Hydrology and
Earth System
Sciences

Discussions

Open Access

HESSD

Interactive
comment

Interactive comment on "Seasonal watershed-scale influences on nitrogen concentrations across the Upper Mississippi River Basin" *by* Michael L. Wine et al.

Michael L. Wine et al.

mlw63@me.com

Received and published: 20 October 2020

We appreciate the referee's timely and thoughtful review. We agree that, particularly in the discussion, a number of factors deserve additional treatment, or in certain cases, minor corrections. We also agree with the referee's recommendation of greater concision.

1. One of the general conclusions of this manuscript, that TN concentrations in the UMRB tend to decline dramatically from spring to summer has been reported for nitrate-N in many individual watersheds in the region. In rivers draining agricultural watersheds, nitrate is frequently the dominant source of N at high flow in the spring. In

summer and fall, nitrate concentrations typically decline because of less drainage from cropland and more in-stream denitrification. Thus, the observed decline in TN is not surprising, although I am not aware of a publication that has demonstrated this for TN over a large region or related the pattern to climate and watershed characteristics as in this manuscript. The conclusion that an increase in wetland area and/or a reduction in cropland area would reduce N concentrations is also not controversial. I am not familiar with some of the statistical methods used, so I can't fully evaluate how appropriately they were applied or interpreted.

Response: We thank the reviewer for their insights on nitrogen transport. We agree with the reviewer that the results are not controversial and add to the timely discussion of drivers (and potential remedies) to freshwater degradation as a consequence of nutrient loading. We also appreciate the reviewer's acknowledgement of the insights our results can provide at the scale of the Upper Mississippi.

2. But I have some concern about independence of observations over time at individual sites and between upstream and downstream sites.

Response: We understand this concern yet can assure the reviewer that observations are largely independent at individual sites. The [TN] data represent grab samples collected during individual low flow events and are well-distributed across the study period – representing fewer than 2% of the days during that time. Therefore, because the measurements are not continuous in time (e.g., event, daily), observational independence is anticipated. We did, however, address serial autocorrelation using three categories of termsâĂŤharmonics (1st and 2nd), lagged wetness index, and seasonal soil nitrogen. All acknowledge the potential for serial autocorrelation in the observations, although limited, and aim to account for it.

With respect to cross-correlation among upstream and downstream sites, we agree that, in certain cases this could exist and aim to reproduce the observed variability at all sites. The majority of watersheds 62% were not nested within other study watersheds.

Moreover, 79% of sites had <50% overlap. Many investigations of watersheds use observational data from nested sites. Here we do not report p-values, which could be biased due to issues of cross-correlation induced by nesting.

To address this further, we calculated the autocorrelation function (autocorrelated variance at a given lag as a proportion of total variance) for each watershed (Fig. 1). With the temporal correlation, 8 out of 82 regions (or about 10%) have lag-1 autocorrelation above 0.5. Since we test "fixed effects" terms, in our opinion, the space/time correlation of residuals is not large enough to invalidate our conclusions. The spatial correlation is low (overall, may be no higher than 20% of overall variation, 0.22 = nugget to 0.27 = the sill of the unidirectional variogram). In other words, the variograms are close to the "pure nugget", i.e., the difference between values as lag distance approaches zero. Also, horizontal and vertical (on the map, so E-W and N-S, respectively) directional variograms are slightly different, but not a lot.

3. The major weakness of the manuscript includes lack of attention to sources of N other than fertilizer, namely animal manure, point sources and mineralization of soil organic N.

Response: As the reviewer notes the focus here is on cultivated areas and fertilizer, the primary cause of N loading. However, please note that we address all sources of N in our analysis (L304-305). We will, however, incorporate more explicit links to them in a revised manuscript. Specifically, our initial screening process for TN sites involved eliminating TN gauges that had point sources within ∼5 km upstream of the sample site (Mengistu et al., 2020). Therefore, point sources generally would not obfuscate this study. Mengistu et al. (2020), upon which this work builds, should serve to clarify aspects of this study. Specifically, we refer to Mengistu et al. (2020) in our manuscript for full explanation of the variables used in our study. Mengistu et al. (2020) states: "We estimated annual TN inputs to each watershed by summing the total atmospheric deposition of N and agricultural N inputs (total N input; kg N/ha/yr). Atmospheric total deposition data were obtained from the 2006 Community Multiscale Air Quality (CMAQ)

Modeling system (USEPA 2006). CMAQ total deposition includes wet plus dry deposition of reduced and oxidized nitrogen species. We further calculated annual agricultural TN inputs to each watershed by summing gridded estimates of fertilizer, manure, and cultivated legumes developed for the 2006 EnviroAtlas national maps derived from county-level data (USEPA 2006; Sobota et al. 2013)."

