PREDICTING PRIVATE WELL WATER QUALITY IMPACTED BY HURRICANE HARVEY USING PRINCIPAL COMPONENT ANALYSIS AND GIS-BASED KRIGING

A Thesis

by

ACHLA OM PRAKASH JHA

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Chair of Committee,	Patricia K. Smith
Co-Chair of Committee,	Diane E. Boellstorff
Committee Members,	Drew M. Gholson
	Raghavan Srinivasan
Head of Department,	Steve Searcy

May 2020

Subject: Biological and Agricultural Engineering

ABSTRACT

When Hurricane Harvey struck Texas in 2017, bringing an unprecedented 102-155 cm of rain to the central and upper Gulf Coast areas and causing widespread flooding, an opportunity developed to conduct a citizen science campaign to investigate water quality in wells impacted during a two-month period following the storm. This study evaluated inorganic contamination in private wells potentially impacted by Hurricane Harvey by 1) characterizing levels of inorganics and fecal indicator bacteria (FIB) in private wells and comparing concentrations to the USEPA drinking water standards for public water systems and 2) assessing the geographic risk variation of contaminants through application of kriging geospatial interpolations. More than 400 well users participated in the seven free well water sampling events offered in 15 hurricane-impacted counties. Arsenic was detected (>1 μg/L) in 80.9% of wells sampled, with 3.4% exceeding 10 μg/L. Lead concentrations exceeded the action level of 15 μg/L in 3.4% of samples and was detected (>1 μg/L) in 25.3% of wells. Iron exceeded 300 μg/L in 22.9% of wells and was detected (>10 μg/L) in 71.0%. Twenty-three percent of the samples exceeded 50 μg/L of manganese and 69.0% of samples had detectable concentrations. The median contaminant concentrations were higher in shallow wells than in deep wells. Also, median contaminant concentrations were higher in wells that had submerged well heads during flooding than wells which did not have submerged well heads. Principal component analysis for elements with secondary drinking water standards with three principal components yielded a cumulative variance of 69.3% in comparison to two principal components with a cumulative variance of 57.5% for elements with primary drinking water standards. Ordinary kriging, universal kriging and empirical Bayesian kriging models were used to interpolate inorganics concentrations and the principal component scores. Empirical Bayesian kriging gave the lowest root mean square error for most variables and was selected as the optimal method. Kriging could have improved if applied to a smaller study area. Study results indicated that about 43.9% of the wells exceeded at least one EPA drinking water standard. Well system characteristics and especially, well head flooding status also likely affected concentrations of inorganic elements. For three kriging methodologies tested, empirical Bayesian kriging was most accurate.

DEDICATION

I would like to dedicate my thesis to my beloved parents, my sister and my niece, Apple.

ACKNOWLEDGEMENTS

I would like to thank my committee chairs, Dr. Patricia Smith and Dr. Diane Boellstorff and my committee members, Dr. Raghavan Srinivasan and Dr. Drew Gholson, for their guidance and support throughout the course of this research.

I would also like to thank Dr. Kelsey Pieper at Northeastern University and Dr. William Rhoads at Virginia Tech for their immense contributions as co-authors for all the conference presentations of my research.

Thanks also go to my friends and colleagues and the department faculty and staff for making my time at Texas A&M University a great experience.

Finally, thanks to my parents and my sister for their encouragement.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a thesis committee consisting of Professor Patricia Smith as the chair of committee and Professor Raghavan Srinivasan of the Department of Biological and Agricultural Engineering at Texas A&M University (TAMU), Associate Professor Diane Boellstorff of the Department of Soil and Crop Sciences at TAMU and Assistant Professor Drew Gholson at Mississippi State University.

The data analyzed for Chapter 3 were provided by Dr. Kelsey Pieper at Northeastern University, Drs. William Rhoads and Marc Edwards at Virginia Tech University and Dr. Diane Boellstorff of Soil and Crop Sciences.

All other work conducted for the thesis was completed by the student independently.

Funding Sources

This work was also made possible in part by National Science Foundation under Grant Number 1760296. The project also was partially supported through Federal Emergency Management Agency funding and through Clean Water Act §319(h) Nonpoint Source funding from Texas State Soil and Water Conservation Board and the United States Environmental Protection Agency under Grant Numbers 13-08 and 17-04. This graduate study was partially funded by the Department of Biological and Agricultural Engineering Graduate Student Scholarship.

NOMENCLATURE

AIC	Akaike Information Criterion
EC	*Escherichia coli*
FIB	Fecal Indicator Bacteria
GIS	Geographic Information Systems
HBSL	Health-Based Screening Level
ICP-MS	Inductively Coupled Plasma-Mass Spectrometry
KMO	Kaiser-Meyer-Olkin
MCL	Maximum Contaminant Level
MDH	Minnesota Department of Health
MPN	Most Probable Number
PC	Principal Component
PCA	Principal Component Analysis
QMRA	Quantitative Microbial Risk Assessment
RMSE	Root Mean Square Error
SMCLs	Secondary Maximum Contaminant Levels
TC	Total Coliform
USEPA	United States Environmental Protection Agency
USGS	United States Geological Survey

TABLE OF CONTENTS

Page

LIST OF FIGURES

LIST OF TABLES

Page

1. INTRODUCTION

Approximately 44 million people in the U.S. (about 13% of U.S. households) rely on domestic wells for drinking water (USGS,2016a; USGS, 2016b). The U.S. Environmental Protection Agency (USEPA) regulates water from public systems, but private water well quality is unregulated. The U.S. Geological Survey's (USGS) National Water Quality Assessment Program evaluated the water quality of 2100 domestic wells in 48 states and reported that one in five of the sampled wells contained one or more contaminants at a concentration greater than either EPA's maximum contaminant level (MCL) or the USGS health-based screening level (HBSL) (DeSimone, 2009). Residents in rural areas with private water systems that are not appropriately installed or adequately maintained could consume drinking water that poses a health risk (Richards et al., 2015).

Texas is 7th in terms of domestic population impacted by arsenic concentrations with roughly 95,000 people affected (Reedy and Scanlon, 2018). Analyses of 10,489 samples collected from 1992 – 2017 in Texas found about 36% of the samples contained arsenic (detection limits varied by years and for different laboratories). About 19% of samples with detectable arsenic and 7% of all samples reported in the study exceeded the current EPA drinking water standard (Reedy and Scanlon, 2018) of 10 μg/L. In addition, Gates et al. (2011) report that nearly 30% of wells in the southern Gulf Coast have arsenic concentrations exceeding the MCL of 10 μg/L.

The World Health Organization notes that testing drinking water for FIB provides a very sensitive method of quality assessment (Edberg et al., 2000). Consumption of well water contaminated with FIB is associated with gastrointestinal illness. *Escherichia coli* (*E. coli*) themselves are usually not pathogenic; however, the presence of *E. coli* in a water sample indicates that it has been contaminated by sewage or animal waste that may contain other disease causing bacteria. Thus, it can be a useful marker for detecting wells that pose a potential health problem in rural areas.

On August 25, 2017 Hurricane Harvey struck Texas, bringing an unprecedented 102-155 cm of rain totaling 76 trillion liters of water to the central and upper Gulf Coast areas and causing widespread flooding, including flooding of private water wells. Well owners became concerned about their well water quality, especially those who knew their well had been flooded. Thus, an opportunity developed to conduct a citizen-science campaign to investigate the water quality of wells impacted by Hurricane Harvey following the storm.

Testing, treatment and maintenance of private water wells is always important as it may prevent contamination, but especially so when a natural disaster such as Hurricane Harvey occurs. Private water wells do not fall under the jurisdiction of the Safe Drinking Water Act, making it the responsibility of the well owner to ensure their water is safe and that they are not contaminating groundwater in the aquifer (Powers et al., 2018). Major barriers to well water testing areas include general awareness, geographic distance to a laboratory and lack of access to water testing (Navas-Acien et al., 2009). To overcome these barriers following the hurricane, collaborators from Virginia Tech, Texas A&M and Louisiana State Universities distributed sampling supplies and survey instruments and conducted laboratory analyses of water samples collected by private well owners from their wells in the counties impacted by Harvey.

In addition, the Texas A&M AgriLife Extension Service utilized its network of county-based personnel and the Texas Well Owner Network, an Extension outreach program, to organize and execute the sampling events. Virginia Tech developed the sampling instructions and results interpretation materials and distributed them to participating citizens. Notification and follow-up with residents was done by both Virginia Tech and Texas A&M AgriLife Extension.

The overall objectives of this study are to a) characterize the groundwater serving private wells by testing drinking water samples and comparing the concentrations to EPA's primary and secondary drinking water standards and b) compare the performance of ordinary kriging, universal kriging and empirical Bayesian kriging as a potential tool to predict contaminant concentrations in the wells of counties affected by hurricane Harvey. In addition, groundwater contaminant concentrations in the flooded wells were compared to groundwater data collected by the Texas Water Development Board (TWDB) for the five years before hurricane Harvey to compare the impacts of pre-Harvey and post-Harvey.

2. MATERIALS AND METHODS

2.1. Study Area

The study area includes counties impacted by Hurricane Harvey in the Gulf Coast aquifer of Texas including, Bexar, Calhoun, Chambers, DeWitt, Hardin, Harris, Goliad, Gonzales, Jefferson, Lavaca, Liberty, Orange, Refugio, Victoria, Waller, and Wharton Counties. This area is composed of discontinuous sand, silt, clay and gravel bed in the Jasper, Evangeline and Chicot Aquifers. The climate of the region is subtropical and influenced primarily by the Gulf of Mexico. Winters are mild and summers are hot, with high humidity in the northeast and semi-arid to arid conditions in the southwest (Larkin and Bomar.,1983). Average high temperatures (°F) range from the 60s in the winter to the 90s in the summer. This region is very active with heavy rainfall, hurricanes and tropical storms due to its location on the Gulf Coast. The amount of rainfall varies throughout the region. Gulf currents bring in over 55 inches of precipitation a year in east Texas. However, the far southwestern portion receives only about 20 inches annually (Mace., 2006). The exact locations of wells surveyed are not shown on the map in Figure 2-1 to ensure participant privacy. The study area showing the participating counties with number of wells tested for inorganics in those counties.

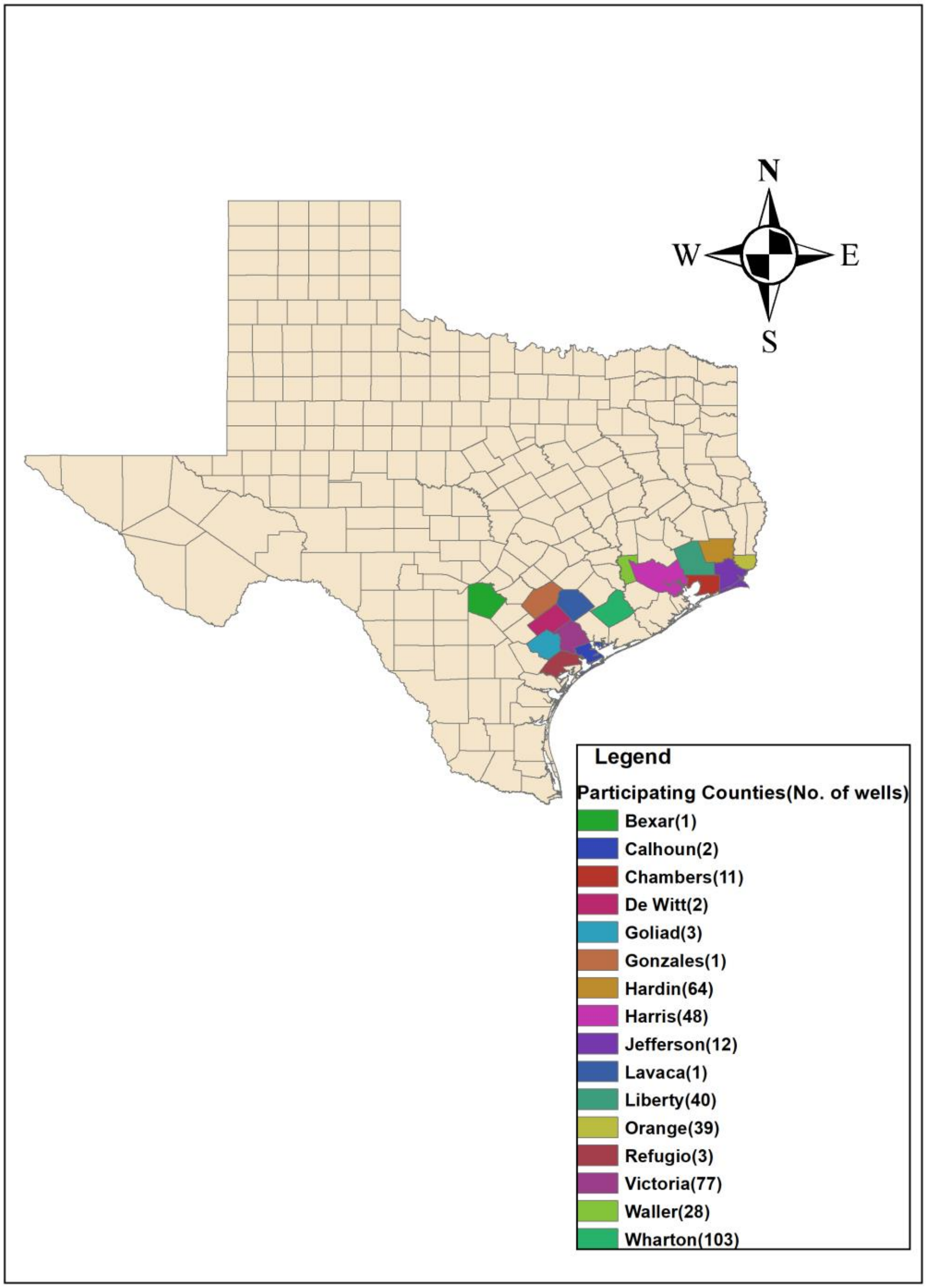


Figure 2-1. Study Area Comprising of 435 Wells Tested in Participating Counties

2.2. Data Collection

Well sampling kits containing a questionnaire to gather information regarding the well, sample collection instructions and collection vessels (IDEXX and large, sterile plastic bottles) were distributed through Texas A&MAgriLife Extension's county offices and other public outlets to private water well owners for this citizen science campaign. As per the sample collection instructions, faucets were rinsed and dried to prevent contamination of the water sample. Participants were instructed flush the system by turning on the cold water tap from the highest flowrate fixture for five minutes. The instructions included the suggestion to use the kitchen tap as the sample collection source and if the kitchen tap was not working, then to use an outdoor tap. After running the cold water tap for one minute, the plastic bottle was filled at a reduced flowrate of the fixture. All samples were collected soon after Hurricane Harvey passed, between 11 September and 16 October 2017.

Around 630 private water wells from 16 hurricane-impacted counties were tested for total coliform bacteria and *E.coli*. Of those, 435 were analyzed with Induced Coupled Plasma-Mass Spectroscopy (ICP-MS), to examine the spatial effects of flooding on well water concentrations in an effort to learn how to better protect private water systems and aquifers. Sample sizes varied for particular analytes. An extension of the sampling effort was conducted to investigate temporal effects, which will be reported elsewhere.

2.2.1. Water Quality Analysis

Well water samples were analyzed at the Water Quality Laboratory in the Department of Civil and Environmental Engineering at Virginia Tech for total coliform and *E. coli* using the IDEXX Colilert test approved by the EPA following the QAQC protocols. Colilert reagent was added to the sample bottles which were then poured into Quanti-Tray/2000 vessels. The Quanti-

Trays were then sealed using the Quanti-Tray Sealer and placed in an incubator at 35°C ± 0.5°C for 24 hours. Yellow wells indicated the presence of total coliforms and yellow/fluorescent wells indicated a positive E. coli result. The most probable number (MPN) per 100 ml of sample for both total coliform bacteria and E. coli was calculated using the MPN table developed by IDEXX. Inorganics were quantified using ICP-MS. Inorganics tested were sodium, magnesium, aluminum, silicon, phosphorus, sulfur, chlorine, potassium, cadmium, titanium, vanadium, chromium, iron, manganese, cobalt, nickel, copper, zinc, arsenic, selenium, boron, strontium, molybdenum, silver, cadmium, tin, barium, lead, uranium and hardness.

2.3. Data Analysis and Comparison to Standards

Concentrations were compared to drinking water primary MCLs and Secondary Maximum Contaminant Levels (SMCLs) for aesthetic standards as per the EPA. Descriptive statistics were calculated only for those constituents for which the contaminant was detected in at least 50% of samples. For samples with concentrations below the detection limit, one-half the detection limit was substituted per EPA protocols (EPA, 2017). Descriptive statistics for samples containing contaminants exceeding EPA standards or objectives, and proportion of samples that were above the detection limit were summarized.

AgriLife Extension distributed a questionnaire in the sampling kits that was developed and used in a prior well recovery project in Louisiana (Dai et al., 2019). Private well owner participants were surveyed to characterize: 1) extent of flooding and well damage; 2) well construction and design parameters; and 3) well water use and consumption patterns including use of water treatment. The survey provides an understanding of resource and recovery needs during natural disasters. Median concentrations for both inorganics with primary drinking water standards and secondary drinking water standards for different survey responses were compared to study the impact of well system characteristics on the contaminants.

2.3.1. Statistical Analysis

All statistical analyses were performed using SAS (SAS Institute Inc., Cary, NC, USA) assuming an $\alpha = 0.05$ as an indication of significance. The data had a non-normal distribution (Shapiro-Wilkes; $p < 0.05$) and so non-parametric statistical tests were performed throughout the study. Potential correlation between different water parameters were evaluated using Spearman's rank coefficient which measures the strength and direction of association between ranked variables. The Kruskal-Wallis test is a rank based non-parametric test that can be used to determine if there are statistically significant differences between two or more groups of independent variables on a continuous or ordinal dependent variable. It is performed to test for differences between three or more conditions. Formally, the null hypothesis is that the population distribution functions are equal for all the conditions which is the responses recorded from private well water owners in the present study. The Kruskal-Wallis test was used to compare the median contaminant concentrations based on the responses obtained from the residents regarding characteristics of the well. This test rejects the null hypothesis that the median rank is equal for all the conditions for $\alpha < 0.05$. However, the Kruskal-Wallis test does not provide indication of which groups are different without also performing post-hoc tests. Hence the Dwass, Steel, Critchlow-Fligner (DSCF) multiple comparisons post hoc procedure was performed to help determine which pairs differ for cases having significant difference in the median ranks of the responses (Higgins., 2004; Conover., 1999). The Mann-Whitney Wilcoxon Test is typically performed to test for differences between two independent conditions. The null hypothesis is the median ranks of both populations are equal while the alternate hypothesis is that the median rank between groups differ (Beall., 1942; Conover., 2004). It was used to compare inorganic concentrations in wells surveyed post- Harvey to concentrations of wells available in Texas Water Development Board (TWDB) database for the five years previous to Hurricane Harvey.

2.3.2. Principal Component Analysis

Principal components analysis (PCA) is a multivariate statistical technique which reduces the dimensionality of a dataset with numerous correlated variables and categorizes those variables into groups based on their covariance (McLeod et al., 2017). It is a variable reduction procedure. PCA is appropriate when data is obtained on several observed variables and there is some correlation in those variables making it possible to reduce the observed variables into a smaller number of PC that will account for most of the variance in the observed variables. PCA has been used in the past to examine and interpret patterns of groundwater quality parameters (Belkhiri et al., 2011; Helena et al., 2000; McLeoad et al., 2017; Sánchez-Martos et al., 2001). It identifies common factor patterns and interprets them with presumed natural and anthropogenic processes that impact groundwater quality. The resulting independent principal components (PC) account for the variance in the observed data, except for those data that are not retained. These independent PC can then be used as predictor variables in subsequent analyses. PCA was performed using SAS (SAS Institute Inc., Cary, NC, USA) and applied to data separated into groups of water parameters identified as United States Environmental Protection Agency (EPA) Primary MCL and EPA Secondary MCL. Descriptive statistics for parameters of interest, along with the proportion of samples exceeding EPA's standards, and the proportion of samples that were above detection limits were summarized.

