Uncertainty Analysis and Visualization for Nitrogen Leaching with the Maize-N Model

Babak Samani[1], Saeideh Samani[1], Haishun Yang[2], Hongfeng Yu[1]
[1]Department of Computer Science and Engineering
[2]Department of Agronomy and Horticulture
University of Nebraska-Lincoln, Lincoln, NE, USA

2020 IEEE International Conference on Big Data (Big Data) | 978-1-7281-6251-5/20/$31.00 ©2020 IEEE | DOI: 10.1109/BigData50022.2020.9378105

***Abstract*—Nitrogen (N) is an essential nutrient for many crops including corn and soybean. However, its leaching into groundwater is a serious cause of concern for environmental and public health. The amount of N leaching is closely linked to soil water drainage and rainfall. Prediction of N leaching in cropping systems is critical to the improvement of crop management. Maize-N is a model for maize yield and N rate recommendation. However, uncertainties in many parameters, such as weather predictions, soil properties, and information entered by users (e.g., applied N fertilizer), can incur uncertainties in N leaching simulation results. We have developed a platform to assist comprehending the relationship between various input parameters and N leaching. Our platform can reveal N leaching with uncertainty analysis and visualization of different parameters.**

***Index Terms*—Uncertainty Analysis and Visualization, Multivariate Visualization, Nitrogen Leaching**

I. Introduction

Nitrogen (N) is an essential nutrient for many crops including corn, soybean, and so on. However, its leaching to groundwater is a cause for serious environmental and public health concerns. Amount of N leaching is closely linked to N fertilizer application, soil water drainage, and rainfall. Prediction of N leaching in crop systems is critical to the improvement of crop management and the reduction of N leaching. Visualization can help reduce uncertainty in prediction of N leaching in soil and water. Testing and validating N leaching predictions can help enhance the understanding of N leaching and soil health, and discover ways to improve fertilizer management and enhance environmental quality. The uncertainty in N leaching has originated from uncertainty in many parameters such as weather prediction, soil properties, and information entered by users (e.g., applied N fertilizer). Here, visualization plays a significant role to reveal the uncertainty for each parameter.

Recently, a web-based application is being developed to predict, in real-time, N leaching across the State of Nebraska in the United States (Nitrogen Leaching Calculator) [17] based on the Maize-N model [27]. To use the application, a user enters information including land characteristics, applied seed, and fertilizer. This application uses the historical and forecast weather data as well. Our objective is to validate the methods of visualizing the uncertainty, by monitoring and evaluating each component of the application. We are interested in addressing two questions using this web application: "How can visualization reveal that the farming data could be uncertain?" and "How we can quantify nitrogen leaching uncertainty?"

In this work, we measure and visualize uncertainty in N leaching with the Maize-N model. We have developed a platform to assist comprehending the relationship between various input parameters and N leaching. Our platform illustrates N leaching ensembles with analysis of uncertainty using different visualization methods.

II. Related Work

A. Uncertainty Visualization

Uncertainty visualization demonstrates the uncertain information of a dataset and becomes an important approach for scientists to understand computational sources and magnifiers of error and uncertainty in their datasets [18]. Different methods have been used in exploring uncertainties, such as quantitative visual interaction, parallel coordinates, sketching, projection-based approaches, scatter plots, incorporating uncertainty information into PCA projections, and k-means clustering [11], [12], [16], [26]. Research in the field of perception and awareness of uncertainty shows that the use of fuzziness can be a good visual variable for uncertainty. For example, Kumpf et al. considered the use of color for uncertainty, which can be used to demonstrate node stress and color labels [13]. Boxplot is another of the most common techniques to illustrate uncertainty [2], [15], [23].

MacEachren et al. [16] named some challenges in geospatial uncertainty visualization, expressing the difference between data quality and uncertainty. They suggested the benefits of hue, saturation, and intensity for representing uncertainty on maps. Ehlschlaeger et al. [8] illustrated how animation is useful to represent uncertainty on elevation data. Wittenbrink et al. [32] suggested involving multivariate glyphs when it comes to environmental flow visualizations. Hengl and Toomanian et al. [28] illustrated color mixing and pixel mixing techniques to visualize uncertainty in soil science research domain. Metaphors, such as glyphs, error-bars, and surface-coloring, can effectively show uncertainty with statistical estimates [6]. Several studies have recognized potential visual characteristics that can help understand uncertainty visualization. Davis and Keller et al. [7] illustrated that texture, color, and value are the features in uncertainty visualization on static maps. Rheingans and Joshi [21] illustrated the uncertainty in molecules postures. Some researchers compared various uncertainty visualization techniques and expressed their theoretical analysis [34], [35].

B. Ensemble Visualization

Ensemble visualization is generated from a series of simulations using different input parameters, and is a particular class of uncertainty visualization comprising a set of practical outputs. Numerical weather prediction ensembles are used for predictive weather forecasting. Simulation runs contain various initial conditions from diverse model parameterizations. Sanyal et al. [25] developed a software tool, Noodles, that enabled scientists to visualize the ensemble outputs and the uncertainty of weather data. To prevent dealing with complex topology and variation of ensemble isocontours, Bo et al. proposed a framework designed with a high density clustering method to increase the efficiency spaghetti plots in ensemble uncertainty visualization [4]. Phadke et al. used sequential animation with screen space subdivision and saturation tinting [19]. Bensma et al. proposed a technique to classify high-variance locations [3]. Ensemble visualization has been used to help users study a large amount of datasets through spatiotemporal analysis [10]. Potter et al. [20] presented an ensemble visualization framework named Ensemble-Vis to help exploring and generating visual results of weather data ensembles.

C. Machine Learning based Algorithms

Lakshminarayanan et al. [14] designed a method to calculate optimal decision tree using Mesh and Stochastic space convergences. This method estimated uncertainty using ensembles, where they believed that the method was better and well calibrated for measuring uncertainty compared to Bayesian Neural Networks. Sanakran et al. [24] introduced a technique to measure the impact of uncertainty in geometry in specific models. In their study, they discovered a good predictor bootstrap aggregated decision tree. Feinberg et al. [9] designed an open source tool for designing methods of measuring uncertainty, which is a Python software toolbox through polynomial chaos expansions and Monte Carlo simulation.