We used this same watershed-scale variable, total N input", in our random forest model (see supplemental information). This variable was not an important predictor variable in the initial random forest model (Table 1). It was not used in the LME modeling because changes in cultivated area translate more directly into changes in wetland area and cultivated area is highly correlated ($R2 = 92\%$) with total N inputs.

Finally, Wu and Lu (2012) soil nitrate data used in our models were derived from Soil and Water Assessment Tool (SWAT) model simulations. Soil nitrate concentration outputs from the SWAT model are a result of a suite of soil N cycling processes such as mineralization, plant uptake, etc. Therefore, soil N processes, including mineralization of organic N, are implicitly included in our suite of variables. We will clarify this in the methodology so that the reader can see that we considered all the various inputs that we have outlined here.

4. There is also a lack of attention to relevant literature on nitrate concentration dynamics in UMRB rivers. These weaknesses seem to contribute to some misunderstanding and misinterpretation of some of their results discussed below.

Response: Our work emphasizes the role of landscape-scale sources of N for optimal management in the UMRB. We used this framework in developing our study and focused explicitly on linking landscape-scale sources and sinks of N to [TN]. However, we recognize that removal of N, e.g., via denitrification, occurs in streams, rivers, and other settings via biogeochemical and hydrogeochemical controls – and that removal becomes less efficient as stream sizes increase (Mulholland et al., 2008). With respect to nitrogen dynamics in UMRB rivers specifically, current work suggests that the Upper

Mississippi River acts "primarily as a passive nitrate transporter" wherein greater than 85% of incoming nitrogen passes through (Loken et al., 2018). This supports our focus on landscape-scale sources.

To address the reviewer's concern, we will, however, incorporate a broader discussion of [TN] dynamics into the introduction and include a paragraph in the discussion about outstanding questions of TN sources and sinks at this large spatial scale.

5. Not including temperature as a predictor variable may also be a weakness, given the role of temperature in many N processes and given the large latitudinal differences in temperatures across the UMRB.

Response: Though past work has often not necessarily viewed temperature variability as a primary driver of river reach-scale denitrification rates (Gomez-Velez et al., 2015;Findlay, 1995), we appreciate this comment and agree that, as shown by the Arrhenius equation, temperature plays a role in governing the rate at which chemical reactions, including those involved in N dynamics, proceed. We (partially) address the potential effects of temperature on both biogeochemical and hydrological processes via a dynamic potential evapotranspiration (PET) variable and seasonal harmonic variables in our models. We discuss PET on Lines 156-157 of the manuscript, stating that, "Potential evaporation (PET) was estimated following Hargreaves (1994), where daily minimum and maximum temperatures were extracted from PRISM for each watershed." The harmonic variables within the LME incorporated the seasonality of the data, which may implicitly represent changes in temperature. As a practical matter, including temperature directly and balancing that with other dynamics, such as varying labile organic carbon and redox potentials, would involve switching to a physically based model. The parsimonious approach here focusing on key drivers allows for computational tractability and respects data limitations, while representing the primary variables that caused the change in [TN] dynamics observed over the preceding six decades.

6. I found the paper somewhat difficult to follow in places, in part due to my unfamiliarity

with some of the methods used, but also due to what seemed to be irrelevant and unnecessary commentary.

Response: In a revised paper, we will remove extraneous commentary and improve discussion of the methods to improve clarity. Overall, our aim is to make this work broadly accessible and concise.