Values of inorganics concentrations for each individual well were used in the PCA of private well data. Natural Log transformed concentrations (μg/L) were converted to standardized z-scores prior to performing the PCA to mitigate the effect of measurement scales and to ensure that highly variable parameters do not dominate the analysis (Cloutier et al., 2008). Kaiser's measure of sampling adequacy (KMO) (Kaiser et al., 1974) and Bartlett's test of sphericity (Bartlett et al., 1950) were performed to assess the sampling adequacy and correlation of these data and therefore their suitability for PCA. The sampling is adequate or sufficient if the KMO value is larger than 0.5 (Field.,2000). Kaiser (1974) recommends a bare minimum of 0.5 and the value between 0.5 and 0.7 are mediocre, value between 0.7 and 0.8 are good, value between 0.8 and 0.9 are great and value between 0.9 and above are superb (Hutcheson & Sofroniou.,1999). The equation for the KMO test is represented by Equation 1 :

$$MO_j = \frac{\sum_{i\neq j} r_{ij}^2}{\sum_{i\neq j} r_{ij}^2 + \sum_{i\neq j} u} \qquad \text{Equation 1}$$

Where:

$R = [r_{ij}]$ is the correlation matrix

$U = [u_{ij}]$ is the partial covariance matrix

In each analysis, PC with an eigenvalue of greater than one were retained (Kaiser, 1960), and subject to varimax rotation to maximize the variation explained by each of the retained principal components and to obtain the final principal component loadings and coefficients. The scree test (Cattell., 1964) is also used as a criterion to decide the number of principal components to be retained. It is basically a plot of eigenvalues and principal components which indicates a PC location having sharp change in the slopes of adjacent line segments to be suitable to be retained for the PCA analysis. The components before the sharp change are assumed to be meaningful and retained; those appearing after the break are assumed to be unimportant and are not retained.

PC scores for each of the retained PC were calculated for the private wells. PCs were obtained for both Primary MCL and Secondary MCL data groups. PC scores for each of the retained PC were calculated for the private wells for use in the geostatistical analysis. The general form for the formula to compute scores on the first component extracted in a principal component analysis is represented by Equation 2:

$$C_1 = b_{11}(X_1) + b_{12}(X_2) + \cdots b_{1p}(X_p) \qquad \text{Equation 2}$$

Where

C_1 = the subject's score on the first principal component extracted

b_{1p} = the regression coefficient (or weight) for observed variable p, as used in creating principal component 1.

X_p = the subject's score on observed variable p.

2.4. Geostatistical Analysis

Kriging is an advanced geostatistical procedure that generates an estimated surface from a scattered set of points. Kriging weights the surrounding measured values to derive a prediction for an unmeasured location. The general formula is formed by the weighted sum of data as shown below in Equation 3:

$$\hat{Z}(S_0) = \sum_{i=1}^{N} \lambda_i Z(S_i) \qquad \text{Equation 3}$$

Where

$Z(S_i)$ = the measured value at the i^{th} location

λ_i = an unknown weight for the measured value at the i^{th} location

S_0 = the prediction location

N = the number of measured values

The basis of kriging is the semivariogram model, which uses the semi-variance between point measures to summarize the spatial relationships in variables. In spatial modeling of the structure of the measured points, an empirical semivariogram is computed with Equation 4 for all pairs of locations separated by distance h:

$$\text{Semivariogram}\ (\text{distance}_h) = 0.5 * \text{average}\ ((\text{Value}_i - \text{Value}_j)^2) \qquad \text{Equation 4}$$

Ordinary kriging and universal kriging are methods for which estimation is based on weighted least squares and the assumption that the calculated semivariogram is the true model for the data. Ordinary kriging is different from universal kriging in that ordinary kriging assumes a constant unknown mean across a given area, while universal kriging assumes a constant trend in the data. The empirical Bayesian kriging method is based on restricted maximum likelihood estimation. Additionally, it allows for uncertainty in the semivariogram model by a process of data subsetting and simulation to estimate a range of semivariogram models (Oliver., 1990; Royle et al., 1981). PC scores and log-transformed concentrations were mapped in ArcMap 10.6 (ArcGIS, ESRI, Redlands, CA, USA) to compare the performance of ordinary, universal and empirical Bayesian kriging for the prediction of concentrations of arsenic along with principal component scores of elements with primary drinking water standards and elements with secondary drinking water standards.

Semivariogram models for each variable were developed using SAS 9.4 software (SAS Institute Inc., Cary, NC, USA). First or second order large scale trends were identified in the data for each variable using regression analysis. Wherever trends were detected, residuals from the regression models were used to develop semivariogram models. The average nearest neighbor distance for the sampling locations for each variable were calculated in ArcGIS 10.6 and set as the lag distance for the semi variograms, and half the study area extent divided by the lag distance was used to calculate the maximum number of lags, which limits the semivariogram lag to half the extent of the sampling area (Olea., 2006). The fits of spherical, exponential, and Gaussian models were compared using the Akaike Information Criterion (AIC) for each variable. The model having the lowest AIC value was selected as the semivariogram model. The presence of anisotropy was evaluated visually by divinding the semivariogram into eight directions, however, for the final analysis, an omnidirectional semivariogram was modeled (McLeod et al., 2017). The order of trend removal and lag distance were set based on the semivariogram models developed in SAS. The calculated maximum number of lags for each parameter exceeded the maximum of 100 which is allowed in ArcMAP, thus restricting the number of lags used for kriging to 100 for each dataset.The root mean square prediction error (RMSE) in the kriging cross-validation analysis performed in ArcMap was used to compare the ability of ordinary, universal, and empirical Bayesian kriging to predict values for each parameter (Robinson and Metternicht., 2006). It gives an estimate of the standard deviation of the residuals (prediction errors). This error indicates how closely the method predicts the measured values. The smaller the RMSE the better (Tziachris et al., 2017). In cross-validation, each point is sequentially omitted from the dataset and the remaining points used to predict a value for that point; the difference between the predicted and the measured value is the prediction error. The method that results in the greatest number of parameters with the lowest RMSE was identified as the optimal kriging method for the water data. The equation for calculating RMSE is shown in Equation 5 :

$$\mathrm{RMSE} = \sqrt{\frac{\sum \hat{Z}(s_i) - Z(s_i)^2}{n}} \qquad \text{Equation 5}$$

Where

$\hat{Z}(S_i)$= predicted value at the i^{th} location

$Z(S_i)$ = the measured value at the i^{th} location

S_0 = the prediction location

n = the number of measured values

3. RESULTS

3.1. Well Constituents Studied

Overall, 43.9% of the well samples contained contaminant concentrations that exceeded at least one primary or secondary drinking water standard. Samples exceeded the standard for total coliform most often as a water quality parameter in 35.3% of the wells, while cadmium and copper were exceeded least often in 0.2% of the wells as shown in Figure 3-1.

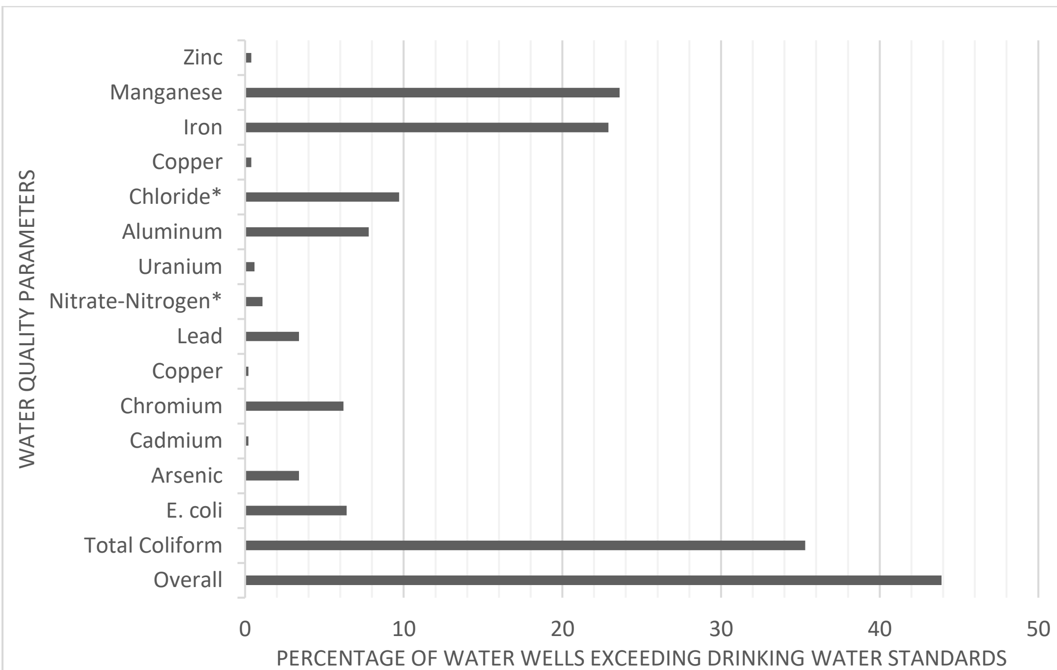


Figure 3-1. Overall Percentage of Private Water Wells That Exceeded at Least One Primary or Secondary Drinking Water Standards and Percentage That Exceeded for Individual Drinking Water Standards

***n=338 for Chloride and Nitrate-Nitrogen**

Arsenic (As) was detected in about 81% of the samples and exceeded the EPA drinking standard for 3.4% of the samples. Copper (Cu) was detected in 80% of the samples but exceeded the action level in only 0.2% of the samples. Similarly, Chromium (Cr) was detected in 61% of the samples while Nitrate (NO_3) was found in 70% of the samples. Lead (Pb) was detected in about 25% of the samples, which is relatively high as samples were collected after five minutes of flushing. Similarly, other elements such as Cadmium (Cd) and Uranium (U) were detected in less than 50% of samples as shown in Table 3-1.

Table 3-1. Descriptive Statistics of Well Water Constituents (in μg/L) with Reference to the U.S. Environmental Protection Agency Primary Drinking Water Standards

n =435	**Mean (SD) Median**	**MCL**	**Above MCL (%)**	**DL**	**Above DL (%)**	**Max**
Arsenic	2.7(4.2) 1.6	10	3.4	0.5	80.9	56.8
Barium	182.6(174.4) 163.7	2000	0.0	1	94	1575
Cadmium	NC	5	0.2	1	0.7	5.7
Chromium	8.3(75.7) 0.5	10	6.2	1	60.6	1233.0
Copper	32.6(123.3) 5.7	1300	0.2	1	80.7	1710.0

Table 3-1. Continued

n =435	Mean (SD) Median	MCL	Above MCL (%)	DL	Above DL (%)	Max
Lead[a]	NC	15[b]	3.4	1	25.3	76.4
Nitrate[c]	2147.7(3047.2) 569.1	10000	1.1	10	70.7	23983.9
Selenium	NC	50	0	2.5	9.7	10.2
Uranium	NC	30	0.6	1	37.2	36.8

SD = standard deviation, max = maximum, DL = detection limit, NC = not calculated, MCL= Maximum Contaminant Level

[a] Water sample collected after flushing the cold water tap for five minutes.

[b] Action level

[c] n = 338 for Nitrate-Nitrogen

Descriptive analysis of elements with secondary drinking water standards shows exceedance of secondary drinking water standards for Chloride (Cl), Iron (Fe), Manganese (Mn) and Zinc. (Zn). Cu exceeded the action level for both primary and secondary drinking water standards. Fe was detected in 71% of samples while Mn was detected in 69% of samples as shown in Table 3-2. Sulfate (SO_4) was detected in around 93% of samples. Aluminum (Al) and Silver (Ag) were detected in less than 50% of the samples.

Table 3-2. Descriptive Statistics of Well Water Constituents (in μg/L) with Reference to the U.S. Environmental Protection Agency Secondary Drinking Water Standards

n =435	Mean (SD) Median	MCL	Above MCL (%)	DL	Above DL (%)	Max
Aluminum	NC	50-200	7.8	5.0	26.2	5945
Chloride[a]	1.03×10^5 (12.30×10^5) 0.58×10^5	2.5×10^5	9.7	10	99.4	10.3×10^5
Copper	32.6(123.3) 5.7	1000	0.4	1.0	80.7	1710
Iron	357.8(986.5) 49.0	300	22.9	10	71.0	8826
Manganese	54.5(154.6) 8.4	50.0	23.6	1.0	69.0	1509
Zinc	166.3(656.5)	2000	0.4	5.0	79.1	10470
Silver	NC	100	0.0	1.0	0.5	7
Sulfate[a]	0.1×10^5 (0.2×10^5) 0.05×10^5	2.5×10^5	0.0	10.0	93.4	2.1×10^5

SD = standard deviation, max = maximum, DL = detection limit, NC = not calculated, MCL = Maximum Contaminant Level

[a] n = 338 for Chloride and Sulfate

Total coliform was detected in around 46% of the samples while *E. coli* was detected in around 12% of the samples as shown in Table 3-3.

Table 3-3. Descriptive Statistics of Total Coliform and *E. Coli* in the Wells (MPN/100 mL)

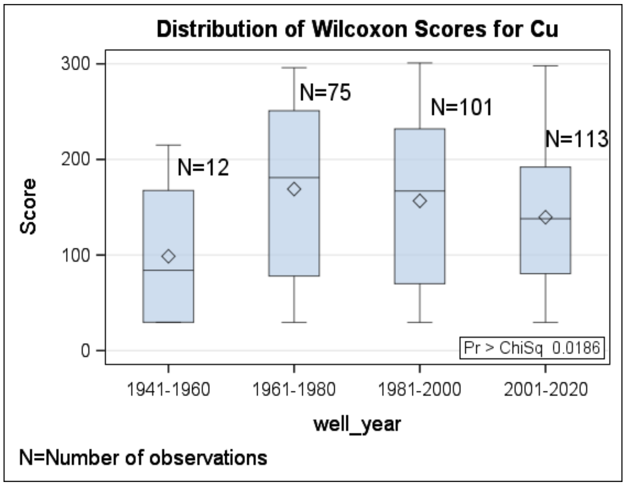
n =630	Total Positive (%)	Mean (SD) Median	Geometric Mean
Total Coliform	46.5	80.6(272.9) 0	0.8
E. coli	11.9	7.8 (103.1) 0	0.1

3.2. System Characteristics

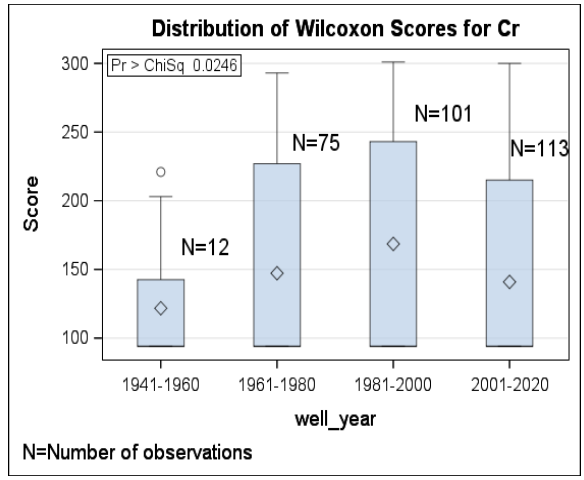
Water quality parameters were analyzed with regard to survey data on well system characteristics. The system characteristics which were taken into consideration were type of well, age of the well, depth of the well, source of water sample collection, submerged well head, damaged well, damaged septic, whether the well system was used during flooding and well water used for drinking. Concentrations of contaminants that were significant based on results of Kruskal-Wallis or Wilcoxon Mann Whitney tests are described below.

3.2.1. Water Well Age

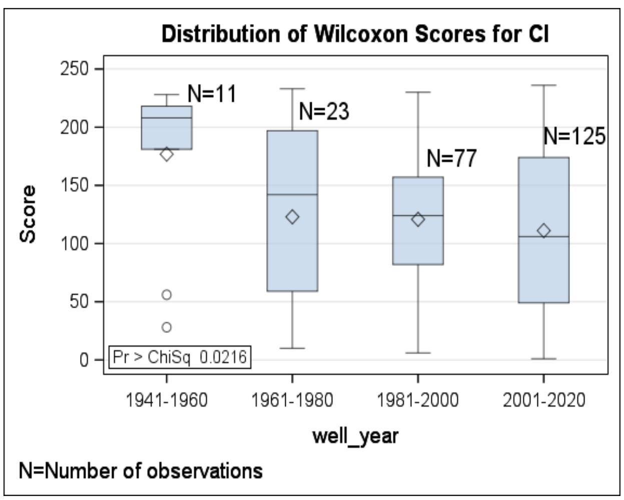
Although Kruskal-Wallis tests indicated that there was a significant difference ($p<0.05$) in element concentrations between various well ages for Cu, Cr, NO_3, and SO_4, with the exception of Cl, no data trends were apparent (Figure 3-2). The median Cl concentration for wells constructed in 1941-1960 (n=11; 176.8 μg/L) was significantly higher (Kruskal-Wallis test, $p<0.05$) than for wells constructed in following years through 2001-2020.



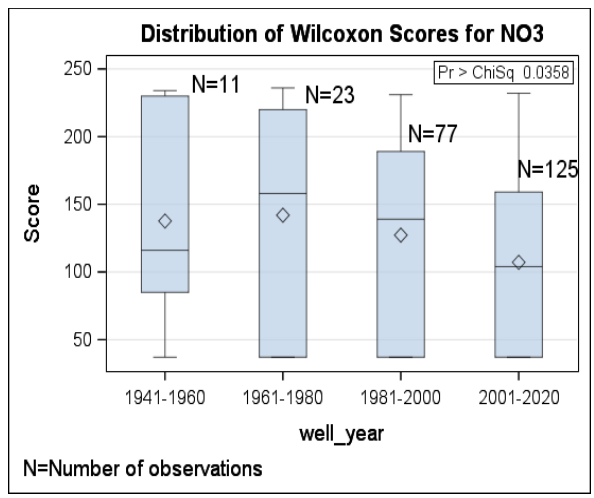
(a)



(b)

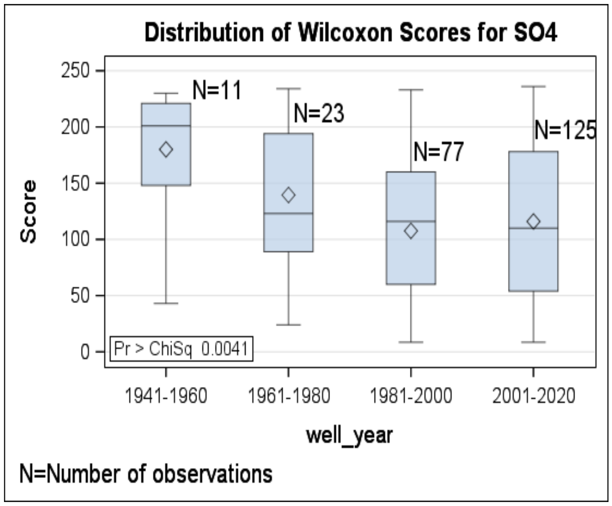


(c)



(d)

Figure 3-2. Effect of Water Well Age on (a) Cu (b) Cr (c) Cl (d) NO_3 (e) SO_4



(e)

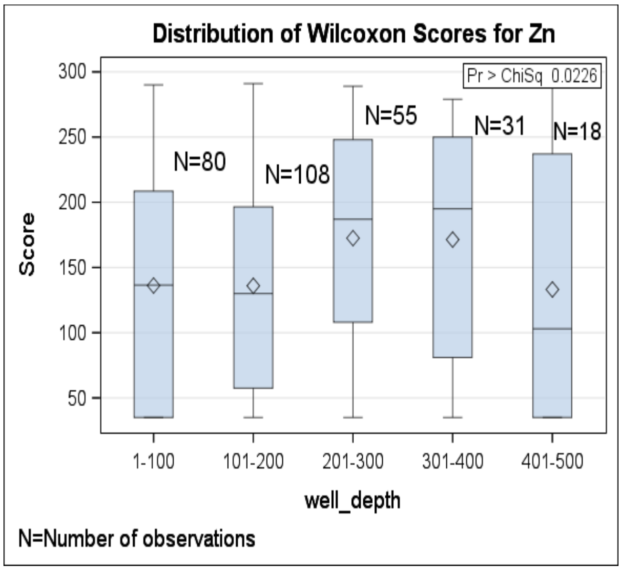
Figure 3-2. Continued

Furthermore, post hoc analysis using the DSCF method of multiple pairwise comparison indicated significant difference occurred in Cl concentrations for wells constructed in 1941-1960 and wells constructed in 1981-2000 ($p=0.03$). Significant difference was also observed when comparing wells constructed in 1941-1960 and wells constructed in 2001-2020 ($p=0.01$). Discarding outliers shown in Figure 3-2 (c) may account for the trend observed. No other elements reported a significant concentration difference for statistical tests based on water well age. Higher median Cl concentration in older wells could be due to the use of thinner casing which may have corroded and older wells more likely needing maintenance attention.

