ERROR PROPAGATION AND UNCERTAINTY IN PREDICTING NONPOINT-SOURCE NITRATE CONTAMINATION IN GROUNDWATER

A thesis submitted to the faculty of

San Francisco State University

In partial fulfillment of

The requirement for

The degree

Master of Science

In

Geosciences

by

ZiZi Angelica Searles

San Francisco, California

May 2014

CERTIFICATION OF APPROVAL

I certify that I have read Error Propagation and Uncertainty in Predictions of Nonpoint-Source Nitrate Contamination in Groundwater by ZiZi Angelica Searles, and that in my opinion this work meets the criteria for approving a thesis submitted in partial fulfillment of the requirement for the degree Master of Science in Geosciences at San Francisco State University at San Francisco State University.

Jason J. Gurdak
Assistant Professor of Earth & Climate Sciences

Leonard Sklar
Associate Professor of Earth & Climate Sciences

Don Hodge at USEPA
Environmental Protection Specialist, Agricultural Program

ERROR PROPOGATION AND UNCERTAINTY IN PREDICTING NONPOINT-SOURCE NITRATE CONTAMINATION IN GROUNDWATER

ZiZi Angelica Searles
San Francisco, California
2014

Nitrate (NO_3^-) is a regulated chemical that often comes from nonpoint sources (NPS) and threatens human health in concentrations above 10 mg/L as Nitrogen (N). To protect groundwater quality from NPS NO_3^- contamination, resource managers need to understand the source, transport, and attenuation factors that control NO_3^- in groundwater and have tools that predict its occurrence and extent at the aquifer scale. Groundwater vulnerability models and maps are promising tools that incorporate knowledge of controlling factors and can be used to make predictions of NO_3^- impacted groundwater. To improve the utility and accuracy of such tools, this research will quantify the sources and propagation of errors in recently (2012) developed vulnerability models and maps for the Basin and Range (BR), Central Valley (CV), Coastal Lowlands (CL), and North Atlantic Coastal Plain (NACP) Principal Aquifers (PA) of the United States. The errors identified will be propagated through the PA vulnerability models using Latin Hypercube Sampling (LHS) and the resulting uncertainty in model predictions will be illustrated.

I certify that the abstract is a correct representation of the content of this thesis

Chair, Thesis Committee

5-15-2014
Date

PREFACE AND/OR ACKNOWLEGEMENT

I completed this thesis while working full time and having a child who is now 3 years old. Finishing this thesis is a great personal accomplishment for me and I am blessed to have made it. At this time in my life I am very happy I choose the Earth Sciences as an area of study because it is a subject matter full of wonder and never boring. My learning about the planet and its inner workings will continue beyond the completion of my studies.

My success is not my own doing. I had a lot of help along the way. I would especially like to thank the Creator and originator of all things for the many miracles when my ability to continue was in doubt. I am forever indebted to my love, Christian Bazinga, who always found ways to give me time so I could would on my thesis without interruption. My late boss Frederick Schauffler deserves much praise for allowing me to have an unorthodox work schedule that allowed me to attend my grad classes. Finally my advisor Jason Gurdak, thesis committee, and SFSU Earth and Climate Sciences faculty will forever be remembered by me for teaching me and giving me a strong foundation to excel in my career.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. Introduction

Nonpoint-source (NPS) nitrate (NO_3^-) contamination is the greatest contaminant of groundwater resources (Spalding and Exner, 1993; Corwin and Wagenet, 1996) and poses well-known human health and ecological risks (Fan and Steinberg, 1996; Galloway et al., 2003). Groundwater-quality monitoring can be used to characterize the controls and spatial patterns of NO_3^- in groundwater (Burow et al., 2010), but monitoring is impractical and cost-prohibitive across all scales consistent with management decisions. Therefore, vulnerability models that predict NPS NO_3^- in groundwater at the watershed and regional scale are often used to help inform management or policy decisions (Gurdak, 2008). Groundwater vulnerability models include the concepts of intrinsic susceptibility, proximity and characteristics of sources of naturally occurring and anthropogenic contamination factors that affect contaminant transport from land surface to the groundwater, and the in situ geochemical conditions (Gurdak and Qi, 2012).

Many recent advances have been made in predictive models of NPS NO_3^- contamination in groundwater (Gurdak, 2014). Yet, few modeling approaches have conceptualized or quantified uncertainty that is inherent to predictions of groundwater vulnerability to NPS contamination (Loague 1991; Loague et al. 1996; Gurdak et al., 2007). If results of predictive groundwater vulnerability models are to carry any weight in resource decision

making, the associated predictive uncertainty needs to be quantified (*Gurdak et al.*, 2007; Gurdak, 2008). Quantifying the inherent error propagation that contributes to uncertainty of model predictions helps users address whether actual NPS contaminants exceed background concentrations or regulatory thresholds. A common approach to groundwater vulnerability modeling involves the coupling of statistical models, such as multivariate logistic regression with Geographic Information Systems (GIS) (*Nolan*, 2001). This approach can be used to develop predictive models and maps of groundwater vulnerability to NPS NO_3^- contamination and to track error propagation and display the associated prediction uncertainty of model results (*Gurdak and Qi*, 2006).

The prediction error, or uncertainty, of the model results is a function of data error from GIS-based explanatory variables and model error of estimated logistic regression coefficients (Gurdak et al., 2007). GIS-based explanatory variables inherently introduce data error into logistic regression models because the GIS data are imperfect representations of reality. The source of error is generally a function of the accuracy and precision of the geospatial data (Mowrer and Congalton, 2000). Accuracy of geospatial data refers to the closeness of represented measurements or computations to their "true" or accepted values, and precision refers to the level of measurement and exactness of descriptions reported in the geospatial data (Gottsegen et al., 1999). The logistic regression coefficients have inherent estimation error (van Horssen et al., 2002). Thus, the errors from the explanatory variables and the model coefficients can propagate through the model calculations and result in spatially variable prediction uncertainty (Gurdak et al., 2007; Gurdak et al., 2009).

Groundwater vulnerability maps and estimates of prediction uncertainty allow users to make the best informed decisions regarding groundwater monitoring plans, best management practices (BMPs), or remediation strategies. For example, a vulnerability map may indicate that a certain area of an aquifer has a 60–80% probability of exceeding background concentrations of NO_3^-. If such a map is used to implement land use BMPs, estimate of the prediction uncertainty can help inform the user about the reliability and effectiveness of the specific BMP type and location in meeting a particular management goal. Furthermore, groundwater vulnerability maps and uncertainty estimates provide users important information about steps to reduce prediction uncertainty in future iterations of the models, such as expanding a groundwater monitoring network or collecting more accurate explanatory data used in the model. Ultimately, groundwater vulnerability maps and uncertainty estimates can help prioritize limited funding and optimize planning, policy, and mitigation objectives.

Gurdak and Qi (2012) used logistic regression coupled with GIS to create groundwater vulnerability models and maps that predict the probability of detecting NPS NO_3^- above background concentrations in recently recharged groundwater of 13 Principal Aquifers (PAs) in the United States (U.S.). Of those 13 PAs, only the model of the High Plains aquifer (Figure 1) has quantified error propagation and prediction uncertainty associated with the probability maps (*Gurdak and Qi*, 2006). Therefore, the primary objective of this thesis is to quantify error propagation and associated prediction uncertainty of additional PA vulnerability models reported by Gurdak and Qi (2012), including models for the Basin and

Range (BR), Central Valley (CV), Coastal Lowlands (CL), and North Atlantic Coastal Plain (NACP) aquifer systems (Figure 1). These four PAs were selected because they are all unconfined, alluvial aquifers, but are located across a range of climates and have explanatory variables typical of the 13 PA models (Gurdak and Qi, 2012). I hypothesize that the PA location (climate) and explanatory variables may have a significant effect on error propagation and the magnitude of prediction uncertainty in the NPS NO_3^- vulnerability models. Additionally, I hypothesize that the relative contribution of variance from the model and explanatory variables can be used to reduce prediction uncertainty in subsequent iterations of the vulnerability models and inform best management practices.

2. Study Areas

The study area includes the CV, BR, CL, and NACP PAs (Table 1and Figure 1). These PA's are all alluvial aquifers with unconfined surface hydrostratigraphic units that allow direct recharge from precipitation, irrigation, and other land-derived sources. These PAs were selected to represent a range of climate regimes; the CV has a Mediterranean climate, the BR a desert climate, the CL a humid subtropical temperate forest climate, and the NACP is a combination of continental humid temperate forest climate with a subtropical climate in the south and a continental climate in the north (Commission for Environmental Cooperation, 1997).

2.1. Central Valley Principal Aquifer

The CV PA in California (Figure 1) is an elongate structural trough bounded by mountains on all sides with the exception of the delta region that drains into the San Francisco Bay. The aquifer resides in an erosional valley filled with sediments derived from the Sierra-Nevada and Coast Ranges (Faunt, 2009). The CV has two distinct aquifer regions; the Sacramento Valley in the north and the San Joaquin Valley in the south. The Sacramento Valley is primarily unconfined and ranges in depth up to 300 m (Faunt, 2009). The San Joaquin Valley has an uppermost unconfined unit and a lower confined/semi-confined unit (Williamson and others, 1989). The 30–60 m thick Cocoran Clay aquitard separates the unconfined and confined systems (Page and Bertoldi, 1983; Farrar and Bertoldi 1988).

2.2. Basin and Range

The BR PA covers a large area of the Southwestern U.S., including most of Nevada, the eastern and southeastern desert regions of California, west and southern Arizona, and western Utah (Figure 1). Generally, the BR consists of several internally draining basins with no outlet to the ocean, with the exception of the Colorado River that that terminates at the Gulf of California. Aquifers in the BR can be alluvial basin-fill, fractured volcanic, and/or fractured carbonate systems (Planert and Williams, 1995).

2.3. Coastal Lowlands Aquifer System

The CL PA is located in the southern parts of Texas, Louisiana, Mississippi, and Alabama (Figure 1). The structural shape of the CL basin is a subsiding wedge that is thin on the northern margin and thickens towards the coast. The CL PA is comprised of sediments deposited by alluvial plain, deltaic, nearshore, and marine environments. The lithologic structure in any given area of the CL both laterally and vertically reflects past fluvial, lagoon, beach, or continental shelf and will have clay, silt and sand sedimentary profiles to reflect these environments (Ryder, 1996). The regressive and transgressive migration of the shoreline over time results in lithology that is complex, with stratigraphic contacts difficult to discern due to lateral facies change within units and the heterogeneous overlapping mixture of sand, silt, and clay, that is characteristic of deltaic depositional environments (Ryder, 1996). The CL PA has five distinct permeable zones and two confining units. The CL aquifers have both confined and unconfined units. Because of the dipping stratigraphy, the thickness and number of hydrostratigraphic units varies across the CL (Ryder, 1996).

2.4. North Atlantic Coastal Plain

The NACP PA covers parts of North Carolina, Virginia, West Virginia, Maryland, Pennsylvania, Delaware, New Jersey and New York's Long Island (Figure 1). It consists

of six vertically stacked aquifer systems separated by four confining units that are underlined by metamorphic and igneous units (Trapp and Horn, 1997). The NACP sediments reflect past depositional regimes of the Atlantic passive margin, including fluvial, deltaic, and marine transgressive environments. Similar to the CL, the structure of the NACP is wedge shape that thins near the base of the Appalachians and thickens towards the Atlantic (Trapp and Horn, 1997). The undulating, arch and trough, topography of the basal rock results in the hydrostratigraphic units of varying thickness throughout the PA. The upper most aquifer is the Surficial Aquifer that consists primarily of unconsolidated gravelly sand of Quaternary age. A confining clay unit separates the Surficial Aquifer from is the underlying sandy Chesapeake aquifer, which is laterally continuous in most areas of the aquifer with the exception of the southern half of North Carolina and parts of Maryland adjacent to the Chesapeake Bay (Trapp and Horn, 1997). Several additional confined units underlie the Chesapeake Bay formation.

3. Methods

In order to achieve the primary objective of quantifying error propagation and prediction uncertainty associated with the NO_3^- logistic regression models (henceforth to be referred to as model); I used the following four-step approach:

1) Quantify the error associated with explanatory variables used as model inputs.
2) Quantify the error associated with the model regression coefficient.
3) Quantify error propagation using a stochastic method called Latin Hypercube Sampling (LHS).

4) Develop uncertainty models that generate a distribution of NO_3^- prediction probabilities based on the propagation of error (step 3) associated with the explanatory variables (step 1) and regression coefficients (step 4).

The fundamental assumption of this approach is that the prediction uncertainty of the models is a function of the uncertainty in both the explanatory variables and regression coefficients that can be expressed as probability distribution functions (Gurdak et al., 2007). The following sections detail each of the four steps in my approach.

3.1. Explanatory Variable Error

To quantify the error associated with the explanatory variables, I first conducted a literature search to consolidate research on the accuracy of the 10 explanatory variables used in the models (dissolved oxygen, soil organic matter, hydric soil, irrigated land, farm fertilizer, seasonally high water table, soil clay, crops, anoxic redox, and well depth). Table 2 shows each of the explanatory variables used in the PA-specific vulnerability models, background concentration of NO_3^- for each PA, the model equation, and the model calibration and validation statistics. I used the results of the literature search to better understand the cause of error associated with the explanatory variables and to assign a reasonable error for each explanatory variable in the models. Table 2 summarizes the error values assigned to each of the 10 explanatory variables, which is described in the following sections.

3.1.1. Dissolved Oxygen

Dissolved oxygen (DO) concentrations in groundwater can vary spatially (laterally and with depth) and temporally within the saturated zone of an aquifer. DO concentrations in groundwater are often a function of the organic carbon concentration in the aquifer, proximity to recharge zones, basin geology, stratigraphy, and basin geomorphology (Rose and Long, 1988). In addition to information from the literature search, I analyzed the DO spatial trends with ESRI's Geostatistical Analyst - Trend Analysis tool (Esri[a], 2012) to gain insight into the controls on DO variability in each PA (Figures 2–21). Among the 87 explanatory variables tested by Gurdak and Qi (2012), DO was statistically significant in more PAs models than any other explanatory variable, which indicates the widespread importance of DO in controlling groundwater vulnerability to NO_3^-.

Before assigning error values to DO, I evaluated the trend plots to determine patterns of spatial variability in the DO concentrations (Figures 2–5). The trend plots indicate that in most PAs, DO concentrations are greater in recharge or higher elevation areas and lower at the terminus or mid-basin. For the purpose of this research, I assumed that the spatial distribution of DO concentrations has been relatively stable during the period of groundwater sampling (1992–2008) (Gurdak and Qi, 2012).

To quantify the error associated with the distribution of DO across each of the four PAs, I used the Empirical Bayesian Kriging (EBK) interpolation method (ESRI's Geostatistical Analyst, 2012). A prerequisite for any kriging method is the presence of spatial

dependency or spatial autocorrelation (Esri[b], 2012). Spatial autocorrelation refers to a statistical relationship between the data value and distance and direction (Esri[b], 2012). Spatial autocorrelation of the DO datasets was observed in the 3D Analysis Trend Plots (Figures 6–12) and the empirical semivariograms in all four PAs. Of the four PA's, the CV has the strongest relative autocorrelation and the BR the weakest. NACL and CL DO datasets are both moderately correlated.

I used the EBK interpolation method because it minimizes variogram modeling uncertainty through the automated quantification of key variogram parameters and variogram error (Krivoruchko, 2012). The EBK process works by generating an initial variogram with its standard prediction error and using this error to generate a second semivariogram (Krivoruchko, 2012). The result of this second semivariogram is assigned a weight (Baye's rule) that accounts for the probability of a semivariogram to give true positive results (Krivoruchko, 2012). Through an iterative process successive variograms are generated with the end result being a range of semivariograms and accompanied by a range values for nugget, slope, and power (Krivoruchko, 2012).

An additional benefit of the EBK method is that the variogram fitting process is automated and performed iteratively using a range of variograms to achieve a best-fit model. I assume that the maps resulting from the EBK interpolation are good approximations of reality to the extent that the a priori assumptions are true. Additionally, I assume that the EBK generated prediction and standard error DO maps are built using datasets

that are representative of the behavior of DO concentrations throughout the aquifer. I also assume that DO concentrations do not vary by much overtime.

The prediction and standard error EBK maps are shown in Figures 14–20. The maps show five ranges of standard error across the aquifer from least error to greatest error. For purposes of the uncertainty modeling (described in section 3.2 Quantifying Uncertainty Using Latin Hypercube Sampling) the 2nd to lowest and 2nd to highest values were selected to represent maximum and minimum DO values. The highest and lowest values were omitted because they represent extreme tails of the distribution and the goal is to assign uncertainty model error based on error values that are reasonable to apply across a majority of a PA. Minimum and maximum DO error values used to inform the uncertainty model can be found in Figures 13–20 and Table 3.

3.1.2. Soil Organic Matter

The explanatory variable soil organic matter (SOM) used in the NO_3^- logistic regression models was extracted from the State Soil Geographic (STATSGO) dataset. STATSGO is a national soil database that aggregates regional soil types and was created by the U.S. Department of Agriculture's (USDA) Natural Resource Conservation Service (NRCS) for public, private, and academic use (USDA, 1994). A complementary database to STATSGO is SSURGO, which is a soils database with a higher resolution of data coverage, is also maintained by the USDA's NCRCS. SSURGO is primarily designed for decision making at the farm/ranch, township, parish, and county level (USDA, 1994),

whereas STATSGO is more appropriate for the regional scale consistent with the four PAs of this study. SSURGO was primarily compiled from field surveys and aerial photo interpretation, whereas STATSGO is a generalization of SSURGO data that is paired with Land Remote Sensing Satellite (LANDSAT) images and other data sources that provide information on topography, geology, vegetation, and climate (USDA, 1994). STATSGO maps were designed for exploring regional soil characteristics at the basin, state, or multistate level (USDA, 1994).

Despite SSURGO's higher resolution, STATSGO has been more popular among researchers as an input for soil organic matter (SOM) because STATSGO requires less preprocessing compared to SSURGO. Currently (2014), SSURGO is of limited use except at landscape scales due to the fragmentation of the data (Zhong and Xu, 2011). While STATSGO is easier to use, SSURGO is the more accurate database when field data is compared to the two databases (Zhong and Xu, 2011). The accuracy of STATSGO SOM data is questionable in many areas compared to SSURGO due to the omission of SOM values (represented as "null") and the use of zero as a valid value for SOM and depths below the surface. Both STATSGO and SSURGO contain a minimum and maximum value to represent SOM. Typically, researchers average the minimum and maximum values to obtain a percentage that is representative of SOM vertically and horizontally over an area (Zhong and Xu, 2011; Amichev and Glabraith, 2003; Homann et al. (1998)).

Null values and SOM fields reported as zero can result in an inaccurate representative SOM estimate for an area of interest. For example, Amichev and Glabraith (2003)

investigated the accuracy of STATSGO SOM values for Maine and Minnesota and revealed that the SOM maximum and minimum fields for Maine contained zero values in 24% and 54% of the records, respectively. For Minnesota, the maximum and minimum zero values comprised 0.2% and 0.4% of the STATSGO dataset, respectively (Amichev and Glabraith, 2003). Zhong and Xu (2011) investigated the rate of SOM null values in STATSGO for Louisiana and found that 75% (261 of 347) of the map units had null values for SOM categories.

SSURGO should not be used validate the accuracy of STATSGO (Zhou and Xu (2011). Zhou and XU (2011) tested SSURGO's viability as a SOM validation tool by comparing actual field measurements of SOM in Louisiana soils (Brupbacher et al. 1973). Field measurements and SSURGO SOM values were well correlated (n=86, R^2=0.635) (Zhou and Xu, 2011), however field measurements and STATSGO SOM values were not well correlated (n=336, R^2=0.013) (Zhou and Xu, 2011). SSURGO's positive correlation with field data indicates the dataset better represents the quantity of SOM in nature compared to STATSGO.. Additionally, Zhou and Xu (2011) compared the 86 SSURGO and STATSGO map units without missing values and found that the agreement between STATSGO and SSURGO SOM values decreases with depth. When all 336 Louisiana SSURGO and STATSGO map units were compared, Zhong and Xu (2011) found that STATSGO underestimated SOM by 9% at the 20 cm level and 36% at depths of 100 cm or more.

The tendency of STATSGO to consistently underestimate SOM when compared to SSURGO was also observed by Davidson and Lefebvre (1993) who found that STATSGO underestimated SOC by 13% for areas in Maine when compared to SSURGO data. Homann et al. (1998) compared the 1994 NRCS soil dataset, comprised of published and unpublished soil surveys, to STATSGO SOM values for Western Oregon. Homann et al. (1998) found that Western Oregon NRCS pedons at the 20 cm level show an SOM average of 68 t ha-1 whereas STATSGO averages 51 t ha-1 for the region, a 25% difference.

To estimate the error associated with the STATSGO SOM for the 4 PAs, I compiled all previously described error values in Table 4 All studies report that STATSGO underestimates the true value, with errors ranging from -4.7 to -36% and average -18.5%, rounded to -20%. Therefore, I used -20% to represent SOM error in the subsequent uncertainty modeling for all 4 PAs.

3.1.3. Hydric Soil

The USDA defines hydric soils as soils that have undergone saturation or flooding for a time sufficient to allow for the development of anaerobic conditions that favor the growth and regeneration of hydrophytic vegetation (USDA, 1985). Accurately identifying the extent and location of hydric soils is important because the anerobic, organic rich hydric zones can promote denitrification and reduced NO_3^- concentrations in groundwater. Hydric soils are frequently characterized by a localized high water table.

Rosenblatt et al, (1996) tested the accuracy of SSURGO and STATSGO datasets to represent riparian areas with hydric soils. To compare field data and SSURGO for hydric soil coverage, Rosenblatt et al, (1996) randomly selected 100 streams in the Rhode Island Pawcatuck Watershed for groundtruthing. At each location, area samples were taken along a 30-m transect perpendicular to the steams. The data collected from these transects were compared to SSURGO maps. SSURGO correctly classified 73% of the 100 stream sites as having hydric soils.

Rosenblatt et al, (1996) additionally defined hydric soils as areas that occupy an area greater than 10 m in width and the absence of groundwater seeps, as seeps indicate a compromised anaerobic environment. SSURGO map results estimated that 37.2% of the hydric soil within 15 m (one half of the 30 m transect) of a stream is hydric and therefore capable of reducing NO_3^- to $N_2(g)$. A STATSGO map of the same area only showed 2% of soil near steams has denitrification potential.

For purposes of assigning error to the STATSGO field representing hydric soil, I assume that SSURGO is a relatively accurate proxy for field measurements. Based on the findings of Rosenblatt et al, (1996), I assume that STATSGO underestimates SSURGO by 37.2% thus yielding an error of 37% in the 4 PAs.

3.1.4. Irrigated Land

Irrigated and non-irrigated lands explanatory variable data was derived from an irrigated lands map of the U.S. created by Pervez and Brown (2010) using 2002 Moderate Resolution Imaging Spectroradiometer (MODIS) satellite imagery and the USDA 2002 Census of Agriculture (COA) county-level irrigated acreage estimates (Pervez and Brown, 2010). To create the MODIS Irrigated Agriculture (MIrAD-US) map, peak Normalized Difference Vegetation Index (NDVI) was identified for different crops. Available moisture in the form of precipitation or supplied water has been found to increases the NDVI in many types of vegetation (Kawabata et al., 2001; Wulder et al., 2004). Thus, irrigated crops tend to have higher peak NDVI values than non-irrigated crops (Ji and Peters, 2003; Wang et al., 2003). The COA's reported irrigated lands acreage acts as a constraint on the amount of peak NDVI cells to be designated as irrigated lands during the MIrAD-US map creation (Pervez and Brown, 2010). Visual comparison of the final map with satellite imagery was also used to adjust MIrAD-US accuracy in representing irrigated land coverage (Pervez and Brown, 2010). The final MIrAD-US map product consists of a 250-m cell national map, or land mask, depicting irrigated and non-irrigated areas (Pervez and Brown, 2010).

Pervez and Brown (2010) calculated two types of error data associated with MIrAD-US; agreement between COA irrigated acreage and MIrAD-US, and agreement between MIrAD and two regional validation datasets. Generally, Pervez and Brown (2010) found good agreement (92%) between COA data and the MIrAD results for western and

mid-western states. There was less agreement (75%) between MIrAD-US and the COA in the eastern U.S., which was attributed to humid climate vegetation dynamics, small county size (a resolution processing issue), and a smaller quantity of irrigated acreage compared to the West (Pervez and Brown, 2010).

Validation of irrigated and non-irrigated map results (Pervez and Brown, 2010) was done by comparing the MIrAD mapped imagery with California Department of Water Resources (CA DWR) agricultural surveys (CA DWR, 2000–2004) and the University of North Dakota (UND) Great Plains field survey cataloging irrigated versus non-irrigated lands (Seelan and Kurz, 2006). A map validation was not conducted for the East Coast because of the lack of a comparable field survey dataset.

The California DWR and MirAD-US, UND and MIrAD validation compares agricultural maps to MIrAD to see how well MIrAD-US matches that have been groundtruthed. The California DWR and MIrAD-US comparison error matrix shows a user's accuracy of 86% for irrigated lands (Pervez and Brown, 2010). The user's accuracy refers to the probability that a pixel labeled as irrigated land on the map is accurately identified if further groundtruthing were to take place. In other words, if an irrigated lands pixel were selected and visited in the field there is an 86% probability that the pixel has been properly classified. The MIrAD-US and the UND error matrix revealed a user's accuracy of 75% irrigated lands explanatory variable although Pervez and Brown (2010) suggest that the lower accuracy associated with this dataset is likely a function the smaller number of pixels compared.

Irrigated lands are only an explanatory variable for the BR PA (Table 2) (Gurdak and Qi, 2012). To obtain an error value for the irrigated lands variable in the BR uncertainty model, I used the average (11%) (Table 4) of the previously described Western U.S. (8%) and the California DWR studies (14%).

3.1.5. Farm Fertilizer

The Association of American Plant Food Control (AAPFCO) and the USDA have national datasets that track the application of nitrogen rich fertilizer to agricultural and other lands. AAPFCO operates at the University of Kentucky and is the central repository for all state data on farm and non-farm fertilizer sales. Reports on state fertilizer sales are published annually. The USDA tracks data on fertilizer use and fertilizer sales through COA census tracking once every five years. COA fertilizer data is available at the county level.

