Spring 2018

Chemical and Statistical Analysis of Karst Groundwater Basin Signatures - Springfield, Mo

Benjamin E. Lockwood
Missouri State University, Lockwood023@live.missouristate.edu

Recommended Citation

Lockwood, Benjamin E., "Chemical and Statistical Analysis of Karst Groundwater Basin Signatures - Springfield, Mo" (2018). *MSU Graduate Theses*. 3254.
https://bearworks.missouristate.edu/theses/3254

CHEMICAL AND STATISTICAL ANALYSIS OF KARST GROUNDWATER BASIN SIGNATURES – SPRINGFIELD, MO

A Masters Thesis
Presented to
The Graduate College of
Missouri State University

In Partial Fulfillment
Of the Requirements for the Degree
Master of Science, Geospatial Sciences in Geography and Geology

By
Benjamin Erwin Lockwood
May 2018

CHEMICAL AND STATISTICAL ANALYSIS OF KARST GROUNDWATER BASIN SIGNATURES – SPRINGFIELD, MO

Geography, Geology, and Planning

Missouri State University, May 2018

Master of Science

Benjamin Erwin Lockwood

ABSTRACT

Springfield, MO is located on the Springfield Plateau physiographic province. The Springfield plateau contains a number of Mississippian aged units and is mainly capped by the Burlington-Keokuk Formation. The Burlington-Keokuk is a highly fossiliferous limestone with nodular and interbedded chert. Beneath the Burlington-Keokuk lies the Elsey, Reeds Spring, and Pierson Formations respectively which comprise the Springfield Plateau aquifer hydrostratigraphic unit. Within the Springfield Plateau aquifer, a well-developed karst system includes springs, sinkholes, and caves. The Springfield Plateau aquifer is the predominant source for springs and seeps in the Springfield area. The purpose of this study was to understand the differences in water chemistry of individual karst groundwater basins. Different land use surrounding these groundwater basins as well as minute differences in the Burlington-Keokuk may lead to different water chemistry for each basin. Sampling was conducted at 12 sites in Springfield, MO from within what are believed to be five separate groundwater basins. Samples were collected over six months, and 11 variables were measured. Field tests included pH, temperature, conductivity, bicarbonate (as $CaCO3$), and flow. Lab analyses included major cations (calcium, magnesium, and sodium) and anions (chloride, sulfate, and nitrate). Statistical analyses were run using SAS 9.4 and included discriminant function analysis, factor analysis, and miscellaneous variable analyses. Results from two models suggest that there is enough difference in water chemistry between groundwater basins to develop statistical models that could accurately classify samples to the correct basin based on water chemistry.

KEY WORDS: groundwater basins, geochemistry, karst, discriminant function analysis, sas 9.4, springfield missouri.

This abstract is approved as to form and content

____
Douglas R. Gouzie, Ph.D.
Chairperson, Advisory Committee
Missouri State University

CHEMICAL AND STATISTICAL ANALYSIS OF KARST GROUNDWATER BASIN SIGNATURES – SPRINGFIELD, MO

By

Benjamin Erwin Lockwood

A Masters Thesis
Submitted to the Graduate College
Of Missouri State University
In Partial Fulfillment of the Requirements
For the Degree of Master of Science, Geospatial Science and Environmental Geology

May 2018

Approved:

Douglas R. Gouzie, Ph.D.

Richard N. Biagioni, Ph.D.

Xin Miao, Ph.D.

Julie Masterson, Ph.D.: Dean, Graduate College

ACKNOWLEDGEMENTS

I would like to thank my thesis committee members for supporting my research: My advisor, Dr. Douglas Gouzie, for his guidance and support through my undergraduate and graduate years. I am very appreciative of the time and effort he has put into helping me succeed; Dr. Richard Biagioni, for his knowledge of chemistry, use of lab space and equipment, and his ability and willingness to troubleshoot the atomic absorption instrument; and Dr. Xin Miao, for his knowledge of statistics, without which I would not have the well-rounded statistical analysis for this thesis. I would also like to thank Elizabeth Robertson, M.Sc. of Springfield's City Utilities Blackman Lab for her work on analyzing many of my samples.

Others that I would like to thank include: Bill Neubert for donating a GPS unit many years ago, it is an invaluable tool and has been used many times in my college career. My peers in the department for the countless hours of laughs, great memories, and motivation. Without you, my thesis would have been completed months earlier. Thank you to those teachers, professors, and instructors throughout my academic career who have pushed and motivated me to be the best that I can. And finally, thank you to my parents and siblings for their love and support, I wouldn't be where I am without you all.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1 – INTRODUCTION

Karst

The term karst applies to unique features found within soluble carbonate bedrock. Carbonate bedrock is mainly comprised of limestone and dolomite (or dolostone), which are easily dissolved when exposed to mildly acidic waters. Dissolution of the carbonate bedrock creates features such as sinkholes, karst windows, karren, pinnacles, and conduits (caves). Combined, these features create an irregular surface that is easily identified, even with topsoil and vegetation. Regions that are rainy and humid usually have well defined karst terrains as compared to dry and arid areas which usually do not have karst features.

The large underground channels where water travels through carbonate bedrock are called conduits. These conduits form when fractures in the bedrock are expanded by acidic water dissolving the soluble rock. This dissolution can occur along faults, fractures, and bedding planes but is limited by the chemistry of the water moving through the system. Waters that are saturated with respect to calcite will not dissolve the surrounding bedrock. A dynamic system that removes saturated waters and brings in undersaturated waters (with respect to calcite) allows for constant dissolution of the surrounding bedrock (White, 2002).

Conduits and caves are vital to groundwater transport within carbonate aquifers. The larger the conduit is, the larger the volume of water that can be transported. An increased number of faults, fractures, and conduits increase the permeability of the bedrock, allowing more water to either recharge the local aquifer or travel a short distance before reemerging at the land surface (Palmer, 2007).

Water enters and exits conduits through two common karst features, sinkholes and springs. Sinkholes are circular depressions in the bedrock surface that act as a collection point for surface water runoff. Springs are the point where water exits a conduit and becomes surface water. Springs can range in size from a small trickle to a large stream emerging from a cave opening.

Often, multiple sinkholes, conduits, and springs are connected and create a network in which surface and groundwater intermix. Usually, one or more sinkholes feed a conduit, allowing water to travel underground for a certain amount of time before reemerging in a spring or series of springs. Multiple sinkholes can capture surface water from a large area and funnel it underground where it can travel miles. Water traveling through conduits either enters a more extensive aquifer or reemerges in one or more springs. Networks of sinkholes and conduits feeding a spring or small group of springs have been referred to as groundwater basins (Thrailkill et al., 1982).

Similar to surface watersheds, many groundwater basins can be found together in an area. These basins can create a large network in which groundwater can easily flow underground through carbonate bedrock. Large networks of groundwater basins have been identified using various tracing methods, including fluorescein or rhodamine WT dyes (Palmer, 2007). Much like surface basins, it is also possible that groundwater basins can also be physically isolated from other basins. In basins that are physically isolated, water does not intermix between two or more basins. This physical isolation leads to the possibility that the chemistry of water flowing through a groundwater basin is unique in respect to another basin or basins.

Water Chemistry

Water chemistry is always changing and in a system where surface and groundwater mix easily, such as carbonate aquifers, it is especially dynamic. The excellent solubility properties of water allow it to erode, dissolve, and absorb chemicals, elements, and compounds which it contacts. This interaction creates a diverse range of water properties and characteristics that can be studied. Some common physical properties include color, total dissolved solids (TDS), turbidity, specific conductivity, and temperature (Manaham, 2010). There are hundreds of elements and compounds that can be found in water. In carbonate environments, common cations and anions include: calcium (Ca), magnesium (Mg), sodium (Na), potassium (K), bicarbonate (HCO3), nitrate (NO3), sulfate (SO4), chloride (Cl), fluoride (F), and some silicon dioxide (SiO2) (Drever, 1997; Bullard, Thomson, and Vandike, 2001). These ions and physical properties come from the interaction of water with its surroundings.

As water travels over the surface, through a sinkhole, into a conduit, and out through a spring, it develops physical and chemical properties. For surface waters, most of the chemistry originates from the soil (Drever, 1997). Here, chemicals from household and industrial use, biologic activity, and inorganic activities are concentrated. As water runs over the surface and percolates downward, it dissolves and transports these chemicals elsewhere. Groundwater chemistry originates from the contact between water and the bedrock it is traveling through. Physical and chemical weathering have several factors that impact the rate in which water erodes the surrounding bedrock. Water velocity, saturation, temperature, carbon dioxide (CO2) levels, and more affect the physical and chemical properties. In carbonate aquifers, calcium, sodium, and bicarbonate are some of the most common ions dissolved from limestones and dolomites based on water temperature, CO2 levels and calcium carbonate saturation. When the water chemistries of surface and ground waters mix, a unique signature is created which could be used to identify a groundwater basin. This is the primary focus of this thesis.

Statistical Analysis

With a large number of physical and chemical properties (variables) available to analyze, a univariate or bivariate statistical analysis is not a viable option for most water chemistry studies. Multivariate statistics provide the ability to use as many variables as needed/wanted for an analysis. For most multivariate statistics such as multivariate analysis of variance (MANOVA), the chemical and physical properties are used as the dependent variables (DV) or predictor variables while the sample sites are used as the independent variables (IV) or groups (Tabachnick and Fidell, 1996). However, for studies where the goal is to predict group (groundwater basin) membership using predictor variables (physical and chemical properties), a unique multivariate statistic may be used.

The discriminant function analysis (DISCRIM) is a unique multivariate statistic that is used to determine group membership based on a set of predictor values. Unlike MANOVA, DISCRIM uses the predictor values as independent variables and the groups as dependent variables (Tabachnick and Fidell, 1996). This function would use the physical and chemical properties to define each groundwater basin.

Goals and Objectives

The primary goal of this study was to see if there was enough difference in water chemistry to develop a geochemical signature for groundwater basins using a statistical model. The secondary goal was to see if it was possible to match unknown samples to their respective basins using the models created from the primary goal. These goals were achieved through four steps:

1. Select groundwater basins using existing dye trace data.
2. Analyze water chemistry of samples from each groundwater basin.
3. Build and test statistical models using chemical data.
4. Test robustness of the statistical models using "blind" samples.

CHAPTER 2 – STUDY AREA

The city of Springfield is in Greene County of southwest Missouri (Figure 1). It is the third largest city in Missouri with an estimated population of ~160,000 (US Census Bureau, 2015). Springfield is a growing city with major urban and industrial areas surrounded by rural farm and pasture lands. Springfield and surrounding towns are somewhat unique in that they are built atop a well-developed karst system. Because there are few cities built on top of karst terrains and karst features (caves and sinkholes), unique studies such as surface-groundwater interaction in an urban environment, contaminant transport in a karst environment, and engineering on and around karst features, can be conducted.

Geology

Springfield sits atop Mississippian-aged Burlington-Keokuk formation (Figure 2) and younger Quaternary alluvium that comprises the Springfield Plateau (Thomson, 1986). While defined as two separate units in other parts of the United States, the Burlington and Keokuk formations are difficult to separate in Greene county and are usually grouped together as one unit (Thomson, 1986). The Burlington-Keokuk is interbedded with chert and contains an abundance of crinoid fossils. In the Springfield Plateau the Burlington-Keokuk is ~200 feet thick and is the dominant unit to outcrop in Greene county. The geologic units within the plateau are almost level, dipping no more than 1-2 degrees (Thomson, 1986). Elevation ranges from 1110 to 1300 feet above sea level (Thomson, 1986). Multiple joint sets in the limestone run parallel in a northwest to southeast trend (Thomson, 1986). Due to the climate and the limestone bedrock, karst features are common in southwest Missouri and Springfield. Undersaturated waters in contact with the limestone results in high dissolution rates of the limestone, and features including sinkholes (Figure 3), seeps, caves, springs (Figure 4), and losing streams can develop (Palmer, 2007). Many of these features are scattered throughout the city. Subsurface dissolution of fractures and bedding planes in the Burlington-Keokuk creates channels or caves that act as the primary flow path for groundwater. These caves can also act as recharge points where surface streams, road runoff, and such flow into the subsurface rapidly. Sinkholes, which are very common in Greene county also act as groundwater recharge points. It is estimated that the sinkhole density in Greene County is 2.2 sinks per square mile (Berglund, 2012). Surface streams on the Springfield Plateau are usually sinking streams, where the combination of a deeper water table and sinkholes moves water underground instead of sustaining overland flow (Palmer, 2007). Groundwater can also flow where the dissolution of bedding planes and fractures creates only small openings and channels. Where sinkholes are generally found in upland areas, springs are usually found in valleys or lowland areas (Thomson, 1986).

Hydrogeology

Springfield and the greater Springfield Plateau is underlain by three aquifers: the Springfield Plateau Aquifer (~450 feet thick), the Ozark Aquifer (~1200 feet thick), and the St. Francois Aquifer (~370 feet thick) (Missouri Department of Natural Resources, 2016). The Springfield Plateau and Ozark aquifers are separated by the Ozark Confining unit, which is comprised of the Northview, Compton, and Bachelor formations and has an

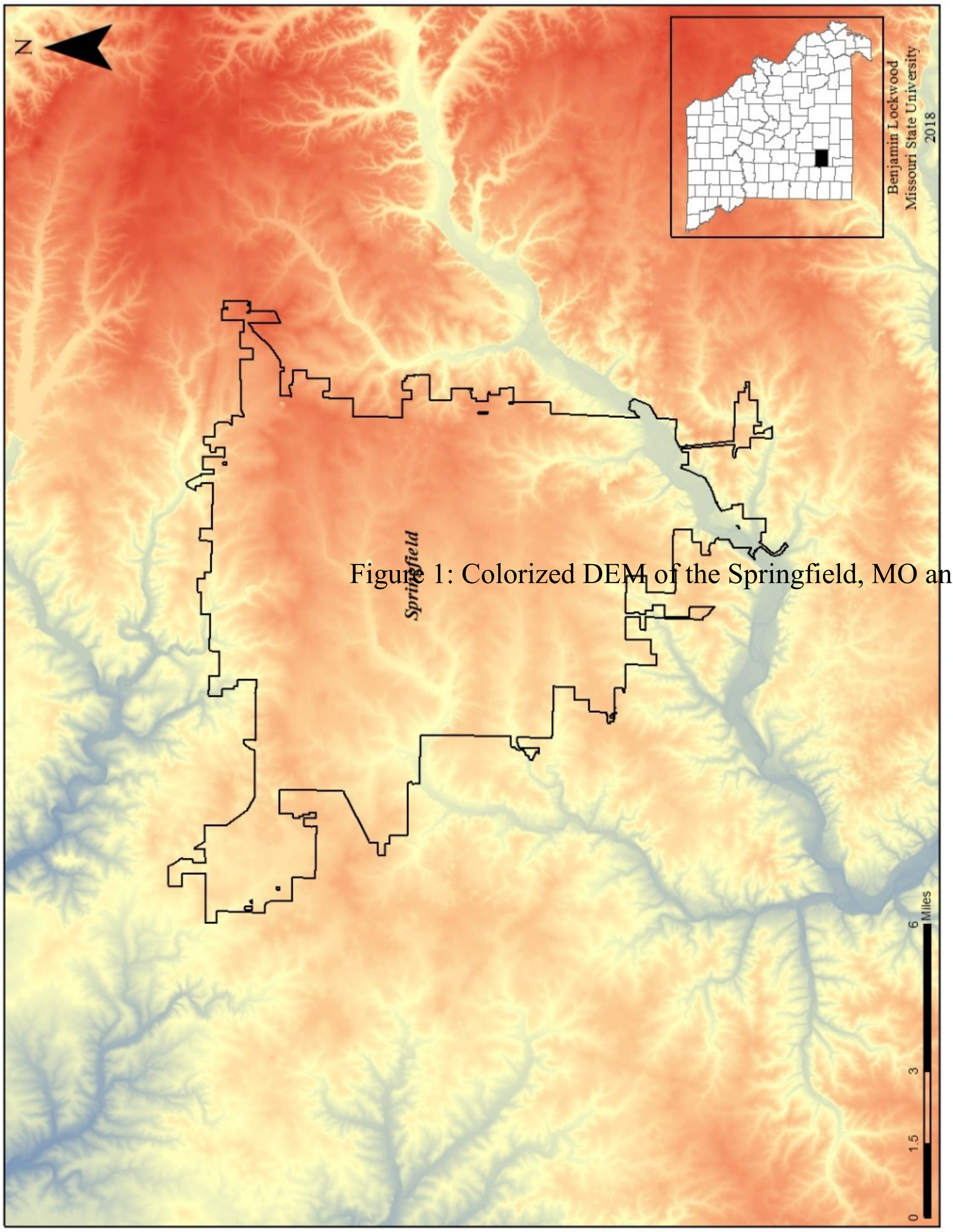


Figure 1: Colorized DEM of the Springfield, MO and the outlaying area.

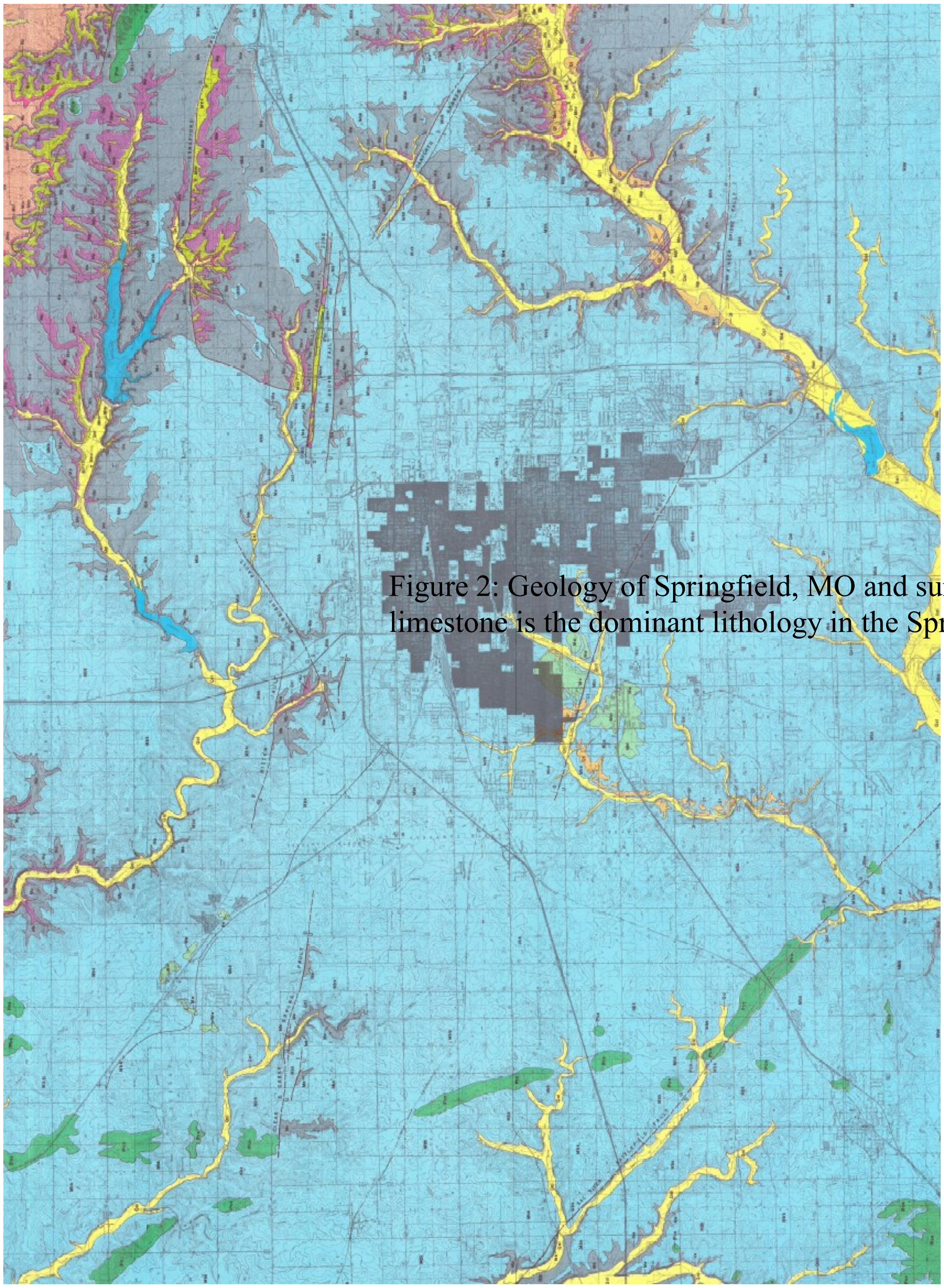


Figure 2: Geology of Springfield, MO and surrounding area. Mississippi
limestone is the dominant lithology in the Springfield Plateau (from Tho

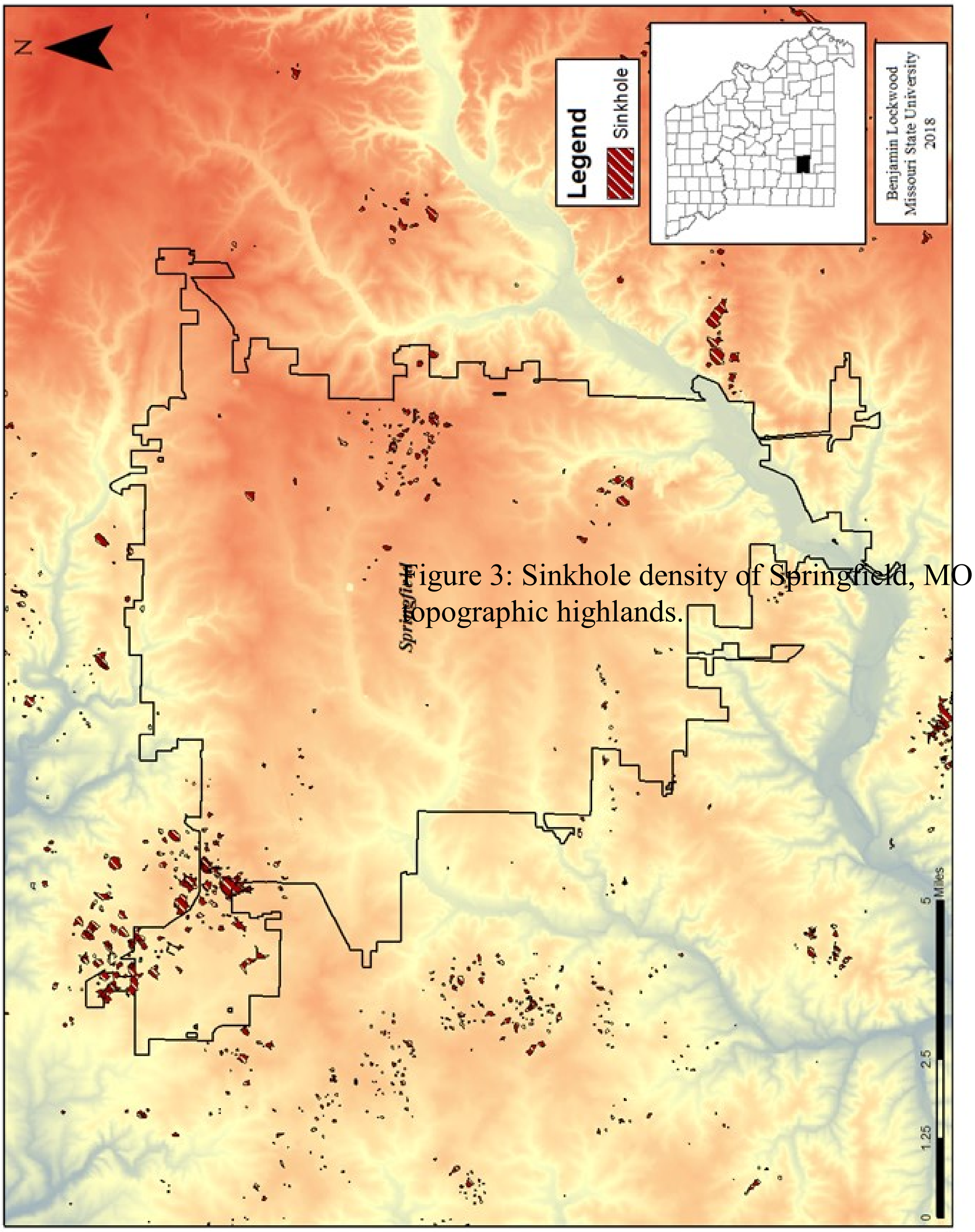


Figure 3: Sinkhole density of Springfield, MO. Majority of sinkholes i
topographic highlands.