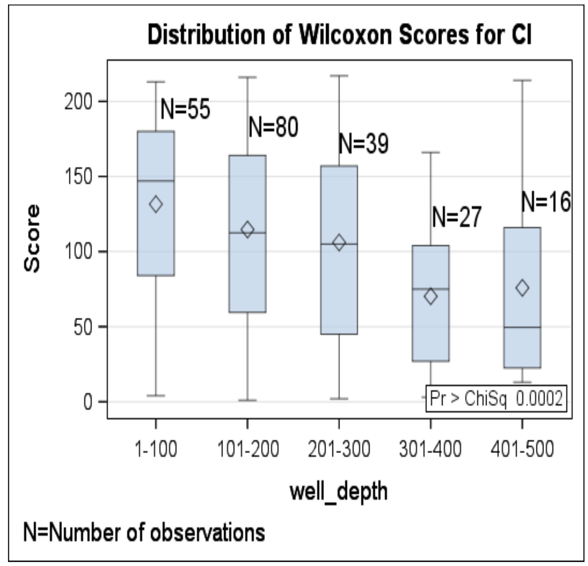
3.2.2. Well Depth

Although Kruskal-Wallis tests indicated that there was a significant difference ($p<0.05$) in element concentrations between various well depths for Zn, Cl, NO_3, and SO_4, no data trends were apparent for Zn (Figure 3-3). Post hoc analysis using DSCF comparisons indicated significant difference in median Cl concentration for wells constructed at 1-100 feet v/s wells constructed at 301-400 feet ($p=0.0005$) and wells constructed at 101-200 feet v/s wells constructed at 301-400 feet ($p=0.01$). Significant difference occurred in median NO_3 concentration for wells constructed at 1-100 feet v/s wells constructed at 300-400 feet ($p=0.004$).A data trend was observed for SO_4 where significant differences were observed in wells constructed at 1-100 feet when compared to wells constructed at 201-300 ($p=0.0003$), 301-400 feet ($p=0.0006$) and 401-500 feet ($p=0.01$). Similarly significant difference in SO_4 concentration was observed in wells constructed at 100-200 feet when compared to wells constructed at 201-300 feet ($p=0.01$) and wells constructed at 301-400 feet ($p=0.02$). Higher SO_4 concentration in shallow wells could be due to application of fertilizers since sulfate is mobile in soil and input to soil could affect shallow groundwater pumped from poorly maintained wells. Significant difference in SO_4 concentration is likely not attributable to flooding as it was found to be higher in the wells not damaged by flooding in comparison to wells damaged by flooding as discussed in section 3.3.5.

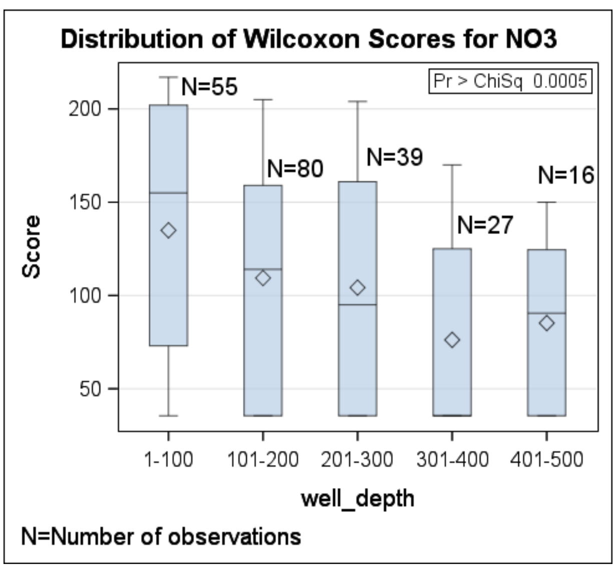
Various other factors that could have contributed to naturally high SO_4 and Cl concentrations in the study area could be seepage of saline water from nearby formations, coastal saltwater intrusion, irrigation return flow, and oil/gas production. Wells deeper than 100 feet are generally safer than shallow wells because the groundwater supplying deep wells has traveled a considerable distance underground, usually over a long time, allowing the soil to filter contaminants. No other elements showed a significant concentration difference for statistical tests based on well depth.



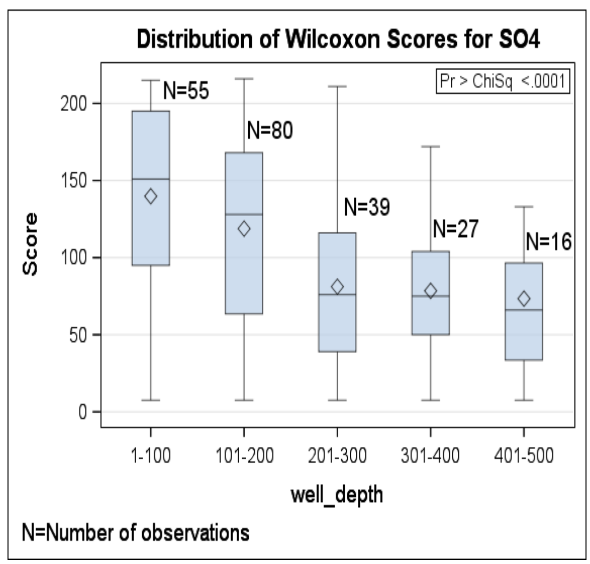
(a)



(b)



(c)

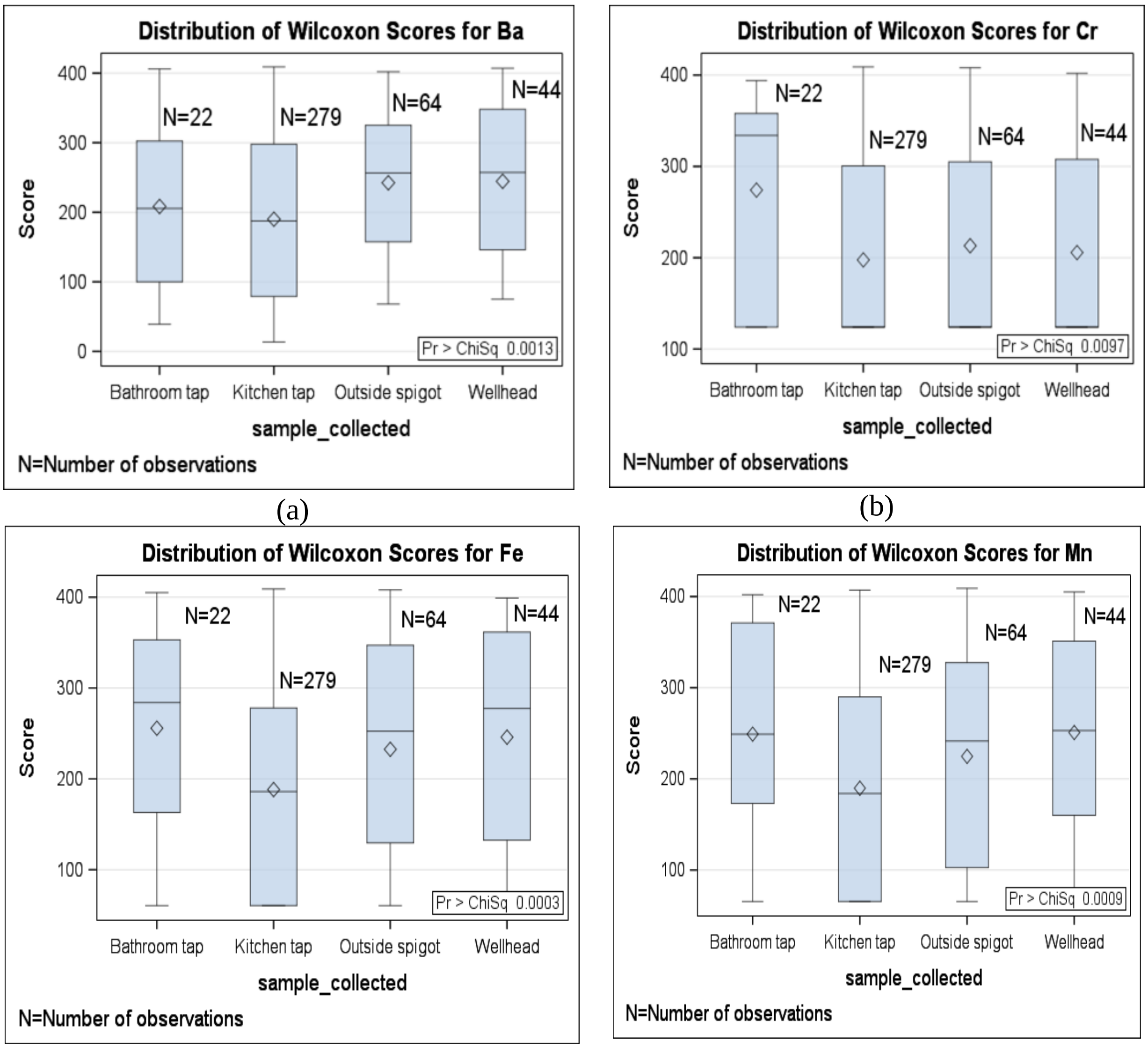


(d)

Figure 3-3. Effect of Well Depth on (a) Zn (b) Cl (c) NO_3 (d) SO_4

3.2.3. Source of Sample Collection

Kruskal-Wallis tests indicated that there was a significant difference ($p<0.05$) in element concentrations due to various sources of sample collections for Ba,Cr, Fe and Mn (Figure 3-4).



(a) (b) (c) (d)

Figure 3-4. Effect of Sources of Sample Collection on (a) Ba (b) Cr (c) Fe (d) Mn

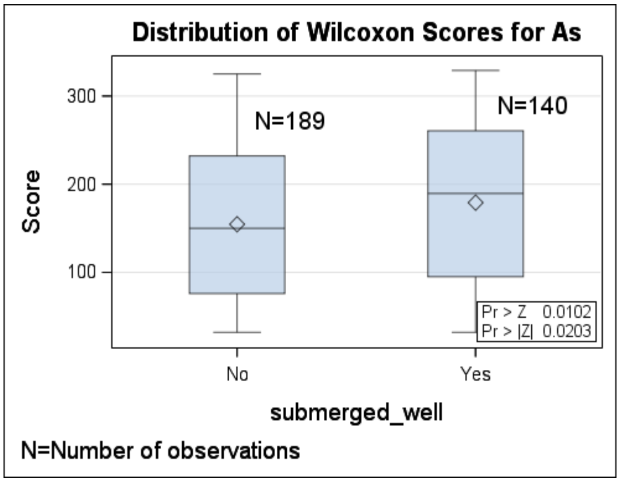
Moreover, post hoc analysis using DSCF comparisons indicated significant difference in median Cr concentration for samples collected from bathroom taps to samples collected from kitchen taps ($p = 0.005$). The median Ba concentration($p=0.02$), Fe concentration ($p=0.03$) and median Mn concentration ($p=0.01$) showed significant differences for paired comparison for samples collected from kitchen taps and well heads. One of the reasons for significant difference in Ba, Fe and Mn concentration for samples collected from well heads v/s kitchen tap could be due to the submergence of the well head as a result of well flooding as discussed in section 3.2.4. No other elements reported a significant concentration difference for statistical tests based on sources of sample collection.

3.2.4. Submerged Well Head

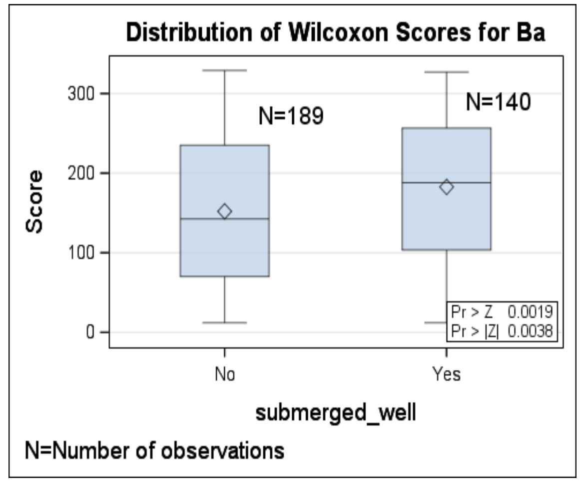
Wilcoxon Mann Whitney tests indicated that concentrations for water samples from wells with submerged wellheads were higher for As ($n= 140$; 179.1 μg/L, $p<0.05$), Ba ($n= 140$; 182.7μg/L, $p< 0.05$), Cr ($n=140$; 182.5 μg/L, $p< 0.05$), Fe ($n=140$; 203.1 μg/L, $p< 0.05$), Mn ($n=140$; 207.3 μg/L, $p< 0.05$), Zn ($n=140$; 187.5 μg/L, $p< 0.05$) and TC ($n=140$; 187.5 μg/L, $p< 0.05$) (Figure 3-5). Thirty-two percent of the wells tested for inorganics ($n=435$) also had a submerged well head. Eleven percent of the wells tested for inorganics ($n=435$) were damaged and 51% of those wells had a submerged well head. Additionally, 18% of the wells analyzed for bacteria ($n=630$) were damaged and 80% of the damaged wells had a submerged well head.

Thirty-four percent of the wells tested for bacteria ($n=630$) had submerged well heads. Higher As concentration in the submerged well could possibly result from leaching of As from some two dozen superfund sites located in the counties affected by Hurricane Harvey (Tabuchi and Kaplan., 2017). Thirteen of 41 superfund sites were flooded and damaged by Harvey according to the EPA. High Cr concentration could be from industrial leaks as Cr is very soluble and moves easily within the environment, and many of the wells surveyed were also near urban areas. High Zn concentration could have resulted from corrosion by flood water (National Research Council (US) Safe Drinking Water Committee, 1982).

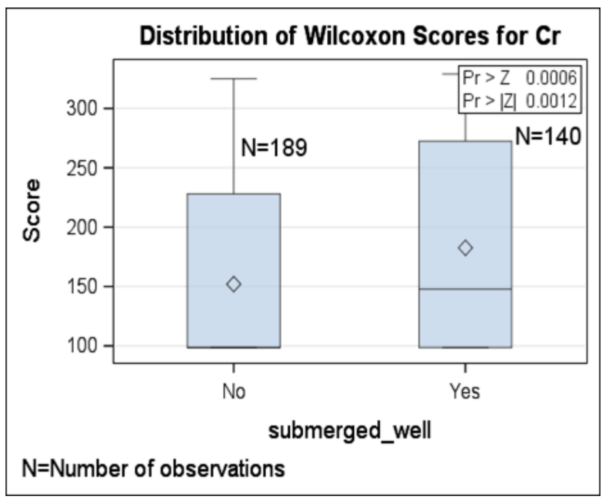
The concentration of Zn in drinking water can often be higher than the concentration in raw water from which drinking water was obtained because Zn may leach from transmission and distribution pipes (Agency for Toxic Substances and Disease Registry, 2005). High fecal coliform bacteria could have resulted from possible causes of contamination such as a faulty well head or improper well construction. Presence of Ba could be due to refineries in the affected areas. Excessive precipitation could have caused the water to percolate through soils and rock and contaminate the wells with the dissolved minerals such as Fe and Mn. Fe could also be present from iron pipe corrosion. No other elements reported a significant concentration difference for statistical tests based on submerged well head.



(a)

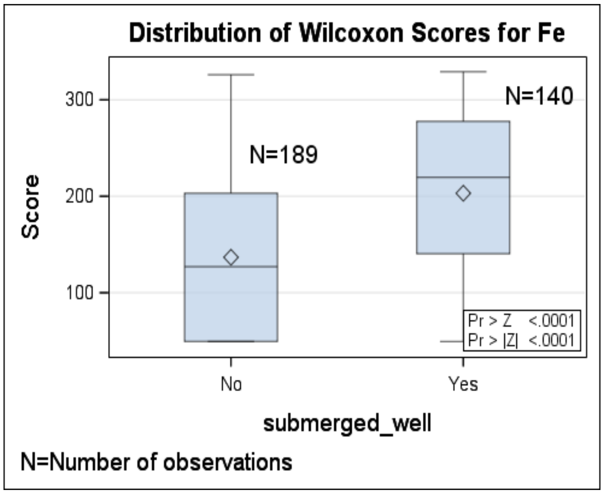


(b)

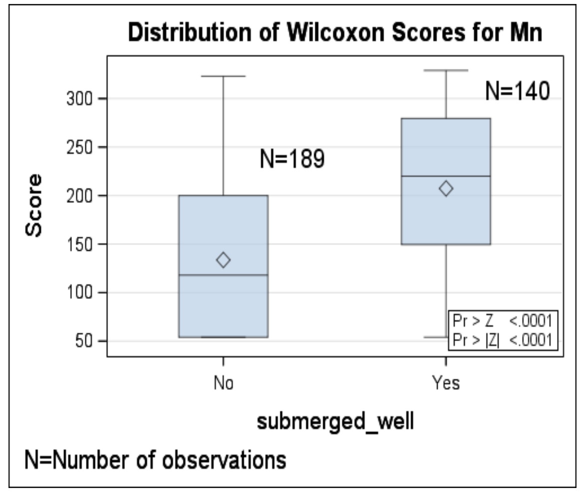


(c)

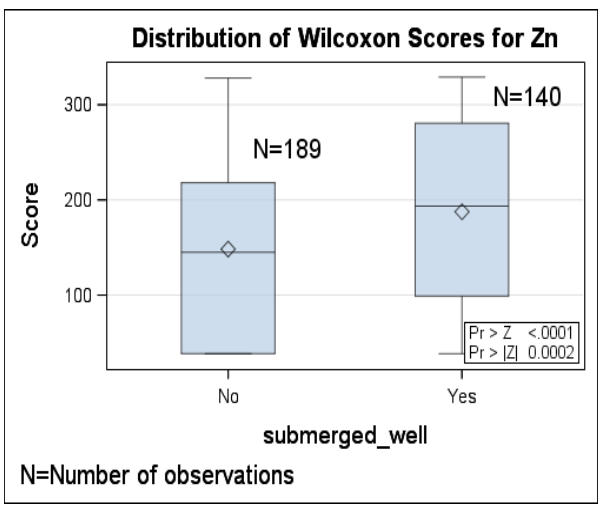
Figure 3-5. Effect of Submerged Well Head due to Flooding on (a) As (b) Ba (c) Cr (d) Fe (e) Mn (f) Zn (g) TC



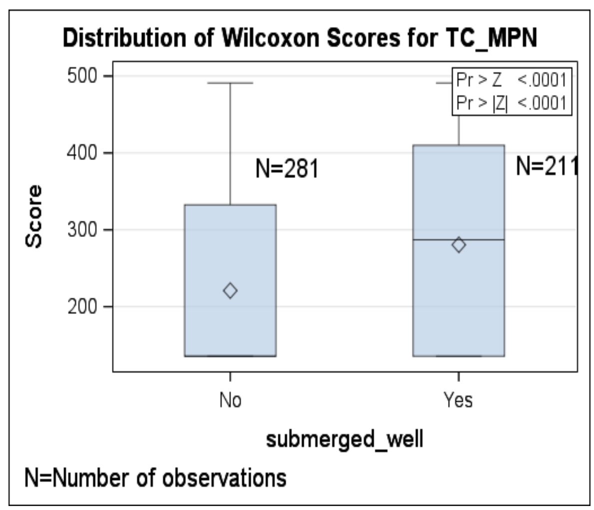
(d)



(e)



(f)



(g)

Figure 3-5. Continued

3.2.5. Damaged Well System

Damaged well system refers to wells having damaged well casing, wiring, and/or piping system resulting from well areas being impacted by the hurricane. The median Mn concentration for wells that were damaged (n= 49; 214.7 μg/L) was significantly higher (Wilcoxon Mann Whitney test, $p<0.05$) than for wells that were not damaged (n=295; 164.6 μg/L) (Figure 3-6). No other elements reported a significant concentration difference for statistical tests based on damaged well system.

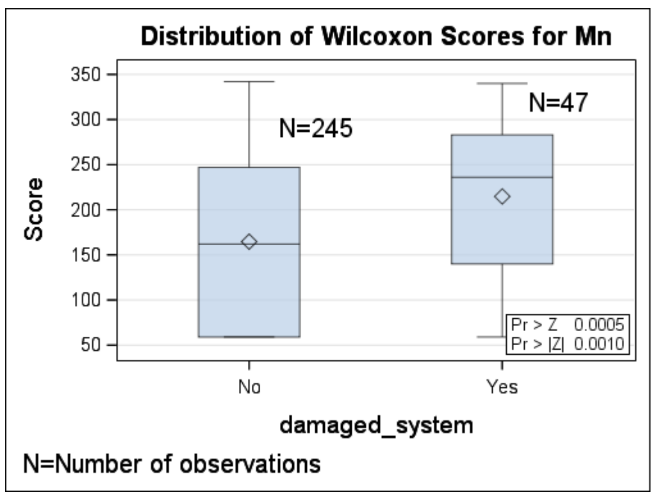


Figure 3-6. Effect of Damaged Well System on Mn

3.2.6. Well Water Use for Drinking

The median SO_4 concentration was significantly higher (Kruskal-Wallis test, $p<0.05$) for wells used for drinking water with filter/treatment (n=89; 186.2 μg/L) than for wells used for drinking water without filter/treatment (n=156; 157.3 μg/L) and the wells which were not used for drinking water (n=81; 150.3 μg/L) (Figure 3-7). The DSCF post hoc analysis found a significant difference in the median SO_4 concentration for wells not used for drinking water to the wells used for drinking water with a filter/treatment ($p=0.03$). No other elements reported a significant concentration difference for statistical tests based on well water usage for drinking. Seventy-Nine percent of the residents that used well water for drinking with treatment/filter had water softners, sediment filter or activated carbon filters as the water treatment device installed in their houses which do not remove sulfate from water according to a report by the Minnesota Department of Health (MDH, 2019). Around 13% of the residents had Reverse Osmosis unit installed in their homes which is considered to be an effective treatment method for removing sulfate from water.

