11-22-2018

The long term effect of agricultural, vadose zone and climatic factors on nitrate contamination in Nebraska's groundwater system

Pongpun Juntakut

Daniel D. Snow

Erin M.K. Haacker

Chittaranjan Ray

Journal of Contaminant Hydrology 220 (2019) 33–48

Contents lists available at ScienceDirect

Journal of Contaminant Hydrology

journal homepage: www.elsevier.com/locate/jconhyd

The long term effect of agricultural, vadose zone and climatic factors on nitrate contamination in Nebraska's groundwater system

Pongpun Juntakut[a,b], Daniel D. Snow[c], Erin M.K. Haacker[d], Chittaranjan Ray[a,d,*]

[a] Department of Civil Engineering, University of Nebraska-Lincoln, Lincoln, NE 68588-6105, United States
[b] Department of Civil Engineering, Chulachomklao Royal Military Academy, Nakhon Nayok 26001, Thailand
[c] Water Sciences Laboratory, University of Nebraska, Lincoln, NE 68583, United States
[d] Nebraska Water Center, University of Nebraska, Lincoln, NE 68588, United States

ARTICLE INFO

Keywords:
Groundwater nitrate contamination
agricultural
vadose zone and climatic factors
CART method
Nebraska

ABSTRACT

A four-decade dataset (1974–2013) of 107,823 nitrate samples in 25,993 wells from western and eastern parts of Nebraska was used to assess long-term trends of groundwater nitrate concentration and decadal changes in the extent of groundwater nitrate-contaminated areas (NO_3-$N \geq 10$ mg N/L) over the entire state. Spatial statistics and regressions were used to investigate the relationships between groundwater nitrate concentrations and several potential natural and anthropogenic factors, including soil drainage capacities, vadose zone characteristics, crop production areas, and irrigation systems. The results of this study show that there is no statistically significant trend in groundwater nitrate concentrations in western Nebraska, in contrast with the increasing trend ($p < .05$) to the east. The spatial extent and nitrate concentrations in contaminated groundwater in center pivot-irrigated areas was less than in gravity-irrigated areas. Areas with a thicker vadose zone and larger saturated thickness of the aquifer have relatively lower nitrate concentrations. The results of a classification and regression tree (CART) model indicate the difference in the influence of physical factors on groundwater nitrate concentrations between western and eastern Nebraska, namely that groundwater nitrate concentrations correspond with vadose zone thickness, effective hydraulic conductivity, and saturated thickness in the west, while in eastern Nebraska, concentrations are correlated with average percent sand in the topsoil (0–150 cm), well depth, and effective hydraulic conductivity.

1. Introduction

Nebraska, an agriculturally intensive state in the mid-western United States (U.S.), has a large number of wells with nitrate concentrations above the drinking water standard of 10 mg NO_3-N/L (NDEQ, 2015). The High Plains/Ogallala Aquifer (HPOA) is a major alluvial aquifer that extends from South Dakota in the north to Texas in the south, and supplies tremendous amounts of water for agricultural, municipal, and industrial uses. About two thirds of the water in the HPOA is in Nebraska, which also contains a number of large rivers with dams and canal diversions. Groundwater irrigation in Nebraska has increased significantly over the past six decades with the adoption of center pivots, which replaced traditional flood irrigation methods. Now, more than 3.4 million hectares of land rely on groundwater from the HPOA to irrigate crops in Nebraska (USDA, 2014). While nitrate may occur naturally in groundwater, a major cause of high nitrate occurrence in Nebraska's wells is the extensive fertilizer application across the state (Stanton and Lynne, 2006; Gurdak and Qi, 2006; Gurdak et al., 2009; Exner et al., 2014), particularly in irrigated fields.

Consumption of water with elevated nitrate concentrations can cause health problems, primarily for infants; its effects are called “blue baby syndrome” or methemoglobinemia, which is caused by the inability of the blood to deliver enough oxygen to the infant's body, as described by Comly (1945). In 1962, the U.S. Public Health Service officially recommended a nitrate standard of 10 mg NO_3-N/L for drinking water (U.S. Public Health Service, 1962). Though other countries have tighter standards, the U.S. has retained this 10 mg NO_3-N/L to the present day (Sattelmacher, 1962; Simon et al., 1964; Kross et al., 1995; NAS, 1995; U.S. EPA, 2004, 2007, 2017; Tiemann, 2017).

Generally, large amounts of nitrogen fertilizers and irrigation are applied annually in agricultural areas of Nebraska to increase and maintain agricultural production and crop yields (Spalding, 1975; Exner and Spalding, 1976; Adelman et al., 1985; Grassini et al., 2012; Ferguson, 2015). Consequently, nitrate contamination in Nebraska's

---

[*] Corresponding author at: Department of Civil Engineering, University of Nebraska-Lincoln, Lincoln, NE 68588-6105, United States.
E-mail address: cray@nebraska.edu (C. Ray).

https://doi.org/10.1016/j.jconhyd.2018.11.007
Received 23 January 2018; Received in revised form 12 October 2018; Accepted 20 November 2018

groundwater occurs primarily from nitrate leaching in agricultural areas. Although these practices and risks are distributed widely across the state, some parts of the HPOA are more vulnerable than others, for example due to soil drainage characteristics at the land surface (Spalding and Exner, 1993; Nolan et al., 1997; Exner et al., 2010, 2014).

Rising nitrate concentrations in groundwater has prompted Nebraska's Natural Resources Districts (NRDs) to begin implementing groundwater management plans for quality and quantity in the mid-1980s (NRD, 2017). The Central Platte NRD (CPNRD) plan is representative of those adopted throughout the state of Nebraska for groundwater quality management, although each of the 23 NRDs is free to enact regulations tailored to local conditions. The CPNRD Groundwater Quality Management Program (GWQMP) was initiated in 1988, and involves a phased approach to nitrogen management (CPNRD, 2016). The CPNRD defined four classes of nitrate contamination: Phase I, II, III, and IV. These classes correspond to nitrate concentrations of < 7.5 mg/L (Phase I), 7.6–15 mg/L (Phase II), > 15 mg/L (Phase III), and areas where the rate of decline in NO_3-N concentrations have not been satisfactory (Phase IV) (CPNRD, 2016). Within these areas, the timing and application rates of nitrogen fertilizer on irrigated agriculture are regulated differently as presented in supplemental information (Table S1).

Although Nebraska NRDs have intensively monitored the extent of nitrate contamination in groundwater, as published in the Nebraska Groundwater Quality Monitoring Report (NDEQ, 2015), and many researchers and managers have made efforts to minimize the impact of irrigated crop production on the occurrence of nitrate in Nebraska's groundwater, it continues to be challenging to control nitrate contamination (Table S2). This is likely due to the complexity of the aquifer system and difficulty in measuring the effectiveness of the best management practices (BMPs) from the GWQMP in reducing areas of nitrate contamination in groundwater. Previous studies by Exner et al. (2010, 2014) have analyzed long-term groundwater nitrate concentration trends in several regions of eastern Nebraska, but their analysis did not include the more rural, semi-arid west, which is also extensively cultivated (Dappen et al., 2007). Additionally, the agricultural areas are expanding in western Nebraska (Dappen et al., 2007; Hiller et al., 2009), potentially impacting the groundwater quality due to significant N fertilizer application (Exner and Spalding, 1994). USGS reports have indicated widespread nitrate contamination in both eastern and western Nebraska (Verstraeten and Ellis, 1994; Verstraeten et al., 1994, 1998).

There is a significant difference in climate, precipitation and irrigation practices across the state. Western Nebraska is drier, with greater temperature extremes. The predominant bedrock is older (Gutentag et al., 1984), which means the landscape is less flat, soil texture is coarser than in many of the easternmost parts of the state, and hydraulic conductivity is lower in the aquifer. One purpose of this study is the analysis of the long-term groundwater nitrate concentration trends in western Nebraska, which will help increase understanding the occurrence of nitrate contamination in groundwater. Examining the commonalities and distinctions between the eastern and western parts of the state will help in planning the policy for the protection of groundwater from nitrate contamination. Furthermore, we employ a statistical classification method to establish the correlations among interrelated factors which may be important to nitrate concentrations in the east and west, which have not been considered in previous studies such as Exner et al. (2014), whose landmark study of eastern Nebraska demonstrated the increasing trends in nitrate over the course of three decades. This assessment goes beyond Nolan et al. (1997) and Nolan and Hitt (2006), who demonstrated that many states, including Nebraska, are threatened with nitrate contamination in groundwater. Several studies have attempted to identify sources of groundwater nitrate contamination based on the local analysis in various parts of the United States (Van der Schans et al., 2009; Lockhart et al., 2013; Murgulet and Tick, 2013) as a supplement to these nationwide surveys. The research presented here shows the spatial-temporal changes of nitrate contamination in groundwater on a local to regional scale for the entire state of Nebraska.

Nitrate may be influenced by many factors, some continuous (e.g., vadose zone thickness, thickness of the aquifer, depth to groundwater, etc.), and others categorical (e.g., type of well —domestic, irrigation, or monitoring; presence of barnyard within the property, older or newer wells, etc.). For the analysis of large data sets, the Classification and Regression Tree (CART) is one of the most commonly used decision tree tools. CART can be used to analyze complex interactions among predictors based on regression equations, particularly when there is a large amount of data with many variables (Zhang et al., 2003; Qi et al., 2010). For example, Burow et al. (2010) used CART to identify the relative significance of N inputs, biogeochemical processes, and physical aquifer properties in explaining nitrate concentrations in groundwater. In this study, we evaluate additional influential factors which were not considered in Burow et al. (2010), such as soil drainage classes, percent sand and organic matter in the topsoil, and weather data. In addition, we develop CART models to predict groundwater nitrate concentrations based on the presence or degree of these factors.

As mentioned above, Nebraska has clear differences in hydrogeology and spatial characterizations between western and eastern parts (e.g., rainfall amounts, soil texture, population growth, and crop varieties). Thus, studies of groundwater nitrate contamination should be considered in each region of Nebraska in order to identify the local causes of leaching and solutions to nitrate occurrence in Nebraska's groundwater.

The objectives of this study include: (i) to estimate the long-term trends of groundwater nitrate concentrations in western and eastern Nebraska; (ii) to examine four decades (1974–2013) of change in the spatial distribution of groundwater nitrate concentrations; and (iii) to evaluate relationships between high groundwater nitrate concentrations (≥ 10 mg NO_3-N/L), and potential natural (e.g., weather, and soil drainage) and anthropogenic (e.g., crop production and price, well type, and irrigation system) factors using CART.

This complements the work by Exner et al. (2014) for the years 1981–2010 in eastern Nebraska, in part by considering a time series that is 33% longer (40 years rather than 30 years) as well as including the western part of the state. This also complements the national-scale analysis of Nolan and Hitt (2006), which contextualizes the local risk of nitrate contamination in groundwater occurring in the High Plains.

