Review of: "Flood Prediction Using Artificial Neural Networks: A Case Study in Temerloh, Pahang"

Alessandro Amaranto

Potential competing interests: No potential competing interests to declare.

The research topic is interesting. However, the description of the experiment can be improved by providing more insights about:

1- the data and their structure.

2- the definition of the architecture of the ANN model: why there was no architecture optimization?

3- Still not clear if it was a regression or a classification problem

4- which threshold (if any) you used to define a flood?

5- prediction lead time specification

Below are some specific comments:

Abstract:

Flood is one of the natural disasters that causes damage to properties, and sometimes loss of lives – perhaps this can be phrased better.

'Despite this' seems inappropriate, since you are not contradicting the previous sentence. Furthermore, at the end of the sentence, you say 'using machine learning'. In general, it seems to me that research on flood prediction needs to be done, but without constraining it using a specific method.

Are highly correlated to floods: Is flood a variable? In that case, it would be best to specify that your problem is a classification one (if this is the case).

A lower temperature has a higher chance of rain and subsequent flooding: I am afraid this contradicts your previous statement

Has produced an accuracy of 0.9909: Using which metric of performance?

Introduction:

Flooding is one of the natural disasters that has been a problem in various parts of the world. Perhaps you can be more specific.

Immensely (?)

However, this modelling. Perhaps you meant 'this model'?

Recent developments in technology have introduced a few techniques which improve flood prediction. One of the developments is a physical-based model that has high effectiveness… Can you provide a name or a reference?

The other machine learning techniques that can be used. It is not correct. There are many others in the literature, and you are just providing an example. Perhaps you can rephrase with 'Other ML techniques', removing the 'The'

Materials and methods:

To achieve all three objectives of this research, a good research procedure needs to be established to produce excellent results. This sentence can be removed.

Fig 1: flood monitoring dashboard. Perhaps it is best to mention the fact that you are developing a dashboard in the introduction.

…a data pre-processing stage to ensure the quality of the data. You can be more specific.

I do not see neither input variable nor lag time selection. In addition, what is the forecasting lead time?

3.2: Model Development. This section can be improved. Plus, I believe it is important to present all the metrics of performance in a separate section. Plus, the criteria used for the ROC and the AUC are not clear.

Results and Discussion.
The first section is data presentation and should be moved above. It is still not clear what you mean by 'flood'. Furthermore, 'There are 10 separate datasets, and each dataset has 13 columns, and 32 rows embody the data collected by days and clustered by month.' No idea what those 13 (*10?) columns are, and what the 32 rows are as well.

The next step is to build the ANN structure in the Python programming. The neural network is constructed to have one input layer, two hidden layers with six neurons in each layer, and one output layer to balance between the model complexity and the process time and required machine capacity. So this means that the architecture has not been optimized?

Where 210 "No Flood" and 4 "Flood" data points are correctly predicted. Where is 214 (or 219) coming from? If the data are daily, this should be 365*2. If the data is monthly, we would have 24 data samples (might be too small to perform ANN). Overall, a better description of the data and the study area would be welcome. Additionally, the criteria for defining a flood is not specified.

Lastly, the performance and error evaluation is done through MAE, MSE, RMSE, which have values of 0.009, 0.009, and 0.096, respectively, indicating the error in the prediction is very low. The R^2 value of 0.768 proves there is a high-variance

relationship between the variables, and 76.8% of the observed variation can be explained by the model's inputs.

I thought that the output of the model was categorical. Did you do both a regression and a classification problem?