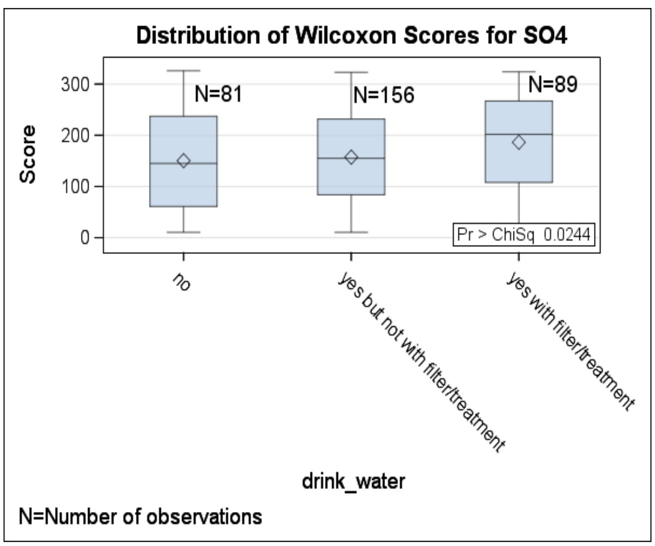
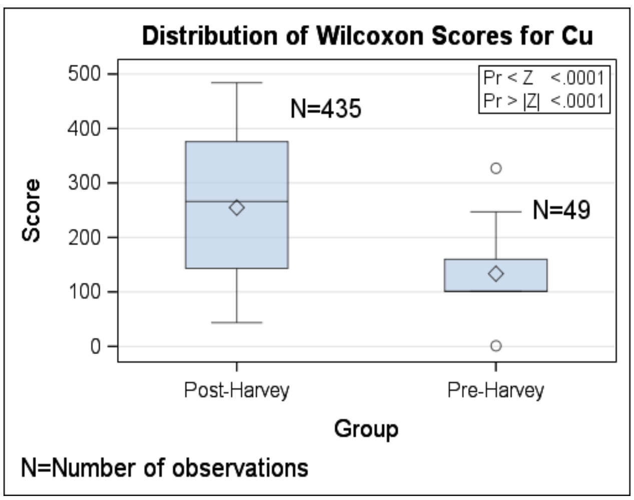


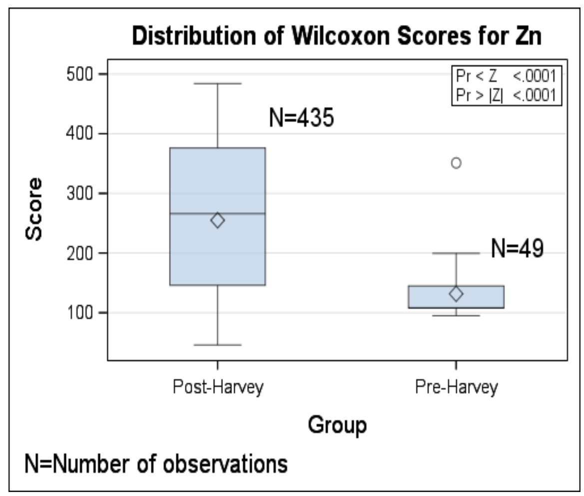
Figure 3-7. Effect of Water Usage for Drinking on SO_4

3.3. Comparative Study of Flooded Wells with Pre-Harvey Conditions

Contaminant concentrations post-Harvey varied significantly compared to the previous five years of water samples recorded in the the study area counties obtained from TWDB for Cu and Zn. No significant difference was found for the other elements reported in this study. The median Cu concentration post-Harvey (n=435; 254.7 µg/L) was significantly higher ($p<0.05$) than during no flood conditions (n = 49; 133.37 µg/L). One of the reasons for higher Cu concentration post-Harvey could be the corrosion of Cu plumbing due to flooding. The median Zn concentration post-Harvey (n=435; 254.9µg/L) was also significantly higher ($p<0.05$) than during no flood conditions (n = 49; 131.7 µg/L). Similarly, higher Zn concentration could be associated with older, galvanized pipes corroding with flood waters (Figure 3-8).



(a)



(b)

Figure 3-8. Effect of Post-Harvey vs Pre-Harvey Conditions on (a) Cu (b) Zn

3.4. Principal Component Analysis

The Kaiser-Meyer-Olkin (KMO) statistical measure of sampling adequacy was above the acceptable value of 0.5 for elements with primary drinking water standards (KMO = 0.52) and for elements with secondary drinking water standards (KMO = 0.54) indicating the data are appropriate for principal component analysis.

The eigenvalue-one criterion was used for solving the number of components retained problem. The rationale for this criteria is that each observed variable contributes to one unit of variance to the total variance in the dataset. Any component that displays an eigenvalue greater than 1 is accounting for a greater amount of variance then had been contributed by one variable and thus retained. The scree test (Cattell, 1966) also helps in deciding the number of principal components to be retained. If the scree plot contains an “elbow” (a sharp change in the slopes of adjacent line segments), that location indicates a good number of PCs to retain. In this study, it is observed at the location of component 2 (Figure 3-9). Hence, two principal components were retained.

Table 3-4. Principal Component Loading, Eigenvalues and Percent of Variance Explained for Each Retained Component of Elements with Primary Drinking Water Standards

	Principal Component 1	**Principal Component 2**
Arsenic	-0.198	**0.748**
Copper	**0.663**	-0.177
Chromium	**0.565**	0.155
Nitrate	0.205	**0.539**
Eigen Value	1.249	1.053
Cumulative Variance	31.2	57.5

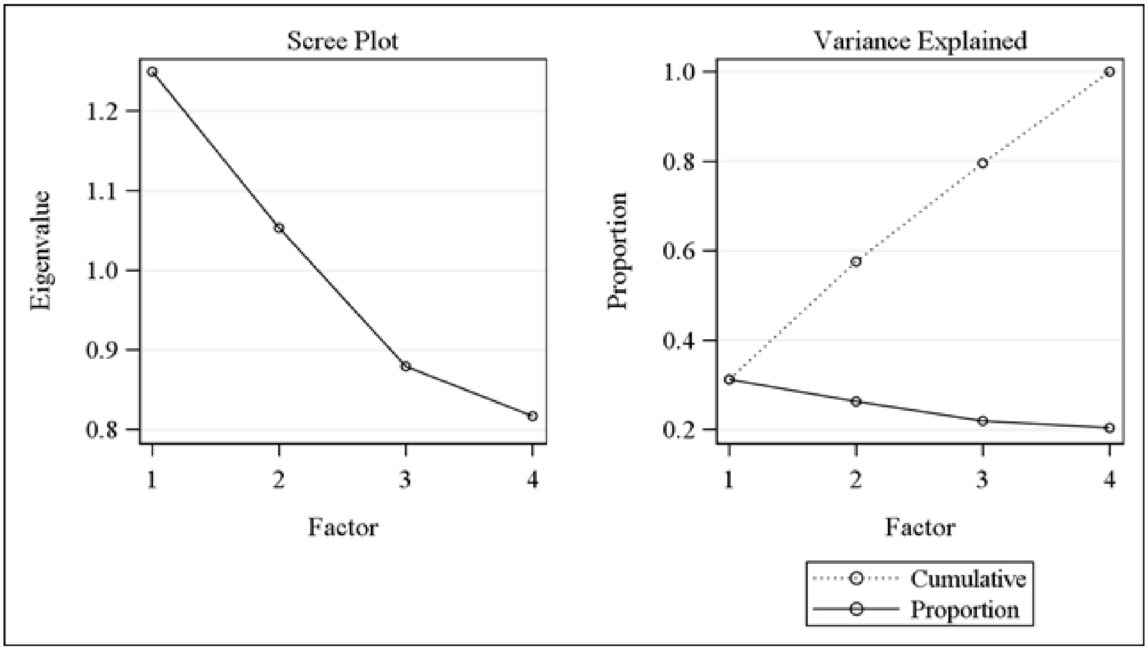


Figure 3-9. Scree Plot and Percentage of Variation Explained for Each Principal Component for the Elements with Primary Drinking Water Standards

The first primary standards principal component (PC1) was strongly associated with copper and chromium because of higher factor loadings while the second primary standards principal component (PC2) was strongly associated with arsenic and nitrate as it had higher factor loadings for these elements than PC1. PCA for elements with primary drinking water standards yielded two PCs accounting for 57.5% of the variance as shown in Table 3-4.

The eigenvalue one criterion or Kaiser's criterion for retaining principal components greater than 1 yielded three PCs accounting for 69.3% of the variance for elements with secondary drinking water standard as shown in Table 3-5.The graphical scree plot for the elements with secondary drinking water standards also indicated the location of component 3 having a sharp change in the slopes of adjacent line segments as a good number of PCs to retain.

Hence using both the eigenvalue-one criterion and the scree test, three PCs were retained for elements with secondary drinking water standards (Figure 3-10).

The first PC of elements with secondary drinking water standards was strongly associated with iron and manganese due to higher loading associated in comparison to PC2 and PC3. Similarly, the second secondary standards principal component was strongly associated with copper and zinc, while the third principal component of elements with secondary drinking water standard was strongly associated with chloride and sulfate because of higher component loading.

Table 3-5. Principal Component Loading, Eigenvalues and Percent of Variance Explained for Each Retained Component of Elements with Secondary Drinking Water Standards

	Principal Component 1	Principal Component 2	Principal Component 3
Chloride	0.144	-0.190	**0.806**
Copper	-0.077	**0.871**	0.034
Iron	**0.874**	0.125	0.074
Manganese	**0.864**	-0.079	-0.135
Sulfate	-0.224	0.189	**0.694**
Zinc	0.472	**0.572**	-0.077
Eigen Value	1.881	1.156	1.121
Cumulative Variance	31.3	50.6	69.3

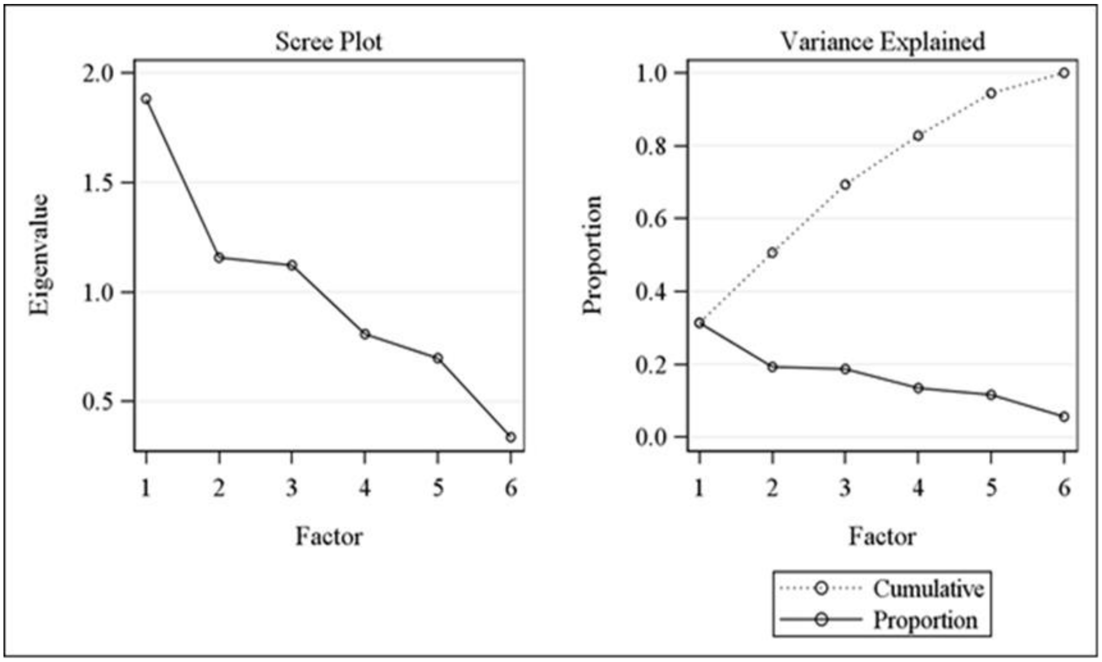


Figure 3-10. Scree Plot and Percentage of Variation Explained for Each Principal Component for the Elements with Secondary Drinking Water Standards

3.5. Kriging

The number of data points,lag distance,and number of lags used as inputs in ordinary and universal kriging is summarized in Table 3-6.

Table 3-6. Summary of Variables Used as Input for Semivariogram Models for Ordinary and Universal Kriging

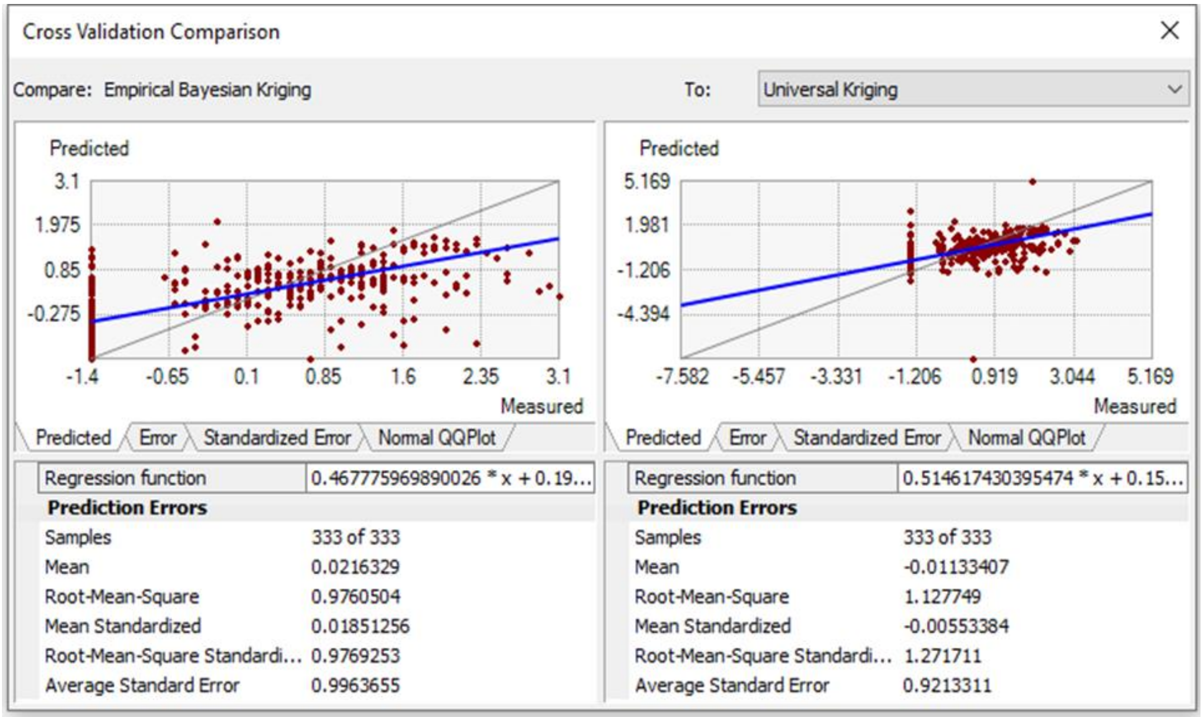
Variable	Number of sites	Lag distance (Feet)	Number of lags	Large Scale trend
Arsenic	333	5841.03	100	1st order
PC1_P	333	5841.03	100	2nd order
PC2_P	333	5841.03	100	1st order
PC1_S	333	5841.03	100	1st order
PC2_S	333	5841.03	100	1st order
PC3_S	333	5841.03	100	1st order

A scatter plot of predicted values true values shown in cross validation plots for log transformed arsenic(Figure 3-11) and cross validation plots for all the other variables are shown in Appendix A (Figure A- 1 through Figure A- 5).The slope is expected to be 1, but it is usually less than 1 as it is property of kriging that tends to underpredict large values and overpredict small values. The fitted line through the scatter of points is given in blue with the equation given just below the plot. The error plot is the same as the prediction plot, except the measured values are subtracted from the predicted values. For the standardized error plot, the measured values are subtracted from the predicted values and divided by the estimated kriging standard errors. All three of these plots show how well kriging is predicting. If all the data were independent (no autocorrelation), all predictions would be the same (every prediction would be the mean of the measured data), so the blue line would be horizontal. With autocorrelation and a good kriging model, the blue line should be closer to the 1:1 gray line. The regression equation below each of these three plots is calculated using a robust regression equation. This procedure first fits a standard linear regression line to the scatterplot. Next, any points that are more than two standard deviations above or below the regression line are removed, and a new regression equation is calculated. This procedure ensures that a few outliers will not corrupt the entire regression equation.

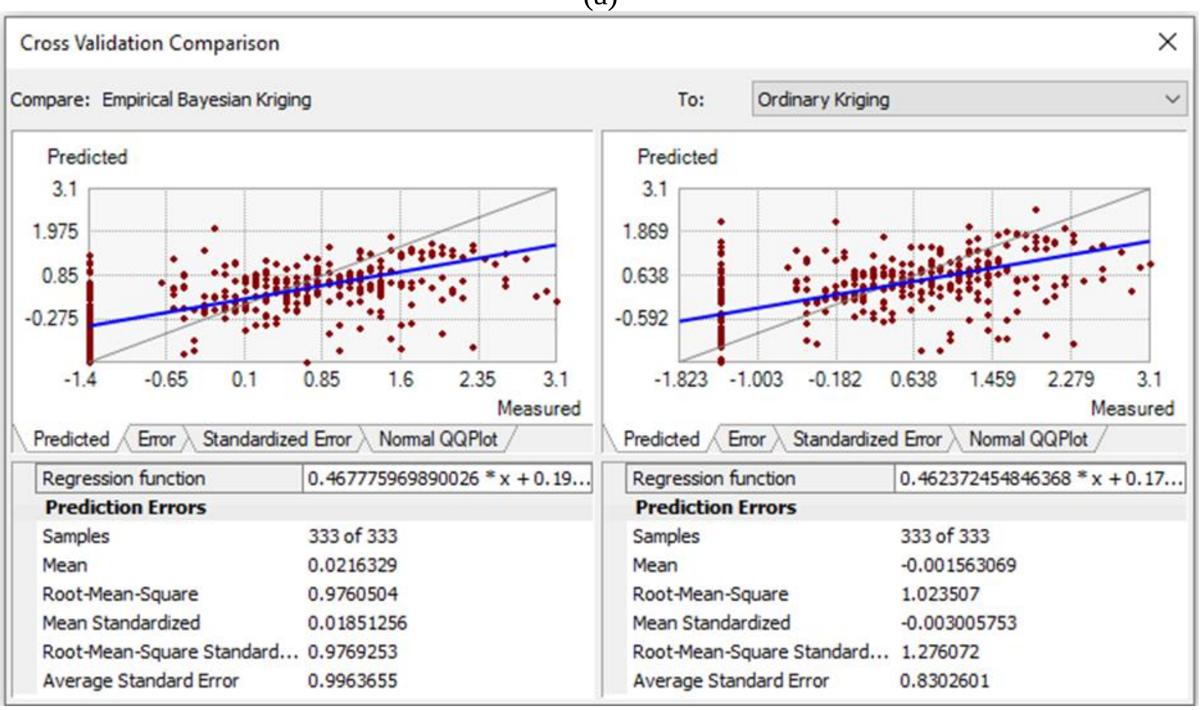
Normal QQ plot for log transformed arsenic obtained from cross validation comparison of (a) empirical Bayesian kriging and universal kriging (b) empirical Bayesian kriging and ordinary kriging (Figure 3-12). The Normal QQ Plot graph shows the quantiles of the difference between the predicted and measured values and the corresponding quantiles from a standard normal distribution. If the errors of the predictions from their true values are normally distributed, the points should lie roughly along the gray line. If the errors are normally distributed, methods that rely on normality can be used as empirical Bayesian kriging in this case.

