Hydrology and Earth System Sciences Discussions

Hydrol. Earth Syst. Sci. Discuss., referee comment RC2
https://doi.org/10.5194/hess-2021-618-RC2, 2022

Comment on hess-2021-618

Anonymous Referee #2

Referee comment on "A global assessment of nitrogen concentrations using spatiotemporal random forests" by Razi Sheikholeslami and Jim W. Hall, Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2021-618-RC2, 2022

The manuscript, titled "A global assessment of nitrogen concentrations using spatiotemporal random forests" by Sheikholeslami and Hall, introduced a machine learning (ML) approach (random forest model - RF) for predicting in-stream nitrogen (NOx-N) concentrations at the global scale. According to the authors, the novelties of this work are (1) its global scale application and (2) the spatio-temporal RF approach proposed in this study. In general, the manuscript was well written. Despite the results (instream NOx-N concentration) look quite well, the are several points regarding the model/approach used in this study that need to be addressed.

General comments:

1) Representation of nitrogen (N) lag times from input to riverine N export: For water quality (e.g., N) modeling, it is expected that there could be significant N accumulated in the soil as biogeochemical legacy and the long travel time within the unsaturated/groundwater zone that could result in a lag time of years to decades between N input and riverine N export signals (e.g., Meals et al., 2010; Van Meter et al., 2017; Chen et al., 2018). It is unclear to me how the proposed RF model could take into account these factors. From my interpretation of the result, the "cumulative month count" variable (Figure 9) somehow could compensate for this kind of effect. However, we should not try to get the right result for the wrong reason.

2) Variable importance: Why are the month of the year and the cumulative month count the most important variables? I am wondering if the data used in the model has a strong seasonality that makes the variable "month of year" really matter. If this is the case, what is the implication for model application/performance in other areas that have less/no clear seasonality? Is the predictor "cumulative month count" highly important because of an increasing trend in the output variables in many areas (lines 495-497)? If yes, what are the implications from this? Why does "fertilizer application" have a low rank? Why is there not much difference in the variables that were ranked 3rd to 15th

(Figure 9)?

3) Spatial unit: For predicting instream nitrogen concentrations, it is not clear to me why the authors did not use river network (instead of grid cell) as a spatial unit. In 1 grid cell (size of ≈ 55 km2) there could be several rivers, so it is unclear if the predicted values are applied for the main or tributary rivers. In addition, with rivers in big basins (e.g., Elbe, Mississippi, Amazon, Mekong River Basins, etc) that are running across multiple grid cells, it would be useful to incorporate the effect of upstream management/catchment characteristics into the model, it is not clear if this was considered in the RF model or not. As I understood from the description, the predictors for instream N concentration prediction currently only cover the properties within the grid cell of interest (no consideration of information from the upstream grid cell for large rivers).

4) Capacities/limitation section: I would suggest including a section describing the model capacities and limitations of the model/approach. Althought the authors mentioned these points briefly in the conclusion section, however, they could be extended in a sperated section.

Specific comments:

Lines 30: "In addition, extensive construction of dams, excessive extraction of groundwater, deforestation, and expanding agricultural land use have altered sedimentary processes, mobilization of salts, and nutrient export to river systems, all of which drive WQ deterioration and groundwater pollution in many parts of the world…". Were these factors considered in the model?

Lines 183-184: What is the temporal resolution of the NOxâ□□N data and how were they aggregated to monthly timestep?

Line 206: Please indicate where the readers could find the list of 27 potential explanatory variables

Line 287: "The second strategy …" The second strategy has not been mentioned before this point

Line 291: "Cumulative Month since 1992": how sensitive are the results to the start of the month count? This is a critical point if someone wants to run the model for other periods

Line 295: "..17 variables" – Please point out the list of 17 variables

Figure 3: I would suggest adding more frames to the "output" panel (as already done in the "input" panel) to reflect that the spatial and temporal properties of outputs

Lines 418-419: Is there a high correlation between elevation and latitude/longitude?

Lines 426 – 430: back to table 1, it is not clear whether these data (livestock population) were assumed to be time-invariant or time-variant

Line 447-448: "This might be partially due to a high correlation between the agricultural fraction of land area and nitrogen fertilizer use" – Why do highly correlated variables have different rankings?

Table 2: How were annual time step data (especially fertilizer application) disaggregated to monthly time step?

Table 1: Please provide full names for the technical terms (e.g., ANN, DT, MLT,…) in Table 1

References:

Meals, D. W., Dressing, S. A., & Davenport, T. E. (2010). Lag time in water quality response to best management practices: A review. Journal of environmental quality, 39(1), 85-96.

Van Meter, K. J., N. B. Basu, and P. Van Cappellen (2017). Two centuries of nitrogen dynamics: Legacy sources and sinks in the Mississippi and Susquehanna River Basins, Global Biogeochem. Cycles, 31, 2–23.

Chen, D., Shen, H., Hu, M., Wang, J., Zhang, Y., & Dahlgren, R. A. (2018). Legacy nutrient dynamics at the watershed scale: principles, modeling, and implications.

Advances in agronomy, 149, 237-313.