Shallower wells are likely to have higher contamination levels than deeper wells. Older wells typically have higher contamination than newer wells and could be linked to construction techniques (Spalding and Exner, 1993). The higher the number of screen zones and the longer the total screen length, the higher the chance for the well to capture the groundwater from the aquifer from all depths. High organic matter tends to preserve soil structure. A higher percentage of sand and organics allows greater infiltration water to the underlying aquifer. Thick vadose zones attenuate the movement of chemicals. A deeper aquifer may have some dilution effect on contaminants reaching the water table. Precipitation prior to planting and during the growing season affects recharge. Temperature is likely to affect plant evapotranspiration. If the land-applied nitrogen load is high at land surface, there will likey be more nitrate available to leach to groundwater as a portion may not be utilized by plants. Many of these factors are not independent: for example, a shallow well is more likely to be installed where the water table is close to the land surface, which may also co-occur with sandier soils that permit more recharge. Monitoring wells may have higher nitrate concentrations by virtue of purposeful installation in locations known to be contaminated. This study highlights the use of CART to identify the relative importance of these interdependent factors, suggesting possible causal mechanisms for nitrate

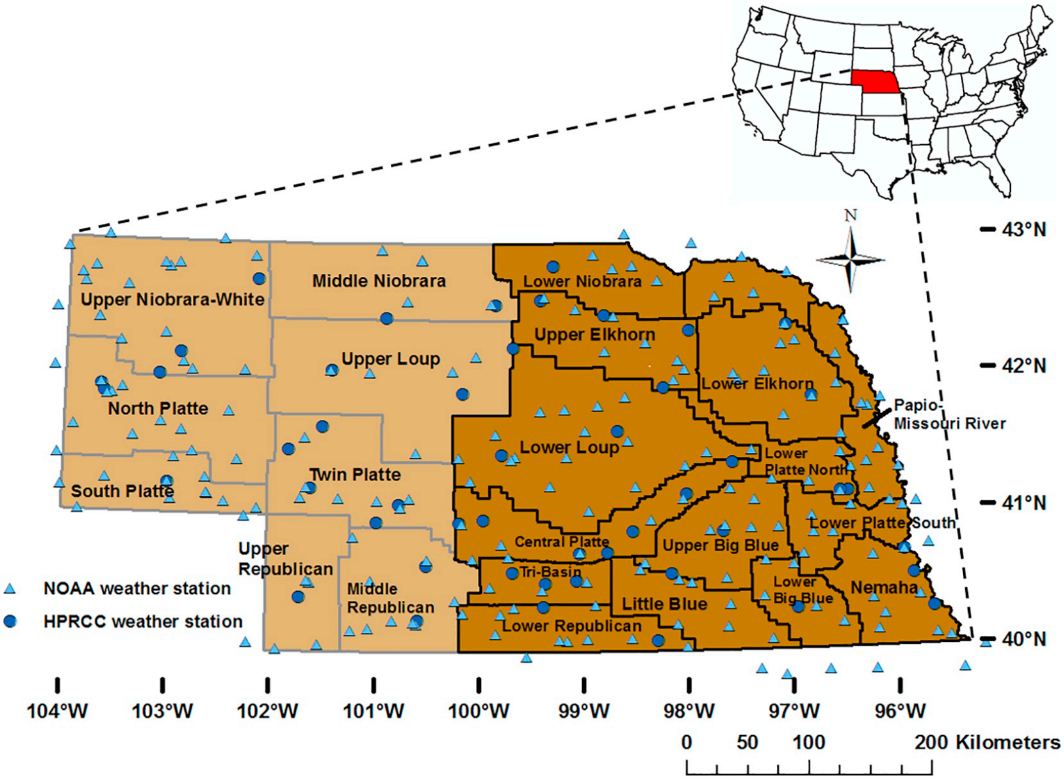


Fig. 1. The location of the study area and associated NRDs in western and eastern Nebraska with weather stations.

contamination in eastern and western Nebraska that can be further investigated using process-based modeling to get specific causal information for smaller areas.

The goal of this paper is to use three statistical techniques – spatial interpolation, pairwise regression, and CART – to identify patterns of nitrate contamination in Nebraska. Each of these techniques can help to demonstrate the factors that are associated with nitrate contamination in both space and time.

2. Materials and methods

2.1. The study area in western and eastern Nebraska

The study area is the state of Nebraska, representing a geographic area of ~200,000 km^2 between latitude 40°N to 43°N and longitude 95° 19′W to 104° 03′W. Nebraska has two major climatic zones: a humid continental climate (average annual precipitation ~750 mm) in the eastern part of the state and a semi-arid climate (average annual precipitation ~350 mm) in the western area of the state (HPRCC, 2016).

In this comparative analysis, the state of Nebraska is divided into west and east regions. The western region includes 8 Natural Resource Districts (NRDs), and the eastern region includes 15 NRDs (Fig. 1). There are 48 weather stations in Nebraska in the network of the High Plains Regional Climate Center (HPRCC) and 201 in the National Oceanic and Atmospheric Administration (NOAA), also shown in Fig. 1.

2.2. Data collection and analysis

2.2.1. Groundwater nitrate data

A total of 107,823 nitrate concentration measurements were obtained from 25,993 wells, distributed across 6282 wells in the western part and 19,711 wells in the eastern part of Nebraska for the time period from 1974 until 2013, from the Quality-Assessed Agrichemical Contaminant Database for Nebraska Groundwater (http://dnrdata.dnr.ne.gov/clearinghouse) (UNL, 2000). Most of the samples were taken in eastern Nebraska (76% of total samples during 1974 to 2013), which has a higher population and greater density of agricultural land. Well types are shown in Table 1.

Irrigation wells in the dataset have the longest screened intervals, with an average of 24 m, compared with 7 m for domestic wells and 5 m for monitoring wells. Most irrigation wells are screened for their entire length, in contrast with other well types which are not screened near the surface. Monitoring wells are commonly shallow, with an average well depth of 24 m. The average well depths of domestic and irrigation wells are 43 m and 69 m, respectively. More comparative descriptions of these well types between western and eastern Nebraska such as well depths, screen zones, pumping capacities, construction details, etc. are in Table 2.

For groundwater nitrate assessment, the NRDs and NDEQ collect samples during July and August every year (NDEQ, 2015). Samples are usually taken from a tap near the well head. Wells which are not in continuous operation are pumped for at least 2 h before water is sampled (Schepers et al., 1991; Exner et al., 2014). All samples are collected in polyethylene bottles and immediately put on ice until delivered for laboratory analysis. Samples are analyzed using the EPA-approved cadmium reduction method or HACH EPA equivalent/compliant methods (Exner et al., 2005; NDEQ, 2015).

To evaluate groundwater nitrate trends, if more than one concentration was reported in a well in a year, the maximum concentration was selected, because the maximum concentration is important for health. The dataset of maximum concentrations for all individual wells was divided into two groups, one for the western half of Nebraska and the other for the eastern part. The concentrations within western and eastern Nebraska were then averaged for each year.

2.2.2. Spatial characterization data

The distribution of irrigated and non-irrigated row crops was available from the MIRAD-US project under the USGS Early Warning and Environmental Monitoring Program (USGS, 2015) and the 2005 Nebraska Land Use Map (University of Nebraska-Lincoln, 2010). The

Table 1
The number of samples and well types including D = domestic, I = irrigation, Q = monitoring, C = commercial and industrial, and S = livestock wells that have been recorded in four decades (1974–1983; 1984–1993; 1994–2003; 2004–2013). Not all of these data were used in each part of this analysis.

Time periods	No. of samples	West					East				
		No. of wells					No. of wells				
		D	I	Q	C	S	D	I	Q	C	S
1974–1983	4748	426 1418	654	1	2	335	1763 3381	1255	134	3	226
1984–1993	15,986	857 1352	402	81	6	6	2048 6336	3308	907	21	52
1994–2003	42,597	669 3534	2045	811	2	7	1821 9970	7053	1030	19	47
2004–2013	44,492	127 2913	1906	849	5	26	923 8984	7406	562	44	49
1974–2013	**107,823**	**1518** **6282**	**3397**	**982**	**14**	**371**	**4499** **19,711**	**13,208**	**1607**	**67**	**330**

Table 2
General descriptions of well types including domestic, irrigation, and monitoring wells, followed by number of wells and number of samples collected. Not all of these measurements were used in every part of this analysis.

WEST (EAST)							
Characteristics	Well types	No. of data	Max	Min	Mean	SD	Source
Well depths (m)	Domestic	2583	265	2	48	32	UNL and NDNR
		-10,645	-296	-0.3	-39	-25	
	Irrigation	12,475	188	6	79	35	
		-52,072	-373	-4	-58	-30	
	Monitoring	13,570	248	3	26	19	
		-20,763	-160	-0.6	-23	-22	
Screen zones (m)	Domestic	299	43	3	8	5	UNL and NDNR
		-1510	-74	-1.2	-5.5	-4.5	
	Irrigation	6889	134	0.6	38	22	
		-42,900	-146	-0.3	-25.6	-20	
	Monitoring	12,590	129.5	0.6	7	8	
		-12,339	-30	-0.01	-3.6	-3.3	
Number of screen intervals (−)	Domestic	299	2/3%	1/ 97%	-	-	UNL and NDNR
		-1510	(3/ 0.07%)	(1/ 96%)	-	-	
	Irrigation	6889	4/ 0.16%	1/ 95%	-	-	
		-42,900	(3/ 1.03%)	(1/ 96%)	-	-	
	Monitoring	12,590	3/ 0.43%	1/ 99%	-	-	
		-12,339	(2/ 1%)	(1/ 99%)			
Pumping capacities (Gallons/Min)	Domestic	2583	2000	1	20	37.2	NDNR
		-10,645	-2000	-1	-19.5	-28.4	
	Irrigation	12,475	5000	2	1028	607.7	
		-52,072	-9020	-3	-876	-331	
	Monitoring	13,570	1334	1	19	74	
		-20,763	-1100	-1	-8.4	-45	
Construction details	All types	Under state regulations of Title 178, Chapter 12, "Water Well Standards" for a variety of intended uses (drinking water, irrigation, livestock watering, geothermal energy, or others), the NDHHS recommends that after drilling, a casing of either plastic (PVC), fiberglass, teflon or steel pipe will be placed in the bore hole. The casing must be extended at least 12 in. (~ 30 cm) above the surrounding land and is capped with a watertight seal on the top. The space between the bore hole and the well casing should be maintained a minimum of 2 in. (~ 5 cm) and must be grouted to protect surface water from running down the casing. A well screen is joined to the casing at one or more intervals in the aquifer's water-bearing zone. Clean sand or gravel that stabilizes the aquifer material, must be placed in the space between the bore hole and the screen while allowing water to move into the well.					UNL (IANR) and NDHHS

NDEQ is the Nebraska Department of Environmental Quality (http://dnrdata.dnr.ne.gov/clearinghouse).
NDNR is the Nebraska Department of Natural Resources (http://www.dnr.ne.gov/groundwater-data).
UNL (IANR) is the University of Nebraska-Lincoln (Institute of Agriculture and Natural Resources) (http://water.unl.edu/wells/design-construct). NDHHS is the Nebraska Department of Health and Human Services (http://dhhs.ne.gov/Pages/default.aspx).

data for irrigation systems in Nebraska were obtained from the University of Nebraska-Lincoln Conservation and Survey Division (School of Natural Resources, 2015) and the Center for Advanced Land Management Information Technologies 2005 Nebraska Land Use map (University of Nebraska-Lincoln, 2010). The spatial map of soil drainage capacities was made in accordance with Exner et al. (2014) by consolidating the seven drainage classifications of the Soil Survey Geographic Database (USDA, 2015) into three groups: excessively well drained, well drained and poorly drained. The spatial maps of corn and soybean production years were created by stacking raster layers of annual data from the National Agricultural Statistics Service (NASS) Cropland Data Layer (USDA-NASS, 2015).

2.2.3. Weather and crop price data

Weather data were collected from 48 stations of the HPRCC (http://hprcc6.unl.edu/cgi-hpcc/home.cgi) automated weather data network (AWDN) and 201 NOAA stations (http://www.ncdc.noaa.gov/cdo-web/datatools/findstation). Daily data included precipitation; wind speed; solar radiation; relative humidity; maximum, average and minimum temperatures; and potential evapotranspiration (ET_p) across the study area during 2004 and 2013.

Soil nitrate, which can be derived from most nitrogen materials in commercial fertilizers, biomass and animal wastes, is highly soluble in water and can easily be transported through soil to groundwater with recharge from agricultural land. Evapotranspiration (ET) is a highly variable and yet significant driving force (USGS, 2000) that is a primary determinant of the amount and timing of recharge. ET affects nitrate concentrations in groundwater through changes to the water balance, especially by decreasing recharge. Because most of the area where groundwater is exposed to nitrate contamination is fully irrigated, and therefore unlikely to experience any long-term moisture deficit, the potential evapotranspiration (ET_p) is a good approximation for the actual ET. In this study, ET_p is calculated on a daily time step using the Penman-Monteith (PM) equation with a fixed surface resistance of $45\,s\,m^{-1}$ and fixed plant height of 0.5 m for a reference surface of grass and alfalfa (Monteith, 1965).

Assuming that N applied per acre is independent of the number of acres in production, increasing crop land areas leads to more total N fertilizer application. Typical N application rates for corn and soybeans in Nebraska are ~180 kg/ha (based on an anticipated yield of 150 bu/ac) and ~100 kg/ha (in a case of nitrogen deficiency in soil), respectively (Shapiro et al., 2008; Ferguson et al., 2006). A higher corn price shifts the corn-soybean rotation in favor of continuous corn and encourages more fertilizer use. Thus, trends in total N fertilizer should be considered in the context of trends in corn price. Historical corn prices were obtained from the Department of Agricultural and Consumer Economics of the College of Agricultural, Consumer and Environmental Sciences at the University of Illinois (farmdoc.illinois.edu/manage/pricehistory/price:history.html).

2.3. Groundwater nitrate-contaminated areas

To differentiate factors controlling leaching and vulnerability to contamination, areas with high groundwater nitrate concentration were first identified and delineated. Outlined areas of nitrate concentrations (≥ 10 mg NO_3-N/L) were created using ArcGIS 10.3.1. The nitrate concentration in each well during each decade (1974–1983; 1984–1993; 1994–2003; 2004–2013) was computed across 2 km by 2 km grid cells using the point-to-raster conversion tool. Spatial analysis tools (interpolation, reclassify and contour) were used on grid cells with average concentration ≥ 10 mg NO_3-N/L. If more than one concentration was reported in a well in a year, the concentrations were averaged, in contrast with the trend analysis (Section 2.2.1), in which the greatest measured concentration was used in a year. The average concentration was more representative for spatial comparison among wells, because spatial interpolations are sensitive to outliers.

