Hydrology and Earth System Sciences Discussions

Hydrol. Earth Syst. Sci. Discuss., referee comment RC1
https://doi.org/10.5194/hess-2022-86-RC1, 2022


# Comment on hess-2022-86

Anonymous Referee #1

---

Referee comment on "Prediction of groundwater quality index to assess suitability for drinking purpose using averaged neural network and geospatial analysis" by Seok Hyun Ahn et al., Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2022-86-RC1, 2022

---

This paper applied GWI to evaluate groundwater quality of 3,552 portable groundwater wells based on the 29 water quality parameters. The authors claimed that ANN and SVM models yielded the best result for GWI prediction. The research has practical applications based on the 2D spatial analysis, but its presentation can be improved as well.

- Please highlight the innovation of your study in Abstract.
- Introduction: It is not clear why GQI is the selected index to evaluate groundwater quality. What are the advantages/potential of using this index?
- Line 35: make “a WQI suitable” to “a suitable WQI”.
- The Introduction section must be written on more quality way. The research gap should be delivered on more clear way with directed necessity for the conducted research work.
- It seems the major contributions of this study are using 47 water quality parameters from 8326 wells to determine the groundwater quality index (GQI) using an averaged neural network and also investigate field applicability with two-dimensional (2D) spatial analysis. I strongly suggest to explain more about these contributions in introduction to enhance the quality of this paper over previous. The novelty of this work must be clearly addressed and discussed in Introduction section.
- Table S1 is not available (line 99).
- Line 114: provide a reference.
- Unclear sentence (line 153): “The models used include averaged neural …”
- Nothing is reported about the distribution of the data, about possible correlations between them. This is to be provided.
- The methods are taken in “model setup” section is not explained in such a way that the strengths and weaknesses of the methods become visible to the reader. Likewise, no reasons are given as to why the selected ensemble methods are favorable for this issue, or why data are split into 5 subsets. It is better for the reader to directly use the cited literature to classify the methods.
- When presenting the RF method for feature selection, no explanation is made about the method, advantages, and disadvantages, no comparison is made to other methods, such as AIC, Gamma Test.

- The Korean groundwater quality standard for each parameter is not provided in the manuscript.
- Again Fig. S1, Fig. S2, Fig. S3, Fig. S4 and Fig. S5 are not available.
- Nothing is said about the results of boxplots (Fig.4).
- The proposed models have different parameters and model structure. How the authors do model parameter, layers, nodes, etc determination? Up to what structure and function of ANN, SVM, and Naïve Bayes models causes these models have high performance?
- Using binning method for special analysis is an applicable and useful result in this study.
- There is not enough explanation on different models' prediction and feature selection results.
- Please cite the following reference:

Uncertainty analysis of water quality index (WQI) for groundwater quality evaluation: Application of Monte-Carlo method for weight allocation. Ecological Indicators, 117, 106653.