Specific Comments 7. For some examples of studies that have shown large seasonal swings in nitrate-N concentrations that are consistent with the seasonal variation in TN concentration presented in the manuscript: see Lucey and Goolsby (1993), David et al. (1997), Mitchell et al. (2000), Keefer et al. (2010). In the pre-fertilizer era, Palmer (1903) commented on a diminution of nitrate concentrations in the Kankakee River in the summer months, which he attributed to uptake by aquatic vegetation and reduced drainage from agricultural fields. More recent studies identify denitrification occurring in stream, river, lake and reservoir sediments as being important factors during warm, low flow periods (see Royer et al., 2004; David et al., 2006; Alexander et al. 2009).

Response: While we agree that many of the factors referenced by the citations presented are at play, it is our understanding that contemporary levels of [TN], including peak spring concentrations, would not persist absent ongoing seasonal basin-scale nitrogen loading. We agree with Loken et al. (2018) that nutrient retention by the Upper Mississippi River is quantitatively of limited primary importance. However, we can discuss our results in the context of some of the aforementioned works in a revised manuscript.

8. Lines 371-3 and Figure 8b.: The discussion and the following sentence in the Figure caption are incorrect: "Inverse relationships [between TN concentration and log of discharge] are observed in watersheds in which 50 percent or more of the area is drained artificially by tiles." Inverse relationships are observed downstream of point sources, where greater flow dilutes the N from the point source. The figure 8b illustrates the inverse relationship for a 192 km2 watershed and the only watershed of that size in

their dataset is the Little Calumet River at Munster, Indiana, which is heavily impacted by wastewater discharge.

The relationship illustrated in Figure 8a is very similar to what I have seen for nitrate in many rivers draining tile drained watersheds in Illinois and Iowa: low concentrations at low flow, increase with the log of flow, up to a moderately high discharge, above which there may be a flattening or a decrease of concentrations at very high flow, probably due to depletion of source N and/or increased surface runoff diluting high nitrate water from tile drains. High flow generally mobilizes more sediment and particulate N, which is likely to render the relationship with TN more linear than with nitrate-N.

Response: While high temporal resolution or event-scale data may record other dynamics in tile-drained watersheds we capture concentration-discharge behavior at a different, broader temporal resolution. Specifically, we have snapshots of TN data through time. With these data, we demonstrate an overall trend toward nitrogen source depletion as discharge increases and tiles activate. Our view from a broader temporal scale perspective may therefore differ than if we were focusing on continuous or event-scale data. Our results are, however, relatively consistent with the inverse leg of several of the piecewise functions fit and presented by Marinos et al. (2020) for heavily tile-drained watersheds in the UMRB. Nonetheless, while we saw this relationship in watersheds with >50% tile drainage, there may be other confounding factors at these sites, such as unpermitted/minor point sources, stormwater, etc. We will further emphasize this point through revision.

9. I wonder how their general results might be different if they conducted their modeling analysis only on the majority of watersheds with positive relationships between TN concentration and log of flow. This would likely exclude the watersheds with significant point source inputs and focus on the predominantly agricultural watersheds.

Response: Our dataset contains only five watersheds that contain moderately to highly inverse concentration-discharge relationships. The data from these watersheds is

a small enough fraction of the overall dataset that their exclusion is not expected to substantially change the dataset nor the results. Moreover, it is relevant to note that the mixed effects modeling framework allows for fitting unique (direct or inverse) concentrations-discharge relationships for each watershed.

10. Figure 8c, illustrates no relationship between TN concentration and flow and is from the Rock River at Afton, which seems to be a mixed-use watershed, with considerable agriculture, natural areas, lakes and urban areas. Lakes, of course, tend to act as an N sink, like wetlands. When data from individual rivers is presented, I think some identifying information would be helpful.

Response: Agree. We will include identifiers for figures where specific watersheds are highlighted.

11. They cite Cao et al 2018 on timing of N fertilizer application, but these estimates are highly speculative, based in part of University recommendations, which are not necessarily adopted. Good data on fertilizer application timing is very limited. Actual fertilizer timing is likely to vary by location and year (see Gentry et al. 2014 for one example).