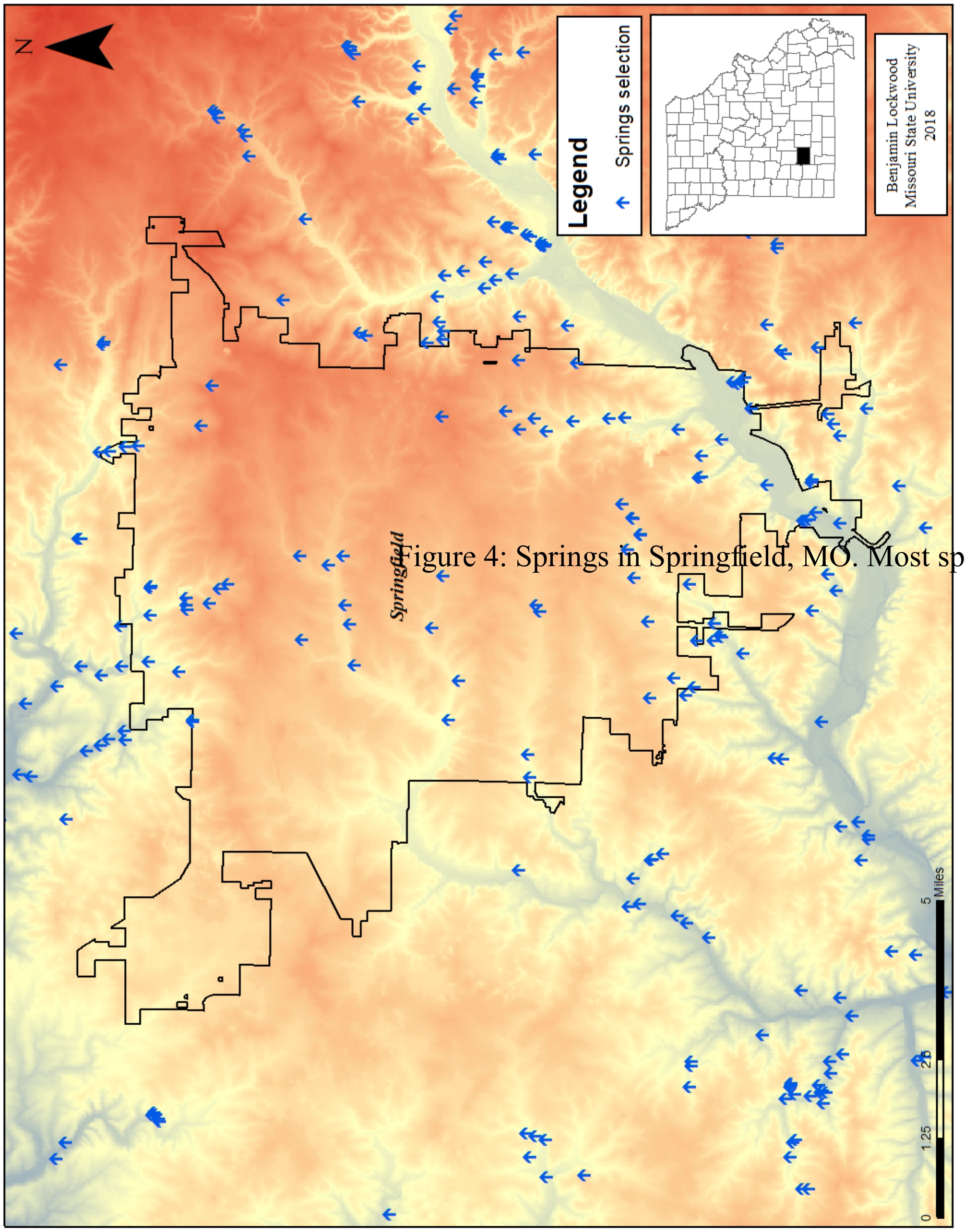


Figure 4: Springs in Springfield, MO. Most springs fall along streams

average thickness of ~80 feet (Missouri Department of Natural Resources, 2016). Little mixing occurs between the two aquifers. Most, if not all, groundwater traced in the Springfield Plateau flows within the Burlington-Keokuk limestone only (Imes, 1989). The Davis formation (shale, siltstone, sandstone, dolomite, and limestone) separates the Ozark and St. Francois aquifers and is roughly 145 feet thick (Missouri Department of Natural Resources, 2016). The St. Francois aquifer is not a significant aquifer in Greene County usage.

Climate

The climate of southwest Missouri is humid subtropical, according to the Köppen-Geiger climate classification with hot, humid summers and cool, mild winters (Peel et al., 2007). Average precipitation is 45.5 inches with an average snowfall of 17 inches in the winter. Temperatures range from an average high of 67°F to an average low of 45°F (U.S. Climate Data, 2016). During the sampling period, temperatures ranged from a high of 98°F in July 2017 to a low of 30°F in November 2017 (Weather Underground, 2018). A total of 39.8 inches of rain fell over the 5 ½ month sampling period. Daily highs, lows, and total daily rainfall are graphed in figure 5.

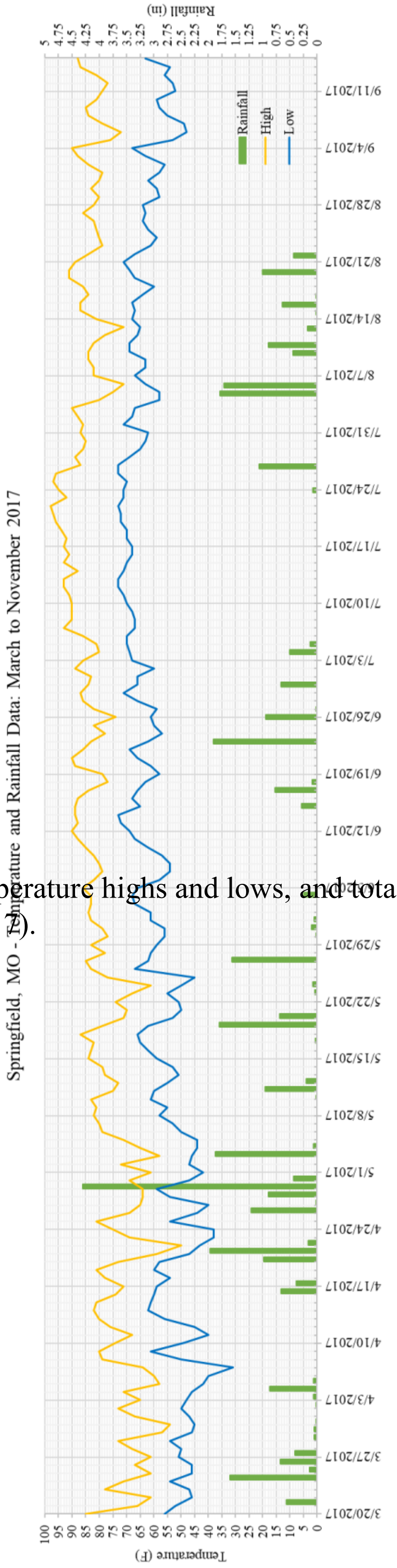


Figure 5: Temperature highs and lows, and total daily rainfall during the sample season (M
September 2017).

CHAPTER 3 – METHODS

Site Selection

Before sampling began sites were picked based on several factors. Sites had to fit the model of a groundwater basin as defined by a series of sinkholes, surface streams, subsurface conduits, and springs. It had to be known that these features were connected by dye traces. Data from the Missouri Department of Natural Resources GeoSTRAT (Missouri Department of Natural Resources) database was used to add dye injection and recovery points and dye path points to Google Earth and ArcMap (Figure 6). It was important to find sampling sites that had a constant supply of water year-round to ensure enough water could be collected and filtered for analysis. Sites also had to be accessible. Majority of the sites were on public land, but one site was on private property, and the land owner granted permission to access the spring.

Sites with limited access during high flow were excluded from the search as well as sites that were difficult to reach during very low (drought conditions). Only the hospital pond provided some difficulty during very dry conditions due to the instability of the bank as water subsided and no samples were collected from the pond that day.

Sites were selected based on the surrounding land use as well. It was intended to select sites where water flowed through areas with different land use (urban, industrial, rural, residential, etc.) with the idea that different land uses may add to the chemical variability of the groundwater. This search filter was added after sites were selected based on year-round flow and ease of access. Discussions with Dr. Gouzie and field scouting of sample sites resulted in the selection of 12 sample sites (five basins) (Figure 7).

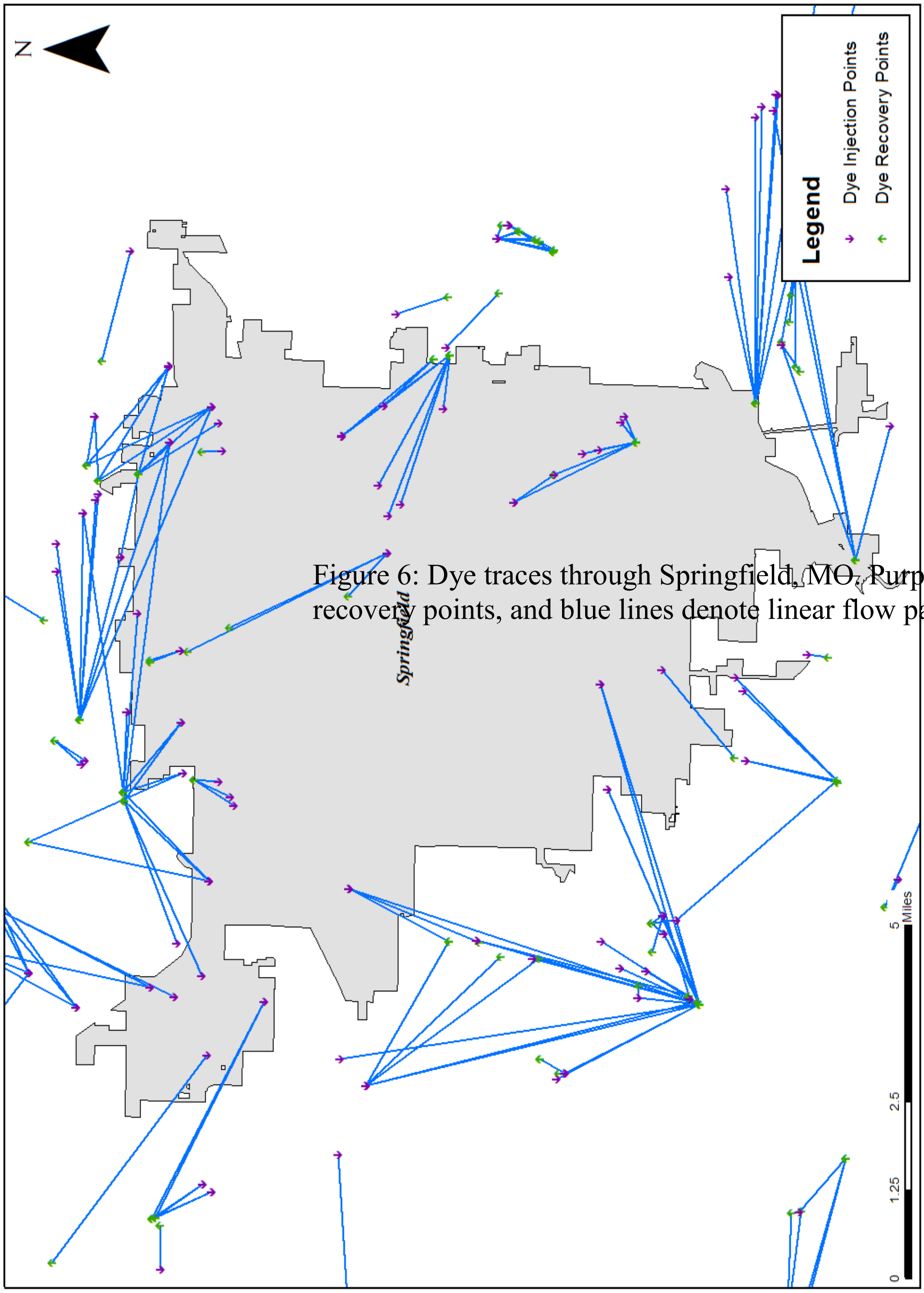


Figure 6: Dye traces through Springfield, MO. Purple arrows are dye injectio
recovery points, and blue lines denote linear flow paths.

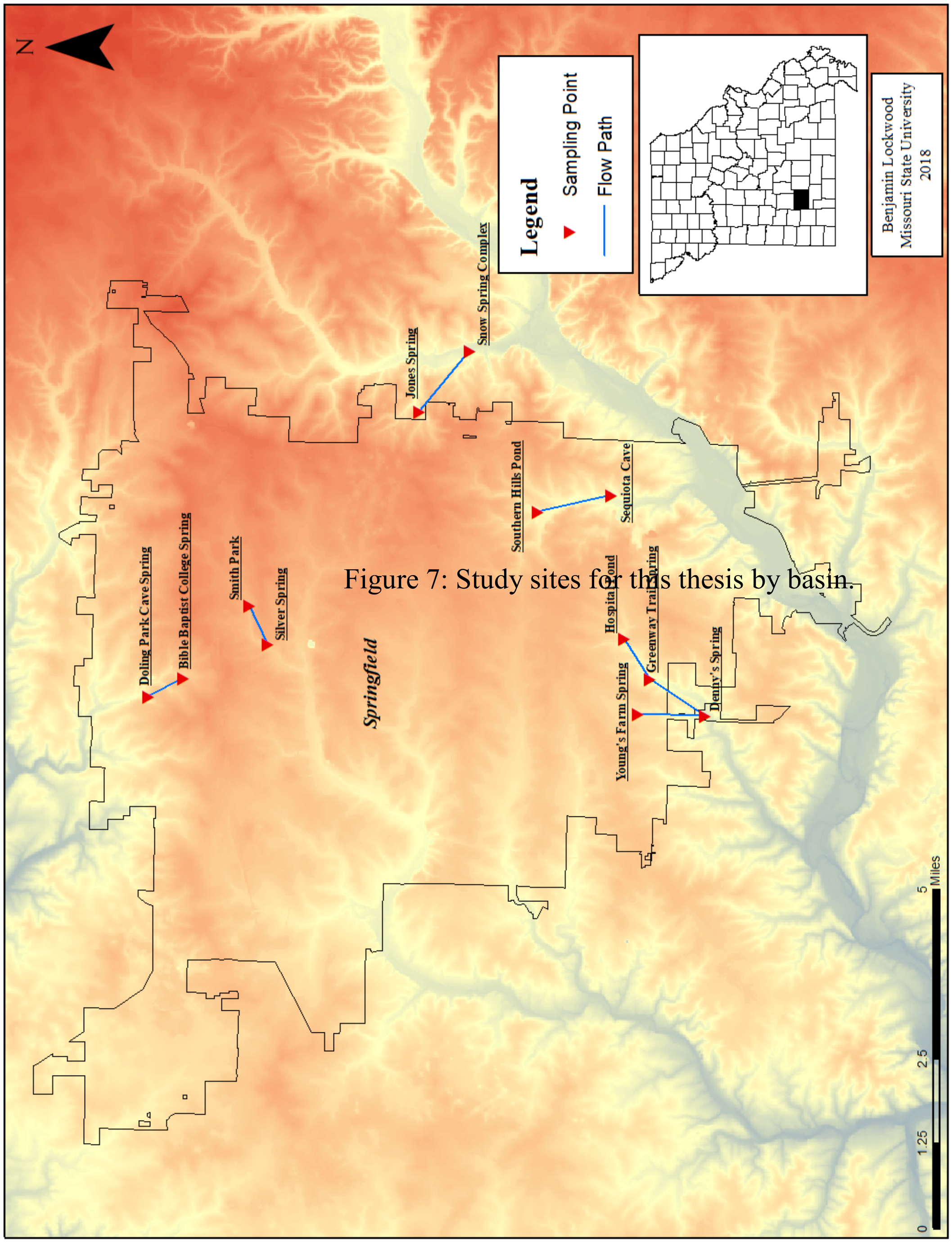


Figure 7: Study sites for this thesis by basin.

Sample Sites

Twelve sites were selected for sampling. These twelve sample sites were divided into five basins based on their interconnectivity (Figure 7). The Denny's spring basin is comprised of the Young's Farm spring, Greenway trail spring, hospital pond sink, and the Denny's spring. Jones basin is comprised of the Jones spring and the Snow spring complex. Sequiota basin is made up of the Southern Hills pond spring and Sequiota cave spring. Doling basin contains the Doling park cave spring and Baptist Bible College spring. Silver spring basin is comprised of Smith park and Silver spring park. GPS coordinates are listed in Table 1.

Table 1: GPS coordinates for every sample site (degrees, minutes, seconds).

	Latitude			**Longitude**		
	D	**M**	**S**	**D**	**M**	**S**
Doling Cave	37	14	47.34	93	17	27.85
BB College	37	14	20.75	93	17	11.04
Smith Park	37	13	29.84	93	16	0.43
Silver Spring	37	13	15.13	93	16	38.34
Jones Spring	37	11	19.16	93	12	52.42
Snow Complex	37	10	40.06	93	11	54.30
Southern Hills Pond	37	9	47.17	93	14	30.08
Sequiota Park	37	8	52.02	93	14	12.94
Hospital Pond	37	8	40.56	93	16	31.24
Greenway Trail	37	8	20.5	93	17	9.40
Denny's	37	7	37.56	93	17	44.49
Young's Farm Spring	37	8	30.29	93	17	44.04

Young's Farm Spring. Comprised of a spring feeding a sinkhole 100 feet away, this system is found ~60 feet east of Campbell road in south Springfield. The spring is surrounded by a crumbling stone springhouse (Figure 8). The spring and sinkhole is a storm water collection point for many roads surrounding it and is inundated with several feet of water during heavy rains. The springhouse is mostly clogged with woody debris

Figure 8: Young's Farm Spring and Sinkhole.

and trash, but a small trickle has forced its way under and around the backside of the springhouse and provides a very slow but steady stream of water to the sinkhole 100 feet downstream. This spring dries up during very dry periods and did go dry for most of the sampling period of this thesis. When possible, water at this site was collected immediately as it left the south side of the ruined springhouse.

Denny's Spring. Denny's spring, previously labeled as Waffle House upwell spring by Berglund (2012) or Ward Branch upwell spring (Tomlin, 2010 and Stanke, 2010) emerges from a low cave just south of the Denny's restaurant at 4760 S. Campbell avenue (Figure 9). The water flows into a small pool, where it is collected, before spilling into Ward Branch ~20 feet from where the spring emerges.

Greenway Trail Spring. Located along the Ward Branch Greenway trail near Republic road and East Monastery street, the spring is found boiling up from below the stream (Figure 10). Upstream of the spring the stream is usually dry most of the year and is only fed with overflow from the hospital pond and storm water runoff. Interestingly, it is easy to spot a fracture within the Burlington-Keokuk where the spring emerges. Unlike Denny's spring, this spring is more robust and only heavy flow from long periods of rainfall might mask the spring's location. It was also noted that a very small spring emerged from the southern bank gravel bar parallel with the bedding plane spring. This spring is perennial and is not visibly affected by drier periods. Water was collected right at the point where it emerges from a small fracture.

Hospital Pond Sink. Northeast of the Greenway trail spring about a half mile is the Hospital Pond. It is connected to the Greenway trail spring via a small drainage channel

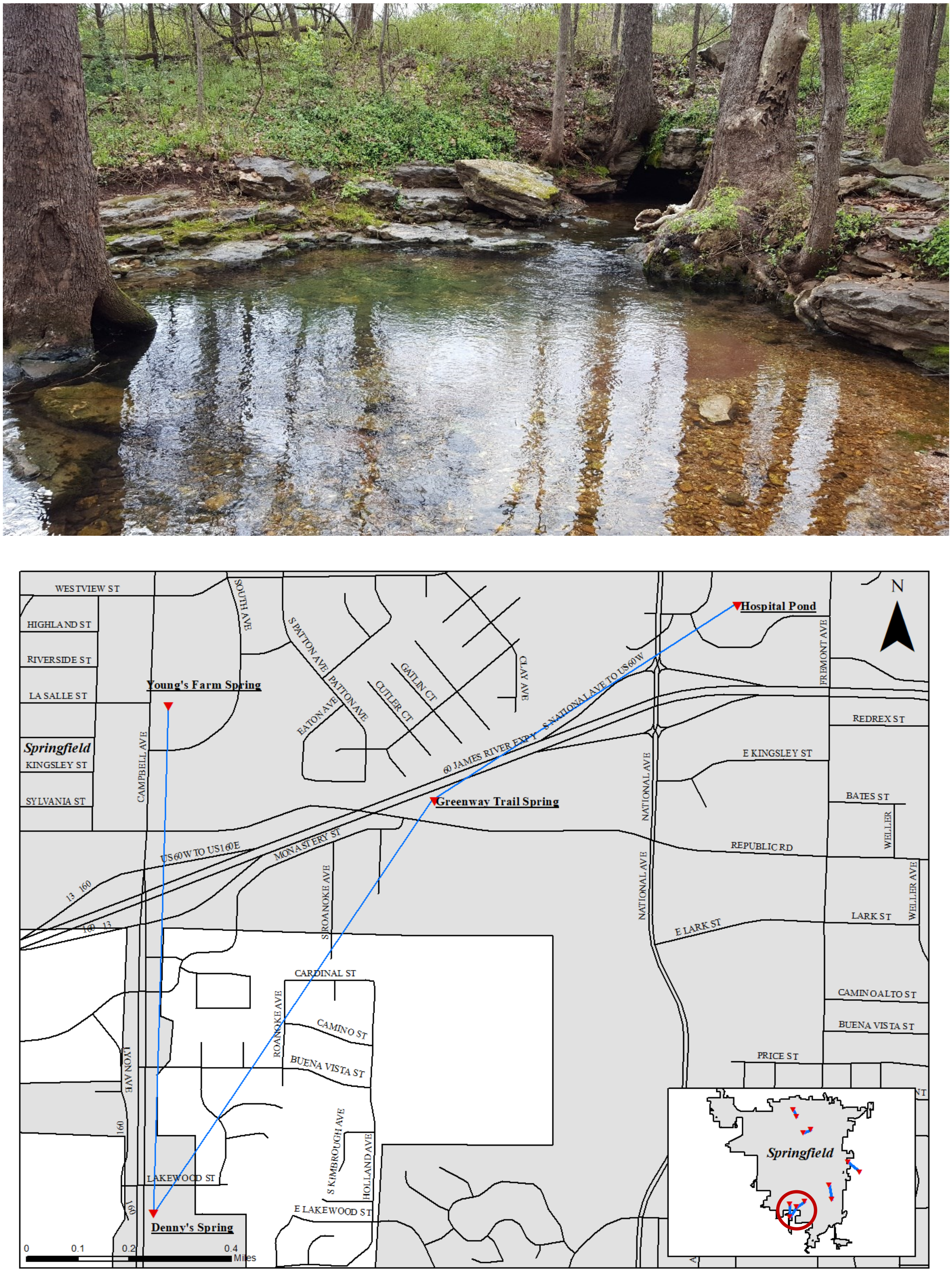


Figure 9: Denny's Spring.

