Response to Reviewers

Determination of vadose and saturated-zone nitrate lag times using long-term groundwater monitoring data and statistical machine learning

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2020-169, 2020.

We are grateful to the reviewers for their thoughtful comments, which will improve the paper. Our initial responses to the first two reviews are indented below and shown in blue text.

We also note that the following disclaimer should be applied to the discussion paper:

This draft manuscript is distributed solely for purposes of scientific peer review. Its content is deliberative and pre-decisional. Because the manuscript has not yet been approved for publication by the U.S. Geological Survey (USGS), it does not represent any official USGS finding or policy.

Reviewer 1

Geology/Hydrogeology is missing. Provide a hydrogeological map, cross section, hydraulic characteristics of the aquifer etc.

> In the revised manuscript we will provide a cross section similar to those available in other publications focused on the Dutch Flats area. We will also add additional hydrogeological descriptions in the text.

Add more information about nitrate and its processes.

> In the revised Section 2.1 (Site Description) we will include more denitrification information, including more detail on findings from prior research in the area. Previous work suggests that denitrification is not extensive in the groundwater in this area.

Fig1.Change the maps. The figure needs to be more attractive. Add coordinate system.

> We will update the figure to include graticules. The figure includes a colored topographic map with appropriate symbology and detail necessary for the paper. We are uncertain what is meant by the suggestion to make the figure more attractive (e.g., overall figure should be changed?, improve resolution?, other?). We will also add a north-south vertical section showing the extent of the aquifer and schematic of groundwater flow directions.

The literature is out of date.

> We agree, as publication of machine learning models has recently been very rapid. We will update the manuscript with literature that has been published while the manuscript was in review.

Discuss the role of Nitrate isotopes for future contribution in this concept. Recent article provide the interaction between surface and groundwater bodies using nitrate isotopes which might be helpful in future works.

> We are aware of some studies involving statistical approaches and N and O isotopes (e.g., https://doi.org/10.1002/2015WR018523; https://doi.org/10.1016/j.jconhyd.2015.07.003) but are unsure if these are the articles referred to by the reviewer.
>
> In general, nitrate isotope ratios in the aquifer are fairly uniform (e.g., d15N = +4 ± 2 per mil) and consistent with recharge beneath fertilized agricultural land elsewhere. Previous work indicated a possible minor downward increase in d15N, which could be related to different recharge sources or historical changes in fertilizer/manure ratios. Evidence of denitrification (from dissolved gases and isotopes) was mostly limited to some of the deepest wells near the bottom of the aquifer. The effect of major canal leakage is considered largely to be nitrate dilution (i.e., relatively little nitrate addition, at least from the upgradient canals). Additional isotope data might be useful for documenting temporal shifts in recharge sources, or irrigation return flows to the river; however, it is difficult to know exactly the location or size of the contributing area for each well, especially the deeper ones. We will clarify some of these points, though a detailed discussion likely is beyond the scope of this paper.

Reviewer 2 (Scott Gardner)

The study presents the environmental setting well in terms of soil, climate, and land use, how-ever, more specific information (cross-sections or maps) on the geologic setting would be useful in evaluating spatial variability in lag times.

> In the revised manuscript we will provide a cross section similar to those available in other publications focused on the Dutch Flats area. We will also add additional hydrogeological descriptions in the text.

The distance between the monitoring wells evaluated and the screens that are sampled to the sources of nitrate (probably fields) are not touched on in the manuscript and might be useful in explaining variance in lag times. Perhaps land use might also be important to consider nearby the wells, as interception, evapotranspiration, and other land use specific processes could be relevant to nitrate lag times.

> Thank you for pointing this out. We do note some general trends over larger spatial areas, where wells north (upgradient) of the canals are lower in nitrate due to the absence of row crop production. The vast majority of wells are surrounded by agricultural fields, and we are lacking detailed year-to-year records of fertilizer application or crop production. We do focus in the paper on the proximity of wells to irrigation canals, which have been shown in past work to substantially impact groundwater nitrate concentrations due to focused recharge of lower-nitrate groundwater. We will add a couple additional sentences to the manuscript to expound on this information.

line 17 - I am not sure you need to include the part about it not being common to have unsaturated velocities slower than saturated, this has been the case in other studies and is not out of the ordinary (fractured bedrock aquifers, karst, etc.)

> We agree that there are environments where this might be expected. We will clarify that this statement is a generalization for unconsolidated surficial aquifers receiving distributed recharge.

line 79 - perhaps provide a reference explaining the importance of canals in the region for readers that are not familiar with the study area.

> Although documented extensively elsewhere, we will insert a brief comment to emphasize the importance of the canals. The impact of canals will also be illustrated in a new figure summarizing the hydrologic setting. Thank you for pointing this out.

line 107 - here and everywhere after it is not clear what is meant by screen length, is this the depth bgs that the screen begins, or the size of the screen?

> In the revised manuscript we will define this as “length of screened interval.”

please clarify line 157 - what is meant by ’bootstrapped’ readers which are unfamiliar with computer science jargon may have trouble with this please clarify.