Unlike other explanatory variables that are based on direct estimates, the application of fertilizer is estimated indirectly by tracking state fertilizer sales or from the USDA COA. There are a number of errors associated with both the AAPFCO and COA tracking methods. When states report sales to AAPFCO, approximately half of the states report separate their reporting of fertilizer sales to farm and non-farm entities and the other half don't separate and report all sales data as a sum without a categorical breakdown indicating sales to farms. The non-standardization in state reporting requirements introduces errors

insofar as it is unknown for half the states the amount of fertilizer than can be attributed to farm operations (Gaither et al., 2004).

AAPFCO reporting states collect data from county level estimates and aggregate them at the state level. Therefore, fertilizer sales may not necessarily reflect the actual location of the fertilizer application. Additionally, not all counties report complete information on an annual basis. In instances where a state submits incomplete data, the AAPFCO will use sales data from the previous year and any monthly state sales data that is available (Gaither et al., 2004). The AAPFCO has not published any error estimates for the annual "Commercial Fertilizer" report.

Another source of farm fertilizer data is the COA. Every five years the USDA queries farms that have over $1,000 in sales to report several categories of information, including fertilizer sales. For each COA year, the USDA allocates AAPFCO data by county. The AAPFCO reports sales in tons by the chemical composition of the fertilizer type (i.e. N, P_2O_5, K_2O). The COA assumes that all fertilizer sold in a year is applied in that same year (USDA, 2004).

Ruddy et al. (2006) created U.S maps depicting annual nutrient (farm and non-farm) by region. Nitrogen inputs from fertilizer and manure were allocated to appropriate 1992 Enhanced National Land Cover Data classes within each county (Ruddy et al., 2006). The following is a summary of fertilizer application by PA (Ruddy et al., 2006):

- Central Valley - For most areas in the Central Valley, 6,000–8,000 kilograms or greater than 8,000 kilograms of fertilizer is applied per square mile.

- Basin and Range - With the exception of parts of Southern Arizona and the Coachella Valley/ Salton Sea Trough, the Basin and Range has relative low fertilizer application rates. In southern Arizona and the Coachella Valley/ Salton Sea Trough, the application rates exceed 8,000-kg per square mile annually.
- Coastal Lowlands - The fertilizer application rates vary considerably by county. Application rates vary between 2,000–8,000 kg per square mile annually.
- North Atlantic Coastal Plain - The fertilizer application rates are high at 6,000–8,000 kg per square mile or greater applied annually.

Since there are no published values quantifying the error associated with the farm fertilizer dataset, I assume that 20% (Table 4) is a reasonable error percentage to inform the uncertainty models.

3.1.6. Seasonally High Water Table

Seasonally high water table (HWT) is defined as a water table that is within 1m of land surface (USDA, 1985). In the STATSGO dataset, HWT is reported as a value in feet representative of the depth of water below land surface. Often the same value will be repeated for different areas of the PA. There are no reported error values for HWT. Therefore, I assume a relatively low error of 5% (Table 3) because HWT is informed by observation than measurement. An area either has a HWT or it does not.

3.1.7. Soil Clay

Bricklemyer et. al (2007) compared STATSGO soil clay % in soil to field data collected at five tillage and no tillage farm sites throughout the Great Plains. The purpose of fielding checking STATSGO data was to test how well STATSGO performed as an input dataset for the Century Model, a model that measures changes to soil carbon. Bricklemyer et al., (2007) found that STATSGO under-reported clay % at the five sites by 28%. This underreporting of soil clay was only identified for the farm sites. Other environments, including non-farm sites were not included as part of the study.

Some other researchers have attempted to identify a relationship between the presence of soil clay and organic carbon as a way for predicting the presence of soil clay. The research to date is not conclusive in demonstrating that such a relationship exists. For example, Davidson (1995) found that organic carbon is positively correlated with clay content in parts of Kansas, but organic carbon is not correlated with clay in Montana. Rasmussen (2006) looked at the correlation between soil clay and organic carbon by biome in the desert environments of Arizona and found that soil clay and organic carbon were positively correlated for all desert biomes studied, with some biomes having a higher carbon and clay content correlation than others.

Given the lack of consistent results in correlating clay with organic matter, I assume 20% (Table 4) as a reasonable error value for clay % in soil as input to the uncertainty models. While Bricklemeyer et al. (2007) does quantify the error for soil clay in STATSGO, the focus on farm sites raised questions about the applicability of this error

quantification across PA's. Farm site soils are disturbed areas, especially tillage sites. In their study, Bricklemeyer et al. (2007) did not control for the fact that these farm sites may have higher clay content due to cultivation practices.

3.1.8. Crops

Crop cover data used in the logistic regression models is from the 2001 National Land Cover Dataset 2001 (NLCD) (Gurdak and Qi, 2012). NLCD 2001 is a 16-class land cover classification scheme, sometimes referred to as Anderson Level 1 (Anderson, 1976) that has been applied consistently across the U.S. at a 30-m resolution. NLCD 2001 relies on the analysis LANDSAT imagery for land-use classification (Homer et al., 2007). Wickham et al., (2001) analyzed the classification of Anderson Level 1 accuracy and found an overall user's accuracy of 85% for Level I categories. Nationwide the Cropland user's accuracy was found to be 82% (Wickham et al., 2001).

Maxwell and Janus (2008) compared the accuracy of 2001 NLCD cultivated cropland Anderson class to 2002 USDA Census data for the upper mid-western U.S. NLCD crop coverage estimates for the entire study area were 1.8% less than COA estimates. While the percent difference between the region values and the NLCD was relatively small at 1.8%, the percent difference at the state level varied considerably. NCLD state-by-state variability was much more pronounced. Of the 14 states in the comparison, the percent difference variability for the NLCD classification of cropland ranged from -37% to 18%.

Areas with 5.0% or more difference between the NLCD and USDA Census values of crop cover share several characteristics. In counties where a cropland area is less than 20% there is less agreement between NLCD and USDA. The presence of grassland tended to obscure crops resulting in NLCD cropland overestimates. Forest dominated landscapes could cause either the under or over estimation of crops. Areas where irrigation is not used for crops are also prone to error especially given the NDVI similarity to grasses and pasture. Regions with small size farms are also more prone to the NDVI reporting error.

For purposes of modeling uncertainty, I assume a 20% error for crop % (Table 4) based on the Wickham et al., (2001) study that reported a user's accuracy of 82% (18% error) for U.S. cropland. The 20% error likely capture factors that cause misrepresentation of the crop category, including proximity to forested land, pastures, and grasslands, humid climate where crops are not irrigated, and small farm areas.

3.1.9. Anoxic Redox

In the logistic regression models (Gurdak and Qi, 2012), the anoxic redox (reduction-oxidation) explanatory variable is represented as a binary value (1 or 0) and based on a redox classification system created by McMahon and Chapelle (2008). The redox classification system is based on concentrations of DO, NO_3^-, manganese, iron, sulfate, and sulfide, and generally categorizes anoxic conditions as water with DO less than 0.5 mg/L (McMahon and Chapelle, 2008). McMahon and Chapelle's (2008) method assumes that

denitrification predominantly occurs below DO concentrations of 0.5 mg/L, but acknowledges exceptions to the rule, such as studies reporting denitrification in water with DO concentration as high as 2 mg/L (Böhlke et al. 2002, 2007; McMahon et al. 2004).

Since anoxic redox is a binary explanatory variable, I assumed that all reported values are accurate and assigned no error value to the explanatory variable (Table 4). If error were to be quantified for this variable it would need to be done within the assumptions of the redox classification system (McMahon and Chapelle, 2008). Knowledge of the microorganisms that catalyze redox processes would also be useful in fully quantifying the error associated with the McMahon and Chapelle (2008) method.

3.1.10. Well Depth

The explanatory variable well depth is assumed to have little or no error associated with the NAQWA and NWIS value used as input to the logic regression models (Gurdak and Qi, 2012). Groundwater Technical Procedures of the U.S. Geological Survey (Cunnignham and Schalk, 2011) require the use of calibrated steel tape to measure well depth. Measurements of well depth are to be repeated several times until a consistent measurement is obtained. The well depth calculation accounts for the length of the steel weight at the end of the tape and a measuring point correction. Therefore, I assume a 5% error for well depth (Table 4) that includes the possibility of misreporting values due to human error.

3.2. Model Coefficient Errors

Following methods outlined by Gurdak et al. (2007), I used the Wald 95% confidence intervals to define a conservative range of errors for the logistic regression model intercept (Table 5). The Wald 95% confidence interval was calculated using the maximum likelihood estimate and the standard error estimate of the logistic regression model coefficients (Hosmer and Lemeshow 2000).

3.3. Quantify Error Propagation with Latin Hypercube Sampling

A modified form of the Monte Carlo sampling method, called Latin Hypercube Sampling (LHS) was used to quantify error propagation within the uncertainty models (described in section 3.4) that I built using @RISK software (Palisade Corp., 2014). Monte Carlo is a stochastic method that works by the repeated sampling of probability distributions associated with the input variables and creates a robust output probability distribution. Monte Carlo methods have a wide application, including generating prediction uncertainty that is associates with logistic regression models (Gurdak et al., 2007) of groundwater vulnerability to NPS NO_3^- for the CV, BR, CL, and NACP PAs. As previously described in Gurdak and Qi (2012), the PA logistic regression models (Helsel and Hirsch, 1992) takes the general form of:

$$P = \frac{e^{(b_o + b_x)}}{1 + e^{(b_o + b_x)}} \quad (1)$$

where P = the probability of NO_3^- background exceedence, e = base of the natural logarithm, b_o = the model coefficient (i.e. logistic regression constant) and b_x = is the vector of slope coefficients and explanatory variables (i.e. explanatory variable constant).

The general Monte Carlo method for generating prediction uncertainty associated with the logistic regression model (Equation 1) is expressed as the following (McKay et al., 1979):

$$P(x) = A_{i1}(x), A_{i2}(x)A_{i3}(x),..., A_{i,n}(x) \tag{2}$$

where P is the probability output of all values that are assigned an error and probability distribution, $A_{i,n}(x)$ is the value at for each data location x (i.e. a monitoring or pumping well) for each PA, where i differentiates the inputs. The following illustrates Monte Carlo applied to the logistic regression model:.

$$PA(x) = \frac{e^{b_o+[b_1*A_{i1}(x)]+[b_2*A_{i2}(x)]+...+[b_n*A_{i,n}(x)]}}{1+e^{b_o+[b_1*A_{i1}(x)]+[b_2*A_{i2}(x)]+...+[b_n*A_{i,n}(x)]}} \tag{3}$$

where $PA(x)$ = the probability of a PA to exceed background concentrations of NO_3^- at a particular well location based on input $A_{i,n}(x)$ selected from the uncertainty probability distributions. Specifically, equation 3 represents the repeated sampling of the error distribution of the model coefficients (b_o and $b_{1\text{-}n}$) and the explanatory variables ($A_{i,n}(x)$) to create an output probability distribution $PA(x)$.

The application of Monte Carlo requires the user to select a sampling strategy for obtaining values from the probability distributions. There are three options for sampling, random, stratified and Latin Hypercube Sampling (LHS) (Helton and Davis, 2003). Random sampling relies on algorithms to repeatedly random sample the input probability distribution intervals a specified number of iterations. The greater the number of iterations, the more accurate the output distribution. The advantage of random sampling is that it is relatively straightforward to execute. The disadvantages are that random sampling is computationally intensive and not good at representing lower probability outcomes associated with the distribution. Stratified sampling allows the probability distribution to be broken into parts (or quantiles of the distribution), so that random sampling is performed within each part or quantile. The advantage of the stratified technique is that it allows a user to subdivide the distribution to ensure distribution intervals of interest are defined and assign weights to sections of the distribution of interest. The disadvantage of this method is that it is burdensome to set up especially if the analysis has many input variables. This method also requires knowledge of the right way to approach stratification, knowledge that the user might not have.

LHS is a blend of random and stratification sampling methodologies that works by dividing the probability distribution into nonoverlapping intervals of equal probability. In other words, LHS divides the distribution into equal parts and samples from each part randomly. For example say you have probability distributions for explanatory variables DO

and SOM. Each variable has it's own range and assigned distribution (i.e. normal and Extreme Value Minimum). LHS samples by dividing the cumulative distribution function (CDF) into intervals of equal probability. The CDF is a graphical product that describes the probability that a real-valued random variable X with a given probability distribution will be found at a value less than or equal to x. For each sampling interval LHS pairs the randomly selected number in that interval with another randomly selected number in that same interval of the paired variable. For example LHS would divide the distributions of DO and SOM into intervals of equal probability. From each interval, for instance the first interval, LHS would select random numbers for the first interval of DO and SOM and use this pair of randomly selected numbers to inform the NO_3^- probability output.

3.4. Development of Uncertainty Models

The PA-specific uncertainty models follow the general form of equation 3. I created the uncertainty models using the @Risk software and LHS sampling of error distribution for the PA-specific explanatory variables and model coefficients (Table 2). Each PA explanatory variable was assigned an error distribution as a percentage except for DO. Since the error for DO was found by using ArcGIS Geostatistical Analyst EBK function, DO error is defined as +/- some value in mg/L. To reflect the minimum and maximum uncertainty associated with DO, each PA where DO was an explanatory variable was run twice, once with the selected minimum EBK value as the error value and once with the selected maximum EBK value. Therefore the CV, BR, and CL have two

uncertainty models each with DO error represent minimum and maximum uncertainty across the aquifer in each model.

The errors for each explanatory variable and model coefficient were assigned normal distributions with the exception of SOM and hydric soils. The literature shows that both of these variables are under-represented in STATSGO. To capture the underrepresentation of the variables, an Extreme Value Minimum distribution (right skew) was assigned to better represent the fact that actual estimates of the data are more likely to be higher than the reported estimate in the dataset.

For each PA, I created four separate uncertainty models to evaluate the contribution of different error components on the overall model prediction uncertainty of a particular well having NPS NO_3^- that exceeds background concentrations. The first uncertainty model follows the form of equation 1 and does not include LHS calculations of error propagation. The second uncertainty model follows the form of equation 3 and uses 1,000 iterations of LHS on the error distributions of each explanatory variable and model coefficient. From PA(x) output distribution (equation 3), I selected the predicted probabilities for the 5^{th}, 25^{th}, 50^{th}, 75^{th}, and 95^{th} percentiles. The third uncertainty model is identical to second model, except LHS is done on the errors of the explanatory variables and not on the model coefficients. The fourth uncertainty model is the opposite of the third model where LHS is done on the errors of the model coefficients and not on the explanatory variables.

These four different uncertainty calculations provide a framework to evaluate the sensitivity of prediction uncertainty to error contributions from explanatory variables versus error contributions from the model coefficients. I used the relative variance contribution (RVC) (van Horssen, 2002; Gurdak et al., 2007) to evaluate the relative error contribution to the overall prediction uncertainty. The RVC due to the regression coefficients is defined as:

$$RVC_r = \frac{\sigma_r^2(P)}{\sigma_t^2(P)} x100\% \quad (4)$$

where RVC_r is a measure of the relative variance contribution due to the model coefficient only, $\sigma_r^2(P)$ is the variance of the predicted probability calculation (equation 3) due to errors from the model coefficients only (i.e., the fourth uncertainty model calculation described above), and σ_t^2 is the total variance. σ_t^2 Is the sum of the variance when both model intercept and explanatory error are accounted for in the logistic regression equation (i.e., the second uncertainty model calculation described above). The RVC due to the explanatory variables:

$$RVC_e = \frac{\sigma_e^2(P)}{\sigma_t^2(P)} x100\% \quad (5)$$

where RVC_e is a measure of the relative variance contribution due to the explanatory variables only, and $\sigma_e^2(P)$ is the variance of the predicted probability calculation (equation 3)

due to errors from the explanatory variables only (i.e., the third uncertainty model calculation described above). If the ratio of RVC_r to RVC_e is greater than one, then errors associated with the model coefficients contributed the most to prediction uncertainty. If the ratio is less than one, then the errors associated with the explanatory variables contributed the most to prediction uncertainty.

4.0 Results

4.1 Output Probability Distributions

Following the identification of input and model error, along with the assignment of the best statistical distribution to use to describe the error (i.e. normal, or extreme value minimum), a 1000 iteration LHS simulation is run on a PA model to propagate error from model inputs to a probability output (Equation 3). The resulting probability output is not a single value, but a range of possible probability values indicating the likelihood of NPS NO_3^- exceeding background concentrations in a given well. With this probability output the following things can be evaluated; 1) the magnitude of uncertainty (low, moderate, high), 2) where the highest uncertainty is concentrated (at higher or lower probabilities), 3) whether the logistic regression maps depicting the probabilities of nitrate exceeding or background are reliable after error is accounted for (low prediction uncertainty, high prediction uncertainty, moderate prediction uncertainty), 4) what explanatory variables the

logistic regression models are the most sensitive to, 5) where the model uncertainty is coming from (i.e. the regression or explanatory variables), and 6) what steps should be taken to improve model performance.

To evaluate a PAs magnitude of uncertainty and where error is concentrated the 5^{th}, 25^{th}, 50^{th}, 75^{th}, and 95^{th} probability outputs were plotted and evaluated for each PA (Figures 22-28). For each PA the 50^{th} percentile output represents the assumption of no input errors in all cases where the input variables are normally distributed. For input variables such as SOM and Hydric soils that have right skewed Extreme Value Minimum (EVM) distributions, the 50^{th} percentile represents the assumption of no errors if the median of the EVM distribution is assumed to be representative of the true value of the variable.

The relative spread of the output probability distributions in Figures 22-28 is an indicator of the magnitude of uncertainty for each PA model. Models where the percentile output lines that are closer together have less uncertainty, whereas models with a large separation of the percentile lines have more uncertainty. The skew of the percentile plots also reveals where the most uncertainty is to be found above and below the 50^{th} percentile output distribution. A pattern of right skewed distributions reveals that there is more prediction uncertainty at higher probabilities (where $P > 50\%$), whereas a pattern of left skewed distributions reveals more prediction uncertainty at lower probabilities (where is $P < 50\%$).

The CV uncertainty outputs models have the largest spread of the percentile outputs when compared to the other PA uncertainty outputs . A comparison of the two CV uncertainty models with varying DO error shows no difference in the spread or magnitude of uncertainty (Figures 24 and 25. Both model percentile outputs follow a similar shape and pattern. The lack of difference between the two CV model percentile output charts suggests that two DO errors used to define the variable uncertainty do not affect model output in a meaningful way. In other words, the uncertainty model outcomes are almost the same whether the DO error is 1.75 mg/L or 2.70 mg/L.

The output probability distributions for the CV models are characterized as step functions when sorted in ascending order of the 50th percentile (Figures 24 and 25). The relatively high percentiles (75% and 95%) of the output distribution converge quickly on a predicted probability of 1, while the relatively low percentiles (5 and 25%) converge quickly on a predicted probability of 0, but jump to a predicted probability of 1 near the tail of the distribution (Figures 24 and 25). A histogram of the 50^{th} percentile probability distribution reveals a bimodal distribution where the output tends to cluster around 0 and 1 (Figure 29). The 50^{th} percentile distribution has a median predicted probability of 88%, which indicates that the model is skews to positive. The median and the right skew of the of the percentile plots affirm that the greatest prediction uncertainty is to be found in probability values above 50%.

Interestingly the 50^{th} percentile distribution is relatively linear in the CV models (Figures 24 and 25). The linearity of the 50^{th} percentile distribution indicates that the

model is sensitive to incremental changes (i.e. increases and decreases) in the value of input explanatory variables, therefor any change to the original values of DO, farm fertilizer (20% error), soil clay (20% error), and a seasonally high water table (5%) can lead to a very high or low probability prediction regarding the presence of NO_3^- above 4 mg/L.

The percentile output distributions from the BR models (Figures 22 and 23) are similar to the output distributions from CL models (Figures 26 and 27), except that the CL models have a slightly larger output distribution and thus greater model uncertainty than the BR models. The output distributions for the two BR DO error models ((Figures 22 and 23) are nearly identical and indicate that like the CV model, the BR logistic regression model is not as sensitive to the DO error inputs. At the highest and lowest probability values (close to 0 and 1) the output distributions from both BR uncertainty models have relatively less spread. The greatest spread, thus greatest uncertainty is found at the mid-ranged predicted probabilities (Figures 22 and 23) . The skew of the percentile outputs is to the right indicating greater uncertainty at probabilities greater than 50%.

Similar to the BR, and CV outcomes, a comparison of CL DO percentile output graphs reveal that the DO uncertainty models, do not result in visually different percentile output probability distributions graphs (Figures 26 and 27). Similar to BR, the CL percentile outputs are closely spaced showing that predictions are more certain at lower probability predictions than higher, but quickly diverge at higher probabilities indicating increased uncertainty at higher probabilities. Similar to the BR and CV, the CL models skew to the right indicating higher uncertainty at probabilities above 50%.

Key explanatory variables for the CL prediction models are DO, hydric soil (37% error), and crops (20% error). Like the BR, explanatory variable error that is greater than the actual input value on the positive side increases the uncertainty of prediction. The potential increased presence of crops results in an increased nitrate load. The presence or absence of the attenuation factors like hydric soil, which is anoxic and organic rich, also impacts where nitrate is found, as hydric areas are places where denitrification occurs. DO is another attenuation factor that also influences whether nitrate persists in the aquifer. Higher DO values in the aquifer increase the probability the accumulation of NO_3^-s in groundwater.

The North Atlantic Coastal Plain has the least uncertainty of all the models (Figure 28). All percentile distributions converge to 1 following the same pattern. The NACP model is an example of a logistic regression model that is ready for use by decision makers, as the model is able to robustly predict where nitrate exceeds background. This statement is based on the fact that the R^2 for the calibration and validation models is 0.893 and 0.803 respectively (Gurdak and Qi, 2012). The R^2 results coupled with the percentile output probability distribution graph shows that this model is a useful tool capable of predicting the presence or absence of nitrate above background in the aquifer.

4.2 90th Percentile Prediction Interval Maps

To investigate the spatial variability of uncertainty surrounding groundwater vulnerability predictions, the 90th percentile prediction interval was calculated at each aquifer well location. Map results of the 90th percentile prediction interval are show in Figures 33–39. The 90th prediction captures the difference between the high end and the low end of the uncertainty output distribution with the low end represented by the 5th percentile output, and the high end represented by the 95th percentile output If the difference between the 5th and 95th percentile outputs is above 67% then the uncertainty is very high. If the difference between the 5h and 95th is between 34-66% the uncertainty is moderate, below 34% the uncertainty low. The prediction interval helps users identify where uncertainty in model predictions is greatest. To understand the utility of prediction interval maps, they must be paired with the logistic regression maps that predict where NO_3^- exceeds background concentrations (Figures 30-32) (Gurdak and Qi, 2012). Maps where the majority of the wells have a low prediction interval indicate a model with low uncertainty. In other words, if the difference between the values at the tails of the uncertainty output distribution is low then the logistic regression model prediction can be considered robust even when uncertainty is accounted for.

A comparison of probability predictions in the BR (Figure 30) to the 90th percentile prediction interval maps with varying DO error (Figures 33 and 34) reveals a high degree of uncertainty for the logistic regression model predictions. With the exception of the Salton Sea Trough, the logistic regression models predict there is a low probability

for nitrate to exceed the background concentration of 1 mg/L. When uncertainty is factored into the BR regression model, the model is shown to be moderately to highly uncertain with many 90th percentile prediction interval wells falling into probability prediction interval above 30% (Figures 33 and 34) . Again uncertainty surrounding explanatory variable inputs of irrigated lands, DO, and SOM influence whether nitrate is found in the aquifer. An increase in any of these variables influences the persistence and presence of NO_3^- in groundwater. With respect to the differences between the two uncertainty models where DO is varied, the observed difference is minimal. The BR prediction interval maps are nearly identical with the exception of a handful of wells between the two maps. Both models are the same in that they clearly show an overall trend of high predictive uncertainty for many wells in the aquifer.

The comparison of the logistic regression CV NO_3^- exceedence probability prediction map (Figure 30) to the 90th percentile prediction interval maps (Figures 35 and 36) shows there is moderate to high uncertainty regarding logistic regression predictions across the PA. There are a few wells (green dots) with a low uncertainty scattered about the PA in similar locations to the uncertain wells (orange dots). A plot of wells by high, moderate, and low 90th percentile prediction intervals reveals that many of the low uncertainty wells share the common property of being relatively deep . Depth is a physical inhibitor to the migration of nitrate. (Appendix A)

The explanatory variables soil clay and DO are important controls on the denitrification process. The presence of clay in the soil lends itself to anaerobic aquifer conditions

as clay can create localized confined conditions in aquifer. The chemistry of clay with Fe^{2+} available as an electron donor can also play a role in the reduction of NO_3^- (Ernsten, 1996). Overall uncertainty regarding the reliability of data on CV model source (fertilizer inputs), and attenuation explanatory variables (DO and soil clay) influences whether nitrate is found in the aquifer. As Figures 24 and 25 show, the CV model is very sensitive to any variation of these variables as inputs.

A comparison of the CL logistic regression NO_3^- exceedence probability prediction map (Figure 31) to the 90th percentile prediction interval maps (Figures 37 and 38) shows moderate to high prediction uncertainty. Similar to the CV and BR, different DO error values did not change the uncertainty outcomes of the two 90th percentile prediction uncertainty maps (Figures 37 and 38). The logistic regression model mostly predicts a low probability ($P< 40$ %, Figure 31) of nitrate exceeding background in this aquifer. In order for there to be more certainty regarding predictions the uncertainty of DO variability and presence of hydric soil (37% error) should be better quantified. To minimize uncertainty for CL NO_3^- predictions map users should accurately identify the extent of hydric soils and variability of DO in the aquifer.

A comparison of the logistic regression NACP model to the 90th percentile prediction probability prediction map (Figure 32) to the 90th percentile prediction map (Figure 39) shows that the uncertainty of the model predictions is low. The results from the 90th percentile prediction map is consistent with the percentile probability distribution output

figure (Figure 28) indicating a robust model capable of giving good predictions despite the inclusion of model and explanatory variable uncertainty.

4.3 Relative Variance Contribution

For all uncertainty models except the NACP, the RVC_r is greater than the RVC_e (Figure 40). For the NACP, the explanatory variable error is a greater contributor to uncertainty. Using the Wilcoxon rank-sum test, the difference between RVC_r and RVC_e was significant for all aquifers ($p< 0.001$). When pairs of RVC_r and RVC_e values for the PA DO error models were compared, no significant difference between the means of the DO error model pairs ($p< 0.001$) was found (Figure 37).

