Hydrol. Earth Syst. Sci. Discuss., 2, S664–S672, 2005
www.copernicus.org/EGU/hess/hessd/2/S664/
European Geosciences Union


Hydrology and Earth System Sciences Discussions



# *Interactive comment on* "Water level forecasting through fuzzy logic and artificial neural network approaches" *by* S. Alvisi et al.

**S. Alvisi et al.**

Received and published: 15 September 2005

1. C: The authors list two different possible neural network training techniques but fail to state which one was used.
R: The Levenberg Marquardt algorithm was used to train the artificial neural network with the ARI input data set (which is characterised by a low number of inputs), whereas the scaled conjugate gradient algorithm was used to train the artificial neural network with the DRI input data set (which is characterised by a high number of inputs). This information is added in the new version of the paper:

p.12, line 16
old version: "...and 1 in the output layer."
new version: "...and 1 in the output layer. The Levenberg Marquardt algorithm is used for training the model since it is characterised by a low number of inputs."

p.18, line 18

old version: “...and 1 in the output layer, while the FL-M...”
new version: “...and 1 in the output layer; the scaled conjugate gradient algorithm is used for training the model, since it is now characterised by a high number of inputs. The FL-M and FL-TS models remain...”

2. C: Full details of the neural network training programmes and training parameters should also be provided in the text.
R: As previously stated, the artificial neural network models were trained by using the Levenberg Marquardt and the scaled conjugate gradient techniques. The software used is available in the Matlab Toolbox. No description of these techniques is given since they are well known and established, and the reader can always refer to the huge amount of scientific and didactic literature part of which is quoted in the submitted paper. In the new version of the paper only the information on the use of the Matlab Toolbox is added:

p.6, line 22
old version: “...thus the training procedure is stopped.”
new version: “...thus the training procedure is stopped. The artificial neural network model was implemented in Matlab environment where both Levenberg Marquardt and scaled conjugate gradient training techniques are available in the Neural Network Toolbox.”

p.8, line 25
“The fuzzy logic models were implemented by using the FORTRAN programming language.”

3. C: The fact that 12 hidden neurons produced the best result for both sets of neural network inputs is of hydrological interest. Model selection was based on trail and error but the question that arises is to what extent were the different models different? Most neural network solutions are rather similar and it is often argued that parsimonious solutions should be selected in favour of models that are selected on the basis of minor

or inconsequential differences that were obtained using traditional measurement statistics. It would therefore be helpful if this paper reported the number of hidden neurons that had been tested in the trail and error model testing operations and provided some numerical statistics on the range of different output results that had been obtained.

R: We agree with the referee that most neural network solutions remain quite similar to each other when the number of hidden neurons are changed slightly. Indeed, in our application, several numbers of hidden neurons were tested, ranging from 5 to 20, and very similar RMSEs and R2 statistics were obtained. For example, for the DRI input data set in the validation phase, the RMSE for 1 hour ahead forecasting was 14.6 mm for 5 hidden neurons, 11.9 mm for 20 hidden neurons and 11.5 mm for 12 hidden neurons, and, for 6 hours ahead forecasting, the RMSE was 29.2 mm for 5 hidden neurons, 30.3 mm for 20 hidden neurons and 28.3 mm for 12 hidden neurons; similarly the R2 statistic for 1 hour ahead forecasting was 0.93 for 5 hidden neurons, and 0.95 for both 20 and 12 hidden neurons and, for 6 hours ahead forecasting, the R2 was 0.72 for 5 hidden neurons, 0.70 mm for 20 hidden neurons and 0.74 mm for 12 hidden neurons. The solution with 12 hidden neurons was selected since we considered it a good compromise between the (lowest) number of hidden neurons and the (highest) model efficiency, although no great differences were observed between the several combinations considered, as is clear from the statistics previously reported. A similar procedure was used to define the number of rules of the fuzzy models, varying them between a minimum and a maximum number and thus selecting that which provided the best values of RMSEs and R2 statistics. This procedure to define the number of hidden neurons (ANN) and number of rules (FL) is widely, if not systematically, adopted in the literature and obviously contains some dose of subjectivity, especially when very similar results are obtained for different values of the controlling parameters (i.e. number of hidden neurons and number of rules). However, both in the old version and in the new version, the authors avoid giving all the details of this process since its logic is well known and applied by all the researchers who use ANN and FL models, Accordingly, the text was only slightly modified:



