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 have significantly improved the manuscript. Our initial responses to the reviewers are indented below and shown in blue text. These are followed by red text, which described the actual changes made to the manuscript and the locations of those changes. Line numbers refer to locations in the clean (unmarked) revised version of the manuscript.

We note that an additional review was provided by Christopher Green (USGS – Menlo Park, CA) for the purpose of internal USGS review. Some minor changes in the revised text reflect responses to this review.

We also note that, per USGS requirements, the following disclaimer should be applied to the **discussion paper** that is permanently posted on the HESS website, if at all possible.

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.
>
> Figure 2 was added to show a cross section and text was added in Line 81. The cross section includes water table elevation and major geologic features. One transect of nested wells is also depicted to illustrate the distribution of wells and screen locations that are typical for the study area. The location of the cross section is also shown in map view in Figure 1.

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.
>
> Lines 137-149 now address some factors affecting nitrate concentrations including the potential for denitrification to removed nitrate from groundwater.

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.

We created a new site map (Figure 1) to improve resolution and clarity, including features described above. Figure 1 also shows the location of the cross section (Figure 2).

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. Recent publications have been added, especially in the Introduction and in our discussion of potential future work.

In the revised manuscript more recent (2019-2020) references have been added in Lines 58-59 and 392.

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.

Lines 137-149 provides a little more background on isotopes as one line of evidence for investigating denitrification. However, available isotope data are much fewer than nitrate concentration data in this study (likely similar to many other monitoring networks) so it is difficult to say how nitrate isotope data would contribute to this particular modeling effort.

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.

Figure 2 was added to show a cross section and text was added in Line 81. The cross section includes water table elevation and major geologic features. One transect of nested wells is also depicted to illustrate the distribution of wells and screen locations that are typical for the study area. The location of the cross section is also shown in map view in Figure 1.

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.

In the revised manuscript, text was added in Lines 76-77, 91-93, 108-111 to emphasize the potential influence of canals and surface irrigation on groundwater in the study area. We also emphasize that crop fields are present across most of the study area and cite Figure S2 as an illustration (Lines 353-354).

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.

The sentence was removed from the abstract, clarified, and reinserted on Lines 341-342.

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.

An existing statement on Lines 84-86 discussed the importance of canals to the region. Additional text was added to Lines 108-111.

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.”

This change was made in Lines 114-119.

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.

The definition was added in Line 183.

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.

This clarification was added in Lines 267-269.

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.

Figure S1 was completely re-worked to show the original data (Figure S1A) and the nitrate data adjusted for total travel time (Figure S1B). These figures are both cited in the main text.

Response to 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).
>
> This change was made in Lines 15-16.

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.
>
> Text added to Line 24 addresses this comment.

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. However, some general introductory sentences were added.
>
> Text was added in Lines 37-39.

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.
>
> The synonyms were added in Line 41 (4^{th} sentence 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.

Example well depths are now depicted in Figure 2 and additional details are included in Lines 114-119.

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.

No changes were made in the revised document.

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. The low correlation may be due to better fertilizer management with agricultural producers applying less fertilizer per hectare than in the past. As a result, we felt this was our best choice for incorporating an important dynamic variable into the study.

Discussion of this issue was expanded in the text between Lines 229-236.

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.

The magnitude of changes is noted in more detail in Lines 229-233.

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).

In Lines 256-261 we clarified that for short-screened wells such as the ones used in this study, the uncertainty (variability) in groundwater age is generally smaller than it might be in long-screen wells. Furthermore, it may be expected that regional changes in nitrate recharge fluxes will be smoothed over a period of years. Thus, our model assumption that individual samples represent approximately discrete travel times.

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.

Example well depths in Figure 2 and additional text in Lines 114 – 119 were added for clarification.

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.

Additional discussion of the potential benefit of a well-defined nitrate input function is included in Lines 409-411. We also note (just for discussion here, not in the manuscript) that in an ideal case, we would have a very well-defined input function *and* long-term time series of groundwater nitrate concentrations. In this case, perhaps recurrent neural networks or other models suitable for time series data could be used for similar study. Unfortunately, we have neither of these data sets. The intermittent sampling of wells is part of the reason we chose the Random Forest model, as traditional time series analyses (or machine learning approaches that leverage time series data) were not suitable for these data.

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.

No change was made to the text in response to this comment. However, we do think the Figures S1A and S1B are interesting and relevant when considering connections between historical nitrate groundwater concentrations and historical N inputs.

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.

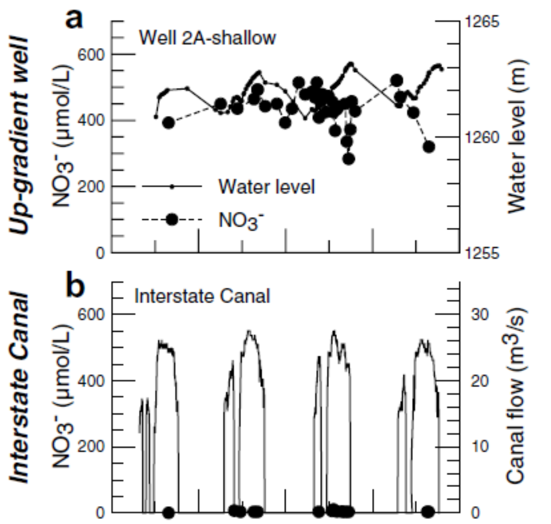
The TT was not calculated based on nitrate, but rather the vertical vadose and saturated zone distance at each well. The rationale was that there is a known relationship between long travel times and low nitrate, and short travel times and high nitrate. Then, we used the Random Forest model to determine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

- Total travel time was estimated for each well as a function of site characteristics (e.g. vadose zone depth) and saturated/unsaturated transport rates.
- Transport rates were varied across specified ranges such that alternative total travel times were constructed for each well
- The model was re-trained and tested for each set of alternative total travel times
- Permutation importance (measured as % increase in MSE or %IncMSE) was calculated for each re-trained model. When calculating permutation importance for total travel time, we are randomly shuffling the total travel time observations across all of the wells to essentially ruin the structure of the dataset. The model is run with this shuffled version of the dataset, and we document that change in error that occurs for the model run with the shuffled data vs. the original correctly-structured data.
- When %IncMSE was high, this indicated the model was sensitive to changes in total travel time.
- The permutation importance of total travel time (%IncMSE) varied depending on the transport rate values used to calculate the total travel time.
- The greatest %IncMSE occurred when the vadose zone transport rate was 3.5 m/yr, and saturated zone transport rate was ~3.7 m/yr. Therefore, we concluded that these were the optimal transport rates for the RF model.

Figure S1B, was added to the Supplemental Information to help illustrate the model outputs, which are consistent with other "reconstructed" input histories for groundwater nitrate (e.g., Puckett et al. 2011). More broadly, we used a machine learning model and appropriate diagnostic tools (%MSE_{inc}, partial dependence plots, etc.) to determine whether the models were reasonable, and we demonstrate that the model captured processes that are consistent with conceptual understanding of the hydrologic system. We also used an independent metric (i.e., independent of our field data; %MSE_{inc}) to select the "optimal" model (and therefore optimal transport rates) and those results were consistent with previous field data. Additional text was also added to Lines 161-164

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?

Previous studies 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 (data collected over a 4-year period with seasonal irrigation flow peaks).



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.

The nitrate concentrations for canal water from previous work were already included in the original manuscript (Line 104 in this revised version).

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 372-373 was changed to state: "the effects of minor and major canals, where groundwater [NO3-] in the vicinity of canals is diluted by canal leakage, and the influence of major canals extends a longer distance when compared to that of minor canals"

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.

The wording was changed to "denitrification" in Line 377.

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

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

Figure 7 is now Figure 8, and this comment is addressed on Line 378 of the revised manuscript.

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.
>
> Reworded to clarify on Lines 381-383

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. Additional data would also help to ensure that the full range of observations are captured in the dataset. Future research could compare results by taking various subsets of the complete data set to provide insight on data requirements.
>
> Lines 444-445 discusses this as a need for future studies to research.

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.
>
> Line 401-404 addresses this issue. We also note that denitrification could also occur in the vadose 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.
>
> The caption for Figure 3 has been re-written. Letters were also added to help distinguish each plot.

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

> There is no difference intended, but the revised manuscript will be updated to maintain a consistent nomenclature for this between the text and figures. Thank you for pointing this out.
>
> This has been done throughout the paper.

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.

Foot note to Table 1 has been corrected.

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.

This is now clarified in Line 712-713 (footnotes of Table 2).

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

Martin J. Wells[1,3], Troy E. Gilmore[2,3], Natalie Nelson[4,5], Aaron Mittelstet[3], ~~John~~ J. ~~Karl~~ K. Böhlke[6],

[1]currently at Natural Resources Conservation Service, Redmond, OR, 97756, USA
[2]Conservation and Survey Division - School of Natural Resources, University of Nebraska, Lincoln, NE, 68583, USA
[3]Biological Systems Engineering, University of Nebraska, Lincoln, NE, 68583, USA
[4]Biological and Agricultural Engineering, North Carolina State University, Raleigh, NC, 27695, USA
[5]Center for Geospatial Analytics, North Carolina State University, Raleigh, NC, 27695, USA
[6]U.S. Geological Survey, Reston, VA, 20192, USA

Correspondence to: Troy E. Gilmore (gilmore@unl.edu)

Abstract. In this study, we explored the use of statistical machine learning and long-term groundwater nitrate monitoring data to estimate vadose-zone and saturated-zone lag times in an irrigated alluvial agricultural setting. Unlike most previous statistical machine learning studies that sought to predict groundwater nitrate concentrations within aquifers, the focus of this study was to leverage available groundwater nitrate concentrations and other environmental variables~~ data~~ to determine mean regional vertical velocities (transport rates) of water and solutes in the vadose zone and saturated zone (3.50 m/year and 3.75 m/year, respectively). ~~Although a saturated-zone velocity that is greater than a vadose-zone velocity would be counterintuitive in most aquifer settings, t~~The statistical machine learning results are consistent with two contrasting primary recharge processes in this Western Nebraska aquifer: (1) diffuse recharge from irrigation and precipitation across the landscape, and (2) focused recharge from leaking irrigation conveyance canals. The vadose-zone mean velocity yielded a mean recharge rate (0.46 m/year) consistent with previous estimates from groundwater age-dating in shallow wells (0.38 m/year). The saturated zone mean velocity yielded a recharge rate (1.31 m/year) that was more consistent with focused recharge from leaky irrigation canals, as indicated by previous results of groundwater age-dating in intermediate-depth wells (1.22 m/year). Collectively, the statistical machine-learning model results are consistent with previous observations of relatively high-water fluxes and short transit times for water and nitrate in the primarily oxic aquifer. Partial dependence plots from the model indicate a sharp threshold where high groundwater nitrate concentrations are mostly associated with total travel times of seven years or less, possibly reflecting some combination of recent management practices and a tendency for nitrate concentrations to be higher in diffuse infiltration recharge than in canal leakage water. Limitations to the machine learning approach include potential non-uniqueness ~~when comparing model performance for~~ of different transport rate combinations when comparing model performance ~~-~~and highlight the need to corroborate statistical model results with a robust conceptual model and complementary information such as groundwater age.

1 Introduction

Nitrate is a common contaminant of groundwater and surface water that can affect drinking water quality and ecosystem health. Responses of aquatic resources to changes in nitrate loading can be complicated by uncertainties related to rates and pathways of nitrate transport from sources to receptors. Lag times for movement of non-point source nitrate contamination through the subsurface are widely recognized (Böhlke, 2002; Meals et al., 2010; Puckett et al., 2011; Van Meter and Basu, 2017) but difficult to measure. Vadose (unsaturated zone) and groundwater (saturated zone~~ zone~~) lag times are of critical importance for monitoring, regulating, and managing the transport of contaminants in groundwater. However, transport time-scales are often generalized ~~spatially and/or temporally~~due to coarse spatial and temporal resolution in data available for ~~for~~ groundwater systems impacted by agricultural activities (Gilmore et al., 2016; Green et al., 2018; Puckett et al., 2011), resulting in a simplified groundwater management approach. Regulators and stakeholders in agricultural landscapes are increasingly in need of more precise and local lag time information to better evaluate and apply regulations and best management practices for the reduction of groundwater nitrate concentrations (e.g., Eberts et al., 2013).

Field-based studies of lag times commonly use expensive groundwater age-dating techniques and/or vadose-zone sampling to estimate nitrate transport rates moving into and through aquifers (Böhlke et al., 2002, 2007; Böhlke and Denver, 1995; Browne and Guldan, 2005; Kennedy et al., 2009; McMahon et al., 2006; Morgenstern et al., 2015; Turkeltaub et al., 2016; Wells et al., 2018). Detailed process-based modelling studies focused on lag times require complex numerical models combined with spatially intensive and/or costly hydrogeological observations (Ilampooranan et al., 2019; Rossman et al., 2014; Russoniello et al., 2016). Thus, efficient but locally-applicable modelling approaches are needed (Green et al., 2018; Liao et al., 2012; Van Meter and Basu, 2015). In this study, an alternative data-driven approach (Random Forest Regression) leverages existing long-term groundwater nitrate concentration (referred to as $[NO_3^-]$ hereafter) data and easily accessible environmental data to estimate vadose and saturated-zone vertical velocities (transport rates) for the determination of subsurface lag times.

Statistical machine learning methods, including Random Forest, have been used successfully for modelling $[NO_3^-]$ distribution in aquifers (Anning et al., 2012; Juntakut et al., 2019; Knoll et al., 2020; Nolan et al., 2014; Ouedraogo et al., 2017; Rodriguez-Galiano et al., 2014; Rahmati et al., 2019; Vanclooster et al., 2020; Wheeler et al., 2015), but there has not been robust analysis of model capabilities for estimating vadose and/or saturated-zone lag times. Proxies for lag time, such as well screen depth, have been used as predictors in Random Forest models (Nolan et al., 2014; Wheeler et al., 2015). Decadal lag times have been suggested from using time-averaged nitrogen inputs as predictors (e.g., 1978-1990 inputs vs 1992-2006 inputs) and by comparing their relative importance in the model (Wheeler et al., 2015). Application of similar machine learning methods suggested groundwater age could be used as a predictor to improve model performance (Ransom et al., 2017). Hybrid models, using both mechanistic models and machine learning, have also sought to integrate vertical transport model parameters and outputs to evaluate nitrate-related predictors, including vadose-zone travel times (Nolan et al., 2018).

The objective of this study is to test a data-driven approach for estimating vadose~~ (unsaturated zone)~~ and ~~groundwater~~ saturated-zone ~~(saturated zone)~~ transport rates and lag times for an intensively monitored alluvial aquifer in western Nebraska (Böhlke et al., 2007; Verstraeten et al., ~~2001b~~2001a, ~~2001a~~2001b; Wells et al., 2018). Results are compared to the hydrogeologic, mechanistic understanding from previous groundwater studies to determine strengths and weaknesses of the approach as (1) a stand-alone technique, or (2) as an exploratory analysis to guide or complement more complex physical-based models or intensive hydrogeologic field investigations.

2 Methods

2.1 Site Description~~s~~

The Dutch Flats study area is located in the western Nebraska counties of Scotts Bluff and Sioux (Fig. 1). The North Platte River delivers large quantities of water for crop irrigation in this region and runs along the southern portion of this study area. Irrigation water is diverted from the North Platte River into three major canals (Mitchell-Gering, Tri-State, and Interstate Canals) that feed a network of minor canals. ~~–~~Several previous Dutch Flats area studies have investigated groundwater characteristics and provided thorough site descriptions of the semi-arid region (Babcock et al., 1951; Böhlke et al., 2007; Verstraeten et al., 2001a, 2001b; Wells et al., 2018). The Dutch Flats area overlies an alluvial aquifer characterized by unconsolidated deposits of predominantly sand and gravel, with the aquifer base largely consisting of consolidated deposits of the Brule, Chadron, or Lance Formation (Verstraeten et al., 1995) (Fig. 2). Irrigation water not derived from the North Platte River is typically pumped from the alluvial aquifer, or water-bearing units of the Brule Formation.

The total area of the Dutch Flats study area is roughly 540 km^2, of which approximately 290 km^2 (53.5%) is agricultural land (cultivated crops and pasture). Most agricultural land is concentrated south of the Interstate Canal (Homer et al., 2015). Due to the combination of intense agriculture and low annual precipitation, producers in Dutch Flats rely on a network of irrigation canals to supply water to the region. From 1908 to 2016, mean precipitation of 390 mm was measured at the nearby Western Regional Airport in Scottsbluff, NE (NOAA, 2017).

