Hydrol. Earth Syst. Sci. Discuss., 2, S82–S87, 2005
www.copernicus.org/EGU/hess/hessd/2/S82/
European Geosciences Union

Hydrology and Earth System Sciences Discussions

Interactive comment on "Constraints of artificial neural networks for rainfall-runoff modelling: trade-offs in hydrological state representation and modelevaluation" *by* N. J. de Vos and T. H. M. Rientjes

e. gaume (Referee)

gaume@cereve.enpc.fr

Received and published: 15 March 2005

Overall evaluation

This paper is composed of three main parts : 1) A review of the application of artificial neural networks in hydrology and a description of ANN implementation, 2) a 'standard' application of ANN on a chosen case study and 3) a test of some alternatives to reduce an observed ANN prediction lag. These three parts are of various lengths and qualities. The first one occupies almost 3/4 of the whole text of the paper. It is to my opinion a well written, clear and documented review which, I think, is welcome and will certainly be appreciated by many HESS readers. I have only two suggestions to formulate. Firstly, the result of section 3.4 (i.e. the superiority of the Levenberg-Marquardt algorithm for the calibration of ANN) does not seem very new to me. It is for instance mentioned in the Matlab Neural network toolbox manual. This should be said somewhere. Secondly, some concepts of the 'neural networks world'; which are not shared by all mathematicians and model developers as the 'overtraining' concept, and their justifications should be presented from a slight more critical or prudent point of view (see my detailed comments on section 2.3). It is necessary to make a clear difference in a review, especially intended for non-specialists, between demonstrated facts and lessons drawn from the experience and trial and error tests. The 'overtraining'; concept and the presented split-sample calibration approach belong, to my point of view, to the second category. The second part of the paper is interesting. Reports on applications of ANN in hydrology are also very welcome. But a reference to the results obtained with other possible forecasting models is missing: linear models using the same explaining variables, lumped conceptual R-R models with either a data assimilation procedure or in combination with a linear prediction model of errors (see Gaume & Gosset, 2004). Without elements of comparisons, it is not possible to draw any valuable conclusions on the ANN usefulness and efficiency. Some other forecasting models should been tested including a R-R model, and their results compared to the ones obtained with the best ANN for each lead time. Their are also too many figures. The scatter plots 9, 11 and 12 do not add information to the R2 criterion. A simple lag time value could replace figures 7 and 8. In general, the figures should be much more commented and used in the text. Finally, if forecasting is the objective of the model, the criterion PI should be preferred to R2. The third part which justifies the title of the paper is less convincing. Firstly, the justification of it 'the prediction lags (have) been remarkably overlooked' is excessive and untrue (see my additional comments). Secondly, the paper does not propose any efficient solution to this 'problem'. Here again, it is necessary to compare the performances of the calibrated ANNs with the performances of a R-R model. My feeling is that ANN can hardly reproduce the rainfall-runoff dynamics and that it is probably the main conclusion to be drawn for this part of the paper. If this is correct, than the performances of ANN which do not use any information on the previous discharges may be much lower than the performances of the R-R models. But I may be wrong. The only way to settle the argument is to compare ANN and R-R models. By the way, their are some strange results in table 2. The efficiency of the ANN sometimes decreases drastically as the number of input data increases (compare model 3 and 6 for instance). The results presented in figure 17, especially on the peaks, are quite good: is it really a flood of the third (validation) data set ?

As a conclusion. This paper is really interesting but it must be completed with the forecasting results obtained using other types of models to be of any real use for hydrologists: linear models, lumped conceptual rainfall-runoff models. I am not sure that such an emphasis, especially in the title of the paper, should be put on the third part of the paper since no real efficient solution is proposed. The comparison with other type of forecasting methods may, by the way, modify significantly the point of view of the authors and their conclusions.

Detailed comments :

Page 369, line 10 : ANN R-R modelling is presented as relatively new. The references show that ANN have been tested in hydrology for at least 10 years.

