Hydrol. Earth Syst. Sci. Discuss., 5, S401–S406, 2008
www.hydrol-earth-syst-sci-discuss.net/5/S401/2008/

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.

N. Q. Hung et al.

Received and published: 20 May 2008

The authors would like to thank the Anonymous Referee #2 for carefully reviewing the paper and providing very constructive comments. This helped us immensely in improving the quality of the paper. We have addressed his/her comments as follows:

General comment: This paper presents a rainfall forecasting result using an artificial neural network model applied to Bangkok in Thailand. Although a new ANN model is proposed, which is trained using consecutive rainfall data including non-rainy days, it is not clear how the ANN model is mathematically improved compared to an existing ANN framework. The manuscript does not include evaluations of the new ANN model compared to the conventional approach that precipitation of only rainy-days is used to train the ANN model. As a scientific paper to evaluate the ANN model performance, it needs in-depth analysis and discussion. I would suggest including discussion that how

Full Screen / Esc
Printer-friendly Version
Interactive Discussion
Discussion Paper

model performance would be improved when non-rainy days are included for training the ANN model. RMSE values in Tables 2, 3 and 4 would be reevaluated only for rainy-days. Also the forecasting results appeared in Figure 10 would be evaluated with some indices such as RMSE to assess matching of spatial distributions at each lead time so that a reader can easily understand the performance of the model.

Response: As summarized in the introduction of the paper, most of the existing ANN models applied in rainfall forecasting were trained using data that contained only rainy period. Mathematically, ANN learning provides a method for learning functions over instances which is described by a group of input attributes. For example, the ANN model in this paper uses attributes which are not only rainfall values at the forecast point but a combination of meteorological parameters and rainfall values from surrounding stations. Results can be improved if the training data is distributed over the entire space of instances so that ANN can approximate target functions better. By using only the data from the rainy period as training data set, ANN models could easily identify the patterns characterizing the rainfall; however, on the other hand any features or characteristics not included within the training data will not be learned by ANN. Translated this means that conventional ANN models could only provide highly accurate forecasting when it rains. This paper presents an ANN model well suited to run a real time task in which rainfall can be forecasted at any point of time. In this situation, the model needs to be trained with not only on rainy data but consecutive data including non-rainy periods in order to acquire a fully representation of both rain and no rain conditions. To satisfy the comment and support our conclusion, another ANN model with the same set up as the final mode (model F in the paper) was trained with rainy period only. Results show that such training model provides high accuracy on testing data containing rainy periods only, but gives very poor forecast result on one month continuous data (i.e. August 1998). Evaluation and discussion of this model was updated in the final manuscript. Because the ANN models were trained in order to run a real time forecast so to evaluate the model performance, we need not only to check the rainy period but also to check if models provide a forecast at right time or not. If we re-evaluate the

HESSD

5, S401–S406, 2008

model with rainy period data only, we may fail to spot the point where model may give a erroneous or missed forecast. Therefore, we would like to keep the Table. 2, 3, and 4 unchanged. Figure 10 which illustrates a series of observed and forecasted interpolating rain maps, with the purpose to create a better visualization for eyes checking on the model performance. The RMSE value for each lead-time was calculated and added in the discussion part as suggested.

Specific comments

1. P. 190, 21th line, "Artificial Neural Network" section: Need description how a newly proposed ANN model using non-rainy days for model training is mathematically different from the existing ANN models.

Response: In the section 3: Artificial Neural Network briefly presented two network types and the Back Propagation algorithm which were used in our work. The final ANN model in this paper uses Generalized Feedforward network type with two hidden layers, hyperbolic tangent activation function, trained with Back Propagation algorithm, which are common set up among other ANN models. However, the difference is illustrated in the training data, which is not only rainfall at the forecast point but a combination of meteorological parameters and rainfall from surrounding stations. Section was revised an update to manuscript as commented.

2. P. 191, 18th line: Need definition of di, i and j.

Response: The term di, as explained already in the paper (p191, 18th line), is the system response, i and j indicate for PE number i and PE number j in different layer. The manuscript, as mentioned, has been updated.

3. P. 191, 24th line: Need definition of xi(n) in Eq. (2).

Response: x(i)n is the transfer function at PEi. Manuscript has been update based on this comment.

4. P. 192, 4th line: The last term in Eq. (3) would be wij(n)?

HESSD

5, S401–S406, 2008

Response: Corrected and manuscript updated. 5. P. 192, 10th line: Need definition of netj and θj in Eq. (4).

Response: Netj is the output of the neurons, and  is a bias. Manuscript updated as to fulfill the suggestions made.

6. P. 195, 25th line, "Results and Discussion" section: The section does not include evaluations of the new ANN model compared to existing ANN models which use only rainy-days data. Discussion on values of the new ANN model with quantitative evaluations should be presented. The output of ANN formulation in the paper is one-hour-ahead rainfall. The lead time is short, so to develop an ANN model only trained with rainy-day would be enough or rather than better in practical forecasting. The authors should show the strong points of a newly proposed ANN formulation.

Response: As explained in the answer for general comment, the ANN presented in our paper was employed to run a real time task hourly, and it has been applied to run forecast at BMA since 2005. It is actually in different class with other existing ANN models which trained with rainfall periods only. So, in the open discussion version of the paper, we did not include the comparison between new ANN model with others existing models. Since the comment requested to make a clear comparison, manuscript has been updated. Discussion with quantitative rainfall value was also presented in revised version of paper. The lead time in the exploratory test (or designing stage) was 1 hour, but the forecast time is from 1 to 6 hours ahead for the final ANN model. In this case, when the model is expected to run forecast at hourly rate, it must be trained with continuous data, not with rainy-day data.

7. P. 196, 1st line and Table 2: If the indices values in Table 2 include non-rainy-days, it is requested to make a similar table that excludes non-rainy-days to evaluate the model appropriately.

8. P. 196, 5th to 7th lines: "This value however ... of training data." RMSE values only for rainy-days are requested to evaluate the model performance.

HESSD

5, S401–S406, 2008

Interactive Comment

Response: Updated in the Rainfall forecasting section.

9. P. 197, 10th line: "overestimation"? Underestimation?

Response: It is "underestimation" and was corrected to manuscript.

10. P. 198, 19st line and Table 3: If the indices values in Table 3 include non-rainy-days, it is requested to make a similar table that excludes non-rainy-days to evaluate the model appropriately.

11. P. 199, 8th line and Table 4: If the indices values in Table 4 include non-rainy-days, it is requested to make a similar table that excludes non-rainy-days to evaluate the model appropriately.

Response: For the comment 7, 10 and 11: Please refer to the answer to general comment where we explained our rationale for keeping table 2. 3 and 4 unchanged. Discussions and RMSE value for the rainy period (in both observed and forecasted data) were included in the discussion section of the revised manuscript.

12. P. 200, 6th line and Figure 10: The forecasting results appeared in Figure 10 would be evaluated with some indices such as RMSE to assess matching spatial distributions at each lead time so that a reader can easily understand the performance of the model.

Response: The RMSE value for each lead-time was calculated and added in the discussion part as suggested.

13. P. 201, 28th line: "Based on ... " The authors would need to show more evidence to conclude the proposed ANN model is "an appropriate predictor".

Response: More discussion points on the results produces by the ANN model and new figures showing a comparison between observed and forecasted rainfall over Bangkok area from one to six hours was included in the manuscript.

Technical corrections

HESSD

5, S401–S406, 2008

P. 193, 26th line: Coulibaly et al.

Response: Corrected and updated in the manuscript

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