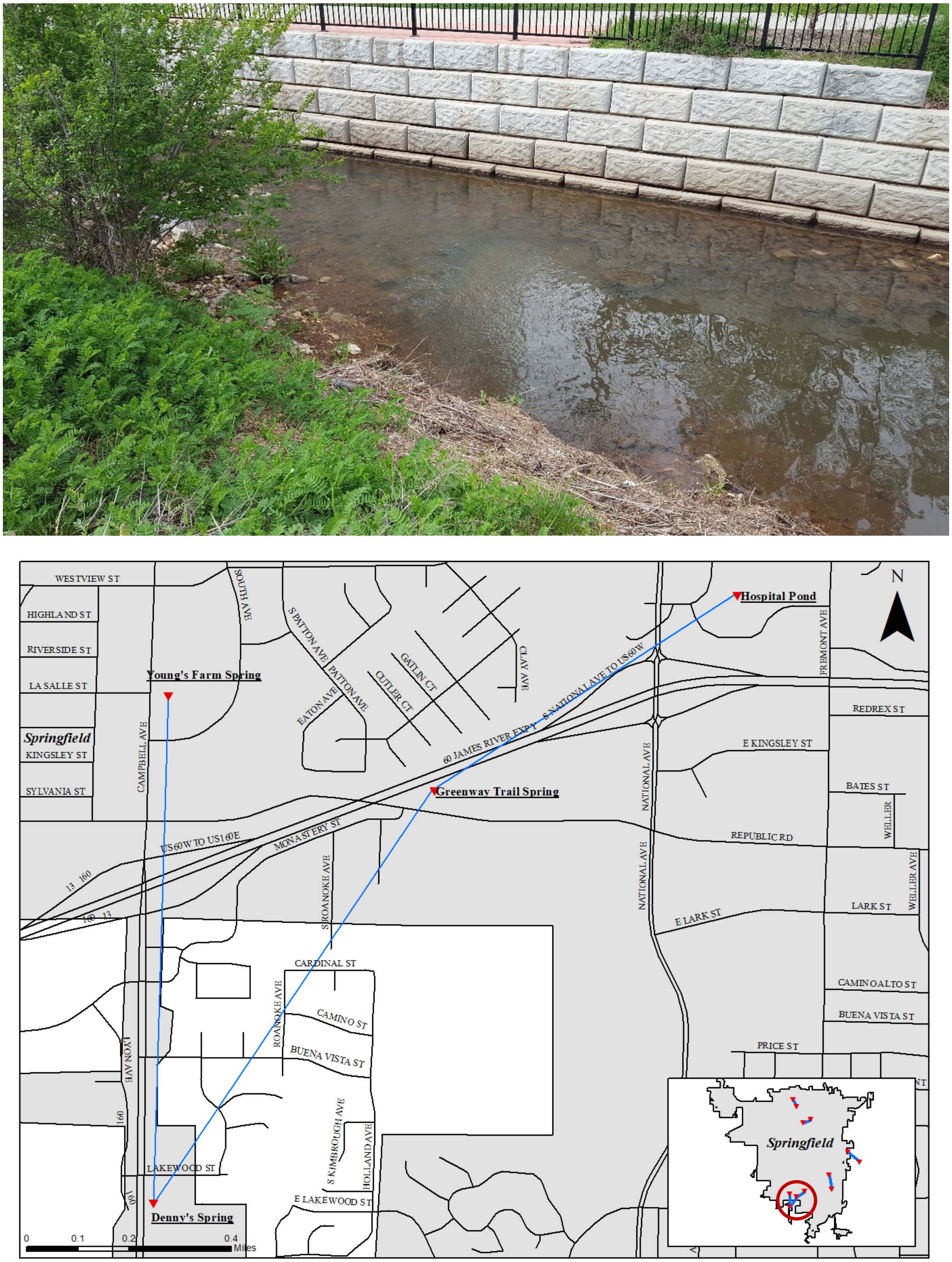


Figure 10: Greenway trail spring.

Which flows only during high outflow from the pond. The pond also has a drainage gully that feeds it from the east during storm events. Samples taken at this site were collected at a point along the southern edge of the pond (Figure 11).

Jones Spring. Located on the east side of Springfield in a rural neighborhood, Jones spring sits on the boundary between two properties. It pours out of a low cave (Figure 12) and flows over a manmade ledge built to support an old grist mill flume (Vineyard and Feder, 1982). Even during dry periods, Jones spring produces a large volume of water and can be heard from about 100 feet away. Despite the high daily flow, the spring pool is clear and aquatic life can be seen. Water from this spring flows downstream, creating Jones Branch, where it both sinks into a groundwater channel and flows overland into Pierson creek (Bullard et al., 2001). Samples were collected from water as it leaves the cave and enters a small pool that forms behind the manmade ledge.

Snow Spring Complex. South of Jones spring is the Snow Spring complex. The Snow Spring complex is two springs that flow into Pierson Creek about 1.5 miles from where Jones Branch (fed by Jones Spring) flows into Pierson Creek (Figure 13). Jones spring contributes a majority of the water in the surface stream. During drier periods, only Jones spring is flowing enough to contribute to the stream, the north snow spring easily runs dry. The sample collection point was located 30 feet south of a small side road bridge along the bank of Pierson Creek.

Southern Hills Pond. Located on the east side of town, one mile southwest of Jones spring is a neighborhood pond that is spring fed. The spring enters at the northeast part of the pond

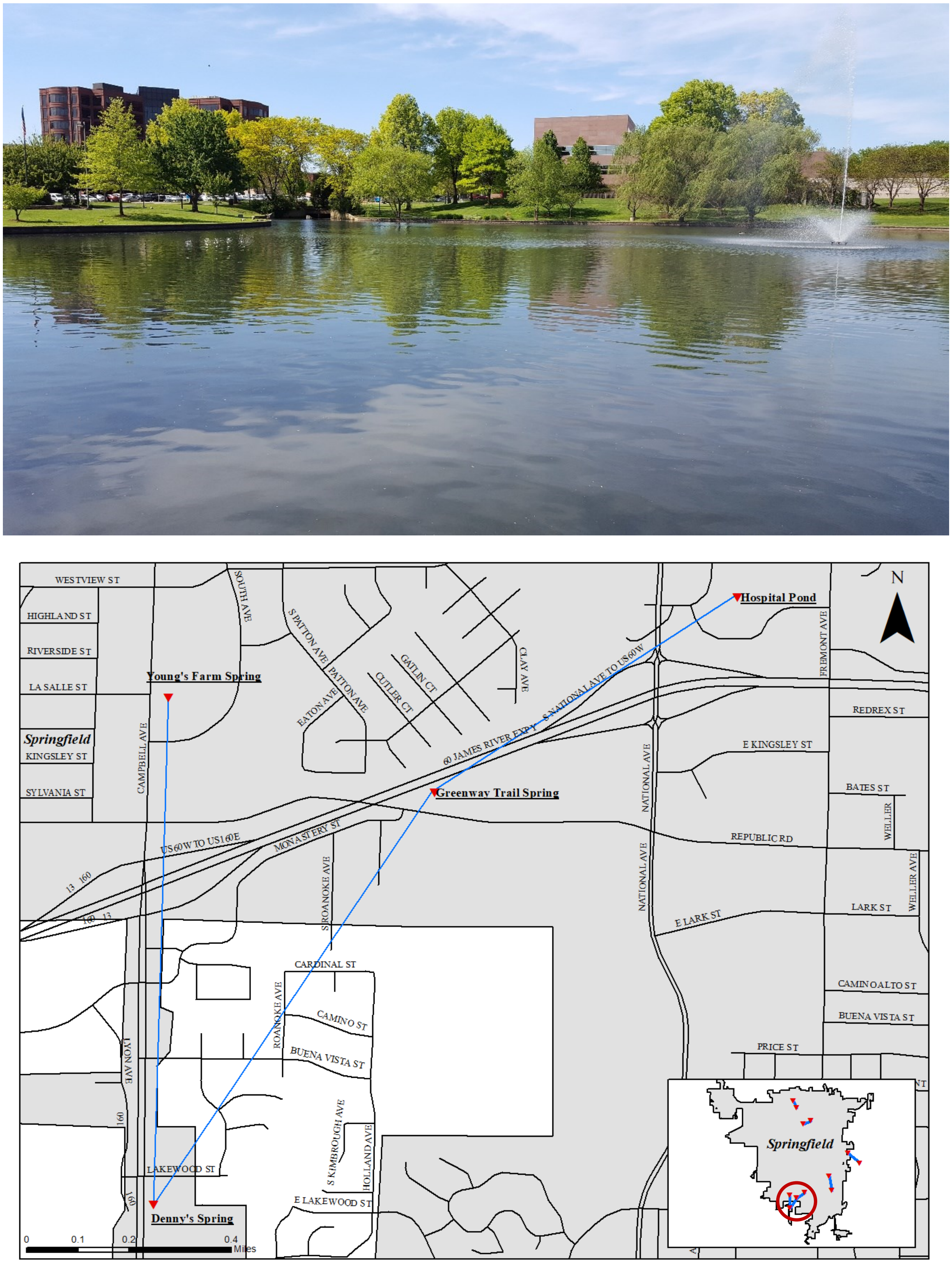


Figure 11: Hospital Pond.

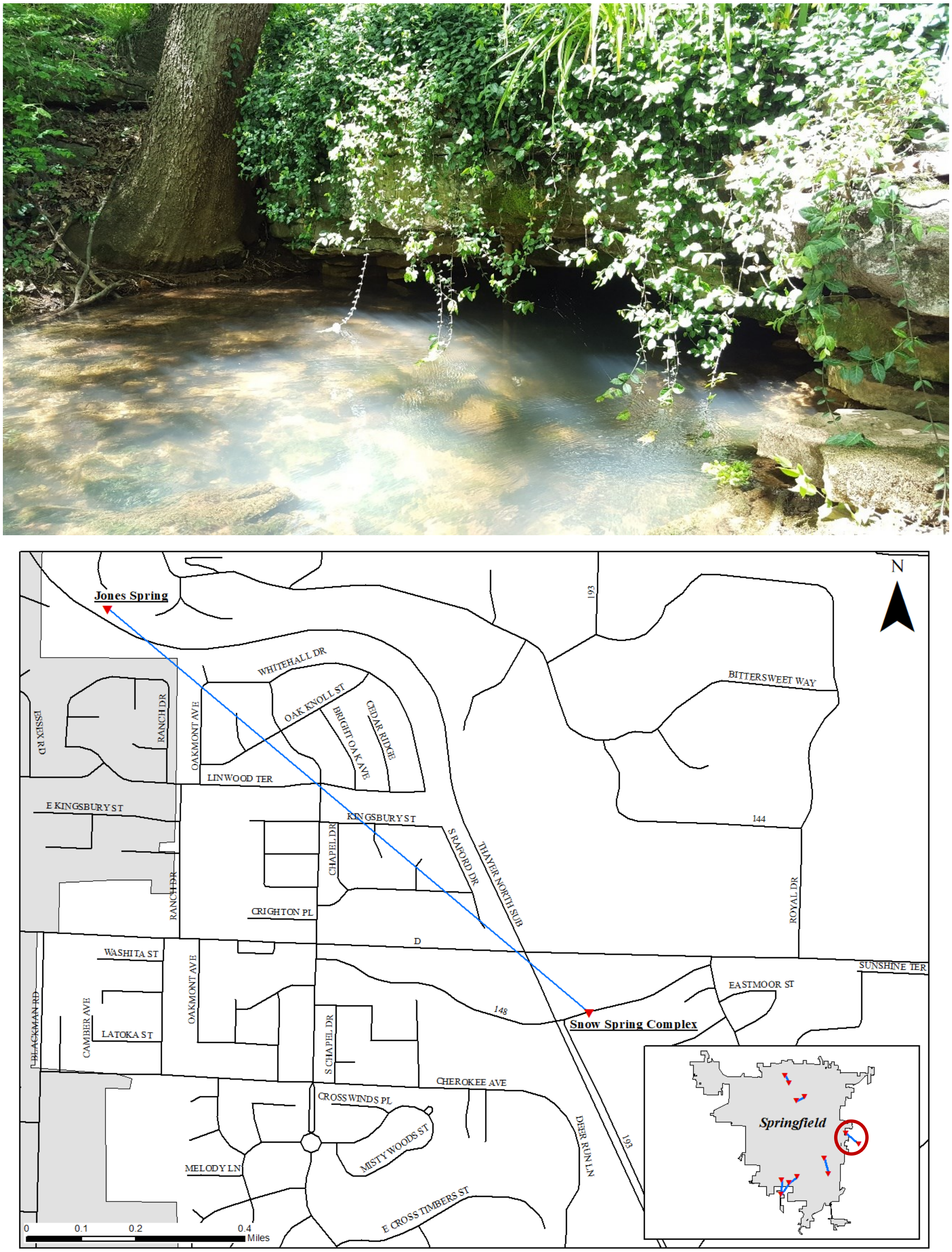


Figure 12: Jones Spring.

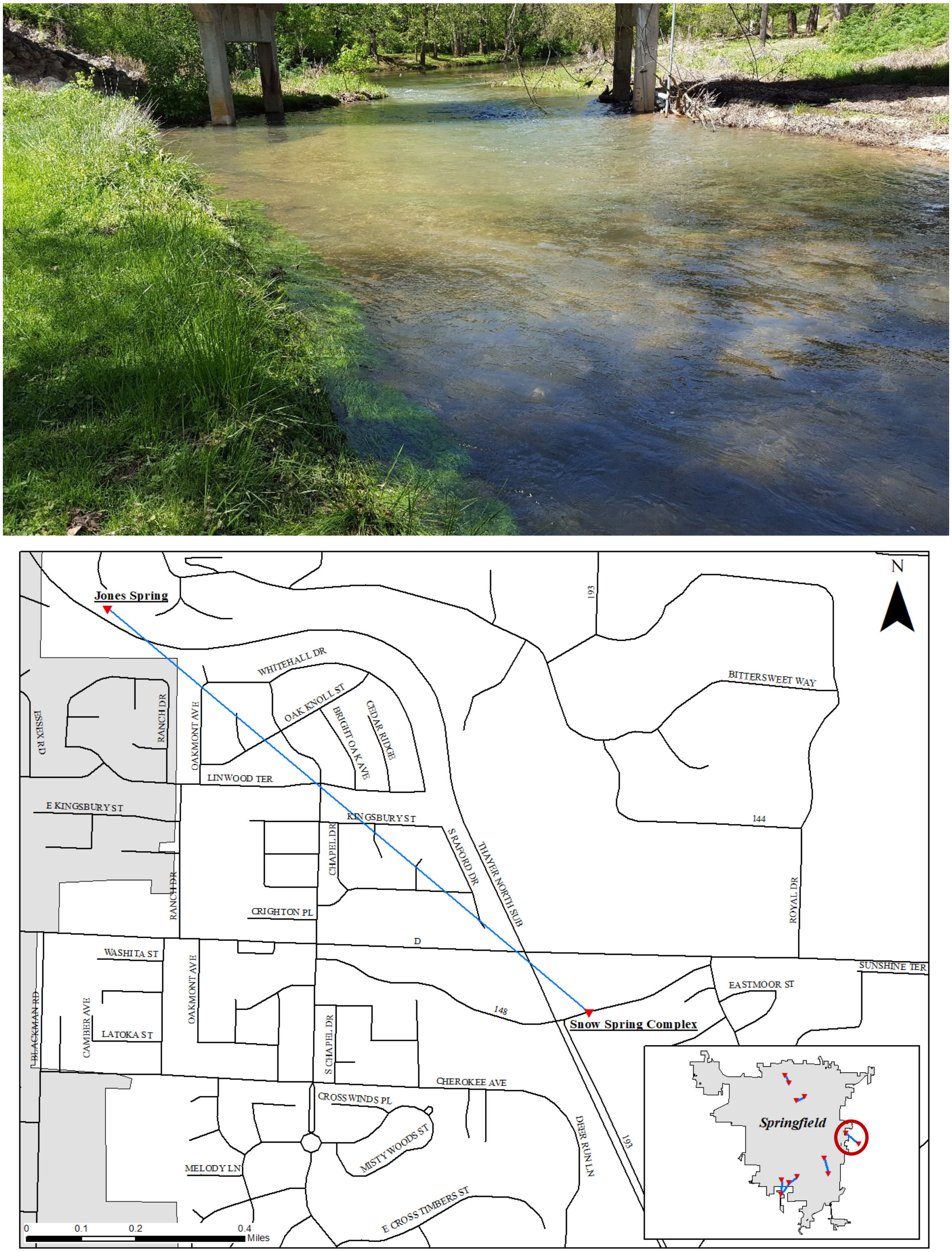


Figure 13: Snow Spring Complex.

and the pond drains over a spillway in the southwest part. The spillway stream meanders for a few hundred feet before disappearing into a sinkhole set into the woods. Water was collected just before it enters the wooded area surrounding the sinkhole (Figure 14).

Sequiota Park. South of Southern Hills pond is Sequiota park. Water flows out of Sequiota cave and into a nearby pond within the park. The stream is roughly 1.5 feet deep as it exits the cave but is much deeper further back in the cave. The stream extends the width of the cave (~15-20 ft.), making any attempt to enter the cave on foot impossible. Samples were collected at the cave opening along the old stone stairs (Figure 15).

Smith Park. Smith Park is a small park in north central Springfield next to the local Boys and Girls club at North Fremont Ave. and Division st. The small park contains several baseball fields, tennis courts, and a playground, all surrounding the north fork of Jordan Creek. Jordan Creek at this point is boxed in by a man-made channel with sloped concrete and stone walls and bottom (Figure 16). Flow in the stream is perennial, but often runs very shallow and is difficult to sample. Samples were collected behind a service building on the north side of Jordan Creek, east of the service bridge.

Silver Spring. Located in Silver Springs Park, the spring pours out of old terracotta pipe one hundred feet southwest of the park amphitheater. To complete the amphitheater the spring was boxed in and rerouted via terracotta pipe to the point where it is today. The channel is mostly open, but watercress forms a braided channel-like system that significantly decreases flow. The spring flows into Jordan creek ~60 feet downstream. Samples at this site were collected right as the water flows out of the terracotta pipe (Figure 17).

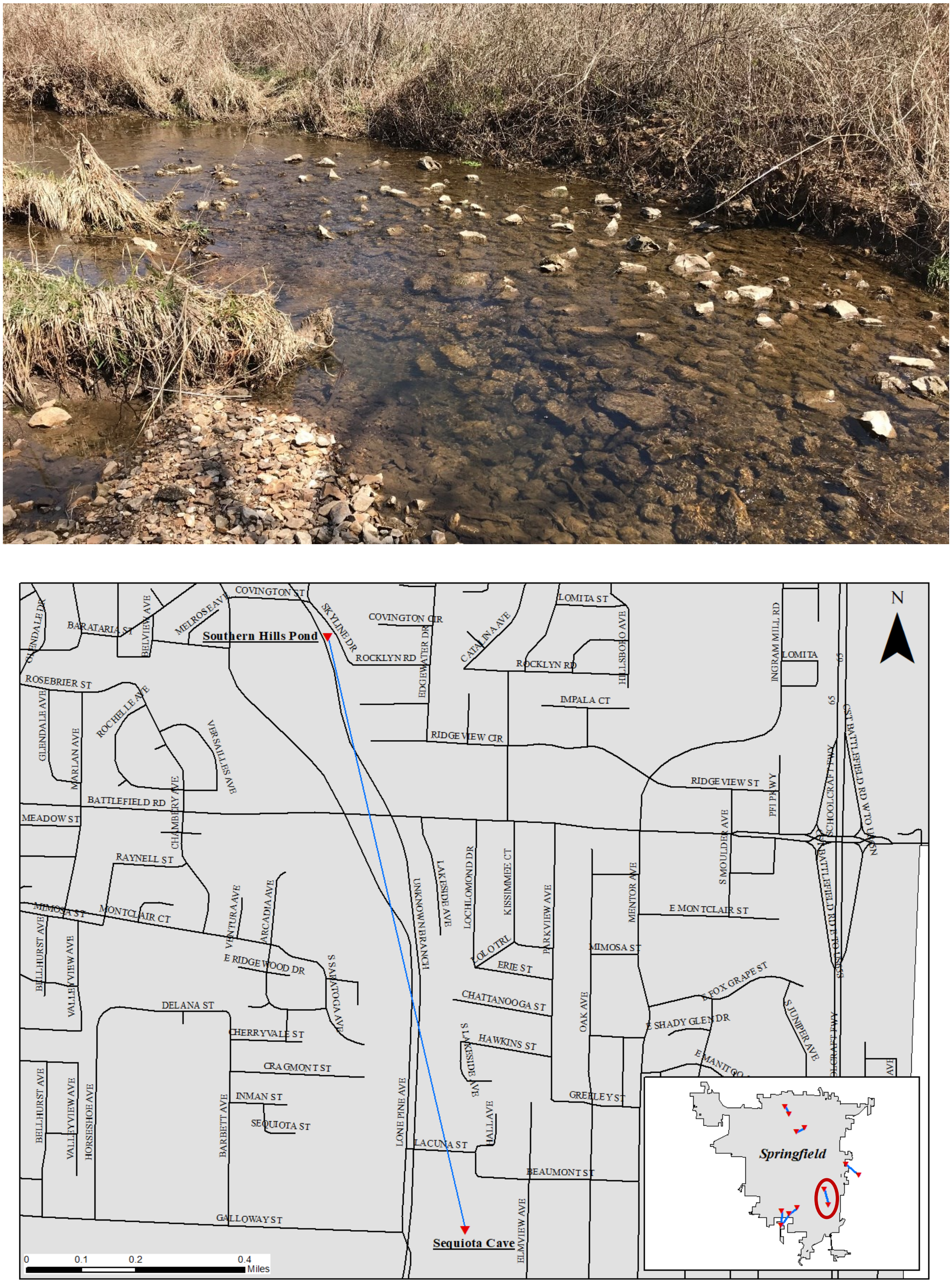


Figure 14: Southern Hills Pond.

Figure 15: Sequiota Park Cave.

Figure 16: Smith Park.

Figure 17: Silver Spring.

Baptist Bible College Spring. The college spring flows out of a large hill next to a maintenance building on campus. Looking into the small opening, it is easy to spot a very small cave (or rather large gap between bedding planes) where the spring emerges. The spring was once well kept as a concrete opening and channel were made to direct the flow into a larger concrete box pond. The samples for this site were collected in the concrete channel as the water emerges from the hillside (Figure 18). Though not directly connected by dye traces, this spring lies along the estimated flow path between an unnamed sinkhole (south) to the Doling Park cave spring (north) and is believed to be related to the system.

Doling Park Cave Spring. Two springs can be found in Doling Park. To the west a small spring trickles out of the hillside, flowing north and following the contour of the park pond. The second, and most prominent, is the Doling Cave spring. Tucked away behind a clump of trees the cave is wide with a large overhang before a bat gate that spans the opening. The spring flows out of the left side of the entrance and into a pond several hundred feet away. Flow is perennial and follows a well-entrenched, natural channel. During periods of high outflow the channel is overburdened and covers most of the trail leading up to the cave. Samples were taken five feet from where the water rushes through the bat gate into the outside channel (Figure 19).

Measurements and Sampling

Of the eleven variables used in this study, five were collected in the field. The remaining six were analyzed in labs using two different techniques. The various instruments and methods used to collect this data are described below.

Figure 18: Baptist Bible College Spring.

Figure 19: Doling Park Cave Spring.

pH. A Horiba pH and temperature electrode attached to a Horiba D-54 multi-parameter meter was used to test for pH and temperature (Celsius) simultaneously. Before every sampling day the probe was calibrated using a pH 7 standard and pH 4 standard. The probe was cleaned using deionized (DI) water and dried before being packed away for field analysis. When in the field the pH meter was turned on, the pH bulb and thermometer were rinsed in the source water for at least 10 seconds, and the refill slider was opened to allow air flow and increase reading stability. The probe was then dipped in the source water at least two inches and held until the meter stabilized. Both pH and temperature in Celsius were documented before moving on to the next test. Any readings that were suspected incorrect (either from instrument error or human error) were retested three more times.