D. N Leaching

Crop producers face the challenge of how to apply sufficient soil nutrients to gain optimal plant growth and at the same time reduce nitrate losses to prevent contamination of ground and surface water. Therefore, it is critical to have an efficient agricultural management system. N leaching to water is caused by physical, chemical, and biological processes in soil and many other factors, such as crop type, soil organic matter levels, hydrology, temperature, drainage, and so on. Thus, it is very important to educate the public about the importance of soil nutrient management to resolve issues of N leaching across various agricultural landscapes [22].

The timing of applying fertilizer is critical in order to minimize N loss and increase fertilizer utilization efficiency. Variability in pre-fertilization conditions will also affect anticipated N leaching losses [30]. There is a positive correlation between applied N rate and its leaching. The most important part of agricultural management focuses on optimization of fertilizer application in terms of timing and amount of fertilizer [27]. Some applications have been developed to estimate N leaching in soil. The Maize-N model has been developed for simulating maize growth and yield in response to climate and nutrients [27]. The main inputs using the model include crop information, soil properties, and daily weather data.

III. Problem

Scientists and researchers have been trying to visualize uncertainty and error in data. Visualizing a discovered result and getting important information from that is not a new approach, and different methods have been used to visualize uncertainty and identify how different results are related together.

In this study, farmers need critical information to attain maximum yield from their Fields. Some of important information include weather data, soil data, as well as the amount of Nitrogen fertilizer needed to be applied during planting season. Some states in the United States, such as Nebraska, Iowa, and South Dakota, have many farms and agriculture is one of the major occupations. Therefore, agricultural management plays an important role in order to increase annual yields. On the one hand, N fertilizer is needed to apply to the fields. On the other hand, excessive amount of applied N can leach to the ground water and cause future problems to human and the environment.

Some methods have been defined in order to estimate the correct amount of N fertilizer needed by the crop. However, there are uncertainties in this prediction and we need to be aware of the factors causing these uncertainties. We have used a method to quantify and visualize uncertainty in the prediction of N leaching. We use the Maize-N model to calculate N leaching in a yearly and long term bases. Uncertain weather data, specifically rainfall and temperature, can lead to the wrong N fertilizer utilization. To prevent or at least reduce this mistake, we defined a method that identifies the N amount, its leaching to ground water, and the factors involved.

IV. Background

A. Maize-N Model

The Maize-N model is an existing robust crop simulation platform. This model has been developed to estimate the required amount of N fertilizer by maize crop. The inputs of this model include daily weather variables, planting date, crop maturity, grain price, population, yield history, cropping methods, and the time of fertilizer application. This model then predicts possible yield, N uptake from soil, N supplies from soil sources, and N fertilizer requirement. Its yield prediction is based on the Hybrid-Maize model [33]. The study also suggests that although climate and amount of available water are the key factors in estimating the attainable yields in different fields, management practices also influence the amount of N uptake. The prediction of N fertilizer is always uncertain because the weather data is not completely accurate and different types of N fertilizer have different efficiencies that causes different levels of losses [27].

The Maize-N model uses the following formula to calculate a required amount of fertilizer F from the Economically

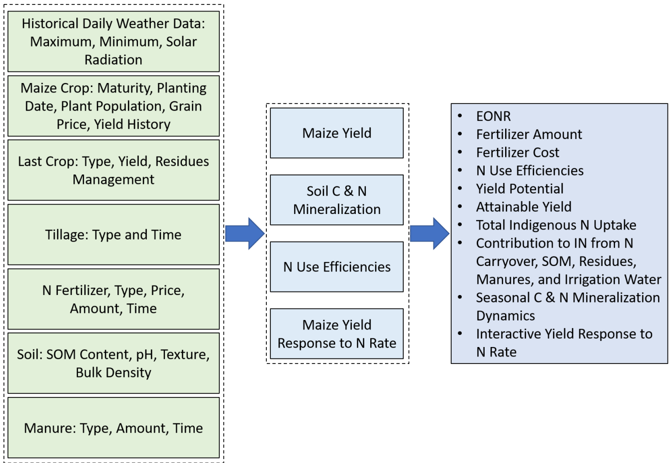


Fig. 1. The framework of the Maize-N model.

Optimum N Rate $EONR$, Recovery Efficiency $RE_{Fertilizer}$ of applied N (i.e., the fraction of N recovered in the crop from N applied), and fertilizer N content $N_{Content\_fertilizer}$.

$$F = \frac{\frac{EONR}{RE_{Fertilizer}}}{N_{Content\_fertilizer}} \tag{1}$$

The required inputs of the simulation include:

- Weather data: contains the weather data for the closest weather station for the intended point or filed;
- Maize crop: contains two options of irrigated and rained with relative maturity, date of planting, plant population, price of maize, and average yield of last 5 years;
- Last crop: contains type of crop, economic yield, total N applied, date of maturity, amount of crop residue left in field, and root-zone soil moisture at crop maturity;
- N fertilizer: contains N fertilizer already applied, type of fertilizer, fertilizer to be applied, N amount from irrigation water, and applied slow release N;
- Tillage: contains type of tillage, date of operation;
- Soil: contains top-soil organic mater content, bulk density, average root-zone texture, soil pH, and root zone depth;
- Measured root zone soil nitrate: contains amount and date of sampling;
- Manuring: contains type of manure, organic N content, inorganic N content, moister content, fresh weight, and date of application (actual of scheduled).

The Maize-N model simulates, on a daily basis, the dynamics of soil organic matter mineralization, crop N uptake demand, soil N balance, soil water balance, and N leaching beyond crop rooting depth. Figure 1 shows the framework of the model including inputs and outputs.

In our study, we use the Maize-N model to calculate N leaching with the required inputs. According to former studies summarized in related work section, variability in weather and the susceptibility of inorganic N from soil and fertilizer causes various N loss processes. The accuracy of N fertilizer requirements before planting nearly always have a degree of uncertainty. Maize-N is a useful tool for scenario analysis to evaluate the impact of different crop and soil management options, and the associated influences on N fertilizer efficiency in a specific cropping system defined by its environment, which includes soil type, climate, crop rotation, and agricultural systems.