> In the revised manuscript we will define this term.

line 234 - what was the reasoning behind selecting 1 standard deviation for an acceptable range of results? If this selection was arbitrary then it should be made clear.

> In the revised manuscript we will note that the range based on 1 standard deviation was considered a reasonable range of recharge rates that might be considered based on prior research in the area.

figure s1 please change the colours on the nitrate concentrations to better contrast the results

> Figure S1 will be updated to provide more distinction between the different results.

Reviewer 3 (Sophie Ehrhardt)

Abstract:

Line 16: Could you add some information about which area/time/well number you averaged the mean? And you did not mention the name or location of the study area in the abstract to which all numbers correspond to. Try to add this to make it more precise and enable the reader to set the study in space.

> We agree, this is good information to add. We will indicate that the mean was with respect to an area (i.e., the Dutch Flats area).

Line 27: Mention that denitrification plays no major role in the study area. Otherwise diffuse recharge could be affected by this process.

> In the revised manuscript we will mention in the abstract the lack of suggested denitrification.

Introduction:

Line 37: Please add a few sentences why research for nitrate contamination is important.

> We feel this material has been heavily documented in nitrate-related research already published – many of which referenced in this paper – and well known by the readers.

Line 63/64: The explanations "vadose (unsaturated)“ and "groundwater (saturated zone)“ could be earlier in the paragraph e.g. Line 38.

> In the revised manuscript we will provide these synonyms in the first paragraph of the introduction.

Methods:

Line 107: In which depths are shallow, intermediate and deep groundwaters? Even more important than the screen length.

> We agree the actual depths are important information. We will add an additional sentence with the range of vadose zone and saturated zone (depth below water table) thicknesses in this study. This will complement the hydrogeologic cross-section, which we will add in response to Reviewers 1 and 2.

Line 123: I did not check the paper, but how can the mean recharge stay the same, if 88% of the rates decrease? Because of highly positive outliers?

> In previous work, the recharge rates were slightly lower in the majority of wells, but the overall mean recharge rate was not statistically different.

Line 203: How strong was the relation between "Area of planted corn“ and "fertilizer application rates“? R2? Should be really high as you substitute the Ninput mass by an area.

> This is a good point. As discussed later in this paragraph, we were not simply substituting a proxy (area of planted corn) for actual fertilizer data. The choice we had to make was between a proxy and “no data” for years prior to 1987. Although the correlation was low for more recent years (R^2 = 0.26), groundwater nitrate concentrations have been closely linked to the area of row crops, including corn, in numerous water quality studies. As a result, we felt this was our best choice for incorporating an important dynamic variable into the study.

Line 204: More information on the reduction- perhaps in brackets "from... to...“ or "by . . ..%“ to estimate the effect (or its potential as marker in case of drastic drop).

> Thank you, this is a good location to give a sense of the magnitude of observed change. In the revised manuscript we will add quantitative information for fertilizer and planted corn, respectively.

Line 230: I am not sure, how to imagine the "apparent“ travel time as I only know about distributions (gamma or log-normal) of TTs. Your TT is the peak TT without any parts of it travelling faster or slower? So, you don’t assume a mixed signal stemming from TTs from different ages (e.g. in 2010 10% signal/NO3 load from 1990, 40% signal from 1991, 50%...)?

> We use the term “apparent” and mentioned imperfect age-dating tracers to address this exact question, which is that a single groundwater age typically represents a mean age reflecting the different recharge year for each water molecule in sample. The equations we present are simplified representations (as are tracers) comparable to piston-flow assumptions (a common simplification when interpreting groundwater age-dating tracer data).

Line 234: Please, define shallow!

> We can understand your frustration here. We will refine our descriptions as stated in the response to the Line 107 comment above. When the cross section is provided, it will show how the terms “shallow”, etc, are tied more to depth below the water table than to total well depth.

Line 252-255: And the fertilizer input (Nsurplus) of 1990? Isn’t this the most important input variable? Perhaps already cleared by Line 203, when adding R2.

> We agree, the fertilizer input certainly would have been a very beneficial variable to include; though, we unfortunately did not have enough data to include this variable in the analysis. Line 275 – 284 discusses dynamic variables and acknowledges stronger dynamic predictors could provide for an interesting follow up study. We will add to this section (i.e., Lines 275 – 284), specifically calling out N loading as a factor to consider in future studies, although these data are very difficult to reconstruct for long-term studies.

Line 263: "historical nitrate groundwater concentrations” or do you mean historical Ninput data?

> Historical groundwater nitrate concentrations are correct here. We unfortunately did not have long-term Ninput data to use for this study.