Conductivity. A Horiba conductivity probe attached to the second channel of the Horiba D-54 meter was used to determine conductivity. The probe was rinsed in the source water before moving upstream a few feet to collect readings. The probe was dipped into the water and rapidly delivers a conductivity (mS/m) and temperature (Celsius) reading.

Temperature. As noted above, temperature readings are documented from both the pH and conductivity probes. The temperature reading on the conductivity probe was consistently higher than on the pH probe. For the statistical analysis, the pH probe temperature reading was used but both readings were recorded in the field to be referred to later if needed.

Flow. The flow measurement was calculated in cubic feet per second = area x velocity. The area was determined by stepping off the width and estimating the depth.

Velocity was calculated using the leaf method, where a leaf was dropped into the water and the time it took to travel a determined distance was logged. The time (in seconds) was multiplied by the length traveled and the channel area to get a value in cubic feet per second (cfs).

Bicarbonate. Bicarbonate was found using an alkalinity titration kit from Hach. In the field, glassware was rinsed three times apiece at the previous site and three times apiece at the new site before being filled for testing. Due to the difficulty of spotting a color change in the field, the phenolphthalein indicator packets included were not used. Instead, a pH meter was used to reach the required pH (Hach, 2013). When finished the flask was dumped downstream and rinsed three times (at least) before being packed away. The delivery tube and pH probe were also rinsed before moving on to the next site.

Water Collection. Water for lab analysis was collected in two bottles – one for City Utilities Blackman lab and one for the Missouri State University (MSU) Chemistry Labs. Polyethylene Nalgene bottles were used to store the water samples. Water was pumped through a peristaltic pump taken from a Sigma portable water sampler and powered with a DeWalt cordless drill. Ten feet of ½ inches vinyl tubing with an inlet filter was attached to the pump intake. Whatman GD/X and XP .45 micron filters were used to remove most organics and sediments from the sample. To allow for the filter to attach on the outlet, the ½ inches diameter tube had to be reduced to ¼ inches tubing. Tubing and filters are connected using tubing adaptors and metal screw clamps. Before a new .45 micron filter was attached at each sample site, water was pumped through the system for 10 seconds to remove any excess water from the previous site to prevent cross contamination. Bottles sent to the MSU chemistry department were acidified using a 1:1

nitric acid solution to prevent attachment of metals to the bottle. Once collected, the bottles were placed in a refrigerator until delivery to the respective labs. Bottles were delivered to the lab and analyzed within no more than a month of collection.

The cation and anion variables could not be tested in the field and required lab instruments to measure. Calcium, magnesium, and sodium were analyzed using the atomic absorption instrument in the Missouri State University Chemistry department. The standard methods for direct air-acetylene flame were used (Eaton et al., 2005). The calcium standard was 5.0mg/L, the magnesium standard was .5mg/L, and the sodium standard was 1.00mg/L. Chloride, sulfate, and nitrate were analyzed using a Thermo ICS 5000 ion chromatograph at the Blackman Lab run by the Springfield's City Utilities. The EPA 300.0 standard method was used.

Statistical Analysis

Multiple statistical packages can run a discriminant function analysis. However, SAS was chosen for two reasons. The first is because SAS is particularly useful for classifying new cases (samples) as compared to SPSS which contains no classification phase at all (Tabachnick and Fidell, 1996). SPSS was not used to classify blind or unknown samples, but instead used to confirm the classification accuracy of the samples used to create the discriminant function in SAS 9.4. The second reason was that this study expanded on work done by Gouzie (1986) using SAS 79 and Lockwood and Gouzie (2015) using SAS 9.4. The SAS 9.4 software package was used to run the DISCRIM procedure and IBM SPSS statistics 24 was used for checking the model and other basic statistics. The Discriminant function was chosen over other classification functions (i.e. ANOVA and MANOVA) because of its ability to develop a discriminant criterion to classify observations into one of a few different groups based on more than one quantitative variable.

Of the 11 variables collected, only 10 were used for the analysis. Flow was not used for the analysis because its purpose was strictly to note periods where flow was either much higher or lower than usual. Unusually high or low flow would likely produce outliers.

When using DISCRIM, there are several limitations and assumptions that can inhibit the classification and predicting capabilities of the model. These limitations are the same limitations that can inhibit multivariate analysis of variance (MANOVA) tests (Tabachnick and Fidel, 1996). The first limitation is the unequal sample size and missing data. While no major problems arise from having an unequal sample size for each basin, it is strongly suggested that the smallest group have more samples than the total number of predictor variables used in the study (Tabachnick and Fidel, 1996). For this study, there are 10 variables, and the smallest number of samples in a basin is eight (Sequiota basin). While this appeared to be an issue before models were run, the eight samples are very close to the minimal 10 and did not appear to have a major impact on the classification.

The assumption of normality is essential to ensuring the robustness to failures. This is the case if the violation of normality is caused by skewness, not outliers within the dataset (Tabachnick and Fidel, 1996). Histograms were created for each variable to test for skewness (Appendix B). It was determined that any value 3σ from the median was classified as an outlier, and 18 outliers were removed, mainly from the Hospital pond sample site. Most of the outliers were caused by very low temperature, conductivity, bicarbonate, nitrate, calcium, and magnesium values. It is likely these values were low because the pond lacked a consistent inlet of water, and acted mainly as a sink. Water was sourced from occasional rainfall or ported in from elsewhere, not from a constant inlet from a spring or stream.

Since this thesis was focused on classification of samples, the homogeneity of variance-covariance matrixes was important. If the test of homogeneity of within variance matrices was significant, then the matrices would be used in the discriminant function (SAS Institute Inc., 2017).

To test the homogeneity of variance-covariance matrices, SAS ran Anderson's test via the POOL=TEST function (Tabachnick and Fidel, 1996). When the POOL=TEST function was run, the chi-squared value was very significant at <.0001, implying that SAS used the matrices in the discriminant function.

The discriminant function (or classification criterion) for each group (basin) was calculated by the measure of generalized squared distance (SAS Institute Inc., 2017). Each observation was then placed into a class (group) in which it had the smallest generalized squared distance and a posterior probability of the observation belonging in that group was calculated. SAS did this through several steps highlighted below:

The squared Mahalanobis distance is used to determine the distance from x to group t. The equation is:

$$d_t^2(x) = (x - m_t)' V_t^{-1} (x - m_t) \qquad \text{(SAS Institute Inc., 2017)}$$

Where Vt = the covariance matrix within group t since the within-group covariance matrices are being used.

Then the group specific density at x from group t was given by:

$$f_t(x) = (2\Pi)^{-\frac{p}{2}} |V_t|^{-\frac{1}{2}} \exp(-0.5 d_2^t(x))$$ (SAS Institute Inc., 2017)

The posterior probability of x belonging to group t was then calculated using Bayes' theorem:

$$p(t|x) = \frac{q_t f_t(x)}{\sum_u q_u f_u(x)}$$ (SAS Institute Inc., 2017)

Then the generalized squared distance from x to group t was defined:

$$D_2^t(x) = d_t^2(x) + g_1(t) + g_2(t)$$ (SAS Institute Inc., 2017)

Finally, the posterior probability of x belonging to group t was found by:

$$p(t|x) = \frac{\exp(-0.5 D_t^2(x))}{\sum_u \exp(-0.5 D_u^2(x))}$$ (SAS Institute Inc., 2017)

In order to create a discriminant function (model) that was statistically significant, not all of the samples collected could be used. It was decided that the dataset be split 80/20, with 80% of the samples (87 samples) used to develop the discriminant function (model), and the remaining 20% samples (25 samples) used as "blind" samples. Since the 25 "blind" samples were not used to create the model, they were new to SAS and could be considered blind. These blind samples were introduced one at a time, as a sample from an unknown basin, into the model, and the percent classified table was used to determine how accurate and robust the basin classification was.

Chemical Analysis

Ten water quality variables that were used in this study are commonly found in many similar studies and publications. The variables can indicate much about the land use and local geology that surround and make up groundwater basins sampled. For this study, the water chemistry was not the main focus, but charge balance errors and saturation index values were calculated and analyzed.

A charge balance error takes the combined charges of the major cations and anions found in a water sample and calculates the percent of the water sample that they comprise. Values ranging from 0-8% suggest that all major ions were accounted for in the analysis. Anything greater than 8% suggests there are other major ion concentrations that contribute to the water quality. Further analysis would be needed to determine what the respective ion or ions were and what their contribution to the water chemistry was. For this study, six ions were analyzed and run through the charge balance error (CBE) equation to calculate the error percent. The CBE was calculated using the equation below:

$$CBE\ \% = \frac{\sum cations + \sum anions}{\sum cations - \sum anions} * 100$$

The saturation index for each basin was calculated using the Lenntech online Langelier saturation index calculator (Lenntech, 2018). The Langelier Saturation Index formula is as follows:

$$LSI = pH + \log\left(\frac{K_a * \gamma_{Ca^{2+}} * (Ca^{2+}) * \gamma_{HCO_3^-} * (HCO_3^-)}{\gamma_{H+} * K_{sp}}\right) \qquad \text{(Lenntech, 2018)}$$

All values from the saturation index are either positive or negative, where negative values are undersaturated with respect calcite and positive values are supersaturated with respect to calcite. The dissolution of calcite increases as the time in which the groundwater is in contact with the source rock increases.

CHAPTER 4 – RESULTS

Data were successfully collected over 5 ½ months from March to September 2017 (Appendix A). Data were grouped by groundwater basin and run through various statistical and chemical analyses. The discriminant function analysis was used to test the idea that each basin has a unique geochemical signature. Charge balance error percent and saturation indices for every sample site was also calculated. Results from these analyses are described below.

Charge Balance Error

The charge balance error percent (CBE) was calculated for every sample site every sample period and graphed below (Figure 20). It was chosen that errors ranging from 0-8% represented a well-rounded and thorough chemical analysis. There were only three occurrences where the CBE lay above 8%. The other values ranged from 3-5.5%, suggesting that the chemical constituents analyzed were the primary components of the water. While the value for each sample site was different, each increased or decreased in a similar trend. Some outliers are apparent during late April and early May. There was no general increase or decrease trend in CBE percent with time over the sample season, however there was a noticeable increase in percent during April and May, while the lowest calculated values occurred in June and July, before increasing again in August.

Saturation Index

Saturation index values were calculated using the Lenntech online Langelier Saturation Index calculator (Lenntech, 2018). The graphs below plot the saturation index value for every site during the sample season (Figure 21, Figure 22, Figure 23, Figure 24, and Figure 25)). Only five values were saturated with respect to calcite, though the values indicate they were only moderately oversaturated. Saturated values ranged from .05 to .76. Undersaturated values ranged from slightly undersaturated (-.01) to very undersaturated (>-2.0). Over the course of the sample season, there was a decreasing trend in saturation levels at every sample site.

Model One Results

Ten variables (pH, temperature, conductivity, chloride, sulfate, nitrate, calcium, magnesium, and sodium) were used in the analysis. When the first model was run, basin classification accuracy ranged from 88% to 100% (Table 2). These percents were satisfactory and foreshadowed success in the blind sample analyses. Each test consisted of adding one blind sample to model one. The model was rerun every time a new sample was added, and the table of misclassified samples of the posterior probability of membership within the analysis output was used to determine if the sample was correctly classified or not. When the first model was run with the blind samples, it was able to correctly place 20 of 25 samples into their respective basin (80% placement accuracy) (Table 3). Of those not correctly classified, the model produced probability values that represented the probability that the sample belonged to another basin. Samples not correctly classified usually had a lower membership probability. In four of the five misclassified samples, the basin with the next highest probability was the correct basin that the sample belonged in.

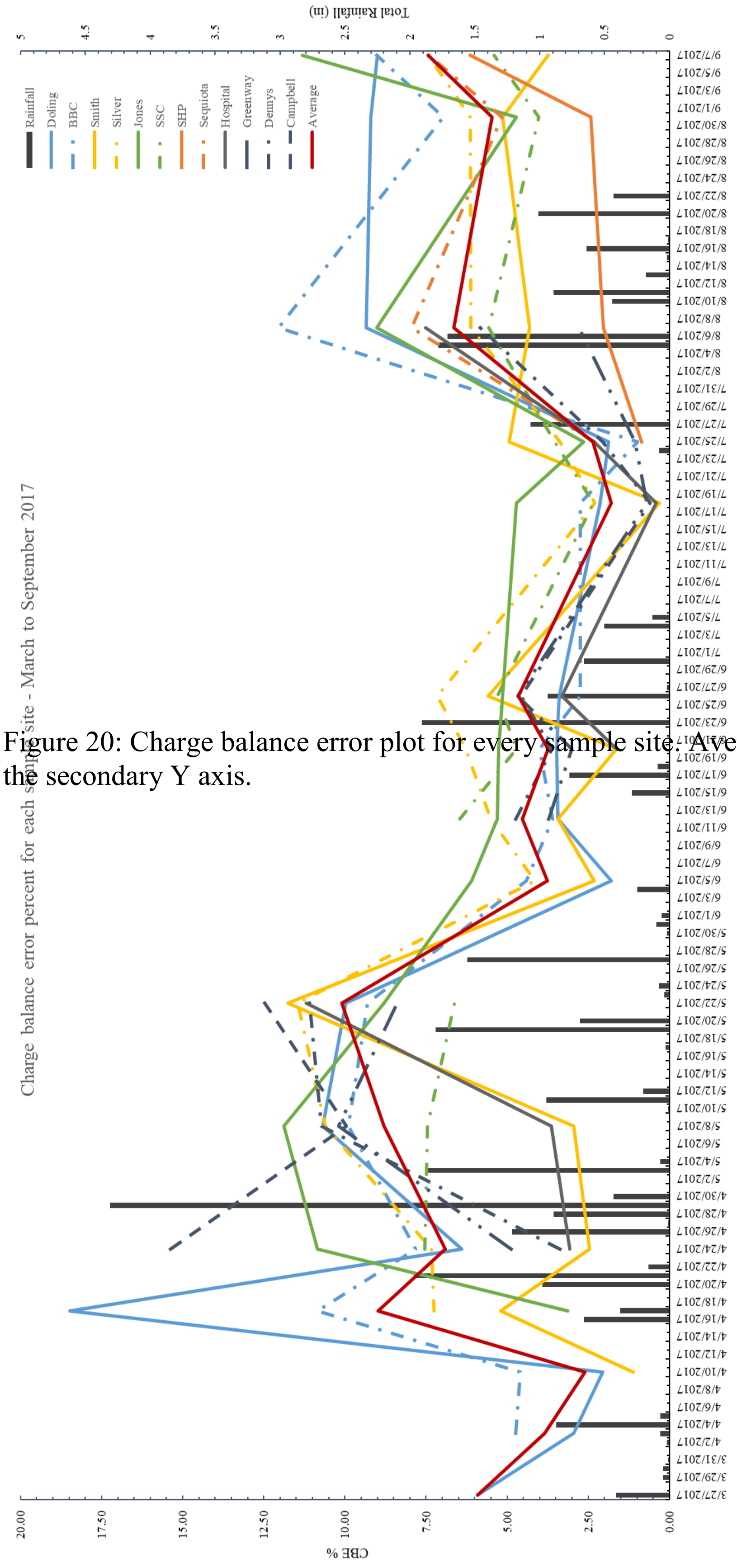


Figure 20: Charge balance error plot for every sample site. Average CBE percent denoted as red
the secondary Y axis.

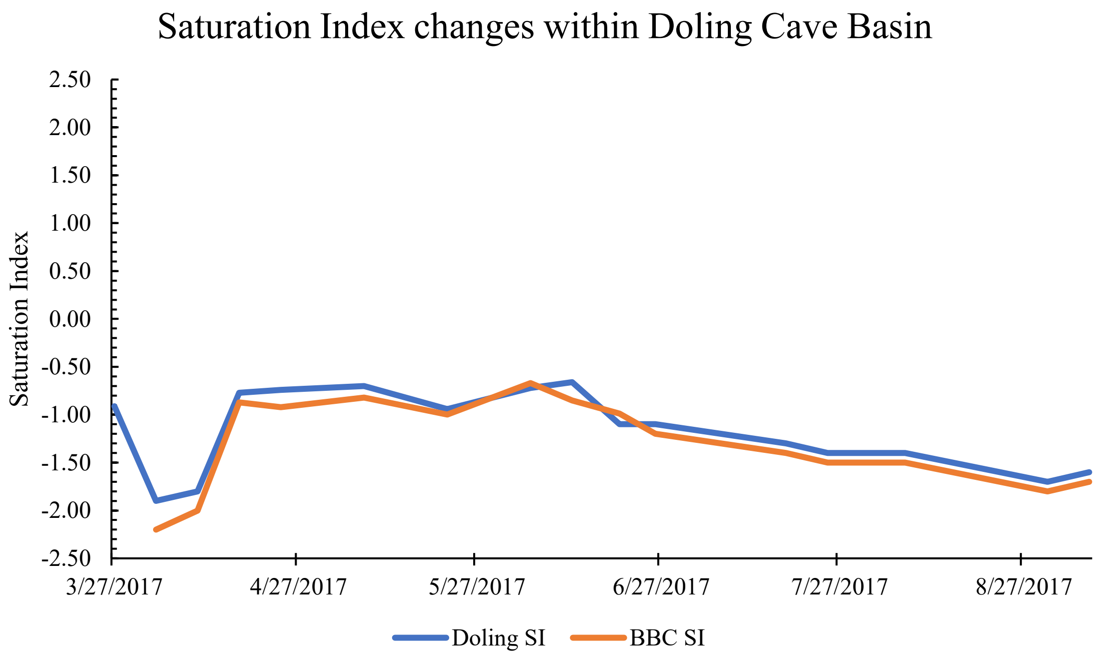


Figure 21: Saturation index for Doling Cave Basin.

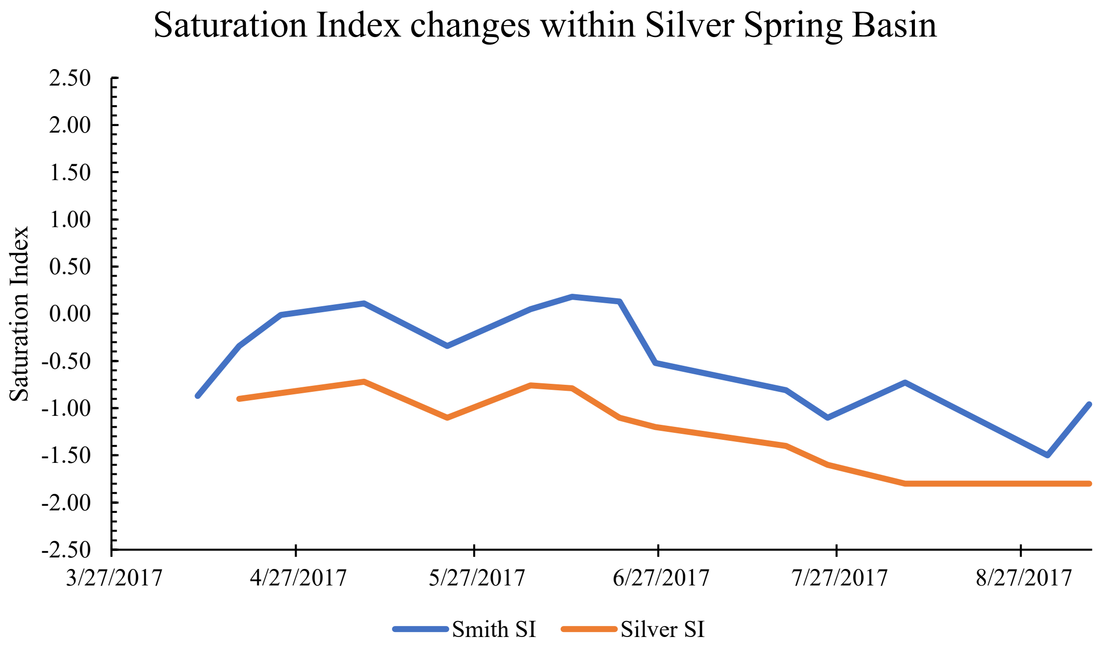


Figure 22: Saturation index for Silver Spring Basin.

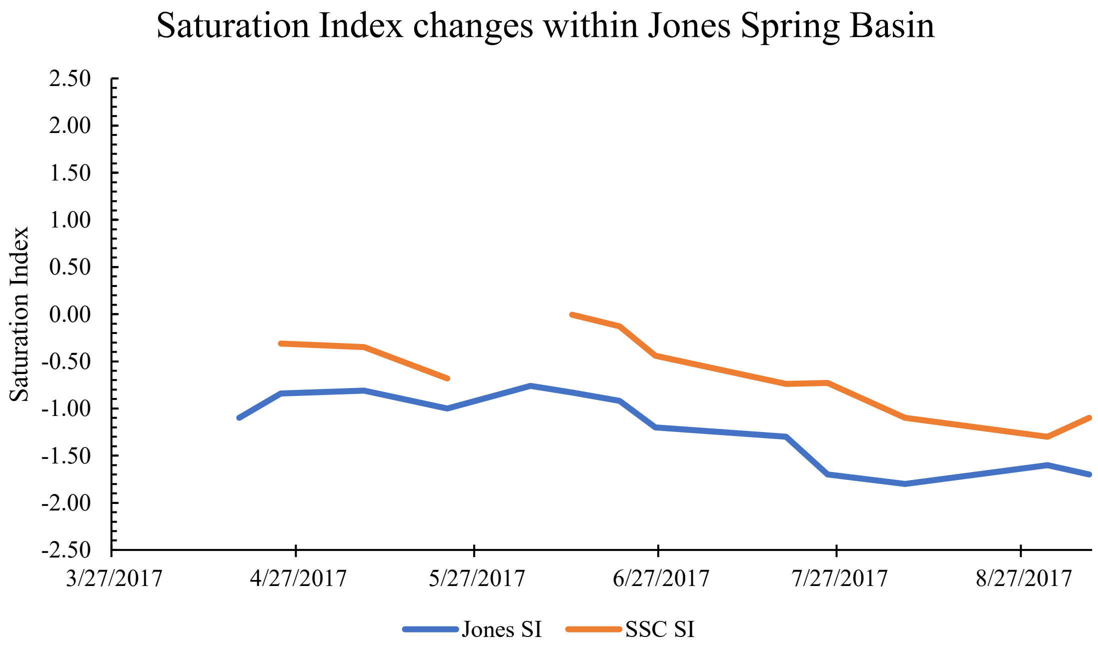


Figure 23: Saturation index for Jones Spring Basin.

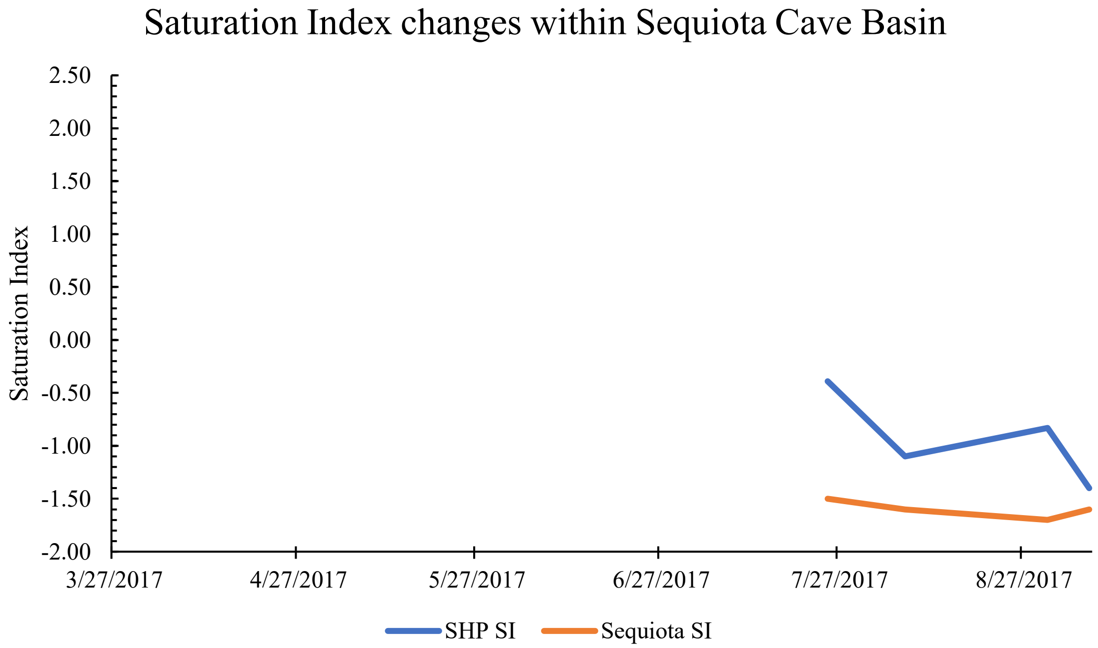


Figure 24: Saturation index for Sequiota Cave Basin.