A relative variance outcome where the regression error is found to be more influential than the explanatory variable error indicates the need for an expansion of the monitoring well network. The larger RVC_r value reveals that more wells are needed in areas of high uncertainty to more adequately capture the spatial variability of nitrate concentrations in recently recharged groundwater. Therefore, an expansion of the set of monitoring wells in areas of greatest uncertainty (Figures 33-38) would be useful in reducing regression model uncertainty for the BR, CV, and CL

The NACP relative variance for explanatory variables is greater than regression variance. To improve the predictive capabilities of this aquifer the focus should be on the quality and accuracy of the data supplied as inputs for the logistic regression. For example NACP explanatory variable farm fertilizer has an uncertainty of 20%. To improve the

predictive capability of the logistic regression model the accuracy of fertilizer input data should be verified.

5.0 Discussion

The three PAs where DO is an explanatory variable of interest offer a best and worse case insight into the probability NO_3- exceeding background. In each of the aquifers the true variability of DO is unknown. In order understand how DO variability impacts the regression models two LHS uncertainty models were created to test model uncertainty with lower and higher DO error. The results show that the CV, BR, and CL output is not very sensitive to the changes in DO error. To improve the utility of the models the well networks should be expanded to include more wells so the models predictive abilities can be improved. LHS can also test the logistic regression models sensitivity to error. The NACP probability output graph and 90^{th} percentile prediction interval map demonstrates this model is the least sensitive to the error propagation.

Accounting for uncertainty in model outcomes allows water resource managers to plan according a more conservative level of acceptable risk. Quantifying uncertainty allows for better resource allocation in that it allows for the prioritization of where to focus on nitrate mitigation efforts. For example in areas with a heavy dependency on domestic wells for drinking water, regulators can work on ensuring DO that the variability of DO is known. Areas with organic rich soils should be delineated and identified for better predictive outcomes.

This methods used in this research quantify the error associated with explanatory variables. Future research that takes this uncertainty data to improve upon the models predicative abilities should consider the error associated with the explanatory variable as part of the selection criteria for variables to be included in the logistic regression. An explanatory variable deemed statistically significant during multivariate logistic regression analysis may not be a desirable explanatory variable for the final vulnerability model if the inherent error is exceptionally large compared to other statistically significant variables. The outcome of such a consideration may be logistic regression models with different explanatory variables.

With respect to the question of uncertainty compared to what or how does one factor uncertainty into decision making I suggest the decision maker use the following approach. Identify the highest priority areas of vulnerability. Evaluate the initial predictions of the logistic regression model. If the logistic regression model suggests that the probability of exceeding the threshold value is above 70% and the uncertainty model's 90^{th} percentile prediction interval uncertainty is in excess of 60% then the decision maker should not use the logistic regression prediction and proceed with collecting more data about the area to get a better understanding of the actual vulnerability of the study area. If the logistic regression model predictions of vulnerability are less that 70% and the 90^{th} percentile prediction interval uncertainty is less than 60% then I would recommend that the decision maker use the logistic regression model as a planning tool keeping in mind

that more wells and/or more accurate datasets are needed to improve predictive capabilities.

6.0 Conclusion

An efficient way to quantify uncertainty, of predictive models is to apply LHS techniques to the error associated with model inputs. LHS is a cost effective, easy to implement method that can be used to improve the predictive capabilities designed to detect a groundwater parameter above a certain threshold. An advantage of the LHS technique as described in this paper is that \ gives the modeler flexibility to constrain explanatory error based on researched best guesses, and to assign a proper distribution to all inputs to errors. At minimum an LHS uncertainty assessment can assist with defining the "best case" and "worse case" probable scenarios for contamination, and identify actions that need to be taken to improve model performance.

The vulnerability and uncertainty maps presented in this research are designed to help resource decision makers prioritize areas for ground water quality monitoring or implement alternative management practices that mitigates human exposure and the release of NO_3^- into groundwater. The appropriate scale to use these maps is 1:250,000, or the scale of the USDA's STATSGO dataset. The intent of these maps is to provide regulators and resource managers a tool to evaluate vulnerability at regional scales. Other field investigative methods should be used to determine the specific sources of NO_3^- contamination.

7.0 References

Amichev, Beyhan Y., and John M. Galbraith. "A revised methodology for estimation of forest soil carbon from spatial soils and forest inventory data sets." *Environmental Management* 33.1 (2004): S74-S86.

Anderson, James Richard. *A land use and land cover classification system for use with remote sensor data.* Vol. 964. US Government Printing Office, 1976.

Böhlke, J. K., et al. "Denitrification in the recharge area and discharge area of a transient agricultural nitrate plume in a glacial outwash sand aquifer, Minnesota." *Water Resources Research* 38.7 (2002): 10-1.

Brupbacher, R. H., and J. E. Sedberry. "Jr. &. WH Willis. The coastal marshlands of Louisiana. Chemical properties of the soil materials." *La. Agr. Exp. Stn. Bull* 672:15, 1973.

Bricklemyer, Ross S., et al. "Sensitivity of the Century model to scale-related soil texture variability." *Soil Science Society of America Journal* 71.3 (2007): 784-792.

Burow, Karen R., et al. "Nitrate in goundwater of the United States, 1991– 2003." *Environmental science & technology* 44.13 (2010): 4988-4997.

California Department of Water Resources county level agricultural GIS maps are available at http://www.water.ca.gov/landwateruse/lusrvymain.cfm Retrieved 3/1/2014.

Commission for Environmental Cooperation (Montréal, Québec)., and Secretariat. *Ecological regions of North America: toward a common perspective*. The Commission, 1997

Comparison of the USGS 2001 NLCD to the 2002 USDA Census of Agriculture for the Upper Midwest United States Wickham, J. D., et al. "Thematic accuracy of the NLCD 2001 land cover for the conterminous United States." Remote Sensing of Environment 114.6 (2010): 1286-1296.

Corwin, Dennis L., and R. J. Wagenet. "Applications of GIS to the modeling of nonpoint source pollutants in the vadose zone: A conference overview." *Journal of Environmental Quality* 25.3 (1996): 403-411.

Cunningham, William L., and C. W. Schalk. "Groundwater technical procedures of the US Geological Survey." *US Geological Survey Techniques and Methods GWPD 11*—measuring well depth by use of a graduated steel tape1-A1. (2011): 95-104

Davidson, Eric A., and Paul A. Lefebvre. "Estimating regional carbon stocks and spatially covarying edaphic factors using soil maps at three scales." *Biogeochemistry* 22.2 (1993): 107-131.

Davidson, Eric A. "Spatial covariation of soil organic carbon, clay content, and drainage class at a regional scale." *Landscape Ecology* 10.6 (1995): 349-362.

ESRI. ArcGIS Help 10.1 Resource Center. Redlands, CA: *Environmental Systems Research Institute*, 2012. Retrieved on 2/22/2014 http://resources.arcgis.com/en/help/main/10.1/

Fan, Anna M., and Valerie E. Steinberg. "Health implications of nitrate and nitrite in drinking water: an update on methemoglobinemia occurrence and reproductive and developmental toxicity." *Regulatory toxicology and pharmacology* 23.1 (1996): 35-43.

Faunt, Claudia C., ed. *Groundwater Availability of the Central Valley Aquifer, California*. U.S. Geological Survey Professional Paper 1766, 2009.

Faunt, Claudia C., Kenneth Belitz, and Randall T. Hanson. "Development of a three-dimensional model of sedimentary texture in valley-fill deposits of Central Valley, California, USA." *Hydrogeology journal* 18.3 (2010): 625-649.

Galloway, James N., et al. "The nitrogen cascade." *Bioscience* 53.4 (2003): 341-356.

Gaither, Kelly, and Terry, D.L., *Uniform fertilizer tonnage reporting system — version 4 — Instruction manual*. Association of American Plant Food Control, 2004.

Gottsegen, Jonathan, Montello, D., and Goodchild, M., 1999, "A comprehensive model of uncertainty in spatial data", in Lowell, K. Lowell, Kim, and Annick Jaton, eds. *Spatial accuracy assessment: land information uncertainty in natural resources*. CRC Press, 2000.

Gurdak, Jason J., et al. "Latin hypercube approach to estimate uncertainty in ground water vulnerability." *Groundwater* 45.3 (2007): 348-361.

Gurdak, Jason J., and Sharon, L. Qi, *Vulnerability of recently recharged ground water in the High Plains aquifer to nitrate contamination*. U.S. Geological Survey Scientific Investigations Report 2006-5050, 2006.

Gurdak, Jason J., and Sharon L. Qi. "Vulnerability of recently recharged groundwater in principle aquifers of the United States to nitrate contamination." *Environmental science & technology* 46.11 (2012): 6004-6012.

Gurdak, Jason J. Qi, Sharon.L., and Michael Geisler. *Estimating prediction uncertainty from geographical information system raster processing: A user's manual for the Raster Error Propagation Tool (REPTool).* U.S. Geological Survey Techniques and Methods 11–C3, 2009.

Helton, Jon C., and Freddie Joe Davis. "Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems." *Reliability Engineering & System Safety* 81.1 (2003): 23-69.

Hosmer Jr, David W., and Stanley Lemeshow. *Applied logistic regression.* John Wiley & Sons, 2004.

Ji, Lei, and Albert J. Peters. "Assessing vegetation response to drought in the northern Great Plains using vegetation and drought indices." *Remote Sensing of Environment* 87.1 (2003): 85-98.

Kawabata, A. K. I. R. A., K. A. Z. U. H. I. T. O. Ichii, and Y. A. S. U. S. H. I. Yamaguchi. "Global monitoring of interannual changes in vegetation activities using NDVI and its relationships to temperature and precipitation." *International Journal of Remote Sensing* 22.7 (2001): 1377-1382.

Krivoruchko, K. *Empirical Bayesian Kriging.* ArcUser, (2012).Retrieved on 2/22/2014 http://www.esri.com/ news/arcuser/1012/empirical-byesian-kriging.html 3

Loague, Keith. "The impact of land use on estimates of pesticide leaching potential: Assessments and uncertainties." *Journal of contaminant hydrology* 8.2 (1991): 157-175.

Loague, Keith, et al. "Uncertainty of groundwater vulnerability assessments for agricultural regions in Hawaii: Review." *Journal of Environmental Quality* 25.3 (1996): 475-490.

McMahon, Peter B., et al. "Occurrence of nitrous oxide in the central High Plains aquifer, 1999." *Environmental science & technology* 34.23 (2000): 4873-4877.

McMahon, Peter B., J. K. Böhlke, and S. C. Christenson. "Geochemistry, radiocarbon ages, and paleorecharge conditions along a transect in the central High Plains aquifer, southwestern Kansas, USA." *Applied Geochemistry* 19.11 (2004): 1655-1686.

McMahon, Peter. B., and Frank H. Chapelle. "Redox processes and water quality of selected principal aquifer systems." *Ground Water* 46.2 (2008): 259-271.

McKay, Michael D., Richard J. Beckman, and William J. Conover. "Comparison of three methods for selecting values of input variables in the analysis of output from a computer code." *Technometrics* 21.2 (1979): 239-245.

Mowrer, H. Todd, and Russell G. Congalton, eds. *Quantifying spatial uncertainty in natural resources: theory and applications for GIS and Remote Sensing*. CRC Press, 2003.

National Atlas (Reston, Virginia). *Precipitation-Average Annual Precipitation 1961-1990.* National Atlas, (2013).

Page, R.W. *Geology of the fresh ground-water basin of the Central Valley, California, with texture maps and sections*. US Geological Survey Professional Paper 1401-C, (1986), 54 p.

Pervez, Md Shahriar, and Jesslyn F. Brown. "Mapping irrigated lands at 250-m scale by merging MODIS data and national agricultural statistics." *Remote Sensing* 2.10 (2010): 2388-2412.

Planert, M., and Williams, J.S., 1995. *Groundwater Atlas of the United States- California and Nevada.* U.S. Geological Survey Professional Paper HA 730-B, Basin and Range section, (1995). Retrieved on 2/1/2014. http://pubs.usgs.gov/ha/ha730/ch_b/B-text2.html.

Rasmussen, Craig. "Davidson, Eric A. "Spatial covariation of soil organic carbon, clay content, and drainage class at a regional scale." *Landscape Ecology* 10.6 (1995): 349-362.

Rosenblatt, A. E., et al. "Identifying riparian sinks for watershed nitrate using soil surveys." *Journal of Environmental Quality* 30.5 (2001): 1596-1604.

Rose, Seth, and Austin Long. "Monitoring dissolved oxygen in ground water: Some basic considerations." *Ground Water Monitoring & Remediation* 8.1 (1988): 93-97.

Zhong, B., and Y. J. Xu. "Scale effects of geographical soil datasets on soil carbon estimation in Louisiana, USA: A comparison of STATSGO and SSURGO." *Pedosphere* 21.4 (2011): 491-501.

Ryder, P. *Groundwater Atlas of the United States- Oklahoma, Texas.* U.S. Geological Survey Professional Paper HA 730-E, Coastal Lowlands Aquifer System section, (1996).http://pubs.usgs.gov/ha/ha730/ch_e/E-text6.html

Ryder, P. *Groundwater Atlas of the United States- Oklahoma, Texas.* U.S. Geological Survey Professional Paper HA 730-E, Coastal Lowlands Aquifer System section, (1996).http://pubs.usgs.gov/ha/ha730/ch_e/E-text6.html

Ruddy, B. C.; Lorenz, D. L.; Mueller, D. K. *County-level estimates of nutrient inputs to the land surface of the conterminous United States, 1982–2001*. US Geological Survey Scientific Investigations Report 2006–5012, 2006.

Seelan and Kurz, – Field data obtained through personnel communication. Santhosh K. Seelan and Bethany A. Kurz of the University of North Dakota, 2006.

Spalding, Roy F., and Mary E. Exner. "Occurrence of nitrate in groundwater—a review." *Journal of environmental quality* 22.3 (1993): 392-402.

Trapp, Henry. Jr., and Marilee A. Horn. *Groundwater Atlas of the United States- Delaware, Maryland, New Jersey, North Carolina, Pennsylvania, Virginia, West Virginia.* U.S. Geological Survey Professional Paper HA 730-E, Northern Atlantic Coastal Plain aquifer system section (1997).
http://pubs.usgs.gov/ha/ha730/ch_l/L-text3.html

Van Horssen, P. W., E. J. Pebesma, and P. P. Schot. "Uncertainties in spatially aggregated predictions from a logistic regression model." *Ecological Modelling* 154.1 (2002): 93-101.

US Army Corps of Engineers. *Wetlands Delineation.* Manualhttp://www.wetlands.com/coe/87manp3b.htm. Retrieved March 15th, 2014.

US Department of Agriculture, Natural Resource Conservation Service. 1994. *State Soil Geographic (STATSGO), Data Use Information.* NSSC Miscellaneous Publication Number 1492, (1994).

United States Department of Agriculture, National Agricultural Statistics Service. *Census of Agriculture National Agricultural Statistics Service Research, Education, and Economics Volume 1.* Geographic Area Series Part 51 AC-02-A-51 (2004).

US Department of Agriculture. National Agricultural Statistics Service. *Watersheds.* 2007 Census of Agriculture. Volume 2. Subject Series, 2007.

US Department of Agriculture, National Resource Conservation Service. *Soil-Field Indicators of Hydric Soils in U.S. 7.0.*, 2010.

Williamson, A.K., Prudic, D.E., and Swain, L.A. *Ground-water flow in the Central Valley, California.* U.S. Geological Survey Professional Paper 1401-D, 1989.

Wang, J., P. M. Rich, and K. P. Price. "Temporal responses of NDVI to precipitation and temperature in the central Great Plains, USA." *International Journal of Remote Sensing* 24.11 (2003): 2345-2364.

Wulder, Michael A., et al. "High spatial resolution remotely sensed data for ecosystem characterization." *BioScience* 54.6 (2004): 511-521

TABLES

Table 1: Principle Aquifers study area for which uncertainty models are created, and selected aquifer properties summarizing geology, annual precipitation, and the percentages of acres that are irrigated within the aquifer.

Principal Aquifer	General Lithology	Precipitation (inches per year)*	Estimated Percent Irrigated Acres by Region†
Basin and Range	Internally draining basins comprised of alluvium, fractured volcanics or fractured carbonate.	5 – 40	Great Basin: 7-12% Upper Colorado: 4-6% Lower Colorado: 4 %
Central Valley	Alluvial trough basin	5-25	13%
Coastal Lowlands	Wedge shaped marine and deltaic deposits. Lithology not well defined.	35-70	Texas: 4% Louisiana: 13%
North Atlantic Coastal Plain	Wedge shaped marine and deltaic deposits. Well defined lithology; aquitards and aquifers easily identified.	35-60	North NACP: 4% Southern NACP: 4-6%

*National Atlas, 2013

†Ruddy et al., 2006

Table 2: Table showing principal aquifer background nitrate concentrations, logistic regression model intercepts, and explanatory variable coefficients, calibration and validation R^2.

Principal Aquifer	Back-ground Nitrate (mg/L)	Explanatory Variables (units)	Model Intercepts and Explanatory Variable (Coefficients)	Calib-ration R^2	Valida-tion R^2
Basin and Range	1	DO (mg/L) Irr (%) SOM (%)	Model Intercept (-2.566) DO (0.5616) Irr (0.0238) SOM (-1.7997)	0.762	0.664
Central Valley	4	DO (mg/L) Fert (kg/km^2) SC (%) HWT (m)	Model Intercept (-13.6456) DO (0.4405) Fert (0.000258) SC (-0.0858) HWT(7.0443)	0.839	0.485
Coastal Lowlands	0.061	DO (mg/L) Hy (%) Crop (%)	Model Intercept (-0.9184) DO (0.2829) Hy (0.0305) Crop (0.0167)	0.757	0.164
North Atlantic Coastal Plain	0.5	Anox (1–Yes,0– No) HWT (m) WD (m) Fert (kg/km^2)	Model Intercept (1.3175) Anox (0.176) HWT (2.7658) WD (0.0738) Fert (0.000159)	0.893	0.803

Acronym Explanation: **Anox** – anoxic redox; **DO**-dissolved oxygen; **Crop** – Crops; **Fert** – fertilizer; **HWT** – seasonally high water table; **Hy** – hydric soil; **Irr**- irrigated lands; **SC** – soil clay; **SOM**- soil organic matter; **WD** – well depth

Table 3: Logistic regression models for the Central Valley, Basin and Range, Coastal Lowlands, North Atlantic Coastal Plain aquifers.

Principal Aquifer	Logistic Regression Model
BR	$$P = \frac{e^{-2.566+(0.5616*DO)+(0.0238*Irr)+(-1.7997*SOM)}}{1+e^{-2.566+(0.5616*DO)+(0.0238*Irr)+(-1.7997*SOM)}}$$
CV	$$P = \frac{e^{-13.6456+(0.4405*DO)+(0.000258*fert)+(-0.0858*SC)+(7.0443*HWT)}}{1+e^{-13.6456+(0.4405*DO)+(0.000258*fert)+(-0.0858*SC)+(7.0443*HWT)}}$$
CL	$$P = \frac{e^{-0.9184+(0.2829*DO)+(0.0305*Hy)+(0.0167*Crop)}}{1+e^{-0.9184+(0.2829*DO)+(0.0305*Hy)+(0.0167*Crop)}}$$
NACP	$$P = \frac{e^{1.3175+(0.0176*Anox)+(2.7658*HWT)+(0.0738*WD)+(0.000159*fert)}}{1+e^{1.3175+(0.0176*Anox)+(2.7658*HWT)+(0.0738*WD)+(0.000159*fert)}}$$

Table 4: Table of error values for logistic regression explanatory variables by principal aquifer used to quantify uncertainty in @RISK. Principal aquifers are listed across the top and explanatory variables are listed vertically on the left. The statistical distribution applied to the dataset in @RISK is listed in the far right column.

Principal Aquifer	CV Error	BR Error	CL Error	NACP Error	@Risk Distribution
Dissolved oxygen min (mg/L)	1.75	2.07	1.07		Normal
Dissolved oxygen max (mg/L)	2.70	2.45	1.97		Normal
Soil Organic Matter (%)		20%			Extreme Value Minimum
Hydric (%)			37 %		Extreme Value Minimum
Irrigated lands (%)		11%			Normal
Farm Fertilizer (%)	20%			20%	Normal
Crops (%)			20%		Normal
Seasonally High Water Table (%)	5%			5%	Normal
Anoxic redox (%)				0%	Normal
Soil Clay (%)	20%				Normal
Well Depth (%)				5%	Normal

Table 5: Sources of data for soil organic matter error. A summary of published literature that has quantified the error associated with SOM values in the STATSGO dataset. Published values were averaged to derive an error value for use in the uncertainty models. Average error associated with STATSGO is approximately -20%.

Location of SOM	**Source**	**STATSGO Error (- represents underestimation)**
Louisiana- SOM at 20 cm	Zhong and Xu, (2011)	-9%
Louisiana- SOM at 100 cm	Zhong and Xu, (2011)	-36%
Louisiana- Soil organic carbon densities at 30 cm	Zhong and Xu, (2011)	-4.7%
Louisiana- Soil organic carbon densities	Zhong and Xu, (2011)	-23%
Maine SOM	Davidson and Lefebvre (1993)	-13%
Western Oregon SOM	Homann et al. (1998)	-25%
		Average Error: -18.5% (Rounded to -20% for use in the uncertainty models)

Table 6: Table of logistic regression model intercepts with Wald 95% confidence limits, used to define the principal aquifer intercept error and distribution in the uncertainty models.

Principal Aquifer	**Model intercept**	**Wald 95% Confidence limits**	
	estimate	**lower**	**upper**
Basin Range	-2.566	-4.5033	-0.6287
Central Valley	-13.6456	-25.844	-1.4472
Coastal Lowlands	-0.9184	-2.5473	0.7105
NACP	-1.3175	-2.7349	0.0998

FIGURES

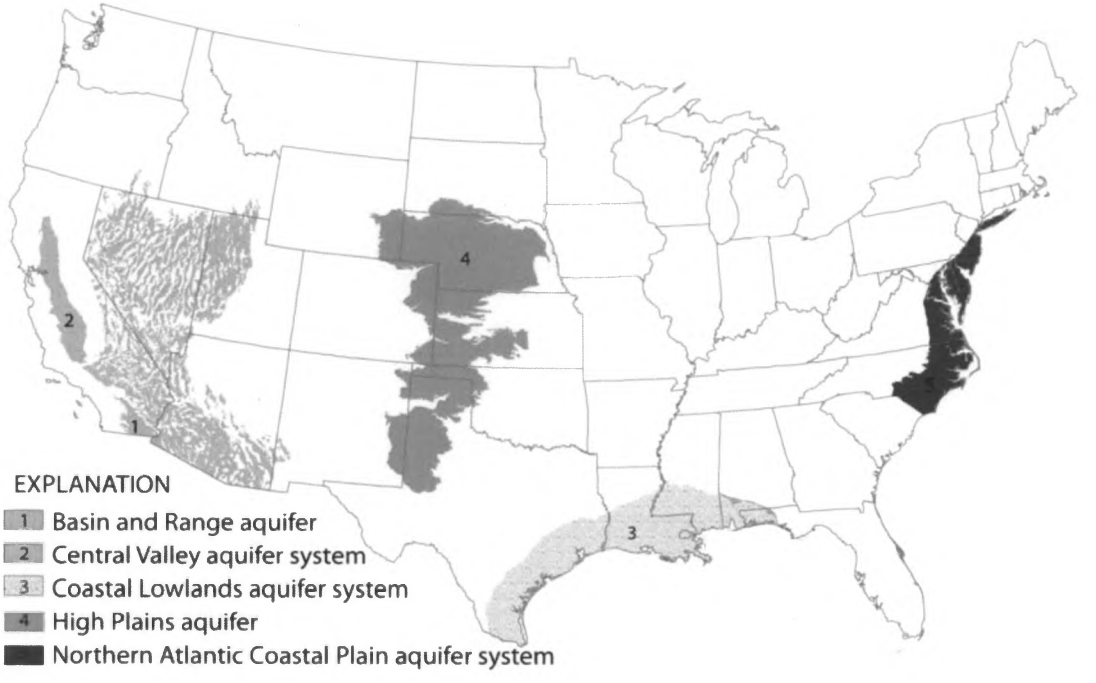


Figure 1: Location of Principal Aquifers where uncertainty models were created that quantify the uncertainty associated with (Gurdak and Qi, 2012) logistic regression models that predict the probability of NO_3^- exceeding background concentrations. Aquifers where uncertainty models were created are (from west to east) the Central Valley, Basin and Range, Coastal Lowlands, and the North Atlantic Coastal Plain.

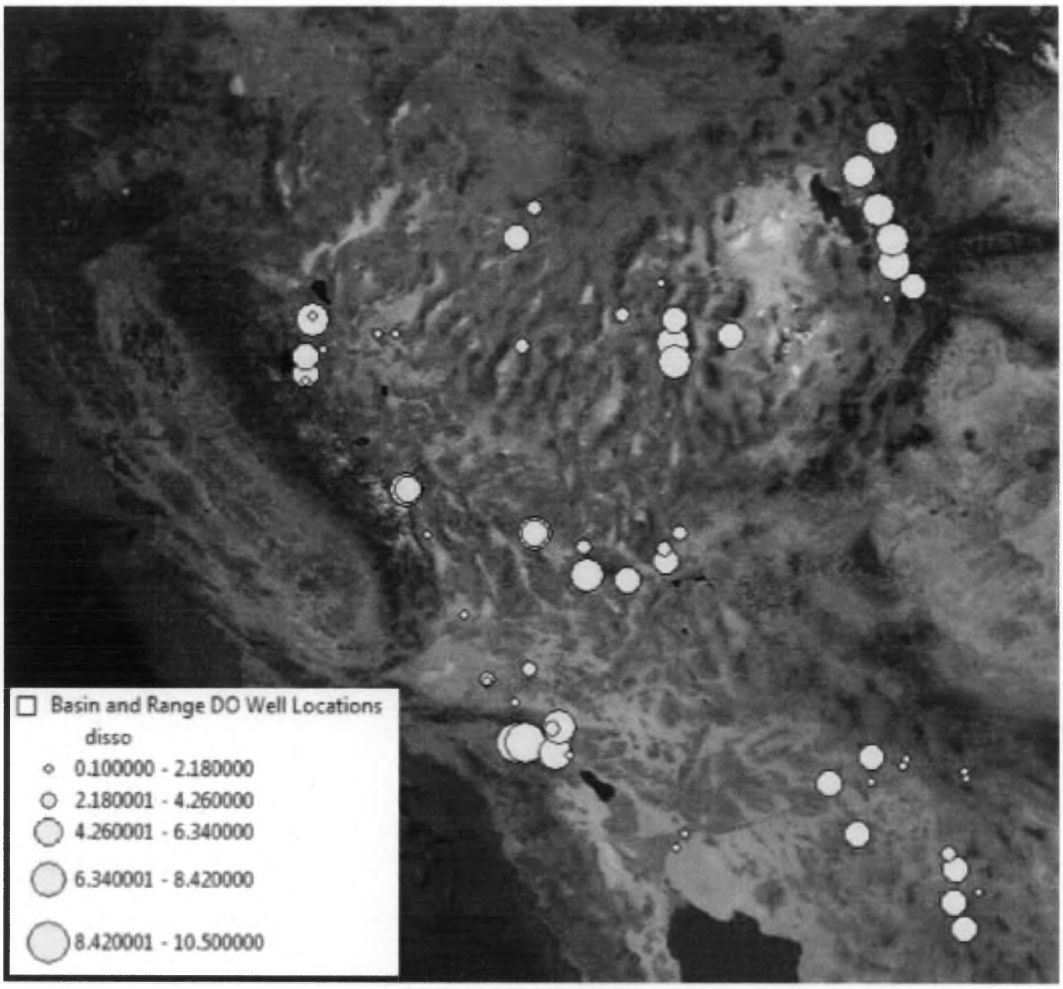


Figure 2: Basin and Range dissolved oxygen well locations symbolized in five concentration ranges in units of mg/L.

