

Integrated versus isolated scenario for prediction dissolved oxygen at progression of water quality monitoring stations

A. Najah, A. El-Shafie, O. A. Karim and O. Jaafar

The authors appreciate the opinion of the referee on our research and thank him for his positive comments that will strengthen the content of our research manuscript. In fact, the revised paper is now significantly improved from the original one by considering all the reviewers' comments. All changes made to accommodate the reviewer comments are underlined in the revised manuscript.

The paper objective sounds interesting: it promises to evaluate the use of Artificial Neural Networks (ANN) for assessing water quality. There are very few applications concerning ANN in aquatic studies and therefore the applicability of ANN in assessing environmental quality is not known as yet. The subject addressed is within the scope of the journal. However, the manuscript, in its present form, should be improved in the light of the following comments in order to be more suitable for the readers. The overall assessment is that the manuscript could be accepted with “minor revision”. Addressing the following comments/modifications could be satisfactory in order to justify recommendation for publication.

Hereafter I am giving my recommended points that should be considered in the revised version of the manuscript.

Comments:

- It is highly recommended to enhance the introduction section by introducing citing several researches for Artificial Intelligence application for environmental and / or other fields**

Reply:

Additional relevant references are now introduced and cited in the revised manuscript.

- **The introduction section should be re-arranged by splitting it into three parts, background, problem statement and objective.**

Reply:

The author highly appreciates the reviewer's feedback. The authors are interested to introduce the introduction section without splitting. The introduction section re-arranged based on reviewer's feedback, background, problem statement and objective in the revised manuscript.

- **A thorough discussion on the advantages and limitations of the method should be provided.**

Reply:

The discussion section has been significantly improved and more rigorous discussion on the advantages and limitations of the method has been added in the revised manuscript.

- **In the description of the study area, it should be better to introduce schematic diagram on the river main stream and tributaries and the location of the monitoring stations under study.**

Reply:

The authors fully agreed with the referee in this point. The schematic diagram on the river main stream and tributaries and the location of the monitoring stations has been drawn in the revised manuscript.

- **The key ANN parameters are not mentioned. The rationale on the choice of the particular set of parameters should be explained. Have the authors experimented with other sets of values? What are the sensitivities of these parameters on the results?**

Reply:

Among many types of ANNs, the most widely used is the three-layer feed-forward neural network which is introduced in this study. LMA algorithm is used for its faster convergence speed and lower error rate to overcome the shortcomings of traditional BP algorithm as slow

to converge and easy to reach extreme minimum value. The architecture of the proposed model giving the optimal result was multi-layer perceptron neural network (MLP-NN) with The hyperbolic tangent sigmoid transfer function between the input and the hidden layers, linear transfer function between the hidden and output layers. The number of neuron is varying in the range 1–20 neurons.17 neurons were selected as the best number of neurons based on the minimum value of MSE of the training data set.

To evaluate the effect of input parameters on the model, two evaluation processes were used. First, the performance evaluation of various possible combinations of the parameters was investigated utilizing Coefficient of Efficiency (CE) and Mean Square Error (MSE) approaches to determine the most effective parameters on the output. The second assessment process was based on partitioning the neural network connection weights in order to determine the relative importance of each input parameter in the network (Garson, 1998; Emad et al., 2010).

- **The discussion section does not really address the significance of these results or set them in the context of the cited. The authors should rewrite the commentary to better explain how the results could be used for better modeling, lessons learned etc. At present the ms reads more like a general description of the work carried out, rather than taking the reader deeper into the processes and mechanisms involved.**

Reply:

The discussion section has been significantly improved and more rigorous discussion on the advantages and limitations of the method has been added in the revised manuscript.

- **Problem statement should be revised to show the contribution of this research in clearer way.**

Reply:

Owing to the reviewer’s feedback, the suggested modifications have been done and underlined in the revised manuscript.

- **Why are the three performance indices (i.e. Coefficient of determination (R^2), Mean Absolute Prediction Error (MAPE) and Coefficient of Correlation (CC).) adopted in this study? What are their advantages over other indices in this case?**

Reply:

Actually, in developing such prediction model using Neural Network, the model could perform well during the training period and might provide higher level of error when evaluating during either testing period. In this context, in this study the authors used these performance indices to make sure of that the proposed model could provide consistent level of accuracy during all periods. The advantages of utilizing these two statistical indices as a performance indicator of the proposed model are as follow:-

- **Coefficient of Efficiency (CE) which is often used to evaluate the model performance, introduced by Nash and Sutcliffe (1970), Mean Absolute Prediction Error MAPE (MAPE) which it usually expresses accuracy as a percentage (Ömer Faruk, 2010) and the coefficient of correlation (CC) which is often used to evaluate the linear relationship between the predicted and measured parameter.**
 - **Consequently, using these indices is guaranteed consistent level of errors which is providing a great potential for having same level error while examining the model for unseen data in the testing period.**
-
- **The author verify the efficiency of the model using field data during the period 2009-2010 using the proposed model using the integrated model ONLY. The author should provide the performance of the isolated scenario for these data as well.**

Reply:

The authors fully agreed with the referee in this point. The performance of the isolated scenario for data set has been discussed in the revised manuscript.

- **In the conclusion section, the limitations of this study, suggested improvements of this work and future directions should be highlighted.**

Reply:

The conclusion section has been significantly improved and more rigorous discussion on the advantages and limitations of the method has been added in the revised manuscript

- **There are a few formatting issues that would improve clarity, which mainly involve figure font sizes.**

Reply:

The formatting of the paper has been improved to provide more clarity to our paper.

- **Line 26, Page 6076, xi: there is no xi in the mathematic formula. Clarify?.**

Reply:

The above mentioned mistake has been corrected

- **There are a few typos in the text that the authors should address.**

Reply:

Once again, we would like to thank the referee for her/his inspiring comments. A new linguistic revision of the whole manuscript will be performed.

On behalf of all co-authors | A. Najah