Four methods of interpolation in ArcGIS 10.3.1 (inverse distance weighting, ordinary kriging, interpolation from contours, and natural neighbor) were compared for spatially estimating averaged values of nitrate concentrations from wells in Table 1. This includes data from some multi-level monitoring wells, including from studies involving nitrates from point sources, which are not representative of nitrates from agricultural drainage. Monitoring wells represented up to 30% of wells in the west (2004–2014) and up to 14% of wells in the east (1984–1993). Nitrate levels from multi-level monitoring wells were averaged across depths before interpolation, since only one value can be used for a given location in any of the above interpolation methods. The selected interpolation methods were cross validated by reserving 10% of known data points from the database for error assessment. The results indicated that the natural neighbor method had the smallest error as calculated by the root mean square error (Fig. S1) and the percent of error (Table S3), and so this method was used to delineate areas of high nitrate concentrations (≥ 10 mg NO_3-N/L). Our methodology is different from the previous study by Exner et al. (2014) which used the "topo to raster" contour interpolation tool for determining the areas of high nitrate concentrations (≥ 10 mg NO_3-N/L) and excluded multi-level monitoring wells.

2.4. CART model for estimating groundwater nitrate concentrations

2.4.1. Methodology for the CART analysis

CART is a useful and popular tool in the field of data science. In this study, the program language "R" was selected to estimate the indicators for splitting nodes, using the package "rpart" (Loh, 2011). According to the description of Breiman et al. (1984), a CART model is principally built for classifying and predicting responses to covariates based on three steps including "tree" growing, pruning, and optimizing (Fig. 2).

"Tree" growing requires two steps, as described in the name of the tool: classification, and regression. The first step uses a classification tree – a recursive partitioning technique – to run several variables against the or dependent variable, to find the most robust and consistent method of sorting the observations into groups based on their similarity. The program checks the greatest improvement of the "purity" score of the resultant nodes (categories), to identify the best splitter in the case of the categorical variable, as well as alternative splitters ("surrogates") that would create similar groups. Thus CART splits the samples into populations with similar attributes, ensuring that the resulting populations are as similar to each other, and as different

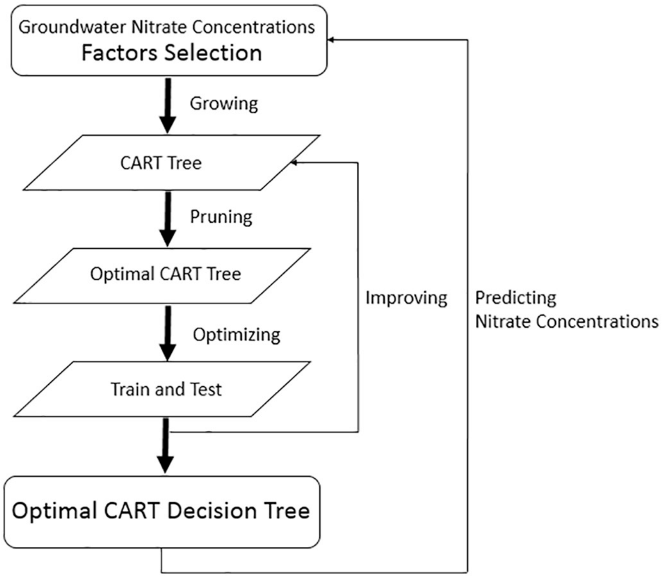


Fig. 2. The CART model for estimating groundwater nitrate concentrations.

Table 3
Principal equations of the CART model (Breiman et al., 1984; Yohannes and Webb, 1999).

Parameter	Equation	Definition
Gini coefficient	$Gini(t) = 1 - \sum_{j=1}^{k} p^2(j \mid t)$	$Gini(t)$ is an indication of the purity at node t, k is the number of categorical predictor variables, $p(j\|t)$ is the probability of a record output being in class j for the node t. When the $Gini(t)$ equals zero, it means all the observations in the node belong to a single group (the most purity).
The reduction of Gini coefficient	$\Delta Gini(t) = Gini(t) - p_L Gini(t_L) - p_R Gini(t_R)$	$\Delta Gini(t)$ is the reduction of Gini coefficient at node t (the greatest value means the best splitter), $Gini(t)$ is the Gini coefficient of output variables before reducing at node t, $Gini(t_L)$ and p_L are respectively the Gini coefficient and the proportions of observations at the left child node t_L, $Gini(t_R)$ and p_R are respectively the Gini coefficient and the proportions of observations at the right child node t_R.
The variance of the regression tree	$R(t) = \frac{1}{N-1} \sum_{j=1}^{k} (x_i(t) - \bar{x}(t))^2$	$R(t)$ is an indicator of the variance at node t, N is the number of observations for the node t, k is the number of categorical predictor variables, $x_i(t)$ is the output variables in class j for the node t, and $\bar{x}(t)$ is the mean of the output variables in class j for the node t.
The reduction of the variance	$\Delta R(t) = R(t) - p_L R(t_L) - p_R R(t_R)$	$\Delta R(t)$ is the reduction of variance at node t (the greatest value means the best splitter), $R(t)$ is the variance of output variables before reducing at node t, $R(t_L)$ and p_L are respectively the variance and the proportions of observations at the left child node t_L, $R(t_R)$ and p_R are respectively the variance and the proportions of observations at the right child node t_R.
The cost-complexity pruning	$R_{\propto}(T) = R(T) + \alpha \, \|\widetilde{T}\|$	$R_{\propto}(T)$ is the cost-complexity pruning of the decision tree, T, $R(T)$ is the error of classification in the decision tree, T, α is the complexity parameter, which will range from 0 to 1 and increase during the pruning process to represent how much additional accuracy is in the tree. When α is increased, the tree will be pruned. $\|\widetilde{T}\|$ is the number of child nodes.

Table 4
Selected factors for the analysis of the CART model.

WEST					
EAST					
Factors	Max	Min	Mean	SD	Source
Well attributes					
Well depth (m)	153	4.5	33	23	UNL and NDNR
	373	6	56	29	
Well age based on sampling date (yr)	75	0	20	14	UNL and NDNR (calculated from the difference between sampling date and completed date of wells)
	88	0	32	12	
Length of wells screen (m)	104	0.15	10	13	UNL and NDNR
	109	1.83	91	28	
Number of screen intervals (−)	2	1	1	0.1	UNL and NDNR
	5	1	2	0.3	
Soil characteristics and physical vadose zone properties					
Average percent sand in 0–150 cm (%)	95.73	9.79	61.87	21.37	NRCS SSURGO
	96.3	2.6	50.93	35.8	
Average percent organic matter in 0–150 cm (%)	1.77	0.27	0.82	0.35	NRCS SSURGO
	5	0.28	0.87	0.51	
Effective hydraulic conductivity (cm/day)	163.54	16.73	101.06	47.45	SNR UNL (calculated based on vertical flow and Pedotransfer Function in ROSETTA database)
	94.53	19.27	59.32	18.68	
Vadose zone thickness (m)	68	2.14	12	9	NDNR
	65	2.14	18	12	
Saturated thickness (m)	132	0.15	21	19	UNL and NDNR (calculated from well depth and vadose zone thickness)
	357	0.3	38.4	25	
Weather data					
Mean montly precipitation during Apr-Sep (mm) based on 2004–2013 data	113.47	27.71	66.04	16.15	PRISM Climate Group
	163.33	37.32	88.92	19.1	
Mean annual max temperature (°C) based on 2004–2013 data	20.27	14.02	17.55	1.02	PRISM Climate Group
	20.61	13.58	16.56	1.41	
Mean annual min temperatue (°C) based on 2004–2013 data	3.83	−0.71	1.69	0.79	PRISM Climate Group
	5.61	0.92	3.24	1.02	
Surface nitrate loading					
Nitrate-N load on surface around each well based on land cover (kg/ha/yr)	50.95	0.5	20.74	21.98	Literature review (see Table 5)
	50.95	1.44	36.67	17.55	

Abbreviations
UNL, University of Nebraska-Lincoln (http://dnrdata.dnr.ne.gov/clearinghouse).
NRCS SSURGO is the USDA Natural Resources Conservation Service and Soil Survey Geography database (https://websoilsurvey.sc.egov.usda.gov/App/HomePage.htm).
SNR UNL is the School of Natural Resources at the University of Nebraska-Lincoln (http://snr.unl.edu/data/geologysoils/NebraskaTestHole/NebraskaTestHoleMap.aspx).
NDNR, Nebraska Department of Natural Resources (https://dnr.nebraska.gov/data/groundwater-data).
PRISM Climate Group is the Parameter elevation Regression on Independent Slopes Model Climate Group (http://www.prism.oregonstate.edu/).

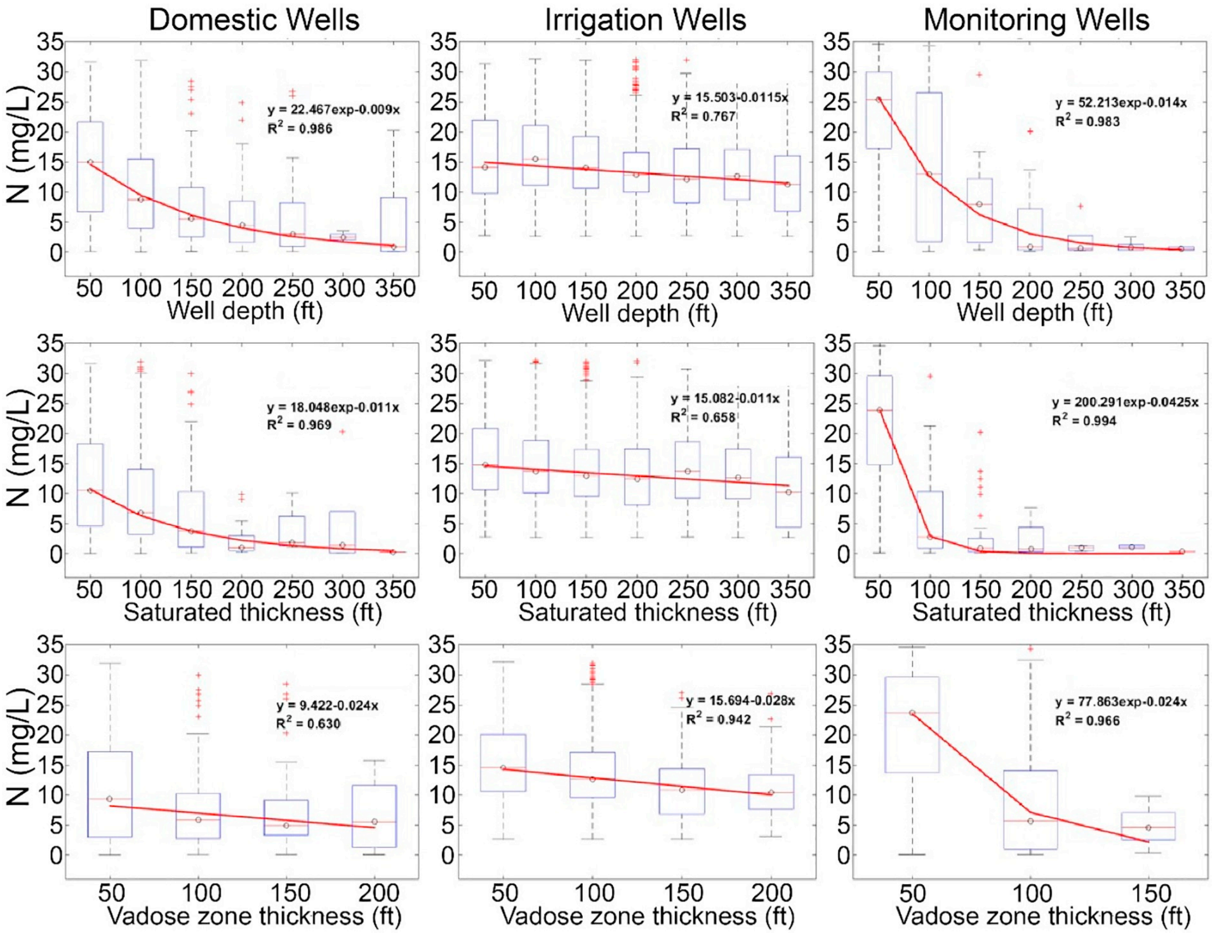


Fig. 3. Vadose zone characteristics (well depth, saturated and vadose zone thickness) with groundwater nitrate concentrations (1974–2013) within the 2004–2013 contaminated area in eastern Nebraska. 1 ft (ft) = 0.3048 m (m). The 2004–2013 contaminated area encompasses most of the contaminated area from previous decades.

from other groupings, as possible. For discrete inputs, such as well type in this analysis, the classification is evaluated using a metric known as a Gini coefficient (Breiman et al., 1984). The regression tree, which is generated for continuous input variables such as screen depth and sand content, has the same procedure with the classification tree, except it uses the variance between groups as the indicator instead of the Gini coefficient. CART automatically splits the observations into a large number of small subgroups with very similar characteristics, but only the first several splits are likely to be statistically significant; therefore it is necessary to get rid of the smallest "branches" in the decision tree.