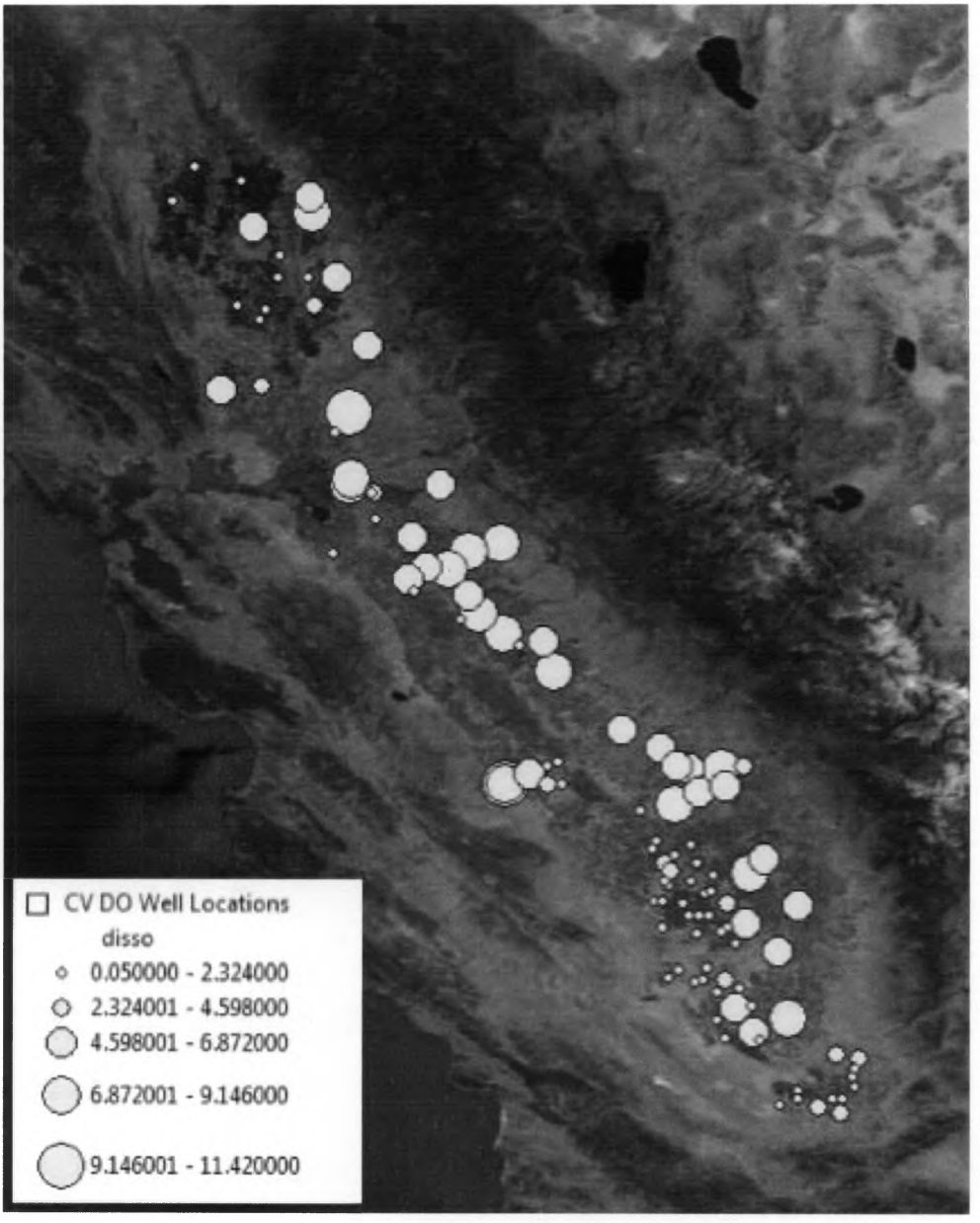


Figure 3: Central Valley dissolved oxygen well locations symbolized, in five concentration ranges in units of mg/L.

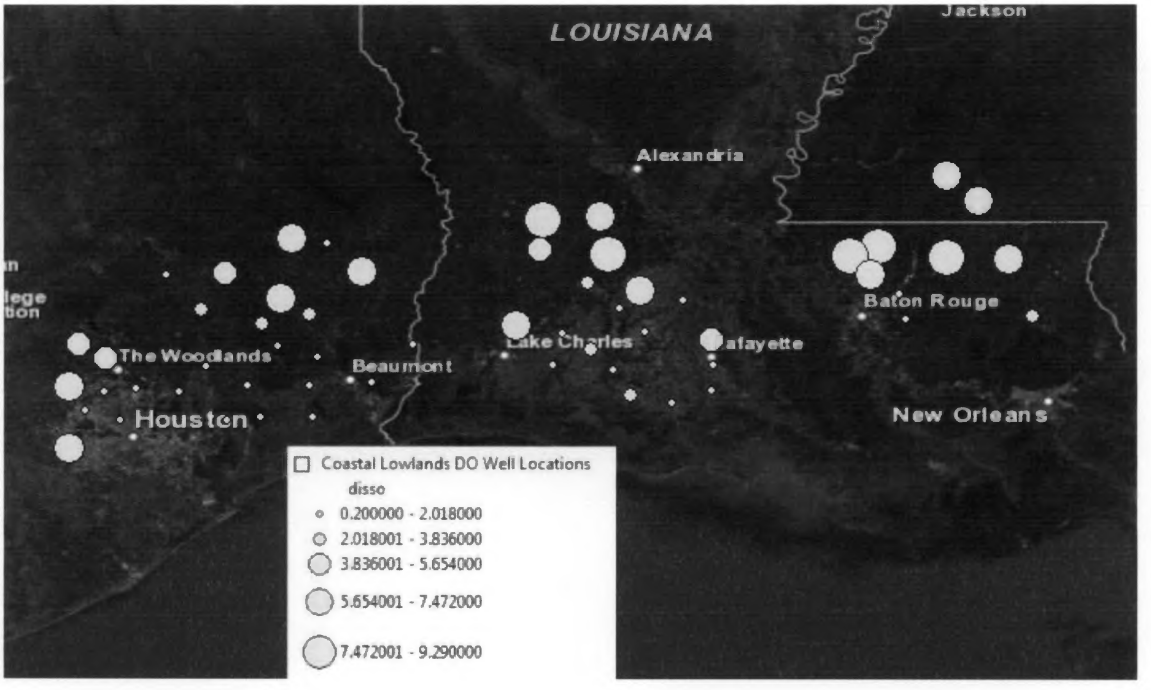


Figure 4: Coastal Lowlands dissolved oxygen well locations, symbolized in five concentration ranges in units of mg/L.

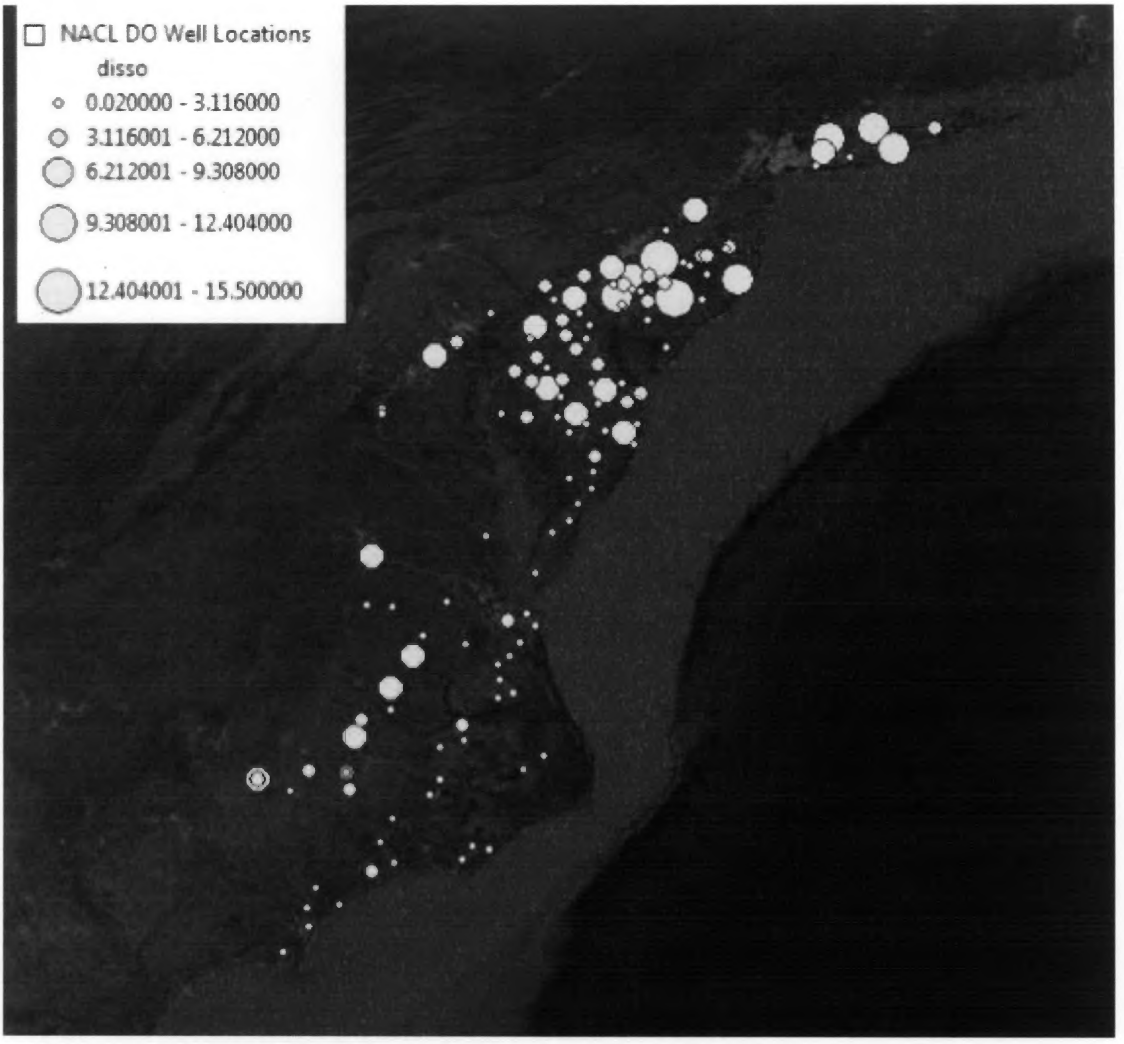
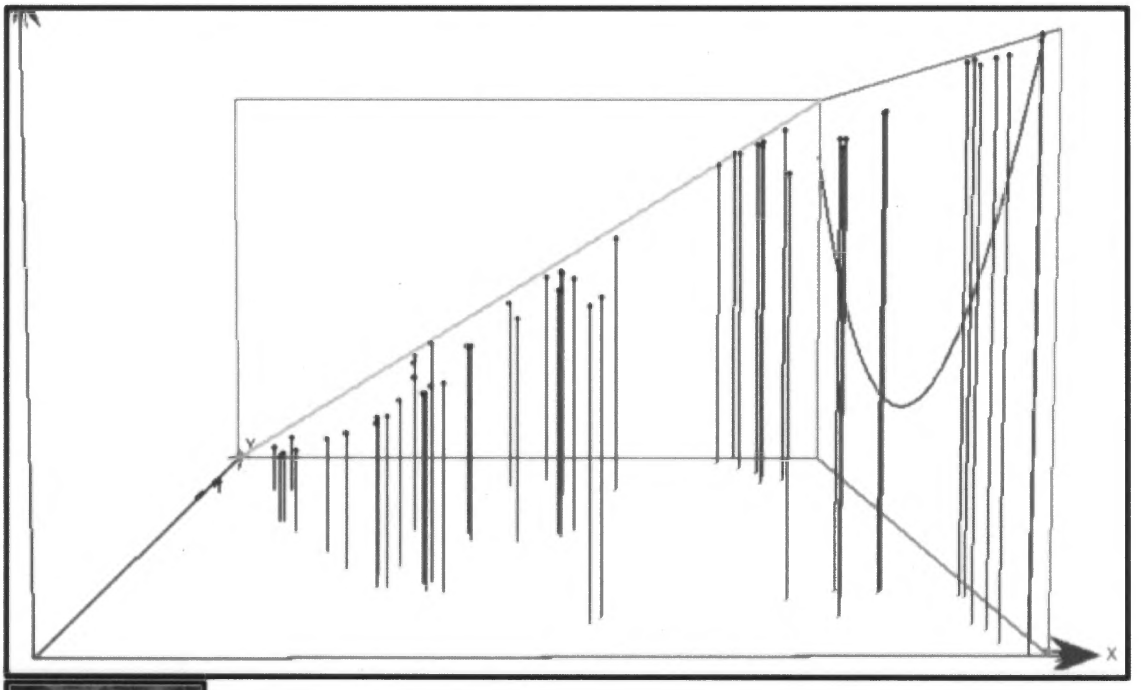


Figure 5. North Atlantic Coastal Plain dissolved oxygen well locations, symbolized in five concentration ranges in units of mg/L.



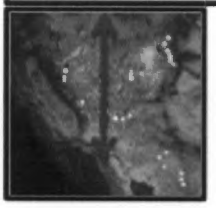
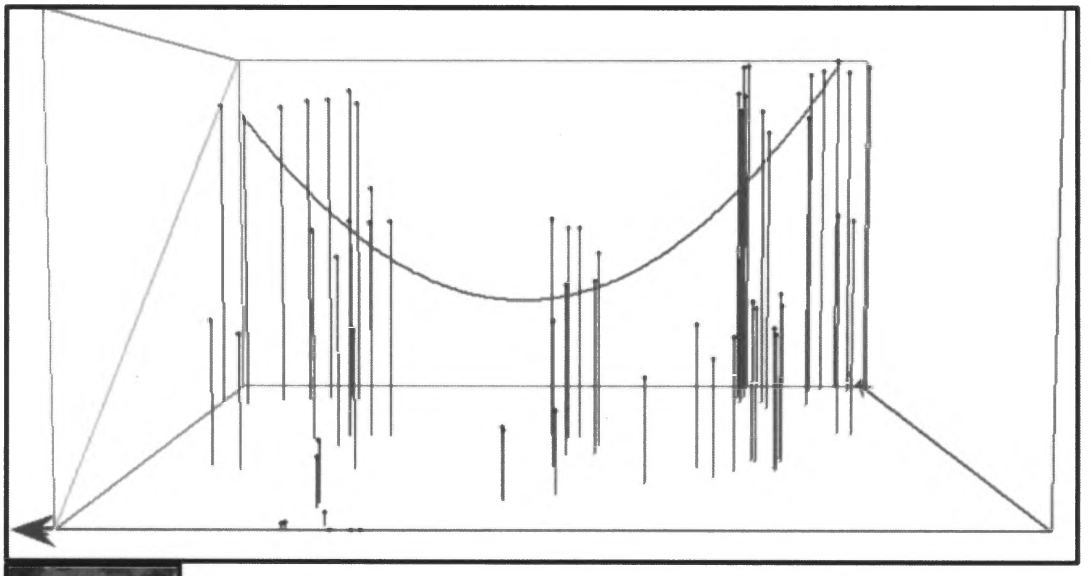


Figure 6: Basin and Range dissolved oxygen (DO) 3D east-west cross section showing major trends in DO concentrations over the entire aquifer. This view shows the trend across an east-west transect of the aquifer. The blue curve shows that dissolved oxygen concentrations are generally higher on the east and west side of the basins adjacent to recharge areas of the Great Salt Lake Basin and Sierra-Nevada/White Mountains respectively.



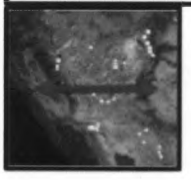


Figure 7: Basin and Range dissolved oxygen (DO) 3D north-south cross-section showing major trends in DO concentrations over the entire aquifer. This view shows the trend across a north-south transect of the aquifer. The green line shows dissolved oxygen concentrations are generally higher in the north, where precipitation is higher, and lower in the south, where precipitation is lower.

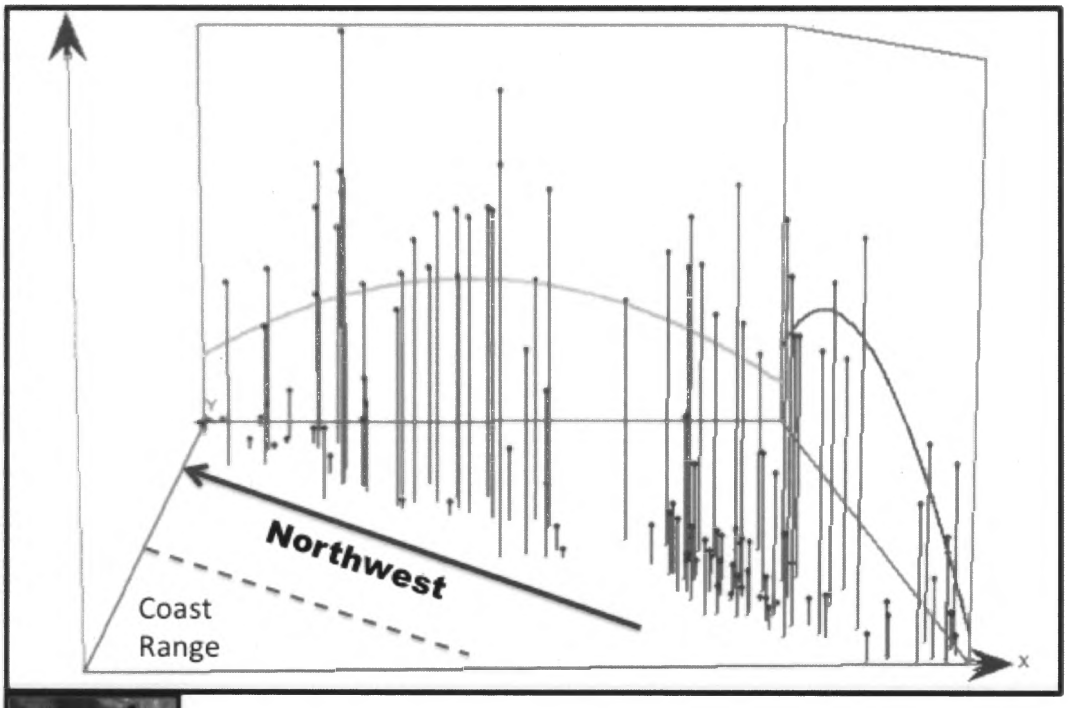
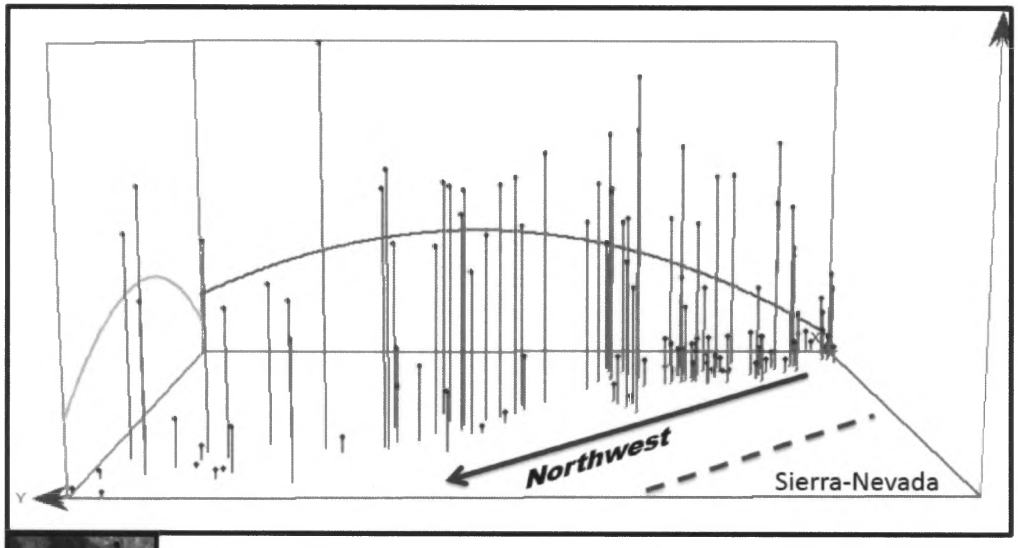


Figure 8. Central Valley dissolved oxygen (DO) 3D northwest-southeast cross section, showing major trends in DO concentrations over the entire aquifer. This view shows the trend from southeast to northwest looking east from the Coast Range Mountains .The green line in the figure represents concentrations trends perpendicular to the view. DO concentrations in this direction are generally higher in the center of the basin and lower at the periphery.



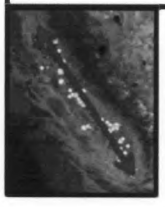


Figure 9: Central Valley dissolved oxygen (DO) 3D northwest-southeast cross section showing major trends in DO concentrations over the entire aquifer. The view shows the trend from northwest to southeast looking west from the Sierra-Nevada Mountains. The blue line in the figure represents concentrations trends perpendicular to the view. DO concentrations in this direction are generally higher in the center of the basin and lower at the periphery, although the trend shows that DO concentrations on the northwest end of the basin are higher than the southern end of the basin.

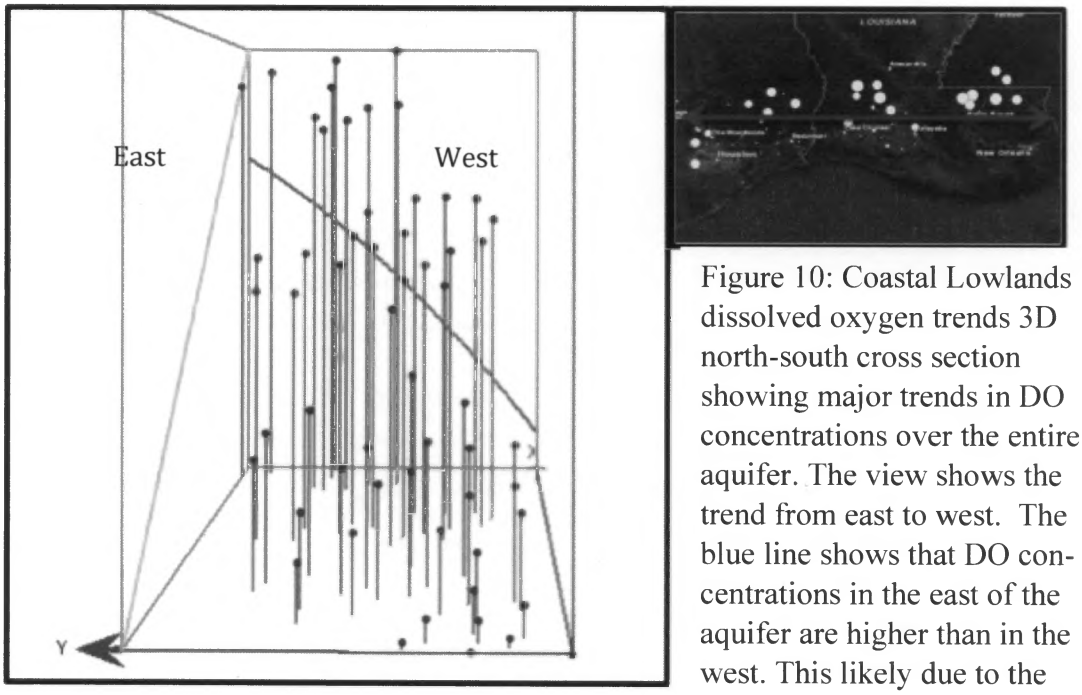
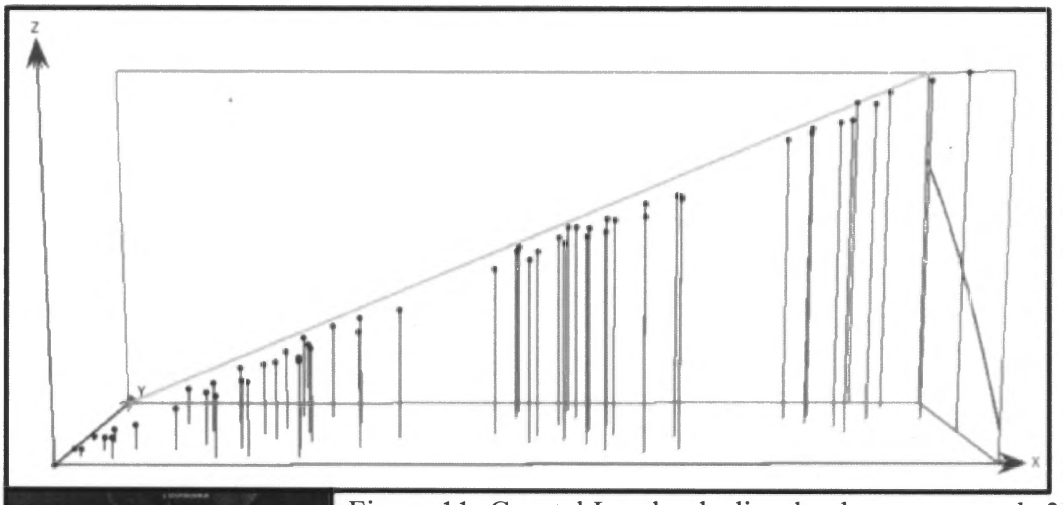


Figure 10: Coastal Lowlands dissolved oxygen trends 3D north-south cross section showing major trends in DO concentrations over the entire aquifer. The view shows the trend from east to west. The blue line shows that DO concentrations in the east of the aquifer are higher than in the west. This likely due to the influence of the Mississippi River in the eastern parts of the aquifer.