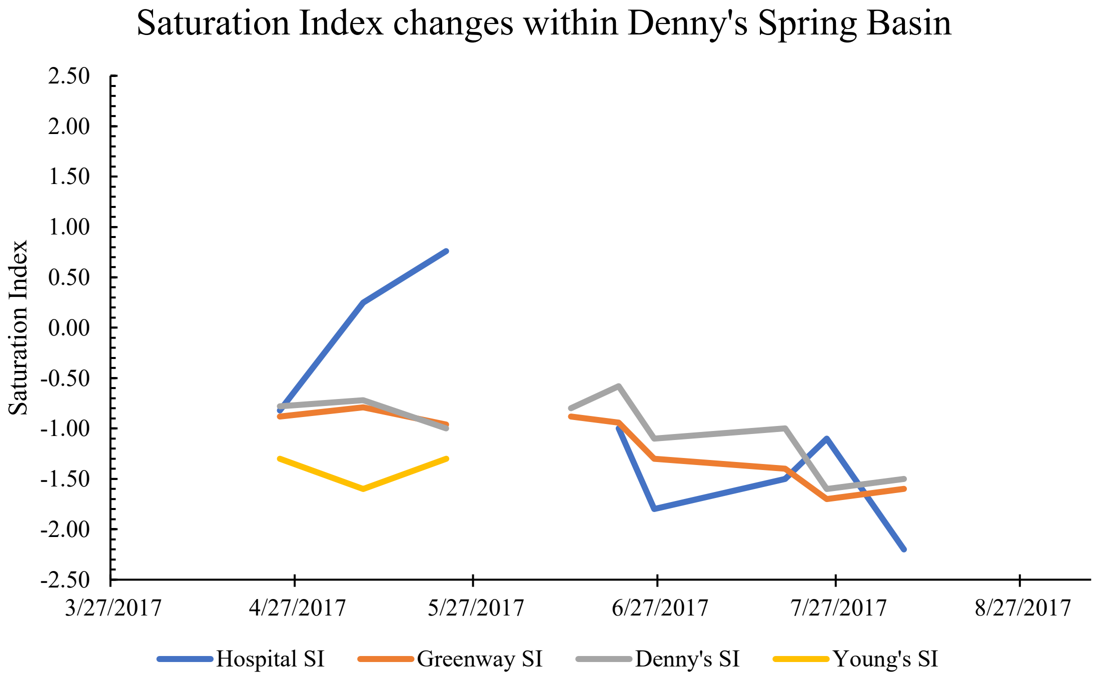


Figure 25: Saturation index for Denny's Spring Basin

Table 2: Percent classified table for model one. Percent ranged from 88% to 100%.

The SAS System

The DISCRIM Procedure
Classification Summary for Calibration Data: WORK.MODELONE80
Resubstitution Summary using Linear Discriminant Function

Number of Observations and Percent Classified into Basin

From Basin	Denny's	Doling	Jones	Sequiota	Silver	Total
Denny's	18	0	0	0	0	18
	100.00	0.00	0.00	0.00	0.00	100.00
Doling	0	24	0	0	0	24
	0.00	100.00	0.00	0.00	0.00	100.00
Jones	0	0	15	2	0	17
	0.00	0.00	88.24	11.76	0.00	100.00
Sequiota	0	0	0	6	0	6
	0.00	0.00	0.00	100.00	0.00	100.00
Silver	0	0	2	0	20	22
	0.00	0.00	9.09	0.00	90.91	100.00
Total	18	24	17	8	20	87
	20.69	27.59	19.54	9.20	22.99	100.00

Table 3: Results from the first blind sample run with model one. Samples colored in red were misclassified.

ID #	Basin	Placed Basin	Confidence	Next Closest	Confidence	Actual Basin	Correct?
21	Doling	Doling	-	-	-	Doling	YES
96	Doling	Jones	0.833	-	-	Jones	YES
254	Doling	Denny’s	0.9972	-	-	Denny’s	YES
214	Doling	Jones	0.9849	-	-	Jones	YES
112	Doling	Silver	0.9958	-	-	Silver	YES
180	Doling	Jones	0.6276	Silver	0.3189	Silver	NO
209	Doling	Jones	0.7822	Silver	0.2170	Jones	YES
49	Doling	Sequiota	0.7467	Jones	0.1952	Jones	NO
208	Doling	Jones	0.9923	-	-	Jones	YES
147	Doling	Silver	0.9992	-	-	Silver	YES
187	Doling	Denny’s	1.000	-	-	Denny’s	YES
218	Doling	Doling	-	-	-	Doling	YES
225	Doling	Jones	0.8780	Silver	0.1100	Silver	NO
186	Doling	Doling	-	-	-	Doling	YES
26	Doling	Jones	0.9919	-	-	Jones	YES
93	Doling	Doling	-	-	-	Silver	NO
99	Doling	Denny’s	1.000	-	-	Denny’s	YES
80	Doling	Sequiota	0.8656	Silver	0.1342	Sequiota	YES
100	Doling	Sequiota	0.4402	Denny’s	0.2635	Denny’s	NO
148	Doling	Doling	-	-	-	Doling	YES
171	Doling	Denny’s	0.7224	Sequiota	0.269	Dennys	YES
46	Doling	Doling	-	-	-	Doling	YES
213	Doling	Doling	-	-	-	Doling	YES
139	Doling	Denny’s	1.000	-	-	Denny’s	YES
83	Doling	Jones	0.7301	Sequiota	0.2687	Jones	YES

CHAPTER 5 – DISCUSSION

Charge Balance Error

The major influence on the charge balance error percent for every sample site was the total amount of rainfall. The relationship between the rainfall and charge balance error percent shows that charge balance errors are affected by rainfall events. There were several noticeable peaks and low points in the graph that correlate with wet and dry periods. For every site there was a major increase in percentage from 4/8/2017 to 6/1/2017, a slight upward trend from 6/21/2017 to 7/3/2017, and another large upward trend from 7/29/2017 to 8/16/2017. Two downward trends occurred from 6/2/2017 to 6/11/2017 and from 7/5/2017 to 7/23/2017. All of these trends correspond to the amount of rainfall (or lack thereof). High rainfall events in late April, late June, and mid-August all result in higher CBE percentages at all sample sites. Periods with low or no rainfall (early June, mid-July, and early September) correspond with much lower CBEs. Where there is little to no rainfall, it can be expected that the major source of ions in the water is the surrounding source rock and soils. When major rainfalls occur, the excess water collects other contaminants and pollutants not usually collected by runoff and flushed through the system. This increases the amount and range of ions in solution and increases the charge balance error percent.

Saturation Index

Almost every value was undersaturated with respect to calcite. There were six values that were saturated with respect to calcite. Two of the saturated values were from hospital pond and were likely caused by other outlier values used in the saturation index calculation (most likely a combination of high pH and very low bicarbonate). The other four saturated values were from Smith park, where water was collected from a very shallow surface stream. Based on the undersaturated values for every sample site for most of the year, and the physical distance between each site, it is likely that water is moving through each basin via conduit flow, not diffuse flow.

When the saturation indices were plotted in comparison to water total CO2 levels of the sample sites, several observations could be made about the correlation between the two. In general, sample sites that were surface streams or ponds had much lower levels of CO2, while spring waters had much higher CO2 levels. This was inversely related with saturation indices, where higher CO2 levels resulted in lower saturation indices. Almost all saturation indices were undersaturated with respect to calcite. Each basin is discussed individually below.

Doling Basin. Doling basin was comprised of two springs along a groundwater flow path. When plotted, it was apparent that both Doling Park cave and the Baptist Bible College (BBC) spring saturation values were very similar, with Doling having a slightly higher saturation index value (Figure 26). This corresponded with Doling having lower CO2 levels than the BBC spring. While both springs were sampled relatively close to where they emerged from the subsurface, the water from Doling Park cave traveled through a larger cave passage for several hundred feet before it was able to be sampled. This increased the water contact with the air and increased the amount of off-gassed CO2. The smaller amount of CO2 in the Doling Park cave water resulted in more saturated water (with respect to calcite) as compared to the BBC spring water.

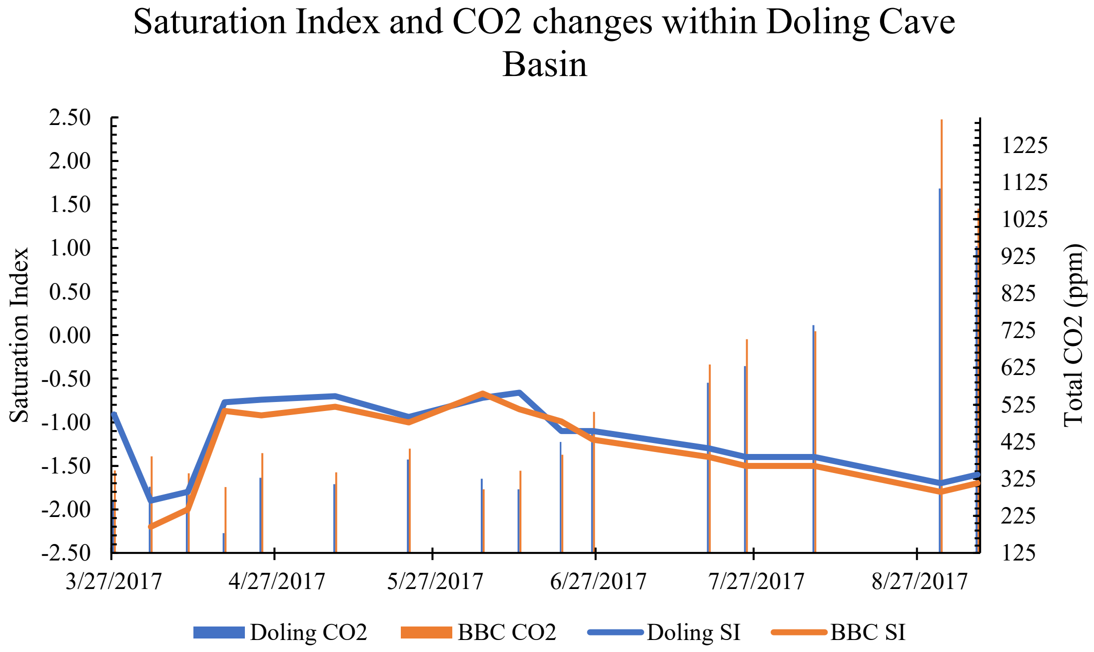


Figure 26: Saturation index compared to CO2 levels in the Doling Cave Basin.

Silver Spring Basin. The two sample sites of Silver Spring basin had consistently different saturation values. Smith park samples were always more saturated with respect to calcite than Silver Spring park (Figure 27). This is probably because the water was sampled at Smith park from a surface stream. Surface streams in the Ozarks are warmer, and higher in pH when compared to groundwater. The lower amount of total CO2 found in surface streams results in a higher pH (which does not dissolve limestone easily) and warmer waters, which also do not dissolve limestone as well as colder waters. Both of these features and the lack of contact with limestone resulted in high saturation levels (and likely very low saturation capacity). Carbon dioxide levels in Silver Spring match CO2 levels seen in other springs sampled in this thesis. The downward trend in CO2 towards late summer seen in Silver Spring is most consistent with Doling Cave Basin to the north. The minor trends in CO2 levels are consistent across both basins, and suggest that Silver Spring may be more connected to Doling Park Basin (Even dye traces suggest they are physically connected).

Jones Spring Basin. In Jones Spring basin, samples taken from Snow Spring complex (SSC) came from Pierson Creek (surface stream), while the samples from Jones Spring came from a spring. Much like Silver Spring basin, the water from the surface stream (SSC) had much lower total CO2 levels as compared to their spring counterpart (Figure 28). The lower CO2 levels resulted in much higher saturation levels throughout the sample season (ranging from ~.25 to -1). The off-gassing of CO2, naturally warmer waters, and high pH all resulted in higher saturation levels and saturation capacity. Water from Jones spring was collected in a large pool after the water left the cave, but before it flowed over

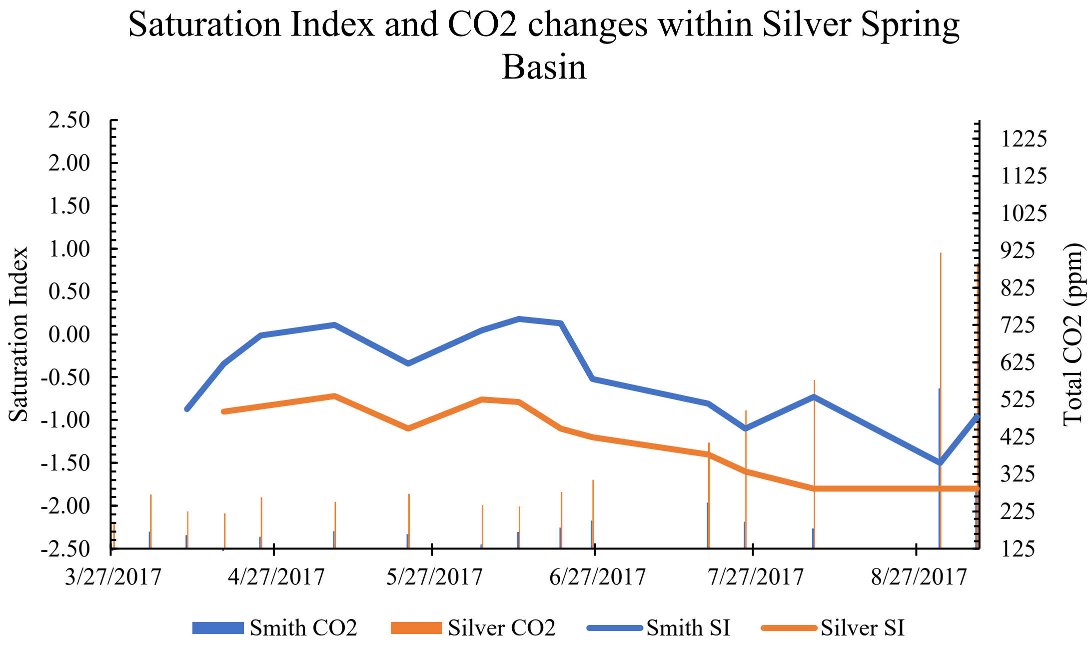


Figure 27: Saturation index compared to CO2 levels for Silver Spring Basin.

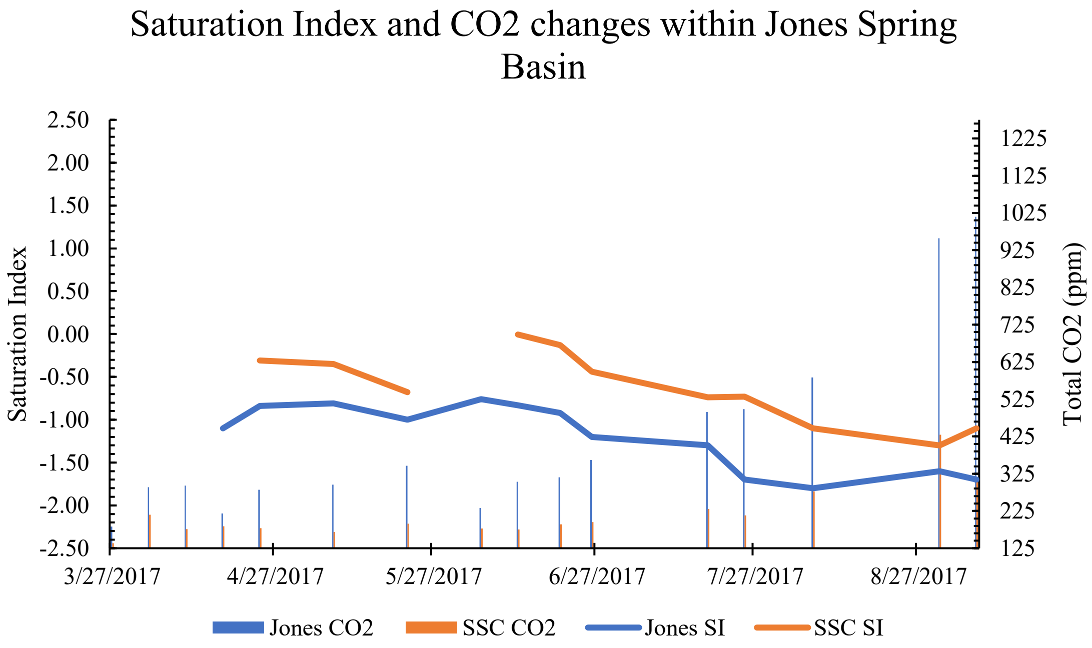


Figure 28: Saturation index compared to CO2 levels for Jones Spring Basin.

a 5 foot dam. This dam likely off gassed much of the CO2 as well, before the water continued downstream towards the SSC sample site.

Sequiota Cave Basin. Even though there were only four sampling periods for Sequiota cave basin, a clear difference between its two sample sites was apparent (Figure 29). The Southern Hills pond had much lower CO2 levels and much higher saturation levels (though still undersaturated) compared to Sequiota Cave. It was difficult to access the pond itself to collect water data, so samples were along an outlet channel ~100 feet from where the pond flows over a small dam. Like Jones spring, this likely off-gasses a lot of the CO2, leaving little to off-gas naturally due to contact with the air. Since this off-gassing did not result in saturated water then it is likely that, either, the spring water feeding the pond was undersaturated or the water sitting in the pond was slightly saturated when it emerged from the spring but precipitated out. Time spent mixing in the pond resulted in equilibrium of CO2 levels to the atmosphere and some calcite precipitated out, dropping the saturation index below 0. Total CO2 levels in water from Sequiota were very high (upwards of 950 ppm) and suggested that water traveling between the two sites collects CO2 rapidly. The low saturation index levels in Sequiota are likely the result of too little time in contact with the Burlington-Keokuk.

Denny's Spring Basin. Hospital pond had the lowest CO2 levels and the highest saturation indices, Greenway Trail spring and Denny's spring were closely related with respect to saturation and CO2 levels, and Young's Farm spring had the highest CO2 levels and lowest saturation indices (Figure 30). This trend of CO2 levels and saturation indices for springs and surface ponds is very similar to the other basins described above. However, through late June and July, the saturation indices for Hospital pond sank below

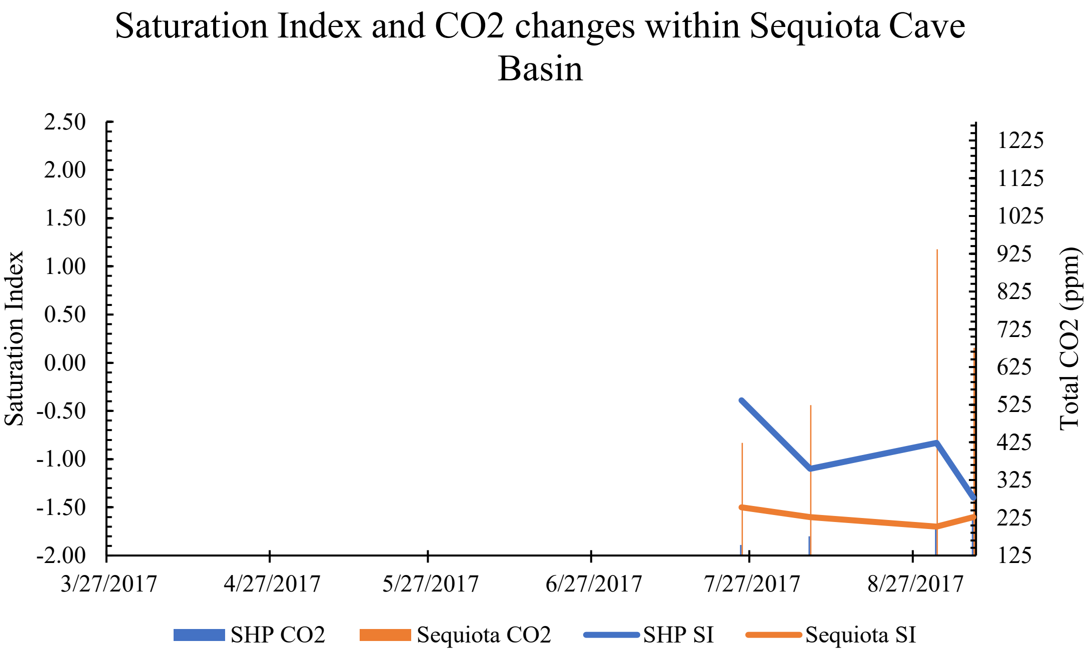

Saturation Index and CO2 changes within Sequiota Cave Basin

Figure 29: Saturation index compared to CO2 levels for Sequiota Cave Basin.

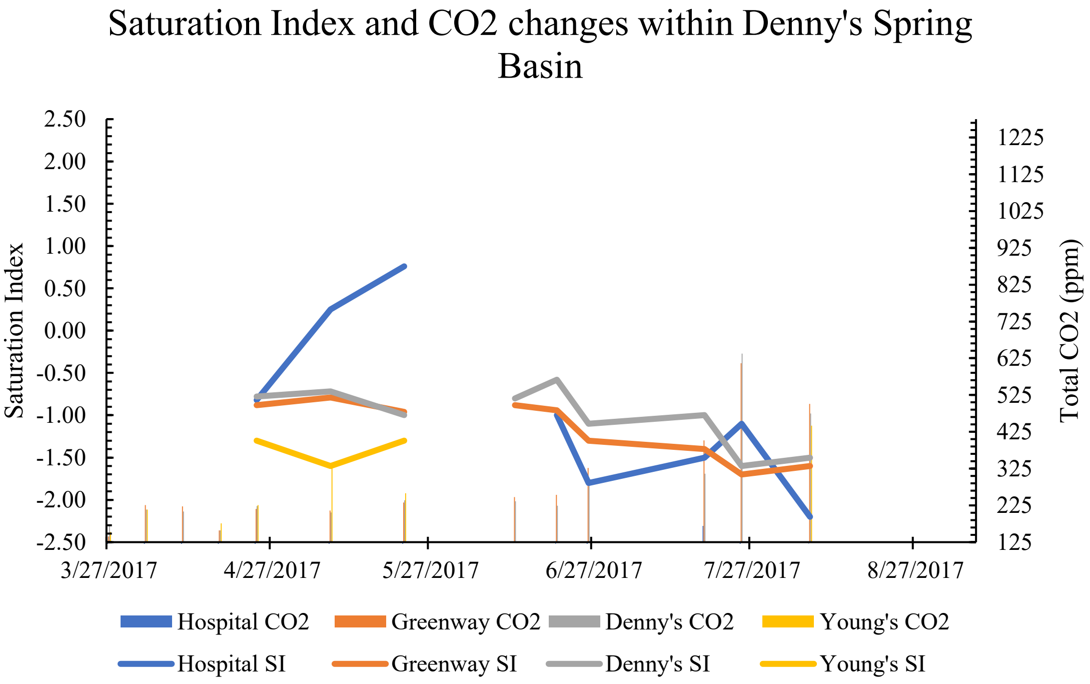

Saturation Index and CO2 changes within Denny's Spring Basin

Figure 30: Saturation index compared to CO2 levels for Denny's Spring Basin.

those of Greenway trail spring and Denny's spring (both springs). This occurred after the water level in the pond dropped several feet (between 6/5/2017 and 6/20/2017). Water was likely added to refill the pond in since it had no natural source of water feeding it. It was possible that water brought in was more saturated. Carbon dioxide levels for hospital pond were similar to CO2 levels in the pond from earlier in the sample season. After looking at other causes, it was apparent that very high temperatures, and lower pH values (drop of ~2 units) likely decreased saturation levels for the pond. The sudden drop in pH and CO2 levels was not apparent in the Southern Hills pond from the Sequiota Basin. Rising summer temperatures resulted in the warmer pond water and the sudden pH drop occurred shortly after pond levels were restored after dropping several feet (between 6/5/2017 and 6/20/2017).