After generating a detailed decision tree from the combination of classification (group membership) and regression (group values) (Fig. S2), the second step of the CART methodology is "tree pruning." Pruning cuts the "branches" of the tree to reduce over-fitting, thereby increasing the ability of new data prediction in the decision tree (Mingers, 1989). In CART, the "minimum cost complexity" pruning method is used to optimize the decision tree. Typically, the cost-complexity pruning threshold of the decision tree is considered equivalent to the decision tree error.

The final step in CART involves evaluating the pruned trees using the split test method, which is one method of examining the optimal tree (Dobbin and Simon, 2011). The observed data are divided into two groups, one for training and the other for testing. The groundwater nitrate concentration dataset within the contaminated area (NO_3-$N \geq 10\,mg\,N/L$) during 2004–2013 was randomly separated into two subsets, 80% of the data (12,880 samples) as the model training observations and 20% of the data (2000 samples) as the testing observations. In this study, the groundwater nitrate concentration was the target value, whereas weather conditions, well and soil characteristics, and surface nitrate-nitrogen loading for each well were chosen as the potential contributing factors based on a literature review (Tables 4 and 5). Groundwater nitrate concentrations were estimated and evaluated by a CART model through the procedure of growing, pruning, and optimizing based on the optimal tree. An overview of CART methodology is presented in Fig. 2 and the principal equations of the CART model are described in Table 3.

2.4.2. Factors for CART modeling

A CART model was created to identify the most significant factors affecting nitrate concentrations in Nebraska's groundwater, beginning with 13 potentially influential factors divided into four groups. These included well attributes, soil and vadose zone characteristics, weather conditions, and surface nitrate-nitrogen load around each well (see Table 4). Selection of these factors was also influenced by the availability of data.

Figs. 3 and 4 present box plots and trend lines of vadose zone characteristics (well depth, saturated thickness, and vadose zone thickness) with nitrate sample data for three major well types, domestic, irrigation, and monitoring, within the contaminated area (2004–2013) in western and eastern Nebraska. The monitoring well data used in these regressions includes data from multi-level monitoring wells for studies in eastern Nebraska, which were not used for CART analysis, although they were used for spatial interpolation.

Wells in a large part of eastern Nebraska exhibited shallower average well depth, less saturated thickness of the aquifer, and smaller vadose zone thickness along with a correlation of increasing nitrate

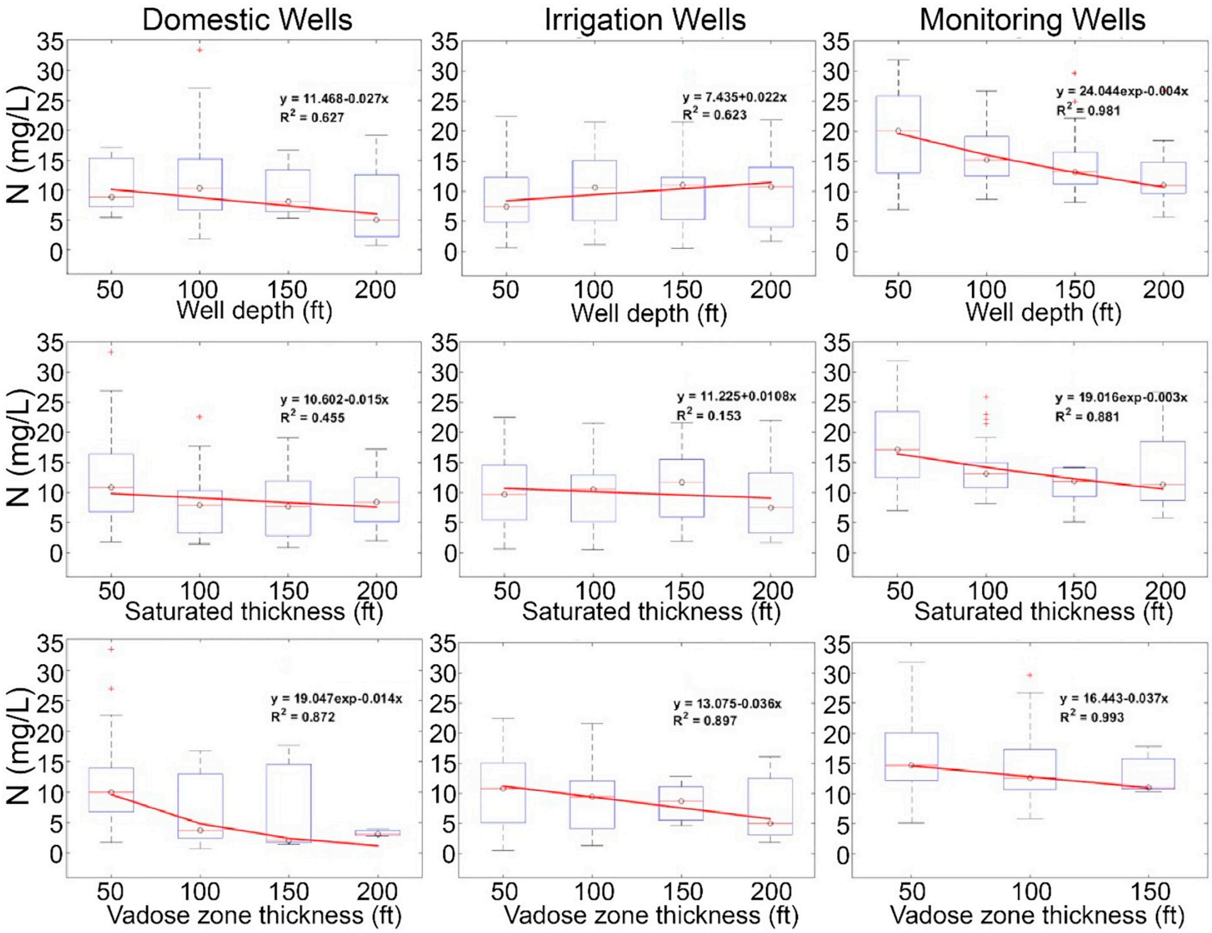


Fig. 4. Vadose zone characteristics (well depth, saturated and vadose zone thickness) with groundwater nitrate concentrations (1974–2013) within the 2004–2013 contaminated area in western Nebraska. 1 ft (ft) = 0.3048 m (m).

Table 5
Nitrate export coefficients assigned to land use codes based on 2004 to 2013 cropland data layer and Keeler and Polasky (2014).

Land use codes	Land use description	Nitrate-N loading (kg/ha/yr)
1	Corn	50.95
4	Sorghum	7.56
5	Soybeans	22.25
6	Sunflowers	7.56
12	Sweet corn	50.95
13	Popcorn	50.95
24	Winter wheat	7.56
26	Double crop winter wheat/soybeans	22.25
27	Rye	7.56
28	Oats	7.56
29	Millet	7.56
36	Alfalfa	10.16
42	Dry beans	7.56
44	Other crops	10.16
141	Deciduous forest	3.72
176	Pasture	3.2
190	Wetlands	1.44
225	Double crop winter wheat/corn	50.95
241	Double crop corn/soybeans	50.95

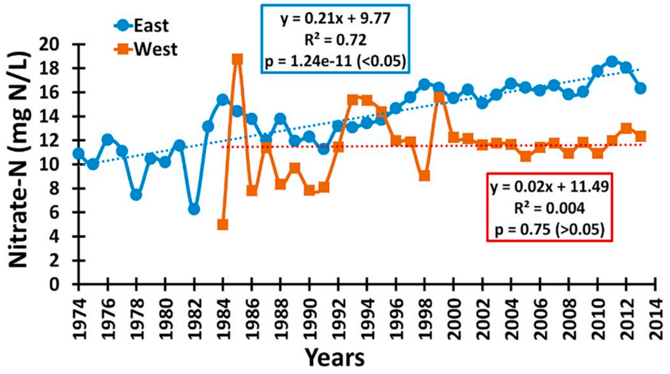


Fig. 5. Long-term trends of groundwater nitrate concentration (1974–2013) in western and eastern Nebraska.

concentrations with depths in wells of all types (domestic, irrigation and monitoring wells) as shown in Figs. 3 and 4. The contrasting shape of the regression in nitrate concentrations in domestic, irrigation, and monitoring wells probably occurs from differences in well attributes such as screen depths, well diameters, etc. The fitted correlation of median well depth, saturated thickness of aquifer, and vadose zone thickness with nitrate concentrations in monitoring wells presents an exponential relationship with R^2 of 0.98, 0.99, and 0.97, respectively. Comparatively, a linear relationship is found only in irrigation wells with R^2 of 0.77 (well depth), 0.66 (saturated thickness), and 0.94 (vadose zone thickness). Deeper well depths and larger thicknesses of saturated and vadose zones are associated with lower nitrate concentrations in groundwater. The shallow well depth and smaller thicknesses of the saturated and vadose zones in the 0–50 ft (i.e., 0–15 m) range of monitoring wells have definitively higher nitrate concentrations than in domestic and irrigation wells. The well attributes, e.g. screen depths and well diameters, likely impact groundwater

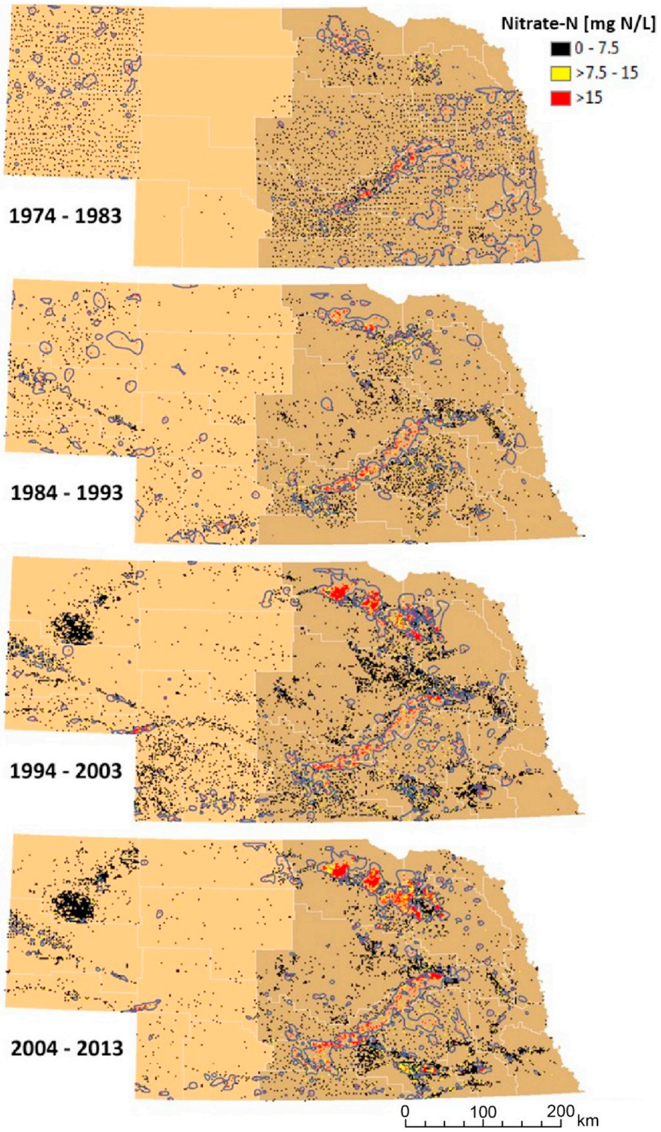


Fig. 6. The distribution of groundwater nitrate concentrations within the blue outlined areas of ≥10 mg NO_3-N/L during 1974–1983, 1984–1993, 1994–2003, and 2004–2013. NRD boundaries, which represent districts for nitrate management policies and which were used to delineate eastern and western Nebraska, are outlined in white. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

nitrate concentrations by affecting the part of the aquifer that the well draws from, particularly in the vertical profile of nitrate.

