The authors wish to express our deep appreciation to both referees for the constructive comments and suggestions, which have led to the making of several corrections and have greatly aided us to improve the presentation of this paper. We are confident that the overall quality of the manuscript has been improved. Detailed responses are listed below.

Response to Dr. J.W. de Vries' Comments:

For 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 empirical formula by Horikawa.

Response:
We have rewritten the full text. We also replotted all figures. The aim of the paper has been described in the Introduction. The first part of this study is aimed to the development of the neural network model for estimating the maximum storm surge. The second part is new added in the revised version, which is aimed to forecast the time variation of storm surges using both multilayer perception and radial basis function methods using a very few major meteorological factors which was based on the model of the first part. We also have described that "The results trained by the neural networks are compared with the empirical formula proposed by Horikawa (1978)." in the Introduction.

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.

Response:
A section has been added, Section 5, including three illustrative examples forecasting the time variation of storm surges using the trained neural networks. The time series of storm surges for nine typhoon events were selected, of which six were used for training and testing the model, and the other three were used for the model forecasting.

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.

Response:

The forecast of storm surges has been made using a new dataset, as shown in the new Section 5. In addition, we have used "forecast" instead of "prediction" throughout.

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 p2/16): Blainetal
   • Marzenna (2003) should be Sztobryn (2003). The author is Marzenna Szto- bryn, 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).

Response:

The references have been revised.

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.

Response:

We have integrated all old 10 figures into one new figure; see new Figs. 2.

6. On the title: "... Development Model ..." should probably better be "... Development of a Model ..."

The title has been changed as "Development of Models for Maximum and Time Variation of Storm Surges at the Tanshui Estuary".

7. Sometimes the authors write storm‑surges. Storm surge should not be hyphen‑ated.

All such errors have been revised.

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.

The word "believe" has been changed to "proposed".

For "In depth comments":

Minor comments have also been responded to and revisions made.

Response to Dr. Bajo's Comments

For General comments:

1. The English needs to be improved and some sentences are not clear. I suggest a check by a natural English speaker after all the corrections.

We have rewritten the full text and had the English checked by a proof reader of scientific English.

2. In many parts of the manuscript citations are missing (see detailed comments). The references are few and rather old, especially those regarding the storm surge forecast.

Response:

Some recent new references have been added to the revised work.

3. A Figure showing the geographical zone of interest and the location of the stations would be useful to the reader.

Response:

We have added a new figure to show the location of the station and the typhoon tracks; see Fig. 1 in the revised manuscript.

4. The authors should give more details when they describe the methods. In particular:
  • They should give a standard definition of storm surge, with a citation (e.g., Pugh, D., 1987. Tides, Surges and Mean Sea Level. Wiley, New York);
  • Which neural network library are they using? Did they develop it? Give some informations and, if possible, some citations;
  • It's not clear what are the NN inputs and the desired output and some other NN settings. What is the temporal frequency of the records? How long are the database? How are computed the pressure differences?
  • From the manuscript it seems that the results are extracted from the training period. The performances of a neural network cannot be evaluated in this phase. This should be done in a testing period, where the data were not used for the training. If the database is too short I suggest to extend it with records from normal weather periods.
  • The choice of the number of neurons in the hidden layer is not clear. This should be explained better and possible overfitting problems should be discussed.

Response:

  • The definition of storm surge has been revised according to Horikawa (1978).

- Although there is some software for evaluating neural networks that could be applied, we developed our computing program.
- The inputs and the desired output are demonstrated in Equations (10)-(13). The training and testing data are listed in Tables 1 and 3. According to Horikawa (1978), the pressure difference is defined by the spatial mean pressure minus the lowest atmospheric pressure on the sea surface during the storm.
- We have added another section, Section 5, to discuss three illustrative examples for the forecast of the time variation of storm surges using the trained neural networks. The time series of storm surges from nine typhoon events were selected, six of which were used for model training and testing the model, and the other three were used for model forecasting.
- The optimum number of neurons in the hidden layer is dependent on the complexities and nonlinearities of the problems and is generally obtained by trial and error.

5. Conclusions should be extended.

Response:

We have rewritten the conclusions.

For Specific comments:

Response:

All these minor comments have been responded to and revisions made accordingly.