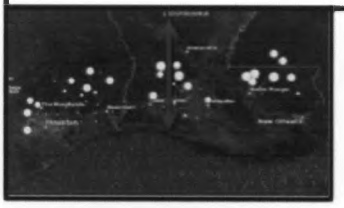


Figure 11: Coastal Lowlands dissolved oxygen trends 3D north-south cross section showing major trends in DO concentrations over the entire aquifer. The view shows the trend from north to south. The green line shows that DO concentrations in the north of the aquifer are higher than in the south. This likely due to conditions along the coast that result in an anoxic groundwater environment.

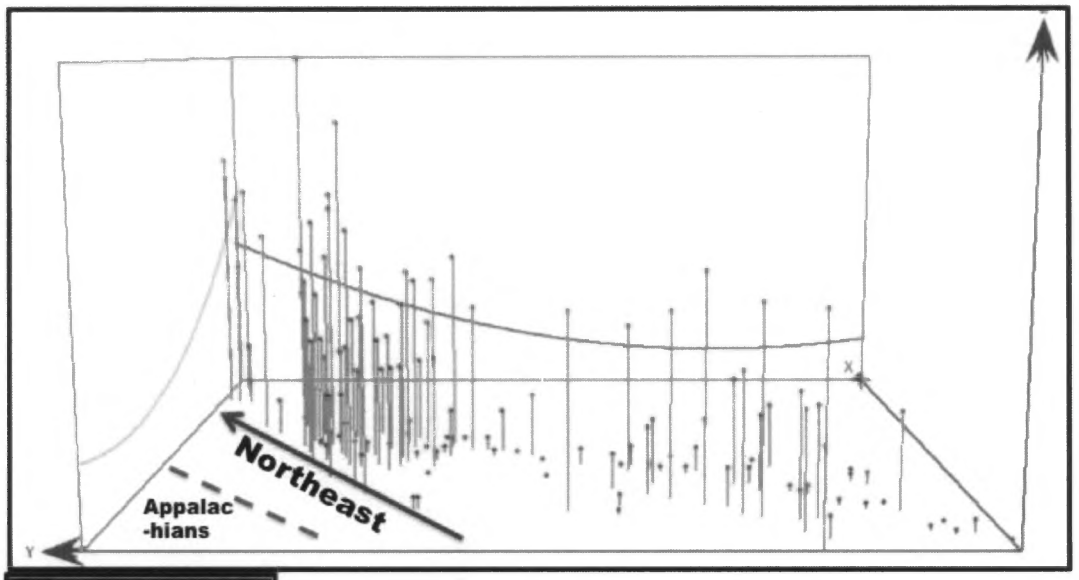


Figure 12: North Atlantic Coastal Plain dissolved oxygen trends 3D northeast-southwest cross section, view looking east from the Appalachian Mountains, showing major trends in DO concentrations over the entire aquifer. The blue line shows DO concentrations perpendicular to the view. DO concentrations are higher in the northwest recharge area.

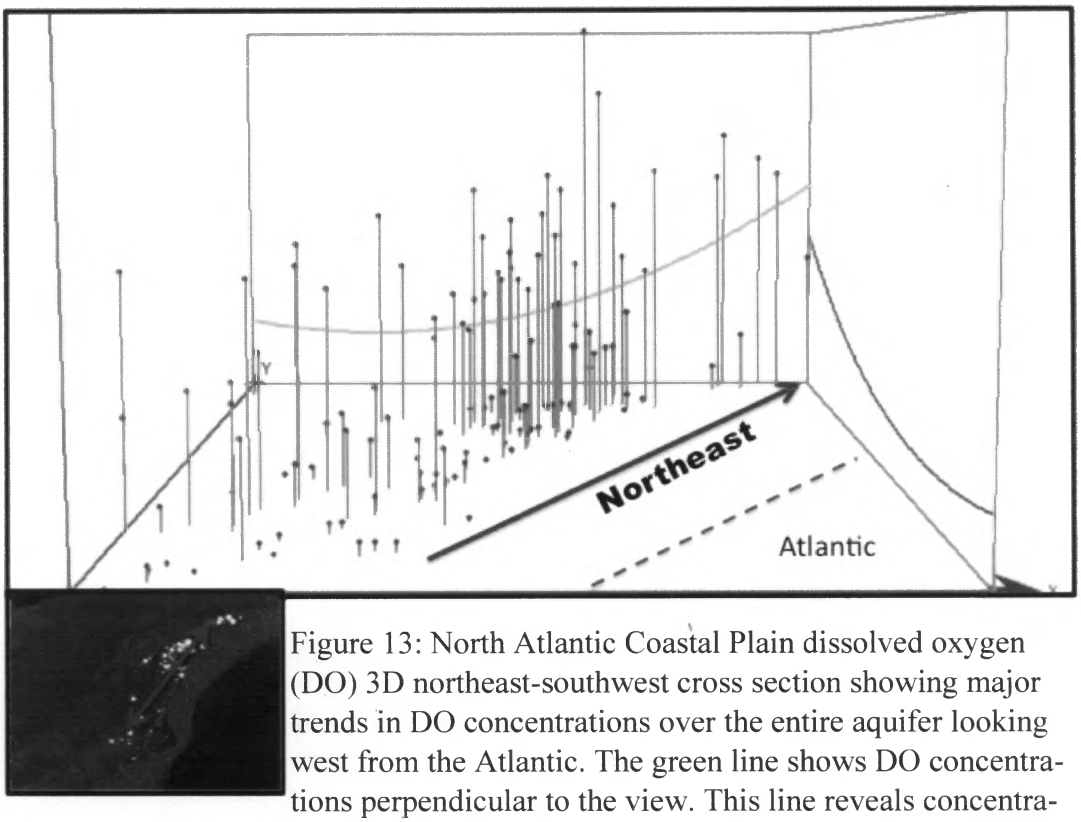


Figure 13: North Atlantic Coastal Plain dissolved oxygen (DO) 3D northeast-southwest cross section showing major trends in DO concentrations over the entire aquifer looking west from the Atlantic. The green line shows DO concentrations perpendicular to the view. This line reveals concentrations are higher in the northeastern part of the aquifer close to the area of higher precipitation and recharge.

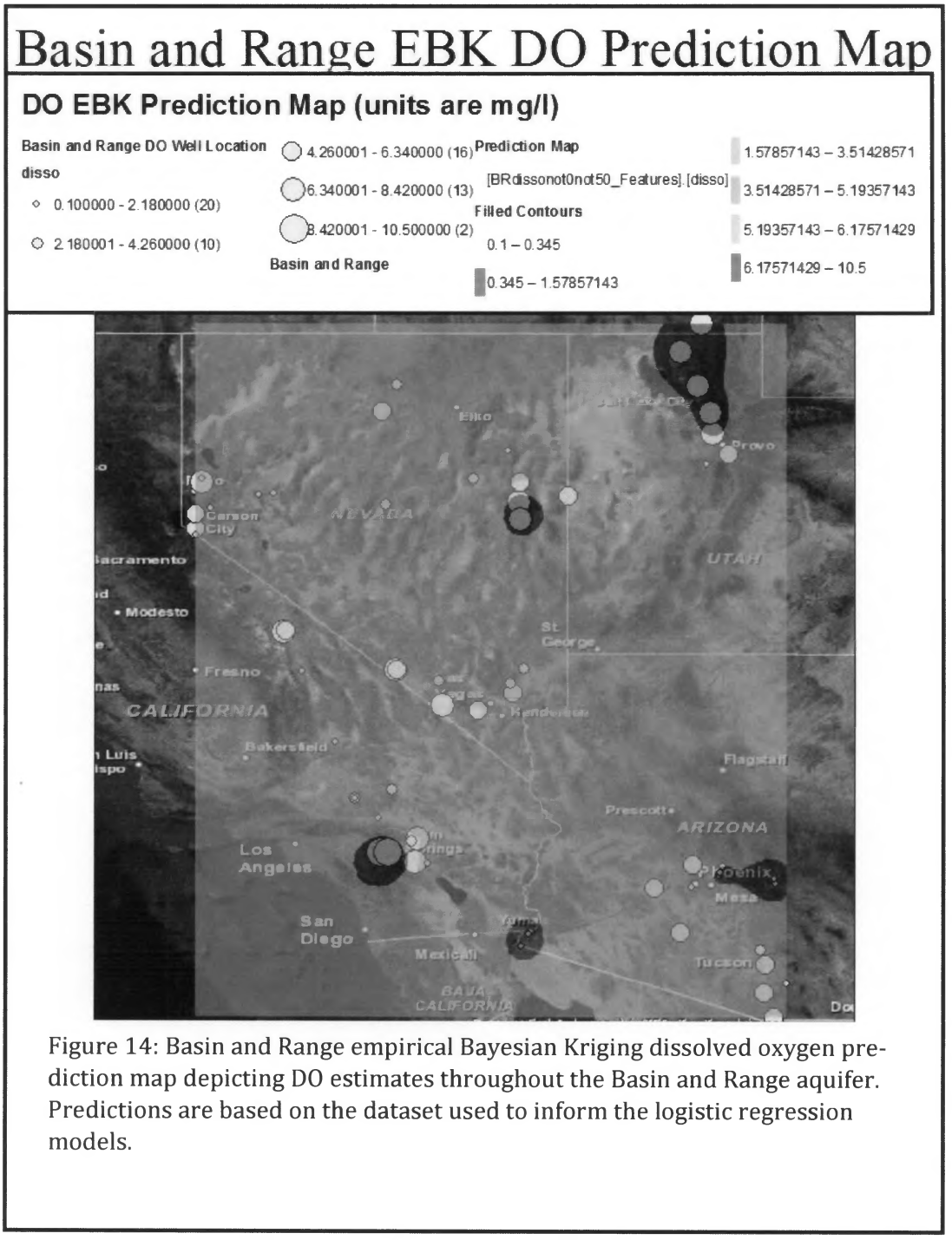


Figure 14: Basin and Range empirical Bayesian Kriging dissolved oxygen prediction map depicting DO estimates throughout the Basin and Range aquifer. Predictions are based on the dataset used to inform the logistic regression models.

Basin and Range EBK DO Standard Error Map

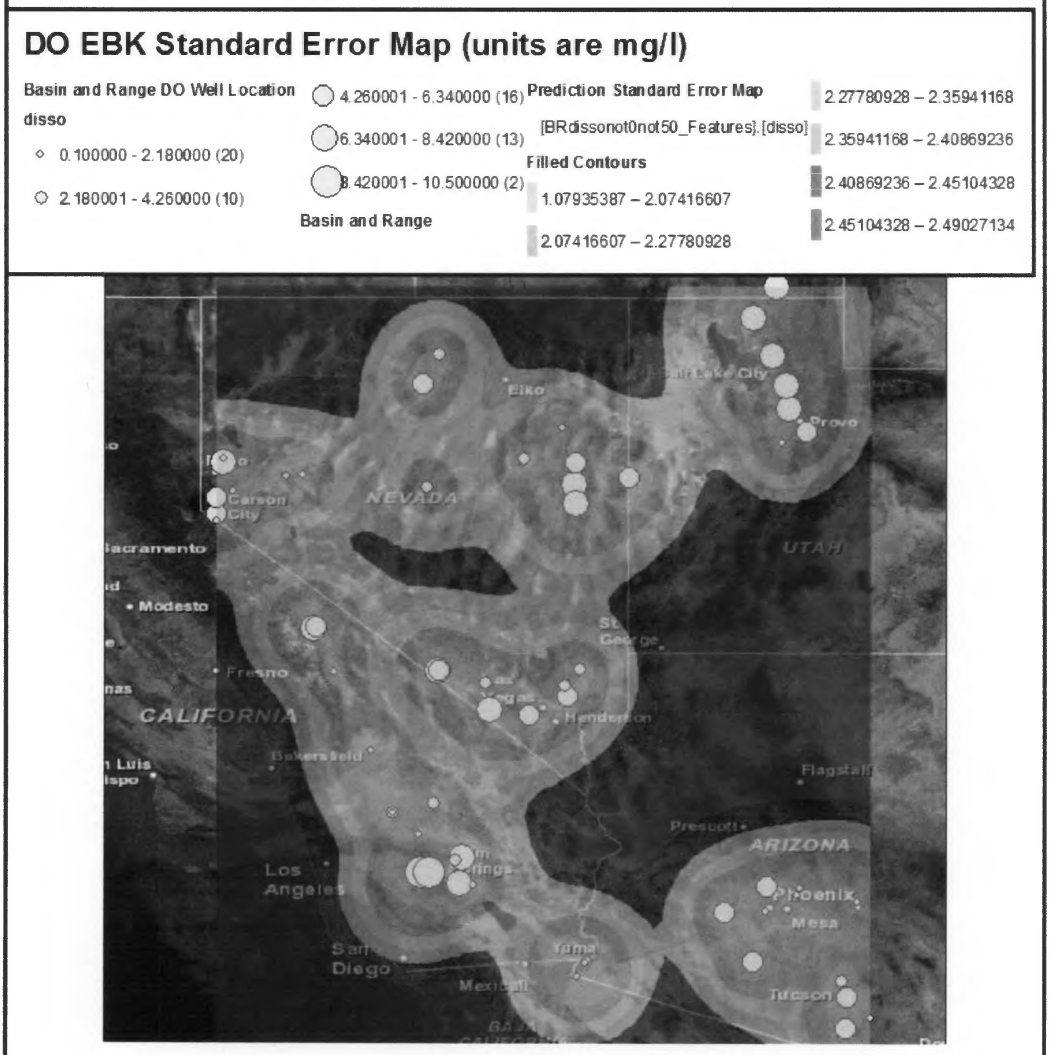


Figure 15: Basin and Range empirical Bayesian Kriging dissolved oxygen standard Error map depicting the error associated with the Basin and Range predictive map (Figure 14). DO estimates throughout the Basin and Range aquifer. The second to highest (2.45 mg/L) and second to lowest error values (2.07 mg/L) were used to define the distribution of outcomes in the uncertainty

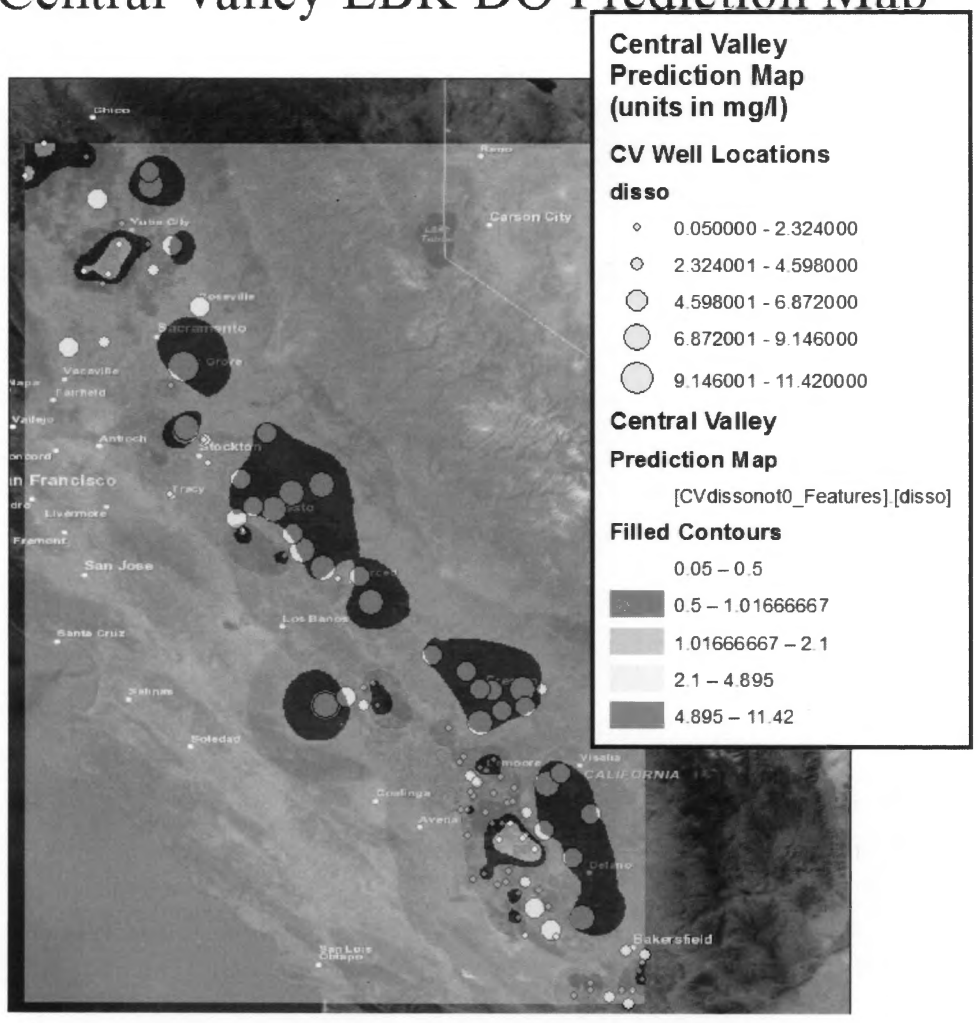


Figure 16: Central Valley (CV) empirical Bayesian Kriging dissolved oxygen (DO) prediction map depicting DO estimates throughout the CV aquifer. Predictions are based on the dataset used to inform the logistic regression models.

Central Valley EBK DO Standard Error Map

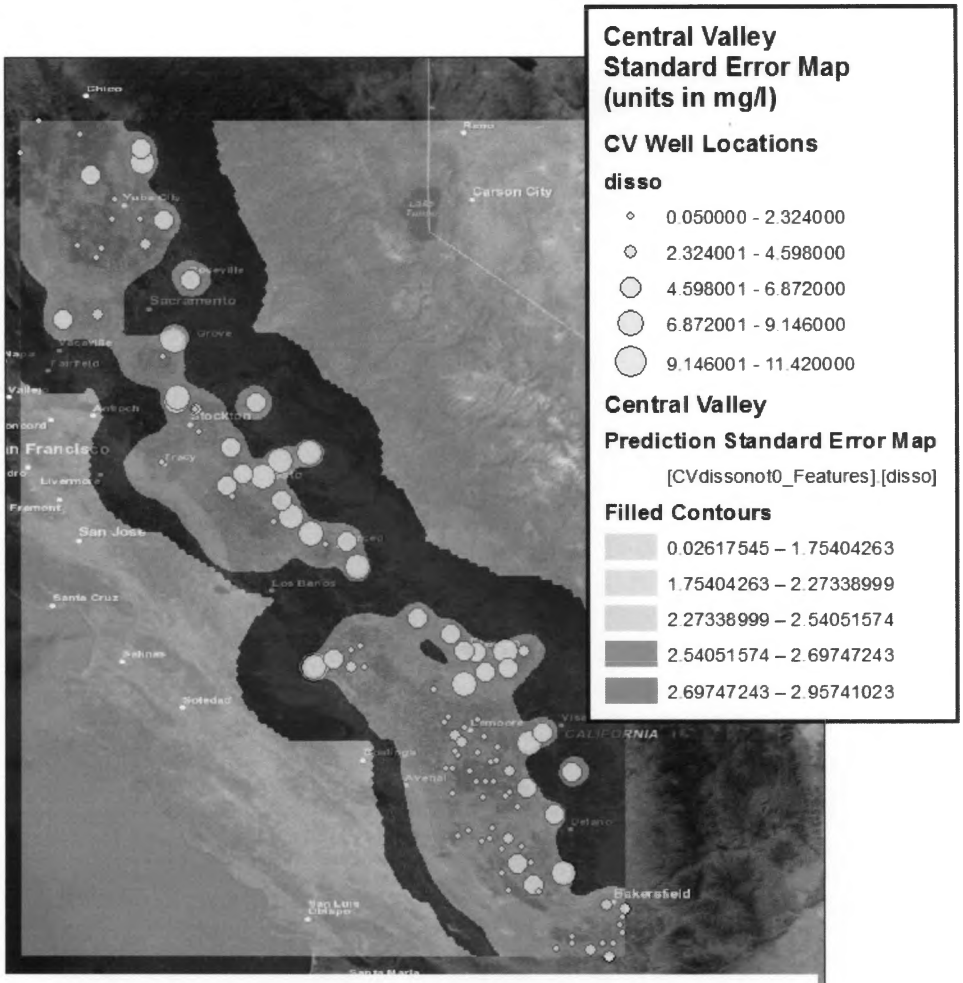


Figure 17: Central Valley (CV) empirical Bayesian Kriging dissolved oxygen (DO) standard error map depicting the error associated with the CV predictive map (Figure 16). DO estimates throughout the CV aquifer. The second to highest (2.70 mg/L) and second to lowest error values (1.75 mg/L) were used to define the uncertainty models.

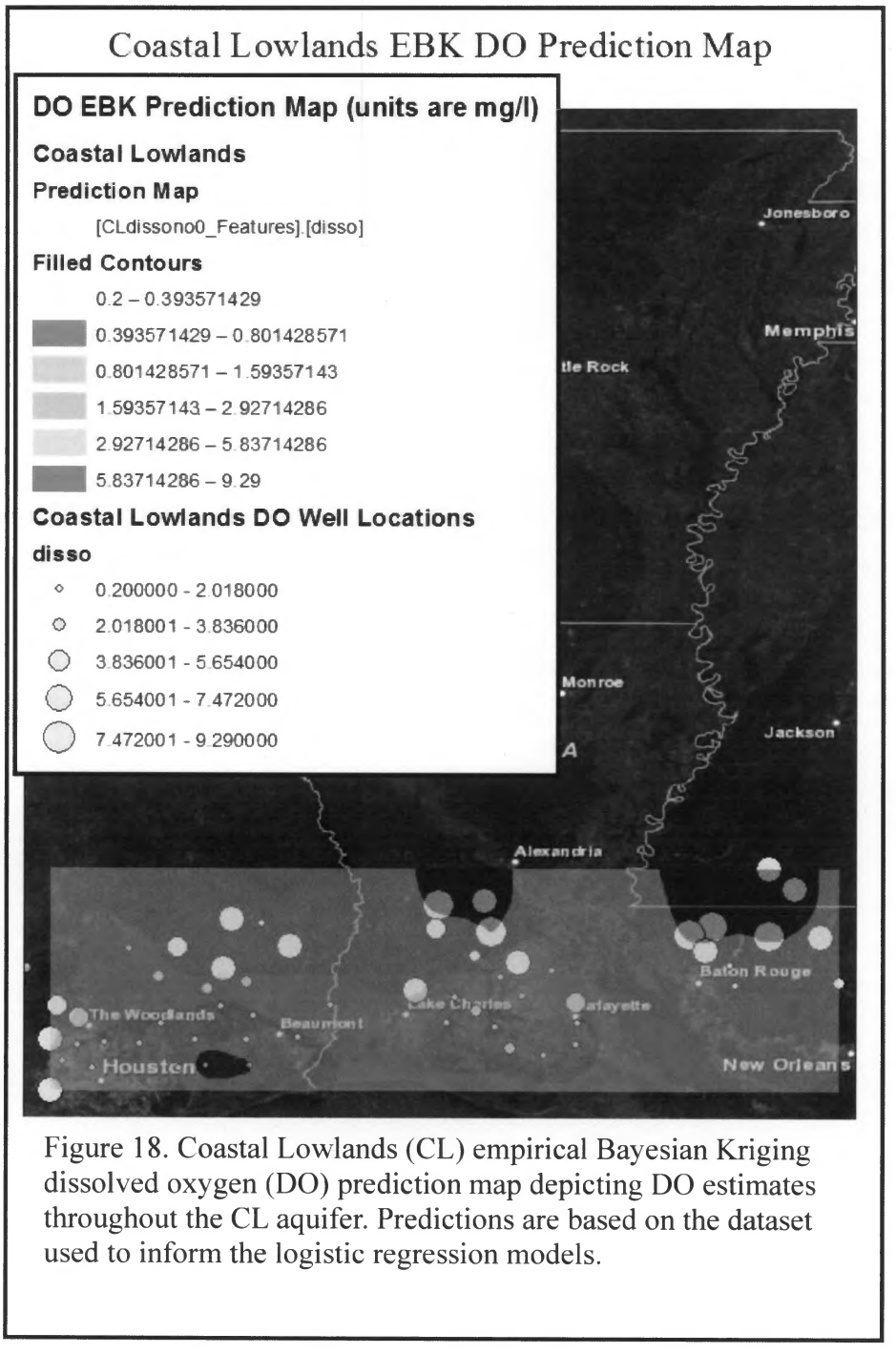


Figure 18. Coastal Lowlands (CL) empirical Bayesian Kriging dissolved oxygen (DO) prediction map depicting DO estimates throughout the CL aquifer. Predictions are based on the dataset used to inform the logistic regression models.