Table 3-7. RMSE Values for Each Variable Obtained from Cross Validation of Ordinary Kriging, Universal Kriging and Empirical Bayesian Kriging. ***The Lowest RMSE for Each Parameter is Bolded***

	Ordinary Kriging	**Universal Kriging**	**Empirical Bayesian Kriging**
Arsenic	1.0235	1.1277	**0.9759**
$PC1_{primary}$	0.9963	1.0022	**0.9599**
$PC2_{primary}$	0.9716	0.9162	**0.8675**
$PC1_{secondary}$	0.8128	0.8144	**0.7895**
$PC2_{secondary}$	0.9885	0.9874	**0.9654**
$PC3_{secondary}$	1.1023	1.1444	**0.9499**

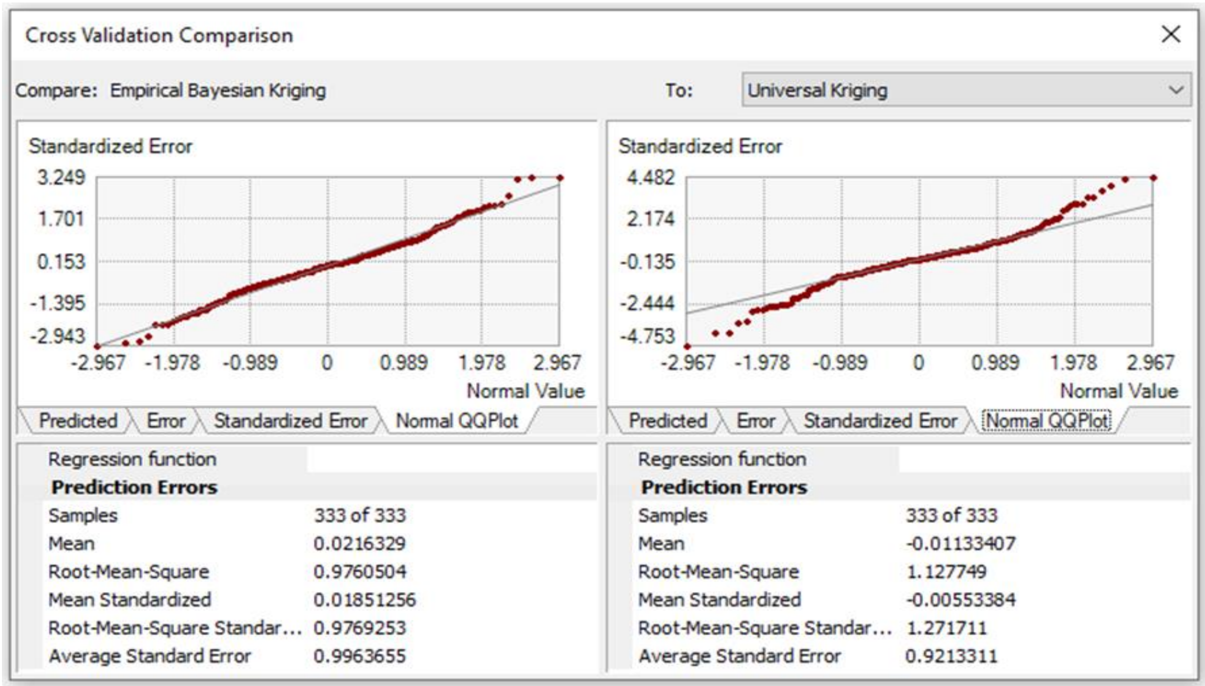


(a)

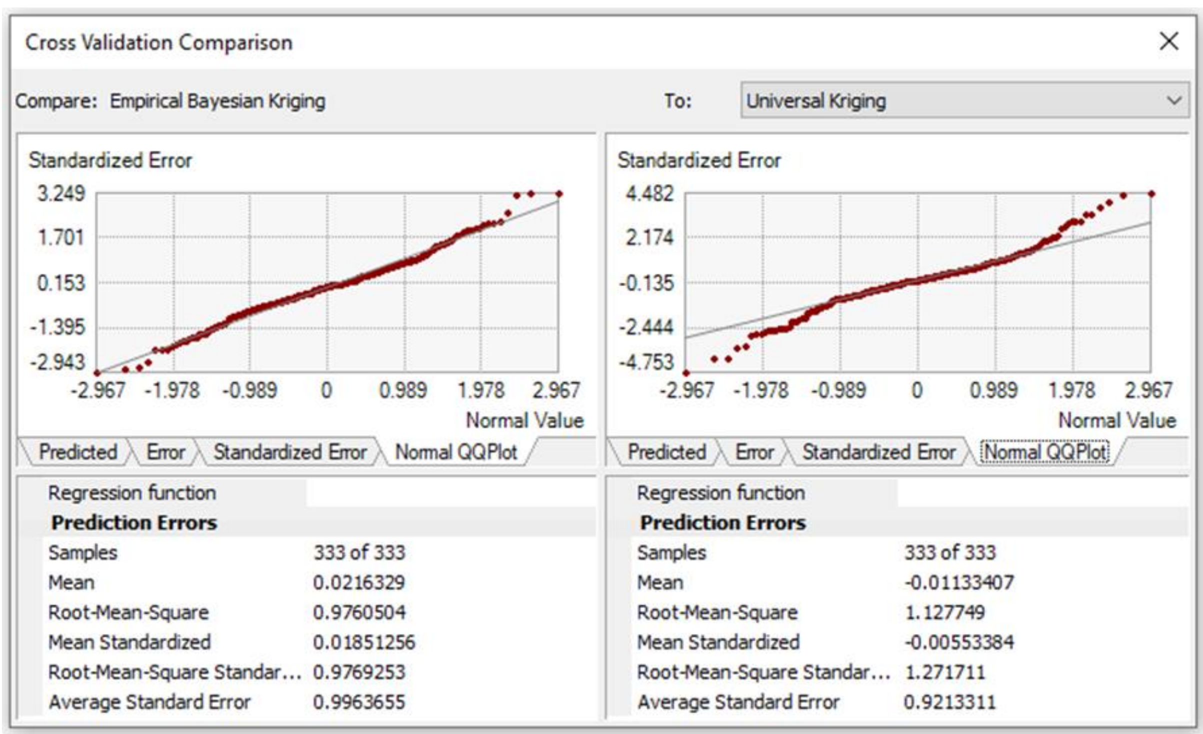


(b)

Figure 3-11. Predicted vs Measured Plots for Log Transformed Arsenic Obtained from Cross Validation of (a) Empirical Bayesian Kriging and Universal Kriging and (b) Empirical Bayesian Kriging and Ordinary Kriging



(a)



(b)

Figure 3-12. Normal QQ Plot of (a) Empirical Bayesian Kriging vs Universal Kriging and (b) Empirical Bayesian Kriging vs Ordinary Kriging

The prediction maps of the three kriging methods for log-transformed arsenic are shown in Figure 3-13 to Figure 3-15. Empirical Bayesian kriging prediction maps for the rest of the variables are shown in Figure 3-16 to Figure 3-20. The RMSE values for each of the kriging methods for each of the variables indicate that overall, the empirical Bayesian kriging resulted in the lowest RMSE values for all the variables as shown in Table 3-7. However, significant differences in the RMSE values were not observed among the three kriging methods. Another study in Texas compared geostatistical methods among regions and found the performance of different methods varied less within a given area than across the different regions (Gong et al., 2014). Because the differences in processes influencing spatial variability of arsenic are at different scales, it is possible that developing kriging models over smaller targeted areas with a high density of samples could have improved the performance of predictions.

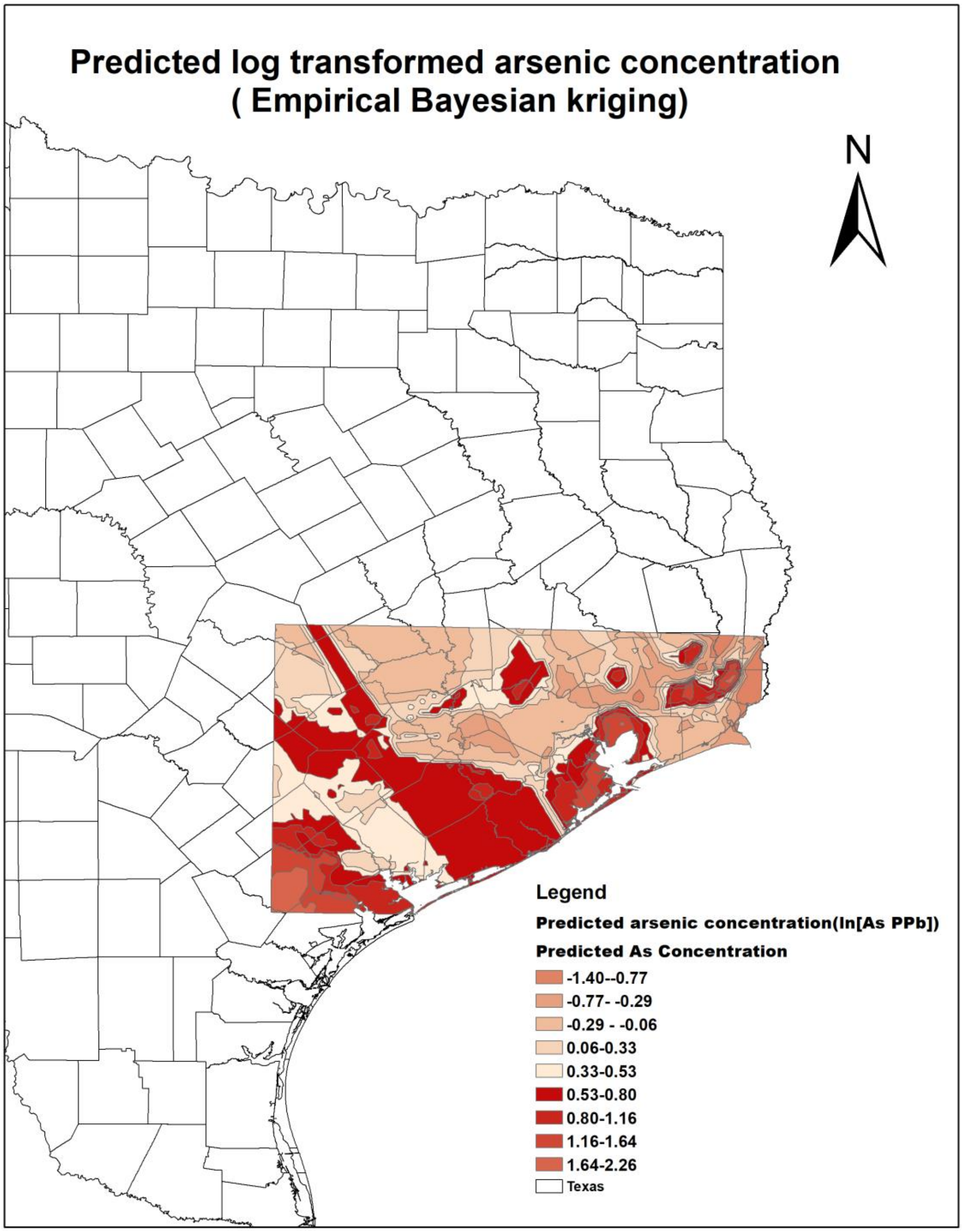


Figure 3-13. Prediction Map of Log Transformed As Using Empirical Bayesian Kriging

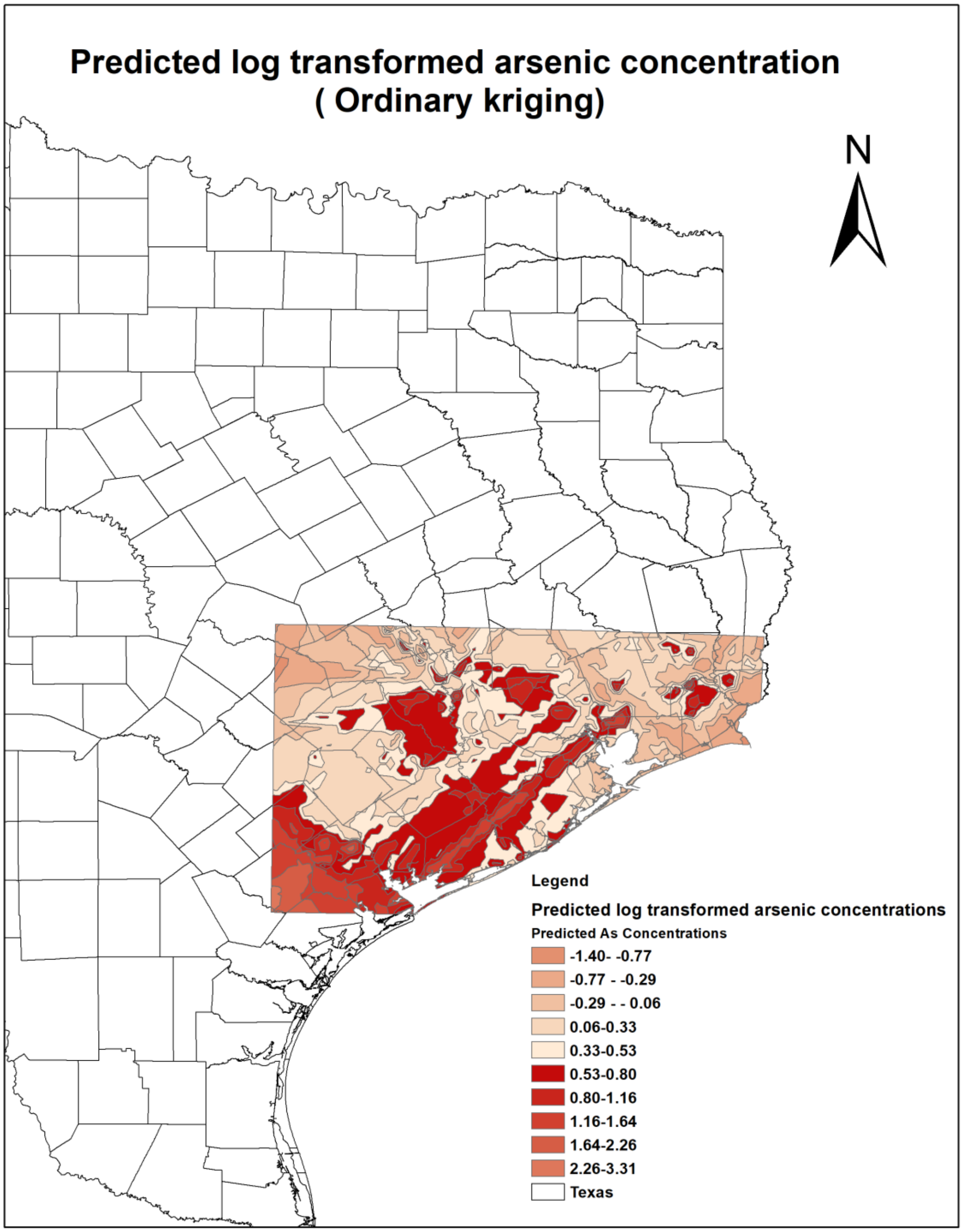


Figure 3-14. Prediction Map of Log Transformed As Using Ordinary Kriging

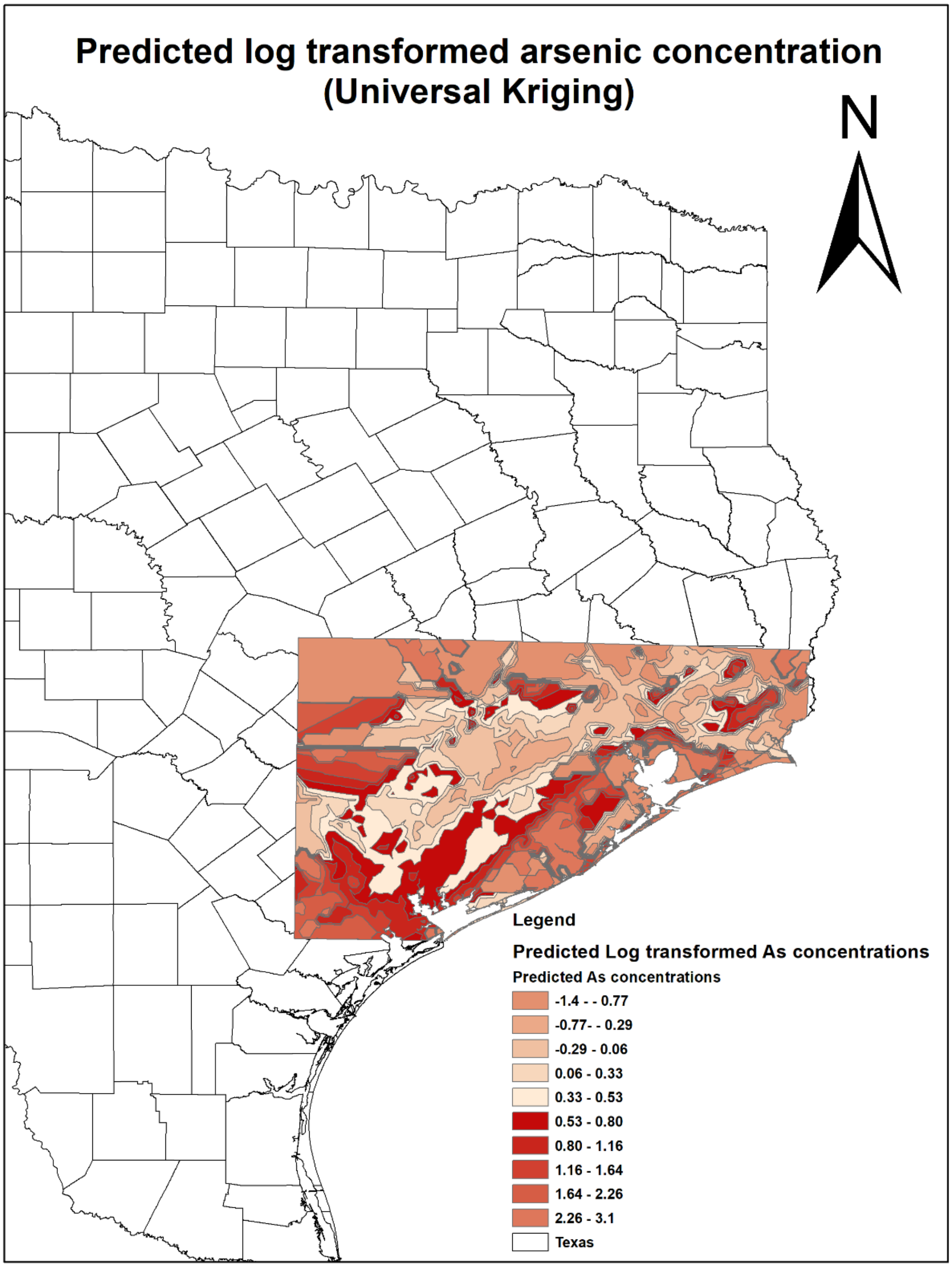


Figure 3-15. Prediction Map of Log Transformed As Using Universal Kriging

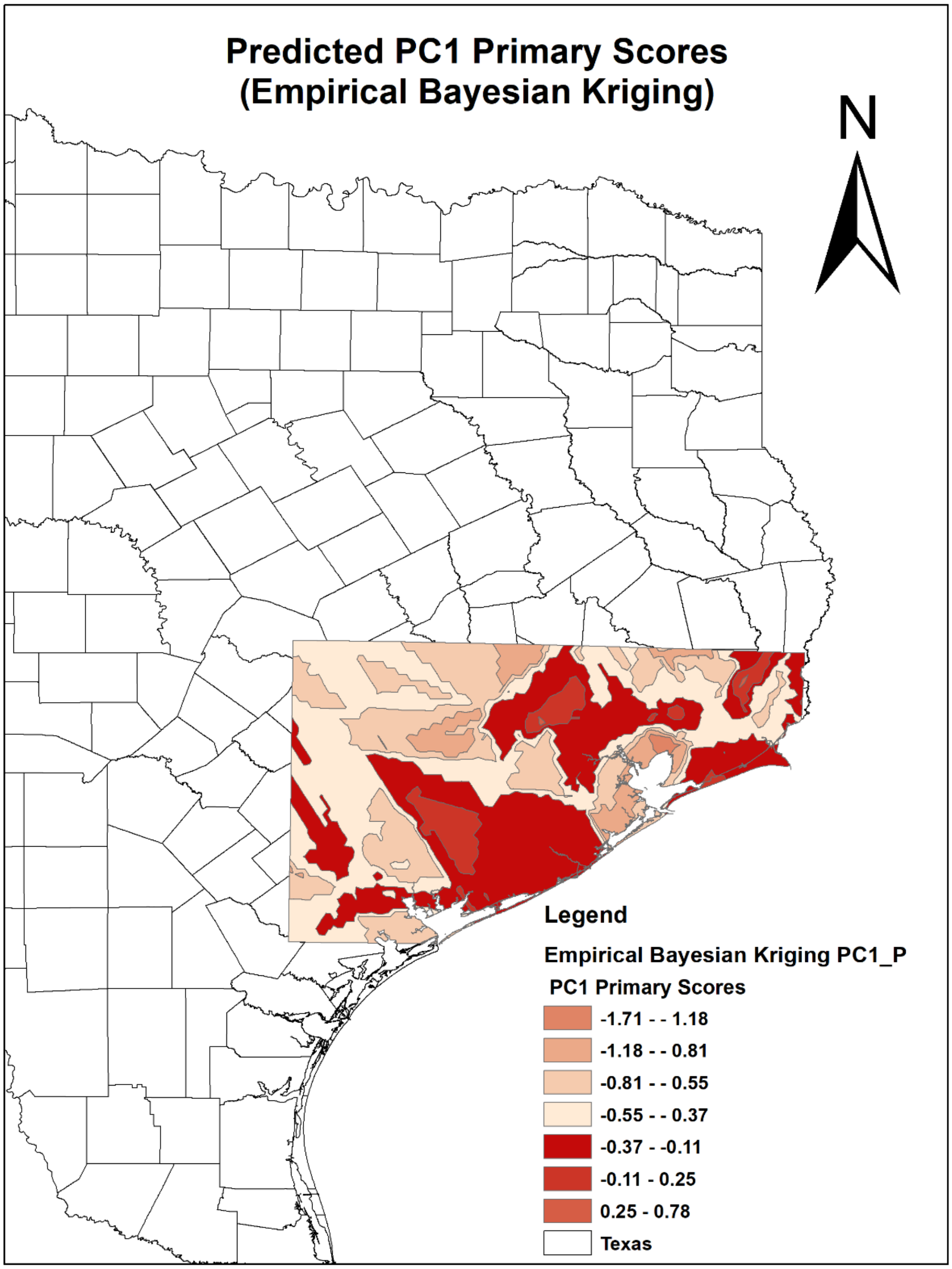


Figure 3-16. Prediction Map of PC1 Primary Scores Using Empirical Bayesian Kriging