HESSD

2, S664–S672, 2005

Interactive Comment

p.12, line 15
old version: “...(number obtained by trial and error procedure)”
new version: “...(number obtained by trial and error procedure, ranging between 5 and 20 neurons)”

p.18, line 14
old version: “...(number obtained by trial and error procedure)”
new version: “...(number obtained by trial and error procedure, ranging between 5 and 20 neurons)”

4. C: Is there a scientific reason for incorporating two different non-linear transfer functions?
R: The combination of one non-linear transfer function in the hidden layer with one linear transfer function in the output layer was also tested and it produced fairly good results, though slightly worse than those obtained with two non-linear transfer functions. For this reason, two different non-linear transfer functions are used in this application. A comment about this is added to the new version of the paper, as follow:

p.5, line 18
old version: “...vectors WPp+bp and WHh+bh. In order to avoid...”
new version: “...vectors WPp+bp and WHh+bh. Other different transfer functions were tested, but the attention was focused on functions (3) and (4) since they produce better results and show high flexibility without increasing model parameterization. In order to avoid...”

5. C: The datasets were ‘standardized’ not ‘normalized’?
R: In literature both the terms are used to describe the rescaling procedure of the inputs to a range (0.1-0.9 or 0.05-0.95). For example, in Campolo et al., 1999, 2003; or See and Openshaw, 1999 the term ’normalized’ is used whereas in Hsu et al., 1995 or Imrie et al.2000, the term ’standardized’ is used. However, the term ’normalized’ (p.5, line 20) is replaced with ’standardized’ in the new version of the paper.

6. C: Fig 1: division of the catchment into two-sub basins is problematic since the dashed divider appears to possess no hydrological relationship to the drainage network? Is it sub-regions not sub-basins?
R: We agree with the referee that the proper term is “sub-region”. In the paper, all the references to the ‘sub-basins’ (p.11, line 10 and p.18, line 8) are changed to ‘sub-regions’ and a more detailed caption to Fig. 1 is given in the new version of the paper:

p.26
old version: “Figure 1. The upper basin of Reno river and position of the rain-gauges considered.”
new version: “Figure 1. The upper basin of Reno river and position of the rain-gauges considered. The dashed line subdivides the basin into two sub-regions used to define the DRI input data set.”

7. C: I wonder if positive and negative water level variations should be encoded into the same ’input stream’ and whether or not traditional neural network standardization protocols are able to cope with such differences in a meaningful manner? Two separate input streams might perhaps be much better?
R: The results obtained confirm that an artificial neural network model is able to cope with positive and negative water level variations encoded together, even using a traditional standardization protocols: clearly negative values of the water level variations are “shifted” into “a lower zone” of the “standardised” range, while the positive values are “shifted” into an “upper zone” of the “standardised” range and subsequently “filtered” by the ANN model in a proper way. If two separate input streams were used for each pattern, one of the two water level variation input would be zero when the other is not. This type of input is odd in our opinion and we thought it unnecessary to test it also because in our paper, the same information is inputted both to the ANN and the FL models and these latter models do not use any standardization protocols. Furthermore, the FL approach emulates the way of human thinking and in that context it is absolutely “natural” to consider the level variations as a single information which

can be both positive and negative. Accordingly, no modifications are made in the new version of the paper with reference to this comment.

8. C: The models were developed on three independent subsets; training, validation and testing. Each subset must contain a comprehensive representation of the different processes that are to be modelled and to ensure that this condition is fulfilled it is essential that neural network papers provide a detailed statistical description of each subset with respect to the different variables that are involved. The authors must record the number of patterns as well as the number of events in their respective subsets.
R: A new table (Table 2) is added in the new version of the paper reporting the main characteristics of each subset and the text is modified:

p.9, line 12
old version: "...this set is split into 2 subsets: 80% of the data are used for training and 20% for validation. A set of 7 flood events observed in the year 2000 is selected and used for testing the models."
new version: "...this set is split into 2 subsets: 36 flood events (observed in the period 10/1994-12/1999) are used for training the model and 9 flood events (observed in the period 09/1993-09/1994) for its validation. A set of 7 flood events observed in the year 2000 is selected and used for testing the models. The main characteristics of training, validation and testing sets are summarized in Table 2."

Table 2. Main characteristics of training, validation and testing sets. dW is the hourly water level variation, Pmax is the maximum hourly areal rainfall and Pcmax is the maximum 12 hour cumulated areal rainfall.



