Hydrol. Earth Syst. Sci. Discuss., 6, S132–S141, 2009
www.hydrol-earth-syst-sci-discuss.net/6/S132/2009/

Hydrology and Earth System Sciences Discussions

Interactive comment on "Daily reservoir inflow forecasting combining QPF into ANNs model" *by* Jun Zhang et al.

Anonymous Referee #1

Received and published: 20 February 2009

Review of 'Daily reservoir inflow forecasting combining QPF into ANNs model' by Zhang et al. published in HESSD 6(2009): 121-150

1. General Comments:

The paper describes an application of the Artificial Neural Network (ANN) model utilizing Quantitative Precipitation Forecasts (QPFs) for producing 1- to 7-day ahead inflow forecasts. For this purpose, the authors have applied a 3-layer ANN structure involving Back-Propagation (BP) learning algorithm and incorporating a self-adaptive training scheme with adapting learning rate and momentum term. The authors have used up to 3-day ahead QPFs, available from a medium range Numerical Weather Prediction system, to forecast reservoir inflows for 'operational planning and scheduling of hydroelectric power system' involving reservoirs.

In my opinion, the application of a rather conventional ANN model to 'river-flow forecasting' is the only contribution to the journal or researchers. The paper has a number of shortcomings described in the following section. Unless these shortcomings are addressed, the paper may not be suitable for publication in the HESS Journal. Except for the requirements of reorganising some details and providing some additional information as indicated in section 3 below, the paper is generally well organized. The language is mostly understandable, but would require editing before publication in the journal; the authors may consult an English editor for this purpose.

2. Shortcomings:

i) From the section on 'Study area and data collection' and Fig. 1 it appears that Shuikou reservoir is located on the Minjiang River and that a number of reservoirs, presumably being used for hydropower generation, exist upstream on the main river and its tributaries. It is therefore highly likely that regulated discharges from all these upstream reservoirs and associated hydroelectric plants have considerable impacts on the pattern of inflow to the Shuikou reservoir and that the inflow to the Shuikou reservoir is highly variable in time. In this context, I feel that, although the inclusion of one antecedent daily discharge on the basis of ACF as input to the selected ANN models may have implicitly accounted for a component of this variability, the highly non-parsimonious ANN structures have resulted because of the attempts to over-fit the highly variable observed flow data. It is noted that, with the 6-12-1 and 8-20-1 structures of the authors' Model(t+0) and Model(t+2), the weights (including bias) of the resulting ANNs are 97 and 201 respectively. These are undoubtedly very large numbers. Because of the lack of parsimony, the resulting models are likely to be very unreliable for real-time forecasting, particularly for input data which may not be within the range of data used for training the networks. An indication of this may be found in the results of ARIMA and Model(t+0) in Table 1, which shows that, despite a relative improvement in CE and R2 in verification in the case of the simplistic ARIMA model,

HESSD

6, S132–S141, 2009

Interactive Comment

there is a reduction in performance in the case of the corresponding (if I am right!) non-linear ANN Model(t+0). Also, from the scales provided in Figs. 5(e, f and g), it can be seen that some of the discharge values, simulated in the validation phase, deviate considerably from the 45 degree line when measured in the unit of flow used, i.e. cumec.

In view of the above, I request that the authors include some details of the upstream reservoirs, e.g. locations, size, mean daily outflows etc., in a tabular form to provide a holistic picture of the hydrologic system that they are modelling. They should also explain their considerations in respect of accounting for the likely variability of the inflow to the Shuikou reservoir caused by the outflows from the upstream reservoirs or hydro-electric plants. I believe that, without these explanations, the contents of the paper a) do not reflect what the title of the paper says, b) report the outcome of a typical (and trivial in the sense of automated application of the ANN structure) river flow forecasting only and c) does not merit classification as a good research publication.