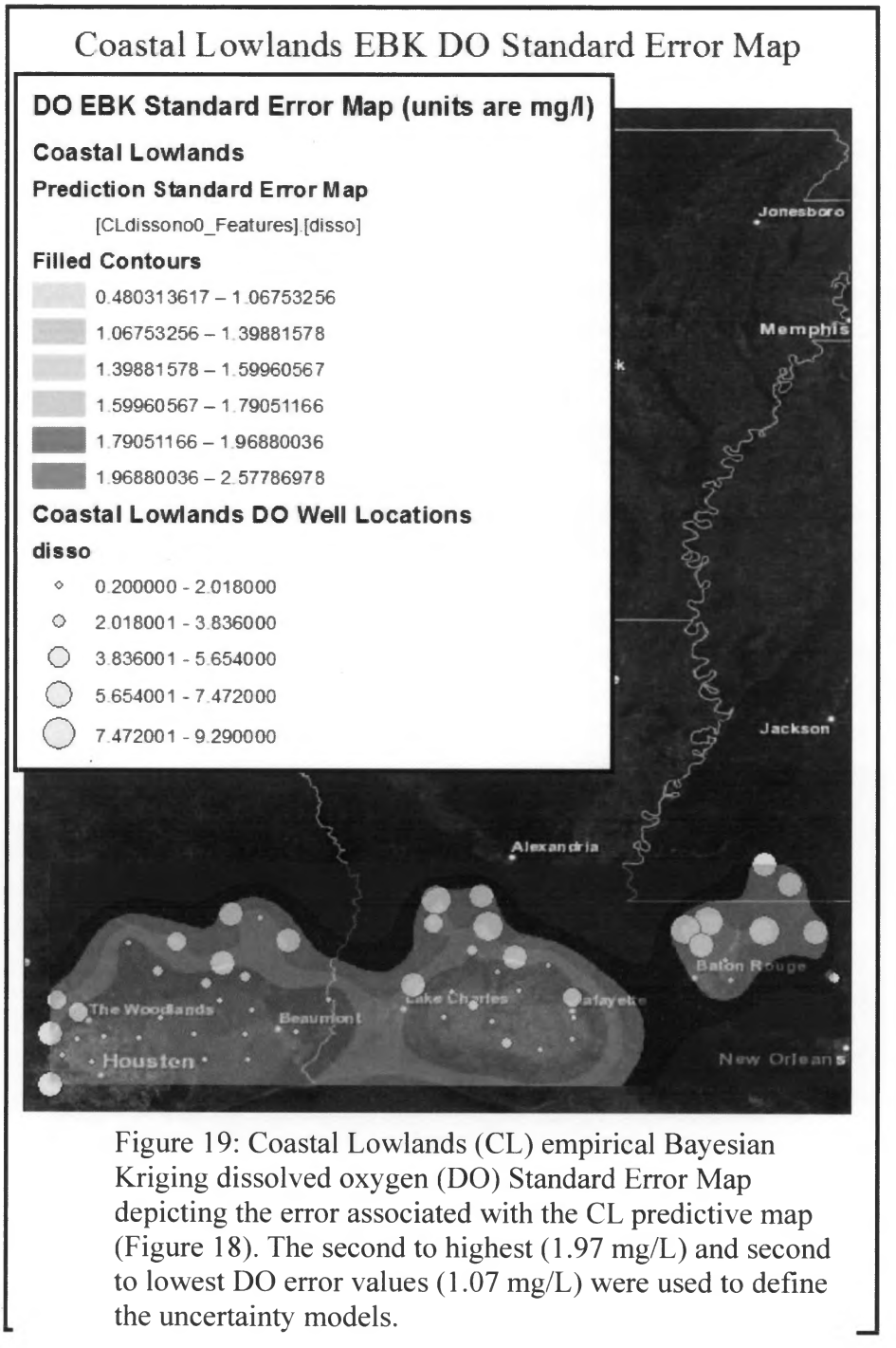


Figure 19: Coastal Lowlands (CL) empirical Bayesian Kriging dissolved oxygen (DO) Standard Error Map depicting the error associated with the CL predictive map (Figure 18). The second to highest (1.97 mg/L) and second to lowest DO error values (1.07 mg/L) were used to define the uncertainty models.

North Atlantic Coastal Plain EBK DO Prediction Map

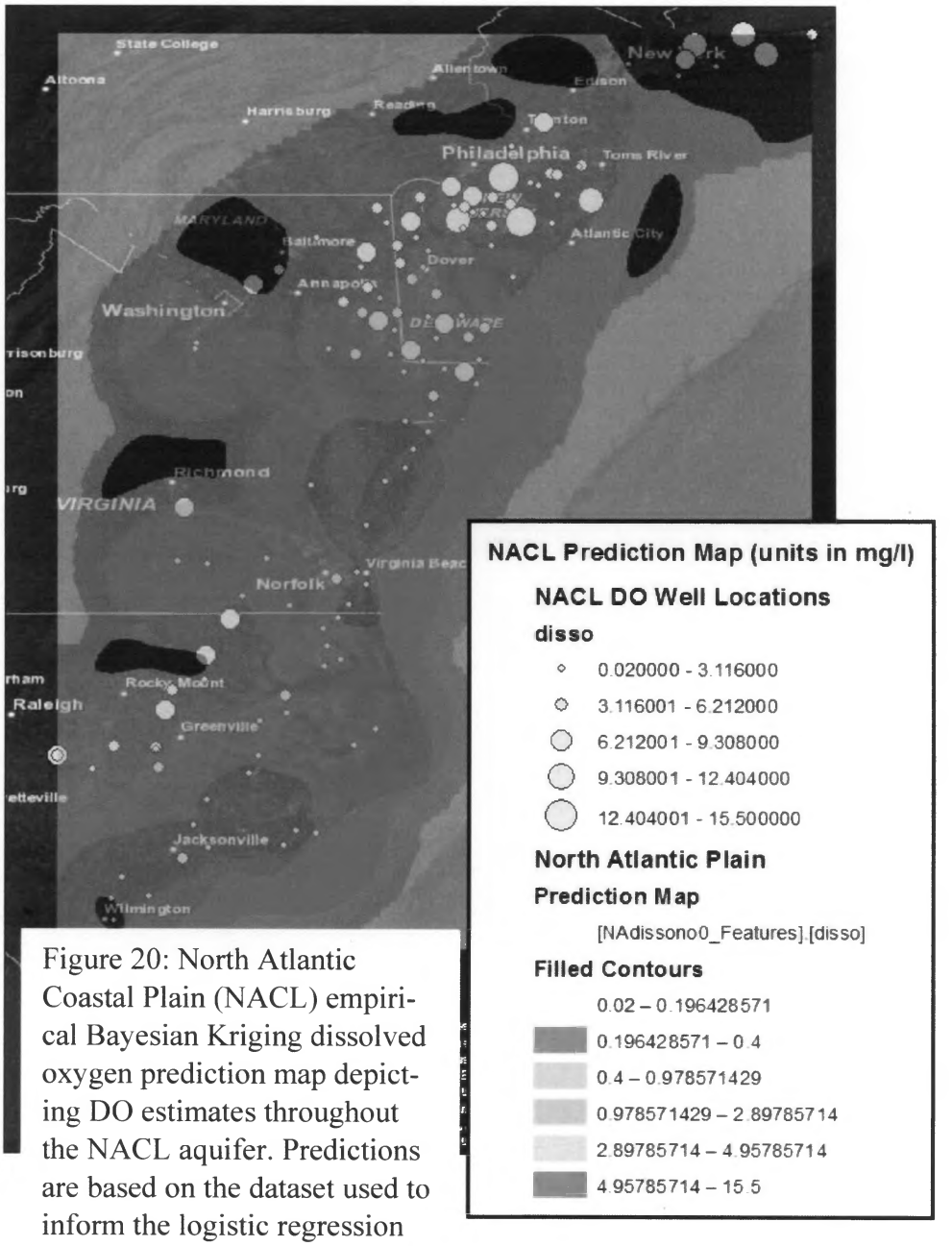


Figure 20: North Atlantic Coastal Plain (NACL) empirical Bayesian Kriging dissolved oxygen prediction map depicting DO estimates throughout the NACL aquifer. Predictions are based on the dataset used to inform the logistic regression models.

North Atlantic Coastal Plain EBK DO Standard Error Map

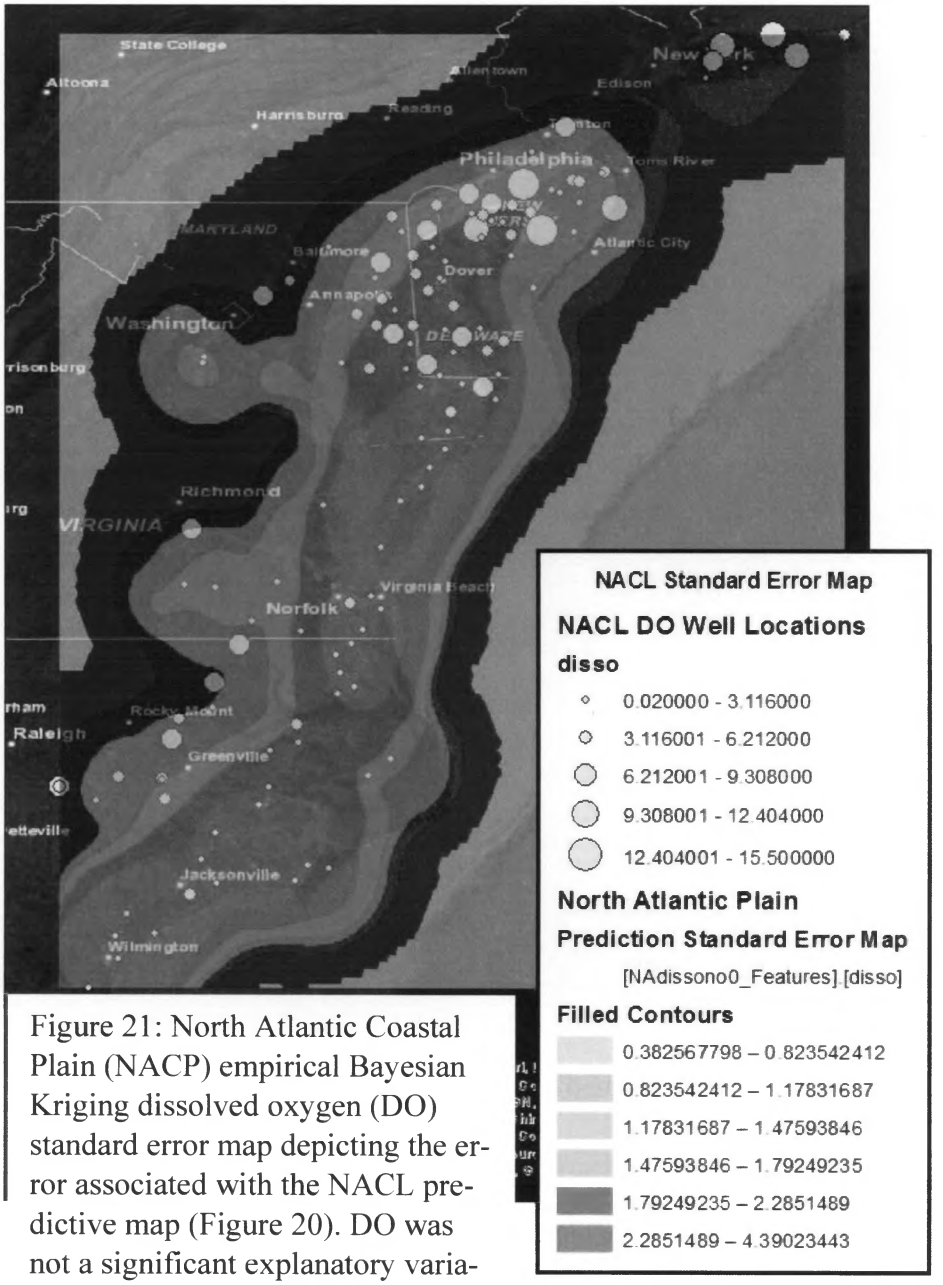


Figure 21: North Atlantic Coastal Plain (NACP) empirical Bayesian Kriging dissolved oxygen (DO) standard error map depicting the error associated with the NACL predictive map (Figure 20). DO was not a significant explanatory variable used to inform the model so standard error was not needed.

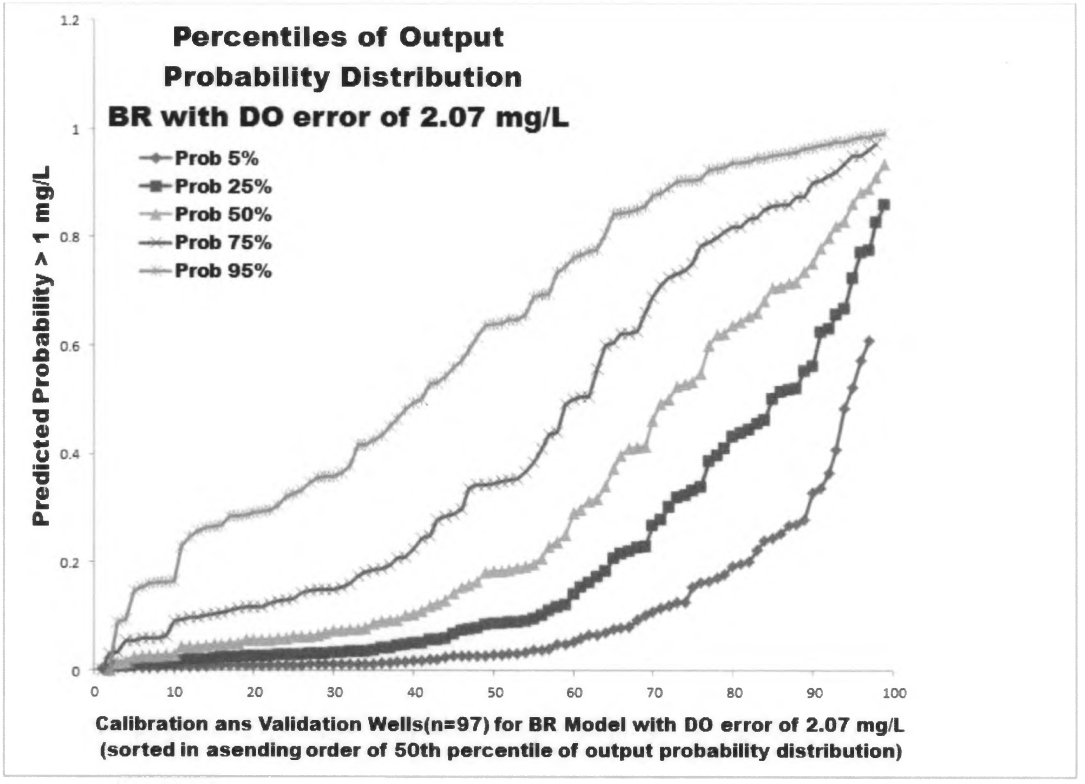


Figure 22: Percentile output probability distribution plots for 5^{th}, 25^{th}, 50^{th}, 75^{th}, 95^{th} percentiles of the uncertainty model for the Basin and Range with dissolved oxygen error of 2.07 mg/L.

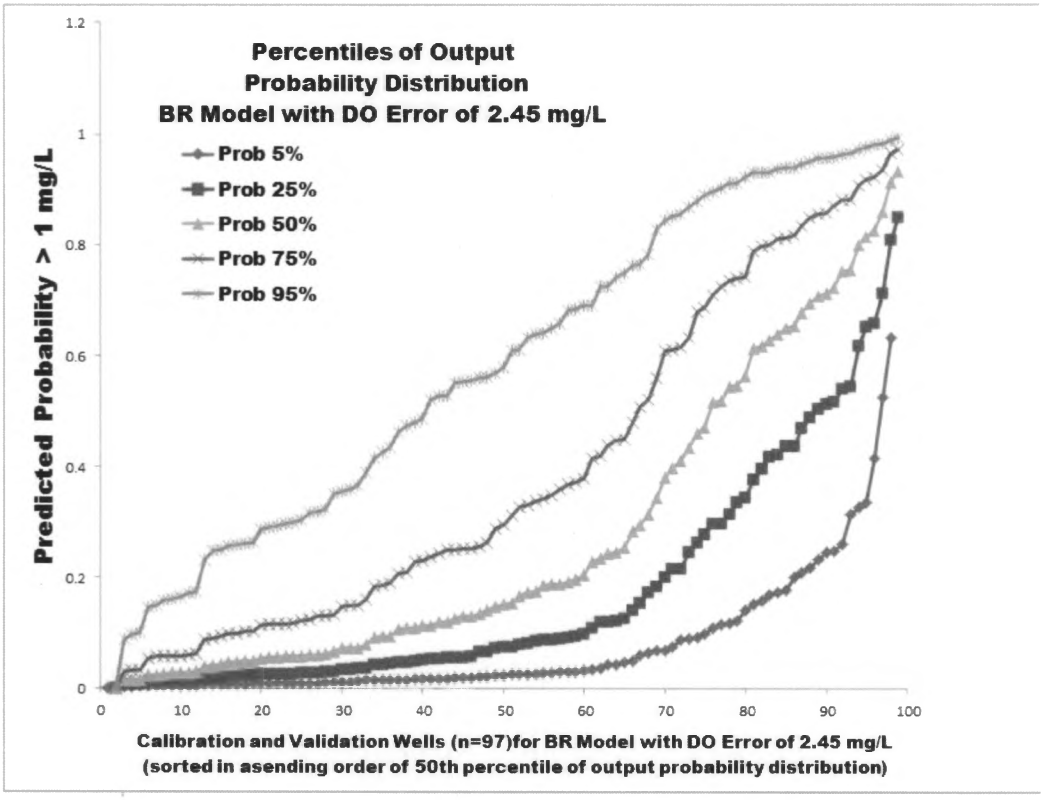


Figure 23: Percentile output probability distribution plots for 5^{th}, 25^{th}, 50^{th}, 75^{th}, 95^{th} percentiles of the uncertainty model for the Basin and Range with dissolved oxygen error of 2.45 mg/

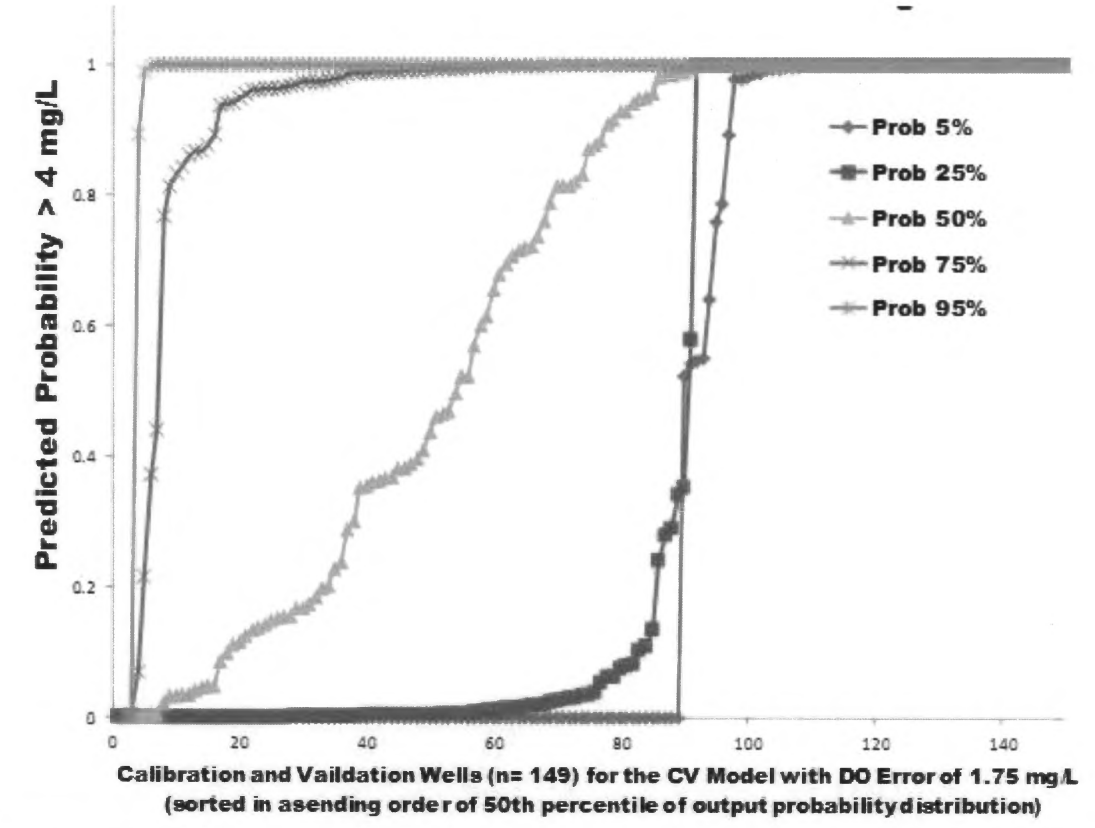


Figure 24: Percentile output probability distribution plots for 5th, 25th, 50th, 75th, 95th percentiles of the uncertainty model for the Central Valley with dissolved oxygen error of 1.75 mg/L.

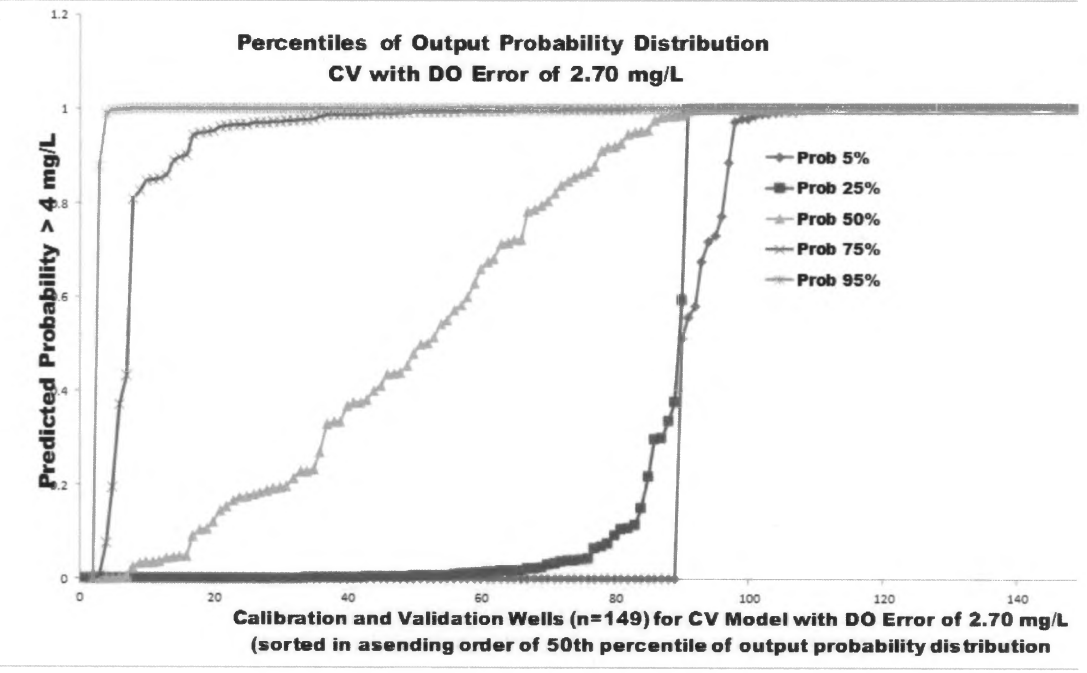


Figure 25: Percentile output probability distribution plots for 5th, 25th, 50th, 75th, 95th percentiles of the uncertainty model for the Central Valley with dissolved oxygen error of 2.70 mg/L.

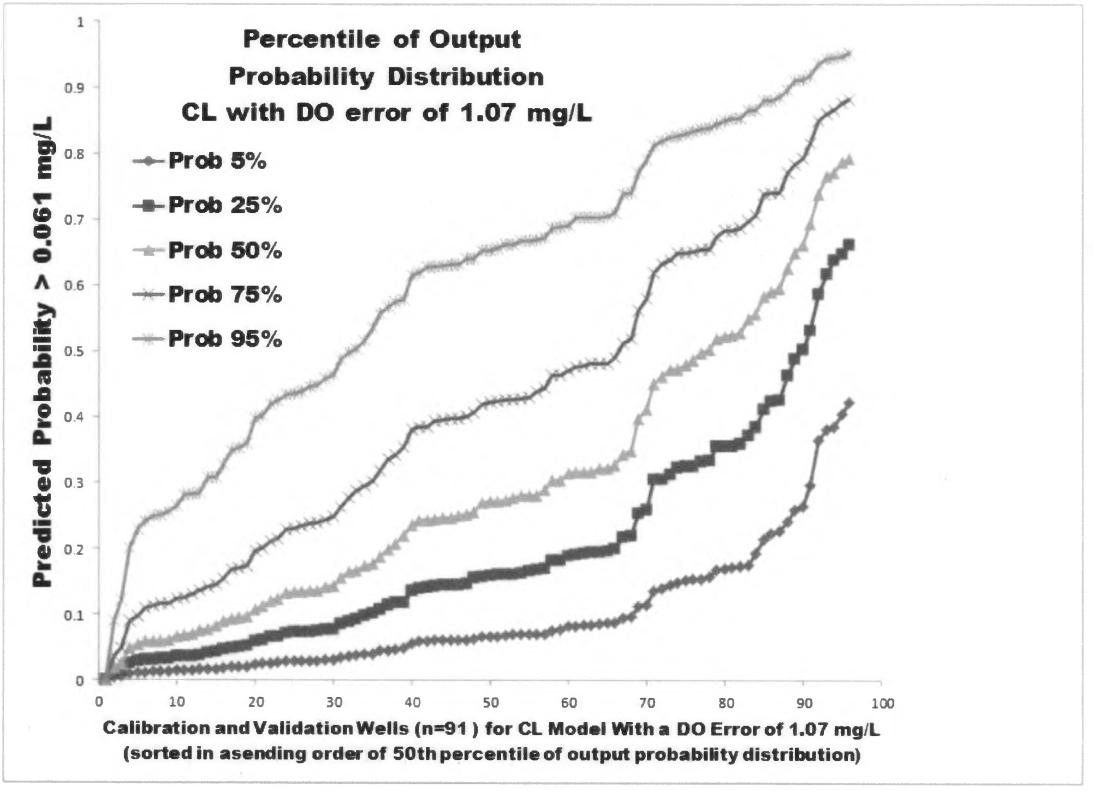


Figure 26: Percentile output probability distribution plots for 5th, 25th, 50th, 75th, 95th percentiles of the uncertainty model for the Coastal Lowlands with dissolved oxygen error of 1.07 mg/L.

