Reply on RC1
Seok Hyun Ahn et al.

Author 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-AC1, 2022

# Reviewer 1

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.

Response:

As the reviewer commented, the previous abstract was written too general. We revised abstract to emphasize the novelty and innovation of our study.

☐ Introduction: It is not clear why GQI is the selected index to evaluate groundwater quality. What are the advantages/potential of using this index?

Response:

We agreed with reviewer’s comments. We revised the introduction to clearly show the advantage of GQI for the groundwater management.

☐ Line 35: make “a WQI suitable” to “a suitable WQI”.

Response:

Corrected.
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.

Response:
We thank you for your thought comments. We revised the introduction to show not only the advantage of GQI but the potential of GQI coupled with 2-D spatial analysis for groundwater management.

Table S1 is not available (line 99).
Response:
You can find Table S1 in Supplement.

Line 114: provide a reference.
Response:

pH is included in the water quality parameters to be measured for groundwater quality, but no water quality standards of pH are presented to determine whether drinking is suitable. Thus it was excluded for GQI calculation because it was difficult to use for GQI, which used the distance from the water quality standard.

Unclear sentence (line 153): “The models used include averaged neural ...”
Response:
We revised this sentence for the clarity.

Nothing is reported about the distribution of the data, about possible correlations between them. This is to be provided.
Response:
We thank you for the pointing out the issues we did not consider. We added the results of comparing individual water quality parameters of groundwater according to suitability in supplementary table and described in the results section. In addition, the results of correlation analysis between water quality parameters were described in the result section.
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.

Response:

We selected 10 classification machine learning models included in the R package ‘caret’ to predict the GQI grades, and selected ANN with the best classification performance as the final model. The ensemble method “Random forest” was used only for feature selection as it was not able to perform feature selection in ANN.

Despite the importance of model setup, we have not written specific details about model setup, so I also agreed with the reviewer’s comments that we can provide inconvenience to readers. We added a supplemental table describing the model setup with detail information (e.g. parameter and function) with the references which help the reader understating the methods clearly.

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.

Response:

RF is a powerful method to select the features that affect classification. We did not use RF to predict GQI grades, but only used it to select features that influence classification of GQI grades. As the reviewer concerned, the selected features may vary from model to model, but it is absolutely difficult to compare because the criteria or indicators selected are different according to the model. We added the supplemental results for variations in the explanatory power of the model according to the combination of features selected by the RF, instead of directly comparing them with other models. This result may be useful information for readers because information on the classification performance according to the number and combination of water quality parameters can be quantitatively known.

The Korean groundwater quality standard for each parameter is not provided in the manuscript.

Response:

You can find this information in Table S1.

Again Fig. S1, Fig. S2, Fig. S3, Fig. S4 and Fig. S5 are not available.

Response:

All these figures were in Supplement.

Nothing is said about the results of boxplots (Fig.4).
Response:

We sorry for the unintentional mistakes. We described in detail the statistical tests for the results.

☐ 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?

Response:

R package 'caret' allow the researcher to test more than 200 classification models with near-automatic cross validation-bootstrapping and parameter tuning, and to find the best predictive model. We used default parameters in R package 'caret' without any modification. As reviewer suggested, the parameters used in each model were summarized in the supplemental table.

☐ Using binning method for special analysis is an applicable and useful result in this study.

Response:

We appreciate your support for our approach.

☐ There is not enough explanation on different models’ prediction and feature selection results.

Response:

It seems that we have not written enough explanations for the results of model prediction and feature selection. We described the both results in more detail in the result section.

☐ Please cite the following reference:


Response:

We added this reference in introduction section.