While some groundwater is withdrawn for irrigation, and some irrigated acres in the study area are classified as commingled (groundwater and surface water source), Scotts Bluff County irrigation is mostly from surface water sources. Estimates determined every five years suggest surface water provided between 76.8% to 98.6% of the total water withdrawals from 1985 to 2015, or about 92% on average (Dieter et al., 2018). Canals transport water from the North Platte River to fields throughout the study area, most of which are downgradient (south) of~~)~~ the Interstate Canal. ~~Large canals in Dutch Flats include the~~ Mitchell-Gering, Tri-State, and Interstate Canals are the major canals in Dutch Flats, with the latter holding the largest water right of 44.5 m^3/s~~.~~ (NEDNR, 2009). Leakage from these canals provides a source of artificial groundwater recharge. Previous studies estimate the leakage potential of canals in the region results in as much as 40% to 50% of canal water being lost during conveyance (Ball et al., 2006; Harvey and Sibray, 2001; Hobza and Andersen, 2010; Luckey and Cannia, 2006). Leakage estimates from a downstream section of the Interstate Canal (extending to the east of the study area; Hobza and Andersen (2010)) suggest fluxes ranging from 0.08 to 0.7 m day^{-1} through the canal bed. Assuming leakage of 0.39 m day^{-1} ~~leakage~~ over the Interstate Canal bed area (16.8 m width x 55.5 km length) within Dutch Flats yields 4.1 x 10^5 m^3 day^{-1} of leakage. Applied over an on-average 151-day operation period (USBR, 2018), leakage from Interstate Canal alone could approach 6.1 x 10^7 m^3 annually, or about 29% of the annual volume of precipitation in the Dutch Flats area.

A 1990s study investigated both spatial and temporal influences from canals in the Dutch Flats area (Verstraeten et al., ~~2001b~~2001a, ~~2001a~~2001b), with results later synthesized by Böhlke et al. (2007). Canals were found to dilute groundwater [NO_3^-] ~~near~~ locally ~~canals~~ with low-[NO_3^-] (e.g., [NO_3^-] < 0.06 mg N L^{-1} in 1997) ~~canal~~ surface water during irrigation season. $^3H/^3He$ age-dating was used to determine apparent groundwater ages and recharge rates. It was noted that wells near canals displayed evidence of high recharge rates influenced by local canal leakage. Data from wells far from the canals indicated that shallow groundwater was more likely influenced by local irrigation practices (i.e., furrows in fields), while deeper groundwater was impacted by both localized irrigation and canal leakage (Böhlke et al., 2007). Shallow groundwater in the Dutch Flats area has hydrogen and oxygen stable isotopic~~e~~ compositions consistent with surface water sources (i.e., North Platte River and associated canals)~~; (Böhlke et al., 2007; Cherry et al., 2020))~~, indicating that most groundwater intercepted by the monitoring well network has been affected by surface-water irrigat~~ed~~ion recharge (Böhlke et al., 2007; Cherry et al., 2020)~~agricultural~~.

The Dutch Flats area is within the North Platte Natural Resources District (NPNRD), ~~which is 1~~one of 23 groundwater management districts in Nebraska tasked with, among other functions, ~~promoting efforts to improve~~improving water quality and quantity. The NPNRD has a large monitoring well network consisting of 797 wells, 327 of which are nested. Nested well clusters are drilled and constructed such that screen intervals represent (1) "shallow" groundwater intersecting the water table (length of screened interval ~~screen length~~ = 6.1 m), (2) "intermediate" groundwater from mid~~-~~-aquifer depths (length of screened interval ~~screen length~~ = 1.5 m), and deep groundwater near the base of the unconfined aquifer (length of screened interval ~~screen length~~ = 1.5 m). Depending on well location within the Dutch Flats area, depths of the water table and base of aquifer are highly variable, such that shallow, intermediate, and deep wells can have overlapping ranges of depths below land surface (Fig. 2).~~, rather classification (shallow, intermediate, deep)~~

Influenced by both regulatory and economic incentives, the Dutch Flats area has undergone a notable shift in irrigation practices in the last two decades. From 1999 to 2017, center pivot irrigated area has increased by approximately 270%, from roughly 3,830 hectares to 14,253 hectares, or from 13% to 49% of the total agricultural land area, respectively. The majority of this shift in technology has occurred on fields previously ~~irrigated by~~under furrow irrigation. Conventional furrow irrigation has an estimated potential application efficiency ("measure of the fraction of the total volume of water delivered to the farm or field to that which is stored in the root zone to meet the crop evapotranspiration needs," per Irmak et al. (2011)) of 45% to 65%, compared to center pivot sprinklers at 75% to 85% (Irmak et al., 2011). ~~-~~Based on improved irrigation efficiency (between 10-40%), average precipitation throughout growing season (29.5 cm for 15 April to 13 October (Yonts, 2002)), and average water requirements for corn (69.2 cm (Yonts, 2002)), converting furrow irrigated fields to center pivot over the aforementioned 14,253 hectares could represent a difference of 1 x 10^7 m^3 to 6 x 10^7 m^3 in water applied. Those (roughly approximated) differences in water volumes are equivalent to 6-28% of average annual precipitation applied over the Dutch Flats area, suggesting the change in irrigation practice does have potential to alter the water balance in the area.

The hypothesis of lower recharge due to changes in irrigation technology was investigated by Wells et al. (2018) by comparing samples collected in 1998 and 2016. Sample sites were selected based on a well's proximity to fields that observed a conversion in irrigation practices (i.e., furrow irrigation to center pivot) between the two collection periods. While mean recharge rate was not significantly different, a lower recharge rate was indicated by data from 88% of the wells. Long-term Dutch Flats [NO_3^-] trends were also assessed in the study, suggesting decreasing trends (though statistically insignificant) from 1998 to 2016 throughout the Dutch Flats area, and nitrogen isotopes of nitrate indicated little change in biogeochemical processes. For additional background, Wells et al., (2018) provides a more in-depth analysis of recent [NO_3^-] trends in this region (see also, Fig. S1A, which shows ~~only~~ the nitrate data used in the present study).

As ~~with~~in other agricultural areas, nitrate in Dutch Flats groundwater is dependent on nitrogen loading at the land surface, rate of leaching below crop root zones, rate of nitrate transport through the vadose and saturated zones, dilution from focused recharge in the vicinity of canals, rate of discharge from the aquifer (whether from pumping or discharge to surface water bodies), and rates of nitrate reduction (primarily denitrification)~~ rates~~ in the aquifer. Based on nitrogen~~N~~ and oxygen~~O~~ isotopes in nitrate and redox conditions observed in previous studies, denitrification likely has a relatively minor or localized influence on groundwater nitrate in the Dutch Flats area (Wells et al., 2018). Evidence of denitrification (from dissolved gases and isotopes (Böhlke et al., 2007, Wells et al. 2018)) was mostly limited to some of the deepest wells near the bottom of the aquifer. ~~The~~ ~~l~~Leakage of low~~-~~-nitrate water in the m~~The effect of m~~ajor canals ~~leakage is considered largely to be~~causes nitrate dilution in the groundwater (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.

Other long-term changes to the landscape were evaluated by Wells et al. (2018) and ~~have~~ included statistically significant reductions in mean fertilizer application rates (1987–1999 vs. 2000–2012) and volume of water diverted into the Interstate Canal (1983–1999 vs. 2000–2016), while a significant increase in area of planted corn ~~was found~~occurred (1983–1999 vs. 2000–2016). Precipitation was also evaluated, and though the mean has decreased over a similar time period, the trend~~it~~ was not statistically significant ~~(Wells et al., 2018)~~.

2.2 Statistical Machine Learning Modelling Framework

Statistical machine learning uses algorithms to assess and identify complex relationships between variables. Learned relations can be used to uncover nonlinear trends in data that might otherwise be overshadowed when using simple regression techniques (Hastie et al., 2009). In this study we used Random Forest Regression, where Random Forests are created by combining hundreds of unskilled regression trees into one model ensemble, or "forest", which collectively produce skilled and robust predictions (Breiman, 2001). Predictors used in the model represent site-specific explanatory variables (e.g., precipitation, vadose-zone thickness, depth to bottom of screen, etc.) that may impact the response variable, groundwater [NO_3^-]. Additionally, a~~A~~s described in detail in Section 2.5, we estimated a range of total travel times (from land surface to the point of sampling) at each of the wells by ~~then~~ vari~~ed~~ying vadose and saturated~~-~~-zone transport rates~~,~~. ~~and re-trained the model for each set of alternative total travel times.~~ ~~and evaluated t~~The relative importance of ~~the~~ total travel time ~~(from land surface to the point of sampling) variable~~as a predictor variable was ultimately used to identify an optimal travel time ~~select the optimal~~and model.

2.3 Random Forest Application

Random Forest regression models of groundwater [NO_3^-] were developed using five-fold cross validation (Hastie et al., 2009), where four folds were used to build the model (training data), and one fold was held out (testing data). The maximum and minimum of the [NO_3^-] and each predictor were determined and placed into each fold for training models to eliminate the potential for extrapolation during validation. Each fold was used as training data four times, and testing data once. This process was repeated five times to create a total of 25 models, similar to the approach used by Nelson et al. (2018). The four folds designated to build the model underwent a nested five-fold cross validation, as specified in the *trainControl* function within the *caret* (Classification and Regression Training) R package (Kuhn, 2008; R Core Team, 2017). Functions in *caret* were used to train the Random Forest models.

To evaluate model performance, Nash-Sutcliffe Efficiency (NSE), permutation importance, and partial dependence were quantified. NSE indicates the degree to which observed and predicted values deviate from a 1:1 line, and ranges from negative infinity to 1 (Nash and Sutcliffe, 1970).

$$NSE = 1 - \left[\frac{\sum_{i=1}^{n}\left(Y_i^{obs} - Y_i^{pred}\right)^2}{\sum_{i=1}^{n}\left(Y_i^{obs} - Y^{mean}\right)^2}\right], \quad (1)$$

where n is the number of observations, Y_i^{obs} is the ith observation of the response variable ([NO_3^-]), Y_i^{pred} is the ith prediction from the Random Forest model, and Y^{mean} is the mean of observations i through n. Values from negative infinity to 0 suggest the mean of the observed [NO_3^-] would serve as a better predictor than the model. When NSE = 0, model predictions are as accurate as that of a model with only the mean observed [NO_3^-] as a predictor. From 0, larger NSE values indicate a model's predictive ability improves, until NSE = 1, where observations and predictions are equal. NSE was calculated for both the training and testing data.

For each tree, a random bootstrapped sample (i.e., data randomly pulled from the dataset, sampled with replacement) is extracted from the dataset (Efron, 1979), as well as a random subset of predictors to consider fitting at each split. Thus, each tree is grown from a bootstrap sample and random subset of predictors, making trees random and grown independent of the others. Observations not used as bootstrap samples are termed out-of-bag (OOB) data ~~(OOB)~~.

When building a tree, all [NO_3^-] from the bootstrap sample are categorized into terminal nodes, such that each node is averaged and yields a predicted [NO_3^-]. The performance and mean squared error (MSE) of a Random Forest model is evaluated by comparing the observed [NO_3^-] of the OOB data to the average predicted [NO_3^-] from the forest. OOB data from the training dataset may be used to evaluate both permutation importance, referred to in the rest of this text as variable importance, and partial dependence. Variable importance uses percent increase in mean squared error ($\%_{inc}$MSE) to describe predictive power of each predictor in the model (Jones and Linder, 2015). During this process, a single predictor is permuted, or shuffled, in the dataset. Therefore, each observed [NO_3^-] has the same relationship between itself and all predictors, except one permuted variable. The $\%_{inc}$MSE of a variable is determined by comparing the permuted OOB MSE to unpermuted OOB MSE. Important predictors will result in a large $\%_{inc}$MSE, while a variable of minor importance does little to impact a model's performance, as suggested by a low $\%_{inc}$MSE value.

Partial dependence curves serve as a graphical representation of the relationship between [NO_3^-] and predictors in the Random Forest model ensemble (Hastie et al., 2009). Each plot considers the effects of other variables in the model, ~~since~~ because predictions of [NO_3^-] are influenced by several predictors when building each tree. In these models, the y-axis of a partial dependence plot represents the average of the OOB predicted [NO_3^-] at a specific x-value of each predictor.

2.4 Variables and Project Setup

Data from 15 predictors were collected and analysed (Table 1). Spatial variables were manipulated using ArcGIS 10.4. The [NO_3^-] dataset for the entire NPNRD had 10,676 observations from 1979 to 2014, and was downloaded from the Quality-Assessed Agrichemical Contaminant Database for Nebraska Groundwater (University of Nebraska-Lincoln, 2016). Spatial locations for each well were included in the original [NO_3^-] dataset and imported into GIS. Wells were clipped to the Dutch Flats model area, resulting in 2,829 [NO_3^-] observations from 214 wells. In order to have an accurate vadose-zone thickness, only wells with a corresponding depth to groundwater record, of which the most recent record was used, were selected (2,651 observations from 172 wells). Over this period, several wells were sampled much more frequently than others (e.g., monthly sampling, over a short period of record), especially during a U.S. Geological Survey (USGS) National Water-Quality Assessment (NAWQA) study from 1995 to 1999. In order to prevent those wells from dominating the training and testing of the model, annual median [NO_3^-] was calculated for each well and used in the dataset. The dataset was further manipulated such that each median [NO_3^-] observation had 15 complementary predictors (Table 1). The selected predictor variables capture drivers of long-term [NO_3^-] and [NO_3^-] lags. After incorporating all data, including limited records of dissolved oxygen (DO), the final dataset included 1,049 [NO_3^-] observations from 162 wells sampled between 1993 and 2013 (Figure S1A). Additional details of the data selection, sources, and manipulations may be found in the supplemental material.

Predictors were divided into two categories; static and dynamic (Table 1). Static predictors are those that either do not change over the period of record, or annual records were limited. DO, for example, could potentially experience slight annual variations, but data were not available to assign each nitrate sample a unique DO value. Instead, observations for each well were assigned the average DO value observed from the well. This approximation was considered reasonable because nitrate isotopic composition and DO data collected in the 1990s and by Wells et al. (2018) did not indicate any major changes to biogeochemical processes over nearly two decades. Total travel time (from ground surface to the point of sampling) was strictly considered a static predictor in this study and was used to link the nitrate-sampling year to a dynamic predictor value.

Dynamic predictors were defined in this study as data that changed temporally over the study period. Therefore, each annual median [NO_3^-] was assigned a lagged dynamic value to represent the difference between the time of a particular surface activity (e.g., timing of a particular irrigation practice) and when groundwater sampling occurred. Dynamic predictors were available from 1946 to 2013 and included annual precipitation, Interstate Canal discharge, area under center pivot sprinklers, and area of planted corn (~~Fig. 2~~Fig. 3). Dynamic predictors were included to assess their ability to optimize Random Forest groundwater modelling and determine an appropriate lag time. Lag times were based on the vertical travel distance through both the vadose and saturated zones (see Section 2.5). Area of planted corn was included as a proxy for fertilizer data, which were unavailable prior to 1987. However, analysis suggests there has been a 17% reduction (comparing the means of 1987-1999 to 2000-2012) in fertilizer application rates per planted hectare, while area of planted corn has increased 16% (comparing the means of 1983-1999 to 2000-2016) in recent decades (Wells et al., 2018). This trend may be attributed to improved fertilizer management by agricultural producers. There was a likely trade-off in using this proxy; we were able to extend the period of record back to 1946, allowing for analysis of a wider range of lag times in the model, but might have sacrificed some accuracy in recent decades when nitrogen management may have improved. Lastly, vadose and saturated-zone transport rates were assumed to be constant over time (Wells et al., 2018).

2.5 Vadose and Saturated-zone Transport Rate Analysis

Ranges of vertical velocities (transport rates) ~~-~~through the Dutch Flats vadose zone and saturated zone were estimated~~,~~ ~~using Equation 2, from~~ from $^3H/^3He$ age-dating derived recharge rates. The vertical velocities were determined from results published for samples collected in ~~-in the Dutch Flats area determined in in both~~ 1998 (~~-~~Böhlke et al., 2007, Verstraeten et al., ~~2001b~~2001a) and 2016 (Wells et al., 2018) as~~using Equation 2:~~.

$$V = \frac{R}{\theta}, \tag{2}$$

where R is the upper and lower bound of recharge rates (m/yr) ~~indicated by groundwater ages~~, and θ is the mobile water content in the vadose zone or porosity in the saturated zone~~is mobile water content and porosity in the vadose and saturated zone, respectively~~. The ~~use of~~ $^3H/^3He$ data ~~was~~ were used in this study solely for constraining the range of potential transport rates to evaluate in the vadose and saturated zones, and as a base comparison to model results. The age-data, however, were not used by the model itself when seeking to identify an optimum transport rate combination. Throughout the text, unsaturated (vadose)-zone vertical transport rates will be abbreviated as V_u, while saturated-zone vertical transport rates will be V_s. In the vadose zone, θ was assigned a constant value of 0.13, which was calibrated previously using a vertical transport model for the Dutch Flats area (Liao et al., 2012). In the saturated zone, θ was assigned a constant value of 0.35, equal to the value assumed previously for recharge calculations (Böhlke et al., 2007). Vadose and saturated-zone travel times (τ) then were calculated using Equation 3:

$$\tau = \frac{z}{V}, \tag{3}$$

where τ is either vadose zone (τ_u) or saturated zone (τ_s) travel time in years, and z is the vadose-zone thickness (z_u) or distance from the water table to well mid-screen (z_s) in meters.