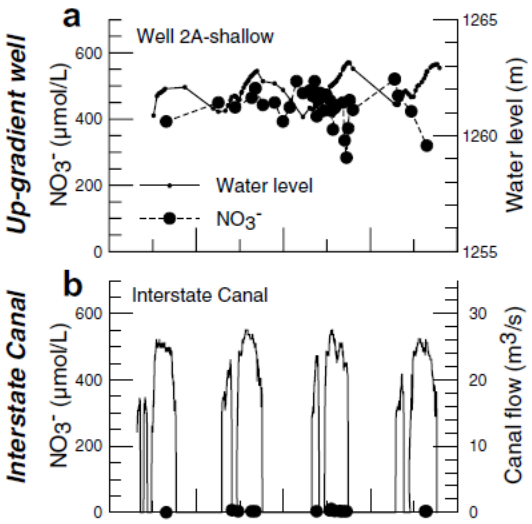
Results:

Line 292: I struggle to understand your differentiation between TTs and evolution of NO3. You don't use NO3 as tracer to derive TTs and therefore you can correlate both? Or don't you use NO3 to derive transport rates? If you calculate one variable based on the other, isn't the correlation useless? Sorry for my confusion. You concept of TTs is quite different from ours.

> The TT was not calculated based on nitrate, but rather the vertical vadose and saturated zone distance at each well. The rationale was that there is a known relationship between long travel times and low nitrate, and short travel times and high nitrate. Then, we used the random forest model to see which TT had the largest influence on the overall model's ability to predict nitrate concentrations.

Line 332: Doesn't your canal leakage has also high NO3 from time to time, based on surface runoff from fertilized fields directly (pipes and drainages)? And can you add some information on the canal system previously? Is it also to drain the fields?

> The previous Böhlke et al. (2007) paper found that when water was flowing through the Interstate Canal (largest canal in this region), nitrate concentrations were less than 0.06 mg N L^{-1}, and did not exhibit large spikes, during their collection period, in nitrate concentrations. Below is an excerpt from Böhlke et al. (2007) showing the nitrate concentrations in the Interstate Canal to be very low.



> While some of the smaller ditches could indeed carry tailwater, the major canals in this region serve as the primary delivery (only) canals in the region. We plan to add additional information regarding the dependence this region has on canals.

Line 332: Why does influence of canals extends further from the canal? Isn't its influence decreasing with distance?

> Thank you for pointing this out, as the wording is not completely clear. The text was intending to state that the influence from canal leakage is exhibited further from major canals than minor canals. We will adjust the text to state: "The effects of minor and major canals, where

groundwater [NO_3^-] in the vicinity of canals is diluted by canal leakage, and the influence of major canals extends further from the canal when compared to minor canal results."

Line 337: "nitrate reduction" add (also known as denitrification)?

Correct, and per comments from Reviewer 1 and 2, we will be incorporating additional discussion and information into the manuscript related to denitrification.

Line 338: "The partial dependence plot" add (Fig. 7)

In the revised manuscript will add "(Fig. 7)" to the text currently on Line 338

Line 342: I am surprised about your conclusion regarding the rapid aquifer response. You mention stratification and a groundwater age of 7years. Doesn't this account for a dampening of changing signals? Or what time do you assume with "rapid"? Or does this only correspond to the shallow, unstratified groundwater?

Our reference point for the term "rapid" is the many previous age-dating studies in shallow unconfined aquifers in agricultural areas where the mean transit time, and therefore the groundwater quality response time, in the aquifer is "decades". As noted earlier in the paper, the random forest model may be strongly influenced by younger groundwater with more pronounced nitrate signals.

Line 355: Do you have a recommendation how many data (stations) we need or how long time series should be to use your ML approach?

Hard to make a recommendation here, but certainly the larger the dataset (and number of stations), the better. Larger datasets provide more data used to train each tree, ultimately giving each tree more data to "learn" from, making the overall forest more robust.

Line 361: Isn't your "may be biased" a bit to optimistic? How can you distinguish a vanished NO3 imprint after denitrification from "stored somewhere in the upper soil"?

This is a good point. We will add that vertical sampling of the vadose zone for nitrate would provide ideal data to address whether this approach "misses" nitrate stored in the unsaturated zone.

Figures

Line 584: Is this pattern clockwise? Don't you need to switch the lower plots then?

In the revised manuscript the text will reflect the correct order of the plots

Line 597-600: Is there a difference between %inc and %Inc? It is not consistent in all figures.

There is no difference, but the revised manuscript will be updated to maintain a consistent nomenclature for this between the text and figures.

Line 622: Is there a space missing at "bData required further analyses“?

> Thank you for your attention to detail – the table will be updated to maintain a consistent format.

Line 625: Why only "some models were ultimately based on <1049 obs”? According to your table all models fit the condition “<= 1049” and some “= 1049 observations”.

> Table 2 reflects the further analyses that were performed on the model when the dynamic predictors were included in the analysis. In the revised manuscript we will add a comment to ensure that readers are aware this table is for the analysis that included dynamic variables. The reason some of the models included <1049 observations is due to the limitation in historical dynamic variable data available, where some data were not present prior to 1946. Therefore, the number of observations were decreased for some of the slower transport rates that result in a total travel time prior to 1946.
>
> For example, if a sample was collected in 2000, and it had a 60-year total travel time, the dynamic variable would be assigned a value from 1940. However, the dataset was limited to 1946, so any observation assigned a dynamic variable year prior to 1946 had to be excluded.