Hydrology and Earth System Sciences Discussions

Hydrol. Earth Syst. Sci. Discuss., author comment AC2
https://doi.org/10.5194/hess-2022-86-AC2, 2022

Reply on RC2

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-AC2, 2022

Reviewer 2

The paper describes the application of data-driven models to predict the groundwater quality index for drinking purposes using multiple water quality parameters. Groundwater quality was assessed by 2D spatial analysis and long-term monitoring results. However, the manuscript needs to be further improved in terms of novelty, literature research, and data handling and interpretation.

☐ Abstract: The authors need to explain more the novelty and importance of their work for international communities.

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: The section must be substantially improved by citing new literature.

Response:

We have partially replaced it with a new literature.

☐ Lines 66-70: I disagree with the authors' statement, there are several recent works about data-driven models predicting comprehensive groundwater quality (WQI, vulnerability, suitability for drinking water,...) as shown in the examples below. Search for recent studies and emphasize improvements of the authors' research from them.

Prediction of groundwater quality using efficient machine learning technique. Chemosphere, 276, 130265. https://doi.org/10.1016/j.chemosphere.2021.130265

Advanced utilization of multi-learning algorithm: ensemble super learner to map groundwater potential for potable mineral water, Geocarto International, DOI: 10.1080/10106049.2022.2025921

Reliability evaluation of groundwater quality index using data-driven models. Environmental Science and Pollution Research, 29(6), 8174-8190.

Response:

We are sorry for missing the recent studies. We revised the introduction sections to emphasize our novelty and innovation from the recent studies with the recent references.

☐ Lines 80-83: There is a doubt that 'Analyzing long-term monitoring results can evaluate the accuracy of the groundwater pollution vulnerability.' Details comments are described in the results section.

Response:

This sentence was deleted because the results of our study were not only unsupportable but could also cause confusion in the overall direction of the manuscript.

☐ Figure 1 needs to be redrawn. Clarify the location of the authors' study area by representing neighboring countries, name of country (or region), and other essential elements of map presentation for the worldwide readers.

Response:

We modified the figure 1 to have more information with neighboring countries, and name of countries. And we zoomed Korean with the original Figure 1.

☐ Lines 105-106: The authors removed 4,774 wells of groundwater quality that are inappropriate for drinking purposes. Considering the title of the manuscript, it is likely more effective to analyze both appropriate and inappropriate groundwater samples to assess suitability for drinking purpose. Is there any reason the authors removed unsuitable groundwater quality for drinking water?

Response:

We thank you for concerning the critical points. Legally, if even one of the groundwater quality indicators exceeds the water quality standard, the groundwater cannot be drink. Since it is already legally determined whether it is suitable for drinking in a dichotomous manner, it is meaningless to provide GQI information to citizens that has been judged to be inappropriate. Therefore, by providing GQI information on the groundwater suitability for drinking, citizens will be able to obtain comprehensive water quality information and determine the possibility of contamination of groundwater in advance. In addition, it will be possible to systematically manage groundwater quality by collecting the regional GQI.

☐ Lines 151-153: In the section 2.3, the method to calculate GQI and to classify GQI into the grade was already developed. Then why the classification models are necessary? I think there is no reason to use learning models including averaged neural network, RF, SVM,… whose accuracy is not perfect. Isn't classifying according to a given classification method the most accurate?

Response:

Of course, GQI graded can be decided by calculating GQI. Nevertheless, there is a reason we used the classification model to predict GQI grade. First, since it is practically difficult (e.g. cost and man power) to measure each of the 48 water quality parameters, it is important to select the minimum water quality indicators that can be determined by GQI grades. Classification model can select water quality parameters that can calculate GQI grades through feature selection. According to our results, we made high performance up to 90% (11 parameters) or 95% (14 parameters) even we used lower number of water quality parameters compared to models made using all parameters. The classification model will dramatically reduce the time, cost, and labor required to measure water quality. We added this information to the results to clearly present the advantage of using the classification model.