Though Equations 2 and 3 do not explicitly consider horizontal groundwater flow, they are believed to adequately model shallow groundwater ages, which are likely to follow approximately linear vertical age gradients near the water table. These simple equations are also suggested to sufficiently estimate groundwater age gradients in wedge-shaped aquifers (Cook and Böhlke, 2000), and Böhlke et al. (2007) found a linear model adequately fit their data in the Dutch Flats area. Discrete transport rates and travel times calculated from Equations 2 and 3 should be considered "apparent" rates and travel times, similar to apparent groundwater ages, which are based on imperfect tracers and may be affected by dispersion and mixing. Nonetheless, the saturated open intervals of the monitoring wells used for this study (< 6.1 m for shallow wells; 1.5 m for intermediate and deep wells) generally were short compared with the aquifer thickness, such that age distributions of individual samples were relatively restricted in comparison to those of the whole aquifer or wells with long screened intervals. ~~Nonetheless, saturated-zone transport rates and travel times calculated from Equations 2 and 3 should be considered "apparent" rates and travel times. By "apparent" we mean that uncertainty is probablyle (i.e., ssimilar to apparent groundwater ages, which are based on imperfect tracers. For monitoring wells with short screens (1-2 meters; very short compared to irrigation wells or municipal wells) uncertainties in the mean age of a groundwater sample are typically plus or minus a few years, depending on the groundwater age and the environmental tracer used).~~ Additionally, it is emphasized that the assumed mobile water content of 0.13 is a calibrated parameter derived previously through inverse modelling and, as suggested by Liao et al. (2012), may have large uncertainties due to the varying site-specific characteristics known to exist from one well to the next.

Because of the influence of canal leakage on both intermediate and deep wells (Böhlke et al., 2007), only recharge rates from shallow wells were used to estimate initial values and permissible ranges of vadose-zone travel times. The mean ($\bar{x}$ = 0.38 m/yr) and standard deviation ($\sigma = \pm 0.23$ m/yr) of all the 1998 (n=7) and 2016 (n=2) shallow recharge rates were calculated. Using $\bar{x} \pm 1\sigma$, a range of recharge rates from 0.15 to 0.61 m/yr were converted to transport rates (V_u) using Equation 2. One standard deviation was selected to constrain the range of rates evaluated, as we considered this method likely encompassed realistic mean field values. ~~This calculation~~Calculated transport rates resulted in 1.15 to 4.69 m/yr as the range of vadose-zone transport rates. Expanding the upper and lower bounds, a minimum vadose-zone transport rate of 1.0 m/yr and maximum of 4.75 m/yr was applied. Vertical transport rates in the vadose zone were increased by increments of 0.25 m/yr from 1.0 to 4.75 m/yr, resulting in 16 possible vadose-zone transport rates to evaluate in the Random Forest model.

Mean ($\bar{x}$ = 0.84 m/yr) and standard deviation ($\sigma = \pm 0.73$ m/yr) of all shallow, intermediate, and deep well recharge rates were included in identifying a range of saturated-zone recharge rates from 0.10 to 1.57 m/yr. A total of 35 and 8 recharge rates were used from the Böhlke et al. (2007)~~1990s~~ and Wells et al. (2018) studies, respectively. Equation 2 was used to calculate saturated-zone transport rates (V_s) of 0.28 and 4.49 m/yr. ~~Saturated~~ Saturated-zone transport rates were increased by increments of 0.25 m/yr, from 0.25 to 4.5 m/yr, resulting in 18 unique saturated-zone transport rates to evaluate in the Random Forest model. The range of transport rates suggested by groundwater age-dating was large (more than an order of magnitude) and ~~would be~~are considered to include rates likely to be expected in a variety of field settings. Presumably, similar~~the same~~ model constraints and results could have been obtained without the prior age data and with some relatively conservative estimates.

Travel times τ_u and τ_s were calculated for each well based on z_u and z_s, respectively. For every possible combination of vadose and saturated-zone transport rates, a unique total travel time, τ_t, was calculated for each well based on the vadose and saturated-zone dimensions of that particular well.

$$\tau_t = \tau_u + \tau_s, \quad (4)$$

The total travel times from Equation 4 were used to lag dynamic predictors relative to each nitrate sample date. For instance, a nitrate sample collected in 2010 at a well with a 20-year total travel time (e.g., τ_u = 10 yrs and τ_s = 10 yrs) would be assigned the 1990 values for precipitation (450 mm), Interstate Canal discharge (0.4 km^3/yr), center pivot irrigated area (2484 hectares), and area of planted corn (8905 hectares).

A total of 288 unique transport rate combinations (corresponding to different combinations of the 16 vadose and 18 saturated-zone transport rates) were joined into a single dataset totalling over 300,000 observations to determine the optimal rate resulting in the maximum testing NSE from the model. Each transport rate combination incorporated up to 1,049 groundwater

$[NO_3^-]$ values. To decrease runtime, Random Forest models were parallel processed through a Holland Computing Center (HCC) cluster at the University of Nebraska-Lincoln.

3 Results and Discussion

This study addressed a relatively unexplored use of Random Forest, which was to identify optimal lag times based on testing a range of transport rate combinations through the vadose and saturated zones, historical $[NO_3^-]$~~nitrate concentrations~~, and the use of easily accessible environmental datasets.

3.1. Relative Importance of Transport Time and Dynamic Variables

In our initial modelling with dynamic predictors, we anticipated that we could use the Random Forest model with the highest NSE to identify the optimal pair of vadose and saturated-zone transport rates. However, no clear pattern emerged among the different models (~~Fig. 3~~Fig. 4). Given the small differences and lack of defined pattern in testing NSE values, we selected ten transport rate combinations (the five top performing models, plus four transport rate combinations of high and low transport rates, and one intermediate transport rate combination) for further evaluation of variable importance and sensitivity to a range of transport rate combinations (Table 2). Median total travel time ranked third in variable importance, while the four dynamic variables consistently had the four lowest rankings (~~Fig. 4~~Fig. 5). Total travel time also had the greatest variability in importance among the fifteen variables, with a range of 18.4% between the upper and lower values, suggesting some model sensitivity to lag times. Excluding total travel time, the remaining variables had an average variable importance range of 6%.

Dynamic variables had little influence on the model, despite common potential linkages to groundwater $[NO_3^-]$ (Böhlke et al., 2007; Exner et al., 2010; Spalding et al., 2001). A pattern emerged among dynamic variables where the stronger the historical trend of the predictor, the greater the importance of the predictor (~~Fig. 2~~Fig. 3; Fig.~~4~~ 5). For instance, center~~-~~ pivot irrigated area (highest ranking dynamic variable) had the least noise and the most pronounced trend, while annual precipitation (lowest ranking variable) was highly variable and lacked any trend over time (~~Fig. 2~~Fig. 3), and also may not be a substantial source of recharge (Böhlke et al., 2007). Further exploration could be done to test more refined variables – for instance, annual median rainfall intensity for the growing season might have a more direct connection to nitrate leaching than total annual precipitation. However, rainfall intensity data are not readily available. Likewise, availability of a long-term, detailed fertilizer loading dataset would be advantageous in providing a more substantiated conclusion regarding the viability of applying dynamic variables to determine vadose and saturated-zone lag. Dynamic variables could be of more use in other study areas that undergo relatively rapid and pronounced changes (e.g., land use). In future work, the model sensitivity to dynamic variables could be tested through formal sensitivity analysis and/or automated variable selection algorithms (Eibe et al., 2016).

Ultimately, results from initial analyses suggest that (1) the dynamic data did little to improve model performance, and (2) Random Forest was not able to relate the four considered dynamic predictors to $[NO_3^-]$ in a meaningful way that could be used to estimate lag time~~. It has also been suggested by Katz et al. (2001) that a monotonic trend in an independent variable is not necessarily linearly related to the dependent variable.~~ It is likely the influence of these dynamic predictors ~~are~~ is dampened as nitrate is transported from the surface to wells such that data-driven approaches are unable to sort through noise to identify relationships.

3.2 Use of Random Forest to determine transport rates

Due to their low relative importance as predictors, all four dynamic predictors were removed in the subsequent analysis. As discussed above, a notable variation in total travel time $\%_{inc}$MSE was observed in ~~Fig. 4~~Fig. 5, suggesting model sensitivity to this variable. Additionally, a relationship between travel time and [NO_3^-] has been suggested in the Dutch Flats area through previous studies (Böhlke et al., 2007; Wells et al., 2018). Therefore, a second analysis of just the 11 static predictors was performed over the full range of vadose and saturated transport rates (i.e., 288 combinations). However, in the second analysis, model sensitivity to total travel time – evaluated with respect to the transport rate combination corresponding to the largest $\%_{inc}$MSE of total travel time – was used to determine a distinguished transport rate combination. In other words, models were re-trained and tested for all transport rate combinations, each of which produced a unique set of values for the total travel time variable. As described in Section 2.3, the $\%_{inc}$MSE value for total travel time was then based on the error induced in the model by permuting the calculated total travel times across all the nitrate observations (i.e., randomly shuffling the total travel time variable, and thus disturbing the structure of the dataset).

Formatted: Indent: First line: 0.5"

The Random Forest models were useful in identifying the relative magnitudes of V_u and V_s that led to high $\%_{inc}$MSE. Based on the heat map of $\%_{inc}$MSE, a band of transport rate combinations with consistently high $\%_{inc}$MSE was visually apparent (~~Fig. 5~~Fig. 6). The upper and lower bounds of the band translate to transport rate ratios (V_s/V_u) ranging from 0.9 to 1.5, and are values that could be useful in constraining recharge and/or transport rate estimates in more complex mechanistic models, as part of a hybrid modelling approach. This is especially important ~~since~~ because recharge is one of the most sensitive parameters in a groundwater model (Mittelstet et al., 2011), yet one with high uncertainty. Whereas a saturated-zone velocity that is greater than a vadose-zone velocity would be unexpected in many unconsolidated surficial aquifers receiving distributed recharge, the statistical machine learning results are consistent with two contrasting primary recharge processes in the Dutch Flats area: (1) diffuse recharge from irrigation and precipitation across the landscape, and (2) focused recharge from leaking irrigation conveyance canals.

The $\%_{inc}$MSE of total travel time in the second analysis ranged from 20.6 to 31.5%, with the largest $\%_{inc}$MSE associated with vadose and saturated-zone transport rates of 3.50 m/yr and 3.75 m/yr, respectively (~~Fig. 5~~Fig. 6), and the top four predictors for this transport rate combination were total travel time, vadose-zone thickness, dissolved oxygen, and saturated thickness (~~Fig. 6~~Fig. 7). Converting those vadose and saturated-zone transport rates to recharge rates yielded values of 0.46 m/yr and 1.31 m/yr, respectively. Such a large difference between the two recharge values ~~would be unexpected in most unconsolidated surficial (water-table) aquifers receiving diffuse recharge, but it~~ is consistent with the hydrologic conceptual model of the Dutch Flats area. In fact, both model recharge rates compare favourably with recharge rates calculated from the previous Dutch Flats studies using ^{3}H/^{3}He age-dating (Böhlke et al., 2007; Wells et al., 2018). For instance, the recharge rate determined from the vadose-zone transport rate in this study (0.46 m/yr) was comparable to the mean recharge rate of 0.38 m/yr (n = 9) from groundwater age-dating at shallow wells, which are most representative of diffuse recharge below crop fields that are present across most of the study area (e.g., Figure S2). Additionally, the recharge rate (1.31 m/yr) determined from the saturated-zone transport rate was consistent with the mean recharge value derived from groundwater ages in intermediate wells (1.22 m/yr, n = 13). Intermediate wells are variably impacted by focused recharge from canals in upgradient areas. Given the similarity in diffuse recharge and focused recharge estimates from both Random Forest and groundwater age-dating, the transport rate ratios (1.2 and 1.1, respectively) were consistent. That is, the Random Forest modelling framework produced transport rates consistent with the major hydrological processes in Dutch Flats both in direct (i.e., transport rate estimates) and relative (i.e., transport rate ratio) terms.

Assuming the Random Forest approach has accurately captured the two major recharge processes (diffuse recharge over crop fields and focused recharge from canals), a comparison of recharge rates from all sampled groundwater wells representative of recharge to the groundwater system as a whole (0.84 m/yr, n = 43) to the recharge rates from Random Forest modelling (0.46

and 1.31 m/yr) would provide an estimate of the relative importance of diffuse versus focused recharge on overall recharge in Dutch Flats. Under these assumptions, diffuse recharge would account for approximately 55%, while focused recharge would account for about 45% of total recharge in the Dutch Flats area. Similarly, Böhlke et al. (2007) concluded that these two recharge sources contributed roughly equally to the aquifer on the basis of groundwater age profiles, as well as from dissolved atmospheric gas data indicating mean recharge temperatures between those expected of diffuse infiltration and focused canal leakage.

Partial dependence plots, which illustrate the impact a single predictor has on [NO_3^-] in the model with respect to other predictors (~~Fig. 7~~Fig. 8), largely reflect the conceptual understanding of the system from previous studies including Böhlke et al. (2007) and Wells et al. (2018). Key features that strengthen confidence in the Random Forest modelling include (1) depth to bottom screen, where groundwater [NO_3^-] is lower at greater depths, (2) 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~~a longer distance when compared to that of minor canals~~ results~~, (3) land surface elevation, where elevations indicating proximity to major canals are associated with relatively lower groundwater [NO_3^-], and (4) DO concentration, where higher DO concentration is linked to higher groundwater [NO_3^-]. We note that decreasing DO and [NO_3^-] with groundwater age can be explained by DO reduction and historical changes in [NO_3^-] recharge, whereas groundwater chemistry and nitrate isotopic data recorded in both this study and previous Dutch Flats studies suggest ~~nitrate reduction~~denitrification was not a major factor in this alluvial aquifer.

The partial dependence plot (Fig. 8) for total travel time exhibits a pronounced threshold, where [NO_3^-] is markedly higher for groundwater with travel time less than seven years. It is possible this reflects long-term stratification of ~~distinct~~ groundwater [NO_3^-], stemming from the suggested patterns stated above as ~~it relates to~~nitrate varies with aquifer depth ~~and~~ due to the influences of diffuse and focused recharge in the region. This seven-year threshold is slightly lower than a previous estimate of mean groundwater age in the aquifer (8.8 years; Böhlke et al., 2007; where groundwater age excludes vadose-zone travel time) and suggests that shallow groundwater can respond relatively rapidly to changes in nitrogen management in the Dutch Flats area.~~lower than a previous estimate of mean groundwater age alone (8.8 years; where groundwater age neglects vadose-zone travel time, Böhlke et al. (2007)) and suggests that rapid aquifer response, compared to previous age-dating studies, to changes in nitrogen management in Dutch Flats is possible.~~

3.3 Opportunities and limitations of Random Forest approach in estimating lag times

Overall, results suggest that in a complex system such as Dutch Flats, Random Forest was able to identify reasonable transport rates for both the vadose and saturated zones, and with additional validation, this method may offer an inexpensive (i.e., compared to groundwater age-dating across a large monitoring well network and/or complex modelling) and reasonable technique for estimating lag time from historical monitoring data. Further, this approach allows for additional insight on groundwater dynamics to be extracted from existing monitoring data. However, this study was conducted in the context of a larger project (Wells et al., 2018) and built on prior research on groundwater flow and [NO_3^-]~~nitrate concentrations~~ in the study area (Böhlke et al., 2007). Therefore, it is critical in future work to ~~consider the “black box” nature of statistical machine learning approach~~incorporate site-specific knowledge, process understanding, and approaches for increasing interpretability of machine learning models (Lundberg et al., 2020, Saia et al., 2020), as highlighted in key considerations below.

Some key considerations for future application of this approach include:

(1) The Random Forest approach might be useful for estimating future recharge and [NO_3^-] using multiple potential management scenarios, as long as considered management scenarios fall within the range of historical observations used

to train the model. This information could be used to inform policy makers of the impact that current and future management decisions will have on recharge and $[NO_3^-]$.