ii) Descriptions and the notations of the ANN models in subsection 3.1 are not clear. It appears that the forecast time origin is t-1, so that the 1 day ahead forecast is indicated as being Q(t+0), i.e. Q(t), the corresponding model being represented by Model(t+0), and that the models have been used in non-updating mode. In this context, Q(t+1) in expression (1) (line 16, page 128, indicated hereinafter in this review by the convention 16/128) should be Q(t+0) and Model(t+0) in expression (4) (16/128) should be Model(t+i). The authors indicate that 'no QPF more than three days are available at present'. For forecasting 'next four days' inflows, they state that Model(t+3) is 'a unified model' given by expression (4). What do the authors mean by the term 'unified model'? The expression for this 'unified model' includes QPF(t) = 0 but does not adequately indicate that the QPFs at times t+3, t+4, t+5 and t+6 are unavailable. What are the inputs to these 'unified models'? How have the authors used the QPFs available for the previous three days? The forecast time origin remaining the same, have the authors consistently used the inputs P(t-2), P(t-1) and Q(t-1) in the models for 4-, 5-, 6- and 7-day ahead forecasts, these inputs being common to the models for 1-, 2- and 3-day ahead forecasts? It is also not clear from subsection 3.2 if the 'unified models' have 12 neurons in the hidden layer like those in Model(t+0). Please clarify. Expressions, similar to those in (1), (2) and (3), will be useful.

In the above context, it is also noted that the graphical displays of all outputs correspond to Model(t+0), Model(t+1) and Model(t+2). No result is provided for Models(t+i), i = 3, 4, 5 and 6. It would perhaps be better to drop all references to Models(t+i), i = 3, 4, 5 and 6 from the paper, and rather concentrate on those models for which QPFs are available for use.

iii) The authors have used the mean of the previous 30 days observed rainfall and flow to incorporate 'the seasonal information'. Why 30? Do the observed rainfall and discharge display any seasonality? If yes, then a graph to display the seasonality, e.g. by plotting the means of rainfall and discharge at each day over the number of years for which data have been used, will be useful. Although seasonality can be expected in the rainfall data series, I am not sure if the discharge data series, influenced by regulated outflows from upstream reservoirs or hydroelectric plants, will display marked seasonality. Authors need to clarify this aspect.

iv) The 'bench-mark' ARIMA model, having a (4,1,2) structure, has been used in the study. What is the basis of selecting this particular structure? Apparently, this model has been used only for 1-day ahead forecast. Therefore, outputs from Model(t+0) can only be compared with those from the ARIMA model. However from the abstract or from the section on 'Introduction' (10/125), the reader gets an impression that outputs from all selected ANN models (each of which is unique!) have been compared with the ARIMA model outputs. This requires clarification.

v) For each of the 2- or more-day ahead forecasts, did the authors try the structure of the model Model(t+0) itself by replacing the observed antecedent flows by the model-simulated flows and the observed antecedent rainfalls by the QPFs? Such replacements would be required only for the day(s) which lie between the forecast origin and the day for which the forecast is required. It is worth investigating the performance of this model for each of the 2- or more-day ahead forecasts vis-à-vis that of the corresponding model in the set of Model(t+i), i=1,2,3,...6. The ARIMA model can also be used in the same way for producing 2- or more-day ahead forecasts, i.e. using the simulated discharges and the QPFs for the day(s), antecedent to the day for which the forecast is required, but beyond the forecast origin. If used in this way, the outputs of the ANN model may be comparable to the outputs of the ARIMA model, although, strictly, these will not satisfy the criterion of 'like-with-like comparison' because of different sets of inputs being used. If not already done, the authors may apply the Model(t+0), as suggested above, and justify the choice of the model forms finally selected for the study.

In the above context, it may also be noted that, although the performance of a single forecasting model in non-updating mode is generally expected to gradually decrease with the increase in lead-time of forecast, such a trend is not obvious from the values of error measures in Table 1 and 2, only because each ANN model, finally selected for the study, is unique. A meaningful comparison is therefore not possible.

vi) The authors have based their selection of antecedent input flows and rainfalls on the basis of the ACF and CCF values. I suggest that for the discharge data series, the authors also provide graphical display of the PACF (Partial auto-correlation function) values to give a better idea of an appropriate ARIMA model.

vii) It is desired that, for each ANN model, the result of the 'experiment with a trial-and-error measure' (17/128), used to decide about the number of neurons in the hidden layer, be graphically presented to show the relative change in the error measure with the number of hidden layer neurons. This is necessary to justify the authors' choice of 12, 15 and 20 neurons in the hidden layer for the models finally selected for 1- 2- and 3-day ahead forecasts. In this context, authors may refer to Fig. 2 in Toth et al., 2000.

