Hydrol. Earth Syst. Sci. Discuss., 5, S181–S184, 2008
www.hydrol-earth-syst-sci-discuss.net/5/S181/2008/


Hydrology and Earth System Sciences Discussions



# *Interactive comment on* "An artificial neural network model for rainfall forecasting in Bangkok, Thailand" *by* N. Q. Hung et al.

**Anonymous Referee #3**

Received and published: 7 April 2008

Elena Toth, University of Bologna, Italy

General comments

The paper describes an application of artificial neural networks (ANN) for the real-time forecasting of rainfall values up to 6 hours ahead. Reliable short-term rainfall forecasts are crucial for the implementation of real-time flood forecasting systems: therefore the topic is certainly suitable for the journal and of broad international interest, not only for researchers but, more importantly, for operational hydrologist and decision makers. In addition, the case study, referring to a real-world application on the Bangkok metropolitan area, is of extreme interest. The methods are correct and sound and the good results obtained by the authors (with model F) confirm that the implemented ANN



Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

are an adequate tool for short-term forecasts. The English should be carefully revised in sections 4 and 5.

I agree with the comment made by Referee #2 that the main weak point of the paper is how it deals with continuous against rainy data: there is a strong need to add a detailed description of the model implementation with and without rainy days. This point is in fact indicated as crucial (and innovative) at the beginning of the paper, but it is not adequately developed in the description of the methods.

Secondly, there is confusion in the description of training and testing data: it is not clear if the year 1998 is part of the first or of the second set and it seems that the validation data are different in the exploratory test (over station E18, section 5.1 and 5.2) and in the test over the whole raingage network (section 5.3) .

In addition, the number of nodes in the hidden layers is very large, especially in the more complex models, and this aspect deserves some comments from the Authors.

Lastly, I would suggest to add a comparison of the results of the models against the use of a simple persistent model, in order to understand the real advantage provided by the models.

Specific comments

p.185 (l. 27)-186 (l. 16): I suggest to divide the literature referring to climatologic forecasting (monthly and seasonal periods, which ma be not strictly needed here, by the way) from that referring to short-term rainfall forecasting.

p. 189 (l. 22): the FFC does not make any use of the SCOUT radar data, but uses only raingauge data for issuing flood forecasts?

p. 190, l. 4: for which period the data were collected?

p. 190, l.7-9: the meteorological data were collected in the same (and all) stations of the raingauge network?



HESSD

5, S181–S184, 2008



Section 3, p. 190-193: there is confusion in the description of the MultiLayerPerceptron (MLP) and the BackPropagation (BP) algorithm, which are not the same thing. In addition, it is not clear which training algorithm was actually used in the applications. The description of the BP (from p191, l. 8 to p. 192, l. 13) should be moved to the end of the section, specifying that it is the one that was used in the study (if it is so…)

p. 194, ll. 11-13: year 1998 is used in training or in validation?

p. 195, ll. 5-8: it is not clear how the Authors dealt with the problem of no-rain periods in models A, B and C…

p. 194-195: the number of hidden nodes seems too large, thus needing an extremely high number of parameters: in model F there are 42 nodes, resulting in 33 biases values and 440 weigths values, if considering only feed-forward connections (and here a generalised model is used, with more connections), to be parameterised. I believe that an Authors’ comment on this point is needed.

Section 5: this section is too detailed (and long) in the description of the “numbers”, that is of the indexes of performance that are all already reported in the tables, whereas it lacks an interpretation of the results.

p. 198: ll. 19-27: this paragraph is far from clear and lacking: an interpretation of the results obtained for the testing data (which are often not consistent with those of the training data) is needed. In addition there are no comments on the results obtained when excluding the rainfall data from station E18 itself (and such results are not presented in Table 3 neither). Furthermore, the analysis of the role of surrounding stations is made excluding all of them and not one by one, which may have given some insights on their relative importance. Lastly, the results obtained excluding the surrounding stations are strongly dependent on the chosen station, therefore, this sensitivity analysis can not be extended to the entire raingauge network. A new sensitivity analysis using the results obtained with model F over the entire network would be more significant.



HESSD

5, S181–S184, 2008

Interactive Comment



p. 199, ll. 4-7: please explain better the cross-validation experiment (how are the 3-years data used?)and why it is different from the validation performed (in the above sections) on the data of station E18 alone.

p. 199, l1-p. 200, l.4: this paragraph is not clear: the RMSE is already a mean value, why citing the number of rainfall observations (patterns)? In addition, please rephrase ll 3-4.

Technical suggestions

p. 189 (l. 19): it may be useful to specify the density of the raingage network (number of operative stations at hourly resolution for km2) p. 193: l- 16-18: this phrase is not clear to me. p. 194, l. 23-25: specify the number of hidden nodes in model C (10+10) p. 195: title of section 5.1: specify that it refers to 1-hour ahead forecasts for one station only p. 199: title of section 5.3: specify that it refers to 1 to 6-hours ahead forecasts for the entire network

Table 1: add in the Input description of Model F that the data from surrounding stations consists of present rainfall data. Tables 2 and 3: add in the caption to which forecast experiment (station E18) they refer. Figures 4, 6 and 8: a period of 24 hours (and with so few rainy hours!) is too short: you may substitute these hyetographs with those referring to a longer period (with 48 hours the graph should still be readable).

---

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 5, 183, 2008.



HESSD

5, S181–S184, 2008

Interactive Comment