(2) The Dutch Flats overlies a predominantly oxic aquifer, where nitrate transport is mostly conservative. In aquifers with both oxic and anoxic conditions and distinct nitrate extinction depths (Liao et al., 2012; Welch et al., 2011), this approach may be biased toward oxic portions of the aquifer where the nitrate signal is preserved. Similarly, vertical profiles of $[NO_3^-]$ and isotopic composition in the vadose zone could provide valuable data to investigate (1) the amount of nitrate stored in the vadose zone, and (2) whether nitrate undergoes any biogeochemical changes while being transported through the vadose zone to the water table.~~Similarly, vertical sampling for nitrate and nitrate isotopes in the vadose zone would provide valuable data to investigate (1) the amount of nitrate stored in the vadose zone, and (2) whether nitrate stored in the unsaturated zone undergoes any biogeochemical changes that may reduce its signal by the time it reaches groundwater.~~

(3) While estimates of vadose and saturated-zone transport rates determined from $\%_{inc}$MSE are consistent with previous studies~~s~~, the predictive performance of the selected model (based on NSE and visual inspection of predicted versus observed nitrate plots) was not substantially different than other models tested. In other words, the "optimal model" was only weakly preferred~~non-unique~~ in terms of predicting $[NO_3^-]$. Testing the approach of using $\%_{inc}$MSE in other vadose and saturated zones, with substantial comparison to previous transport rate estimates, is warranted. This would be especially valuable in an area with a well-defined input function for nitrate that could be compared to a reconstructed input function from the model. Further, ~~(i $[NO_3^-]$ adjusted for total travel time, e.g., Figure S1B).~~in aquifer settings with relatively evenly distributed recharge, optimized travel times to wells could be used to estimate the infiltration date of samples, thus providing an optimized view of historical variation of $[NO_3^-]$ entering the subsurface, as illustrated in Figure S1B. In the Dutch Flats area, however, such an analysis is complicated by effects of subsurface nitrate dilution by local recharge from canal leakage.

(4) Despite potential non-uniqueness in prediction metrics, the heat map of $\%_{inc}$MSE did reveal an orderly pattern suggesting consistent transport rate ratios. For modelling efforts where recharge rates are a key calibration parameter, identification of a range of reasonable recharge rates, and/or the ratio of recharge rates from diffuse and focused recharge sources for a complex system will reduce model uncertainty and improve results. This statistical machine learning approach, which essentially leverages nitrate as a tracer (albeit with an unknown input function in this case), may provide valuable insight to complement relatively expensive groundwater age-dating or vadose-zone monitoring data, or as a standalone approach for first-order approximations.

(5) The demonstrated statistical machine learning approach is apparently well-suited for drawing out transport rate information from a site with two distinct recharge sources (diffuse versus focused recharge sources) driving the groundwater nitrate dynamics. Further testing is needed at sites where recharge and nitrate dynamics are more subtle.

4 Conclusions

The Dutch Flats area ~~consists~~ exhibits ~~of~~ large variations in $[NO_3^-]$ throughout a relatively small region in western Nebraska. Long-term groundwater $[NO_3^-]$ monitoring and previous groundwater age-dating studies in Dutch Flats provided an opportune setting to test a new application of statistical machine learning (Random Forest) for determining vadose and saturated-zone transport rates. Overall results suggest Random Forest has the capability to both identify reasonable transport rates (and lag time) and key variables influencing groundwater $[NO_3^-]$, albeit with potential for non-unique results. Limitations were also identified when using dynamic predictors to model groundwater $[NO_3^-]$. Utilizing only static predictors, and Random Forest's

Formatted: Subscript

Formatted: Superscript

ability to evaluate variable importance, ~~a~~ vadose-zone and saturated-zone transport rates ~~was~~ were selected based on model sensitivity to changing the total travel time predictor. In other words, total travel time variable importance was evaluated for 288 different transport rate combinations, and the combination with a total travel time having the largest influence over the model's ability to predict $[NO_3^-]$ was selected for additional examination. This analysis identified a vadose-zone and saturated-zone transport rate combination consistent with rates previously estimated from $^3H/^3He$ age-dating in Böhlke et al. (2007) and Wells et al. (2018), indicating a combination of distributed and focused sources of irrigation recharge to this aquifer~~in both direct and relative terms.~~

Future studies should include assessments of the proper conditions for application of dynamic predictors and include comparisons of data-driven analyses with complementary datasets. Despite noted limitations, partial dependence plots and relative importance of predictors were largely consistent with previous findings and mechanistic understanding of the study area, giving greater confidence in model outputs. The influence of canal leakage on groundwater recharge rates and $[NO_3^-]$, for example, was consistent with previous Dutch Flats studies. Partial dependence plots suggest a threshold of higher $[NO_3^-]$ for groundwater with total travel time (vadose and saturated-zone travel times, combined) of less than seven years, indicating the potential for relatively rapid groundwater $[NO_3^-]$ response to widespread implementation of best management practices. Additionally, research is needed to determine the minimum number of observations needed to effectively apply the framework shown here.

Author contribution: TG, AM, and NN were responsible for conceptualization. MW and NN developed the model code and MW performed formal analysis. MW prepared the manuscript from his M.S. thesis with contributions from all co-authors, including JKB. TG was responsible for project administration and funding acquisition.

Acknowledgements: The authors acknowledge the North Platte Natural Resources District for providing technical assistance and resources, including long-term groundwater nitrate data accessed via the Quality-Assessed Agrichemical Contaminant Database for Nebraska Groundwater. We ~~also~~ thank Steve Sibray and Mason Johnson for their support in field sampling efforts and Les Howard for cartography. We also thank Christopher Green, Sophie Ehrhardt, Scott Gardner, and an anonymous reviewer for helpful comments on earlier versions of the paper. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Funding: This work was supported by the U.S. Geological Survey 104b Program (Project 2016NE286B), U.S. Department of Agriculture—National Institute of Food and Agriculture NEB-21-177 (Hatch Project 1015698), and Daugherty Water for Food Global Institute Graduate Student Fellowship.

Code and Data Availability: Code is available on request. Data used in the random forest model and described in the supplemental information is available via the University of Nebraska – Lincoln Data Repository (https://doi.org/10.32873/unl.dr.20200428).

References

Anning, D. W., Paul, A. P., McKinney, T. S., Huntington, J. M., Bexfield, L. M. and Thiros, S. A.: Predicted nitrate and arsenic concentrations in basin-fill aquifers of the Southwestern United States, Report, United States Geological Survey. [online] Available from: https://pubs.usgs.gov/sir/2012/5065/, 2012.

Babcock, H. M., Visher, F. N. and Durum, W. H.: Ground-water conditions in the Dutch Flats area, Scotts Bluff and Sioux Counties, Nebraska, with a section on chemical quality of the ground water, Report. [online] Available from: http://pubs.er.usgs.gov/publication/cir126, 1951.

Ball, L. B., Kress, W. H., Steele, G. V., Cannia, J. C. and Andersen, M. J.: Determination of canal leakage potential using continuous resistivity profiling techniques, Interstate and Tri-State Canals, western Nebraska and eastern Wyoming, 2004, Report, United States Geological Survey. [online] Available from: http://pubs.er.usgs.gov/publication/sir20065032, 2006.

Böhlke, J. K.: Groundwater recharge and agricultural contamination, Hydrogeology Journal, 10(1), 153–179, doi:10.1007/s10040-001-0183-3, 2002.

Böhlke, J. K. and Denver, J. M.: Combined Use of Groundwater Dating, Chemical, and Isotopic Analyses to Resolve the History and Fate of Nitrate Contamination in Two Agricultural Watersheds, Atlantic Coastal Plain, Maryland, Water Resources Research, 31(9), 2319–2339, doi:10.1029/95WR01584, 1995.

Böhlke, J. K., Wanty, R., Tuttle, M., Delin, G. and Landon, M.: Denitrification in the recharge area and discharge area of a transient agricultural nitrate plume in a glacial outwash sand aquifer, Minnesota: Denitrification in recharge and discharge areas, Water Resources Research, 38(7), 10-1-10–26, doi:10.1029/2001WR000663, 2002.

Böhlke, J. K., Verstraeten, I. M. and Kraemer, T. F.: Effects of surface-water irrigation on sources, fluxes, and residence times of water, nitrate, and uranium in an alluvial aquifer, Applied Geochemistry, 22(1), 152–174, doi:10.1016/j.apgeochem.2006.08.019, 2007.

Breiman, L.: Random Forests, Machine Learning, 45(1), 5–32, doi:10.1023/A:1010933404324, 2001.

Browne, B. A. and Guldan, N. M.: Understanding Long-Term Baseflow Water Quality Trends Using a Synoptic Survey of the Ground Water–Surface Water Interface, Central Wisconsin, Journal of Environment Quality, 34(3), 825, doi:10.2134/jeq2004.0134, 2005.

Cherry, M., Gilmore, T., Mittelstet, A., Gastmans, D., Santos, V. and Gates, J. B.: Recharge seasonality based on stable isotopes: Nongrowing season bias altered by irrigation in Nebraska, Hydrological Processes, doi:10.1002/hyp.13683, 2020.

Cook, P. G. and Böhlke, J. K.: Determining Timescales for Groundwater Flow and Solute Transport, in Environmental Tracers in Subsurface Hydrology, edited by P. G. Cook and A. L. Herczeg, pp. 1–30, Springer US, Boston, MA., 2000.

Dieter, C. A., Maupin, M. A., Caldwell, R. R., Harris, M. A., Ivahnenko, T. I., Lovelace, J. K., Barber, N. L. and Linsey, K. S.: Estimated use of water in the United States in 2015, Report, Reston, VA., 2018.

Eberts, S. M., Thomas, M. S. and Jagucki, M. L.: The quality of our Nation's waters—Factors affecting public-supply-well vulnerability to contamination—Understanding observed water quality and anticipating future water quality, U.S. Geological Survey Circular 1385. [online] Available from: https://pubs.usgs.gov/circ/1385/, 2013.

Efron, B.: Bootstrap Methods: Another Look at the Jackknife, The Annals of Statistics, 7(1), 1–26, doi:10.1214/aos/1176344552, 1979.

Eibe, F., Hall, M. A. and Witten, I. H.: The WEKA Workbench, in Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques," Morgan Kaufmann., 2016.

Exner, M. E., Perea-Estrada, H. and Spalding, R. F.: Long-Term Response of Groundwater Nitrate Concentrations to Management Regulations in Nebraska's Central Platte Valley, The Scientific World Journal, 10, 286–297, doi:10.1100/tsw.2010.25, 2010.

Gilmore, T. E., Genereux, D. P., Solomon, D. K. and Solder, J. E.: Groundwater transit time distribution and mean from streambed sampling in an agricultural coastal plain watershed, North Carolina, USA: Groundwater transit time, Water Resources Research, 52(3), 2025–2044, doi:10.1002/2015WR017600, 2016.

Green, C. T., Liao, L., Nolan, B. T., Juckem, P. F., Shope, C. L., Tesoriero, A. J. and Jurgens, B. C.: Regional Variability of Nitrate Fluxes in the Unsaturated Zone and Groundwater, Wisconsin, USA, Water Resources Research, 54(1), 301–322, doi:10.1002/2017WR022012, 2018.

Harvey, F. E. and Sibray, S. S.: Delineating Ground Water Recharge from Leaking Irrigation Canals Using Water Chemistry and Isotopes, Ground Water, 39(3), 408–421, doi:10.1111/j.1745-6584.2001.tb02325.x, 2001.

Hastie, T., Tibshirani, R. and Friedman, J. H.: The elements of statistical learning: data mining, inference, and prediction, 2nd ed., Springer, New York, NY., 2009.

Hobza, C. M. and Andersen, M. J.: Quantifying canal leakage rates using a mass-balance approach and heat-based hydraulic conductivity estimates in selected irrigation canals, western Nebraska, 2007 through 2009, Report, United States Geological Survey. [online] Available from: http://pubs.er.usgs.gov/publication/sir20105226, 2010.

Homer, C. G., Dewitz, J., Yang, L., Jin, S., Danielson, P., Xian, G. Z., Coulston, J., Herold, N., Wickham, J. and Megown, K.: Completion of the 2011 National Land Cover Database for the conterminous United States – Representing a decade of land cover change information, Photogrammetric Engineering and Remote Sensing, 81, 345354, 2015.

Hudson, C. (NPNRD): Personal Communication with M.J. Wells, University of Nebraska, Lincoln, NE, USA, 2018.

Ilampooranan, I., Van Meter, K. J. and Basu, N. B.: A Race against Time: Modelling Time Lags in Watershed Response, Water Resources Research, doi:10.1029/2018WR023815, 2019.

Irmak, S., Odhiambo, L., Kranz, W. L. and Eisenhauer, D. E.: Irrigation Efficiency And Uniformity, And Crop Water Use Efficiency, Extension Circular, University of Nebraska – Lincoln, Lincoln, NE. [online] Available from: Available at http://extensionpubs.unl.edu/, 2011.

Jones, Z. M. and Linder, F. J.: Exploratory Data Analysis using Random Forests, in 73rd Annual MPSA Conference, pp. 1–31. [online] Available from: http://zmjones.com/static/papers/rfss_manuscript.pdf (Accessed 25 May 2018), 2015.

Juntakut, P., Snow, D. D., Haacker, E. M. K. and Ray, C.: The long term effect of agricultural, vadose zone and climatic factors on nitrate contamination in Nebraska's groundwater system, Journal of Contaminant Hydrology, 220, 33–48, doi:10.1016/j.jconhyd.2018.11.007, 2019.

Formatted: Line spacing: single

Formatted: Normal

~~Katz, B. G., Böhlke, J. K. and Hornsby, H. D.: Timescales for nitrate contamination of spring waters, northern Florida, USA, Chemical Geology, 179(1–4), 167–186, doi:10.1016/S0009-2541(01)00321-7, 2001.~~

Kennedy, C. D., Genereux, D. P., Corbett, D. R. and Mitasova, H.: Spatial and temporal dynamics of coupled groundwater and nitrogen fluxes through a streambed in an agricultural watershed: Groundwater and nitrogen fluxes in a streambed, Water Resources Research, 45(9), doi:10.1029/2008WR007397, 2009.

Knoll, L., Breuer, L. and Bach, M.: Nation-wide estimation of groundwater redox conditions and nitrate concentrations through machine learning, Environmental Research Letters, 15(6), 064004, doi:10.1088/1748-9326/ab7d5c, 2020.

Formatted: Normal

Kuhn, M.: Building Predictive Models in *R* Using the caret Package, Journal of Statistical Software, 28(5), doi:10.18637/jss.v028.i05, 2008.

Liao, L., Green, C. T., Bekins, B. A. and Böhlke, J. K.: Factors controlling nitrate fluxes in groundwater in agricultural areas: Factors controlling nitrate fluxes in groundwater, Water Resources Research, 48(6), doi:10.1029/2011WR011008, 2012.

Luckey, R. R. and Cannia, J. C.: Groundwater Flow Model of the Western Model Unit of the Nebraska Cooperative Hydrology Study (COHYST) Area, Nebraska Department of Natural Resources, Lincoln, NE. [online] Available from: ftp://ftp.dnr.nebraska.gov/Pub/cohystftp/cohyst/model_reports/WMU_Documentation_060519.pdf, 2006.

Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N. and Lee, S.-I.: From local explanations to global understanding with explainable AI for trees, Nature Machine Intelligence, 2(1), 56–67, doi:10.1038/s42256-019-0138-9, 2020.

Formatted: Line spacing: single

Formatted: English (United States)

Formatted: Normal

McMahon, P. B., Dennehy, K. F., Bruce, B. W., Böhlke, J. K., Michel, R. L., Gurdak, J. J. and Hurlbut, D. B.: Storage and transit time of chemicals in thick unsaturated zones under rangeland and irrigated cropland, High Plains, United States: Chemical storage in thick unsaturated zone, Water Resources Research, 42(3), doi:10.1029/2005WR004417, 2006.

Meals, D. W., Dressing, S. A. and Davenport, T. E.: Lag Time in Water Quality Response to Best Management Practices: A Review, Journal of Environment Quality, 39(1), 85, doi:10.2134/jeq2009.0108, 2010.

Mittelstet, A. R., Smolen, M. D., Fox, G. A. and Adams, D. C.: Comparison of Aquifer Sustainability Under Groundwater Administrations in Oklahoma and Texas: Comparison of Aquifer Sustainability Under Groundwater Administrations in Oklahoma and Texas, JAWRA Journal of the American Water Resources Association, 47(2), 424–431, doi:10.1111/j.1752-1688.2011.00524.x, 2011.