viii) The authors have used an adaptive learning algorithm for training. It would be worth, for the sake of completeness, to include a comparison of the training phase considering a fixed learning rate and a fixed momentum term (e.g. 0.5 and 0.5). For this purpose, the comparison may be drawn in terms of either the number of epochs or the time taken in training a network by both non-adaptive and adaptive procedures.

ix) The application reported in the paper is for reservoir flow forecasting for 'operational planning and scheduling of hydroelectric power system' involving reservoirs, as distinct from an application for flood forecasting flood. For this purpose, the models, developed in the current study, will be expected to be reasonably good in forecasting flows across a wide range of flow variability, i.e. for high, medium and low flows. Although, it is recognized that no model can successfully simulate both the high and low flows, some indication of the degree of fit of the simulated flows with the observed flows in different ranges of flows will be relevant to the study reported in the paper. The global values of error measures, as given in Tables 1 and 2, are not very useful. It is suggested that the authors produce additional values of the error measures separately either for each of the high, medium and low ranges of flow or for each of the four parts divided by three quartiles of the observed discharge series. Graphical displays of the selected error measures may be useful. In this context, authors may refer to Fig. 3 in Toth et al., 2000.

Also, the authors may agree that Figs. 3 and 4 in the paper, meant to visually display the degree of fit of the observed and the simulated flows for three lead-times, fail to serve the purpose. For easy visual comparison, the scales for both the observed and simulated discharges in these figures should be the same. Even with the modification of the scales, the plot area in such a figure will be too small to justify the inclusion of data for the whole calibration or validation period. It is suggested that, for each of the calibration and validation periods, plots showing the degree of fit of the highest flow, the second highest flow, a flow in the middle of the range of flows and a low flow, including, in each case, a few days before and after the occurrence of the particular

HESSD

6, S132–S141, 2009

Interactive Comment

observed flow considered, be presented for each lead-time. In this context, authors may refer to Figs. 3 and 4 in Goswami et al., 2005. Although it is expected that, for each lead time, the simulated flows may not be able to reproduce the corresponding observed flows in magnitude and time of occurrence of the observed flows, these plots will give a better indication of the degree of fit across the whole range of flows. The author's statement: 'The simulated curves in both Figs. 3 and 4 clearly indicate that not only the rising trends and the falling trends in the hydrograph are picked up by Model(t+0), Model(t+1) and Model(t+2) but also excellent goodness of fit performances are achieved' (26-27/131 and 1-2/132) is very bold and is not enough. Similarly, the statement: 'From the scatter diagrams in Fig. 5, it is obviously that both of the low values and the high values are close to the exact fit line and this result suggests that there is no evident overestimate or underestimate occurs during the simulation' (11-13/132), in addition to being grammatically wrong, is inappropriate.

x) The authors must highlight the sources of uncertainties in the inflow forecasts being produced by the selected models. Particularly the uncertainty associated with the estimation of the QPFs may considerably influence the uncertainties in the inflow forecasts.

3. Organization of the paper:

i) The periods of calibration and validation have been specified in the 'Results and discussions' section (17-18/131). These specifications should be included in the section describing the methodology.

ii) The structure of the 'bench-mark' ARIMA model has been briefly provided in the 'Results and discussions' section (19-20/132). However, a separate subsection under section 3: 'Methodology and modelling', providing the basic description of an ARIMA model, the structure of the model finally selected for the study and the mode of application of this model for lead-time forecasting in the current study, will be desirable.

iii) It may be more appropriate to provide the details on 'Software implementation' in an

Appendix.