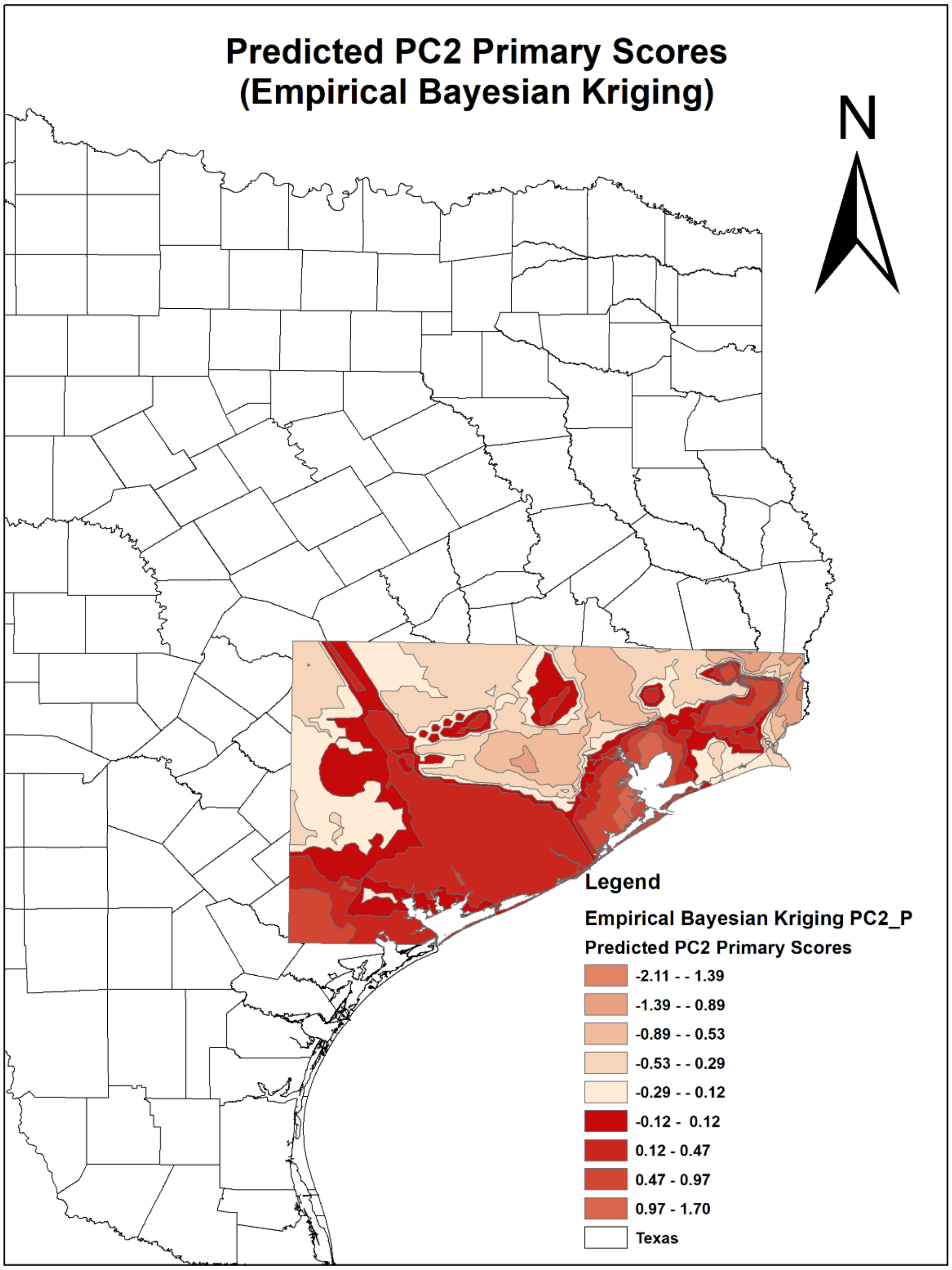


Figure 3-17. Prediction Map of PC2 Primary Score Using Empirical Bayesian Kriging

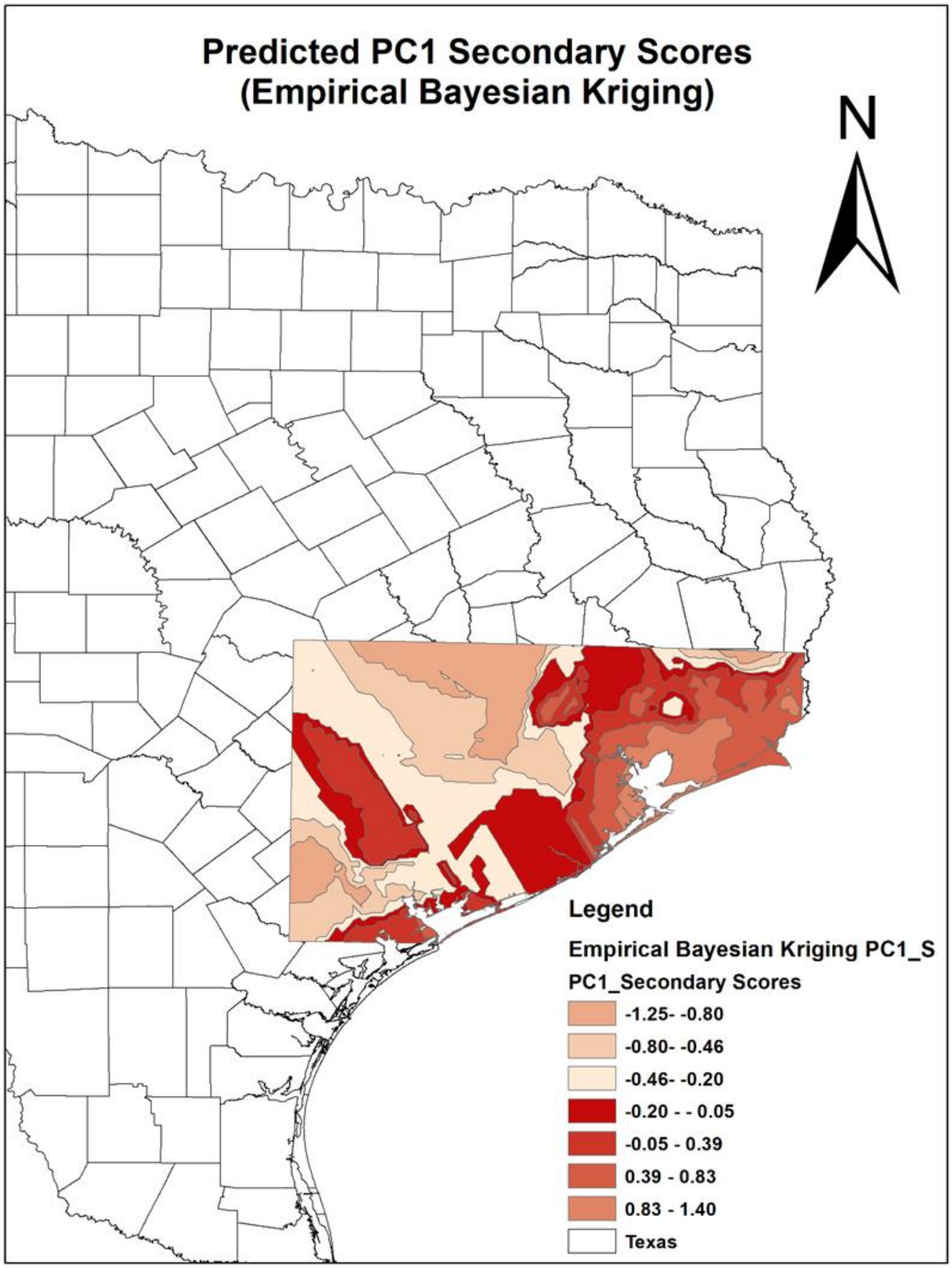


Figure 3-18. Prediction Map of PC1 Secondary Score Using Empirical Bayesian Kriging

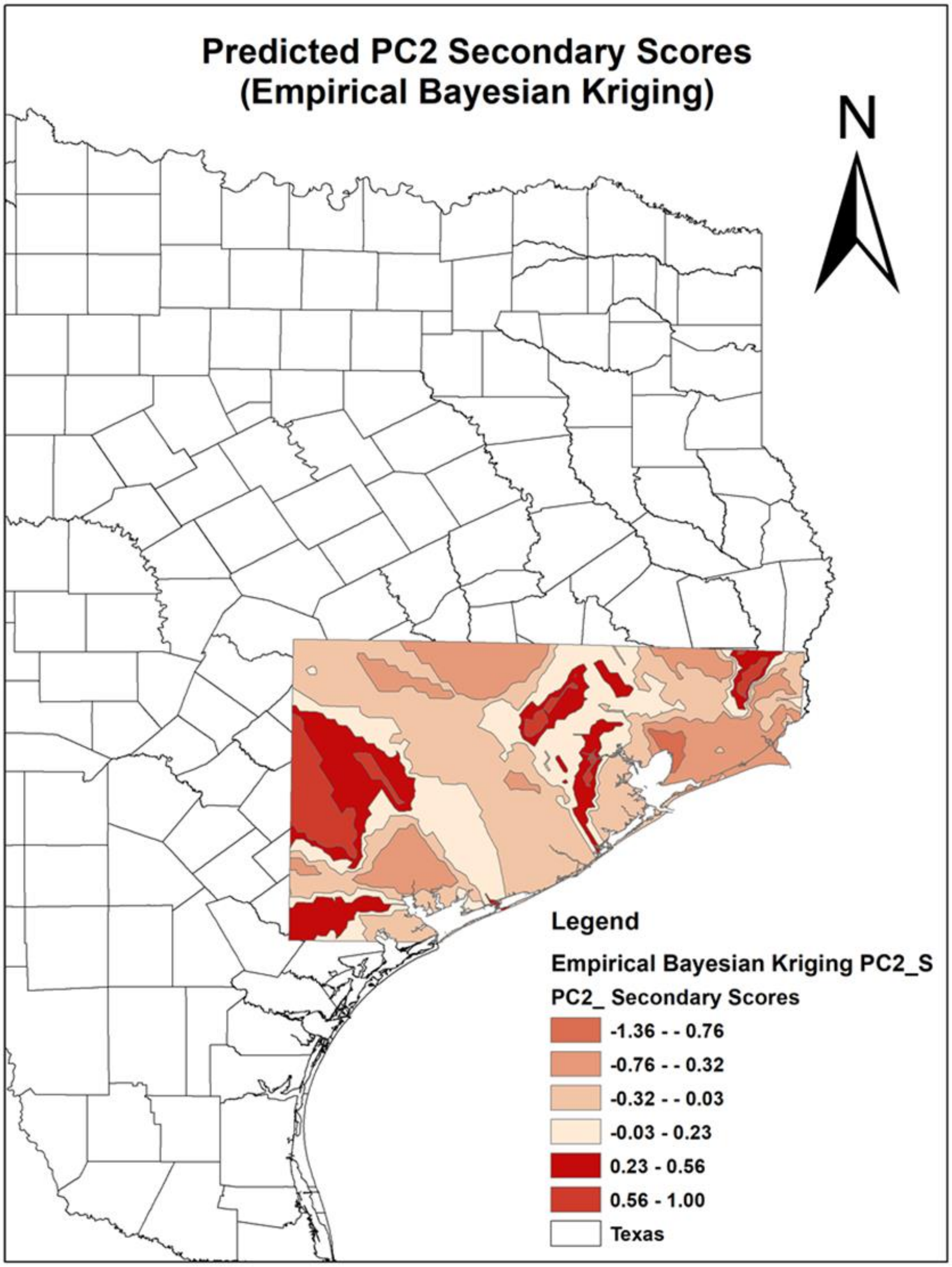


Figure 3-19. Prediction Map of PC2 Secondary Score Using Empirical Bayesian Kriging

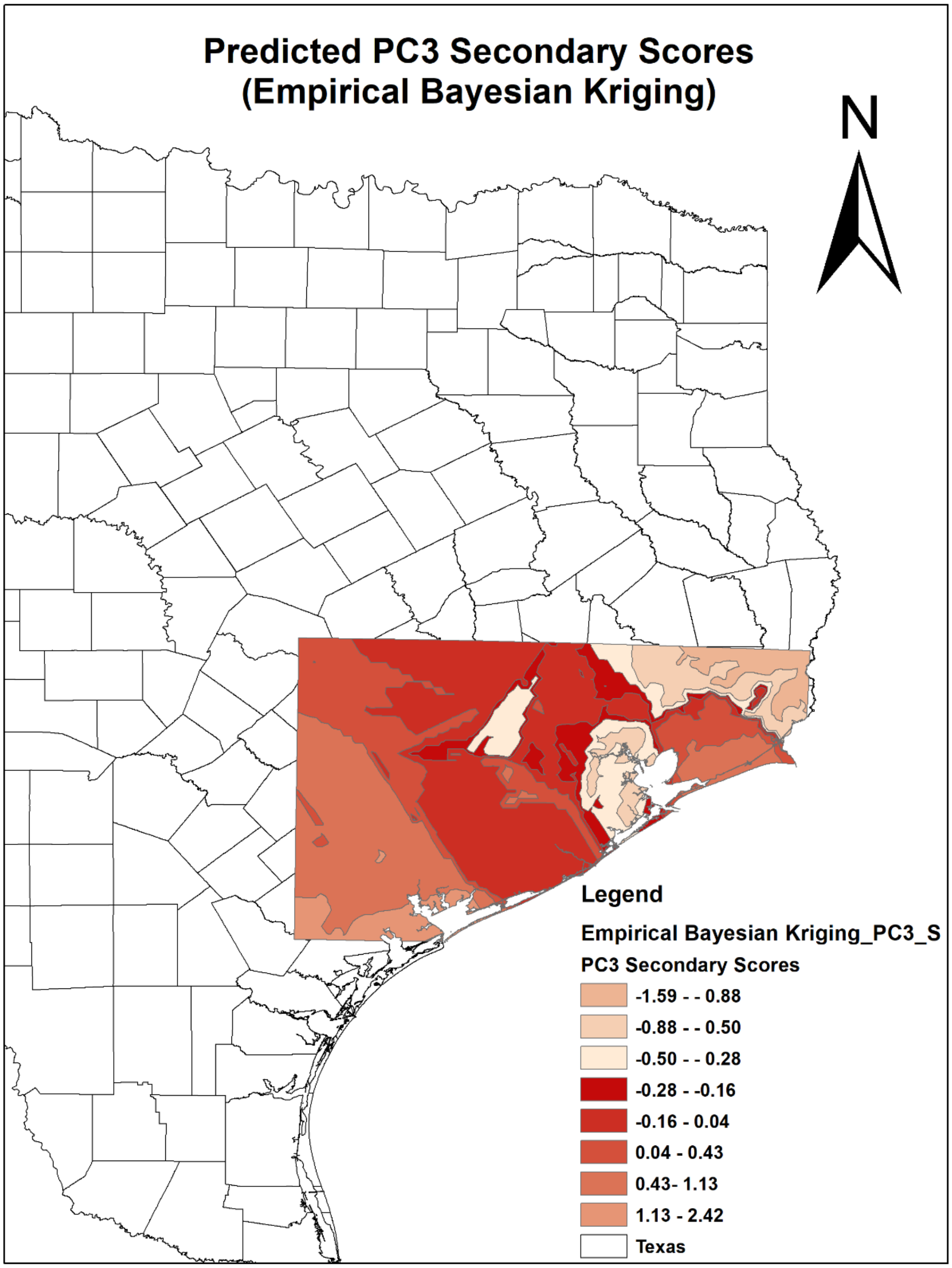


Figure 3-20. Prediction Map of PC3 Secondary Score Using Empirical Bayesian Kriging

4. DISCUSSION

4.1. Well Constituents

Most of the water samples analyzed did not exceed the primary MCL for most of the elements; however, many samples tested positive for arsenic, barium and copper, highlighting the significance of testing drinking water after a catastrophic event such as Hurricane Harvey to avoid health risks.

Arsenic in drinking water increases the risk of getting cancer (Celik et al., 2008). Previous studies conducted in water systems in the major aquifers of Texas from 1992-2017 indicated the highest concentrations of arsenic occurring in the southern Gulf Coast aquifer. Concentrations of arsenic were found to decrease downdip from the Catahoula Formation, consistent with Miocene volcanic ash being the primary source of groundwater arsenic in this region (Gates et al., 2011). About 43,000 people are potentially affected by arsenic from this aquifer, making them the most arsenic-affected population in Texas (Reedy et al., 2018). Spatial distribution of the groundwater arsenic concentration in the Gulf Coast Aquifer showed arsenic concentration to increase towards the south of the aquifer making it an arsenic hotspot. About 74% of the area has very low to moderate probability of having arsenic > 5 µ/L. Most of the wells for which arsenic was tested were in the low to moderate probability area and there was a 3.4% chance of the arsenic concentration exceeding the EPA standards (Reedy et al., 2018). (Hudak, 2003) also found that geology better predicted the occurrence of arsenic than agricultural practices in a study done on 69 wells in the Gulf Coast Aquifer.

Ninety-four percent of samples tested positive for barium in this study with a mean concentration of 182.6 µ/L (median 163.7 µ/L). No samples exceeded the MCL for barium. Drinking water should be monitored regularly for Ba as studies have shown elevated levels of barium to be associated with high blood pressure and higher death rates related to cardiovascular disease (Brenniman et al., 1979). An elevated level of barium in drinking water used without filter or treatment could most likely be due to presence of industrial waste, mixing of natural saline and brine water and saltwater intrusion. The median barium concentration was also higher for wells having submerged well heads.

Pizarro et al. (1999) found elevated levels of copper to have acute gastrointestinal effects on the human health, with no association between aggregate copper, in 0-5 mg/L drinking water and diarrhea. Three mg/L Cu concentration was associated with nausea, abdominal pain or pain. Copper was detected in 81% percent of samples, but exceeded primary drinking water standards for 0.2 % of the samples. The median copper concentration was higher for well systems damaged during Harvey. A significant difference in median copper concentration was observed for wells affected by Harvey compared to data from wells under no flood conditions.

Chromium in drinking water is considered to have carcinogenic effects and exposure of animals to Cr (VI) drinking water induces tumors in the alimentary tract (Zhitkovich, 2011). The median chromium concentration was higher for wells with submerged well heads than for well heads not submerged.

A recent study shows the association of nitrate in drinking water to increased risk of colorectal cancer due to endogenous transformation into carcinogenic N-nitroso compounds (Schullehner et al., 2018). Shallow wells were found to be susceptible to nitrate contamination compared to deeper wells in a study conducted in south-central Kansas attributable to land use practices and sub-surface clay lenses (Townsend and Young, 1995). An inverse correlation between nitrate and well depth was found in a study of 69 wells in the Gulf Coast Aquifer of South-Central Texas (Hudak, 2003). Nitrate contamination was found to be higher in wells of Minnesota that were 15 - 30 years old in comparison to wells constructed less than 15 years previously (Lewandowski et al., 2008).