Morgenstern, U., Daughney, C. J., Leonard, G., Gordon, D., Donath, F. M. and Reeves, R.: Using groundwater age and hydrochemistry to understand sources and dynamics of nutrient contamination through the catchment into Lake Rotorua, New Zealand, Hydrology and Earth System Sciences, 19(2), 803–822, doi:10.5194/hess-19-803-2015, 2015.

Nash, J. E. and Sutcliffe, J. V.: River flow forecasting through conceptual models part I — A discussion of principles, Journal of Hydrology, 10(3), 282–290, doi:10.1016/0022-1694(70)90255-6, 1970.

NASS: USDA/NASS QuickStats Ad-hoc Query Tool, [online] Available from: https://quickstats.nass.usda.gov/ (Accessed 15 February 2018), 2018.

~~NEDNR: 7.5 Digital Elevation Models, Elevation Data [online] Available from: https://dnr.nebraska.gov/data/elevation-data, 1997.~~

NEDNR: Fifty-fifth biennial report of the Department of Natural Resources, Nebraska Department of Natural Resources, Lincoln, NE. [online] Available from: https://dnr.nebraska.gov/sites/dnr.nebraska.gov/files/doc/surface-water/biennial-reports/BiennialReport2005-06.pdf, 2009.

Nelson, N. G., Muñoz-Carpena, R., Phlips, E. J., Kaplan, D., Sucsy, P. and Hendrickson, J.: Revealing Biotic and Abiotic Controls of Harmful Algal Blooms in a Shallow Subtropical Lake through Statistical Machine Learning, Environmental Science & Technology, 52(6), 3527–3535, doi:10.1021/acs.est.7b05884, 2018.

NOAA: National Climatic Data Center (NCDC), [online] Available from: https://www.ncdc.noaa.gov/cdo-web/datatools (Accessed 4 August 2017), 2017.

Nolan, B. T., Green, C. T., Juckem, P. F., Liao, L. and Reddy, J. E.: Metamodeling and mapping of nitrate flux in the unsaturated zone and groundwater, Wisconsin, USA, Journal of Hydrology, 559, 428–441, doi:10.1016/j.jhydrol.2018.02.029, 2018.

Nolan, B. T., Gronberg, J. M., Faunt, C. C., Eberts, S. M. and Belitz, K.: Modeling Nitrate at Domestic and Public-Supply Well Depths in the Central Valley, California, Environmental Science & Technology, 48(10), 5643–5651, doi:10.1021/es405452q, 2014.

~~Nolan, B. T., Green, C. T., Juckem, P. F., Liao, L. and Reddy, J. E.: Metamodeling and mapping of nitrate flux in the unsaturated zone and groundwater, Wisconsin, USA, Journal of Hydrology, 559, 428–441, doi:10.1016/j.jhydrol.2018.02.029, 2018.~~

NRCS: Web Soil Survey. [online] Available from: https://websoilsurvey.sc.egov.usda.gov/ (Accessed 16 November 2017), 2018.

Ouedraogo, I., Defourny, P. and Vanclooster, M.: Validating a continental-scale groundwater diffuse pollution model using regional datasets, Environmental Science and Pollution Research, doi:10.1007/s11356-017-0899-9, 2017.

Preston, T. (NPNRD): Personal Communication with M.J. Wells, University of Nebraska, Lincoln, NE, USA, 2017.

Puckett, L. J., Tesoriero, A. J. and Dubrovsky, N. M.: Nitrogen Contamination of Surficial Aquifers—A Growing Legacy †, Environmental Science & Technology, 45(3), 839–844, doi:10.1021/es1038358, 2011.

R Core Team: R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria. [online] Available from: https://www.R-project.org/, 2017.

Rahmati, O., Choubin, B., Fathabadi, A., Coulon, F., Soltani, E., Shahabi, H., Mollaefar, E., Tiefenbacher, J., Cipullo, S., Ahmad, B. B. and Tien Bui, D.: Predicting uncertainty of machine learning models for modelling nitrate pollution of groundwater using quantile regression and UNEEC methods, Science of The Total Environment, 688, 855–866, doi:10.1016/j.scitotenv.2019.06.320, 2019.

Ransom, K. M., Nolan, B. T., A. Traum, J., Faunt, C. C., Bell, A. M., Gronberg, J. A. M., Wheeler, D. C., Z. Rosecrans, C., Jurgens, B., Schwarz, G. E., Belitz, K., M. Eberts, S., Kourakos, G. and Harter, T.: A hybrid machine learning model to predict and visualize nitrate concentration throughout the Central Valley aquifer, California, USA, Science of The Total Environment, 601–602, 1160–1172, doi:10.1016/j.scitotenv.2017.05.192, 2017.

Rodriguez-Galiano, V. F., Mendes, M. P., Garcia-Soldado, M. J., Chica-Olmo, M. and Ribeiro, L.: Predictive modeling of groundwater nitrate pollution using Random Forest and multisource variables related to intrinsic and specific vulnerability: A case study in an agricultural setting (Southern Spain), Science of The Total Environment, 476–477, 189–206, doi:10.1016/j.scitotenv.2014.01.001, 2014.

Rossman, N. R., Zlotnik, V. A., Rowe, C. M. and Szilagyi, J.: Vadose zone lag time and potential 21st century climate change effects on spatially distributed groundwater recharge in the semi-arid Nebraska Sand Hills, Journal of Hydrology, 519, 656–669, doi:10.1016/j.jhydrol.2014.07.057, 2014.

Russoniello, C. J., Konikow, L. F., Kroeger, K. D., Fernandez, C., Andres, A. S. and Michael, H. A.: Hydrogeologic controls on groundwater discharge and nitrogen loads in a coastal watershed, Journal of Hydrology, 538, 783–793, doi:10.1016/j.jhydrol.2016.05.013, 2016.

Saia, S. M., Nelson, N., Huseth, A. S., Grieger, K. and Reich, B. J.: Transitioning Machine Learning from Theory to Practice in Natural Resources Management, Ecological Modelling, 435, 109257, doi:10.1016/j.ecolmodel.2020.109257, 2020.

Spalding, R. F., Watts, D. G., Schepers, J. S., Burbach, M. E., Exner, M. E., Poreda, R. J. and Martin, G. E.: Controlling Nitrate Leaching in Irrigated Agriculture, Journal of Environment Quality, 30(4), 1184, doi:10.2134/jeq2001.3041184x, 2001.

Turkeltaub, T., Kurtzman, D. and Dahan, O.: Real-time monitoring of nitrate transport in the deep vadose zone under a crop field – implications for groundwater protection, Hydrology and Earth System Sciences, 20(8), 3099–3108, doi:10.5194/hess-20-3099-2016, 2016.

University of Nebraska-Lincoln (UNL): Quality-Assessed Agrichemical Contaminant Database for Nebraska Ground Water, [online] Available from: https://clearinghouse.nebraska.gov/Clearinghouse.aspx (Accessed 5 September 2016), 2016.

USBR: Hydromet: Archive Data Access, [online] Available from: https://www.usbr.gov/gp/hydromet/hydromet_arcread.html (Accessed 22 May 2018), 2018.

USDA: NAIP and NAPP Imagery, [online] Available from https://dnr.nebraska.gov/data/digital-imagery (Accessed 14 August 2017), 2017

U.S. Geological Survey [USGS]: National Elevation Dataset (NED), [online] Available from: https://datagateway.nrcs.usda.gov/ (Accessed 08 October 2020), 1997.

U.S. Geological Survey [USGS]: LANDSAT Imagery, [online] Available from: https://earthexplorer.usgs.gov/ (Accessed 14 August 2017), 2017.

U.S. Geological Survey [USGS]~~USGS~~: NHDPlus High Resolution, [online] Available from: https://nhd.usgs.gov/NHDPlus_HR.html (Accessed 29 June 2018), 2012.

Van Meter, K. J. and Basu, N. B.: Catchment Legacies and Time Lags: A Parsimonious Watershed Model to Predict the Effects of Legacy Storage on Nitrogen Export, edited by Y. Hong, PLoS ONE, 10(5), e0125971, doi:10.1371/journal.pone.0125971, 2015.

Van Meter, K. J. and Basu, N. B.: Time lags in watershed-scale nutrient transport: an exploration of dominant controls, Environmental Research Letters, 12(8), 084017, doi:10.1088/1748-9326/aa7bf4, 2017.

Vanclooster, M., Petit, S., Bogaert, P. and Lietar, A.: Modelling Nitrate Pollution Vulnerability in the Brussel's Capital Region (Belgium) Using Data-Driven Modelling Approaches, Journal of Water Resource and Protection, 12(05), 416–430, doi:10.4236/jwarp.2020.125025, 2020.

Formatted: Normal

Formatted: Normal

Formatted: Normal

Verstraeten, I. M., Sibray, S. S., Cannia, J. C. and Tanner, D. Q.: Reconnaissance of ground-water quality in the North Platte Natural Resources District, western Nebraska, June-July 1991, Report, United States Geological Survey. [online] Available from: http://pubs.er.usgs.gov/publication/wri944057, 1995.

Verstraeten, I. M., Steele, G. V., Cannia, J. C., Böhlke, J. K., Kraemer, T. E., Hitch, D. E., Wilson, K. E. and Carnes, A. E.: Selected field and analytical methods and analytical results in the Dutch Flats area, western Nebraska, 1995-99, Report, United States Geological Survey, Reston, VA. [online] Available from: http://pubs.er.usgs.gov/publication/ofr00413, 2001~~b~~a.

Verstraeten, I. M., Steele, G. V., Cannia, J. C., Hitch, D. E., Scripter, K. G., Böhlke, J. K., Kraemer, T. F. and Stanton, J. S.: Interaction of surface water and ground water in the Dutch Flats area, western Nebraska, 1995-99, Report, United States Geological Survey. [online] Available from: http://pubs.er.usgs.gov/publication/wri014070, ~~2001a~~2001b.

~~Verstraeten, I. M., Steele, G. V., Cannia, J. C., Böhlke, J. K., Kraemer, T. E., Hitch, D. E., Wilson, K. E. and Carnes, A. E.: Selected field and analytical methods and analytical results in the Dutch Flats area, western Nebraska, 1995-99, Report, United States Geological Survey, Reston, VA. [online] Available from: http://pubs.er.usgs.gov/publication/ofr00413, 2001b.~~

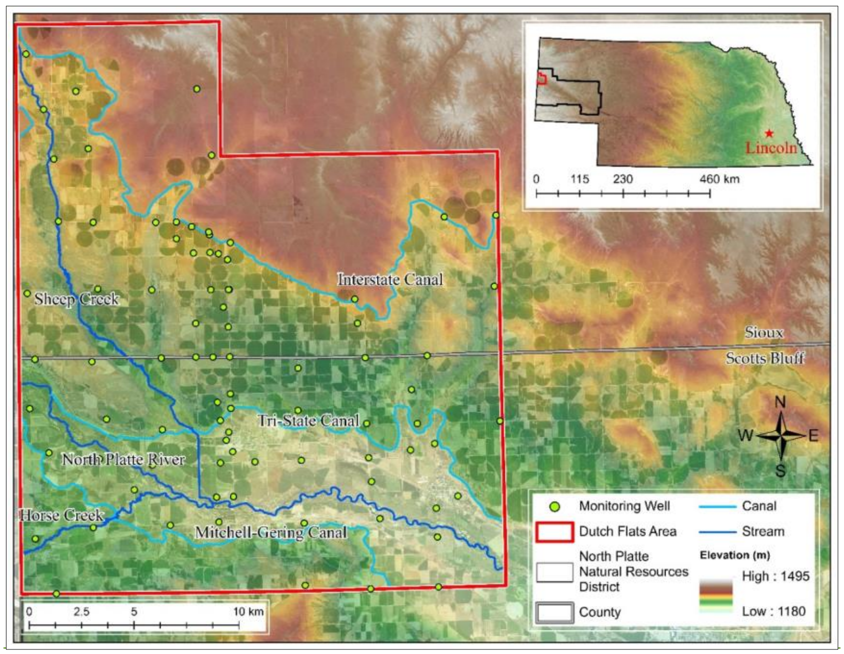
Welch, H. L., Green, C. T. and Coupe, R. H.: The fate and transport of nitrate in shallow groundwater in northwestern Mississippi, USA, Hydrogeology Journal, 19(6), 1239–1252, doi:10.1007/s10040-011-0748-8, 2011.

Wells, M., Gilmore, T., Mittelstet, A., Snow, D. and Sibray, S.: Assessing Decadal Trends of a Nitrate-Contaminated Shallow Aquifer in Western Nebraska Using Groundwater Isotopes, Age-Dating, and Monitoring, Water, 10(8), 1047, doi:10.3390/w10081047, 2018.

Wheeler, D. C., Nolan, B. T., Flory, A. R., DellaValle, C. T. and Ward, M. H.: Modeling groundwater nitrate concentrations in private wells in Iowa, Science of The Total Environment, 536, 481–488, doi:10.1016/j.scitotenv.2015.07.080, 2015.

Yonts, D.: G02-1465 Crop Water Use in Western Nebraska, University of Nebraska-Lincoln Extension [online] Available from: https://digitalcommons.unl.edu/extensionhist, 2002.

Young, L.A. (UNL): Personal Communication with M.J. Wells, University of Nebraska, Lincoln, NE, USA, 2016.


Lincoln
0 115 230 460 km
Interstate Canal
Sheep Creek
Sioux
Scotts Bluff
Tri-State Canal
North Platte River
Horse Creek
Mitchell-Gering Canal
N
W
E
S
Monitoring Well
Canal
Dutch Flats Area
Stream
North Platte Natural Resources District
Elevation (m)
High : 1495
Low : 1180
County
0 2.5 5 10 km

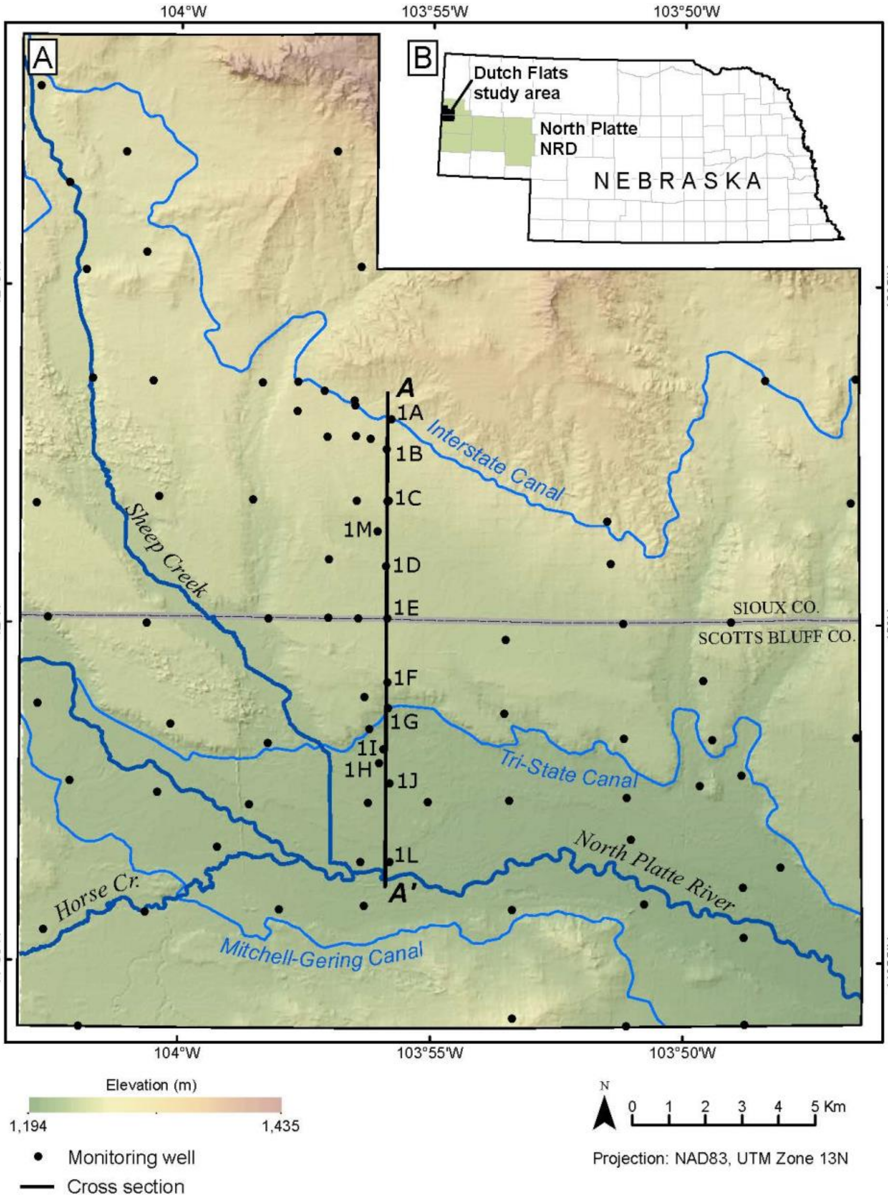


Figure 1: Dutch Flats study area (A) overlain by 30 m Digital Elevation Model (~~NeDNR~~ USGS, 1997). The study area is located within the North Platte Natural Resources District of western Nebraska (B). Depending on data availability, multiple wells (well nest) or a single well may be found at each monitoring well location. Transect A-A' represents the location and wells displayed in the Fig. 2 hydrogeologic cross-section.