B. Inverse-Distance Weighting Spatial Interpolation

Inverse Distance Weighting (IDW) [1] is a popular interpolation technique in spatial interpolation, as it is ease to use and straightforward to compute. The attribute value of an unknown point z_p is the weighted average of its known neighboring points. The calculated weights are related to the distances between the unknown point and the known points. To tune the diminishing strength while distance increases, IDW can be modified by a constant power q. The IDW interpolation can be expressed as:

$$z_p = \frac{\sum_{i=1}^{n}\left(\frac{z_i}{d_i^q}\right)}{\sum_{i=1}^{n}\left(\frac{1}{d_i^q}\right)} \tag{2}$$

where n is the number of known neighboring points, z_i is the attribute value of the ith known point, and d_i is the distance between the unknown point and the ith known point.

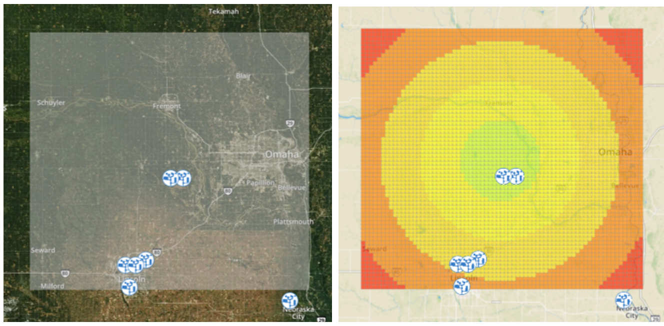
IDW weights the points closer to the intended point more than those further away. To use the IDW method, a certain points or all points in a certain radius from the intended point can be used to define the value for each location. The IDW method is a moving average interpolator that is usually applied to highly variable data. In some situations, it is possible to go back and record a new value if the result value is statistically different compare to values in that area. The IDW method has an assumption that points close to each other have similar values compare to those farther away. The IDW function is used when some locations are close enough to capture the extent of local surface variation needed for analysis. IDW determines cell values using a linear-weighted combination set of sample points. Some or all of the data points can be used in the interpolation process. The node value is calculated by averaging the weighted sum of all the points.

C. Weather Data

Weather data includes two type of daily data: historical weather data and forecast weather data. The historical weather data is obtained from High Plains Regional Climate Center (HPRCC) [31] weather data center and includes historical weather variables of maximum and minimum temperature, solar radiation, wind speed, relative humidity, ET (Evapotranspiration), and precipitation for about 200 weather stations in the high plains region. Based on the geographical coordinates of a location (i.e., latitude and longitude), the forecast weather data is retrieved using APIXU API [29], and stored on the database for the purpose of N leaching prediction.

D. Soil Texture Data

Soil texture data is one of the most important parameters in the calculation of N leaching prediction. To define soil texture



(a) (b)

Fig. 2. (a) Our research area. (b) N leaching uncertainty distribution in our research area.

in a filed, there are eight different categories including loamy sand, sandy loam, silt loam, loam, sandy clay loam, silt clay loam, clay loam, clay, and silt clay. As a field can be divided into several sections based on the soil texture polygons, to have an accurate visualization for N leaching the intersection of the soil texture polygons and field polygons needs to be stored in the database appropriately. Therefore, it is practical to store this dataset in a format of TopoJSON, GeoJSON, or ShapeFile. The primary source of soil data is obtained from USDA [5] that provides soil texture polygons in the ShapeFile format. The GeoJSON format facilities us to add more properties (such as clay, sand, and silt percentages of a soil) to each GeoJSON object.

V. Approach

The possible uncertainties that we need to consider include N leaching engine input uncertainty, engine uncertainty, weather data uncertainty, and soil texture data uncertainty. In this study, we only consider weather data uncertainty. We first explore the effect of weather data uncertainty on N leaching uncertainty for one weather station. Second, we employ the IDW interpolation method to get the interpolated weather data for calculating N leaching at a location. Third, we compute simulated N leaching data using weather data with uncertainty as an input parameter. Then, we compare the difference between simulation data and measured data. Finally, we visualize the spatial distribution of each parameter on the map. We can explore the relationship across the uncertainty in the input weather data, the output N leaching data and its uncertainty, and other parameters.

A. Effect of Weather Data Uncertainty on N Leaching Uncertainty for a Weather Station

At the first phase of our work, we explore the effect of weather data uncertainty on N leaching uncertainty. We choose one weather station located in Nebraska with 41.15, 96.48, 366 as latitude (deg), longitude (deg), and elevation (m). We take one year of weather data for this location and use the N leaching engine to calculate N leaching at this spot. We also calculate N leaching for all the locations with distance of 50 miles in radius from this location by assuming that there are uncertainties in precipitation in locations farther away from our weather station.

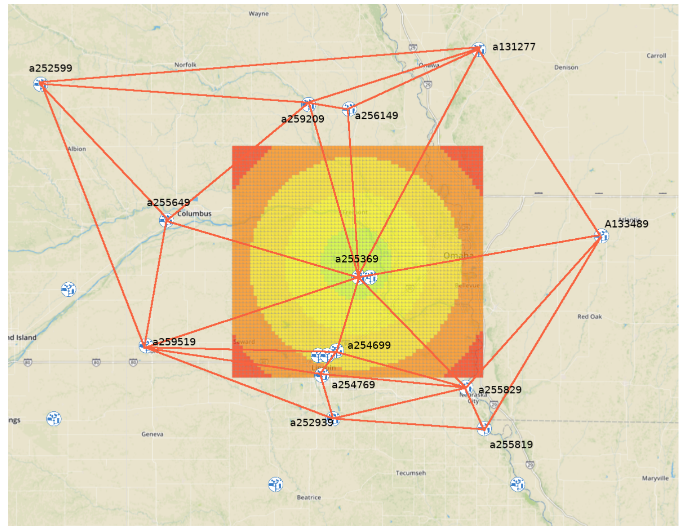


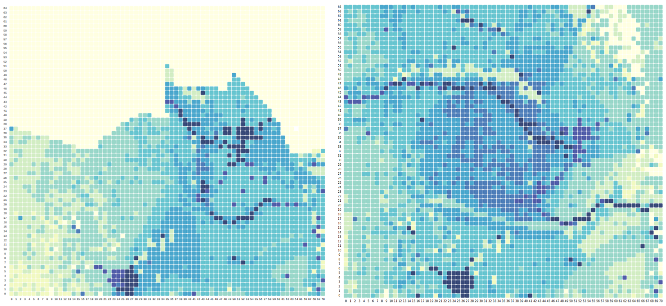
Fig. 3. Research area triangulation.