Concentrations of aluminum, copper, iron, manganese, zinc and silver exceeded established SMCLs for some individual wells. Onufrak et al. (2014) found that the aesthetic quality of tap water could affect health by increasing consumption of water alternatives, which may include sugar-sweetened beverages. Iron and manganese exceedances can cause water to be unpalatable. Although a primary (health-related) drinking water standard for manganese has not been established, active research continues in this area as Mn concentrations above the SMCL may be related to insomnia, weakness of legs and even nervous system disorders (EPA, 2015c). Gauthier et al. (2000) confirmed the importance of aluminum compounds in drinking water and consideration of effects on genetic characteristics related to aluminum exposure and Alzheimer's disease. Children, transients and the elderly are such populations because of the potentially high risk of dehydration from diarrhea that may be caused by high levels of sulfate in drinking water. Drinking water containing ≥750 mg of sulfate per liter was associated with a self-reported laxative effect, whereas water containing < 600 mg/ L was not. The presence of sulfate in drinking-water can also result in a noticeable taste; the lowest taste threshold concentration for sulfate is approximately 250 mg/liter as the sodium salt. Sulfate may also contribute to the corrosion of distribution systems (World Health Organization, 2004).

High concentrations of dissolved iron within the well bore may lead to growth of iron bacteria. These bacteria may coat the inside of the casing or any other submerged part of the plumbing in the well and may cause problems. Median iron concentration was higher for wells with submerged well heads than for well heads not submerged.

Manganese concentration was significantly and positively associated with the prevalence of delayed milestones and hearing loss in the children of North Carolina (Langley et al., 2015). It is also believed to cause aesthetic and economic damage and imparts brownish stains to laundry. Manganese affects the taste of water and causes dark brown or black stains on plumbing fixtures. Median manganese concentrations were higher for wells with submerged well heads than for well heads not submerged.

A report published by World Health Organization indicated higher concentrations of zinc in drinking water cause an undesirable, astringent taste. Pipeline water supplies with galvanized plumbing material record higher levels of zinc (World Health Organization, 1996), which could possibly be a reason for higher zinc concentrations in well systems damaged due to flooding than wells system which were not damaged. The concentration of zinc in drinking water may increase as a result of the distribution system and household plumbing.

Chloride increases the electrical conductivity of water and thus increases its corrosivity. In metal pipes, chloride reacts with metal ions to form soluble salts, thus increasing levels of metals in drinking water. In lead pipes, a protective oxide layer is built up, but chloride enhances galvanic corrosion. It can also increase the rate of pitting corrosion of metal pipes (World Health Organization, 2003).

A significant difference in zinc concentration for wells impacted by Harvey compared to no flood conditions is likely caused by corrosion resulting from flooding for submerged well heads and damaged well systems.

The World Health Organization guidelines for drinking water quality recommend that there should be no detectable *E. coli* bacteria or Fecal Indicator Bacteria in a 100 ml drinking water sample (World Health Organization, 1997). However, a significant presence of *E. coli* after the massive flood indicates that water from some of the wells would not be suitable for drinking. Increased incidence of contamination could have resulted from the well head being submerged under flood waters. Damage to well systems could have also led to contamination. Around 34% of the well heads tested for bacteria were submerged, which likely also led to bacterial contamination.

Gonzales (2008) found that well head protection does play a role in bacterial contamination of water wells. Gonzales reported that a strong statistical difference in total coliform bacteria contamination for wells with varied well head protection. It was found that all the wells having poor well head protection tested positive for total coliform, while 60% of the wells having fair well head protection tested positive for total coliform. Moreover, nine percent of wells having good well head protection tested positive for total coliform. Around 18% of the well systems in the samples analyzed for bacteria were damaged during Hurricane Harvey with 80% of those wells having submerged well heads, which likely contributed to the bacterial contamination indicated when these wells were tested immediately after the hurricane.

4.2. Principal Component Analysis

PCA has been used to interpret patterns of groundwater quality parameters. It is used to identify common factor patterns and interpret them with respect to presumed natural and anthropogenic processes which impact groundwater quality and mostly focus on ions such as chloride, magnesium, and sulfate, all of which have secondary drinking water standards. PCA analysis of groundwater mostly includes nitrate as a marker for anthropogenic influences on groundwater (Belkhiri et al., 2011). However, previous studies lack consistency with parameters, especially inclusion of trace elements, making it difficult to compare with the present study.

The low cumulative variance of 57.5% for elements with primary standards can be attributed to the low Kaiser's measure of sampling adequacy of 0.52. A KMO statistic of 0.5 and above is acceptable for performing PCA, but according to (Kaiser 1970) a value between 0.6-1 is recommended to get better results. Removing the parameter with the lowest individual measure of sampling adequacy (barium) did not improve the KMO value. If possible, increasing the sample size could have been a remedial measure to improve the results for elements with primary standards.

PCA gave higher cumulative variance of 69.3% for elements with secondary standards as three PCs were retained in comparison to two PCs retained for elements with primary drinking water standards according to the eigenvalue one criterion and the scree plot. One of the other reasons for higher cumulative variation for elements with secondary standards is a higher KMO value (KMO=0.54) when compared to elements with primary drinking water standards (KMO=0.52).

4.3. Kriging

In the present study, empirical Bayesian kriging had the lowest RMSE for the greatest number of variables and was considered the optimal method based on our data. However, a significant difference was not observed in the RMSE values of ordinary kriging and universal kriging to that of empirical Bayesian kriging. Hence, there does not appear to be much difference in these methods in the accuracy of predicted values on cross validation. Similar results were observed in a study done on the wells in Saskatchewan, Canada (McLeod et al., 2017). Gong et al. (2014) compared inverse distance weighed interpolation with kriging using Gaussian and spherical models as well as cokriging in predicting arsenic concentrations over various regions in Texas, and found regional differences in the performances of kriging, and concluded that kriging over smaller areas was more accurate than over large geographic regions. James et al. (2014) evaluated the performance of various kriging methods (ordinary, universal, simple kriging with varying means, kriging with external drift, cokriging with ordinary kriging and cokriging with universal kriging) over a relatively small area in Colorado and found that ordinary kriging performed best.

All the wells in the present study were located in the Gulf coast aquifer and we did not find a trend in arsenic concentrations with well depth. Hence, well depth was not considered while performing kriging. Furthermore, a previous study demonstrated that including well depth in cokriging did not necessarily improve the ability of kriging to predict arsenic levels. Gong et al. found that incorporating well depth in cokriging did not necessarily improve the correlation between predicted and actual values. Yu et al. (2003) investigated factors affecting arsenic at different geographic scales and concluded that much of variability in arsenic concentrations at a scale of less than 3 km could be explained by well depth, while geology was the most important factor at scales greater than 10 km. Hence, for the large area of the present study in comparison to the reported studies, it is unlikely that adding well depth as a covariate would have improved our models.

Interpretation of mapped results of PCA is less straightforward because the values are a representation of parameters that contribute to the PCA components. The areas with high values of the first PC for elements with primary drinking water standards represents higher predicted concentrations of the contributors of this component namely copper and chromium. Moreover, the areas with high values of the second PC for elements with primary drinking water standards represents higher predicted concentrations of arsenic and nitrate since these concentrations were strongly related to the second PC. Similarly, the areas with high values of first PC for elements with secondary drinking water standards represents higher predicted concentrations of iron and manganese. Moreover, the areas with high values of second PC for elements with secondary drinking water standards represents higher predicted concentrations of copper and zinc while the areas with high values of third PC for elements with secondary drinking water standards represents higher predicted concentrations of chloride and sulfate.

However, this analysis was intended to estimate the mean arsenic concentrations along with principal components representing drinking water quality in private wells in the area impacted by Hurricane Harvey. Previous studies have used geostatistical methods to map the PCA scores to predict factors that may have impacted groundwater quality, such as pollution or saltwater intrusion. However, a sufficient number of studies have not been conducted to compare the performance of kriging to accurately predict PCA scores making it difficult to compare the results obtained in the present study.

5. CONCLUSIONS

In this study, the use of PCA and kriging was investigated to predict the groundwater quality of private wells in the counties affected by hurricane Harvey. About 43.9% of the wells exceeded at least one EPA drinking water standard. As was detected in 81% of the samples and it exceeded the EPA drinking water standard in about 3.4 % of the wells.

Pb, with an action level of 15 µ/L, was detected in about 25% of the wells (samples were collected after five minutes of flushing), while Copper, with an action level of 1300 µ/L, was detected in about 81% of the samples. Iron and manganese exceeded the secondary drinking water standards in 22.9 % and 23.6% of the wells, respectively and it is likely those waters were unpalatable. Total coliform was detected in 46.5% of the wells, while *E. coli* was detected in 11.9% of the wells.

System characteristics of the wells also likely affected concentrations of those elements with primary and secondary water standards. While there was a significant difference ($p<0.05$) in element concentrations between various well ages for Cu, Cr, NO_3, and SO_4, with the exception of Cl, no data trends were apparent. SO_4 concentration showed a trend with well depth as it varied inversely with well depth. Contaminant concentrations were significantly different for Fe and Mn when collected from a well head in comparison to kitchen tap collection. A submerged well head led to increased As, Ba, Cr, Fe, Mn, Zn and total coliform concentrations when compared to wells which did not have submerged well heads. Mn concentrations were higher in wells damaged due to Harvey than for undamaged wells. Wells used as a source for filtered drinking water had higher SO_4 concentrations than wells were not used for drinking water.

Cu and Zn showed a significant difference in the wells affected by Hurricane Harvey when compared to previous five year data as the median Cu and Zn concentrations were higher in wells in the Hurricane Harvey affected area.

For the PCA analysis, all the elements with secondary drinking water standards were included which had a cumulative variance of 69.3% explained by three PCs, while all elements with primary drinking water standards except for barium was included which had a cumulative variance of 57.5% explained by two PCs, as the latter had a low KMO statistic.

Three kriging methods including ordinary, universal and empirical Bayesian kriging were compared for predicting water quality focusing on arsenic and PC scores across the study area. Empirical Bayesian kriging resulted in the greatest accuracy of predicted values with the lowest RMSE values for most of the variables.

6. RECOMMENDATIONS AND FUTURE WORK

Immersion of the well head by flood waters increased well water As, Ba, Cr, Fe, Mn, Zn and total coliform concentrations. This result emphasizes the importance of protecting the well head to prevent future well contamination. If a well head has been flooded, the height of the well head casing should be increased to prevent future flooding of the well head. Typically, well water casing extends 1 to 2 feet above surrounding land, preventing surface water from running down the casing or on top of the cap and into the well (Harris et al., 1996).Texas Water Code requires that at least 12 inches of casing pipe extend above the top of the slab. The code also requires that the casing extend at least 2 feet above flood level (Texas Administrative Code Chapter 76, 2014). Well heads that were flooded during Hurricane Harvey should be retrofitted to be in compliance with this requirement of the Texas Water code.

At a minimum, well owners should have their well tested yearly by a certified laboratory for *E.Coli*. It is also recommended that well water is tested for nitrate-nitrogen and total dissolved salts every few years. If the water is a drinking water source, arsenic testing is also recommended for wells in the Gulf Coast Aquifer . Well owners and private water well drillers are required to follow the Texas Water Code regarding the location of the well head (Uhlman et al., 2012). Specifically, a well head should be atleast 50 feet from a septic tank, 100 feet from any leach field, 150 feet from any pesticide or fertilizer storage area and 250 feet from liquid waste disposal system.

Although statistically significant differences were not detected in this study,well age is probably an important factor in predicting the likelihood of contamination. A well constructed more than 50 years ago is likely to be near the center of the property, and it could be a relatively shallow well probably surrounded by many potential contamination sources. Older well pumps are more likely to leak lubricating oils into the well. Older wells also are more likely to have thinner casing that may be corroded. Even modern casing in wells 30 to 40 years old is subject to corrosion and perforation. Thus, older wells should be inspected by a qualified well driller (Harris et al.,1996).

Good maintenance means testing the water every year, keeping the well area clean and accessible, keeping pollutants as far away as possible and periodically (every few years) having a qualified well driller or pump installer check well mechanics.

Older wells without proper casing, sealing and protective slabs, or with other problems, should be brought up to current standards. Wells should be upgraded by removing well pits, installing caps, extending casings as appropriate, and moving such activities as pesticide mixing, tank rinsing or gasoline storage farther from the well.

In the present study, well water quality was analyzed for up to six months post-Harvey. Future research work investigating the longer- term, temporal effects would be beneficial. Studies related to potential risk characterization due to exposure of the FIB post-Harvey could be achieved by the Quantitative Microbial Risk Assessment modeling.

REFERENCES

Agency for Toxic Substances and Disease Registry. (2005). *Toxicological Profile of Zinc*. Retrieved from https://www.atsdr.cdc.gov/toxprofiles/tp60-c6.pdf

Bartlett M.S. (1950). Tests of significance in factor analysis. *British Journal of Mathematics and Statistical Psychology*,3, 785.doi:10.1111/j.20448317.1950.tb00285.x

Beall G. (1942). The Transformation of Data from Entomological Field Experiements.*Biometrika*,29,243-262.

Belkhiri L., Boudoukha A., Mouni L. (2010) A multi variate statistical analysis of groundwater chemistry data. *International Journal of Environmental Research*,5, 537-544.doi:10.22059/IJER.2011.338

Brenniman G., Namekata T., Kojola W., Carnow B., & Levy P. (1979).Cardiovascular disease death rates in communities with elevated levels of barium in drinking water. *Environmental Research*,20(2),318-324. doi : 10.1016/0013-9351(79)90007-0

Chang K. (2009). Introduction to Geographic Information Systems with Data Files and CD-ROM, 5th edition. *McGraw-Hill Science/Engineering/Math*, ISBN 978-0-07-729436-6.

Cattell R.B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research,1,*245-276.

Celik I., Gallicchio L., Boyd K., Lam T.K., Matanoski G., Tao X., Shiels M., Hammond E., Chen L., Robinson K.A., et al. (2008). Arsenic in drinking water and lung cancer: A systematic review. *Environmental Research*,108,48-55. doi:10.1016/j.taap.2006.12.032

Cloutier V., Lefebvre R., Therrien R., Savard M. M. (2008). Multivariate statistical analysis of geochemical data as indicative of the hydro geochemical evolution of groundwater in a sedimentary rock aquifer system. *Journal of Hydrology*,353, 294-313. doi:10.1016/j.jhydrol.2008.02.015

Conover W.J. (1994). *Practical Nonparametric Statistics*. New York, NY: John Wiley & Sons, Inc.

Dai D., et al. (2019). Molecular survey of *Legionella* and *Naegleria fowleri* in private well water and premise plumbing following the 2016 Louisiana flood. *Environment Science: Water Research & Technology*,5,1464-1477. doi :10.1039/C9EW00109C

DeSimone L. (2009). The quality of our nations waters-quality of water from domestic wells in principal aquifers of the United States,1991-2004- Overview of major findings *U.S. Geological Survey Circular*,1332,48.

Edberg S.C., Rice E.W., Karlin R.J., Allen M.J. (2012). *Escherichia coli*: the best biological drinking water indicator for public health protection, *Journal of Applied Microbiology*. doi :10.1111/j.1365-2672.2000.tb05338.x

Environmental Protection Agency. (1984). *Methods for chemical analysis of water and wastes*. Retrieved from https://archieve.epa.gov/water/archieve/web/html/132.html

Environmental Protection Agency. (2015b). *National Primary Drinking Water Regulations*. Retrieved from http://water.epa./drink/contaminants/index.cfm#Primary

Environmental Protection Agency. (2017). *Regional Guidance on Handling Chemical Concentration Data near the Detection Limit in Risk Assessments.* Retrieved from https://www.epa.gov/risk/regional-guidance-handling-chemical-concentration-data-near-detection-limit-risk-assessments

Field A. (2000). *Discovering Statistics using SPSS for Windows*. London-Thousand Oaks-New Delhi: Sage Publications.

Gates et al. (2011). Arsenic enrichment in unconfined sections of the southern Gulf Coast aquifer system, Texas, *Applied Geochemistry*,26,421-431. doi:10.1016/j.apgeochem.2011.01.002

Gauthier E., Fortier I., Courchesne F., Pepin P., Mortimer J., & Gauvreau, D.(2000) Aluminum forms in drinking water and risk of Alzheimer's disease. *Environmental Research*,84(3),234-46. doi :10.1006/enrs.2000.4101

Goldberg D. W. (2011). Improving geocoding match rates with spatially-varying block metrics. *Transactions in GIS* ,15,6,829-850. doi: 10.1111/j.1467-9671.2011.01295.x

Gong G., Mattevada S., O'Bryant S. E. (2014) Comparison of kriging and IDW interpolations in estimating groundwater arsenic concentrations in Texas, *Environment Research*,130,59-69. doi:10.1016/j.envres.2013.12.005

Gonzales T.R. (2008). The effects that well depth and well head protection have on bacterial contamination of private water wells in the Estes Park Valley,Colorado. *Journal of Environmental Health*,71(5),17-23. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/19115719

Goodchild M.F. (1993). The state of GIS for environmental problem-solving, *Environmental Modeling with GIS*,8-15.

Güler C., Thyne G.D., McCray J.E., Turner K. A. (2002). Evaluation of graphical and multivariate statistical methods for classification of water chemistry data. *Journal of Hydrogeology*,10,455-474. doi :0.1007/s I0040-002-0196-6.

Harris B.L., Hoffman D.W.,& Mazac F.J, Jr.(1996) Reducing the Risk of Ground Water Contamination by Improving Wellhead Management and Conditions. Retrived from http://publications.tamu.edu/WATER/PUB_water_Reducing%20the%20Risk%20of%20Groundwater%20Contamination.pdf

Helena B., Pardo R., Vega M., & Barrado E., Fernandez J.M., Fernandez L. (2000). Temporal evolution of groundwater composition in the alluvial aquifer (Pisuerga River, Spain) by principal component analysis. *Water Research*, 34,807-816. doi:10.1016/S0043-1354 (99)000225-0

Higgins J.J. (2004). *Introduction to Modern Nonparametric Statistics*, Pacific Grove,CA:Brooks/Cole,Thomson Learning,Inc.

Hudak P. (2003). Arsenic, nitrate, chloride and bromide contamination in the gulf coast aquifer, south-central Texas, USA. *International Journal of Environment Studies*,60,123-133. doi:10.1080/00207230304736

Hutcheson G.D.,and Sofroniou N.(1999). *The Multivariate Social Scientists: an introduction to generalized linear models*. Sage Publications.

James K.A., Meliker J.R., Buttenfield B.E., Byers T., Zerbe G.O., Hokanson J.E., Marshall J.(2014). Predicting arsenic concentrations in groundwater of San Luis Valley, Colorado: Implications for individual-level lifetime exposure assessment. Environment Geochemical Health,36,773-782. doi:10.1007/s10653-014-9595-6

Johnson C.D., Nandi A., Joyner A.T., & Luffmann I. (2018). Iron and Manganese in Groundwater: Using Kriging and GIS to locate High Concentrations in Buncombe County, North Carolina.*Groundwater*,56,87-95.: doi: 10.1111/gwat.12560

Kaiser H.F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*,20,141-151. doi:10.1214/aos/1009210544

Kaiser H.F., Rice J. (1974). Little Jiffy, Mark IV. *Educational and Psychological Measurement*,34(1), 111-117.doi:10.1177/001316447403400115

Langley R.L., Kao Y., Mort S.A., Bateman A., Simpson B.D., & Reich B.J. (2017). adverse neurodevelopmental effects and hearing loss in children associated with manganese in well water, North Carolina, USA. *Journal of Environment Occupational Science*,4(2),62-69. doi :10.5455/jeos.20150403060427.

Larkin T.J., & Bomar G.W. (1983). Climatic Atlas of Texas. Retrieved from http://hdl.handle.net/1969.3/20098

Lewandowski A.M., Montgomery B.R., Rosen C.J., & Moncrief J.F. (2008). Groundwater nitrate contamination costs: A survey of private well owners. Retrieved fromhttps://www.house.leg.state.mn.us/comm/docs/JSWC2008CostofNitrateContamination.pdf

Mace R.E., Davidson S.C., Angle E.S., and Mullican III W.F. (2006). Aquifers of the Gulf Coast. Retrieved from http://www.twdb.state.tx.us/publications/reports/GroundWaterReports/GWReports/R365/R365_Composite.pdf.

McLeod L., Bhardwaj L., Epp T., Waldner C. L. (2017). Use of Principal Component Analysis and kriging to predict Groundwater-sourced Rural Drinking water Quality in Saskatchewan. *International Journal of Environmental Research and Public Health*,14, 1065.doi:10.3390

Minnesota Department of Health. (2019). Sulfate in Well Water. Retrived from https://www.health.state.mn.us/communities/environment/water/docs/wells/waterquality/sulfate.pdf

National Research Council (US) Safe Drinking Water Committee. (1982). III, Chemical Quality of Water in Distribution System. *Drinking water and Health*, 4. Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK216607/

Navas-Acien A., Sharrett, A.R., Silbergeld E.K., Schwartz, B.S., Nachman K.E., Burke, T.A., Guallar E. (2005). Arsenic exposure and cardiovascular disease: A systematic review of the epidemiologic evidence. *American Journal of Epidemiology*. 162, 1037-1049.doi:10.1093/aje/kwi330

Olea R.A. (2006). A six-step practical approach to semivariogram modeling. *Stochastic Environmental Research and Risk Assessment*,20,307-318. doi :10.1007/s00477-005-0026-1.