HESSD

2, S664–S672, 2005



Full Screen / Esc

Print Version

Interactive Discussion

Discussion Paper

EGU

| Phase | n° of events | n° of patterns | Peak flood (cm) | | dW (cm) | | Pmax (mm) | Pcmax (mm) |
|---|---|---|---|---|---|---|---|---|
| | | | min | max | min | max | | |
| Training (10/94-12/99) | 36 | 1549 | 54 | 357 | -46 | 119 | 15.20 | 66.09 |
| Validation (09/93-09/94) | 9 | 386 | 79 | 250 | -33 | 151 | 12.15 | 63.25 |
| Testing (03/00-12/00) | 7 | 343 | 80 | 317 | -42 | 126 | 14.43 | 73.25 |

9. C: How were the events selected? What constituted the start and end points?
R: The new version of the paper contains this clarification:

p.9, line 14
old version: "Each event considered is "complete", in the sense that each flood event is described in its complete evolution."
new version: "The events selected are all characterised by a water level peak equal to or greater than 50 cm above the reference zero of the measuring station at Casalecchio di Reno. The initial and the final points of the flood events were marked in the water level time series where the periodic oscillations, due to anthropic activities, terminate and begin again respectively."

10. C: The authors have not provided hydrometeorological particulars or morphological descriptions with regard to the nature and pertinent characteristics of this catchment. Their findings cannot in consequence be put into a hydrological context.
R: Some further hydrometeorological and morphological information is given in the new version of the paper by inserting a new table (Table 1). However, it is worth recalling that the models considered are data-driven and thus are "blind" with respect to specific physical information: they only deal with data (e.g. rainfall and levels) which indirectly contain the "integral" behaviour of the system to be modelled. Accordingly, the new

version of the paper is modified thus:

p.9, line 5
old version: “...of the basin is about 12 hours.”
new version: “...of the basin is about 12 hours. Further hydrometeorological and morphological information is reported in Table 1”

p.22, line 10
old version: “In order to confirm these preliminary results, similar analyses are currently being developed with reference to catchments with different hydrological and morphological characteristics. This aspect is under investigation and the results will be presented in due course.”
new version: “Finally, it is worth stressing that the aim of this study was the comparison between two different data-driven approaches for water level forecasting when equal input sets are used. These models, for their very nature, are “blind” with respect to a specific physical information: they only deal with data (e.g. rainfall and levels) which indirectly contain the “integral” behaviour of the system to be modelled. As a consequence, since the presented comparison was performed with reference to a basin of a humid region, it is expected that similar results (in term of comparison between the two approaches) can be obtained when other basins located in similar regions are considered. However, when basins located in dry regions are considered, different inputs might be selected and thus different structures for both the models (ANN and FL) might be identified, so the relative performances of the two models (in term of comparison) might change. Thus, further analyses are currently being developed with reference to catchments in dry regions. The results will be presented in due course.”

Table 1. Hydrometeorological and morphological information about the Reno river basin at Casalecchio di Reno (Bologna, Italy). Qpeak is the mean value of the annual maximum peak discharges, Hm is the mean altitude of the basin respect to the basin outlet, L is the length of the Reno river upstream from Casalecchio di Reno, Tc is the time of concentration, S is the mean bed slope and Pyear is the yearly average areal

| Basin | Qpeak (m3/s) | Area (km2) | Hm (m) | L (km) | Tc (h) | S (%) | Pyear (mm) |
|---|---|---|---|---|---|---|---|
| Reno at Casalecchio | 767.71 | 1051 | 581 | 84.2 | 12 | 1.18 | 1336.00 |

rainfall depth.

Bibliography quoted in the reply to the referee:
Campolo, M., Andreussi, P. and Soldati, A.: River flood forecasting with neural network model, Water Resour. Res., 35(4), 1191-1197, 1999.
Campolo, M., Andreussi, P. and Soldati, A.: Artificial neural network approach to flood forecasting in the river Arno, Hydrol. Sci. J., 48(3), 381-398, 2003.
See, L. and Openshaw, S.: Applying soft computing approaches to river level forecasting, Hydrol. Sci. J., 44(5), 763-778, 1999.
Hsu, K. L., Gupta, H. V. and Sorooshian S.: Artificial neural network modeling of the rainfall-runoff process, Water Resour. Res., 31(10), 2517-2530, 1995.
Imrie, C. E., Durucan, S. and Korre, A.: River flow prediction using artificial neural networks: generalizations beyond the calibration range, J. Hydrol., 233, 138-153, 2000.

Interactive comment on Hydrology and Earth System Sciences Discussions, 2, 1107, 2005.