Model Modification

With the success of the first model, the next step was to determine what variables were contributing significantly to the model. The model could be more efficient if fewer variables could be used to get the same or similar results. The FACTOR procedure in SAS 9.4 was used to run a principal component analysis using the first model's 87 samples. The eigenvalue of the correlation matrix table was used to determine the most significant variables (Table 4). According to the factor procedure, the first five variables (pH, temperature, conductivity, bicarbonate, and chloride) account for 95% of the variance.

Table 4: FACTOR procedure results. The cumulative column shows the first 6 variables (Eigenvalues) result in 97% of the variance in the first model.

The SAS System

The FACTOR Procedure
Initial Factor Method: Principal Components

Prior Communality Estimates: ONE

Eigenvalues of the Correlation Matrix: Total = 10 Average = 1

	Eigenvalue	Difference	Proportion	Cumulative
1	5.25843213	3.28858743	0.5258	0.5258
2	1.96984470	0.91051706	0.1970	0.7228
3	1.05932764	0.24549787	0.1059	0.8288
4	0.81382977	0.40248796	0.0814	0.9101
5	0.41134182	0.19467717	0.0411	0.9513
6	0.21666465	0.08801152	0.0217	0.9729
7	0.12865312	0.03763133	0.0129	0.9858
8	0.09102179	0.05026306	0.0091	0.9949
9	0.04075873	0.03063308	0.0041	0.9990
10	0.01012565		0.0010	1.0000

Model Two

The second model was almost identical to model one. To generate the model 87 samples (80%) from the original dataset were used. The same 25 samples that were not used in the model were used as blind samples to once again test the model's classification accuracy. The only difference was that five variables were used instead of the ten from the first model. Temperature, pH, conductivity, bicarbonate, and chloride were used to run the model since they were defined as contributing to 95% of the variance (see the factor analysis above). When the model was run, classification accuracy ranged from 76% to 100% with one basin scoring 41% (Silver Spring). While this was a decrease in classification accuracy, only two basins saw minor decreases, and it was decided to run the second model anyway. All 25 samples were run individually through the model and 19 samples (76%) were correctly classified (Table 5). The 76% classification accuracy was only one less sample than the first model (20/25 samples or 80% accuracy). The six misclassified basins had confidence intervals that were very small, and in four instances, were very close to the next closest classification. For example, sample ID 180 was classified as Jones basin with a confidence of .4645 (which was not the correct basin). The next closest basin it was classified into was Silver with a confidence of .4290 (which was the correct basin). The two remaining samples were vastly misclassified with confidences greater than .64 between the first and second classifications (Table 5). The 80% classification accuracy of model one and 76% classification accuracy of model two is very high. While the model was most successful when using all ten variables, it is interesting to note that very similar results could be found using only five variables.

Table 5: Blind sample placement results for model two with reduce variables used. Samples in red were misclassified.

ID #	Basin	Placed Basin	Confidence	Next Closest	Confidence	Actual Basin	Correct?
21	Doling	Doling	-	-	-	Doling	YES
96	Doling	Jones	0.6982	Silver	0.2549	Jones	YES
254	Doling	Denny’s	0.9400	-	-	Denny’s	YES
214	Doling	Jones	0.7609	Silver	0.2077	Jones	YES
112	Doling	Silver	0.6259	Jones	0.3375	Silver	YES
180	Doling	Jones	0.4645	Silver	0.429	Silver	NO
209	Doling	Jones	0.7947	Doling	0.1186	Jones	YES
49	Doling	Sequiota	0.3623	Jones	0.3088	Jones	NO
208	Doling	Silver	0.51	Jones	0.4727	Jones	NO
147	Doling	Silver	0.4755	Jones	0.3017	Silver	YES
187	Doling	Denny’s	0.955	-	-	Denny’s	YES
218	Doling	Doling	-	-	-	Doling	YES
225	Doling	Jones	0.5724	Silver	0.3939	Silver	NO
186	Doling	Doling	-	-	-	Doling	YES
26	Doling	Jones	0.5502	Silver	0.4375	Jones	YES
93	Doling	Doling	-	-	-	Silver	NO
99	Doling	Denny’s	0.9983	-	-	Denny’s	YES
80	Doling	Sequiota	0.9389	-	-	Sequiota	YES
100	Doling	Denny’s	0.5232	Silver	0.2348	Denny’s	YES
148	Doling	Doling	-	-	-	Doling	YES
171	Doling	Silver	0.7817	Sequiota	0.1427	Denny’s	NO
46	Doling	Doling	-	-	-	Doling	YES
213	Doling	Doling	-	-	-	Doling	YES
139	Doling	Denny’s	0.9984	-	-	Denny’s	YES
83	Doling	Jones	0.5019	Silver	0.3528	Jones	YES

Misclassified Samples

There were several basins that were constantly mistaken as the correct basin for some blind samples. The results from both models' blind sample analysis showed that samples were commonly misclassified between Jones basin, Sequiota basin, and Silver Spring basin. This was likely because of similarities between the discriminant functions for each basin. To test this idea, linear discriminant function values from each basin were used to create a regression equation for each basin. This plots the values as a single score on a number line, making it easier to visualize any overlap in the discriminant functions. While each basin had unique values that were used in the equation, the general formula was:

(constant) + 59.50(pH) + 10.06(temp) -.72(conductivity) + .36(bicarbonate) +
.73(chloride) -.19(sulfate) + 31.83(nitrate) + 1.29(calcite) – 2.95(magnesium) –
2.60(sodium) = regression score.

Where the constant was a value calculated by SAS.Linear
Scores were plotted for comparison (Figure 31).

Doling basin was unique, due to its range of scores being higher than any other basin. Values for Denny's overlapped with every basin but Doling. Despite this, only one blind sample was misclassified as Denny's between both models. Jones basin, Sequiota basin, and Silver basin all had values that overlapped as well. This was expected and confirms the idea that the discriminant functions for each basin have some similarities since the discriminant function analysis continued to misclassify blind samples between the three basins. To understand why Denny's basin only had one misclassified sample despite a wide range in scores that overlapped with every other basin and why Jones

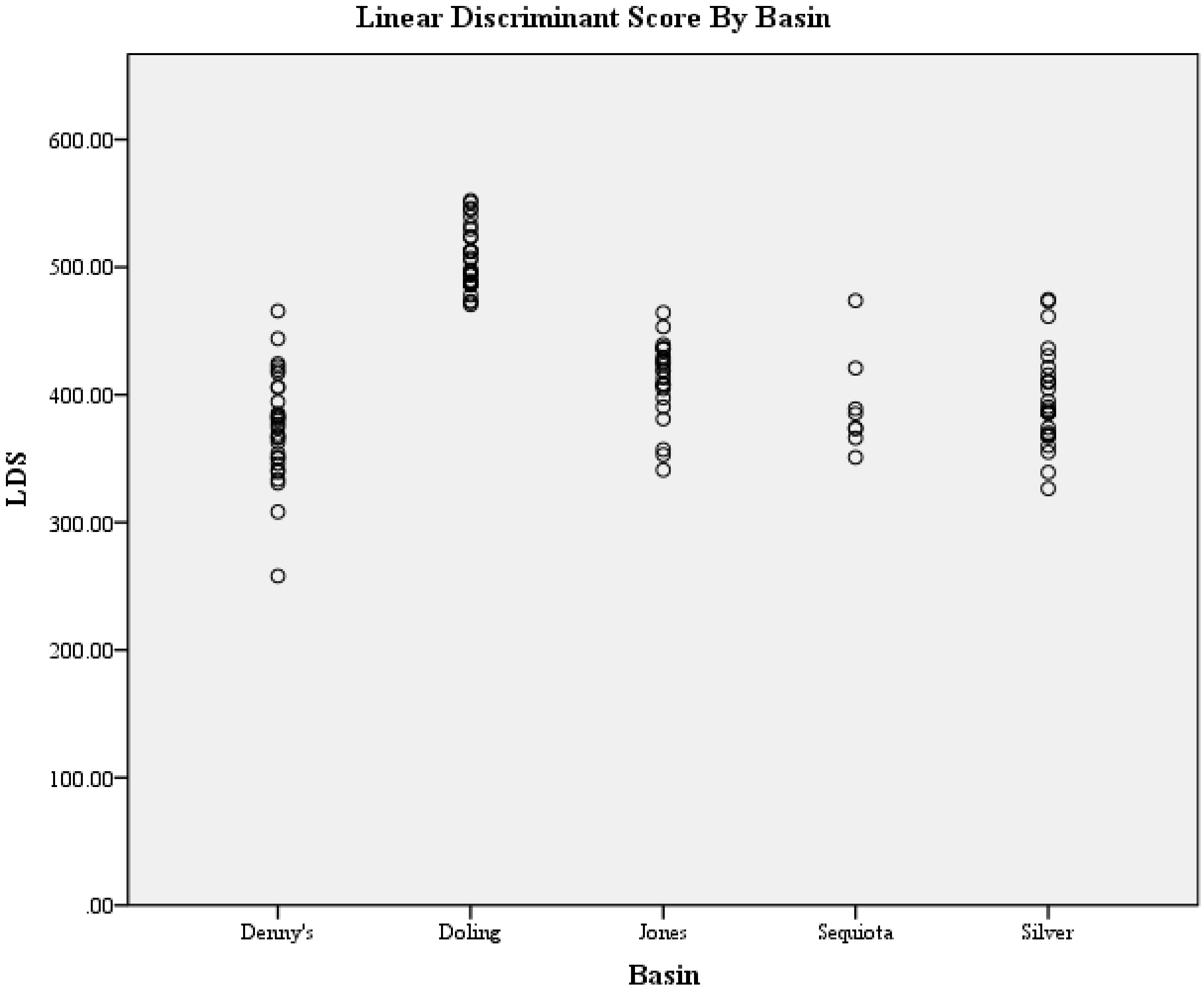


Figure 31: Linear discriminant plot for each basin calculated from the equation above.

basin, Sequiota basin, and Silver basin had many misclassified samples with overlapping scores, further analyses were run to test the significance of each variable individually.

Variable Analysis

The FACTOR procedure found that five of the ten variables used in the first model were the most significant to the discriminant function analysis. The second model showed that those five variables were significant enough to classify most samples into four of the five basins. However, one basin's classification accuracy dropped very low. To further understand the significance of each variable, two additional variable analyses were run. The first was a stepwise removal of one variable between each run. All ten variables were added, the model was rerun, then one model was subsequently removed before running the model again. Sodium was the first variable removed, followed by magnesium, then calcium, and so on. This was done until variables were removed, with pH the only remaining variable (since the discriminant function analysis needed at least one variable to run). The changes in classification accuracy in the number of observations and percent classified into basin table for each basin were noted and used to describe the significance each variable had on basin classification. Table 6 shows the results of this stepwise removal of variables. Basins that were not affected by the removal of a variable would have 0% change in their classification accuracy. For most basins, variables that were least useful included calcium, magnesium, sodium, and nitrate (0 to -18% change). Only for Silver spring were sodium, magnesium, and nitrate moderately useful. Variables that were most significant were bicarbonate, sulfate, chloride, and temperature (bicarbonate being the most useful).

Table 6: Results from the stepwise removal of variables from the discriminant function analysis.

	Stepwise Removal									
	pH	**Temperature**	**Conductivity**	**Bicarbonate**	**Chloride**	**Sulfate**	**Nitrate**	**Calcium**	**Magnesium**	**Sodium**
Denny's	0	-9.74	-6.76	-73.69	5.28	-5.27	-5.26	0	0	0.00
Doling	0	-80	-16	8.33	-8.33	0	0	0	0	0.00
Jones	0	-17.64	-5.89	-23.52	-11.76	-11.77	0	0	0	0.00
Sequiota	0	-9.53	0	-33.33	0	0	0	0	0	0.00
Silver	0	18.19	0	-4.55	0	-13.64	-13.63	0	-18.18	-4.55

The second analysis was similar to the stepwise removal described above, however only one variable was added at a time. This analysis was more effective in describing the significance each variable had on a basin's classification accuracy and it is described in greater detail here. The number of observations and percent classified into basin table showed how many observations (samples) the discriminant function was able to classify into the correct basin using only that one variable. Table 7 lists the effect each individual variable had on the classification accuracy for each basin. It can be inferred that higher classification accuracy caused by a variable or set variables likely indicates that the values for each variable were uniquely higher or lower than other basins. For example, if calcium accounted for 86% of the classification accuracy of Denny's basin, then the values at Denny's basin were likely unique (either higher or lower than other basins). Further analysis of the significance of variable for every basin is discussed below.

Denny's Basin. The variable with the largest effect on Denny's basin samples classification was chloride. It was found that chloride values for Denny's basin were consistently higher than the other basins. Conductivity, bicarbonate, and sulfate were moderately useful in the classification (30-45% accuracy). Conductivity values for Denny's basin fell between .14 and .67 mS/cm. These ranges were much lower than any other basin. Bicarbonate levels ranged from 50 to 200 ppm, were most like Silver Spring basin, Jones basin, and Sequiota basin, and much lower than Doling Park basin. Sulfate levels were much lower than Doling Park basin but very similar to the other three basins. pH, temperature, sulfate, nitrate, calcium, and magnesium were insignificant to the classification accuracy.

Table 7: Results from the individual addition of variables to the discriminant function analysis.

	Single Addition									
	pH	**Temperature**	**Conductivity**	**Bicarbonate**	**Chloride**	**Sulfate**	**Nitrate**	**Calcium**	**Magnesium**	**Sodium**
Denny's	4.55	22.73	45.00	40.00	65.22	30.43	20.00	5.00	0.00	56.56
Doling	4.00	84.00	100.00	87.50	24.00	100.00	96.00	96.00	68.00	60.00
Jones	17.65	11.76	17.65	23.53	70.59	17.65	17.65	29.41	0.00	64.71
Sequiota	57.14	33.33	0.00	0.00	28.57	57.14	0.00	42.86	0.00	14.29
Silver	54.55	4.55	31.82	13.64	13.64	72.73	50.00	4.55	54.55	4.55

Doling Park Basin. For the classification of Doling Park basin samples, temperature, conductivity, bicarbonate, sulfate, nitrate, calcium, magnesium, and sodium were the most significant variables for sample classification (68 to 100% accuracy). Temperatures at Doling Park basin were, on average, lower than other basins by ~1 degree. Conductivity, bicarbonate, sulfate, nitrate, calcium, and magnesium values were much higher than in the other basins with very little overlap in values. Sodium values were higher than most basins but overlapped some with Silver basin values. Chloride and pH were insignificant to the classification accuracy. The larger values likely had a greater impact on the discriminant function which resulted in a very high classification accuracy for the samples in Doling Park basin. Because there were many variables that could be used to accurately classify samples into the basin, only a few variables would be needed for Doling Park basin's discriminant function to be effective. This was unusual compared to the other basins analyzed.

Jones Basin. The two variables most significant to the classification of Jones basin samples were chloride and sodium (70.59% and 64.71% accuracy respectively). Values for chloride within Jones basin were lower compared to other basins. There were some chloride values that were similar to Silver and Doling Park basins. Sodium levels for Jones were also much lower than any basin except Silver Spring basin, where the average was only slightly lower (~3 mg/L). The remaining eight variables (pH, temperature, conductivity, bicarbonate, sulfate, nitrate, calcium, and magnesium) all had little impact on classification accuracy, ranging from 0 to 29.41%.

Sequiota Basin. Sulfate and pH were the largest contributors to the discriminant function for Sequiota basin, though they could only classify 57.14% of the samples.

While lower than the values from Doling Park basin, the sulfate values for Sequiota basin overlapped with values from Jones and Denny's basins. Variables calcium, chloride, and temperature contributed a moderate amount (33.33%, 28.57%, and 42.86% respectively) but values for these variables also overlapped with values from other basins, decreasing their significance overall to the classification for Sequiota basin. Conductivity, bicarbonate, nitrate, magnesium, and sodium were very insignificant to the classification accuracy (0-14.29%).

Silver Spring Basin. Silver spring, like Doling Park basin, had several variables that were significant to the classification of samples. Sulfate, pH, nitrate, and magnesium were the most significant variables, ranging from 50 to 72.73% accuracy (sulfate being the most significant at 72.73%). Sulfate values at Silver Spring basin were the second highest out of the other basins. Only Doling Park basin had larger values (and they were uniquely larger with no overlap of values between Doling Park and Silver Spring basins). The remaining variables: temperature, conductivity, bicarbonate, chloride, calcite, and sodium were not significant to the classification (ranging from 4.55% to 31.82%).

From the variable analyses, it was apparent that the most significant variables to the classification accuracy varied between basins. Significant variables had unique value ranges, where values for a basin that were larger, smaller, or did not overlap with values of another basin played a greater role in creating a discriminant function for a basin. The sources for these variables are dependent on the variables but could generally be described as affect by the soils, land use, and local geology. The intensity or range of the values was likely affected by the time and amount of water that was in contact with the soils, local geology, and the different land uses.

Land Use and Sample Site Location

While most of Springfield can be classified as urban, there were minute differences between the land use overlaying the five groundwater basins in this study. Minor differences in land use and sample site location could result in unique values for the chemical and physical variables used for the statistical analysis. As mentioned in the previous section, values for some variables were either higher or lower in some basins as compared to others, and an analysis of the land use surrounding those basins and site location would likely show what was influencing the values. The impact of land use and site location on the value ranges for each variable is discussed below.

pH. The pH of surface and groundwaters is most dependent on the total dissolved CO2 in the water. Surface waters generally have much lower CO2 levels since they are at or near equilibrium with the atmosphere which has very low CO2 levels. CO2 levels are most often higher in groundwater systems due to the downward movement of CO2 from the biological activity that produces CO2 in the soils above. When compared, Hospital pond from Denny's basin, Smith Park of Silver basin, and Snow Spring Complex from Jones basin all had the highest pH values. Inversely, these three sample sites also had the lowest total CO2 levels. The Baptist Bible College spring, Jones spring, Silver Spring and Doling Cave spring all had relatively low pH ranges and very high total CO2 levels.

Temperature. The surface stream and pond sample sites were always warmer than the spring sample sites. Hospital pond, Southern Hills pond, Smith Park, and Snow Spring Complex all had very high water temperatures that were likely affected by the air temperature. Each had a general warming trend that corresponded with the increase in average daily high air temperature for Springfield as the summer progressed. Much like the air temperatures in caves, the temperature of the spring waters remained relatively constant throughout the summer months regardless of changes in outside air temperature.

Conductivity. The conductivity of water is affected by the amount of ions present in solution (Drever, 1997). Sample sites with the highest conductivity values were Doling Cave spring and the Baptist Bible College spring (both springs and both a part of Doling Park Basin). Chloride, sulfate, nitrate, calcium, magnesium, and sodium values for Doling Park cave and the Baptist Bible College spring were some of the highest compared to other sample sites. These six ions influenced the high conductivity values for Doling basin. Sites with the lowest conductivity values were the Hospital pond, Smith Park, and the Southern Hills pond. All three of these sample sites were ponds or surface streams. Chloride, sulfate, nitrate, calcium, magnesium, and sodium values at these three sample sites were also very low. Individual analysis of each ion is described in greater detail in the following sections.

Bicarbonate. The formation of bicarbonate results from the combination of dissolved CO2 and water in contact with limestone. When the sample sites were sorted highest to lowest with respect to bicarbonate, values were reviewed for calcium, CO2, and pH as well. As bicarbonate values decreased, calcium and CO2 values did too. Bicarbonate levels were inversely related to pH levels and showed a moderate increase as bicarbonate levels decreased. This was expected since CO2 played a major role in bicarbonate values. Because of their high CO2 and calcium levels, Baptist Bible College spring, Jones Spring, and Doling cave all had very high bicarbonate values. Hospital pond and Smith park had the lowest bicarbonate values since their CO2 and calcite values were so low as well.

Chloride. The highest chloride levels were found at the Greenway Trail spring and at the Denny's Spring. The lowest values were found at the Snow Spring complex, the Hospital pond, and Silver spring. It is most likely that anthropogenic sources result in high chloride levels at the Greenway Trail spring and Denny's Spring. A common source of chloride contamination is road salt used to de-ice roads in the winter. Both sites are located along major roads that are heavily salted. This salt can continue to influence sodium levels in surface and ground water months after it was laid. The salt, mixed with rain or snow melt can runoff into the soils and become trapped. Trapped salt can have a slower dissolution, which can create elevated sodium levels for longer periods of time compared to systems where the salt is flushed through quickly. This likely caused elevated chloride levels at the Greenway Trail spring, and subsequently Denny's Spring downstream. Chloride levels were surprisingly low at Silver spring, where a chlorinated pool was located less than 200 feet from the spring.

Sulfate. Sulfate levels were the highest at Doling Park cave and the Baptist Bible College spring. Levels at those sites were 3x to 4x higher than any other sample site. While sulfates can come from household uses like detergents, the most common source is naturally from various minerals (World Health Organization, 2017). In southwest Missouri, lead and zinc are two of the most common minerals. Thomson (1986) listed several locations in Greene county where lead and zinc deposits were discovered in the Burlington-Keokuk limestone. One such location was in the same area as the two sample sites for Doling Cave basin and it is possible that some groundwater was interacting with the ore deposit, resulting in elevated sulfate levels.

Nitrate. Nitrate is commonly found in fertilizers and human waste from septic systems (World Health Organization, 2017). Higher nitrate levels should be expected at sites where farms or agriculture centers are nearby and in areas where septic systems are old or leaky. Of the 12 sample sites, the Baptist Bible College spring and Doling Park cave had the highest levels of nitrate. Doling basin is located on the north side of Springfield, which is one of the oldest parts of town. It could be suggested that the older homes and businesses had old septic systems that were leaking nitrates into the groundwater. Jones spring occasionally had elevated levels as well. These levels likely correlate with the increased use of fertilizers on a nearby farm as the weather began to warm and favor agricultural activities. The other sample sites had very low nitrate levels. The sites are in newer parts of Springfield (relative to the area surrounding Doling basin) and it is probable that the residences and businesses were connected to a sewer system immediately instead of first using septic systems.

Calcium. Calcium is naturally found in groundwater where the bedrock is comprised of limestone or other carbonates. Groundwater systems will usually have higher levels of calcium compared to surface water because of the increased contact with the bedrock. Doling Park cave and the Baptist Bible College spring had the highest levels of calcium while Hospital pond, Southern Hills pond, and Smith park had the lowest calcium levels. The Doling Park cave and Baptist Bible College springs were likely part of a larger groundwater network that spent more time underground, in contact with the Burlington-Keokuk compared to the other sample sites that were springs. The Hospital pond, Southern Hills pond, and Smith park sample sites are all surface waters and did not have much (if any) contact with the Burlington-Keokuk limestone. This would prevent calcium levels from increasing at those sites.

Magnesium. Magnesium is also a common cation found in groundwater that flows through carbonate bedrock. The alteration of limestone to dolomite occurs with the substitution of magnesium ions. Doling basin had the highest magnesium levels relatively, but no basin had levels that were significantly higher than another basin. This suggests that the groundwater flow through the basins stays within the Burlington-Keokuk limestone and does not travel through the closest dolomite unit (the Cotter dolomite).

