1-3), 85-96.

Maier H.R., Sayed T. and Lence B.J. (2000) Forecasting cyanobacterial concentrations using B-spline networks. Journal of Computing in Civil Engineering, ASCE, 14(3), 183-189.

Olden, J.D., Jackson, D.A., 2002. Illuminating the 'black box': a randomization approach for understanding variable contributions in artificial neural networks. Ecological Modelling 154 (1–2), 135–150.

Kingston G.B., Maier H.R. and Lambert M.F. (2003) Understanding the mechanisms modelled by artificial neural networks for hydrological prediction. Modsim 2003 - International Congress on Modelling and Simulation, Modelling and Simulation Society of Australia and New Zealand Inc, Townsville, Australia, 14-17 July, Vol. 2, pp.825-830.

Kingston G.B., Maier H.R. and Lambert M.F. (2005) Calibration and validation of neural networks to ensure physically plausible hydrological modeling. Journal of Hydrology, 314(1-4), 158-176.

Kingston G.B., Maier H.R. and Lambert M.F. (2006) Forecasting cyanobacteria with Bayesian and deterministic artificial neural networks. IEEE World Congress of Computational Intelligence, Vancouver, Canada, July 16-21.

Kingston G.B., Maier H.R. and Lambert M.F. (2006) A probabilistic method to assist knowledge extraction from artificial neural networks used for hydrological prediction. Mathematical and Computer Modelling, 44(5-6), 499-512.

6. Given that there has been quite a bit of work in this area using different methods, the quality of this paper would be enhanced significantly if the proposed approach was compared with that of others. For example, the overall connection weight approach

is very simple to apply. However, if no such comparisons are performed, it would be useful to have some qualitative comparison of different methods, including advantages and disadvantages, as this would be extremely useful for people working in this space.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 10, 145, 2013.