Nat. Hazards Earth Syst. Sci. Discuss., 1, C2364–C2369, 2014
www.nat-hazards-earth-syst-sci-discuss.net/1/C2364/2014/

Natural Hazards and Earth System Sciences
Open Access
Discussions

Interactive comment on "Storm-surge prediction at the Tanshui estuary: development model for maximum storm surges" *by* C.-P. Tsai et al.

J.W. de Vries (Referee)

hans.de.vries@knmi.nl

Received and published: 22 January 2014

The manuscript describes an interesting method to calculate the height of storm surges, applied to the Tanshui estuary in Taiwan. The authors show that the method has potential, although they do not (yet) apply it in forecast mode.

The work is certainly interesting enough to be published, but the manuscript still needs work to make it more clear and precise.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

General comments

1. The reader is not well guided through the different experiments. It would greatly help to have a short outlook in the introduction which describes the aim of the paper: develop the neural networks and compare their results, also with the emperical formula by Horikawa.

2. The different models are generated from the complete dataset, and afterwards verified with the same dataset. As the number of hidden layers in the neural networks is apparently chosen to give the best results, this means that these models are just fitted to these particular data and no conclusions can be drawn on the general validity of the models for storm surge calculation.

   The models would be much more valuable if they were generated from and tested on separate partitions of the dataset. Alternatively, the results could be checked for sensitivity to leaving out parts of the dataset in the generation step.

3. In the manuscript, the word 'predict' is commonly used to denote the calculation of storm surges. However, it seems that no prediction (I would prefer forecast) for storm surges has yet been made. It would be worthwhile to dedicate a paragraph on the possible use and application of the methods described.

4. The list of references should be carefully checked, also their use in the text. I noted several errors, but did not check systematically:

   - Page 2, line 16 (henceforth p2l16): Blainetal
   - Marzenna (2003) should be Sztobryn (2003). The author is Marzenna Sztobryn, Marzenna being the first and Sztobryn the last name.
   - Several references to Cornner, but the list gives Conner.
   - Isozaki (1966) (p9l2) should be Unoki and Isozaki (1966).

NHESSD

1, C2364–C2369, 2014

5. On the figures:
    - Figures 2, 5 and 6 do probably not show predicted vs observed correlation coefficients, whatever that may be, but, I guess, predicted vs observed storm surge. If that is the case, the axes should have units as well.
    - For the other figures, it would be helpful (and reduce the number of figures) to combine them in fewer figures, e.g. 3 separate figures for models A, B and C, with the forecasts from MLP and RBF, together with the empirical formula.
    - Actually, the different types of figures (2, 5, 6 vs the rest) contain the same information. It would not do much harm to leave out 2, 5 and 6 completely.
6. On the title: "... Development Model ..." should probably better be "... Development of a Model ..."
7. Sometimes the authors write storm-surges. Storm surge should not be hyphenated.
8. In a few places the word "believe" is used. I think that is not in place in the natural sciences. People either argue that something is valid or they assume it.

In depth comments

1. p2l5: the use of "previous" makes it unclear. What is meant is, I think, "... based on observed data from 22 historical storms ..."
2. p2l10: The last sentence of the Abstract does not make much sense:
    - "These" does not refer to anything (results were not mentioned earlier, certainly not that they were good)

NHESSD
1, C2364–C2369, 2014

- prediction=forecast here?
- What is meant by “time series data”? Forecasts? Will they be used to build the neural network or to run the neural network in forecast mode? This needs elaboration further on in the text, see my General comment 3 above.

3. p2l24: “than that of the AR model”, should probably be without ‘that of’, and shortly introduce “the AR model” with something like “the AR model which is commonly used for...”.

4. p3l10: ‘overall’ water level, is that maximum water level, or total water level envelope?

5. Sec 2.1: First sentences are puzzling. I had not come across “irregular weather water level” before, and it does not really seem relevant here. The second sentence probably means that storm surges are more interesting when they are higher. And, if storm surges are also(!) related to weather conditions, what else?

6. p4l3: “referenced” might be not correct; with “actual data” probably observations of storm surges and meteorological parameters are meant.

7. p4l5: “Their empirical formula...” needs some more introduction, like “They derived an empirical formula to calculate storm surges”.

8. p4l10: “empirical constants” = parameters fitted to the data?

9. Sec 2.2, 2.3: I am not too familiar with neural networks, so I am a bit puzzled here. But I suspect that unclarities could well make the paper ambiguous, also for those who are more experienced in this technique.

   First of all, I would be helped with the definition of “neuron”, “neural layer” and “neural input signal”. E.g. in p5l19-20, is “jth layer” the same as “jth neuron”?

NHESSD
1, C2364–C2369, 2014

Also, p8l5-7: does this mean that the output variable is directly coupled to the output neuron, is y_j the output neuron?

Furthermore, it should be specified how the X_i are determined. Are they the observations in the input layer? How are they calculated in the hidden layers?

10. Eq. 3, 4, 5: make the summations consistent. Just the subscript of the summation index will do.

11. p6l17: "whether they had caused serious storm surges" might be a little dangerous: when you derive and validate a model on actual high surges only, you should also make sure that it does not generate high surges when there are not.

12. p7l13: "kth estimation" should be "estimate of the kth sample point" (or simply "its estimate").

13. Eq. 9 does not contain $\hat{y}_k$ itself, only their average. Is that correct? Moreover, the average of y_k (etc) should not be $\bar{y}_k$, but simply $\bar{y}$, as it is no longer dependent on k.

14. Sec 3.3 might be better in place as Sec 2.4.

15. p8l9-10: The indication "generalized least squares method" is too vague. Does it mean that A and B were fitted to the data of these typhoons?

16. p8l14: refer to Table 1, as that lists the characteristics of the 22 typhoons.

17. p9l5: "As can be seen..." suggests that the table lists a number of possible configurations with their results. I assume the authors varied the number of hidden layers to get an optimum RMSE and CC, but I could not find that confirmed in the text. And it can definitely not be seen from Table 2.

18. p10l19: "Model D" should be "Model C".

NHESSD

1, C2364–C2369, 2014

19. p10l24-25: I would use “input variable(s)” instead of “factor(s)”.

20. Sec 5: the argument to choose Model B over Model C is very fuzzy. Probably, there are to many aspects squeezed into too few sentences. Although my conclusion was that the method is not (yet) fit for forecasting (see general comment 2), this argument seems to be mainly inspired by the applicability in a forecasting environment. Moreover, the argument of missing data and large errors is not mentioned in Sec 4.3.

21. p11l6-7: It is not clear what is meant here with “the best input factors”. It is unlikely to be “input variables”, as above, because then the qualification ‘best’ does not make sense, as the choice of input variables for Model B is fixed to pressure and wind speed.

Interactive comment on Nat. Hazards Earth Syst. Sci. Discuss., 1, 7333, 2013.