4. Some specific comments:

i) The acronyms, i.e. QPF, NWP, MLP-ANNs and ARIMA, may be avoided in the Abstract. Rather, acronyms are to be provided in the main body of the paper with their first appearance provided alongside the respective expanded form.

ii) In the 'Introduction' section, the authors have stated that '...the black-box models...are widely applied to forecast the streamflow because of their requirement of little data...' (12-15/123). This is not quite right. Whereas the physically distributed and some conceptual models of the mechanistic category of hydrological models require large data of different types, all hydrological models, irrespective of these being mechanistic or black-box, require sufficient data of relevant hydrological variables for effective calibration and for being reliable in forecasting applications. ANNs which are non-linear black-box models also require adequate (and not 'limited') data for being trained.

iii) The references in 19-26/123 and elsewhere may be provided in chronological, rather than in alphabetical, order.

iv) The caption of section 2 may be changed to 'Study area and data characteristics'. The study area in section 2 describes the Fujian province in great details and includes reference to all eight 'hydrographic basins'. Some of these details, which are not relevant to the study, may be dropped. Instead, as indicated in Sl. No. (i) under 2: Shortcomings, more details specific to the Minijiang River and the reservoirs and hydropower plants upstream of the Shuikou reservoir should be provided. No information is available either about the rainfall stations in the study area or about the rainfall data used in the study. These details are required. Some statistics of the series of observed rainfall and discharge data and a plot of the seasonality of these variables, as indicated in Sl. No. (iii) under 2: Shortcomings, may be included for facilitating a better understanding of the data characteristics. Information about the data for the period 1990-2000 (11-14/126), which have not been used in the study, is superfluous

HESSD

6, S132–S141, 2009

Interactive Comment

and may be dropped.

v) Figures 1 to 5 require improvements. In addition to the comments in Sl. No. (ix) under 2: Shortcomings for the plots in Figs. 3 and 4, the texts in Fig. 1 and Figs. 3-5 are too small to be intelligible. Authors must take care to see that the texts in these plots are made sufficiently large to be readable in the journal format. In Fig. 1, the names of the country and the province in the respective map and an index to indicate the rivers and reservoirs, together with their captions, will be required.

vi) 'Hydropower power' (20/124)? Not sure about the inclusion of 'ecological destruction' in 22/124 as one of the justifiable motivations for laying more emphasis on hydropower generation in comparison with other conventional forms of energy production!

vii) All symbols in page 129 are to be defined. For example, 'Delta'E(k), E(k), 'Delta'wji(k+1), 'Delta'qji(k+1), dpj etc. have not been defined.

viii) Which two 'existed' systems in 15/134?

ix) What is 'approbatory simulation' in 29/134?

x) As stated in the last sentence of 1: General Comments, the language needs improvement. Some examples are as follows. The use of the term 'severed' in 16/122 or in 17/135 is not appropriate. Replace 'relationship' in 18/130 and 1/131 by 'degree of fit', 'dynamic, uncertain, and nonlinear' in 23/122 by 'dynamism, uncertainty and non-linearity' and '...hydropower is strongly advocated by the...' in 22-23/124 by '...the development of hydropower is strongly advocated in the...'. Drop 'So' in 1/126. Replace 'statistic' in 6/125 by 'statistical, 'approximate' in 22/125 by 'approximately', 'downriver' in 11/126 by 'downstream', 'generational' in 17/126 by 'generation', 'revealing' in 23/128 by 'establishing', 'obviously' in 11/132 by 'obvious', 'affection' in 28/132 by 'effect', 'two existed system' in 15/133 by 'two existing systems', 'By the aid' in 20/133 by 'With the aid', 'accomplished' in 21/133 by 'obtained', 'popularly' in 11/134 by 'popular', 'input' in

HESSD

6, S132–S141, 2009

Interactive Comment

27/134 by 'inputting' and so on.... Replace lines 20-21/126 by 'The application of the models, developed in the study, will be illustrated by considering the Shuikou reservoir which is the most important reservoir in the Fujian province'. The term 'general quantitative scope' in 3/132 is inappropriate.

5. References:

i) Goswami, M., O'Connor, K.M, Bhattarai, K.P., Shamseldin, A.Y., 2005. Assessing the performance of eight real-time updating models and procedures for the Brosna River. Hydrology of Earth System Sciences, 9(4), 394-411.

ii) Toth, E, Brath, A., Montanari, A., 2000. Comparison of short-term rainfall prediction models for real-time flood forecasting. J. Hydrology, 239, 132-147.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 6, 121, 2009.