We assume that each 10 miles from that point the precipitation might increase or decrease based on different parameters such as elevation. By decreasing and increasing precipitation, we compute N leaching. Our purpose is to compute uncertainty in N leaching while there is uncertainty in precipitation at those locations. Primarily, We have a square with 50 by 50 miles in size with about 4200 grids that we study.

Figure 2(a) shows the research area, and Figure 2(b) visualizes N leaching uncertainties. Visualization shows that N leaching uncertainty increases if the distance from the weather station increases as we have uncertainty in those points weather data. Green color shows that uncertainty is almost zero at weather station point and when we depart from that point uncertainty increases and we see orange color when we have 50% uncertainty in weather data.

To make sure that N leaching calculator is accurate enough, we have used a ground truth data of soil moisture content from our previous study, and compared with the corresponding simulated results. The main points of these ground-truth data are as follow:

First, the experiment was conducted at two fields (Site 1 and Site 2) close to each other at Mead, Nebraska (30 miles from Lincoln, Nebraska). The crop planted at these sites is maize. Site 1 is irrigated while Site 2 is not irrigated (also called rainfed). Beside the difference in irrigation, Site 1 is continuous maize from 2001 to 2005, whereas Site 2 is maize and soybean rotation, meaning one year of maize, and next year for soybean. Thus, for Site 2 we have data of 2001, 2003 and 2005 as 2002 and 2004 was soybean and we do not use the data to calculate N leaching for soybean because of different plant types behavior [33]. N leaching from continuous corn averaged 24 mg, while that from the corn-soybean rotation averaged 42 mg. Total yearly nitrate leaching loss averaged 52 kg for continuous corn and 91 kg for the rotation. This represents the equivalent of 27% and 105% of the amount of N fertilizer applied over the



(a) (b)

Fig. 4. N leaching output (a) before and (b) after completing the data.

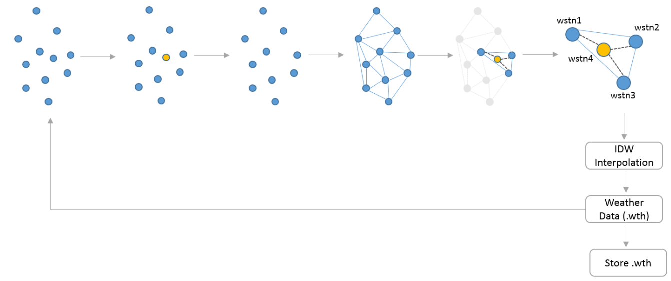


Fig. 5. Weather data estimation by the IDW method.

six years of study. Soybean in Nebraska receives no or very little N fertilizer. N leaching, however, can still be very strong, because part of the N leaching comes from soil organic matter mineralization, which is the same for corn crop and soybean. When calculating the percentage of N leaching, we can easily get a high value because of the small value of denominator.

Second, the soil moisture was monitored continuously in each field during the cropping season at four depths of 10, 25, 50, and 100 cm. The model output for soil moisture is for the layers of 0-30, 30-60, and >60 cm (60 -100). In order to match the depth of measurement and simulation, the mean of 10 and 25 cm for 0-30cm, the mean of 25 and 50 cm for 30-60 cm, and the mean of 50 and 100 cm for 60-100 cm was taken.

Third, we used the calculated total amount of water in the entire depth of 0-100 cm for both measured and simulated results.

Finally, for the uncertainty, we used the total amount of water in the 0-100 cm in depth. Uncertainty of the simulation comes from two sources: the error of rainfall in the weather data, and the error in the equations in the model. The weather data that we have has some missing data in some days. We first needed to fix the errors and fill those missing data. To do that we filled out the missing data using available data from the Automated Weather Data Network (AWDN) of HPRCC.

B. Interpolating Weather Data using IDW

Inverse Distance Weighting (IDW) predicts values at each point by averaging the amount of sample data points in the neighborhood of each processing point. The closer a point is to the location being predicted, the more weight it has in the averaging process. After visualizing N leaching in a specific point explained in previous section, we used interpolation methods to study how we can decrease uncertainties in weather data and visualize in a more accurate way. There were a number of weather stations in neighboring area of the studied location. We developed a python script to make triangles from every three weather stations, use the IDW interpolation method to get the interpolated weather data for each grid point, and then calculate N leaching in that area. Figure 3 shows how triangles are generated with the surrounding stations.

The result of this study at this point is shown in Figure 4(a). However, we did not have perfect weather data. There were some missing weather data or soil texture data that we needed to find and fill out in our database. After completing the data, we repeated this step and result is shown in Figure 4(b). The visualization illustrates that in the area that there are rivers or creeks there is almost 100% chance of N leaching. The darker color line means that there is a creek or river. Lighter colors show less leaching. According to our findings, if the area is with more rain and the soil texture type is silt clay loam, there are more chance of N leaching to the ground water compared to other soil textures.

C. Collecting More Data

With the precipitation from weather data, soil type, applied N, and the amount of irrigation at a certain location, we can get the value of possible N leaching using the Maize-N model. In order to obtain more accurate results and conclude our study, we need to have more data. To do this and based on different studies on IDW accuracy, we decided to use the IDW method and make data to feed in our platform. In our database, we have 420 weather stations that are located mostly in the Midwest of the U.S.. We studied 8 different sites in the states of Nebraska, Iowa, Kansas, and Colorado, and named these sites from 1 to 8. We chose 32 weather stations in these states.

We have developed another script to select one point at each time and use this point to create a triangle around this point. The steps of choosing the points and applying the IDW method have been shown in Figure 5. Given a set of weather stations (the blue points in Figure 5), we process the location of each station. For a weather station location (the yellow point in Figure 5), we choose three closest weather stations around this location point and create a triangle in our platform. After that, we use the triangular interpolation method to measure weather data (e.g., rainfall). The results of running this script give us 420 files (.wth files) of interpolated weather data.

