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

Mohammad Javad Shaibani[1]

1 Tehran University of Medical Sciences

Potential competing interests: No potential competing interests to declare.

This paper addresses flood prediction using ANN in Temerloh as a case study. The general idea behind the research is good. However, I have several points regarding the paper:

- The key words should be altered as they seem inappropriate for the subject: Machine Learning, Flood in Pahang, Flood Monitoring Dashboard.
- The literature survey is incomplete. Please justify why you are using AI methods instead of statistical methods and why only 8 papers are selected from the mass of literature in the research area.
- I think the last phrase of the introduction should be removed or modified as the literature is full of ANN-based prediction methods: ”Lastly, the study will benefit academicians as it will be one of the references that can be added to the list of the latest technology in predicting floods.”
- Regarding the input data: Please explain more about the dataset or attach it as a file to the article. Besides, why some important data inputs like geographical coordinates are not considered?
- Why do you not use dimension reduction methods such as PCL instead of correlation analysis? Especially when some data points are generated via the linear interpolation method. Please indicate whether you could build the model with the same (or close) accuracy and use less data. I think some simple feature selecting approaches like PCL should be applied.
- In modeling ANN: please justify why you are using a shallow ANN and not using a deep ANN. Besides, why you are not using random forest classification methods as they seem to be more efficient for such classification problems.
- I think the obtained results should be compared with some benchmark models.
- Regarding your setup for the ANN: please justify why you're using such a setup, like two hidden layers with 6 nodes. First, why you are using two hidden layers, as it is a well-known fact that an ANN with one hidden layer is enough for addressing such a problem…
- Please explain more about the training algorithm as it has a significant influence on the accuracy of the ANN.
- In the Model Development section: Why you are discussing the learning rate and left it incomplete: ”Another method to be applied to the model in order to minimize the differences between the predicted and actual output is the learning rate.”
- Please explain how the “Flood Monitoring Dashboard” is relevant to the ANN and if not, why it has been brought into the paper.
- I think the conclusion section is long and should be written concisely.