Sodium. Sodium levels were highest at the Greenway Trail spring and Denny's Spring and moderately high at Doling Park cave and the Baptist Bible College spring. At the Snow Spring Complex and Silver Spring, sodium levels were the lowest. At the Greenway Trail, Denny's Spring, Doling Park cave, and Baptist Bible College sodium levels remained constant, with little variation, for the entire sample season. Sodium levels Snow Spring Complex and Silver Spring decreased during the sample season. The proximity of the Greenway Trail spring and Denny's Spring to major roadways that are heavily salted during the winter seasons is a probable source for the elevated sodium levels at those sites. Even though sampling occurred months after the last snowfall, the salt may have become trapped in the soil, resulting in a slower dissipation of the salt over time.

CHAPTER 6 – CONCLUSIONS

Review

Groundwater basins in Springfield, MO were chosen based on existing dye trace data. Five basins were chosen, each with 2-4 sample sites for each basin. Twelve sample sites were used to collect water quality data during a 5 ½ month period from March to September 2017. Ten water quality variables were analyzed and used to develop a database that was used for the statistical analysis of the groundwater basins. The discriminant function analysis in SAS 9.4 was used to build models that would be used to describe the fingerprint or signature of each basin. The discriminant function analysis was also used to test the uniqueness of each basin's fingerprint by predicting the membership probability of a blind sample to a groundwater basin.

Two models were created and tested. The first model was able to correctly predict the membership probability for 80% (20/25) of the samples. The second model was used to see how few variables could be used to achieve similar results to model one. A factor analysis was run to determine those variables most significant to the first model's results. Five variables (pH, temperature, conductivity, bicarbonate, and chloride) accounted for 95% of the variance for model 1 and were considered most significant. These variables were used to create the second model and test the blind samples. When the sample set of blind samples were run through model 2, it was found that 76% (19/25) of the sample's membership was correctly predicted. The membership probabilities for models 1 and 2 were higher than those from previous work by Gouzie (1986) and Lockwood and Gouzie

(2015). The concept that groundwater basins have unique signatures that can be used to differentiate them via statistical models is shown to work in this thesis.

Limitations and Future Work

There were several limitations that could be improved upon from this study. Near the end of the sampling season, Springfield experienced very dry conditions and many springs and streams began to run very low (one spring dried up). Sampling during a wetter season (or areas where rainfall is higher or more consistent) may yield different results. The length of the sampling season was only six months. While six months was enough to get an 80% classification accuracy, a longer sampling season may increase the range of values for a groundwater basin, improving its unique signature; one year may be enough to do this.

Overall, the total number of basins was very small. While the number of samples for each basin (~20) seemed to be the right amount, an increased number of basins for the study area could improve the robustness of the classification results. If the percent classified for blind samples is the same (or better) with six or more basins in a study site as compared to five, then it could be said that the model is robust.

While it would be difficult to do in Springfield, finding basins with more than two sample sites could produce different results. This study had a basin with four sample sites, and its classification accuracy was relatively high. Developing a similar study with an even mix of basins made up of two sample sites and basins made up of four or more sample sites could help determine if increasing the number of sample sites per basin is needed. This study found that five variables could be used to reach classification accuracies upwards of 75%. Though it was noted that different variables had varying significances for each basin and different areas may have different variables that are most significant.

REFERENCES

Berglund, J.L., 2012, An applied karst study of the Ward Branch watershed near the James River freeway/South Campbell interchange [Masters thesis]: Springfield, Missouri State University.

Bullard, L., Thomson, K.C., and Vandike, J.E., 2001, The Springs of Greene County Missouri, Missouri Department of Natural Resources, Geological Survey and Resource Assessment Divison, Water resources Report No. 68, p. 120.

Drever, J.I., 1997, The geochemistry of Natural Waters 3rd edition, Prentice-Hall, New Jersey.

Eaton, A.D., Clesceri, L.S., Rice, E.W., and Greenberg, A.E., 2005, Standard Methods for the examination of water & wastewater, 21st edition, American Public Health Association, Washington, DC.

Gouzie, D.R., 1986, Carbonate geochemistry and chemical discrimination of ground water basins in portions of the Inner Bluegrass karst region, Kentucky [Ph.D. thesis]: Lexington, University of Kentucky.

Hach, 2013, Digital Titrator Manual Method 8203, Hach Company, p. 45-51.

Imes, J.I., 1989, Analysis of the effect of pumping on ground-water flow in the Springfield plateau and Ozark Aquifers near Springfield, Missouri: U.S. Geological Survey Water Resources Investigations Report 89-4079: https://pubs.usgs.gov/wri/1989/4079/report.pdf, p. 1, 12-13.

Lenntech, 2018, Langelier Saturation Index Calculator: This calculator helps you determine the scaling potential of the water by using the Langelier Saturation Index:https://www.lenntech.com/calculators/langelier/index/langelier.htm#ixzz5EPGSo8KIhttps://www.lenntech.com/calculators/langelier/index/langelier.htm (accessed December 2017).

Lockwood, B.E., and Gouzie, D.R., 2015, Statistical analysis of geochemical markers in karst groundwater of Central Kentucky, Geological society of America Abstracts with Programs, v.48, no. 5

Manahan, S.E., 2010, Environmental Chemistry, 9th edition: Florida, Taylor and Francis Group, p. 51, 687-688.

Missouri Department of Natural Resources, Springfield plateau groundwater province: http://dnr.mo.gov/geology/wrc/groundwater/education/provinces/springfieldplatprovince.htm (accessed November 26th, 2016).

Missouri Department of Natural Resources, Missouri Geological Survey Geosciences Technical Resource Assessment Tool (GeoSTRAT): http://dnr.mo.gov/geostrat/ (accessed September 2017).

Missouri Spatial Data Information Service, 2018, MSDIS Data List with Descriptions: http://www.msdis.missouri.edu/data/datalist.html#lg (accessed August – December, 2017).

Palmer, A.N., 2007, Cave Geology, Dayton Ohio, CAVE BOOKS.

Peel, M.C., Finlayson, B.L., and McMahon, T.A., 2007, Updated world map of the Köppen-Geiger climate classification, Hydrology and Earth System Sciences v. 11, p. 1633-1644.

SAS Institute Inc., 2017, SAS/STAT 14.3 User's Guide, SAS Institute Inc., Cary, NC, p. 2346-2435.

Stanke, B.D., 2010, Time-Series Relationships in an Urabn Karst Watershed in Springfield, Missouri [Masters thesis]: Springfield, Missouri State University.

Tabachnick, B.G., and Fidell, L.S., 1996, Using Multivariate Statistics, HarperCollins College Div, California State University, Northridge, 3rd edition, p. 507-574.

Thomson, K.C., 1986, Geology of Greene County: Springfield, MO, Watershed Management Coordinating Committee, Springfield, Missouri, p. 87.

Thrailkill, J., Spangler, L.E., Hopper Jr, W.M., McCann, M.R., Troester, J.W., and Gouzie, D.R., 1982, Groundwater in the Inner Bluegrass Karst Region, Kentucky, University of Kentucky Water Resources Research Institute, Research Report No. 136.

Tomlin, K.M., 2010, Quantitative Dye-Tracing Investigation of Groundwater Flow in the Ward Branch Watershed, Springfield Missouri [Masters thesis]: Springfield, Missouri State University.

Vineyard, J.D., Feder, G.L., 1982, Springs of Missouri: Missouri Department of Natural Resources, Division of Geology and Land Survey, p. 173.

US Census Bureau, 2015, Quick Facts Springfield city, Missouri: http://www.census.gov/quickfacts/table/PST045215/2970000 (accessed November 2016).

Weather Underground, 2018, Weather History for KSGF – March to September 2017: https://www.wunderground.com/history/airport/KSGF/2017/3/20/CustomHistory.html?dayend=28&monthend=8&yearend=2017&req_city=&req_state=&req_statename=&reqdb.zip=&reqdb.magic=&reqdb.wmo (accessed September 2017).

White, W.B., 2002, Karst hydrology: recent developments and open questions: Engineering Geology, v. 65, p. 85-105.

World Health Organization, 2017, Guidelines for Drinking-water Quality Fourth Edition incorporating the first addendum, Creative Commons Attribution-NonCommercialShareAlike 3.0 IGO licence (CC BY-NC-SA 3.0 IGO).

APPENDICES

Appendix A - Data

Date	Location	pH	Temp	Conduct	Bicarb (CaCO3)	Flow	Chloride	Sulfate	Nitrate	Ca	Mg	Na
			(°c)	(mS/m)	(mg/l)	(cfs)	(mg/L)	(mg/L)	(mg/L)	(mg/L)	(mg/L)	(mg/L)
3/27/2017	Doling Cave	6.76	15.1	61.5	217	1.80	27.459	33.218	3.363	99.60	3.666	14.80
	BB College	6.62	15.5	70.9	256	0.41	-	-	-	-	-	-
	Smith Park	7.51	13.6	28.6	98	9.58	-	-	-	-	-	-
	Silver Spring	6.75	16.3	39.9	160	12.94	-	-	-	-	-	-
	Jones Spring	6.74	14.8	37.5	147	13.42	-	-	-	-	-	-
	Snow Complex	7.36	13.9	36	143	86.25	-	-	-	-	-	-
	Hospital Pond	7.28	14.9	21.7	53	-	-	-	-	-	-	-
	Greenway Trail	6.98	15	42	129	2.40	-	-	-	-	-	-
	Denny’s Spring	6.9	14.7	42.8	152	28.75	-	-	-	-	-	-
	Young’s Farm Spring	6.83	15.7	39.3	130	1.80	-	-	-	-	-	-
	Lark Spring	-	-	-	-	-	-	-	-	-	-	-
4/3/2017	Doling Cave	6.87	15.1	71.7	264	2.88	37.840	41.009	3.253	114.8	3.968	18.6
	BB College	6.57	15.9	71.0	272	0.16	28.493	43.943	3.360	118.70	3.480	19.20
	Smith Park	7.94	15.5	54.9	190	2.99	-	-	-	-	-	-
	Silver Spring	6.69	16.5	51.6	210	7.67	-	-	-	-	-	-
	Jones Spring	6.82	15.4	62	244	9.58	-	-	-	-	-	-
	Snow Complex	7.51	14.4	46.1	229	-	-	-	-	-	-	-
	Hospital Pond	7.74	16.0	35.2	106	-	-	-	-	-	-	-
	Greenway Trail	6.84	15.2	66.8	193	0.96	-	-	-	-	-	-
	Denny’s Spring	7.08	15.2	63.3	204	4.75	-	-	-	-	-	-
	Young’s Farm Spring	6.98	15.8	53.2	196	0.21	-	-	-	-	-	-
	Lark Spring	-	-	-	-	-	-	-	-	-	-	-

Date	Location	pH	Temp	Conduct	Bicarbonate (CaCO3)	Flow	Chloride	Sulfate	Nitrate	Ca	Mg	Na
			(°c)	(mS/m)	(mg/l)	(cfs)	(mg/L)	(mg/L)	(mg/L)	(mg/L)	(mg/L)	(mg/L)
4/10/2017	Doling Cave	7.02	15.2	71.0	276	2.40	36.107	41.742	3.341	105.00	3.924	18.60
	BB College	6.75	15.5	70.2	275	0.24	28.155	43.856	3.382	118.70	3.438	19.80
	Smith Park	8.15	18.9	51.9	182	2.00	41.263	21.158	0.300	68.30	2.996	24.20
	Silver Spring	7.03	16.7	51.5	211	7.67	-	-	-	-	-	-
	Jones Spring	6.93	15.4	61.1	263	7.67	-	-	-	-	-	-
	Snow Complex	7.84	16.3	43.1	195	83.06	-	-	-	-	-	-
	Hospital Pond	8.87	20.0	25.1	86	-	-	-	-	-	-	-
	Greenway Trail	6.98	15.8	64.9	204	1.92	-	-	-	-	-	-
	Denny’s Spring	7.17	15.7	61.7	205	5.59	-	-	-	-	-	-
	Young’s Farm Spring	-	-	-	-	-	-	-	-	-	-	-
	Lark Spring	-	-	-	-	-	-	-	-	-	-	-
4/17/2017	Doling Cave	6.98	15.4	64.6	164	2.00	32.347	35.261	2.759	110.80	3.854	16.80
	BB College	6.69	15.5	70.9	235	0.12	30.014	44.775	3.178	123.80	3.358	18.80
	Smith Park	7.81	18.1	30.8	129	4.19	21.918	7.051	0.213	37.80	1.858	13.00
	Silver Spring	6.85	16.6	45.0	189	4.11	19.291	8.421	1.722	79.10	2.354	9.40
	Jones Spring	6.78	15.7	40.7	180	10.22	21.669	5.838	1.626	68.50	1.624	11.20
	Snow Complex	7.76	16.4	43.0	202	62.29	-	-	-	-	-	-
	Hospital Pond	7.22	21.0	15.7	77	-	-	-	-	-	-	-
	Greenway Trail	6.9	16.7	41.8	139	1.60	-	-	-	-	-	-
	Denny’s Spring	6.98	16.2	38.6	144	14.38	-	-	-	-	-	-
	Young’s Farm Spring	6.82	16.1	39.8	149	1.26	-	-	-	-	-	-
	Lark Spring	-	-	-	-	-	-	-	-	-	-	-

Date	Location	pH	Temp	Conduct	Bicarbonate (CaCO3)	Flow	Chloride	Sulfate	Nitrate	Ca	Mg	Na
			(°c)	(mS/m)	(mg/l)	(cfs)	(mg/L)	(mg/L)	(mg/L)	(mg/L)	(mg/L)	(mg/L)
4/24/2017	Doling Cave	6.76	15.4	69.2	267	3.99	30.139	41.527	3.325	123.10	3.874	16.60
	BB College	6.57	15.3	68.1	278	0.40	24.061	38.857	3.474	123.60	5.264	17.20
	Smith Park	7.88	15.6	47.1	174	2.24	32.463	19.160	0.660	66.30	3.050	19.80
	Silver Spring	6.79	16.8	49.4	218	5.75	19.114	12.337	2.167	92.50	2.398	9.60
	Jones Spring	6.76	15.4	56.2	230	15.33	25.805	9.221	2.695	104.60	3.430	12.80
	Snow Complex	7.47	15.6	41.3	189	56.22	10.931	7.039	2.108	75.10	3.124	6.80
	Hospital Pond	7.36	19.1	31.4	106	-	25.141	6.383	1.132	41.10	2.526	13.00
	Greenway Trail	6.89	15.9	58.9	189	1.20	63.427	10.137	2.571	86.90	2.846	28.60
	Denny’s Spring	6.95	15.7	55.6	202	6.39	45.593	10.700	3.164	87.60	2.268	20.00
	Young’s Farm Spring	6.77	16.1	51.6	185	0.77	42.157	9.653	2.354	46.90	3.642	16.60
	Lark Spring	-	-	-	-	-	-	-	-	-	-	-
5/8/2017	Doling Cave	6.81	15.4	68.2	261	2.99	23.245	35.366	2.723	124.00	3.908	17.20
	BB College	6.69	15.3	64.9	266	0.15	17.674	31.562	2.855	120.60	3.168	15.80
	Smith Park	7.86	19.6	48.5	190	4.60	26.186	18.789	0.685	70.20	3.070	18.20
	Silver Spring	6.89	16.9	52.1	220	0.55	18.017	12.823	2.051	98.00	2.712	11.00
	Jones Spring	6.76	15.3	56.0	241	14.38	19.734	8.688	2.362	108.20	2.852	12.00
	Snow Complex	7.48	16.3	39.5	178	172.50	8.468	5.791	1.520	68.10	3.124	6.20
	Hospital Pond	8.27	25.1	38.0	113	-	30.243	8.144	0.979	44.20	2.858	16.60
	Greenway Trail	6.94	16.4	63.7	190	5.75	60.316	9.628	2.243	95.10	3.314	30.40
	Denny’s Spring	7	16.1	58.6	191	-	42.756	9.558	2.646	94.10	2.934	21.00
	Young’s Farm Spring	6.25	16.2	50.2	170	0.45	34.087	8.773	2.092	80.60	2.588	17.00
	Lark Spring	-	-	-	-	-	-	-	-	-	-	-

Date	Location	pH	Temp	Conduct	Bicarbonate (CaCO3)	Flow	Chloride	Sulfate	Nitrate	Ca	Mg	Na
			(°c)	(mS/m)	(mg/l)	(cfs)	(mg/L)	(mg/L)	(mg/L)	(mg/L)	(mg/L)	(mg/L)
5/22/2017	Doling Cave	6.57	15.5	66.4	266	2.00	22.320	33.216	2.492	120.60	4.728	17.00
	BB College	6.49	15.3	65.5	266	0.18	19.049	31.733	2.731	118.40	3.668	16.40
	Smith Park	7.43	18.2	41.3	172	6.47	19.733	10.397	0.448	74.90	3.076	13.60
	Silver Spring	6.59	16.6	48.0	195	2.52	16.044	10.891	1.631	88.10	2.354	9.80
	Jones Spring	6.57	15.4	53.2	244	15.33	17.748	8.218	2.151	102.00	2.360	10.60
	Snow Complex	7.14	15.6	37.8	186	158.13	7.981	5.833	1.436	68.20	3.938	6.00
	Hospital Pond	9.09	24.5	18.4	69	-	10.110	3.979	0.505	29.50	1.836	6.80
	Greenway Trail	6.78	16.9	56.5	192	1.28	45.787	8.435	2.017	85.40	3.504	24.60
	Denny’s Spring	6.71	16.7	53.2	188	-	32.533	8.579	2.434	88.70	2.974	17.40
	Young’s Farm Spring	6.51	16.6	49.1	172	0.60	30.317	8.247	1.924	83.90	3.080	15.40
	Lark Spring	6.98	16.9	57.6	205	0.13	38.073	9.416	1.600	92.00	2.024	19.60
6/5/2017	Doling Cave	6.82	15.6	67.5	275	2.40	27.466	41.102	3.129	108.00	4.840	18.32
	BB College	6.90	15.3	65.4	262	0.16	22.703	38.697	3.416	107.00	3.880	19.98
	Smith Park	7.92	22.5	36.8	151	5.75	23.443	7.523	0.350	52.00	2.740	14.96
	Silver Spring	6.92	16.7	49.4	216	-	20.935	11.310	2.021	82.80	3.320	11.20
	Jones Spring	6.95	16.0	51.4	211	14.38	23.671	8.758	2.368	84.20	2.960	13.90
	Snow Complex	7.81	17.6	43.6	196	43.70	-	-	-	-	-	-
	Hospital Pond	9.85	27.6	14.4	-	-	-	-	-	-	-	-
	Greenway Trail	-	-	-	-	-	-	-	-	-	-	-
	Denny’s Spring	-	-	-	-	-	-	-	-	-	-	-
	Young’s Farm Spring	-	-	-	-	-	-	-	-	-	-	-
	Lark Spring	-	-	-	-	-	-	-	-	-	-	-

Date	Location	pH	Temp	Conduct	Bicarbonate (CaCO3)	Flow	Chloride	Sulfate	Nitrate	Ca	Mg	Na
			(°c)	(mS/m)	(mg/l)	(cfs)	(mg/L)	(mg/L)	(mg/L)	(mg/L)	(mg/L)	(mg/L)
6/12/2017	Doling Cave	6.90	15.7	69.5	262	2.00	32.884	40.669	3.238	109.60	4.980	19.78
	BB College	6.70	15.3	66.6	271	0.12	24.806	39.566	3.341	108.80	4.240	20.24
	Smith Park	7.80	26.3	50.5	187	0.30	36.894	11.474	0.636	70.40	3.960	19.68
	Silver Spring	6.90	16.6	52.2	211	-	25.536	11.096	2.241	86.40	3.480	11.42
	Jones Spring	6.77	15.6	60.8	249	13.42	31.897	10.311	2.908	100.40	3.380	16.34
	Snow Complex	7.71	19.6	44.7	191	43.13	16.161	7.658	1.834	72.80	4.880	8.52
	Hospital Pond	-	-	-	-	-	-	-	-	-	-	-
	Greenway Trail	6.80	18.1	66.3	207	0.96	72.043	11.673	2.569	90.80	4.240	32.52
	Denny’s Spring	6.88	17.5	62.4	207	9.58	53.123	11.286	3.122	90.20	3.960	24.00
	Young’s Farm Spring	-	-	-	-	-	-	-	-	-	-	-
	Lark Spring	6.93	19.2	63.8	225	-	59.459	11.051	1.141	93.40	2.820	26.08
6/20/2017	Doling Cave	6.46	15.9	70.1	270	-	34.682	40.180	3.101	112.80	4.980	20.16
	BB College	6.55	15.3	67.2	270	0.12	26.901	40.737	3.259	111.80	4.520	18.98
	Smith Park	7.74	25.3	50.6	199	0.30	36.425	10.337	0.516	70.00	3.860	20.10
	Silver Spring	6.62	16.7	51.4	204	-	25.135	10.659	2.049	84.40	3.460	11.70
	Jones Spring	6.69	15.6	59.7	245	13.42	33.525	10.371	2.811	99.60	3.380	16.52
	Snow Complex	7.56	19.3	46.1	203	31.05	18.314	7.839	1.745	73.80	5.240	9.64
	Hospital Pond	7.13	27.5	24.6	92	-	19.540	5.460	0.043	31.80	2.520	10.36
	Greenway Trail	6.75	18.4	63.3	205	1.92	68.785	10.714	2.240	87.20	4.020	30.90
	Denny’s Spring	7.09	17.8	62.5	215	2.56	54.285	11.150	2.749	89.20	3.980	25.44
	Young’s Farm Spring	-	-	-	-	-	-	-	-	-	-	-

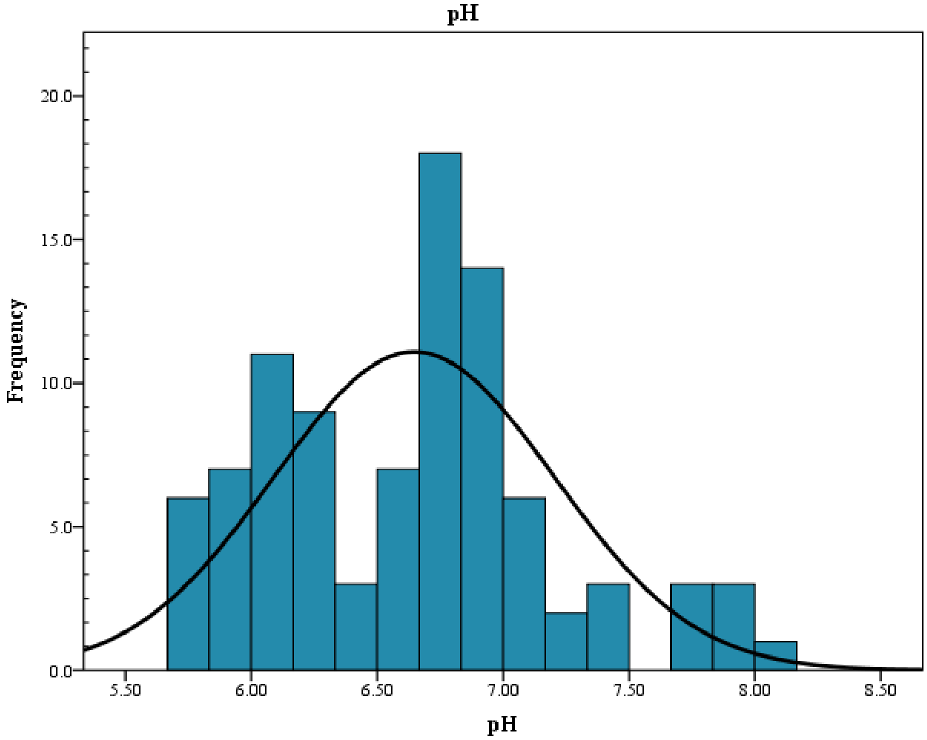
Date	Location	pH	Temp	Conduct	Bicarbonate (CaCO3)	Flow	Chloride	Sulfate	Nitrate	Ca	Mg	Na
			(°c)	(mS/m)	(mg/l)	(cfs)	(mg/L)	(mg/L)	(mg/L)	(mg/L)	(mg/L)	(mg/L)
6/26/2017	Doling Cave	6.4	15.9	69.9	269	-	33.420	43.509	3.187	113.20	5.300	19.42
	BB College	6.32	15.3	67.8	275	0.06	26.989	40.624	3.311	111.00	4.260	18.64
	Smith Park	7.15	20.7	50.2	197	0.27	33.750	10.907	0.618	77.00	3.760	18.86
	Silver Spring	6.52	16.6	50	209	1.92	22.824	10.170	1.920	84.20	3.500	13.80
	Jones Spring	6.46	15.7	55.4	230	15.33	27.946	9.296	2.411	91.40	3.140	14.48
	Snow Complex	7.29	17.7	45.3	199	31.05	17.073	7.845	1.922	74.00	4.860	9.14
	Hospital Pond	6.71	23.9	15.65	63	-	8.191	2.859	0.373	22.00	1.560	4.92
	Greenway Trail	6.4	18.4	59.9	195	1.44	61.609	10.108	2.217	82.00	3.880	32.26
	Denny's Spring	6.58	17.7	58.5	207	3.07	51.303	10.848	2.583	86.40	3.840	25.68
	Young's Farm Spring	-	-	-	-	-	-	-	-	-	-	-
7/18/2017	Doling Cave	6.19	16.4	70.3	269	-	36.799	41.369	2.996	111.00	4.52	20.34
	BB College	6.14	15.4	68.2	-	0.08	28.304	41.964	3.146	108.80	3.70	18.86
	Smith Park	6.79	24.3	52.4	207	0.21	37.706	10.499	0.332	73.20	3.24	18.40
	Silver Spring	6.31	16.8	53.5	220	0.51	27.865	11.177	2.023	85.20	2.96	11.26
	Jones Spring	6.27	15.6	64.1	251	8.63	39.051	11.675	3.187	105.20	2.88	17.30
	Snow Complex	6.93	20.8	48.6	207	16.43	21.730	8.656	1.818	73.20	4.98	10.40
	Hospital Pond	6.49	28.8	30.4	111	-	22.621	4.772	0.046	37.80	2.08	10.98
	Greenway Trail	6.31	18.8	68.8	216	1.92	79.274	10.953	2.456	87.80	3.66	32.24
	Denny's Spring	6.6	18.5	64.7	225	0.96	58.222	10.938	2.733	88.20	3.66	25.52
	Young's Farm Spring	-	-	-	-	-	-	-	-	-	-	-