Based on these regressions, it is a valid concern that data from monitoring, irrigation, and domestic wells might not be comparable. In particular, monitoring wells are often put into place in areas of known nitrate contamination. However, CART analysis explicitly shows that in western Nebraska, monitoring wells in areas with a thin vadose zone and high hydraulic conductivity have more in common (in terms of nitrate concentrations) with irrigation and domestic wells in areas with a thin vadose zone and high hydraulic conductivity than with other monitoring wells that do not share these physical and geographical characteristics. No monitoring wells were included in the CART analysis for eastern Nebraska.

In the western part of Nebraska, domestic and monitoring wells have lower nitrate concentrations where the vadose zone is thicker. Except for irrigation wells, there are no obvious correlations of nitrate concentration with well depth and saturated thickness of the aquifer. Eastern Nebraska has higher nitrate concentrations in groundwater than in the west. Interestingly, at the same depth to water table or vadose zone thickness, groundwater nitrate concentrations in the east are higher than in the west. The difference in soil properties and other conditions between western and eastern Nebraska, as well as higher rainfall, recharge, and N inputs in the east, likely plays a significant role in predicting nitrate concentrations in groundwater.

Note that, while trends in groundwater concentrations over time were different between domestic, irrigation, and monitoring wells, the type of well was not a significant factor in absolute nitrate concentrations according to the CART model. The type of well was included as an input factor when creating the CART model, and if the mean groundwater nitrate concentration were significantly different among the types of wells, CART would have identified this factor as part of the classification tree. Instead, well type was identified in the west as a surrogate split in node 22, whereas everything beyond node 5 was considered statistically insignificant. In eastern Nebraska, no monitoring well data was used for CART analysis.

The depth from the land surface to the water table, referred to as vadose zone thickness, was estimated from a map of the depth to water. This was combined with the groundwater nitrate sample database of the University of Nebraska-Lincoln (UNL) between 1974 and 2013 using the extract values to points tool in ArcGIS 10.3.1. The depth to water map was created using the kriging interpolation tool using water depth data from 206,061 registered wells collected by the Department of Natural Resources (DNR) in Nebraska (http://www.dnr.ne.gov/groundwater-data). Saturated thickness was estimated from the difference between the well depth and the estimated depth to water table. From the estimation of the thickness of saturated zone, we found that many wells do not extend to the bottom of the aquifer in Nebraska. While the HPOA is known for its rapid depletion due to unsustainable use (Scanlon et al., 2012; McGuire, 2017; Haacker, et al., 2016), the depth to water in Nebraska has remained much more stable than the depth to water in other areas that rely on the aquifer for irrigation, such as the Kansas and Texas High Plains.

Average percent of sand and organic matter (OM) in the top soil (0–150 cm) were estimated from a map of a soil layer with the Soil Survey Geography (SSURGO) Mapunit Key (mukey). Each mukey is associated with specific soil characteristics (sand, silt, clay, organic matter, etc.). The soil layer was downloaded from SSURGO using the Web Soil Survey (WSS) operated by the USDA Natural Resources Conservation Service (NRCS) (WSS, 2017). Effective hydraulic conductivity (K_{eff}) was calculated from soil texture layers in the test hole, collected by the School of Natural Resources (SNR) at the University of Nebraska-Lincoln (UNL) (SNR-UNL, 2017). The K_{eff} was estimated based on vertical flow through soil layers. The saturated hydraulic conductivity (K_s) in each soil layer was predicted by the pedotransfer function in the ROSETTA database in HYDRUS-1D based on soil textures, which were obtained from the UNL-CSD test hole database.

Weather data, including monthly precipitation and maximum and minimum air temperature, were collected from the Parameter Elevation Regression on Independent Slopes Model (PRISM) Climate Group provided by the Oregon State University (PRISM, 2017). Land-use nitrate export coefficients were used for determining the annual average nitrate loading around each well under each year during 2004–2013. The land use data layers (2004–2013) were collected from the U.S. Department of Agriculture (USDA-NASS, 2016). The nitrate export coefficient for each land cover, which can indicate the amount of nitrates available for leaching from the surface due to inputs of fertilizers, crop residues, and atmospheric deposition (Reckhow and Simpson, 1980), were found using literature sources such as Parn et al. (2012) and Keeler and Polasky (2014). In this study, we used the nitrate export coefficients from Keeler and Polasky (2014) as shown in Table 5.

3. Results and discussion

3.1. Long-term trends of groundwater nitrate concentration

Fig. 5 presents long-term trends of groundwater nitrate concentration (1974–2013) in western and eastern Nebraska. Average annual groundwater nitrate concentrations in eastern Nebraska are increasing ($p < .05$), probably due to an increase in the intensity of crop production and the adoption of center pivots. Nitrate concentrations remain stable in the west, despite increasing irrigation intensity. Many of Nebraska's groundwater management programs implemented since 1988 may have reduced the loading contributing to increasing nitrate concentrations. Interestingly, it is only after the programs started that a clear upward trend begins in eastern wells, possibly a result of legacy nitrates reaching the water table.

3.2. Spatial distribution of groundwater nitrate contamination

Fig. 6 shows the contaminated areas of groundwater nitrate concentrations ≥ 10 mg NO_3-N/L during each of the previous four decades in Nebraska. Well depth, screen depth, and sampling date were obtained from the well database of the University of Nebraska-Lincoln (UNL, 2000). Within the groundwater nitrate contaminated areas during 2004–2013 (outlined in Fig. 6), there are a total of 40,758 available nitrate samples: 36,835 samples in the eastern part and 3923 samples in the western part of Nebraska. These data are based on nitrate samples in the 40-year database from 1974 to 2013, which covers only part of the state. Another 67,065 nitrate samples were excluded from this study because their highest nitrate concentrations were under the MCL of 10 mg NO3-N/L. Based on the evaluation of average nitrate concentrations across the state, the groundwater nitrate-contaminated

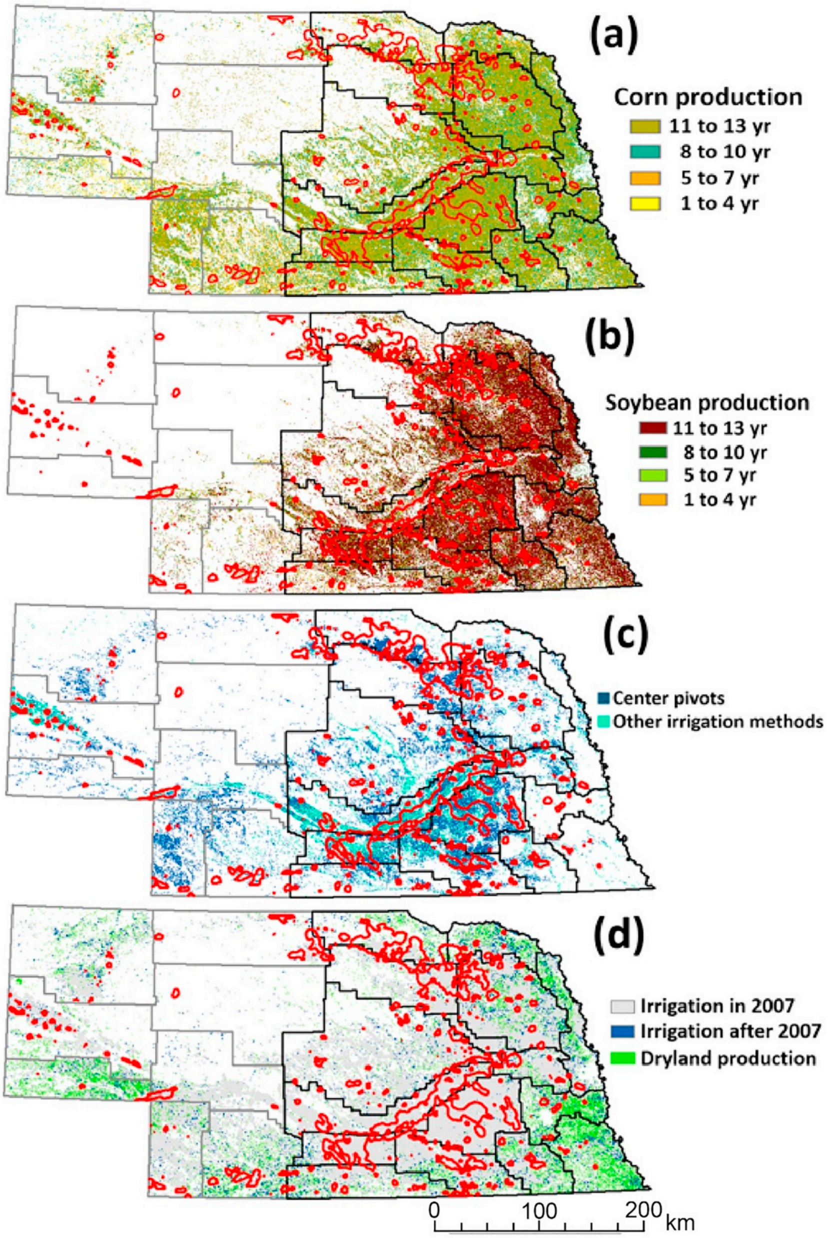


Fig. 7. Anthropogenic factors: (a) years of corn production between 2002 and 2014, (b) years of soybean production between 2002 and 2014, (c) irrigation systems, and (d) irrigated and non-irrigated (dryland) crops, with groundwater nitrate concentrations within the red outlined areas of ≥ 10 mg NO_3-N/L during 2004–2013. The original map is of sufficient resolution (30 m) to show individual fields planted to corn or soybeans. Natural factors: (e) Soil drainage capacities, (f) depth to water table, and (g) saturated thickness with groundwater nitrate concentrations within the red outlined areas of ≥ 10 mg NO_3-N/L during 2004–2013. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

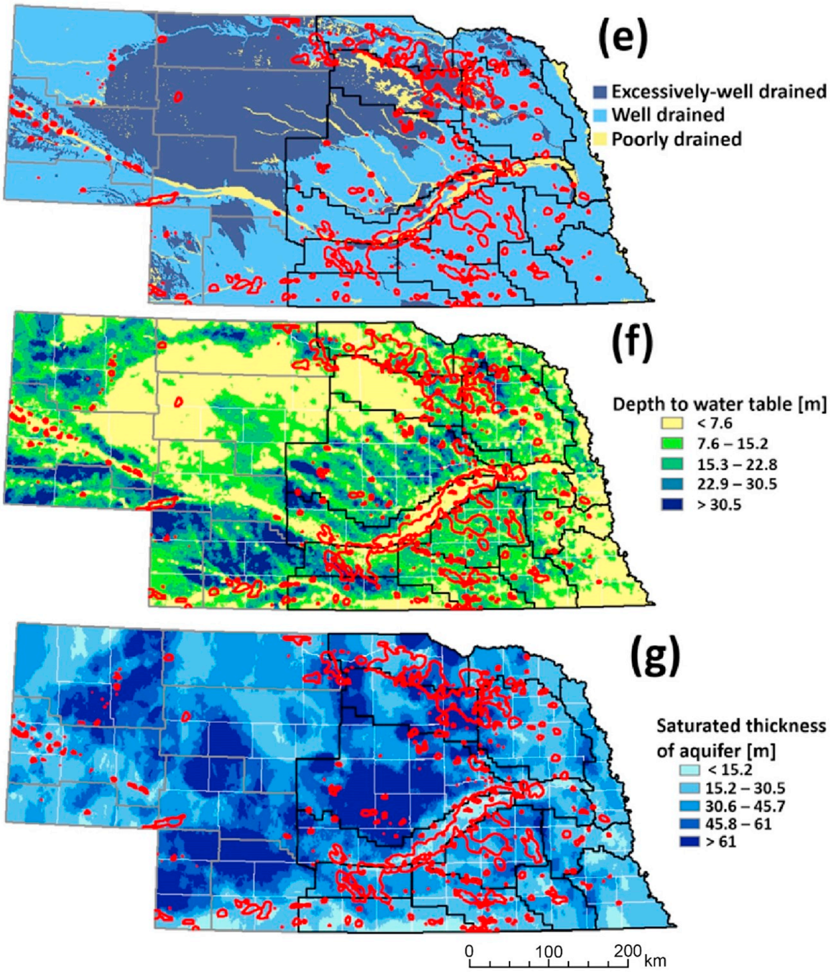


Fig. 7. (*continued*)

area in the west seems to be much smaller and more stable than in eastern Nebraska. On the other hand, the contaminated area (2004–2013) in the east part of Nebraska was more expansive than the previous decades, which makes sense given the long-term trends of groundwater nitrate concentration in Fig. 5.