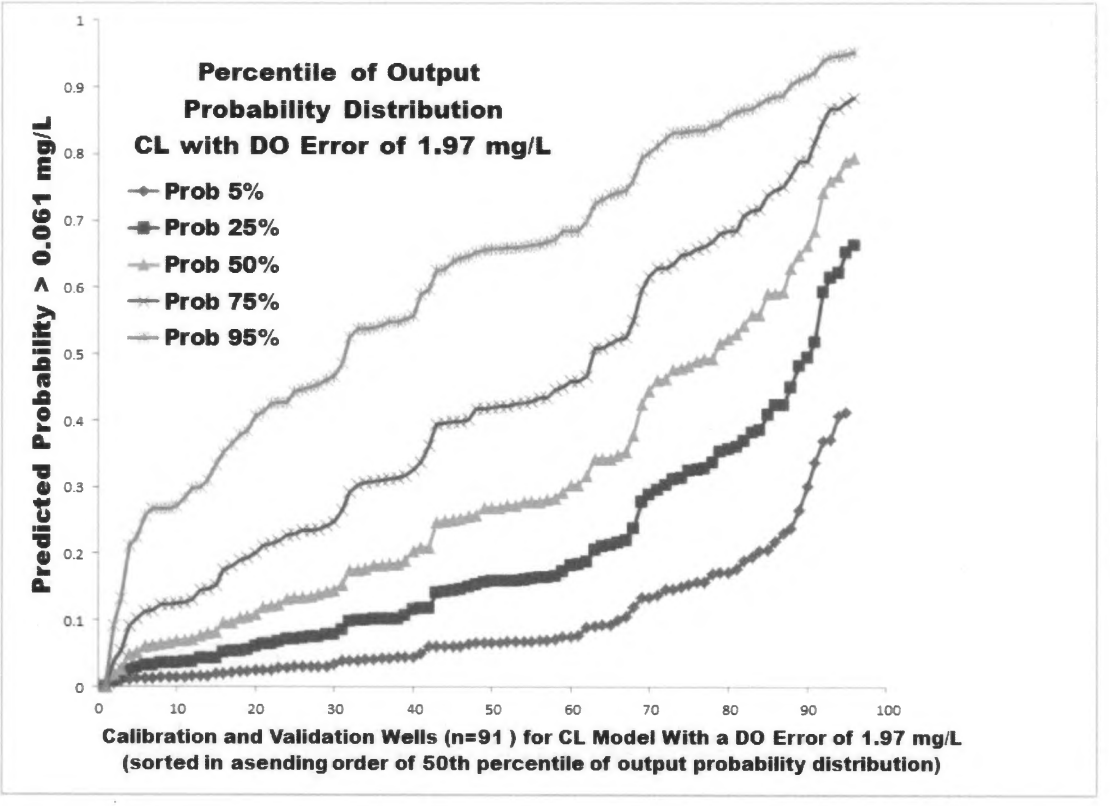


Figure 27: Percentile output probability distribution plots for 5th, 25th, 50th, 75th, 95th percentiles of the uncertainty model for the Coastal Lowlands with dissolved oxygen error of 1.97 mg/L.

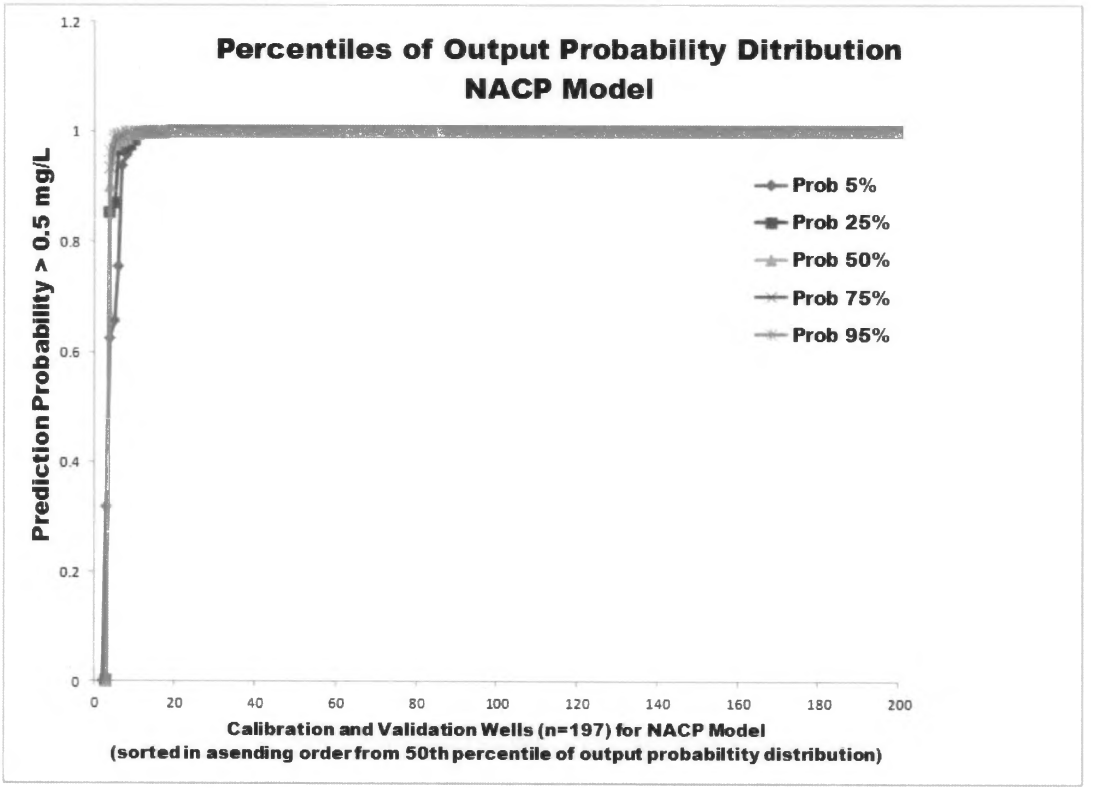


Figure 28: Percentile output probability distribution plots for 5th, 25th, 50th, 75th, 95th percentiles of the uncertainty model for the North Atlantic Coastal Plain.

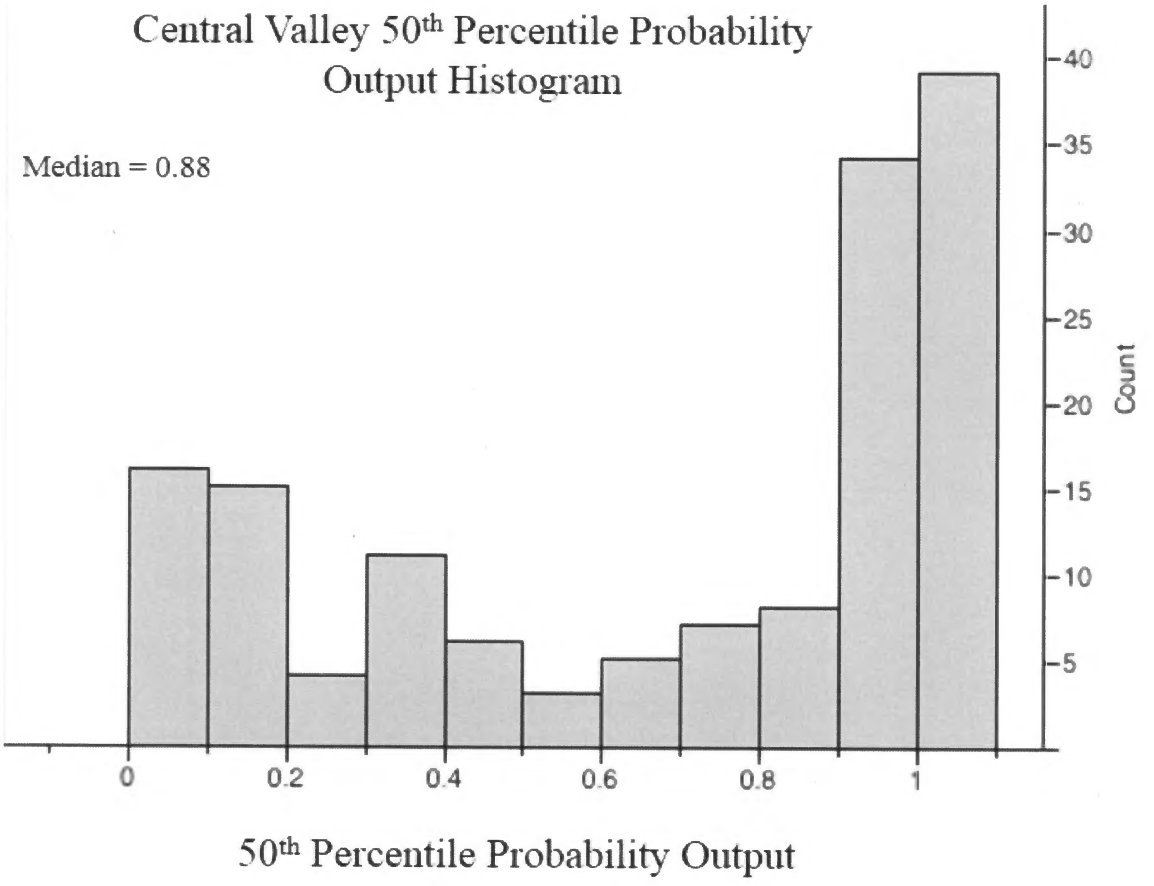


Figure 29: Central Valley histogram of 50% percentile probability output showing the distribution has a bimodal character with probabilities clustered near 0 and 1.

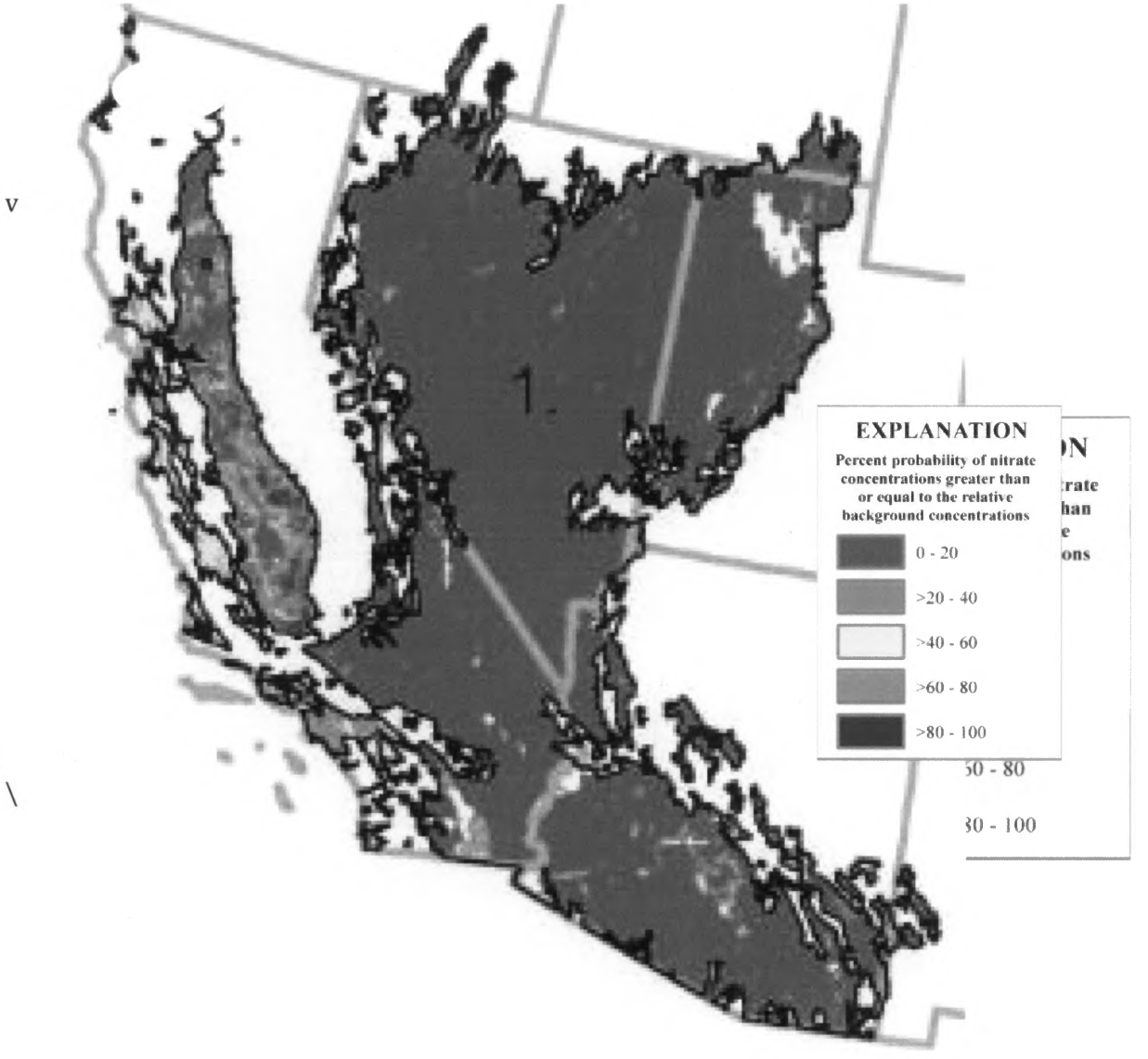


Figure 30: Central Valley and Basin and Range principal aquifer logistic regression prediction maps from Gurdak and Qi (2012) showingt he percent probability of nitrate concentrations being greater than background.

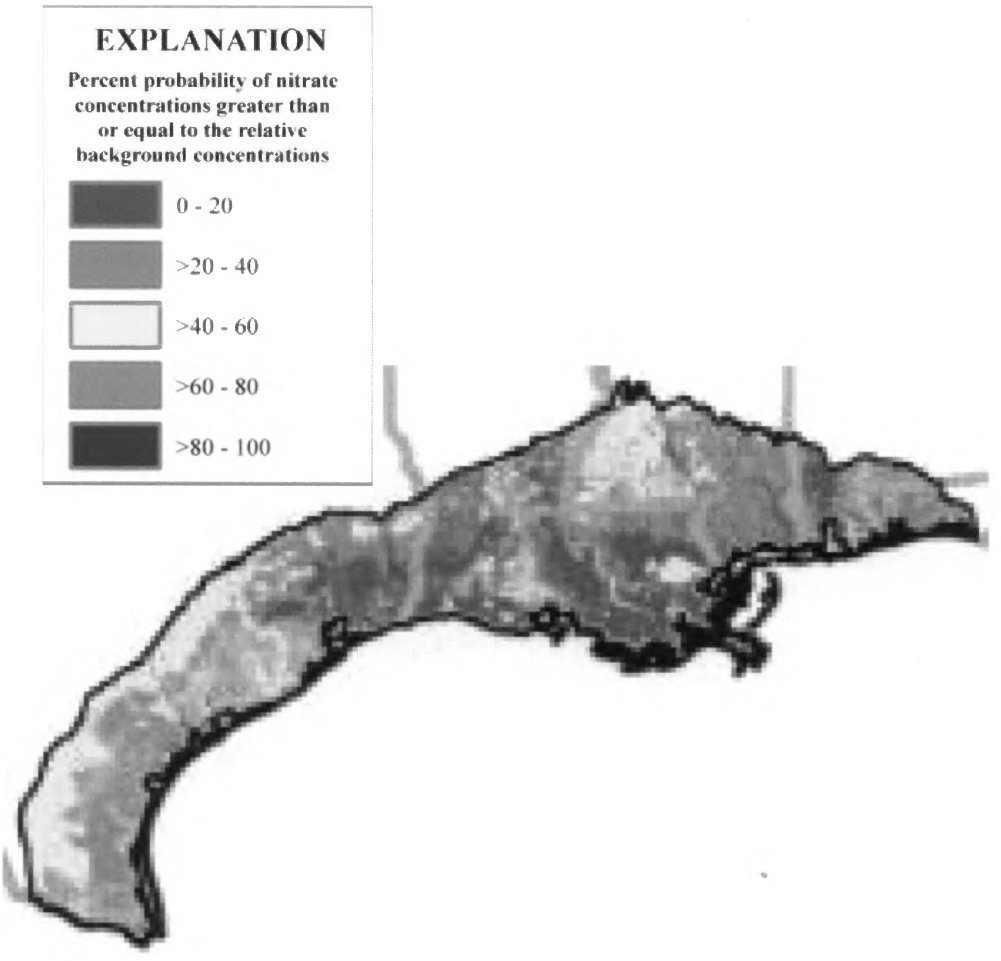


Figure 31: Coastal Lowland a logistic regression prediction map from Gurdak and Qi (2012) showing the percent probability of nitrate concentrations being greater than background.

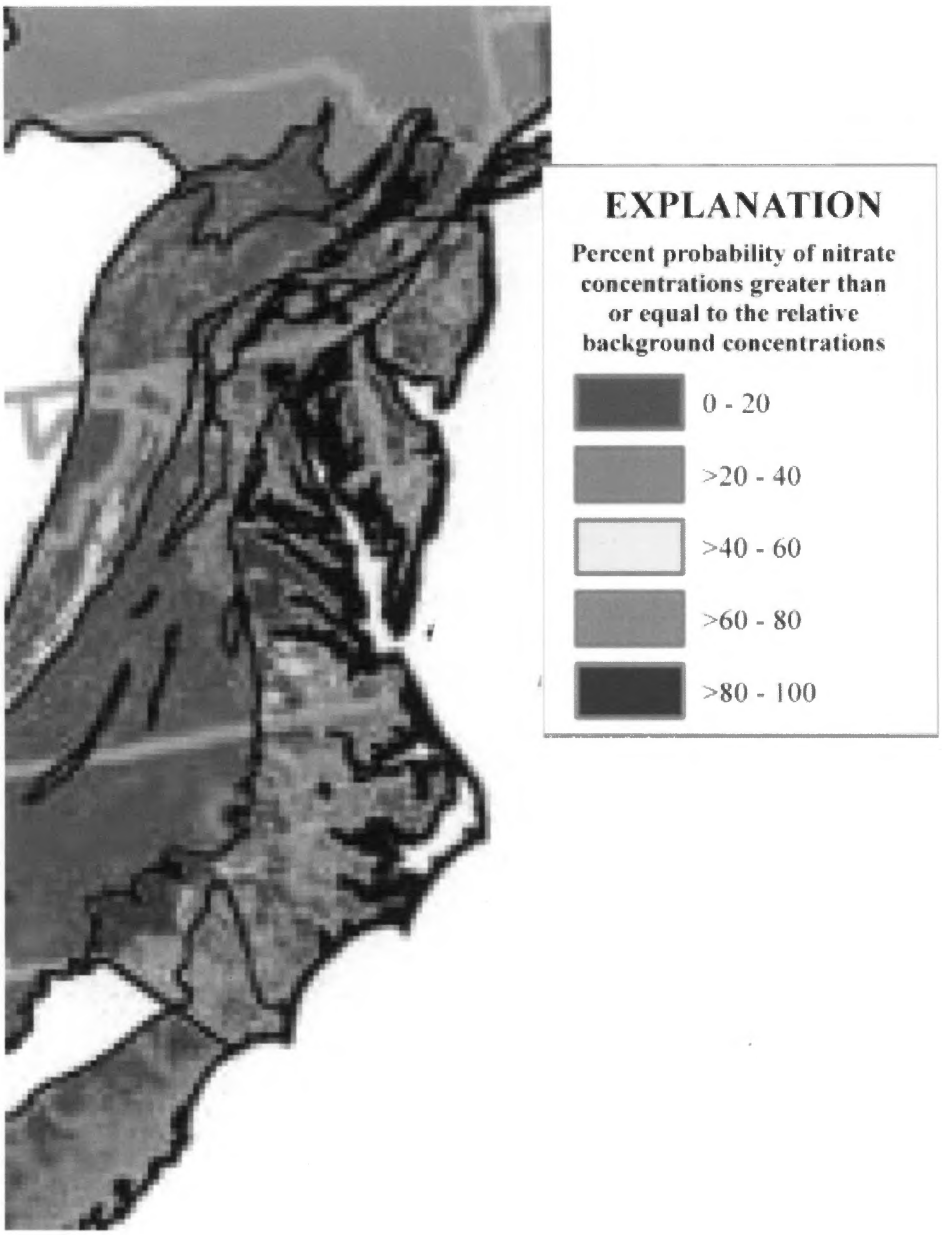


Figure 32: North Atlantic Coastal Plain logistic regression prediction map from Gurdak and Qi (2012) showing the percent probability of nitrate concentrations being greater than background.

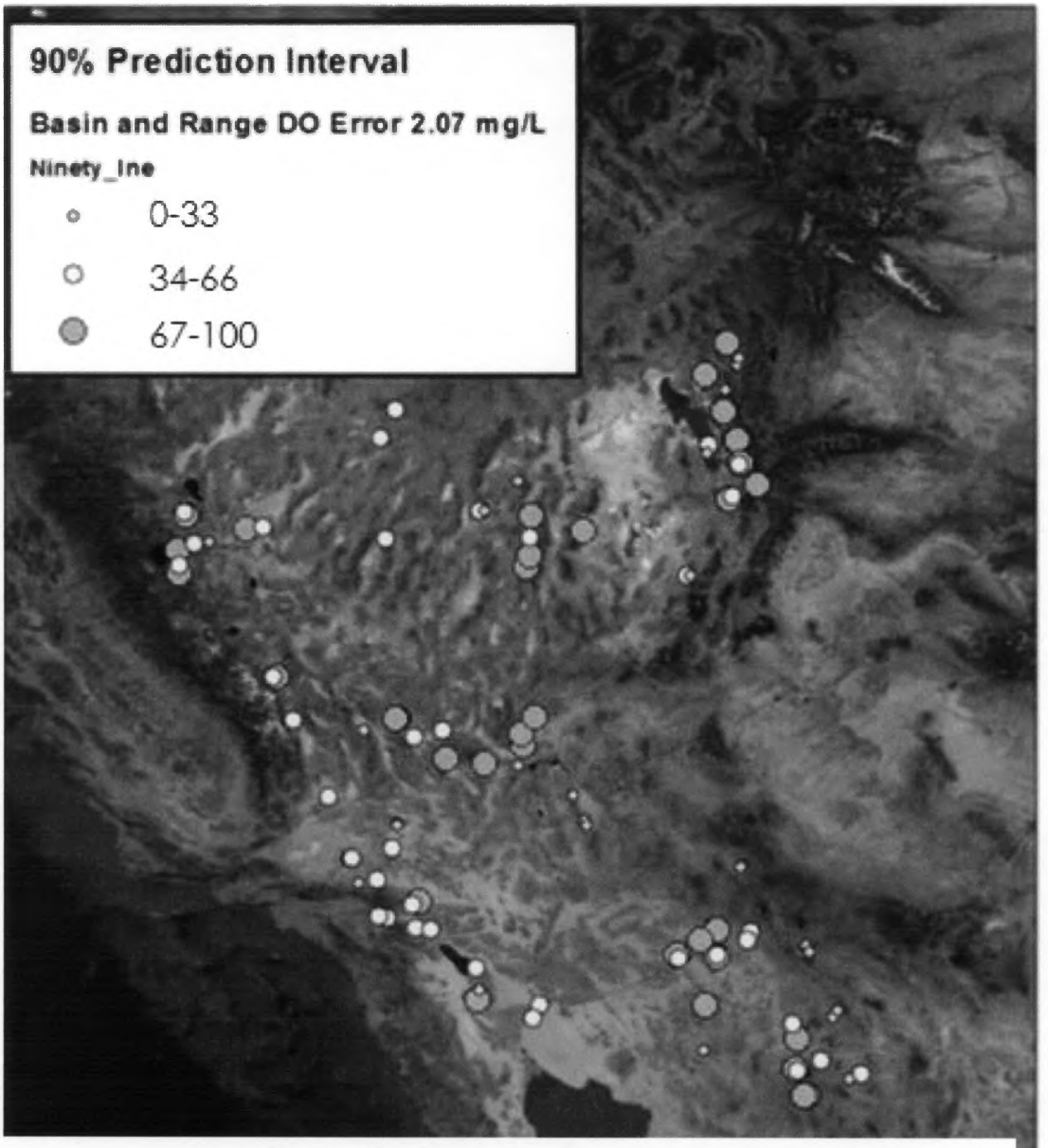


Figure 33: Basin and Range 90th percentile prediction interval map, with dissolved oxygen error of 2.07 mg/L, measuring the difference between the 5th and 95th percentile probability predictions. The green dots represent good agreement, yellow represents moderate agreement, and orange poor agreement between the original logistic regression model predictions and uncertainty model predictions.

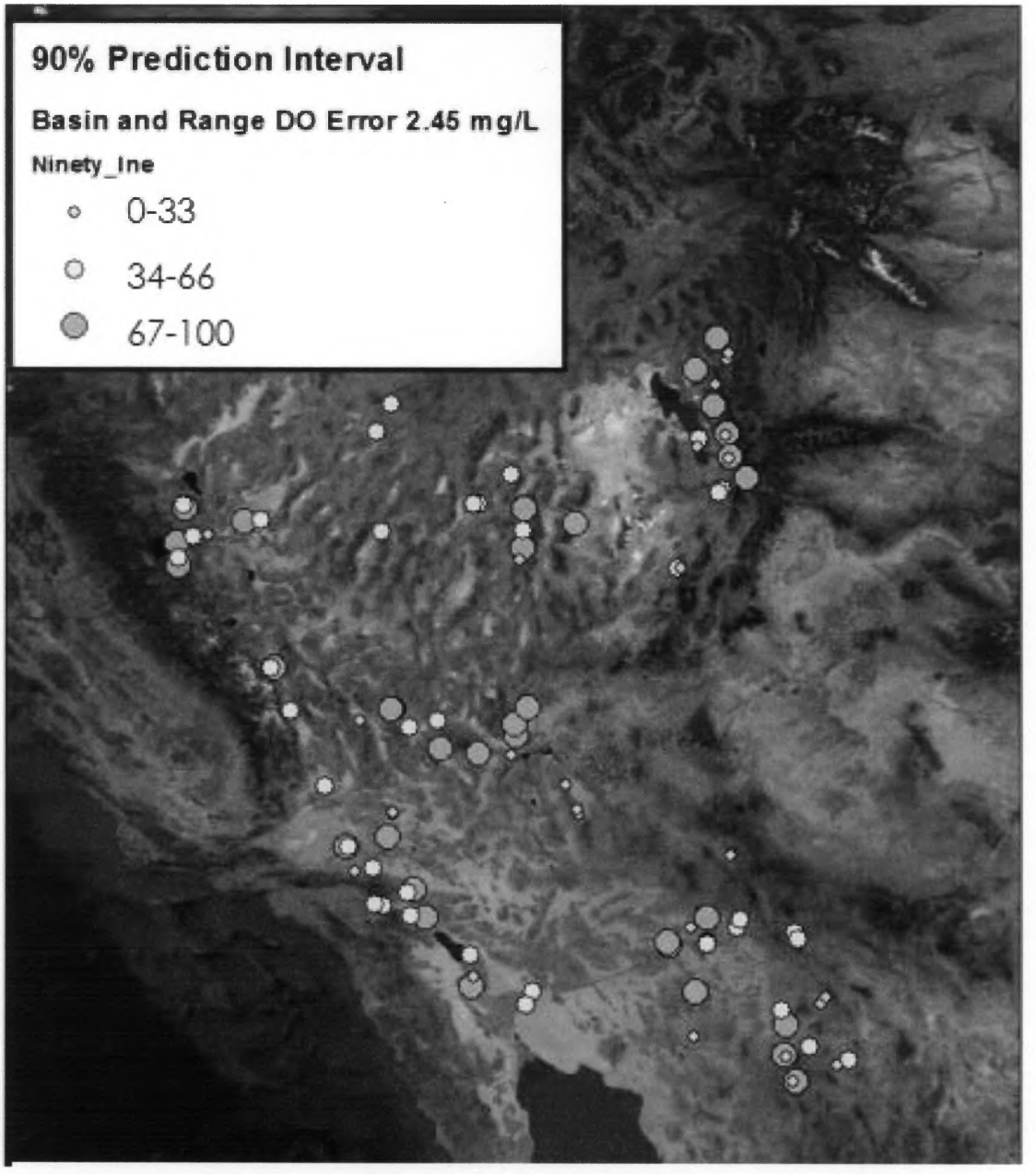


Figure 33: Basin and Range 90th percentile prediction interval map, with dissolved oxygen error of 2.45 mg/L, measuring the difference between the 5th and 95th percentile probability predictions. The green dots represent good agreement, yellow represents moderate agreement, and orange poor agreement between the original logistic regression model predictions and uncertainty model predictions.