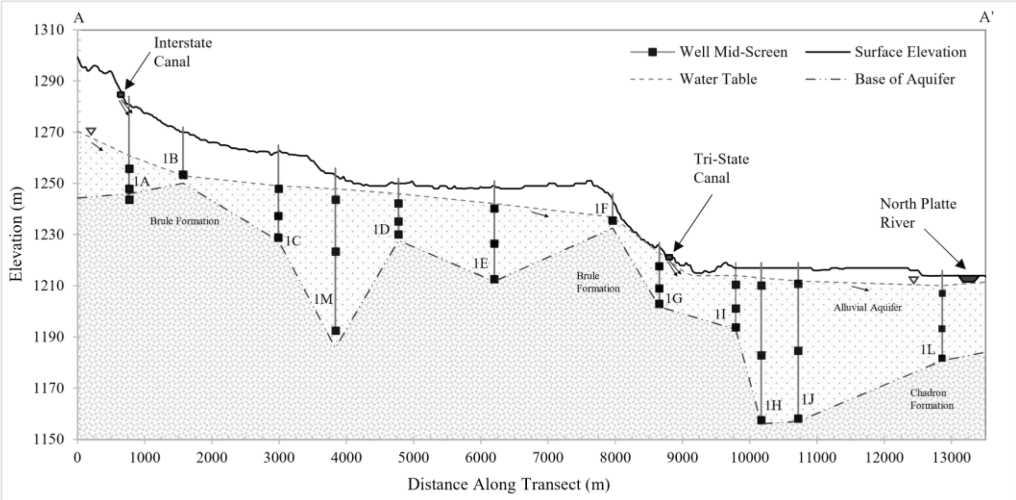


Figure 2: Cross-section along representative well transect (see Fig. 1) within the Dutch Flats area. Surface elevation data were derived from a 30-meter ~~surface~~ Digital Elevation Model ~~(DEM) raster~~ (USGS, 1997). Water ~~s~~Surface and ~~b~~Base of ~~a~~Aquifer ~~e~~Elevations were sourced from a 1998 Dutch Flats ~~s~~Study (Böhlke et al., 2007, Verstraeten et al., 2001a, 2001b). Small black arrows beneath the surface indicate general groundwater flow direction.

Formatted: Centered

Formatted: Font: Bold

Formatted: Font: Bold

Formatted: Font: Bold

Formatted: Font: Bold

Formatted: Font: Bold

Formatted: Font: Bold

Formatted: Normal

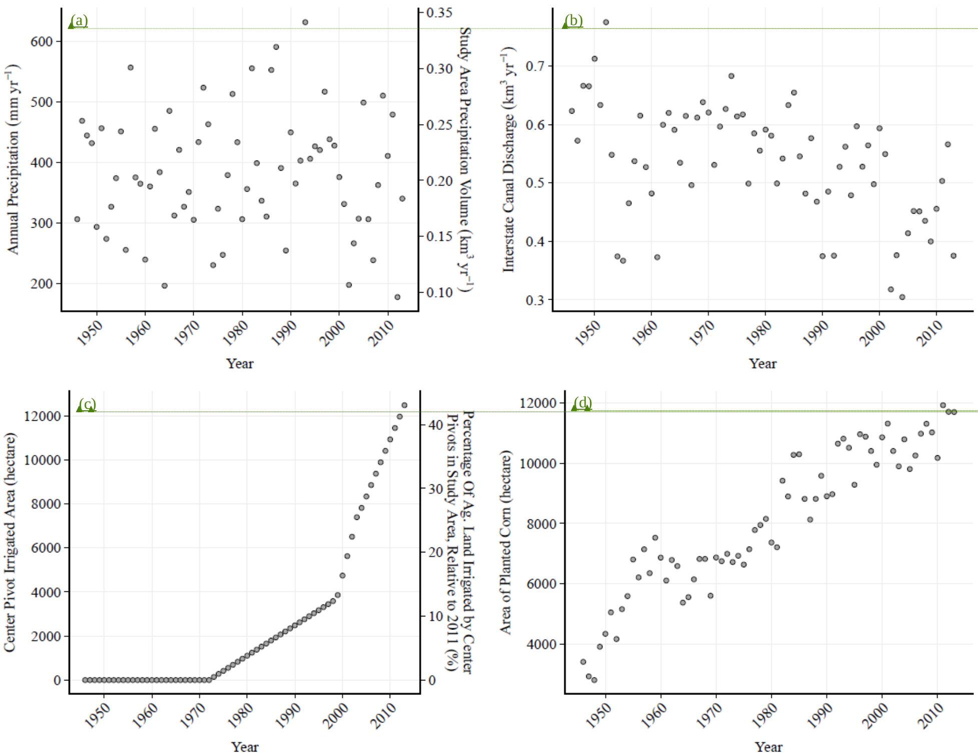


Figure 3~~2~~: Time series plots of all four dynamic predictors. Figures represent (a) annual precipitation, (b) Interstate canal discharge, (c) center pivot ~~irrigation~~ irrigated area, and (d) area of planted corn from 1946 to 2013.

~~**Figure 2: Time series plots of all four dynamic predictors. Starting in the upper left and moving clockwise, figures represent annual precipitation, canal discharge, center pivot irrigation and area of plant corn from 1946 to 2013.**~~

Formatted: Font: 8 pt

Formatted: Font: 8 pt

Formatted: Font: 8 pt

Formatted: Font: 8 pt

Formatted: Font: 8 pt

Formatted: Font: 8 pt

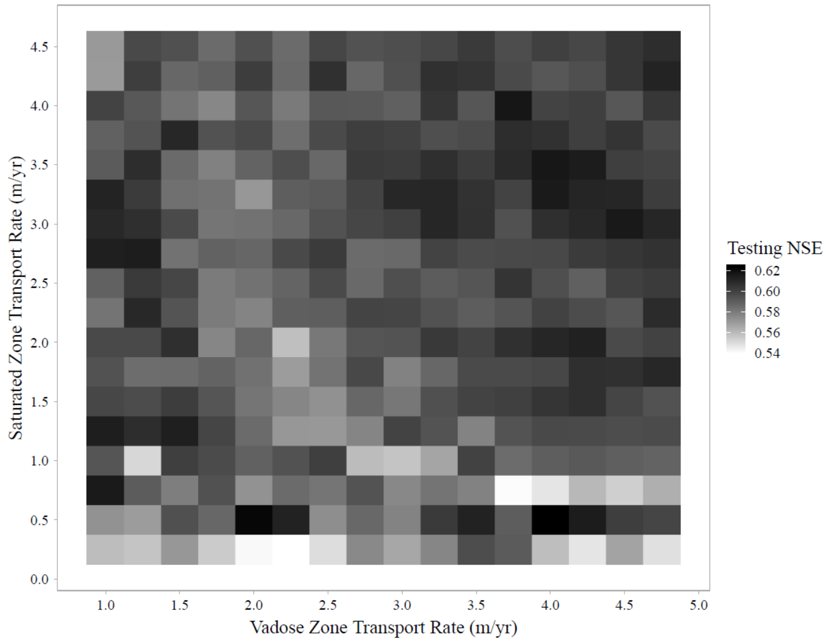
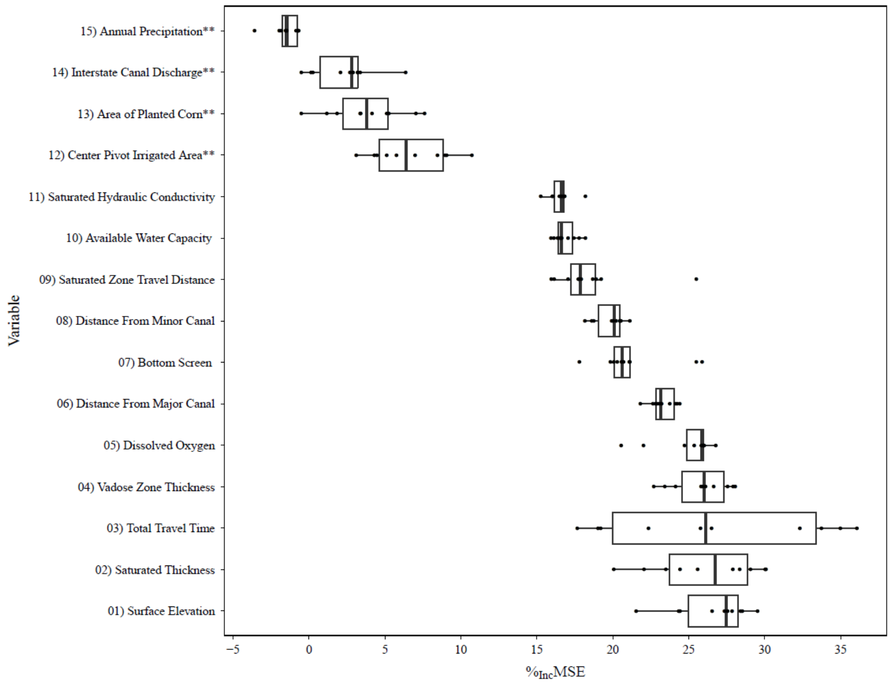


Figure ~~3~~4: Heat map of testing NSE results from 288 vadose and saturated-zone transport rate combinations. Testing NSE in this figure is the median of all 25 model outputs from each of the 288 transport rate combinations. No clear pattern of optimal vadose and saturated-zone transport rate combinations was observed.


15) Annual Precipitation**
14) Interstate Canal Discharge**
13) Area of Planted Corn**
12) Center Pivot Irrigated Area**
11) Saturated Hydraulic Conductivity
10) Available Water Capacity
09) Saturated Zone Travel Distance
08) Distance From Minor Canal
07) Bottom Screen
06) Distance From Major Canal
05) Dissolved Oxygen
04) Vadose Zone Thickness
03) Total Travel Time
02) Saturated Thickness
01) Surface Elevation
Variable
−5
0
5
10
15
20
25
30
35
$\%_{Inc}$MSE

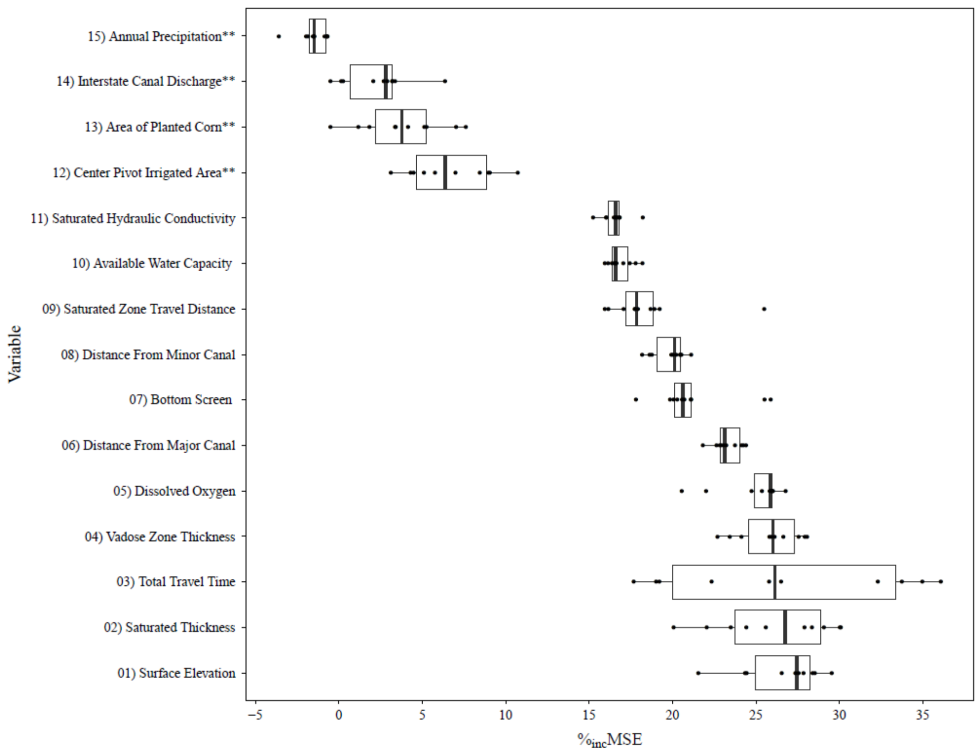
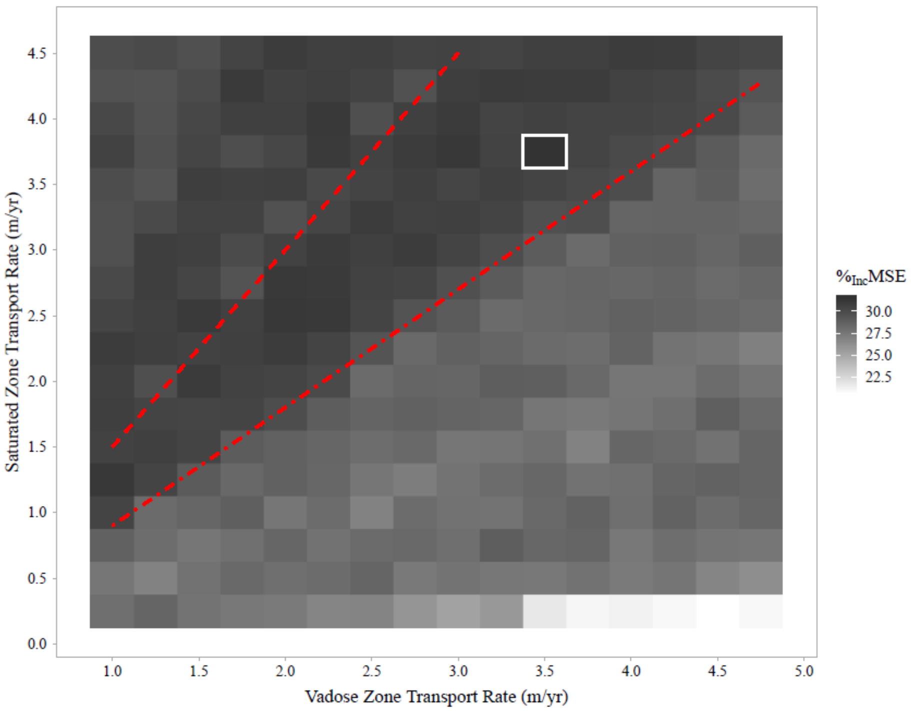


Figure ~~4~~5: Boxplot of the $\%_{inc}$MSE from the ten transport rate combinations shown in Table 2. Each boxplot has ten points for each transport rate combination, representing the median $\%_{inc}$MSE from the 25 models (five-fold cross validation, repeated 5 times). A larger $\%_{inc}$MSE suggests the variable had a greater influence on a model's ability to predict [NO_3^-]. **Denotes dynamic predictors.