As we need ground truth data to compare our results, we cannot randomly select points on the map and get interpolated weather data and calculate N leaching. This is because we do not have any way of proving our method and compute uncertainty on the weather data and at the end get N leaching uncertainty. To address this, we used the Maize-N engine to calculate the amount of N leaching. Figure 6 shows an example of the engine outputs. The outputs include total nitrogen in soil (green), nitrogen from soil organic matter (red), nitrogen

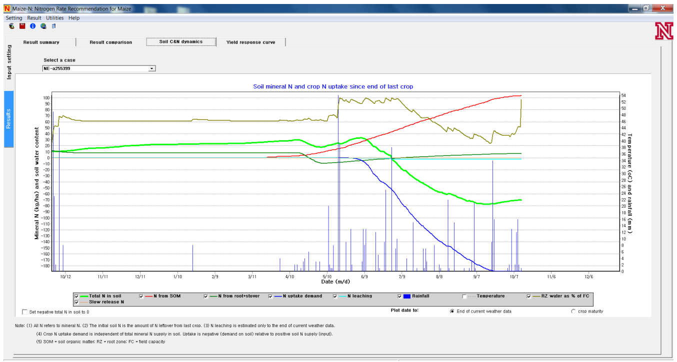


Fig. 6. Maize-N model soil mineral N and crop N uptake in a weather station location in Nebraska.

from root (dark green), nitrogen uptake by plant (light blue), and rainfall (dark blue) with mineral uptake and temperature as the *x* and *y* axes.

D. Quantifying Uncertainty Error

After having the option of calculating N leaching using the IDW method, we need to calculate the amount of uncertainty for this method for our selected points and get the error of the results. We select a weather station location as a study point. We choose 3 closest weather stations around this location point and creates a triangle in our platform. After that, we used the triangular interpolation method to measure rainfall. We know the accurate amount of rainfall for this point because this point has a weather stations and our goal is to measure the amount of estimated rainfall using the interpolation method and compare both data together. As this point is the weather station location, we assume the precipitation data at this location is accurate.

We use this data and calculate N leaching at this point with the Maize-N model. Next, we presume there is no weather station at this location and use the interpolation method to get precipitation data. Then, we calculate N leaching again. Last, we compute the error between these two N leaching results and get the amount of uncertainty.

We use the IDW interpolation method and Equation 3 to calculate the error:

$$error(\%) = \frac{(\mu_i - \mu_e) * 100}{\mu_e} \tag{3}$$

where μ_i is the result of N leaching from the interpolated method at point i and μ_e is the result of N leaching from the estimated data from weather stations. As shown in Figure 7, we also use RMSE to assess the error of uncertainty with Equation 4:

$$RMSE = \sqrt{\frac{1}{n}\Sigma(\mu_i - \mu_e)} \tag{4}$$

where n is the number of data points.

E. Uncertainty Visualization

We design a visualization approach with multiple components, as shown in Figure 8.

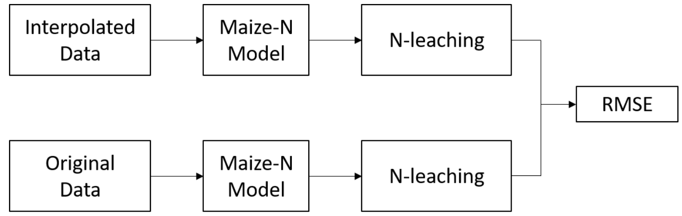


Fig. 7. The process to compute RMSE.

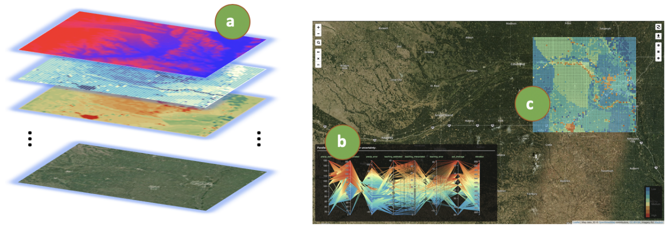


Fig. 8. Main components of our visualization approach.

First, we visualize each parameter as an individual layer that can be superimposed on the map layer. For each layer, our approach supports different visualization methods, such as heat map and contour map, to show the different properties of the corresponding parameters, as shown in Figure 8(a). For example, Figure 9 (a) illustrates an N leaching heat map that has been made using Python and QGIS. We also visualize soil drainage values through heat map in Figure 9(b).

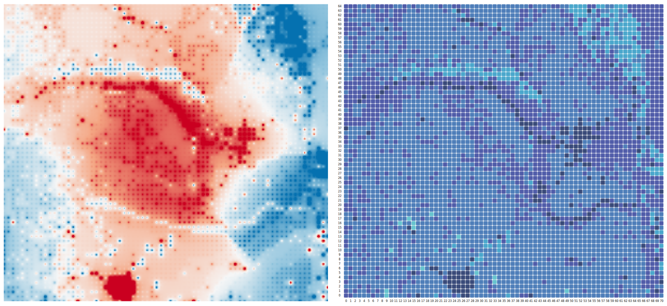
Second, we use parallel coordinates to visualize the relationship among the parameters, as shown in Figure 8(b). More importantly, a user can brush the parallel coordinates plot to select certain interesting portion of parameters and get more detailed relationship among them.

Third, apart from their relationships, the parameter ranges selected by a user can be also examined on the map, as shown in Figure 8(c). In this way, we cannot only see the possible relationship between values of different parameters, but also explore their spatial distributions. For example, a higher weather uncertainty value may cause a higher N leaching uncertainty for a location with a certain soil type; however, this would not be held for different soil types at other locations.

Through these main components, our visualization approach allows users to display different parameters with different visualization styles. A user can interactively brush parallel coordinates to examine the relationship among the parameters. The corresponding selected data ranges can be updated on the map. Our visualization can make it easy and intuitive for a user to explore the effects of parameter uncertainties across different geolocations.

VI. Results and Discussions

To experiment our system, we consider 80 sites from the database. Using the IDW method, we calculate N leaching using the N leaching engine. Therefore, we have 80 weather-stations with available weather data, 80 values of N leaching for these sites, and 80 interpolated values of weather and N leaching data.



(a) (b)

Fig. 9. (a) N leaching heat map. (b) Soil drainage values.