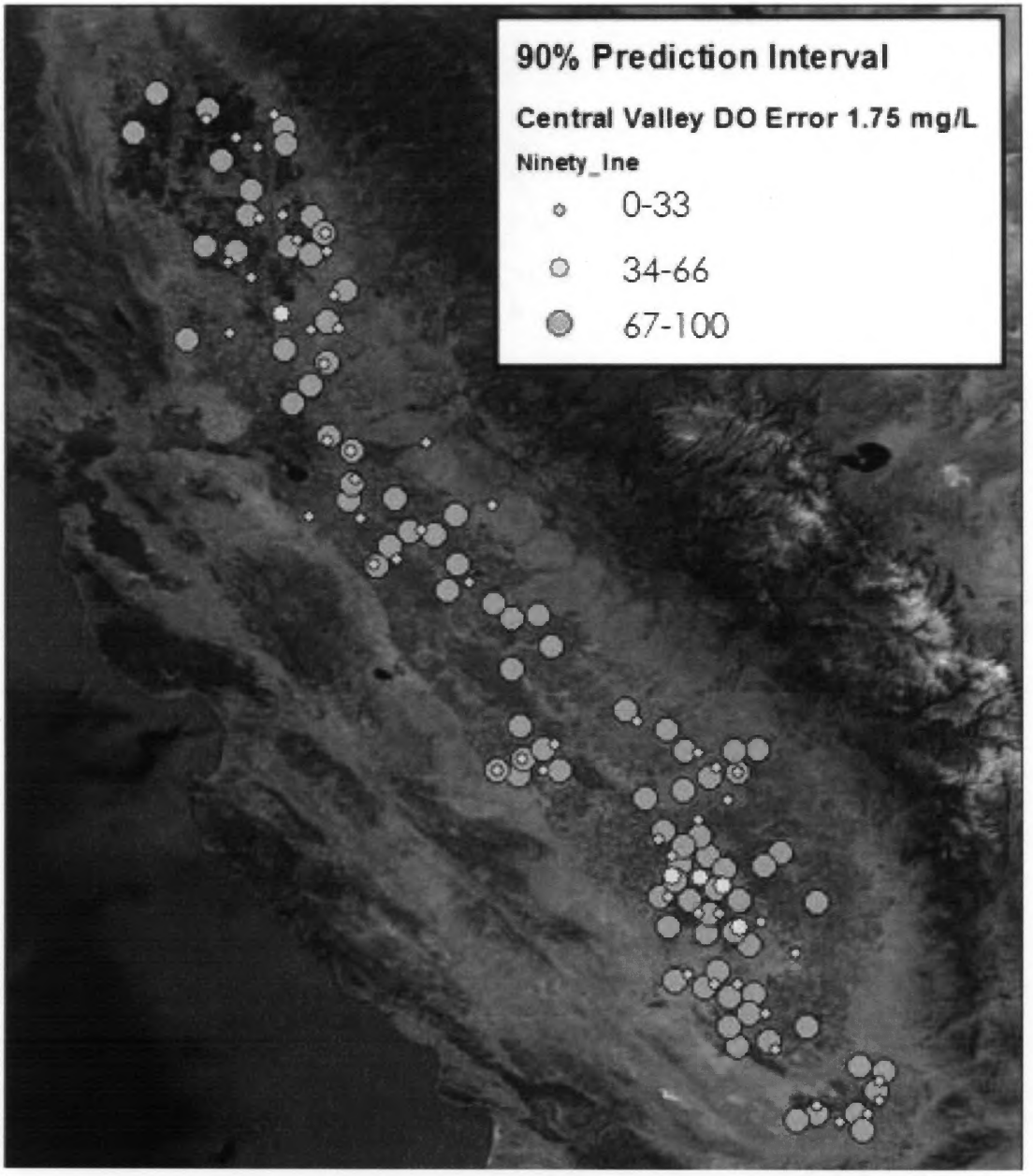


Figure 35: Central Valley 90th percentile prediction interval map, with dissolved oxygen error of 1.75 mg/L, measuring the difference between the 5th and 95th percentile probability predictions. The green dots represent good agreement, yellow represents moderate agreement, and orange poor agreement between the original logistic regression model predictions and uncertainty model predictions.

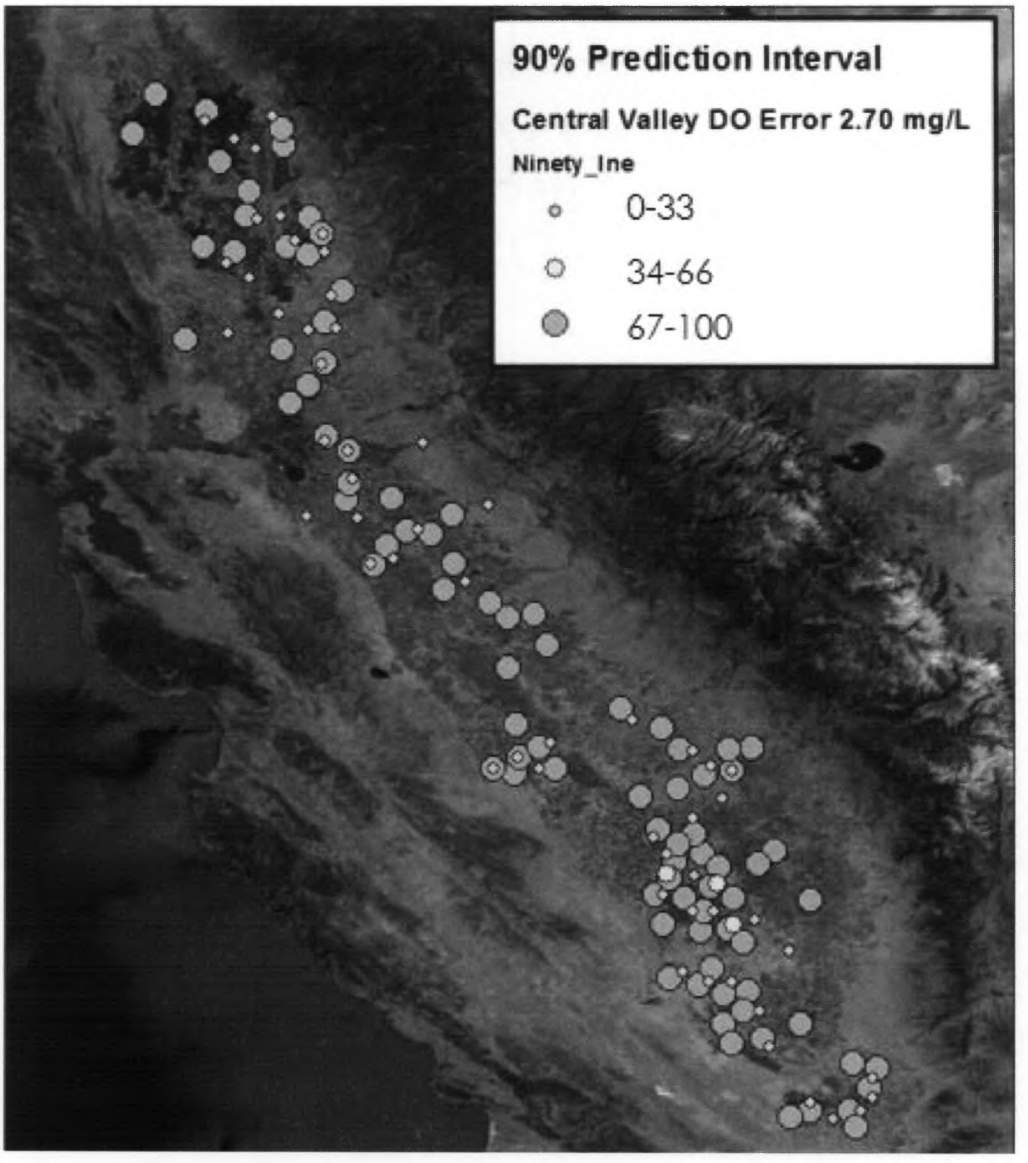


Figure 36: Central Valley 90th percentile prediction interval map, with dissolved oxygen error of 2.70 mg/L, measuring the difference between the 5th and 95th percentile probability predictions. The green dots represent good agreement, yellow represents moderate agreement, and orange poor agreement between the original logistic regression model predictions and uncertainty model predictions.

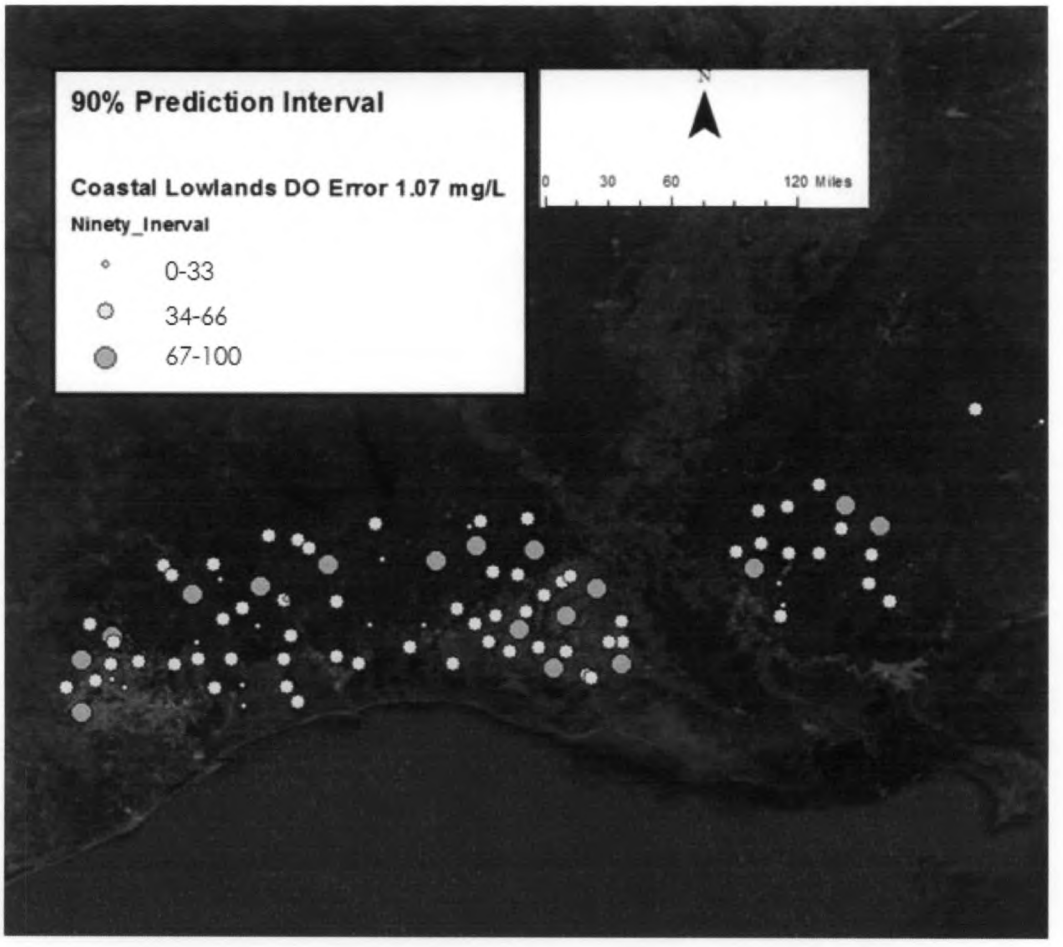


Figure 37: Coastal Lowlands 90th percentile prediction interval map, with dissolved oxygen error of 1.07 mg/L, measuring the difference between the 5th and 95th percentile probability predictions. The green dots represent good agreement, yellow represents moderate agreement, and orange poor agreement between the original logistic regression model predictions and uncertainty model predictions.

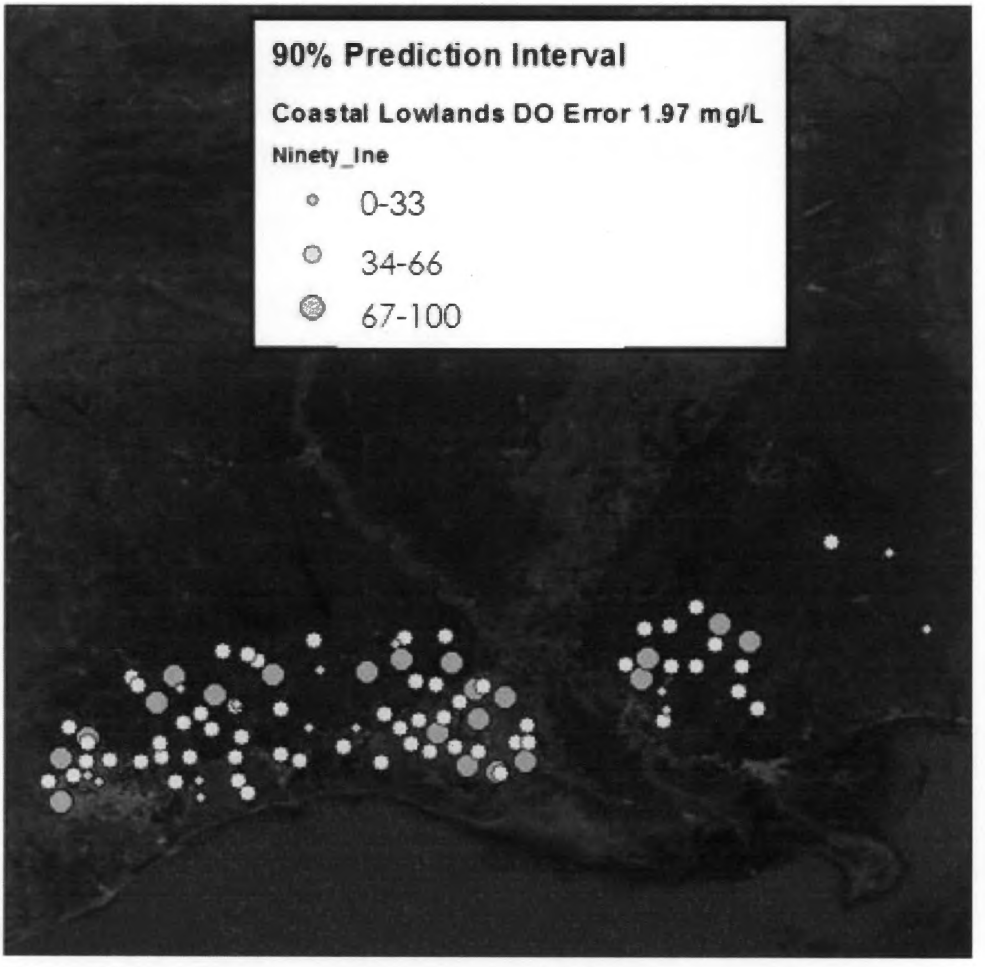


Figure 39: Coastal Lowlands 90th percentile prediction interval map, with dissolved oxygen error of 1.97 mg/L, measuring the difference between the 5th and 95th percentile probability predictions. The green dots represent good agreement, yellow represents moderate agreement, and orange poor agreement between the original logistic regression model predictions and uncertainty model predictions.

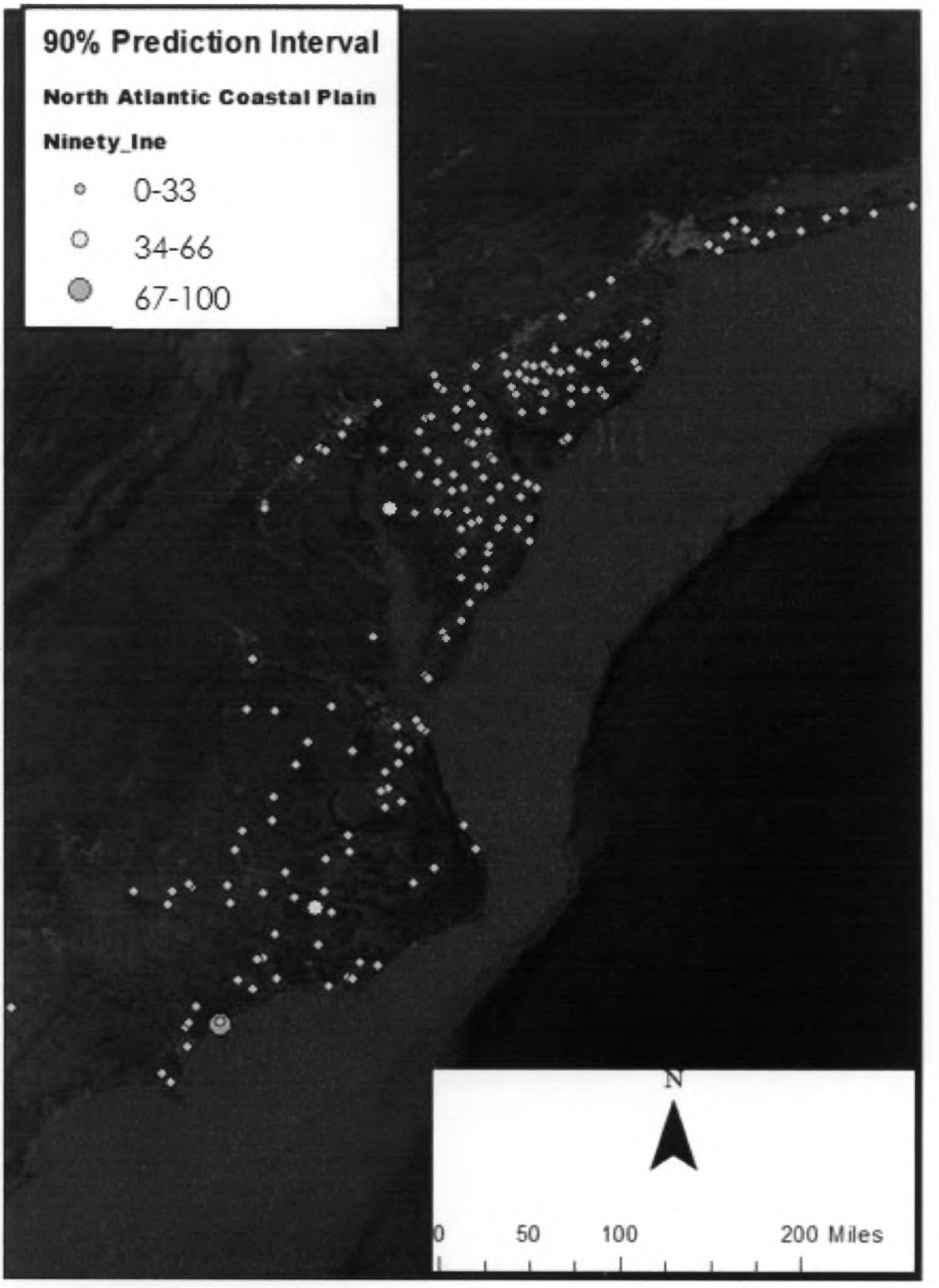


Figure 39: North Atlantic Coastal Plain 90th percentile measuring the difference between the 5th and 95th percentile probability predictions. The green dots represent good agreement, yellow represents moderate agreement, and orange poor agreement between the original logistic regression model predictions and uncertainty model predictions.

Comparison of Relative Variance Contribution by Location

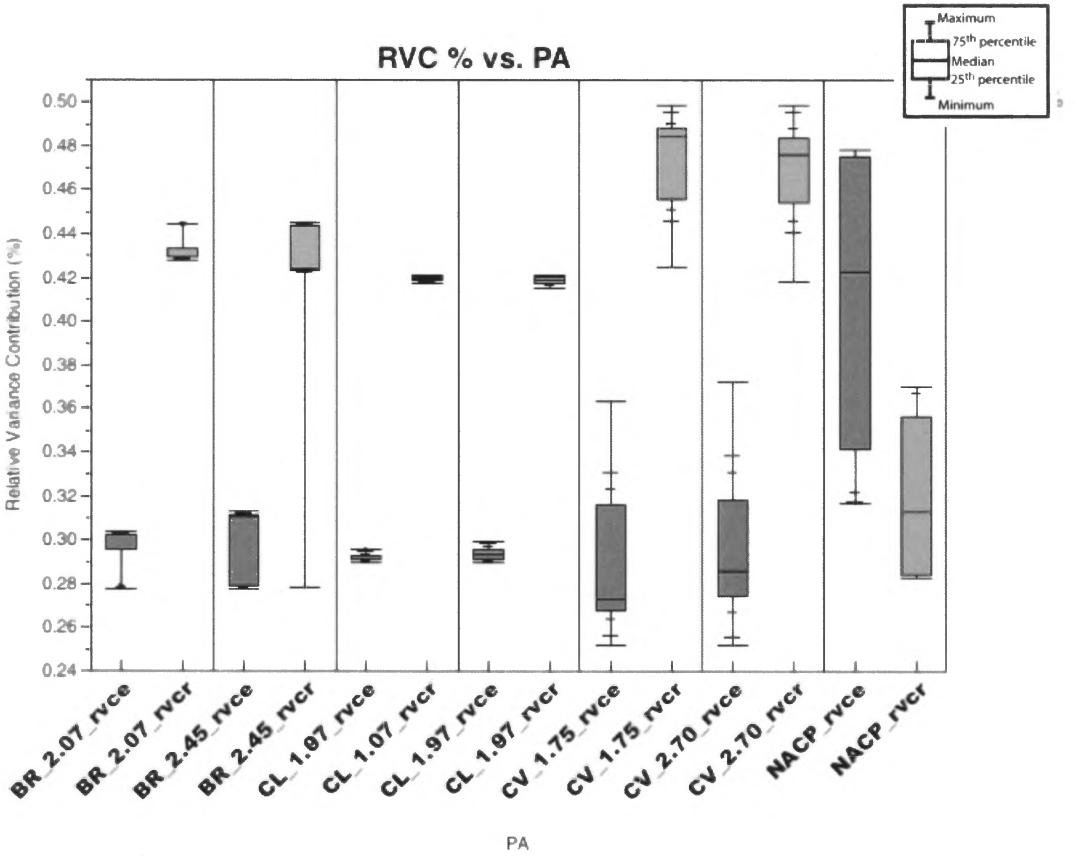


Figure 40: Relative Variance Contribution (RVC) boxplots for each principal aquifer (PA) boxplot showing the value and distribution for the RVC for explanatory variables and model coefficients. The vertical axis is the RVC a decimal scaled so the maximum value is no more than 0.5 or 50%. The horizontal scale represents principal aquifers. The name of each PA is followed by an underscore and dissolved oxygen error value specific to that model. The second underscore is followed by the acronym rvcr and rve. RVCr stands for relative variance contribution for the regression and RVCe stands for relative variance contribution for the explanatory variables. RVCr boxplots are blue and RVCe boxplots are pink.

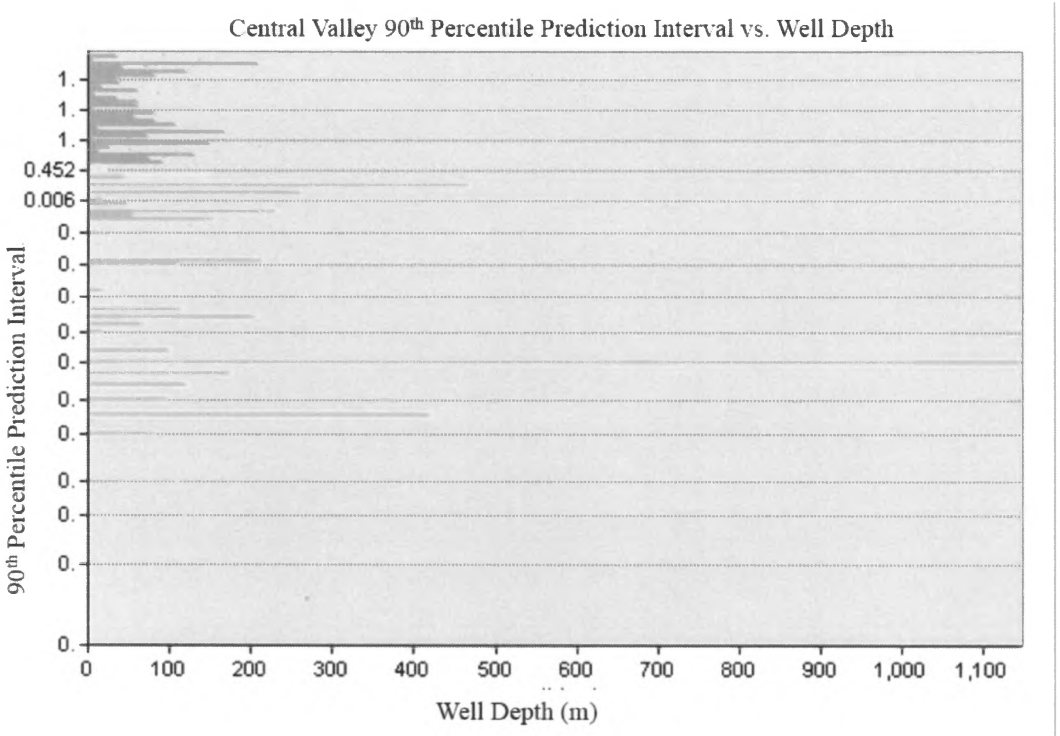


Figure 41: Graph of 90th percentile prediction interval categories (green=0-33, yellow=34-66, orange= 67-100) compared to well depth in the Central Valley aquifer. The vertical axis represents the 90th percentile prediction interval value and the horizontal axis represents well depth in feet. The graph suggests that there is more uncertainty in in shallower wells (see orange bars) compared to deeper wells (see green bars).