☐ Section 2.4: References and details are required when describing the models.

Response:

We summarized the function and parameter of the models with the references in the supplemental table.

☐ Section 2.5 also needs to be described in more detail.

Response:

We describe the function and parameter of the RF for clarity in section 2.5.

☐ Lines 198-200: It is unclear why Chungcheongbuk-do was analyzed with high resolution unlike the other regions. Is there a specific reason?

Response:

Even we used the groundwater qualities from 3,552 wells and binning methods, it is still insufficient data considering a national scale (e.g. Republic of Korea). We were able to determine the GQI grades for some provinces less than 30% of total province area. We wanted to show the feasibility of combining GQI with spatial analysis by selecting Chungcheongbuk-do, where the GQI grades was determined in a significant area through the binning methods among the province. And we tried the to compare the GQI results with other independent long-term water quality results. Although we gave a brief reason selecting Chungcheongbuk-do shortly in result section in original version, we revised this sentence for clarity.

☐ Lines 278-288: The GQI grade was estimated using groundwater quality data sampled at one time for each well. Then, what is scientific basis of the statement that long-term trends confirm the reliability of the authors' results? What is relation between increasing (or decreasing) trend and the grade? In addition, the authors only showed a long-term trend in only one area corresponding to each grade. I disagree with the authors on the following three points.

- There is no relation between long-term trend and the GQI grade.

- It is impossible to confirm the reliability of the authors' results by checking only one area per each grade.

- In Figure 6B, Site A (worrisome grade) showed increasing trend in chloride but the value is lower than Site C (very good grade) showing decreasing trend. It is not up to expectation.

Response:

Due to the nature of groundwater, pollution is localized and residual, so it takes a very long time to observe the pollution of groundwater. As mentioned earlier, the GQI we developed was developed for potable groundwater, and through the binning method, we present information on the groundwater quality level of a specific area. The fact that the GQI of a particular region is selected as "worrisome" means that the pollution is not simply going on in one well of the region, but that the pollution is going on for a long time due to the pollution source in the region. Therefore, we think it is a very important approach to find the characteristics of GQI grades by comparing them with long-term water quality measurement data from national groundwater networks. Through this process, the GQI grades may not simply provide information on the current state of water quality, but can help rapidly select areas that require groundwater quality management.

Of course, I also agree that selecting an area for each GQI grade and matching GQI grade with a long-term trend in national groundwater network may be considered insufficient to explain the characteristics of the grade. There are 688 wells for total national groundwater network, and 48 wells in Chungcheongbuk-do. Unfortunately, selecting an independent region (where the GQI rating results are not present on either side) to confirm the characteristics of the region's GQI grades is very limited in terms of the wells overlapping with that of the national groundwater network. We further described the contents of these technical difficulties in the discussion as future studies.

As groundwater move slowly through an aquifer, the groundwater qualities should be determined by the type and concentration of the geological materials and nutrients in the aquifer. The change in groundwater quality is due to external environmental factors such as human activity, and in particular, nitrate are very likely to be factors of human activity around them. In the case of chloride, it may be high due to regional characteristics, but nitrate increases due to surrounding pollutants, and generally, the concentration of chloride increases at the same time. Simply increasing one water quality indicator does not increase GQI rapidly. Because GQI increases as multiple water quality indicators approach water quality standards, it is reasonable to analyze the characteristics of GQI ratings to consider the trends of Chloride and nitrate simultaneously.

☐ Discussion: Most of the discussion is a repeat of the previous sections. More in-depth discussion is required.

Response:

We agreed that there is a more space to improve the discussion. We revised the discussion so that the proposed GQI could be more innovative and useful compared to previous studies.

☐ Figure S4 is not referred in the manuscript.

Response:

We are sorry for the unintentional mistake. We added the Figure S4 in the proper positions.