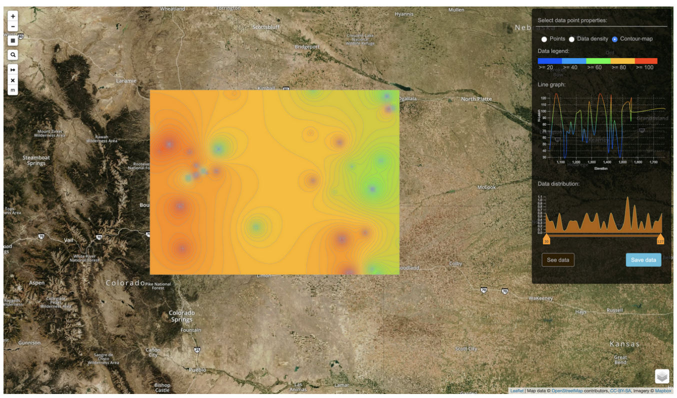


Fig. 10. Our platform showing N leaching and visualization parameters.

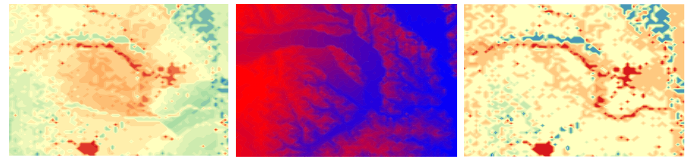
Since we have all of the required data to compare the results and calculate uncertainty, we use Equation 3 to compute the value of the error at each location and Equation 4 to calculate RMSE. The calculated results are shown in Table (I).

% of Error Range in Entire Study Area	
Error Range	**IDW**
-10% to 10%	11.25
-20% to 20%	10
-50% to 50%	38.75
else	40
RMSE	3.69

TABLE I
ROOT MEAN SQUARE ERROR RESULTS.

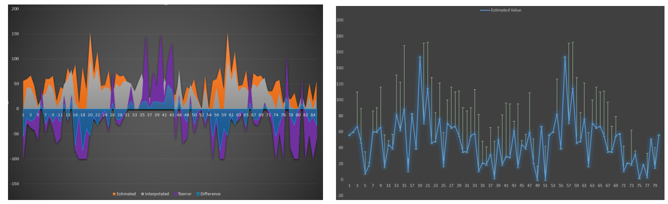
We compare our results from the interpolation method with the N leaching engine results to quantify uncertainty. The difference between the estimated data and the interpolated data gives us the uncertainty. We visualize this uncertainty for the region that we select and use in all of our visualization steps. We use different visualization methods including heat map and contour map. Our visualization approach can illustrate different parameters on top of the map. Figure 10 shows an example view of our platform for N leaching.

Figure 11(a) shows the contour map of N leaching visualized using our approach. We also visualize the elevation of the research area to study the effect of elevation in N leaching, as shown in Figure 11(b). Blue means lower elevation while



(a) (b) (c)

Fig. 11. (a) N leaching contour map. (b) Research area elevation. (c) Soil drainage classification contour map.



(a) (b)

Fig. 12. N leaching comparison chart (a) and uncertainty error bar chart (b) for different weather stations.

red means higher elevation. Based on this visualization, in our N leaching results we should see higher N leaching in red areas. Figure 11(c) classifies different soil types in the research location based on water drainage into the soil. Different soil type has different ability to hold or drain water. The drainage value has a direct effect on N leaching. This layer of visualization can be merged with other layers of data to create powerful results.

Figure 12(a) shows the comparison among the data that we gathered for estimated, interpolated, and error of N leaching. Figure 12(b) shows the estimated N leaching values with the error bars that we calculated from N leaching of the interpolated results. For both plots, the horizontal axis presents the weather stations and the vertical axis represents the N leaching values. As a result, the N leaching value varies in different locations with diffident uncertainties, which are dependant on different parameters including weather data, elevation, and soil data.

To get a better understanding of the relation between the parameters and the N leaching, we use parallel coordinates plots. We have implemented these plots in JavaScript. The vertical axes are corresponding to estimated precipitation, interpolated precipitation, estimated leaching, interpolated leaching, soil drainage value, precipitation error, N leaching error, and elevation data. Each vertical axis contains different values for an individual parameter.

The results show that there are different leaching values for different values of measured precipitation, and there might be different amount of leaching for interpolated precipitation values. By closely examining the plots, we can see this is because there are other parameters influencing these values. These parameters include soil drainage value and elevation. As shown in Figure 13, if we select the high leaching values, they are highly correlated to a specific soil drainage value and relatively lower elevation values.

Another observation that we can obtain is that the highest

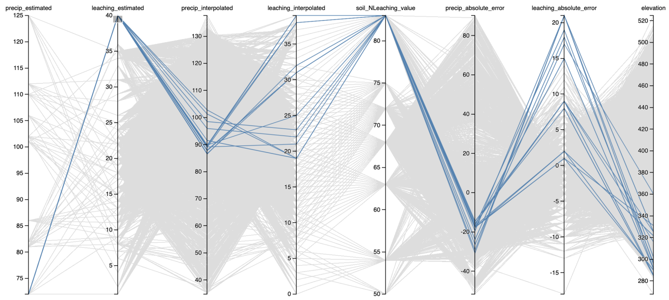
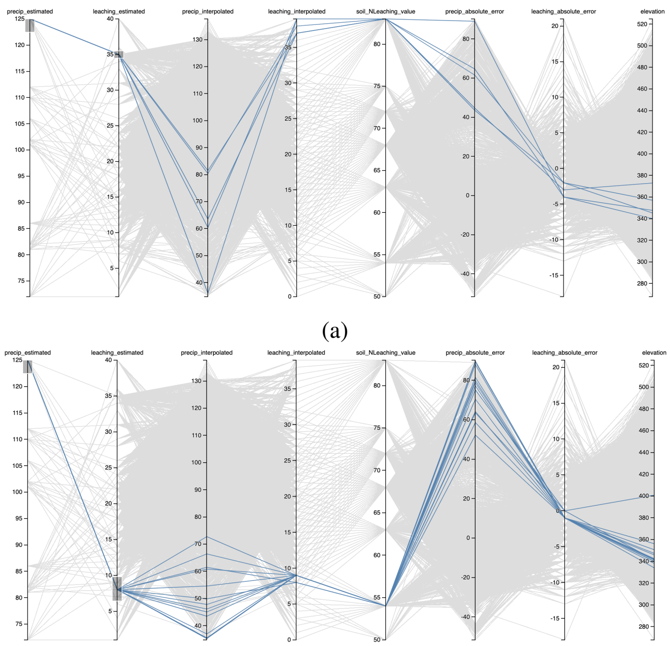


Fig. 13. Relationship between high N leaching values, soil drainage, and elevation.