Response: With respect to timing of fertilizer application Cao et al. (2018) reports using state-level US Department of Agriculture-Economic Research Survey data of producer fertilizer practices. Cao et al. (2018) provides 5 km x 5 km annual rates of fertilizer applications across the US from 1850-2015. Therefore, while spatial and temporal uncertainty exists, the location and timing of fertilization is captured by the data at a scale which reasonably fits with our broad temporal scale of season/month as well as the broad spatial scale of large watersheds. While all fertilizer estimates beyond the plot scale are "best estimates", it is unclear which aspect of these surveys the reviewer characterizes as speculative. We welcome further discussion.

12. In addition to fertilizer, animal manures are applied, and mineralization of soil organic N increases as soils warm up in the spring, which are not discussed in this paper.

The manuscript frequently attributes high river TN concentrations to recent fertilizer applications, which may be a factor in some settings, but N concentration at any time is likely to be from a variety of sources and ages. The highest N concentrations typically occur in a wet spring following a drought during the previous growing season that depressed corn yields and N fertilizer uptake (see Loecke et al. 2017). Consequently, much of the elevated N in such a spring is not necessarily from current year fertilization but may be from the previous year as well as mineralization of soil organic N (Gentry et al. 2009).

Response: The importance and inclusion of multiple N sources was addressed in our response to question #3. With the broader comment of timing and accumulation, we somewhat agree; this is an interesting comment that warrants discussion. Our study encompasses 13 years; therefore, Figure 2 represents monthly median [TN] (solid horizontal line) across this period. Based on our measured data, Figure 2 shows a consistent decline in median [TN] between the peak in fertilization (June) and September, the end of the growing season. We suggest that this is indicative of contemporary, sources of TN reaching the stream. More research and resources are required (as we recommend on Lines 398-399) to determine the extent to which one year's [TN] in the stream reflects that same year's seasonal inputs to the landscape. In a revised paper, we would clarify this discussion, noting how we define the terms "legacy" (recent decades) and "contemporary" (recent months or years). As a practical matter, much of what is asserted regarding nitrogen dynamics is based on highly simplified and uncertain models as opposed to critical experiments, which are not deemed practical at the basin scale. In the absence of critical experiments, the reviewer's assertions, though correct qualitatively, are uncertain from a quantitative perspective. This uncertainty is clear, for example, from those studies such as Ilampooranan et al. (2019) that have modeled N dynamics both considering and neglecting legacy effects and found similar model performance. The results of Gentry et al. (2009) cited by the reviewer may explain why year achieved high importance in random forest and reduced (improved) AIC in the mixed effects model. It is also interesting to think about interannual dynamics

from the context of shortcomings of those frameworks, such as mixed effects that do not attempt to account for changes in storage.

13. On their counterfactual modeling: it would be informative to specify the number of hectares or the percentage of cropland converted to wetlands or other land uses. The conversion appears to be rather extensive and if so, they are extrapolating well beyond the data used to develop the model, resulting in highly uncertain projections.

Response: With respect to specifying the area converted, we agree that this is necessary.

Regarding extrapolating beyond the existing data: We used the counterfactual modeling to explore a tradeoff between an increase in wetland area and a presumed corresponding decrease in cultivated area. Based upon the LME model fit, we demonstrate that reduction in cultivated area is of primary importance to reducing [TN] in the stream. Given that cultivated areas in the 82 watersheds ranged from <1% to 92% of the land areas, the percent cultivated areas used in the counterfactual modeling remained within the range of observation (i.e., they were not substantially extrapolated to unrealistic values). We agree with the reviewer that the same would not necessarily be true of wetlands – meaning that wetland restoration has not yet occurred at 50% and 100% of historic distributions across the UMRB. However, our goal here was to explore scenarios to demonstrate the utility of wetlands for reducing [TN] at watershed scales. This requires modeling, and thereby extrapolating, because the critical field experimentâĂŤactually restoring the wetlands in the UMRBâĂŤis unrealized.

14. Furthermore, wetland denitrification is influenced by temperature, and that is not considered in their model.

Response: Site specific factors such as temperature are considered implicitly, through both seasonal fixed effects and random effects in the LME model. Further, we included temperature in our random forest model via a dynamic (monthly) potential evapotranspiration (PET) variable. As in our response to question #5, we discuss this on Lines

156-157 of the manuscript, stating that, "Potential evaporation (PET) was estimated following Hargreaves (1994), where daily minimum and maximum temperatures were extracted from PRISM for each watershed." See also response to (5).