Oliver M.A. (1990). "Kriging: A Method of Interpolation for Geographical Information Systems." *International Journal of Geographic Information Systems*,4,313-332.

Onufrak S.J., Park S., Sharkey J.R., Sherry B. (2014) The relationship of perceptions of tap water safety with intake of sugar-sweetened beverages and plain water among US adults. *Public Health Nutrition*,17(1),179-185. doi: 10.1017/S1368980012004600.

Pizarro F., Olivares M. Uauy R., Contreras P., Rebelo A., Gidi V. (1999). Acute gastrointestinal effects of graded levels of copper in drinking water. *Environmental Health Perspective*,107(2),117-21 doi:10.1289/ehp.99107117.

Powers M., et al. (2018). Arsenic in Groundwater in private wells in rural North Dakota and South Dakota: Water quality assessment for an intervention trail. *Environmental Research*,168, 41-47.doi: 10.1016/j.envres.2018.09.016

Raina P.S., Pollari F.L., Teare G.F. (1999) The relationship between *E. coli* indicator bacteria in well-water and Gastrointestinal illnesses in rural families, *Canadian Journal of Public Health*,90,172. doi: 10.1007/BF03404501

Richards C.L., Broadway S.C., Eggers M.J., Doyle J., Pyle B.H., Camper A.K., &Ford T.E. (2015). Detection of Pathogenic and Non-Pathogenic Bacteria in Drinking Water and associated Biofilms on the crow Reservation, Montana, USA. *Environment Microbiology*,76,52-63. doi :10.1007/s00248-015-05956

Reedy R.C., & Scanlon B.R. (2018). *Assessment of Arsenic in Groundwater and Water Supply Systems in Texas*. Retrieved from http://www.beg.utexas.edu/files/content/beg/research/water/TCEQ%20Arsenic%20Report%20EPA%20Aug%2010%202018.pdf

Robinson T.P.,& Metternicht G. (2006). Testing the performance of spatial interpolation techniques for mapping soil properties. *Computers and Electronics in Agriculture*,50,97-108. doi :10.1016/j.compag.2005.07.003

Royle A., et al. (1981). “Practical Universal Kriging and Automatic Contouring.” *Geoprocessing*,1,377-394.

Stafford, D.B. (1991). Civil engineering applications of remote sensing and geographic information systems. *American Society of Civil Engineers*,10,360.

Schullehner J., Hansen B., Thygesen M.,Pedersen C.B.,Sigsgaard T.(2018) *International Journal of Cancer*,143(1),73-79. doi: 10.1002/ijc.31306.

Tabuchi H., & Kaplan S. (2017). A Sea of Health and Environmental Hazards in Houston Floodwaters. New York Times. Online Edition. Retrieved from https://www.nytimes.com/2017/08/31/us/houston-contaminated-floodwaters.html

Texas Administrative Code. (2014). Water Well Drillers and Water Well Pump Installers. Retrieved from https://texreg.sos.state.tx.us/public/readtac$ext.TacPage?sl=R&app=9&p_dir=&p_rloc=&p_tloc=&p_ploc=&pg=1&p_tac=&ti=16&pt=4&ch=76&rl=100

Tziachris P., et al. (2017). Spatial Modelling and Prediction Assessment of Soil Iron using Kriging Interpolation with pH as Auxiliary Information, International Journal of Geo-Information, 6,283.doi : 10.3390/ijgi6090283

Townsend, M., &Young, D.P. (1995). Factors affecting nitrate concentrations in ground water in Stafford county, Kansas. Retrieved from http://www.kgs.ku.edu/Current/townsend.pdf

U.S. Geological Survey (2016b). *Water Quality of potential concern in U.S. Private Wells*. Retrieved from http://www.usgs.gov/newsroom/article.asp?ID=2173

U.S. Geological Survey (2016). *Contamination in U.S. Private Wells*. Retrieved from http://www.usgs.gov/newsroom/article.asp?ID=2173

Uhlman K., Boellstorff D.E., McFarland M.L.,Clayton B., Smith J.W.(2012). Texas Well Owner Network Well Owner's Guide to Water Supply. Retrieved from http://agrilife.org/water/files/2013/02/texas-well-owner-network-guide-to-water-supply.pdf

Sánchez-Martos F., Jiménez-Espinosa R., Pulido-Bosch A. (2001). Mapping groundwater quality variables using PCA and geostatistics: A case study of Bajo Andarax, southeastern Spain.

World Health Organization (1996). Zinc in drinking water. Retrieved from https://www.who.int/water_sanitation_health/dwq/chemicals/zinc.pdf

World Health Organization (1997). Guidelines for drinking-water quality. Retrieved from https://www.who.int/water_sanitation_health/dwq/2edvol3a.pdf

World Health Organization (2003). Chloride in drinking water. Retrieved from https://www.who.int/water_sanitation_health/dwq/chloride.pdf

World Health Organization (2004). Sulfate in drinking water. Retrieved from https://www.who.int/water_sanitation_health/dwq/chemicals/sulfate.pdf

Yang Q., Jung H.B., Culbertson C.W., Marvinney R.G., Loiselle M.C., Locke D.B., Cheek H., Thibodeau H., Zheng Y. (2009). Spatial pattern of groundwater arsenic occurrence and association with bedrock geology in Greater Augusta, Maine, USA. *Environment Science & Technology*,43(8),2714-2719. Doi:10.1021/es803141m

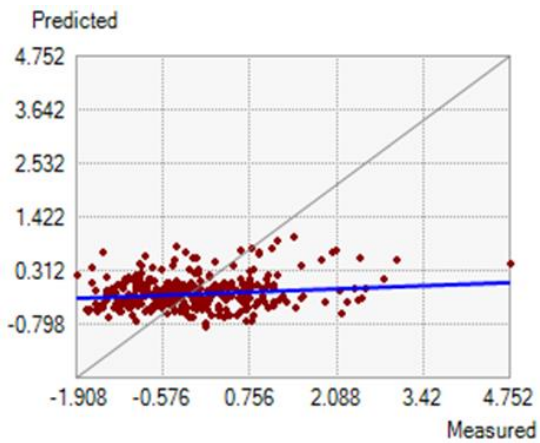
Yu, W.H., Harvey, C.M., Harvey, C.F. (2003). Arsenic in groundwater in Bangladesh: A geostatistical and epidemiological framework for evaluating health effects and potential remedies. *Water Resources*,39,1146, doi:10.1029/2002WR001327

Zandbergen P.A. (2008). A comparison of address point, parcel and street geocoding techniques. *Computers, Environment and Urban Systems*, 32,214-232. doi : 10.1016/j.compenvurbsys.2007.11.006

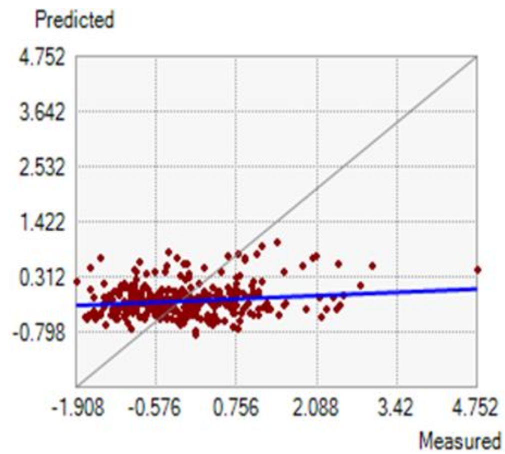
Zhitkovich, A .(2011). Chromium in drinking water: sources, metabolism, and cancer risks. *Chemical Research in Toxicology*,24(10),1617-1629.doi:10.1021/tx200251t

APPENDIX A

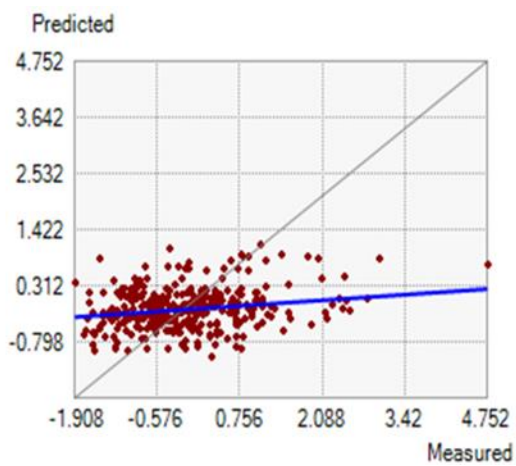
PREDICTED VS MEASURED PLOTS



(a)

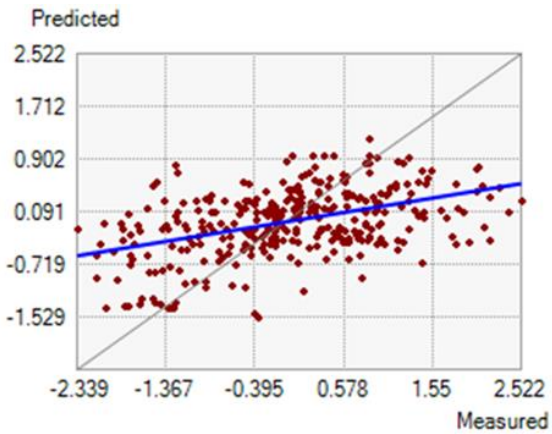


(b)

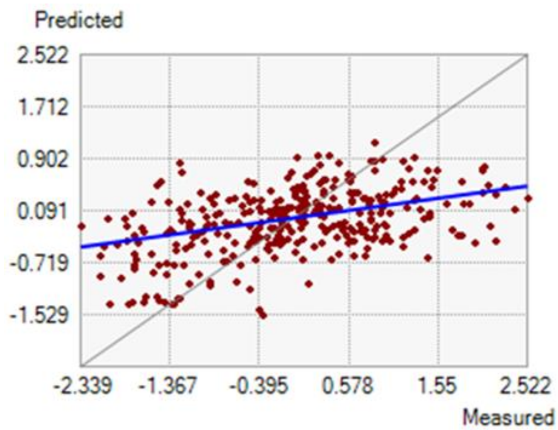


(c)

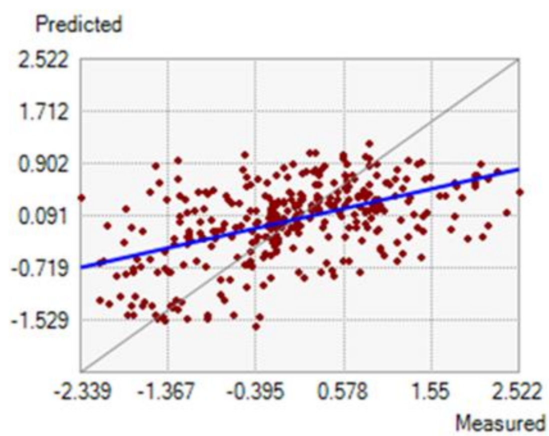
Figure A- 1. Predicted vs Measured Values for PC1 Scores for Elements with Primary Standards Obtained from Cross-Validation Results for (a) Ordinary, (b) Universal and (c) Empirical Bayesian Kriging



(a)



(b)



(c)

Figure A- 2. Predicted vs Measured Values for PC2 Scores for Elements with Primary Standards Obtained from Cross-Validation Results for (a) Ordinary, (b) Universal and (c) Empirical Bayesian Kriging

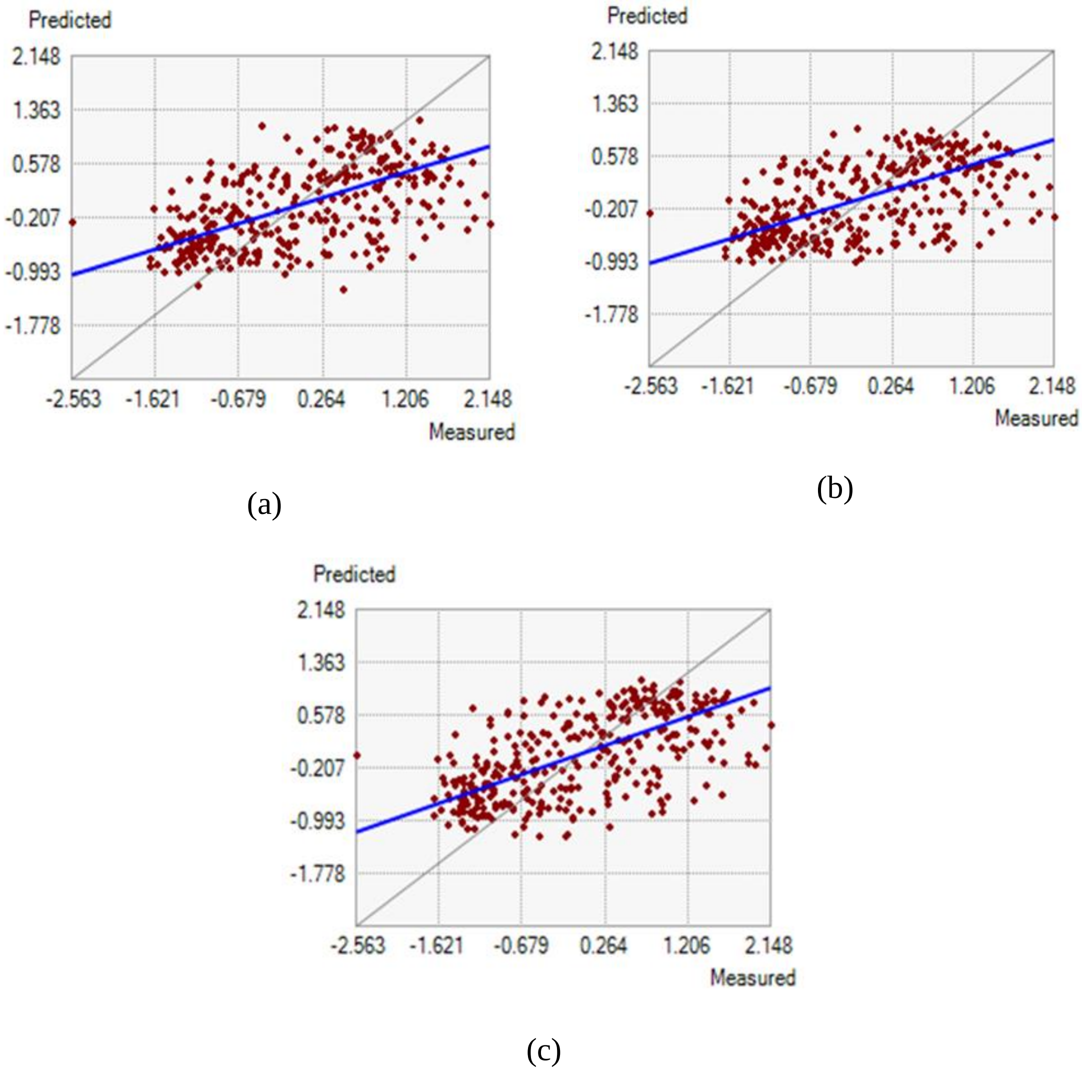
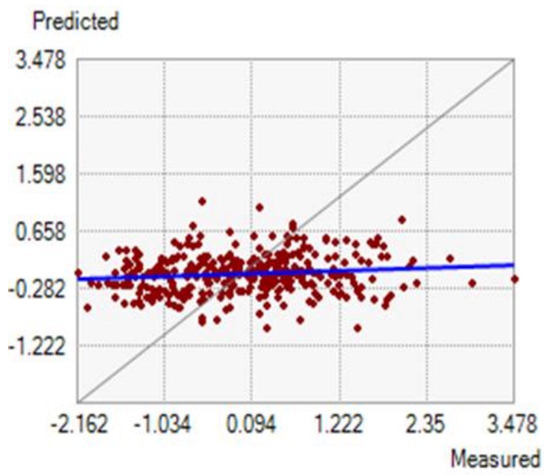
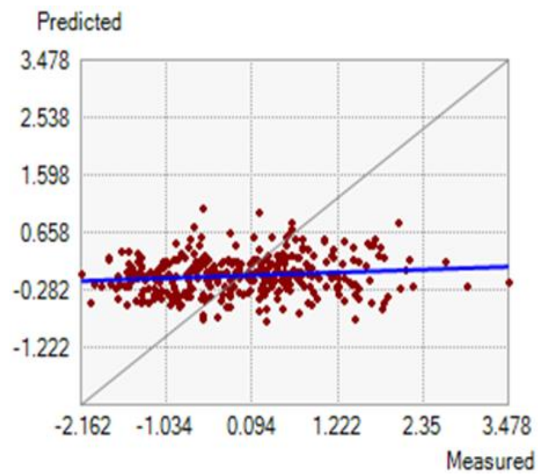


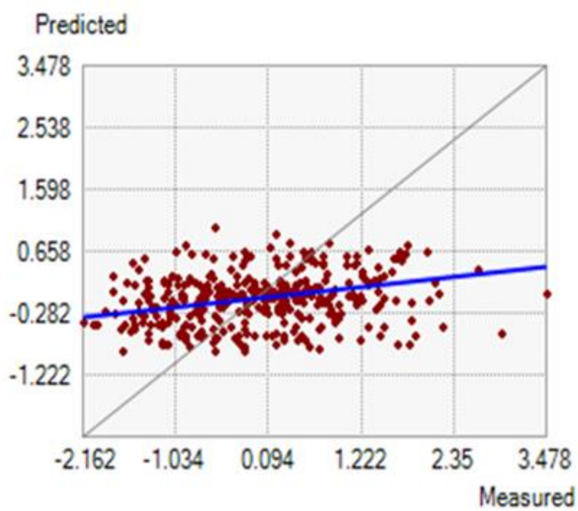
Figure A- 3. Predicted vs Measured Values for PC1 Scores for Elements with Secondary Standards Obtained from Cross-Validation Results for (a) Ordinary, (b) Universal and (c) Empirical Bayesian Kriging



(a)

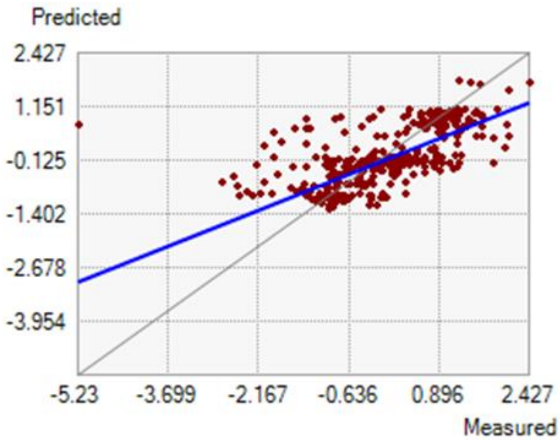


(b)

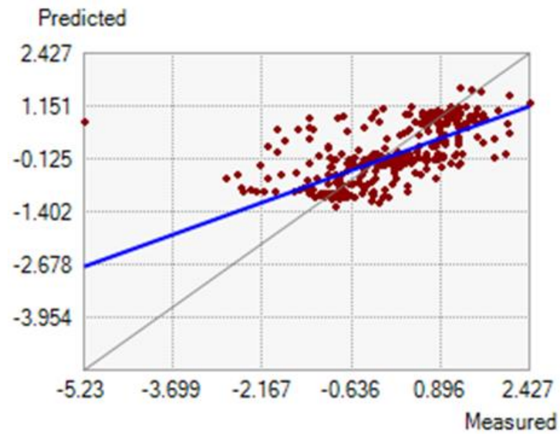


(c)

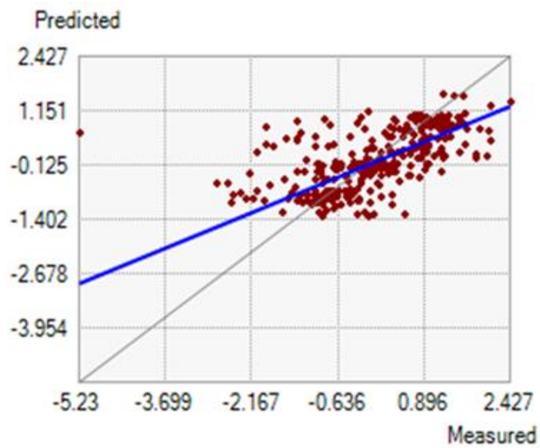
Figure A- 4. Predicted vs Measured Values for PC2 Scores for Elements with Secondary Standards Obtained from Cross-Validation Results for (a) Ordinary, (b) Universal and (c) Empirical Bayesian Kriging



(a)



(b)



(c)

Figure A- 5. Predicted vs Measured Values for PC3 Scores for Elements with Secondary Standards Obtained from Cross-Validation Results for (a)Ordinary, (b) Universal and (c) Empirical Bayesian Kriging