(a)

(b)

Fig. 14. Relationship between highest level of precipitation and other parameters.

level of precipitation may not necessarily cause highest N leaching. As shown in Figure 14(a) and (b), the highest value of estimated precipitation that is 125 and results different N leaching values, where one value is 7 and is low. However, if we continue looking at the values of the plot parameters in Figure 14(b), we can see that the soil drainage is also very low for these low leaching amounts. Moreover, we can see that the interpolation values for precipitation that we calculated have high errors.

Figure 15 shows the relationship between the high N leaching errors and other parameters. We can see that they are highly correlated to lower elevation values and negative precipitation errors, and have associated with different estimated precipitation values.

Finally, we add a user interaction feature in our system for visualizing N leaching. On the map, a user can draw a polygon region, and a parallel coordinates plot will appear based on the

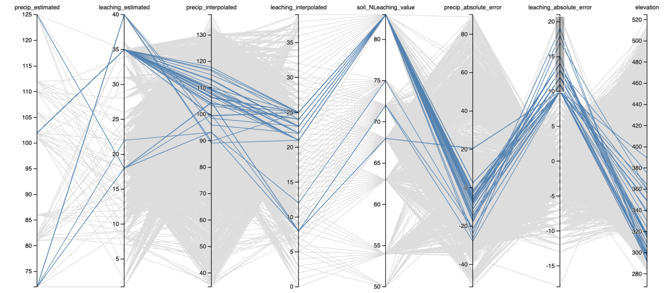
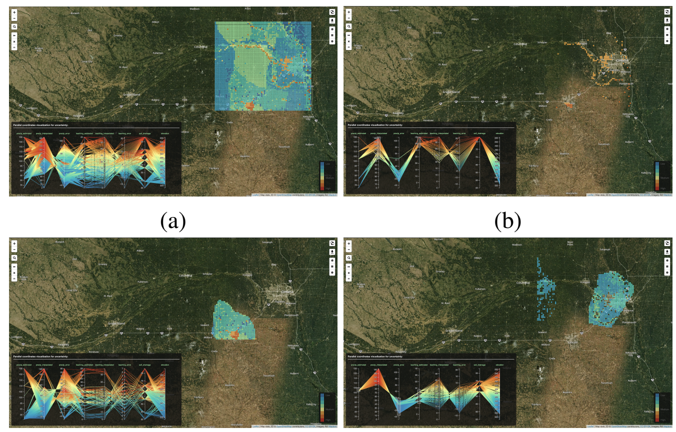


Fig. 15. Relationship between high N leaching error and other parameters.



(a) (b)

(c) (d)

Fig. 16. A parallel coordinates plot is displayed for a user selected region (a). A user can brush different axes of the parallel coordinates plot to explore their spatial distributions (b)-(d).

calculations of the parameters (i.e., N leaching, precipitation, error, elevation, etc.) in the selected region. Colors on a plot represent the values for each parameter (i.e., red for higher and blue for lower values). The user can brush different axes of the parallel coordinates plot, and the spatial distribution of the selected data ranges of parameters can be displayed on the map, as shown in Figure 16.

VII. Conclusion

It is important and helpful for scientists and farmers to understand N leaching to the soil and to ground water. We have developed a platform that demonstrates N leaching using different visualization methods. We have created a database for weather data within the regions including Nebraska, Iowa, Colorado, and Kansas. We use the Maize-N engine to calculate N leaching. Some input parameters are set according to the common ranges used by the farmers, such as the amount of applied fertilizer, the date of planting, and so on. We have selected 300 weather stations inside of these states and fed the weather data into our database. We have used the Maize-N model to calculate N leaching for year 2017 for all of these stations. As we do not have enough weather stations in those area that are close together, we can have uncertainty in weather data. Therefore, we gain uncertain N leaching values for the points far from weather stations locations.

To compute uncertainty at these locations, we have used the IDW interpolation method, and compared the estimated values and interpolated values to calculate the uncertainty. We assume that the result of N leaching in the location of a weather station is accurate because the precipitation is more accurate than other points. We also assume that the data for soil texture and the input data that we use to calculate N leaching is certain. We first calculate N leaching using these data, and then use the same data except for precipitation that we use interpolated value to calculate nitrogen leaching. Afterwards, we compute the differences between these two values to determine uncertainty. The results show that the uncertainty in locations that are not very close to weather stations varied based on distance of the studied points from other weather stations. The calculated RMSE for theses studied points is 3.69. We visualized the results of uncertainty to have a better understanding of N loss in different areas with different elevation, soil type, and weather datasets. Our visualization includes parallel coordinates plots with interactive operations on a map.

For future work, we would like to use Markov Chain Monte Carlo (MCMC) and Bayesian Inference to calculate uncertainty and compare those results with our current result. We plan to visualize uncertainty in various methods. We will compare them and expand this nitrogen leaching uncertainty visualization in more extensive area such as illustrating the leaching for all states in the United States.

Acknowledgment

This research has been sponsored in part by the National Science Foundation through grants IIS-1423487 and CNS-1739705, and by the Layman award of the University of Nebraska Foundation.

References

[1] P. M. Bartier and C. P. Keller. Multivariate interpolation to incorporate thematic surface data using inverse distance weighting (IDW). *Computers & Geosciences*, 22(7):795–799, 1996.

[2] R. Becker and W. Cleveland. Brushing scatterplots. *Technometrics*, 29(2):127–142, 1987.

[3] K. Bensema, L. Gosink, H. Obermaier, and K. I. Joy. Modality-driven classification and visualization of ensemble variance. *IEEE transactions on visualization and computer graphics*, (10):2289–2299, 2016.

[4] M. Bo and A. Montazeri. An interactive framework for visualization of weather forecast ensembles. *IEEE transactions on visualization and computer graphics*, (1):1091–1101, 2018.

[5] California Soil Resource Lab, UCDAVIS, https://casoilresource.lawr.ucdavis.edu.