3.3. *The relationships between groundwater nitrate concentrations and potential natural and anthropogenic factors*

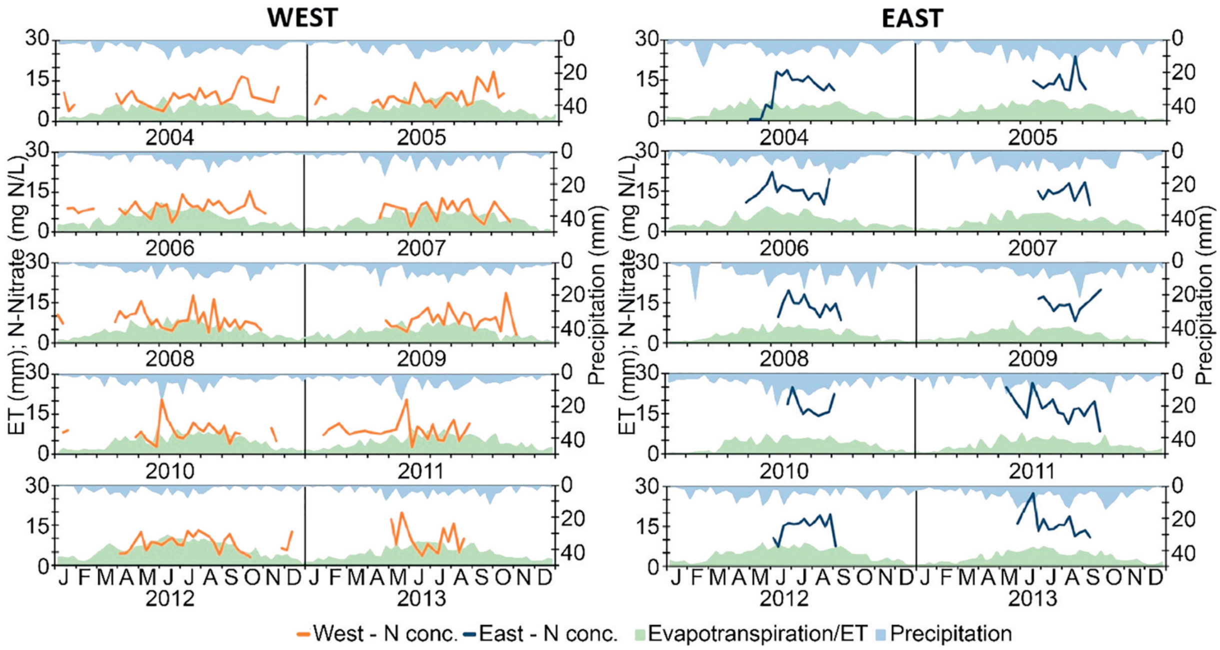
3.3.1. *Characterization of nitrate-contaminated areas*

As shown in Fig. 7a and b, extensive production of corn and soybean coincides with groundwater nitrate-contaminated areas. Currently, the total area of corn production is growing both in western and eastern Nebraska. Limited production of corn and the absence of soybean in western Nebraska resulted in limited groundwater nitrate-contaminated areas in western Nebraska compared to eastern Nebraska.

Groundwater nitrate-contaminated areas are primarily irrigated with gravity irrigation systems in the eastern central part of Nebraska. This contamination is likely due to more water ponding in this type of irrigation, both within furrows and as tailwater. Contaminated areas tend not to occur in non-irrigated (dryland) agricultural areas, where less water recharges than in the irrigated areas, particularly in south-eastern Nebraska. In the west, smaller groundwater nitrate-contaminated areas were found in areas of dryland production and larger saturated thickness of the aquifer, as shown in Fig. 7c and d.

The soil drainage capacity can be an important factor in system response to nitrate applications. In the north-eastern part of the state, especially around the boundary between the Upper Elkhorn and Lower Elkhorn NRDs (Fig. 7e), groundwater under well-drained and excessively well-drained soils frequently exceeds the maximum contaminant level (MCL) of 10 mg NO3-N/L. This is in accordance with Spalding and Hirsh (2012), who mention that the north-eastern area of the state has faced groundwater nitrate contamination beneath intensively spray-irrigated areas with coarse-textured soils. In fact, high levels of nitrate concentration occur in some areas where there is poor soil drainage, particularly in riparian areas in the central-eastern part of the state, such as the Platte River Valley in the Central Platte NRD (Fig. 7e). In the past, gravity irrigation dominated these areas. Although soil drainage is low compared to sand, extended ponding time within furrows could have caused greater leaching (Spalding et al., 1978). Extensive groundwater-surface water connectivity may also influence nitrate concentrations, and tributaries feeding the Platte River may also contribute nitrate to groundwater. The application rates of anhydrous ammonia fertilizer declined during the 1980's in the central area of the state due to lower crop prices (Exner et al., 2010). In comparison with the edges of the state, more of central Nebraska is well drained to excessively well-drained, to the extent that row crop production is not economically feasible. This part of the state is known as the Sand Hills and is visible as the large area of low crop production in Fig. 7. Except for a few intensively farmed areas near Alliance, Nebraska, the north central part of the state has not been threatened as intensively from high groundwater nitrate contamination, despite the vulnerability of the aquifer in areas with sandy soils.

The depth to the water table plays a significant role in predicting



	N conc. (west)	Precipitation (west)	ET (west)
N conc. (west)	**1**		
Precipitation (west)	**0.274**	**1**	
ET (west)	**-0.027**	**0.131**	**1**

	N conc. (east)	Precipitation (east)	ET (east)
N conc. (east)	**1**		
Precipitation (east)	**0.266**	**1**	
ET (east)	**0.061**	**-0.159**	**1**

Fig. 8. Precipitation and ET_p with groundwater nitrate data in the high nitrate areas (≥ 10 mg NO_3-N/L) during 2004 and 2013 considering weekly trends in western and eastern Nebraska (top) and the Pearson's correlation coefficient (bottom).

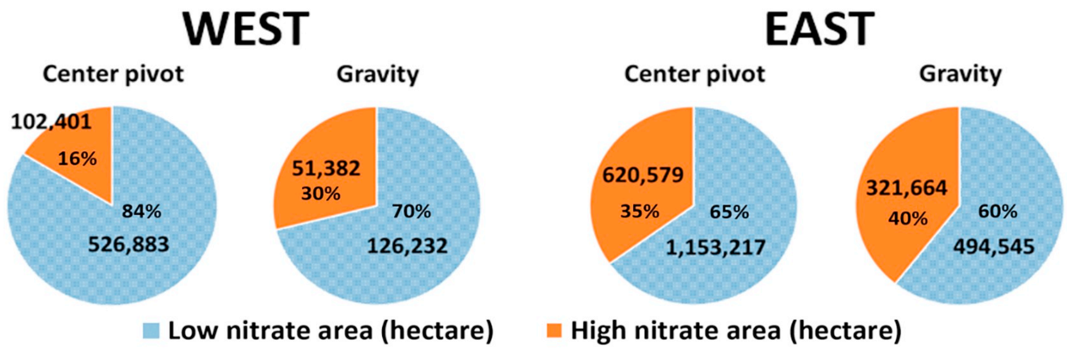
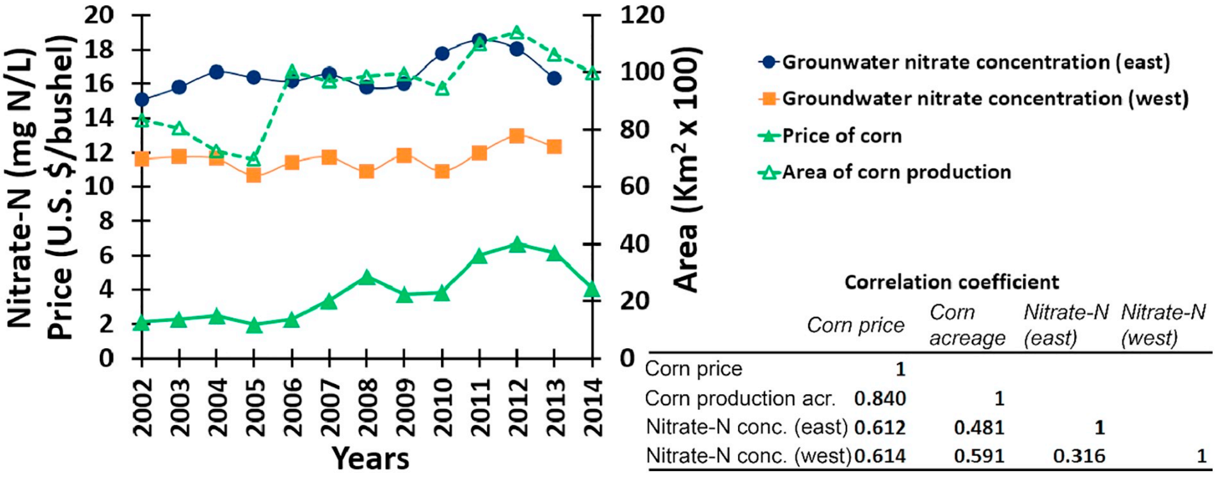


Fig. 9. Areas of high (exceeding or at the maximum contaminant level, MCL) nitrate concentrations (≥ 10 mg NO_3-N/L) and low (below the MCL) nitrate concentrations (< 10 mg NO_3-N/L) under different irrigation systems in western and eastern Nebraska.

groundwater nitrate concentration. A deeper depth to water table reduces the probability for nitrate contamination compared with a shallower water table (i.e. thin vadose zone). This was found in both western and eastern Nebraska (Fig. 7f) and is consistent with the previous studies such as Spalding and Hirsh (2012) and Exner et al. (2014). In addition, the saturated thickness of the aquifer also influences groundwater nitrate concentrations, likely due to dilution and mixing. Areas with greater saturated thickness in both western and eastern Nebraska have lower nitrate concentrations in groundwater (Fig. 7g).

3.3.2. Assumptions of the validity of the extent of nitrate contamination in groundwater

After outlining the area of high nitrate concentrations (≥ 10 mg NO_3-N/L) in the last decade (2004–2013), factors controlling groundwater nitrate concentrations within the contaminated area were considered and compared between western and eastern Nebraska. These factors include depth to water table, saturated thickness of the aquifer (the distance from the top to the bottom of the water-bearing sediment), precipitation, evapotranspiration, cropland areas and prices of corn and soybeans.



Correlation coefficient

	Corn price	Corn acreage	Nitrate-N (east)	Nitrate-N (west)
Corn price	1			
Corn production acr.	0.840	1		
Nitrate-N conc. (east)	0.612	0.481	1	
Nitrate-N conc. (west)	0.614	0.591	0.316	1

Fig. 10. The comparison of corn prices, production area and groundwater nitrate concentrations in western and eastern Nebraska, 2002 to 2014.