15.Fortunately, the manuscript does not devote much attention to the quantitative model predictions, but to the extent that it does, perhaps a few words about extrapolation and uncertainty are in order.

Response: We will include a statement regarding extrapolation and uncertainty in a revised manuscript. We note that the take-home message for decision-makers from our scenarios is that wetland restoration is, from a scientific perspective, an auspicious means to substantially reduce [TN] UMRB. Modeling approaches are currently our main tool to assess the effects of wetland restoration on [TN] at watershed scales, because field experimentation a priori to implementation has not been executed at large spatial scales.

16. Interpreting their results for impact on N loads is difficult because concentration reductions do not directly translate to load reductions (Royer et al. 2006). It would be difficult to estimate loads at all the sites in the dataset, with some of the sites having as few as 2 samples per year on average.

Response: We focused on concentration changes because of (1) data availability and (2) the ready translation of concentrations to maximum contaminant levels for drinking water, as shown in Figure 3. However, due to issues of sampling sparsity, we elected to not extend the analysis to loads. This could be a viable future research direction to articulate, given that the question is an important one and the model is already built.

17. The manuscript seems unnecessarily long, in part, because a considerable amount of irrelevant, and sometimes incorrect, background is presented in the introductory and methods paragraphs. The analysis focuses on total N, but much of the literature review discusses "nutrients" (N and P) rather than focusing on N.

Response: We agree that there are parts of the Introduction could be trimmed, which will be addressed in a revised manuscript. We also agree that most reference to P is unnecessary, with the possible exception of in framing the global nutrient challenges in the introduction.

18. On line 94, they state that the Mississippi River is the longest river in the US, which is incorrect and irrelevant.

Response: We will omit this sentence in the revised manuscript.

19. On lines 95-6 they state that the UMRB is the largest contributor of "residual" N to the Gulf of Mexico. I am not sure what is meant by "residual". UMRB typically has higher N yields than other parts of the MRB, but the Ohio River typically carries a higher load. URMB loads are less than half of the overall loads to the Gulf of Mexico, so ranking URMB as the highest depends on how other portions of the MRB are divided up.

Response: The term "residual" references N that was not taken up by plants. We will clarify that in the revised manuscript. We agree that both UMR and the Ohio river are the largest N contributors in the Mississippi River basin.

References Cao, P., Lu, C., and Yu, Z.: Historical nitrogen fertilizer use in agricultural ecosystems of the contiguous United States during 1850–2015: application rate, timing, and fertilizer types, Earth System Science Data, 10, 969-984, 10.5194/essd-10-969-2018, 2018. Findlay, S.: Importance of surface-subsurface exchange in stream ecosystems: The hyporheic zone, Limnol. Oceanogr., 40, 159-164, 10.4319/lo.1995.40.1.0159, 1995. Gentry, L. E., David, M. B., Below, F. E., Royer, T. V., and McIsaac, G. F.: Nitrogen mass balance of a tile-drained agricultural watershed in East-Central Illinois, J Environ Qual, 38, 1841-1847, 10.2134/jeq2008.0406, 2009. Gomez-Velez, J. D., Harvey, J. W., Cardenas, M. B., and Kiel, B.: Denitrification in the Mississippi River network controlled by flow through river bedforms, Nat. Geosci., 8, 941-945, 10.1038/ngeo2567, 2015. Ilampooranan, I., Van Meter, K. J.,