Formatted: Caption, Left, Indent: First line: 0", Don't keep with next

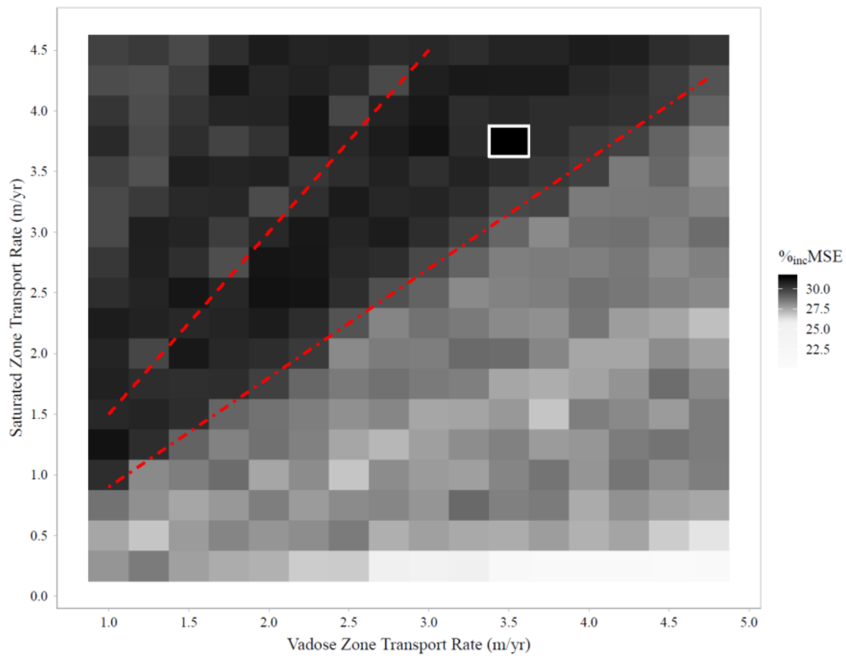
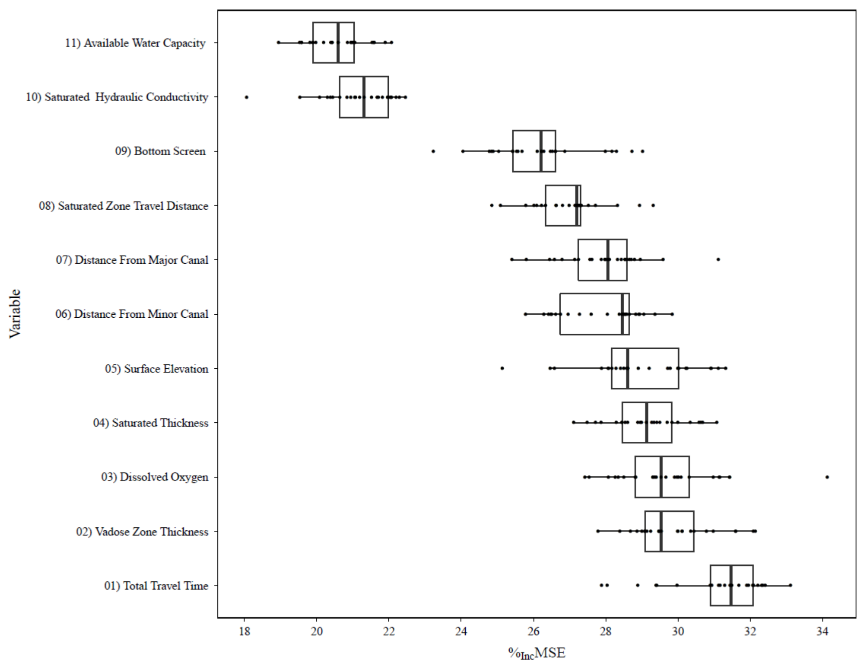


Figure 56: Heat map of $\%_{inc}$MSE (median from 25 models) from variable importance of total travel time for each of the 288 transport rate combinations evaluated. Red dashed lines indicate upper (V_u / V_u = 1.5, long dashes) and lower (0.9, short dashes) bounds of the band of transport rate combinations with consistently higher $\%_{inc}$MSE. The white square highlights the single transport rate combination with the highest $\%_{inc}$MSE.



Formatted: Indent: First line: 0"

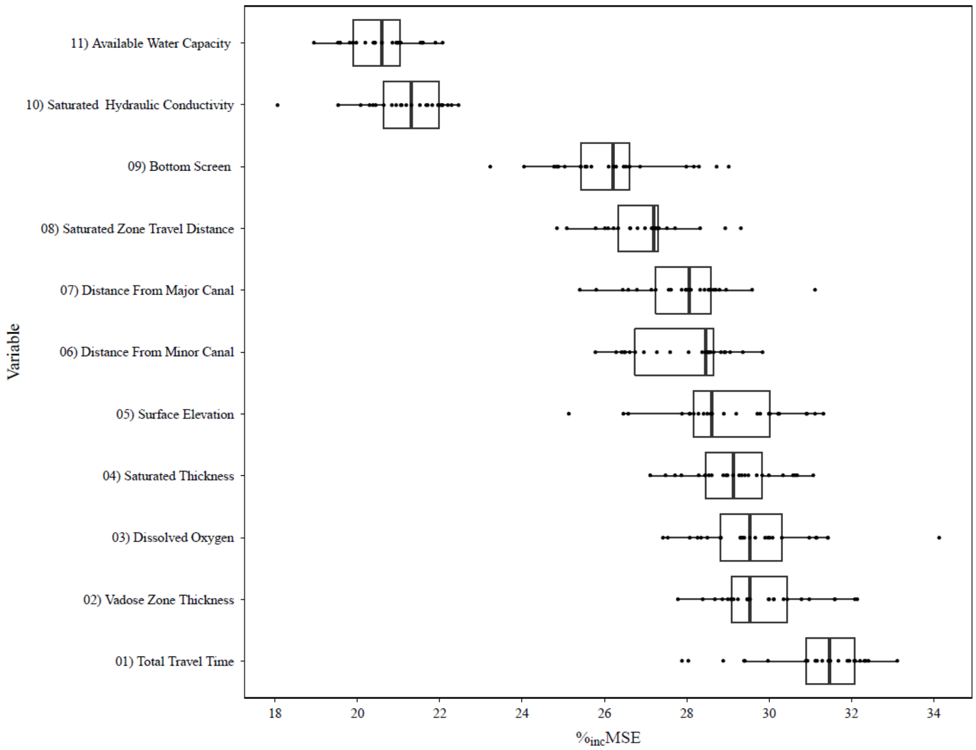


Figure ~~6~~7: Plot from secondary analysis exploring variable importance of the transport rate combination with the largest median %$_{inc}$MSE in total travel time (V_u = 3.5 m/yr; V_s= 3.75 m/yr). Each point is from one of 25 Random Forest models run for this evaluation. A larger %~~$_{inc}$MSE~~ $_{inc}$MSE suggests the variable had a greater influence on a model's ability to predict [NO_3^-].

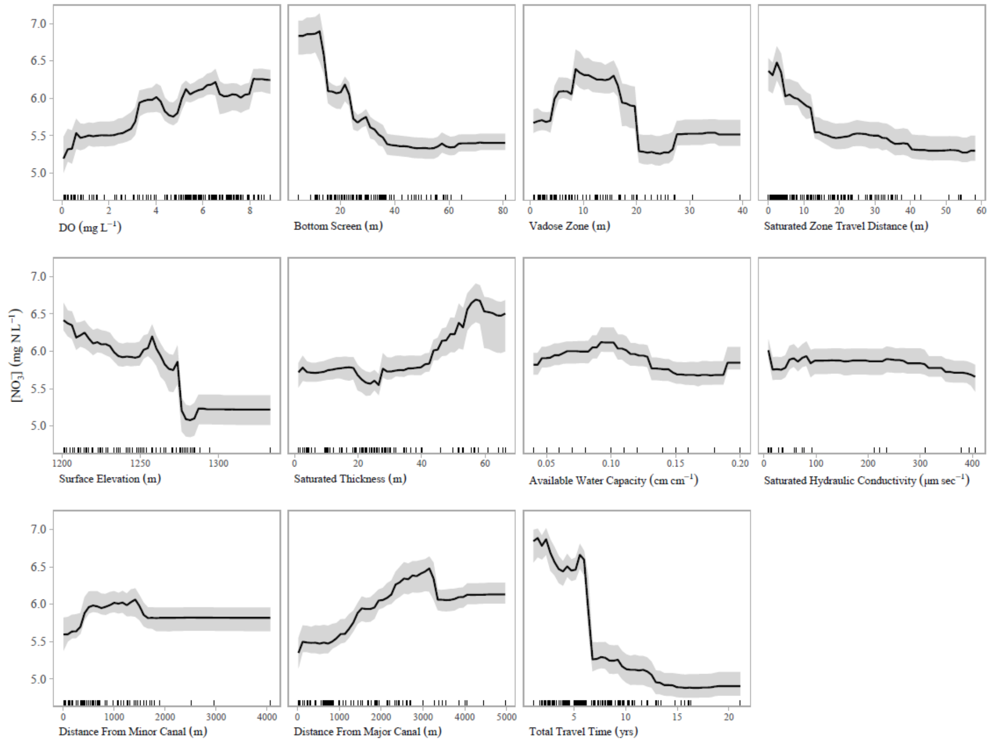


Figure ~~7~~8: Partial dependence plot for model evaluating transport rate combination of $V_u = 3.5$ m/yr and $V_s = 3.75$ m/yr. Tick marks on each plot represent predictor observations used to train models.

Table 1. List of the 15 predictors used for Random Forest evaluation. Average (avg.) and median (med.) values are shown.

Predictor	Units	Predictor Type	Source
Center Pivot Irrigated Area (avg. = 2618; med. = 1037)[a]	hectare	Dynamic	NAIP; NAPP; Landsat-1,5, 7, 8[b]
Interstate Canal Discharge (avg. = 0.53; med. = 0.55)[a]	$km^3\ yr^{-1}$	Dynamic	USBR (2018)
Area of Planted Corn (avg. = 8065; med. = 7869)[a]	hectare	Dynamic	NASS (2018)
Precipitation (avg. = 384; med. = 377)[a]	$mm\ yr^{-1}$	Dynamic	NOAA (2017)
Available Water Capacity (avg. = 0.1; med. = 0.1)	$cm\ cm^{-1}$	Static	NRCS (2018)
Dissolved Oxygen (avg. = 4.6; med. = 5.4)	$mg\ L^{-1}$	Static	C. Hudson, Personal Communication (2018)
Distance from a Major Canal (avg. = 1462.2; med. = 1161.4)	m	Static	USGS (2012)[b]
Distance from a Minor Canal (avg. = 633.2; med. = 397.6)	m	Static	USGS (2012)[b]
Bottom Screen (avg. = 26.9; med. = 24.4)	m	Static	~~NEDNR~~ UNL (2016)[b]
Saturated Hydraulic Conductivity (avg. = 68; med. = 28)	$\mu m\ sec^{-1}$	Static	NRCS (2018)
Saturated Thickness (avg. = 30.2; med. = 27.6)	m	Static	T. Preston, Personal Communication (2017)[b]
Saturated-Zone Travel Distance (avg. = 13.3; med. = 7)	m	Static	~~NEDNR~~ UNL (2016)[b]
Surface Elevation (DEM) (avg. = 1244; med. = 1248)	m	Static	~~NEDNR~~ USGS (1997)
Total Travel Time (avg. = 6.4; med. = 5.7)[c]	years	Static	~~NEDNR~~ UNL (2016)[b]
Vadose-Zone Thickness (avg. = 9.9; med. = 7.3)	m	Static	T. Preston, Personal Communication (2017); A. Young, Personal Communication (2016)

[a] Average and median span from 1946 to 2013

[b] Data required further analysis to yield calculated values; data sources are USDA (2017) and USGS (2017)

[c] Average and Median reflects transport rates of V_u = 3.5 m/yr and V_u = 3.75 m/yr

Table 2. Summary of ten vadose and saturated-zone transport rate combinations selected from 288 unique potential combinations from the analysis including dynamic variables.

	Vadose-zone Transport Rate (m/yr)	Sat. Zone Transport Rate (m/yr)	Test NSE	$[NO_3^-]$ Observations[a]	Total Travel Time (yrs) Mean (±1σ)	Total Travel Time (yrs) Median
Five Top-Performing Transport Rates	4.00	0.50	0.623	878	19.9 (± 15.8)	11.3
	2.00	0.50	0.622	861	21.6 (± 15.0)	16.5
	3.75	4.00	0.617	1049	6 (± 3.7)	5.4
	4.00	3.50	0.617	1049	6.3 (±4.1)	5.7
	4.50	3.00	0.616	1049	6.7 (± 4.7)	5.7
Extreme and Midrange Transport Combinations	4.75	4.50	0.608	1049	5.1 (± 3.2)	4.6
	2.75	2.25	0.599	1049	9.6 (± 6.3)	8.5
	1.00	4.50	0.570	1049	12.6 (± 7.7)	10.8
	1.00	0.25	0.559	607	26.7 (± 13.3)	20.6
	4.75	0.25	0.548	664	21.3 (± 15.0)	14.9

[a] In cases with slow transport rates, lag times were relatively long and not all ~~historical~~ $[NO_3^-]$ data could be used in the model. For example, a slow transport rate combination resulting in a lag time with the infiltration year prior to 1946 could not be included. Thus, some models were ultimately based on <1,049 observations.

Formatted: Font: Not Bold

Formatted: Line spacing: single

Supplemental Information – Random Forest Variable Analysis

A brief summary of how each variable was evaluated is provided below.

Note: Dynamic predictors were downloaded and analyzed from 1946 (first year Interstate Canal discharge was available) to 2013 (last year of available nitrate data (Fig. S1)).

Center Pivot Irrigated Area:

Aerial imagery was used to digitize center pivot irrigated fields (Fig. S2). When available, the analysis utilized NAIP and NAPP imagery, with LANDSAT, though at a lower resolution, providing additional imagery when the former two were unavailable (Table S1). Annual center pivot irrigated area was plotted (Fig. S3), with 1999 and 2003 identified as breaks (i.e., observed notable shifts in irrigation practice) in the data, and used to interpolate irrigated area before, between, and after the breaks. That is, estimate area for years before the shift (<1999), during the major shift (1999 – 2003), and years following the major shift (>2003). Linear regression was used to estimate center pivot irrigated area for years that were not digitized, providing a dynamic dataset for the study. Before the shift, center pivot fields were installed at approximately 138 hectares per year. During the largest conversion from furrow to center pivot, fields increased from 1999 – 2003 at a rate of 883 hectares per year, followed by 518 hectares per year thereafter. In 1999, area under center pivot irrigation was approximately 3,830 hectares. As of 2017, irrigated hectares increased by nearly 270%, to 14,253 hectares. Center pivot irrigated area was set to zero from 1946 to 1972.

Center pivot and furrow irrigated fields are believed to follow an inverse trend. In other words, as center pivot irrigated fields increase over time, fields irrigated by less efficient furrow methods are assumed to decrease. With this, it is possible less nitrate would leach through the root zone with improved water application efficiency; however, it is also possible that decreased application rates reduce the dilution of [NO_3^-] in the aquifer.

Interstate Canal Discharge:

Annual Interstate Canal discharge was retrieved from the Bureau of Reclamation's Hydromet data archive (USBR, 2018). Data were downloaded, and records were converted from $ft^3\ sec^{-1}$ to $km^3\ year^{-1}$. Since unique annual values were available for each year over the period of record, this was a dynamic dataset. Datasets were also downloaded for both the Tri-State and Mitchell-Gering Canal, however, only an Interstate Canal discharge value was assigned to each nitrate observation. This approach was justified because it is unknown which canal influences a well, regardless of the distance from a canal. Further, annual canal discharge from the Tri-State and Mitchell-Gering Canals were also compared to the Interstate Canal, in which it was determined canals follow similar annual trends, and only a relative influence was needed for the purpose of this analysis. Each nitrate observation was assigned a lagged canal discharge value depending on the transport rate combination and total travel time.

Due to the known influence of canal leakage on groundwater [NO_3^-], annual volume of water diverted into the Interstate Canal was used as a dynamic predictor to investigate if the model could identify the impact of annually high or low stream diversions on groundwater [NO_3^-].

Area of Planted Corn:

Because the study area includes parts of two counties (i.e., Scotts Bluff and Sioux), planted corn area was analyzed in three steps. First, the total annual planted area was downloaded from the USDA National Statistics Service (NASS) for both Scotts Bluff and Sioux Counties from 1946 to 2015 (NASS, 2018). Second, the USDA-NASS georeferenced Cropland Data Layer was downloaded through Geospatial Data Gateway (https://gdg.sc.egov.usda.gov), and projected into ArcMap 10.4. From 2002 to 2015, estimated corn area was determined for each year in the two respective counties. The area of planted corn was also determined for each respective county within the boundary of the Dutch Flats area. Next, a ratio for every year was estimated by comparing the area of planted corn for each county within Dutch Flats, to the planted corn area for the entirety of each county. Ratios were averaged from 2002 to 2015, with the Dutch Flats county-level planted corn area having ratios of 0.18 (± 0.01) and 0.77 (± 0.03) compared to total planted corn in Scotts Bluff and Sioux Counties, respectively. Third, estimated annual planted corn area within Dutch Flats from 1946 to 2015 was determined by multiplying each year of total county-level planted corn by the respective ratio and summing up areas.

This variable was used as a dynamic predictor, and was included as a proxy for the limited amount of long-term fertilizer data available since corn requires high fertilizer inputs. Statistical analysis of the long-term dataset suggested there has been a significant decrease in fertilizer application in Scotts Bluff County, while planted corn area has significantly increased. However, it is also possible there are more uncertainties related to application surveys associated with fertilizer than planted corn area.