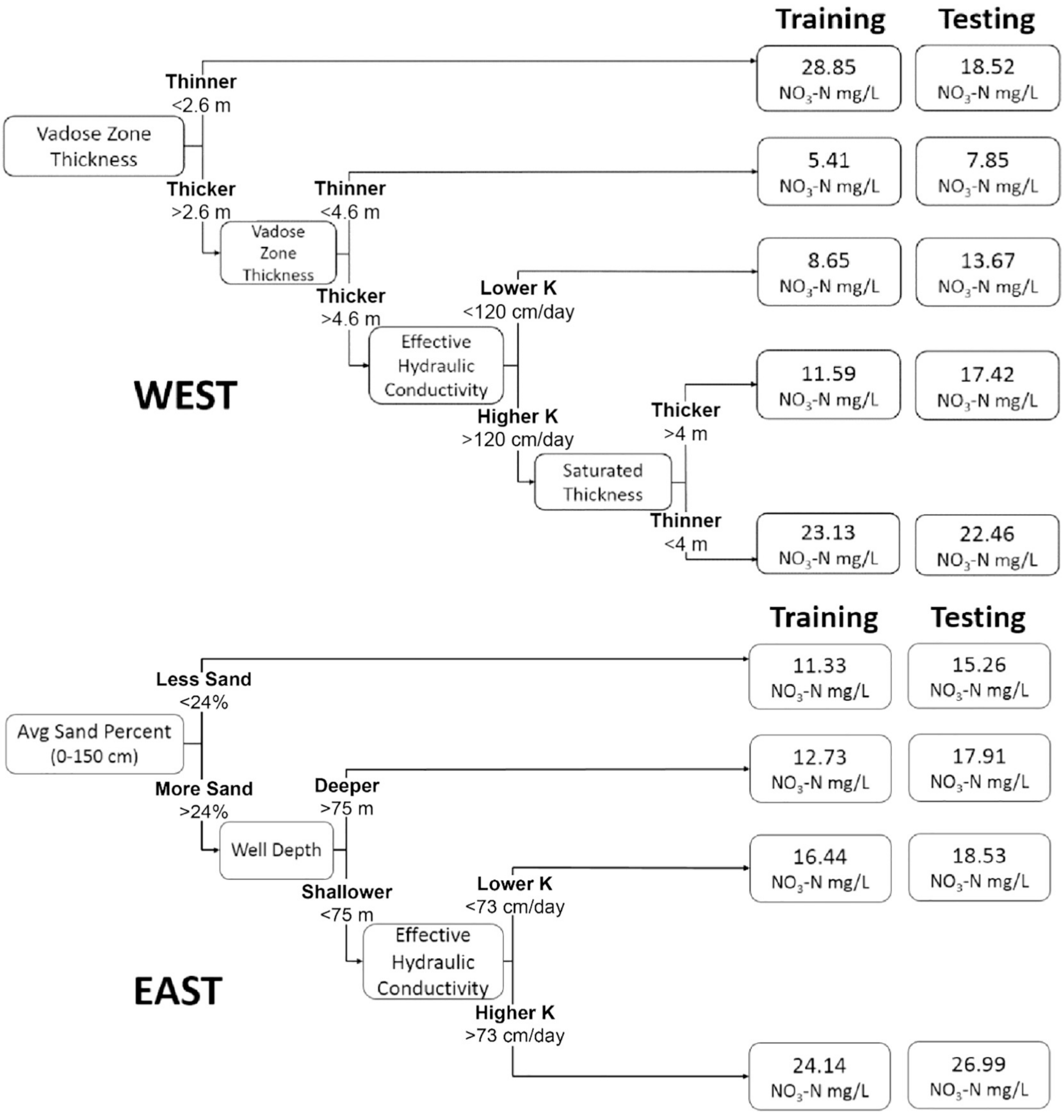


Fig. 11. The output of CART pruning tree for groundwater nitrate concentration prediction in western and eastern Nebraska.

Table 6
The error of the CART model after the validation.

Data	WEST				EAST			
	No. observations	NO_3-N conc.	Absolute error (mg N/L)	Relative error (%)	No. observations	NO_3-N conc.	Absolute error (mg N/L)	Relative error (%)
	Node # 1				Node # 1			
Training	421	5.41	2.44	31.08	3,941	11.33	3.93	25.75
Testing	10	7.85			564	15.26		
	Node # 2				Node # 2			
Training	1,274	8.65	5.07	37.08	1,796	12.73	5.18	28.92
Testing	136	13.67			266	17.91		
	Node # 3				Node # 3			
Training	812	11.59	5.83	33.46	3,548	16.44	2.09	11.28
Testing	83	17.42			685	18.53		
	Node # 4				Node # 4			
Training	44	23.13	0.67	2.98	1,025	24.14	2.85	10.56
Testing	13	22.46			199	26.99		
	Node # 5							
Training	19	28.85	10.33	55.77				
Testing	44	18.52						
Min	10	5.41	0.67	2.98	29	11.33	2.09	10.56
Max	1,274	28.85	10.33	55.77	3,941	40.09	5.18	28.92
Mean	286	15.75	4.87	32.08	1,202	21.57	3.51	19.13

Fig. 8 presents precipitation and potential evapotranspiration (ET_p) with groundwater nitrate data in the nitrate-contaminated (≥ 10 mg NO_3-N/L) area during 2004 to 2013 by considering weekly trends in western and eastern Nebraska. Precipitation has a statistically significant correlation ($p < .05$) with nitrate concentrations in both western and eastern Nebraska (Fig. S3). Total annual precipitation in the eastern part of Nebraska is higher than in the western region (Fig. 8). The ET_p differs slightly between western and eastern Nebraska when considering irrigated cropland. ET_p is not significantly related with groundwater nitrate concentrations ($p > .05$) (Fig. S3). However, precipitation and ET_p each show a small Pearson's correlation coefficient (~0.27 between precipitation and nitrate concentrations in both west and east, and −0.03 and 0.06 between ET_p and nitrate concentrations in west and east, respectively) (Fig. 8).

Fig. 9 presents high nitrate (≥ 10 mg NO_3-N/L) and lower nitrate (< 10 mg NO_3-N/L) areas under different irrigation systems in western and eastern Nebraska. Based on this study, center pivot irrigated areas have less risk of groundwater nitrate contamination than gravity irrigation systems. In eastern Nebraska (2,590,005 ha of irrigated area), there is about three times more irrigated area than in the west (806,898 ha of irrigated area), and the increased density of irrigated crops is directly related to high nitrate in groundwater. These graphs were created by overlaying a map of center pivot and gravity irrigated land with the map of nitrate-contaminated areas in 2004–2013. The total amount of land for each irrigation technology was then compared with the irrigated land from 2012 from MODIS-MIrAD falling within an identified nitrate-contamination zone (Fig. 6). Irrigation technology was obtained from USGS (earlywarning.usgs.gov/USirrigation).

Commodity prices are a strong driver of crop and nutrient management, and very likely play a leading role in increasing groundwater nitrate concentrations. Fig. 10 compares corn prices, area of production and groundwater nitrate concentrations in western and eastern Nebraska during 2002 to 2014. Corn price and production area have a correlation coefficient of 0.84. Application of nitrogen fertilizer is higher when corn production areas increase, but the trend of groundwater nitrate concentrations shows a smaller correlation coefficient (0.48 for eastern Nebraska and 0.59 for western Nebraska) with corn production areas, as compared to corn prices (0.61 for both east and west). This implies that nitrogen is applied more intensively (less crop rotation) when corn prices are high, which may lead to increases in nitrate leaching to groundwater.

3.4. *Estimation of groundwater nitrate concentrations based on CART model*

Groundwater nitrate concentrations were forecast in western and eastern Nebraska through the optimized CART model. We found that vadose zone thickness, effective hydraulic conductivity, and saturated thickness were the most significant factors influencing groundwater nitrate concentrations in western Nebraska with an explanatory power of 30%, 16% and 12%, respectively, i.e. vadose zone thickness accounted for 30% of the total variability in nitrate concentrations. In eastern Nebraska, the most influential factors were average percent of sand in the top 0–150 cm of soil (21%), well depth (18%), and effective hydraulic conductivity (14%) (Fig. 11). The explanatory power is a measure of how much of the nitrate concentration can be attributed to each of these factors. The results of the model show the mean absolute error in predicting nitrate concentrations are 4.87 mg NO_3-N/L (west) and 3.51 mg NO_3-N/L (east). The relative errors are 32% (west) and 19% (east) as presented in Table 6. This indicates that the CART model can be used to predict groundwater nitrate concentrations most accurately in eastern Nebraska, which is unsurprising given the larger quantity of data available in the east.

4. Conclusions

This study confirms that nitrate contaminated areas are expanding and new areas continue to emerge beneath irrigated cropland in Nebraska, particularly in the east. It is possible that some wells within the identified high-nitrate areas are still below MCL as all existing wells were not included in the database. However, nitrate concentrations in additional wells can be predicted using the optimal CART model of this study, with an expected accuracy of about 70–80%. The trends of increasing groundwater nitrate concentrations have occurred only in eastern Nebraska following an increase in the intensity of crop production and irrigation. While the rate of increase of average nitrate groundwater concentration has slowed in some areas under the Nebraska's GWQMP, intense irrigation increases the rate of nitrate leaching to groundwater.

This study additionally shows that the areal extent and growth of contaminated groundwater in predominately center pivot-irrigated areas is lower than beneath gravity-irrigated areas. Converting from gravity irrigation to center pivot may help to protect against the expansion of nitrate-contaminated groundwater. Based on this study, the spatial differences in climate, soil, cropping, irrigation and vadose zone characteristics (e.g. precipitation, ET, soil drainage capacities, depth to water table) significantly predict groundwater nitrate concentrations in western and eastern Nebraska. Thus, the investigation of soil nitrogen processes and nitrate flux through soil into groundwater under climate variability and the complexities of aquifer and vadose zone characteristics are key for analyzing the occurrence of nitrate contamination in Nebraska's groundwater.

The CART model was used to identify the relative importance of well attributes, soil and vadose zone characteristics, weather conditions, and nonpoint-source N inputs for each well with groundwater nitrate concentrations. Physical characteristics – geography and well construction – were found to be significant, irrespective of well type (irrigation, monitoring, or domestic). This supports Burow et al.'s (2010) use of monitoring wells for predicing groundwater nitrate concentrations in CART modeling. Vadose zone thickness and well depth were found to be the most significant factors affecting groundwater nitrate concentrations in western Nebraska, with an explanatory power of 30%, 16%, and 12%, respectively. The most influential factors included average percent of sand in the 0–150 cm topsoil (21%), well depth (18%), and effective hydraulic conductivity (14%) in eastern Nebraska. After testing the model, we conclude that the CART model can be applied to predict groundwater nitrate concentrations from those influential factors. The CART model proved to be useful for both prediction of groundwater nitrate concentrations, and for identifying potential factors that could place areas at greater risk for groundwater contamination.

Acknowledgments

This paper is the result of an effort between the Nebraska Water Center (NWC) and the Water Sciences Laboratory at the University of Nebraska-Lincoln to protect Nebraska's drinking water resources. The authors are thankful for all postdoctoral researchers under the NWC who gave scientific guidance, participated in discussions, and provided expertise that greatly assisted the study.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jconhyd.2018.11.007.

References

Adelman, D.D., Schroeder, W.J., Smaus, R.J., Wallin, G.P., 1985. Overview of nitrate in Nebraska's ground water. Trans. Nebraska Acad. Sci. Affiliat. Soc. 220. http://digitalcommons.unl.edu/tnas/220.

Breiman, L., Friedman, J.H., Olshen, R., Stone, C.J., 1984. Classification and Regression Tree. Wadsworth & Brooks/Cole Advanced Books & Software, Pacific California.

Burow, K.R., Nolan, B.T., Rupert, M.G., Dubrovsky, N.M., 2010. Nitrate in groundwater of the United States, 1991–2003. Environ. Sci. Technol. 44, 4988–4997.

Comly, H.H., 1945. Cyanosis in infants caused by nitrate in well water. JAMA 129, 112–116.

CPNRD (Central Platte Natural Resources District), 2016. Groundwater Management Plan. Grand Island, Nebr. http://cpnrd.org/groundwater-quality/.

Dappen, P., Merchant, J., Ratcliffe, I., Robbins, C., 2007. Delineation of 2005 land use patterns for the state of Nebraska department of natural resources. Final Rep. Lincoln. NE. http://www.calmit.unl.edu/2005landuse/data/2005_NE_landuse_finalreport.pdf.

Dobbin, K.K., Simon, R.M., 2011. Optimally splitting cases for training and testing high dimensional classifiers. BMC Med. Genet. 4 (31). http://www.biomedcentral.com/1755-8794/4/31.

Exner, M.E., Hirsh, A.J., Spalding, R.F., 2014. Nebraska's groundwater legacy: nitrate contamination beneath irrigated cropland. Water Resour. Res. 50, 4474–4489. https://doi.org/10.1002/2013WR015073.

Exner, M.E., Perea-Estrada, H., Spalding, R.F., 2010. Long-term response of groundwater nitrate concentrations to management regulations in Nebraska's central platte valley. Sci. World J. 10, 286–297. https://doi.org/10.1100/tsw.2010.25.

Exner, M.E., Spalding, R.F., 1976. Groundwater quality of the central Platte region, 1974. Lincoln, University of Nebraska-Lincoln, Conservation and Survey Division Resource Atlas No. 2, (47 p).

Exner, M.E., Spalding, R.F., 1994. N-15 identification of nonpoint sources of nitrate contamination beneath cropland in the Nebraska Panhandle: two case studies. Appl. Geochem. 9 (1), 73–81. https://doi.org/10.1016/0883-2927(94)90054-X.

Ferguson, R.B., 2015. Groundwater quality and nitrogen use efficiency in Nebraska's Central platte river calley. J. Environ. Qual. 44, 449–459. https://doi.org/10.2134/jeq2014.02.0085.

Ferguson, R.B., Shapiro, C.A., Dobermann, A.R., Wortmann, C.S., 2006. Fertilizer Recommendations for Soybean. University of Nebraska-Lincoln Extension educational programs. http://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=2712&context=extensionhist last accessed on July 2, 2018.

Grassini, P., Yang, H., Irmak, S., Rees, J., Cassman, K.G., 2012. Evaluation of Water Productivity and Irrigation Efficiency in Nebraska Corn Production. University of Nebraska-Lincoln Extension EC105.