and Basu, N. B.: A Race Against Time: Modeling Time Lags in Watershed Response, Water Resour. Res., 55, 3941-3959, 10.1029/2018WR023815, 2019. Loken, L. C., Crawford, J. T., Dornblaser, M. M., Striegl, R. G., Houser, J. N., Turner, P. A., and Stanley, E. H.: Limited nitrate retention capacity in the Upper Mississippi River, Environ. Res. Lett., 13, 10.1088/1748-9326/aacd51, 2018. Marinos, R. E., Van Meter, K. J., and Basu, N. B.: Is the River a Chemostat?: Scale Versus Land Use Controls on Nitrate Concentration-Discharge Dynamics in the Upper Mississippi River Basin, Geophys. Res. Lett., 47, e2020GL087051, 10.1029/2020GL087051, 2020. Mengistu, S. G., Golden, H. E., Lane, C. R., Christensen, J. R., Wine, M. L., D'Amico, E., Prues, A., Leibowitz, S. G., Compton, J. E., Weber, M. H., and Hill, R. A.: Wetland Flowpaths Mediate Nitrogen and Phosphorus Concentrations across the Upper Mississippi River Basin, JAWRA Journal of the American Water Resources Association, 10.1111/1752-1688.12885, 2020. Mulholland, P. J., Helton, A. M., Poole, G. C., Hall, R. O., Hamilton, S. K., Peterson, B. J., Tank, J. L., Ashkenas, L. R., Cooper, L. W., Dahm, C. N., Dodds, W. K., Findlay, S. E. G., Gregory, S. V., Grimm, N. B., Johnson, S. L., McDowell, W. H., Meyer, J. L., Valett, H. M., Webster, J. R., Arango, C. P., Beaulieu, J. J., Bernot, M. J., Burgin, A. J., Crenshaw, C. L., Johnson, L. T., Niederlehner, B. R., O'Brien, J. M., Potter, J. D., Sheibley, R. W., Sobota, D. J., and Thomas, S. M.: Stream denitrification across biomes and its response to anthropogenic nitrate loading, Nature, 452, 202-205, 10.1038/nature06686, 2008.
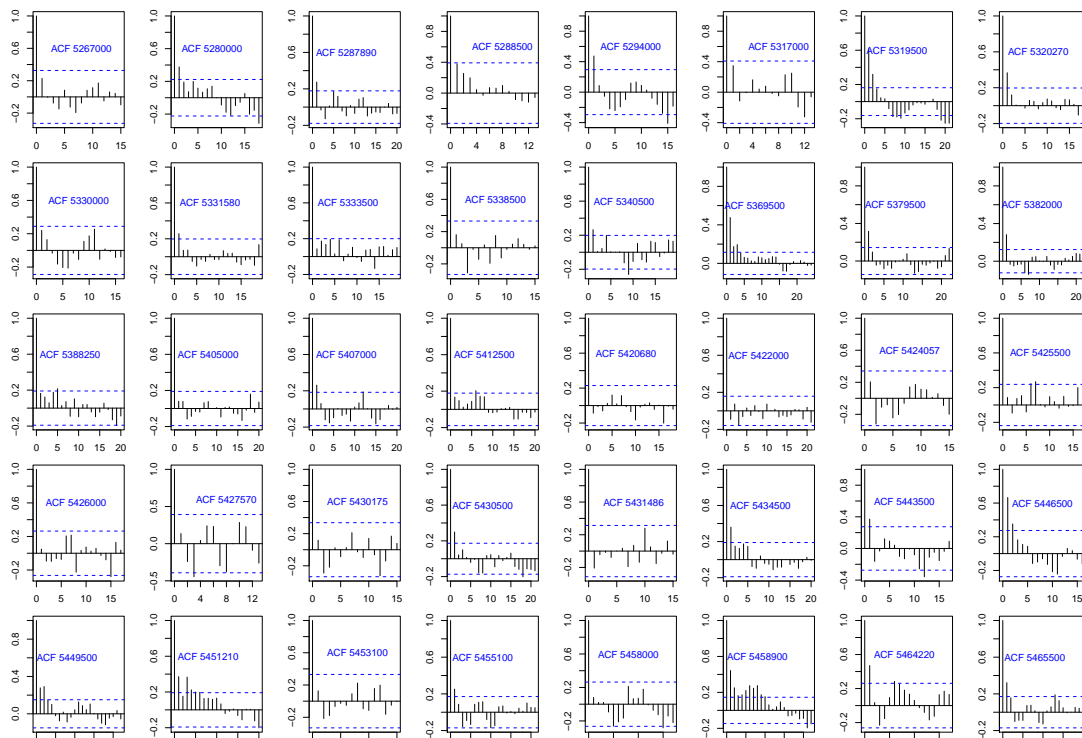
Fig. 1. Representative autocorrelation functions by watershed. Unity indicates perfect correlation, negative unity indicates perfect inverse correlation, and zero indicates no correlation.