Precipitation:

Annual sums for precipitation were downloaded from the National Climatic Data Center of National Oceanic and Atmospheric Administration (NOAA) at the Scottsbluff W.B. Heilig Field Airport, NE US (NOAA, 2017). With several years of consistent data, precipitation was used as a dynamic predictor, and lagged with respect to the sample collection date. Higher years of precipitation could potentially lead to increased leaching or fertilizer runoff.

Available Water Capacity and Saturated Hydraulic Conductivity:

Available water capacity (AWC) and saturated hydraulic conductivity (K) were retrieved from Web Soil Survey, which is maintained by the U.S. Department of Agriculture (USDA) Natural Resources Conservation Service. Spatial data were downloaded from Geospatial Data Gateway (https://gdg.sc.egov.usda.gov). AWC is the amount of water that is retained by soil and available for plant uptake. Saturated hydraulic conductivity describes the movement of water through saturated soil. Data were retrieved for both Sioux and Scotts Bluff Counties, and evaluated with the *Soil Data Viewer* tool developed by NRCS. Values of each respective predictor were extracted to each well.

Available water content influences water storage and infiltration rate, and thus is a variable influencing the amount of time that irrigated water will reside within the root zone. Lower values of AWC could influence leaching, as more nitrate may leach below the root zone before plants can assimilate the nutrients. This study incorporated saturated hydraulic conductivity to evaluate how the ability for which water moves through a saturated vadose zone may impact [NO_3^-] in the Dutch Flats area. Higher K would suggest water moves more quickly through the soil, potentially transporting nitrate to groundwater at higher rates.

Bottom Screen & Saturated-Zone Travel Distance:

Bottom screen depth was already within the nitrate dataset retrieved from the Quality-Assessed Agrichemical Contaminant Database (UNL, 2016). Saturated-zone travel distance is the distance below the water table to the screen midpoint. If the screen crossed the water table, common for shallow wells, the saturated-zone travel distance was the screen midpoint between the bottom screen and water table. In deeper wells, the saturated-zone travel distance value was found as the distance from the water table to the midpoint between the top and bottom of the screened interval.

Bottom Screen depth has been found an important variable in other Random Forest studies (Wheeler et al., 2015). This predictor is a factor encompassing the movement of water through both the vadose and saturated zones, and evaluates how traveling through both these zones impacts [NO_3^-]. Saturated-zone travel distance assesses how the distance water travels from the water table to screen midpoint influences groundwater [NO_3^-].

Surface Elevation (DEM), Vadose-Zone Thickness, and Saturated Thickness:

To estimate saturated thickness, wells were only selected for the model dataset if they had both a [NO_3^-] and depth to groundwater value. Depth to groundwater records for the NPNRD (n = 49,765; 1929 – 2016) were retrieved from the University of Nebraska Conservation and Survey Division (CSD) (A. Young, Personal Communication, 2016). Depth to groundwater records from 2017 (n = 806) were also sent from the NPNRD GIS Coordinator (T. Preston, Personal Communication, 2017). Wells from the nitrate database were joined with the 2017 NPNRD monitoring well data. [NO_3^-] that did not have a 2017 depth to groundwater value were checked with the CSD database for additional depth to groundwater records. Of the 2,829 nitrate samples in the Dutch Flats area, 2,651 samples had a well matching a depth to groundwater value from the 2017 monitoring well and/or CSD depth to groundwater dataset. Under the assumption shallow wells were in an unconfined aquifer, the most recent depth to groundwater record at a well nest's shallow well were assigned to the entire well nest.

Base of aquifer contours were acquired from NPNRD (T. Preston, Personal Communication, 2017). A 30-meter base of aquifer surface was interpolated using the *Topo to Raster* tool in ArcMap 10.4. A base of aquifer and 30-meter surface Digital Elevation Model (DEM) raster value were extracted to each well point, representing the surface elevation and base of aquifer elevation at each well. Depth to groundwater at each well was used to determine the water table elevation. This in turn was used to estimate saturated thickness of the aquifer by subtracting the interpolated base of aquifer elevation from the water table elevation. One well (8C-S) returned a negative aquifer thickness value. For this specific well, the saturated thickness was estimated by taking bottom of the well and subtracting it from the water table elevation, and assuming the bottom of the well was located near, or at, the top of the confining layer. The resulting saturated thickness was consistent with estimated thicknesses in the surrounding area.

Surface elevation was used, in part, to evaluate how moving further from the North Platte River (NPR) influenced groundwater [NO_3^-]. Because the land surface generally slopes toward the river, locations of higher elevation generally would indicate wells that are further from the NPR. This predictor may also assess the impact of canals at their respective elevations. Vadose-zone thickness has been used previously to evaluate how quickly a contaminant may reach the water table, and was assessed by Rodriguez-Galiano et al. (2014). Vadose-zone thickness may also play a role in nitrate storage in the unsaturated zone.

Distance from major and minor canals:

Canal spatial data were from the USGS National Hydrography Dataset (USGS, 2012). The three largest canals were identified, based on canal discharge, in the Dutch Flats area as the Interstate, Tri-State, and Mitchell-Gering Canals. These canals were used to determine the "Distance from Major Canal" variable. Further analysis was conducted via aerial imagery to digitize lower order irrigation canals that were not in the NHD database. These canals were used to determine the "Distance from minor canal" variable. The *Near* tool in ArcMap 10.4 was used to calculate the distance each well was from the closest, irrespective of direction, major canal and minor canal.

Since Interstate Canal discharge is a temporally dependent variable, the distance from major and minor canals was used to evaluate the spatial component of canals in the region. Due to the high leakage potential of these canals, the variable was included to evaluate the influence of a well's proximity to canals in Dutch Flats on groundwater [NO_3^-].

Dissolved Oxygen (DO):

This dataset was obtained from NPNRD (C. Hudson, Personal Communication, 2018). The dataset had several DO values per well, however, there was not a perfect 1:1 match between the collection of a nitrate sample, and the collection of DO. Therefore, all DO measurements for each well were averaged. Thus, each well was assigned a single unique DO value that was associated with all the annual median [NO_3^-] values for the well. Analysis of DO and isotopes of nitrate suggest there has not been a major change in nitrate reduction within Dutch Flats from 1998 to 2016 (Wells et al., 2018).

DO largely drives biological processes impacting groundwater [NO_3^-], in which case low DO could be associated with low [NO_3^-], and high DO with higher [NO_3^-] because of NO_3^- reduction in suboxic parts of the aquifer. Alternatively, in the absence of NO_3^- reduction, DO and [NO_3^-] could be correlated independently with groundwater age through DO reduction and historical change in NO_3^- recharge.

Total Travel Time:

While this variable has a time component, it was treated as a static variable because it was held constant from one sample year to the next for each well. Total travel time was determined through several steps, also discussed in depth within Section 2.5 of the main paper, though a brief stepped summary is provided below.

(1) The mean and standard deviation for recharge (R) data from Böhlke et al., 2007 and Wells et al., 2018 was determined and used to calculate a range of vertical velocities (V), or transport rates, through both the vadose and saturated zone. Equation 1a was assigned a mobile water content value of θ = 0.13 and 0.35, depending on whether rates were being calculated for the vadose and saturated zone, respectively.

a. $V = \frac{R}{\theta}$

(2) Each [NO_3^-] had a depth component (z) in the dataset, whether that be the vadose zone depth, or the vertical distance from the water table to the well mid-screen (i.e., saturated-zone vertical travel distance). From the vertical velocity determined in equation 1a and respective zone depth, equation 2a was used to independently estimate travel time through the vadose and saturated zones.

a. $\tau = \frac{z}{V}$

(3) Lastly, based on the travel time determine for the vadose zone (τ_u) and saturated zone (τ_s), equation 3a provided a unique estimate for the total travel time at each well location.

a. $\tau_t = \tau_u + \tau_s$

The total travel time variable was used as a means to link dynamic variables to the date a NO_3^- sample was collected, or in essence, establish a lag time between surface activities when nitrate entered the system, and when it was sampled years later. Additionally, this variable, due to the known inverse relationship (e.g., as travel time increases, [NO_3^-] decreases) observed in previous Dutch Flats studies, was utilized in the secondary analysis discussed in Section 3.2.

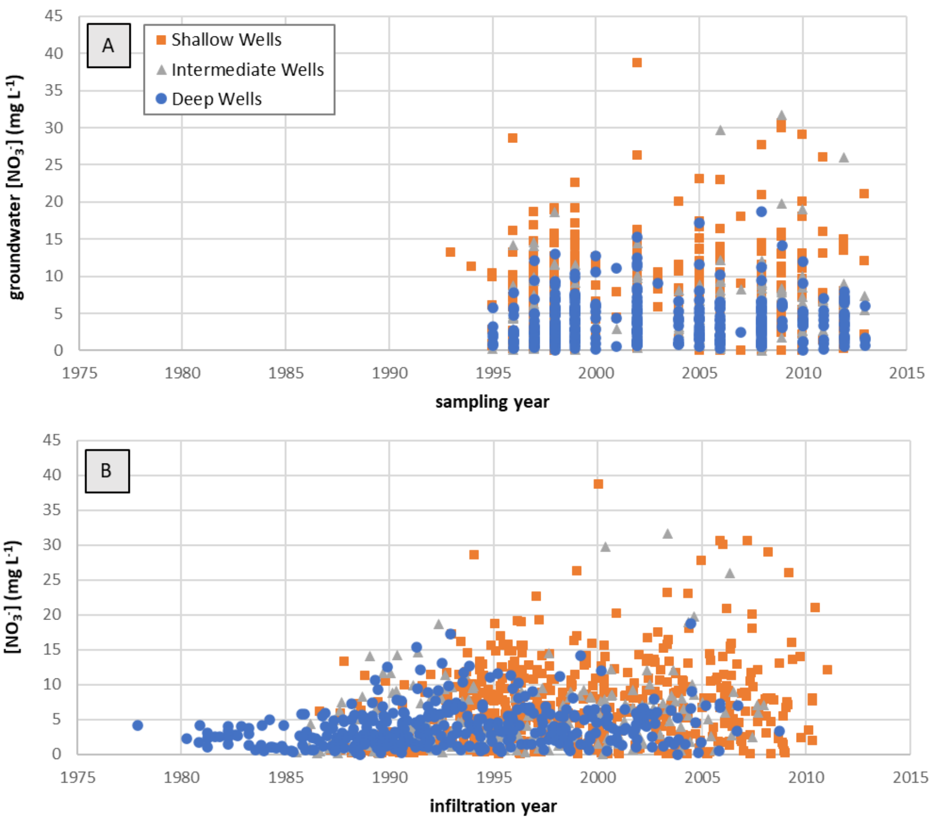


Figure S1: (A) Plot of nitrate data used in the analysis from 1993 to 2013 (n = 1,049), differentiated by well depth labels. Note that well depth is determined within individual well nests, and therefore not an absolute indicator of well depth from the land surface (see discussion in main text). (B) Plot of nitrate data adjusted for total travel time when calculated using optimal transport rates identified in the study. The adjusted nitrate data are the modelled input of nitrate to the system, starting at the land surface (infiltration). The overall pattern of input over time is similar to other studies where nitrate data were adjusted for groundwater age (e.g., Puckett et al. 2011) although inputs are more typically plotted based on recharge year rather than infiltration year.

Commented [TG1]: Figure S1 and caption were reworked.

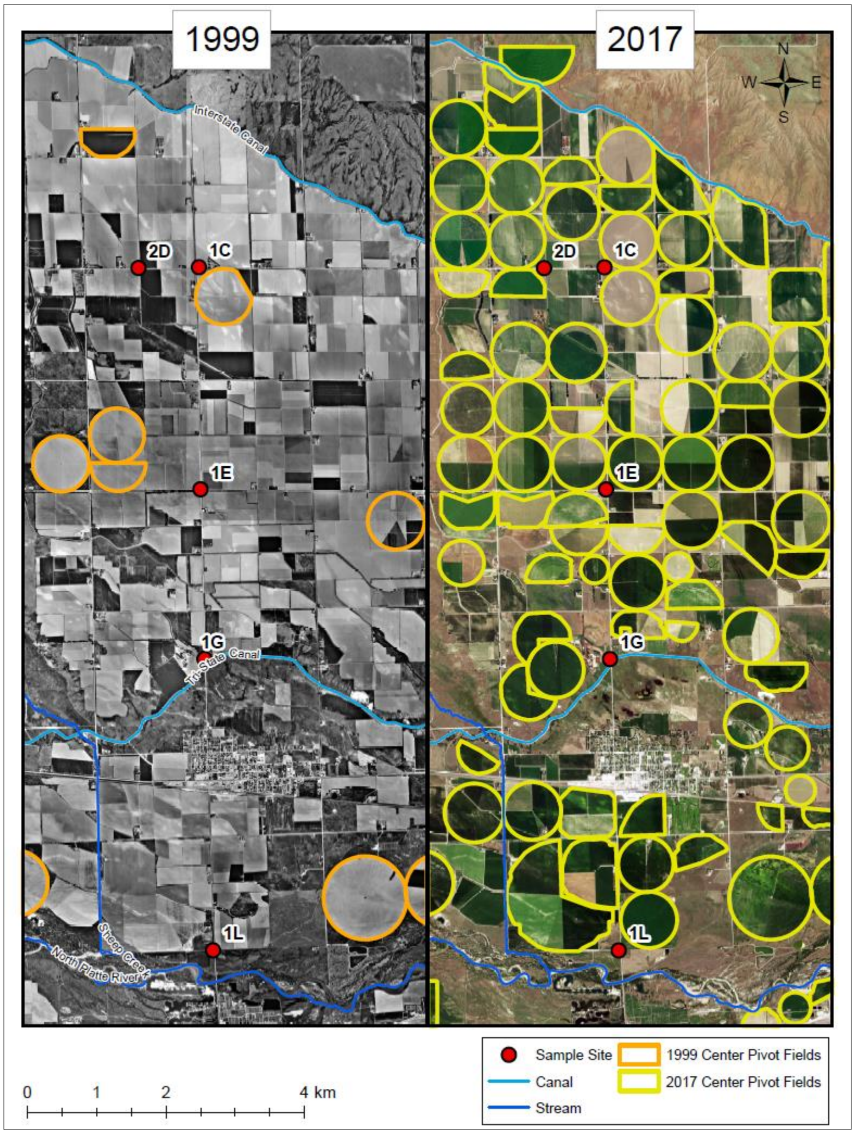


Figure S2: Visual comparison of 1999 Dutch Flats center pivot irrigated fields to 2017 center pivot irrigated fields using NAIP, NAPP, (USDA 2017) and LANDSAT imagery (USGS 2017) (Table S1). Sample sites shown in the figure are from a representative subset of wells selected for comparison in the Wells et al., 2018 study.

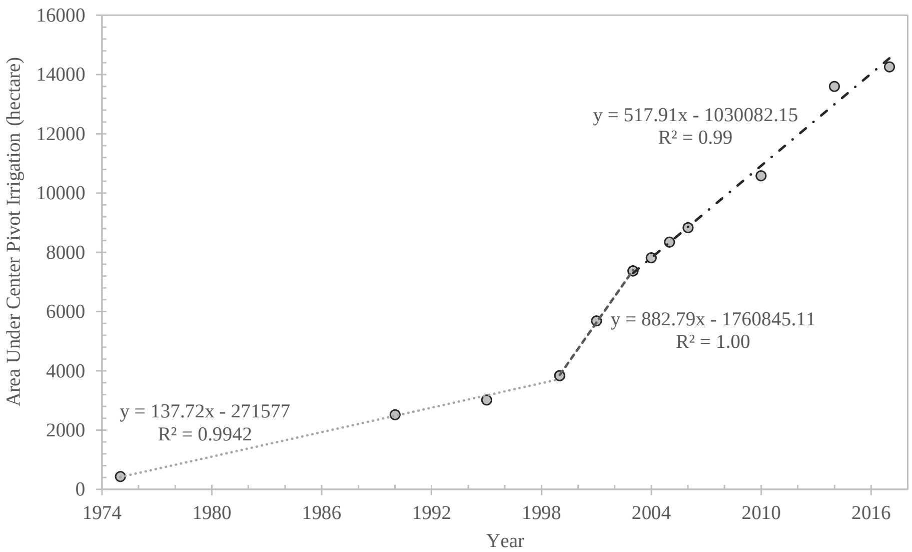


Figure S3: Center Pivot Irrigated Area based on observations from aerial imagery. The years 1999 and 2003 were used as breaks for determining linear regression equation.

Table S1: Years when aerial imagery was utilized to digitize center pivot irrigated fields.

Year	Estimated Center Pivot Irrigated Area (hectare)
1975	429
1990	2507
1995	3015
1999*	3830
2001	5685
2003*	7361
2004*	7804
2005*	8341
2006*	8822
2010*	10577
2014*	13591
2017	14253

*Years analysed with NAIP or NAPP imagery