Gurdak, J.J., McMahon, P.B., Dennehy, K.F., Qi, S.L., 2009. Water quality in the High Plains aquifer, Colorado, Kansas, Nebraska, New Mexico, Oklahoma, South Dakota, Texas, and Wyoming, 1999–2004. U.S. Geol. Survey Circular. 1337, 63.

Gurdak, J.J., Qi, S.L., 2006. Vulnerability of recently recharged ground water in the High Plains aquifer to nitrate contamination. U.S. Geol. Survey Sci. Invest. Rep. 2006–5050, 39.

Gutentag, E.D., Heimes, F.J., Krothe, N.C., Luckey, R.R., Weeks, J.B., 1984. Geohydrology of the High Plains Aquifer in parts of Colorado, Kansas, Nebraska, New Mexico, Oklahoma, South Dakota, Texas and Wyoming. U.S. In: Geological Survey Professional Paper 1400-B, . https://pubs.usgs.gov/pp/1400b/report.pdf.

Haacker, E.M., Kendall, A.D., Hyndman, D.W., 2016. Water level declines in the High Plains Aquifer: predevelopment to resource senescence. Groundwater 54 (2), 231–242.

Hiller, T.L., Powell, L.A., McCoy, T.D., Lusk, J.J., 2009. Long-term agricultural land-use trends in Nebraska. 1866–2007. Great Plains Res. 19, 225–237.

HPRCC (High Plains Regional Climate Center), 2016. HPRCC Climate Summaries. https://hprcc.unl.edu/climatesummaries.php last accessed on July 2, 2018.

Keeler, B.L., Polasky, S., 2014. Land-use change and costs to rural households: a case study in groundwater nitrate contamination. Environ. Res. Lett. 9, 074002. https://doi.org/10.1088/1748-9326/9/7/074002.

Kross, B.C., Olson, M.L., Ayebo, A., Johnson, J.K., 1995. In: Rechchigl, J.E. (Ed.), In Soil Amendments: Impacts on Biotic Systems. Lewis Publishers, Boca Raton, pp. 153–214.

Lockhart, K.M., King, A.M., Harter, T., 2013. Identifying sources of groundwater nitrate contamination in a large alluvial groundwater basin with highly diversified intensive agricultural production. J. Contam. Hydrol. 151, 140–154. https://doi.org/10.1016/j.jconhyd.2013.05.008.

Loh, W.Y., 2011. Classification and Regression Trees. John Wiley & Sons, Inc. WIREs Data Mining Knowl. Discov. 1. pp. 14–23. https://doi.org/10.1002/widm.8.

McGuire, V.L., 2017. Water-level and recoverable water in storage changes, High Plains aquifer, predevelopment to 2015 and 2013–15 (No. 2017–5040). US Geological Survey.

Mingers, J., 1989. An empirical comparison of pruning methods for decision tree induction. Mach. Learn. 4, 227–243.

Monteith, J.L., 1965. Evaporation and the environment. In: Proc. 19th Symp. Society for Experimental Biology: The State and Movement of Water in Living Organisms. Cambridge University Press, Cambridge, U.K., pp. 205–234.

Murgulet, D., Tick, G.R., 2013. Understanding the sources and fate of nitrate in a highly developed aquifer system. J. Contam. Hydrol. 155, 69–81. https://doi.org/10.1016/j.jconhyd.2013.09.004.

NAS (National Academy of Sciences), 1995. Board on Environmental Studies and Toxicology. Nitrate and Nitrite in Drinking Water. National Academy of Sciences, Washington, D.C.. https://www.nap.edu/read/9038/chapter/3 last accessed on June 22, 2018.

NDEQ (Nebraska Department of Environmental Quality), 2015. 2015 Nebraska Groundwater Quality Monitoring Report. Annual Rep. Lincoln, NE. http://deq.ne.gov/publica.nsf/Pages/WAT230 last accessed on May 12, 2018.

Nolan, B.T., Hitt, K.J., 2006. Vulnerability of shallow groundwater and drinking-water wells to nitrate in the United States. Environ. Sci. Technol. 40 (24), 7834–7840.

Nolan, B.T., Ruddy, B.C., Hitt, K.J., Helsel, D.R., 1997. Risk of nitrate in groundwaters of the United States – a national perspective. Environ. Sci. Technol. 31 (8), 2229–2236.

NRD (Nebraska's Natural Resources Districts), 2017. NRD Groundwater Management Summary. https://www.nrdnet.org/sites/default/files/groundwater_management_summary_2017_1.pdf last accessed on June 20, 2018.

Parn, J., Pinary, G., Mander, U., 2012. Indicators of nutrients transport from agricultural catchments under temperate climate: a review. Ecol. Indic. 22, 4–15.

PRISM, 2017. The Parameter Elevation Regression on Independent Slopes Model (PRISM) Climate Group. http://prism.oregonstate.edu/.

Qi, C.S., Yu, Z.H., Hou, Z., 2010. Study on quality prediction of the complex production based on CART algorithm. Modul. Mach. Tool Automatic Manuf. Tech. 3, 94–97.

Reckhow, K., Simpson, J., 1980. A procedure using modeling and error analysis for the prediction of lake phosphorus concentration from land use information. Can. J. Fish. Aquat. Sci. 37, 1439–1448.

Sattelmacher, P.G., 1962. Methemoglobinemia from nitrates in drinking water, Schriftenreiche des Vererins fur Wasser Boden und Lufthygiene. (No. 21).

Scanlon, B.R., Faunt, C.C., Longuevergne, L., Reedy, R.C., Alley, W.M., McGuire, V.L., McMahon, P.B., 2012. Groundwater depletion and sustainability of irrigation in the US High Plains and Central Valley. Proc. Natl. Acad. Sci. 109 (24), 9320–9325.

Schepers, J.S., Moravek, M.G., Alberts, E.E., Frank, K.D., 1991. Maize production impacts on groundwater quality. J. Environ. Qual. 20, 12–16.

School of Natural Resources, 2015. Water GIS data, Institute of Agriculture and Natural Resources. College of Arts and Sciences. University of Nebraska-Lincoln, Lincoln. Nebr. http://snr.unl.edu/data/geographygis/water.aspx last accessed on July 2, 2018.

Shapiro, C.A., Ferguson, R.B., Hergert, G.W., Wortmann, C.S., Walter, D.T., 2008.

Fertilizer suggestions for corn. University of Nebraska-Lincoln Extension educational programs. http://extensionpublications.unl.edu/assets/pdf/ec117.pdf.

Simon, C., Manzke, H., Kay, H., Mrowetz, G., 1964. Occurrence, pathogenesis, and possible prophylaxis of nitrite induced methemoglobinemia. Z. Kinderheilkunde 91, 124–138.

SNR-UNL, 2017. The School of Natural Resources at the University of Nebraska Lincoln. Test Hole Map Viewer. http://snr.unl.edu/data/geologysoils/NebraskaTestHole/NebraskaTestHoleMap.aspx last accessed on July 2, 2018.

Spalding, R.F., 1975. Effects of land use and river seepage on groundwater quality in Hall County. Nebraska: Nebraska Water Survey Paper No. 38, p. 95.

Spalding, R.F., Exner, M.E., 1993. Occurrence of nitrate in groundwater-a review. J. Environ. Qual. 22 (3), 392–402. https://doi.org/10.2134/jeq1993.00472425002200030002x.

Spalding, R.F., Gormly, J.R., Curtiss, B.H., Exner, M.E., 1978. Nonpoint nitrate contamination of ground water in Merrick County, Nebraska. Groundwater 16 (2), 86–95. https://doi.org/10.1111/j.1745-6584.1978.tb03207.x.

Spalding, R.F., Hirsh, A.J., 2012. Risk-managed approach for routing petroleum pipelines: keystone XL pipeline, Nebraska. Environ. Sci. Technol. 46, 12754–12758. https://doi.org/10.1021/es303238z.

Stanton, J.S., Lynne, F., 2006. Ground-water quality beneath irrigated cropland of the northern and southern High Plains aquifer. Nebraska and Texas. 2003–04. U.S. Geological Survey Scientific Investigations Report 2006–5196, 95 p.

Tiemann, M., 2017. Safe drinking water act (SDWA): a summary of the act and its major requirements. In: Congressional Research Service, . https://fas.org/sgp/crs/misc/RL31243.pdf.

U.S. EPA (Environmental Protection Agency), 2004. Understanding the Safe Drinking Water Act. U.S. Environmental Protection Agency, Washington, D.C. https://www.epa.gov/sites/production/files/2015-04/documents/epa816f04030.pdf.

U.S. EPA (U.S. Environmental Protection Agency), 2007. Nitrates and nitrites, TEACH Chemical Summary. U.S. Environmental Protection Agency, Washington, D.C. http://www.epa.gov/teach/chem_summ/Nitrates_summary.pdf.

U.S. EPA (U.S. Environmental Protection Agency), 2017. National Primary Drinking Water Regulations. U.S. Environmental Protection Agency, Washington, D.C. https://www.epa.gov/ground-water-and-drinking-water/national-primary-drinking-water-regulations#Inorganic.

U.S. Public Health Service, 1962. Public Health Service Drinking Water Standards. Publ. 956. U.S. Dept. of Health. Education and Welfare. Public Health Service, Washington, D.C, pp. 47–51. http://archive.org/details/gov.law.usphs.956.1962.

UNL (University of Nebraska-Lincoln), 2000. University of Nebraska-Lincoln. Quality-assessed Agrichemical Contaminant Database for Nebraska Ground Water. https://clearinghouse.nebraska.gov/Clearinghouse.aspx last accessed on July 2, 2018.

UNL (University of Nebraska-Lincoln), 2010. 2005 Land use mapping. Lincoln, Nebr. http://calmit.unl.edu/2005landuse/statewide.shtml last accessed on July 2, 2018.

USDA, 2014. 2012 CENSUS of Agriculture, Nebraska State and County Data. Geogr. Area Ser. Part 27 1 U.S. Depth of Agriculture and National Agricultural Statistics Service. Washington, D.C. https://www.agcensus.usda.gov/Publications/2012/Full_Report/Volume_1,_Chapter_1_State_Level/Nebraska/nev1.pdf.

USDA, 2015. Web soil survey, Natural Resources Conservation Service Soils. United States Department of Agriculture (USDA. https://websoilsurvey.sc.egov.usda.gov/App/WebSoilSurvey.aspx.

USDA-NASS, 2015. CropScape – cropland data layer, United States Department of Agriculture (USDA). In: National Agricultural Statistics Service (NASS), . https://nassgeodata.gmu.edu/CropScape/.

USGS, 2000. Water use in Nebraska, 2005. U.S. Geological Survey. Nebraska Water Science Center, Lincoln, Neb. http://ne.water.usgs.gov/infodata/wateruse/.

USGS, 2015. Moderate resolution imaging spectroradiometer (MODIS) irrigated agriculture dataset for the United States (MIrAD-US). U.S. Geological Survey. Early Warning and Environmental Monitoring Program. https://earlywarning.usgs.gov/USirrigation.

Van der Schans, M.L., Harter, T., Leijinse, A., Mathews, M.C., Meyer, R.D., 2009. Characterizing sources of nitrate leaching from an irrigated dairy farm in Merced County, California. J. Contam. Hydrol. 110, 9–21. https://doi.org/10.1016/j.jconhyd.2009.06.002.

Verstraeten, I.M., Ellis, M.J., 1994. Reconnaissance of ground-water quality in the Papio-Missouri river natural resources district, Eastern Nebraska, July through September 1992. USGS Water-Resour. Invest. Rep. 94–4197.

Verstraeten, I.M., McGuire, V.L., Heckman, K.L., 1998. Hydrogeology and subsurface nitrate in the Upper Big Blue natural resources district, Central Nebraska, July 1995 through September 1997. USGS Water-Resour. Invest. Rep. 98–4207.

Verstraeten, I.M., Sibray, S.S., Cannia, J.C., Tanner, D.Q., 1994. Reconnaissance of ground-water quality in the North Platte natural resources district, Western Nebraska, June–July 1991. USGS Water-Resour. Invest. Rep. 94–4057.

WSS (Web Soil Survey), 2017. The United States Department of Agriculture (USDA) and the Natural Resources Conservation Service (NRCS). https://websoilsurvey.sc.egov.usda.gov/App/HomePage.htm last accessed on July 2, 2018.

Yohannes, Y., Webb, P., 1999. Classification and Regression Trees, CARTTM: A User Manual for Identifying Indicators of Vulnerability to Famine and Chronic Food Insecurity. International Food Policy Research Institute0-89629-337-8 ISBN.

Zhang, L.B., Zhang, Q.Q., Xu, F., 2003. Research and application of the statistical models based on CART. J. Zhejiang Univ. Technol. 30 (4), 315–318.