[6] A. Cedilnik and A. Rheingans. Procedural annotation of uncertain information. In *Proc. Visualization 2000*, pp. 77–84, 2000.

[7] D. Davis and C. Keller. Modeling and visualizing multiple spatial uncertainties. *Computers & Geosciences*, 23(4):397–408, 1997.

[8] C. R. Ehlschlaeger, A. M. Shortridge, and M. F. Goodchild. Visualizing spatial data uncertainty using animation. *Computers & Geosciences*, (4):387–395, 1997.

[9] J. Feinberg and P. Langtangen. Chaospy: An open source tool for designing methods of uncertainty quantification. *Journal of Computational Science*, pp. 46–57, 2015.

[10] O. Harald and I. J. Kenneth. Future challenges for ensemble visualization. *IEEE Computer Graphics and Applications*, (3):8–14, 2014.

[11] C. Johnson. Top scientific visualization research problems. *IEEE Computer Graphics and Applications*, (4):13–17, 2004.

[12] C. Johnson and A. Sanderson. A next step: Visualizing errors and uncertainty. *IEEE Computer Graphics and Applications*, (5):6–10, 2003.

[13] A. Kumpf, B. Tost, M. Baumgart, M. Riemer, and M. Westermann, Rudiger Rautenhaus. Visualizing confidence in cluster-based ensemble weather forecast analyses. *IEEE transactions on visualization and computer graphics*, (1):109–119, 2018.

[14] B. Lakshminarayanan, A. Pritzel, and C. Blundell. Simple and scalable predictive uncertainty estimation using deep ensembles. *In Advances in Neural Information Processing Systems*, pp. 6402–6413, 2017.

[15] A. Love, D. Pang, and D. Kao. Visualizing spatial multivalue data. *IEEE CG & A*, 25(3):69–79, 2005.

[16] A. MacEachren, A. Robinson, S. Hopper, S. Gardner, R. Murray, M. Gahegan, and E. Hetzler. Visualizing geospatial information uncertainty: What we know and what we need to know. *Cartography and Geographic Information Science*, (3):139–160, 2005.

[17] Nitrogen Leaching Calculator Web Application, https://hprcc-agron1.unl.edu/.

[18] A. T. Pang, C. M. Wittenbrink, and S. K. Lodha. Approaches to uncertainty visualization. *The Visual Computer*, (8):370–390, 1997.

[19] M. N. Phadke, L. Pinto, O. Alabi, J. Harter, R. M. Taylor, X. Wu, H. Petersen, S. A. Bass, and C. G. Healey. Exploring ensemble visualization. *In Visualization and Data Analysis 2012, International Society for Optics and Photonics*, 8294:82940B, 2012.

[20] K. Potter, A. Wilson, P.-T. Bremer, D. Williams, C. Doutriaux, V. Pascucci, and C. R. Johnson. Ensemble-vis: A framework for the statistical visualization of ensemble data. *In Data Mining Workshops, ICDMW'09. IEEE International Conference.*, pp. 233–240, 2009.

[21] P. Rheingans and S. Joshi. Visualization of molecules with positional uncertainty. In *Proc. Data Visualization*, pp. 299–306, 1999.

[22] W. Riley, I. Ortiz-Monasterio, and P. Matson. Nitrogen leaching and soil nitrate, nitrite, and ammonium levels under irrigated wheat in northern mexico. *Nutrient Cycling in Agroecosystems 61*, (3):223–236, 2001.

[23] J. Roberts. State of the art: Coordinated and multiple views in exploratory visualization. *5th International Conference on Coordinated and Multiple Views in Exploratory Visualization*, pp. 61–71, 2007.

[24] S. Sankaran and A. L. Marsden. A stochastic collocation method for uncertainty quantification and propagation in cardiovascular simulations. *Journal of biomechanical engineering*, (3), 2011.

[25] J. Sanyal, S. Zhang, J. Dyer, A. Mercer, P. Amburn, and R. J. Moorhead. Noodles: A tool for visualization of numerical weather model ensemble uncertainty. *IEEE transactions on visualization and computer graphics*, 16(6):1421–1430, 2010.

[26] S. Sethuraman, L. Grady, and C. A. Taylor. Impact of geometric uncertainty on hemodynamic simulations using machine learning. *Computer Methods in Applied Mechanics and Engineering*, pp. 167–190, 2015.

[27] T. D. Setiyono, H. Yang, D. T. Walters, A. Dobermann, R. B. Fergusen, D. F. Roberts, D. J. Lyon, D. E. Clay, and K. G. Cassman. Maize-N: A decision tool for nitrogen management in maize. *Agronomy journal 103*, (4):1276–1283, 2011.

[28] H. Tomislav and N. Toomanian. Maps are not what they seem: representing uncertainty in soil-property maps. In *Proc. Accuracy*, pp. 805–813, 2006.

[29] United States Department of Agriculture, https://gdg.sc.egov.usda.gov.

[30] J. Van Groenigen, G. Velthof, K. J. Van Groenigen, O. Oenema, and C. V. Kessel. Towards an agronomic assessment of n2o emissions: a case study for arable crops. *European Journal of Soil Science 61*, (6):903–913, 2010.

[31] Weather forecast API, APIXU, https://www.apixu.com.

[32] C. M. Wittenbrink, A. T. Pang, and S. K. Lodha. Glyphs for visualizing uncertainty in vector fields. *IEEE transactions on Visualization and Computer Graphics*, (3):266–279, 1996.

[33] H. Yang, A. Dobermann, J. L. Lindquist, D. T. Walters, T. J. Arkebauer, and K. G. Cassman. Hybrid-maize: A maize simulation model that combines two crop modeling approaches. *Field Crops Research*, 87(2-3):131–154, 2004.

[34] F. Zhou, M. Chen, Z. Wang, F. Luo, X. Luo, W. Huang, Y. Chen, and Y. Zhao. A radviz-based visualization for understanding fuzzy clustering results. In *Proceedings of the 10th International Symposium on Visual Information Communication and Interaction*, pp. 9–15, 2017.

[35] T. Zuk and S. Carpendale. Theoretical analysis of uncertainty visualizations. In *Visualization and data analysis 2006*, vol. 6060, p. 606007. International Society for Optics and Photonics, 2006.