Hydrology and Earth System Sciences Discussions

Interactive comment on "On-line multistep-ahead inundation depth forecasts by recurrent NARX networks" *by* H.-Y. Shen and L.-C. Chang

Anonymous Referee #3

Received and published: 4 December 2012

General comments:

The authors report results of a study to predict inundation depths for an urban flood application with artificial neural networks. I agree with the authors that the ANN technique can be a feasible method for on-line forecast applications due to the computational time requirements of conventional models. The authors considered 3 network configurations T-NARX (ideal case, assumes all input are known), O-NARX and R-NARX (practical case). The authors used results from a flow simulation model, generated for 24 design and 2 actual storms for training the model. The ANN model was tested against simulation results for 29 actual storm events. Model performance was presented in terms of the rmse error statistic. The overall results are consistent with expectations.

Especially welcome is the finding that R-NARX results are not too far from T-NARX results and especially that R-NARX are better than O-NARX results. I therefore would like to recommend acceptance of this paper.

Specific Comments:

The use of rainfall as input is interesting, since the authors opt to choose an NARX model. How significant is the effect of rainfall used as an input to the model? Does rainfall have as much an impact on the model on short term forecasts compared to longer term forecast?

Page 12005 top line states: ". . . fed back to the input layer in both training and testing phases . . .". I may have misunderstood this sentence but is it really "both training and testing phases" for the O-NARX model? It seems that outputs are only fed back in the testing phase for O-NARX.

The better results obtained by the R-NARX model over the O-NARX model is an interesting result, although it seems to me to be counter-intuitive. The authors attempt at providing a reason for this in page 12012, line 5 onwards. Further elaboration on this point will be helpful. For example, why does the R-NARX exhibit smaller error accumulation and propagation compared to O-NARX?

The authors use RMSE which is calculated on an entire time series as a measure of model performance. While the rmse statistic is useful, in flood situations, the parameters of real practical concern are time to peak, peak inundation depths and time at which cessation begins. The authors should consider providing some error statistics in terms of these parameters so that further evaluation on the model performance can be made.

Technical comments:

Is Figure 6 a repetition of the data provided in Table 4? Suggest to delete Figure 6 if it is.