Page 371 lines 10-25, the authors reproduce, without any distance, the justification for the split-sampling method which appears in many papers and text books on neural networks. The concept of overtraining, used essentially in the field of neural networks, is highly questionable. The modelling conventional approach consists in limiting the number of parameters of a model to be calibrated (Beck, 1987; Jorgensen, 1988; Perrin et al 2003) rather than finding 'tips' to handle over-parameterized models. 'The network will start learning the noise in the training data and lose its generalisation capability'. Has this been demonstrated ? To my knowledge no. Moreover, it has been shown that 'overtraining' or over-parametrization occurs even while calibrating ANN on signals with no noise (Gaume & Gosset, 2004). There is no miracle ! ANN are not able to distinguish noise and the deterministic part of a signal: the calibration does not begin to fit the deterministic signal and than the noise. Everything is mixed. But the split-sampling method aims at selecting the calibration trials which seem to have cached the greater part of the deterministic signal: it is in a way a trial and error method. I wonder if the efficiency of the split-sampling method combined with under-calibration has been compared to that of a standard calibration parameter parsimonious method.

Page 375 : line 19, it would be interesting to summarise the characteristics of each period in a table: mean annual rainfall amount, maximum daily and hourly rainfall intensities, mean annual discharge, maximum peak and daily discharge...

Page 376 : lines 20 to 22, I am not convinced by this square root relation between hidden and input neurones. Moreover, the number of parameters of a x-y-z architecture of a neural network is y*(z+1). The ANN in tables 2 and 3 have between 96 and 140 parameters ! I am not sure that the term 'parsimonious' (line 20) is well suited to these models.

Page 380 : line 11, 'reasonably good forecasts'. Nothing supports this statement. The performances of other more 'conventional' models (linear models, lumped R-R models coupled with an AR model on the errors ...) should be given as an element of comparison. The absence of this reference is a real lack of the paper and should absolutely be included. The efficiency of ANN can only be evaluated in reference to alternative forecasting models. The PI criterion (line 17) is not so good indicating, contrary to what is stated, poor forecasting performances. Nash and R2 criteria are clearly not suited for the evaluation of short term forecasting models.

Page 381 : line 20 'this issue has been remarkably overlooked'. This is excessive: 1) many authors have mentioned the lag in the ANN forecasts, Gaume & Gosset (2004) can by the way be added to the list, 2) no efficient solution is proposed in this paper. This lag is an important drawback of neural networks as well as of linear models which are both not able to make an efficient use of the rainfall data. Therefore the 'auto-regressive' component dominates in both models.

Page 382 : this part of the paper is really disappointing, none of the tested approaches prove to be efficient. The conclusions drawn are not really convincing. It appears clear that what the authors call the 'auto-regressive' component is the explanation for the 'reasonably good forecasts' of the ANN. But they partly fail in representing the R-R relation: this is to my point of view the main conclusion which can be drawn from this part of the paper.

Page 383 : The results reported in Table 2 corresponding to the model (P,Qma) seem to be incorrect. It is strange that this model has much better performances than the (P,Qma,Pma,SM) model including the same variables plus Pma and SM. Figure 18 is not referred to in the text. As in figures 11 and 12, it would be interesting to mention the values of the criterions (R2, Nash and PI). The figures 16 and 17 are not really commented in the text. Are all these figures necessary ? They seem to show the results obtained on one of the most important floods. Are they representative of the whole series ?

Page 385 : what is the aim of this section here in the paper, based only on a literature review ? The authors should test some of the proposed methods or remove this section.

Page 387 : line 6 'ANNs are alternatives for traditional R-R modelling'. Such a conclusion can not be drawn without having compared the performances of both approaches. Line 8 'has hitherto been neglected by hydrologists': I understand that the authors want to put forward the originality of their paper, nevertheless this statement is excessive and therefore untrue. Line 25 'complementary conceptual models can be valuable additions to ANN' : this may be true but the results shown in the paper are not really encouraging.

Page 388 : last sentence. This conclusion is interesting but not particularly original as mentioned before.

HESSD

2, S82–S87, 2005

Interactive Comment

Various typing errors :

The values of the Nash criterions should appear in tables 2 and 3. A legend is missing in figure 15 making it very difficult to understand. The units and detailed labels are missing in fig 18.

References:

Jorgensen S.E., 1988. Fundamentals of ecological modelling: developments in environmental modelling, volume 9. Elsevier. Beck, M.B., 1987. Water quality modeling: a review of the analysis of uncertainty. Water Resources Research 23, 1987: 1393-1442

Interactive comment on Hydrology and Earth System Sciences Discussions, 2, 365, 2005.