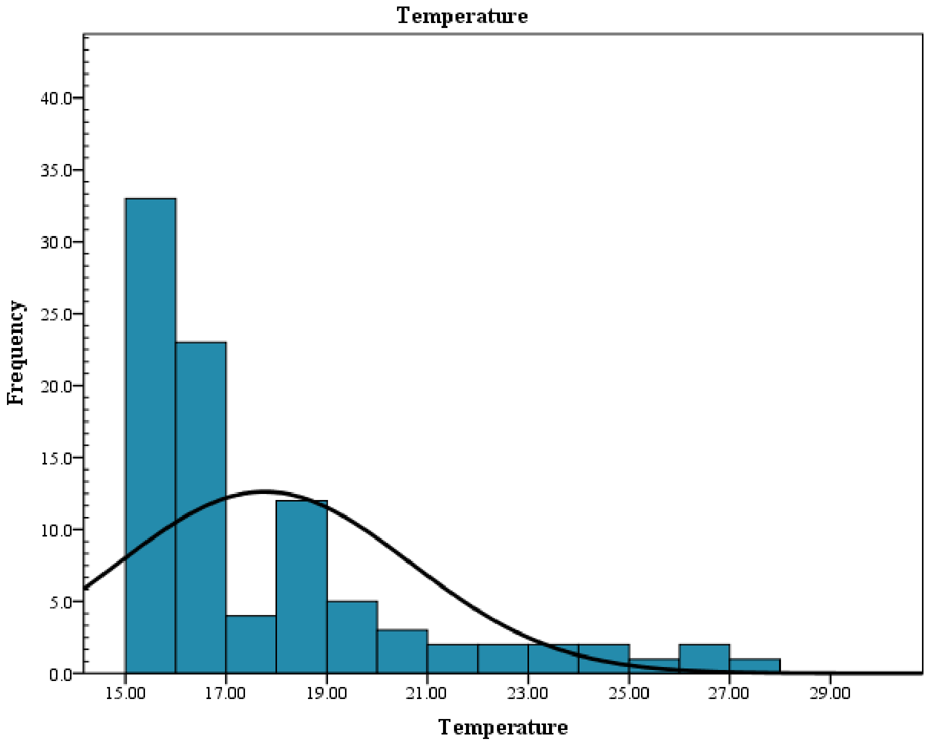
Date	**Location**	**pH**	**Temp**	**Conduct**	**Bicarbonate (CaCO3)**	**Flow**	**Chloride**	**Sulfate**	**Nitrate**	**Ca**	**Mg**	**Na**
			(°c)	**(mS/m)**	**(mg/l)**	**(cfs)**	**(mg/L)**	**(mg/L)**	**(mg/L)**	**(mg/L)**	**(mg/L)**	**(mg/L)**
7/25/2017	Doling Cave	6.15	16.8	70.9	274	2.57	36.683	40.744	2.812	111.80	4.36	20.16
	BB College	6.08	15.4	68.9	276	0.07	29.304	43.318	3.156	109.40	3.58	19.34
	Smith Park	6.61	26.4	61.2	144	0.36	25.263	124.142	0.495	85.60	3.56	13.42
	Silver Spring	6.10	16.9	49.1	201	0.36	24.733	9.683	1.961	79.00	2.76	10.56
	Jones Spring	6.09	16.1	49.2	199	5.75	25.928	8.978	2.632	78.40	2.24	11.52
	Snow Complex	6.95	21.9	46.9	193	19.17	21.248	8.283	1.942	71.00	4.78	9.88
	Southern Hills Pond	7.30	31.2	39.1	156	1.15	28.923	4.436	0.230	52.20	2.64	13.66
	Sequiota Park	6.23	18.4	53.8	206	12.78	35.624	8.182	2.377	79.20	3.70	15.64
	Hospital Pond	6.89	33.5	29.2	100	-	24.871	4.299	0.025	36.40	2.38	12.20
	Greenway Trail	5.99	19.0	68.7	209	0.85	78.245	10.133	2.359	88.20	3.64	32.86
	Denny’s Spring	6.01	18.3	67.2	225	1.92	62.481	11.429	2.896	94.40	3.48	26.08
	Young’s Farm Spring	-	-	-	-	-	-	-	-	-	-	-
8/7/2017	Doling Cave	6.02	16.4	68.4	265	3.59	28.8335	44.707	2.7712	127.20	5.50	17.68
	BB College	6.00	15.6	68.0	251	0.14	27.76	42.13	3.23	128.80	4.24	19.82
	Smith Park	7.03	21.5	41.1	168	2.00	20.92	17.09	0.51	65.40	2.78	13.18
	Silver Spring	5.95	16.9	42.5	185	0.81	14.65	9.74	1.62	73.20	2.84	8.30
	Jones Spring	5.97	16.2	47.8	193	7.19	20.61	7.69	2.25	83.20	2.54	11.54
	Snow Complex	6.59	18.7	44.5	199	56.35	14.26	7.94	2.46	75.40	4.22	8.30
	Southern Hills Pond	6.79	23.3	34.1	146	2.88	21.77	4.77	0.37	51.80	2.50	10.74
	Sequiota Park	6.09	18.5	52.0	209	30.49	26.27	7.49	2.30	86.60	3.50	14.80
	Hospital Pond	6.27	22.3	17.5	62	-	9.00	3.79	0.06	24.60	1.34	5.34
	Greenway Trail	6.08	19.5	57.9	197	7.99	54.32	10.38	2.06	82.40	3.44	28.54
	Denny’s Spring	6.14	18.6	56.2	204	7.67	41.33	10.46	2.51	85.60	3.66	21.42
	Young’s Farm Spring	6.10	19.2	50.4	179	0.36	37.23	10.32	2.38	82.00	2.68	16.66

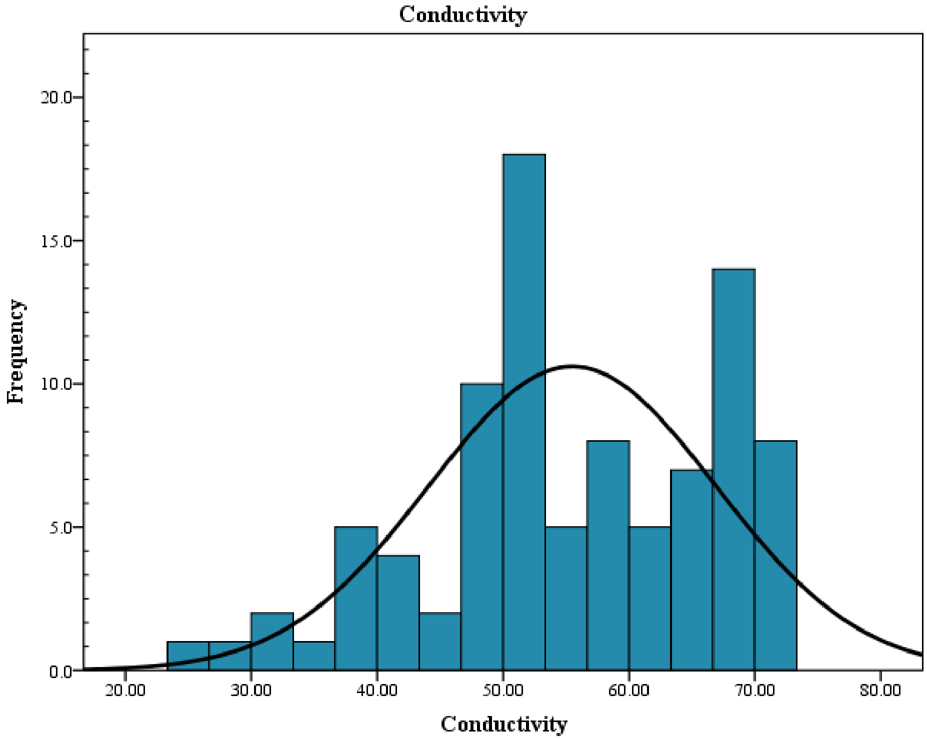
Date	Location	pH	Temp	Conduct	Bicarbonate (CaCO3)	Flow	Chloride	Sulfate	Nitrate	Ca	Mg	Na
			(°c)	(mS/m)	(mg/l)	(cfs)	(mg/L)	(mg/L)	(mg/L)	(mg/L)	(mg/L)	(mg/L)
8/31/2017	Doling Cave	5.78	16.4	70.1	264	2.88	33.8385	41.592	3.0082	128.20	4.88	19.32
	BB College	5.72	15.8	69.3	276	0.07	26.335	41.986	3.300	123.40	4.14	18.90
	Smith Park	6.10	20.6	53.2	225	0.45	33.497	10.969	0.738	86.60	3.66	18.72
	Silver Spring	5.80	16.8	52.4	227	0.81	23.091	10.523	2.231	92.60	3.18	11.04
	Jones Spring	5.89	15.9	61.7	277	7.19	33.878	10.891	3.013	111.40	3.16	15.96
	Snow Complex	6.32	18.6	50.5	234	31.31	20.615	9.072	2.251	86.60	4.94	9.90
	Southern Hills Pond	6.86	26.0	39.9	171	0.72	21.280	7.461	0.535	61.00	2.80	11.36
	Sequiota Park	5.83	18.6	58.5	244	11.18	35.546	8.755	2.472	97.40	4.32	16.50
9/7/2017	Doling Cave	5.88	16.3	70.2	271	2.88	34.158	40.702	2.904	129.60	4.90	19.64
	BB College	5.81	15.8	68.0	265	0.07	27.623	43.100	3.243	126.40	4.10	19.20
	Smith Park	6.69	18.5	52.7	219	0.24	35.157	10.687	0.579	82.60	3.68	18.12
	Silver Spring	5.81	16.8	53.3	224	0.54	25.905	11.148	2.267	96.40	3.22	11.06
	Jones Spring	5.79	15.8	63.1	246	7.19	35.210	11.535	3.182	118.60	3.26	16.00
	Snow Complex	6.52	18.6	49.5	212	31.70	20.991	8.831	1.976	80.40	5.22	10.22
	Southern Hills Pond	6.41	25.0	36.6	134	0.72	24.181	8.091	0.284	54.20	2.78	12.16
	Sequiota Park	5.99	17.8	59.5	231	8.39	39.051	9.051	2.480	99.40	4.56	17.20

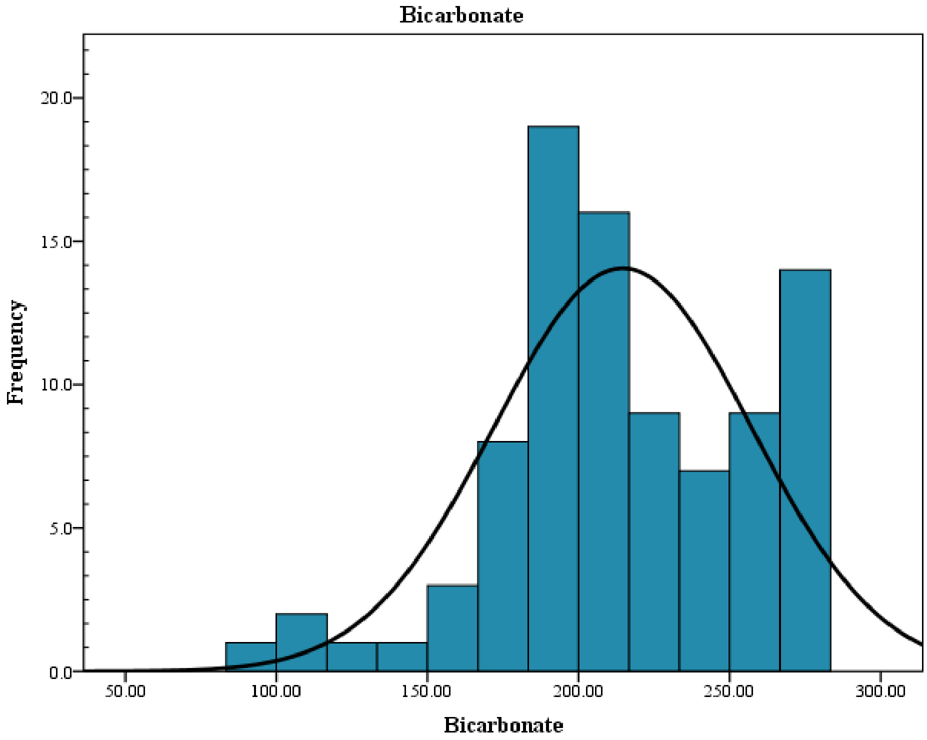
Appendix B – Statistics

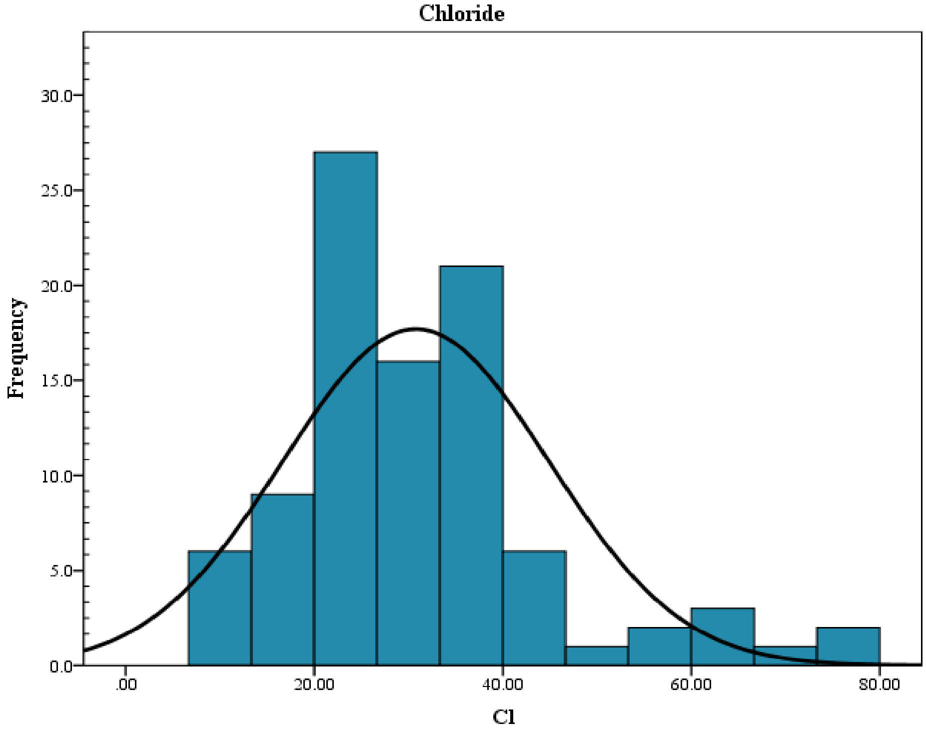
Normalcy Histograms.

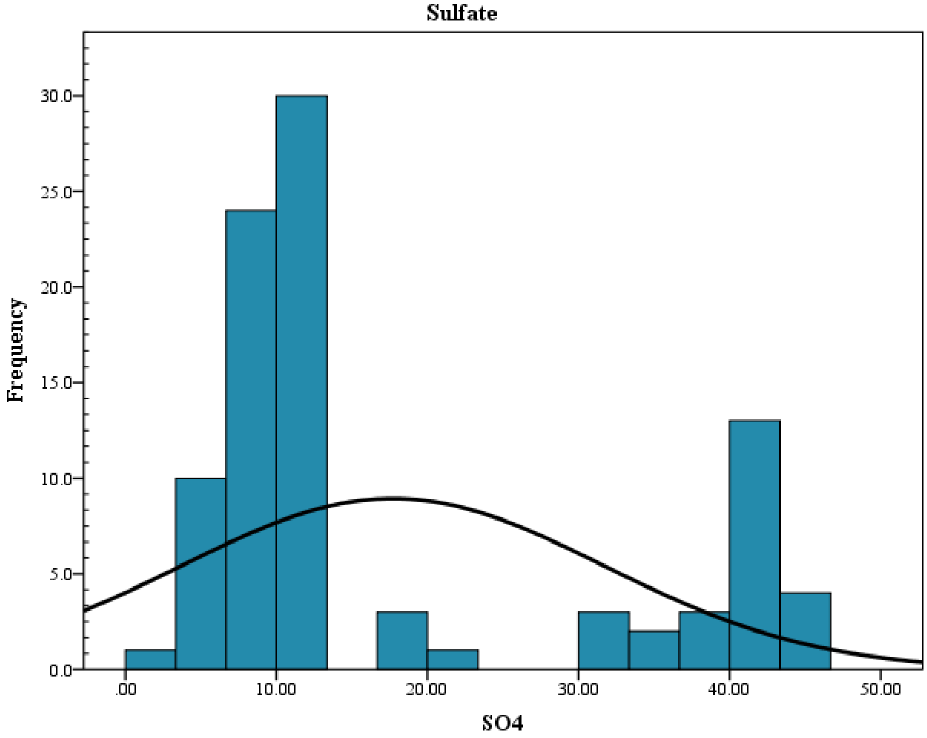


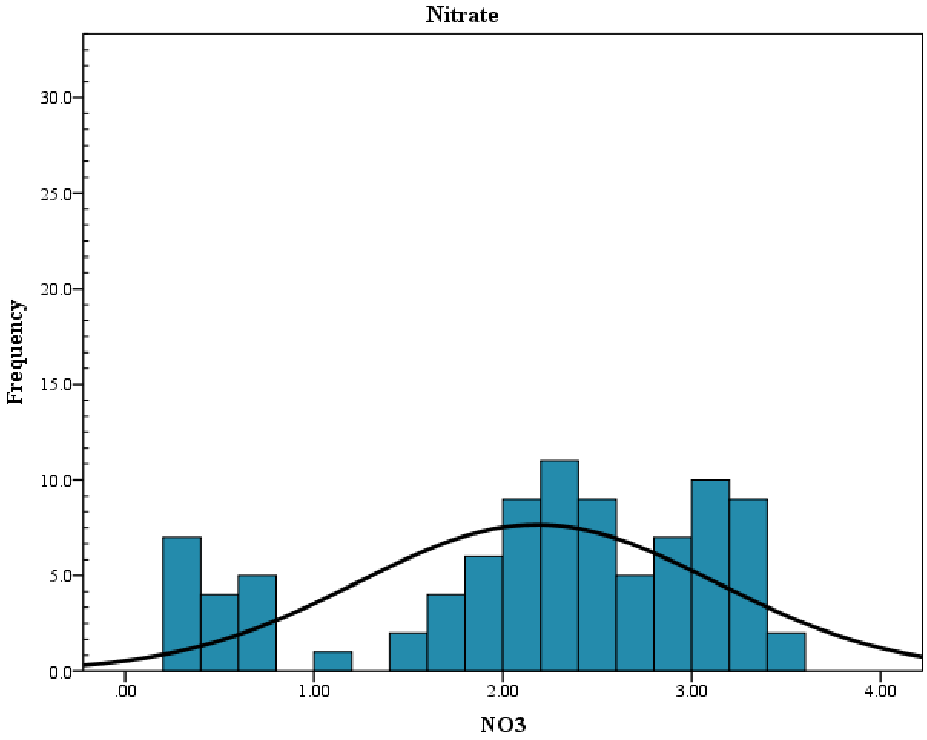


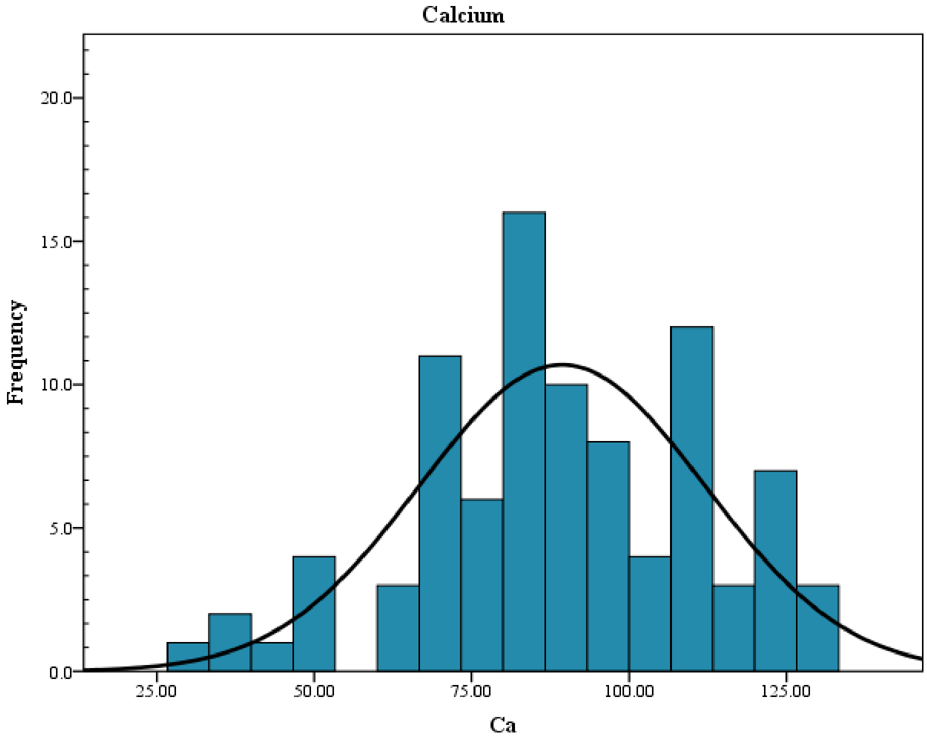

Conductivity
Frequency
20.0
15.0
10.0
5.0
0.0
20.00
30.00
40.00
50.00
60.00
70.00
80.00
Conductivity

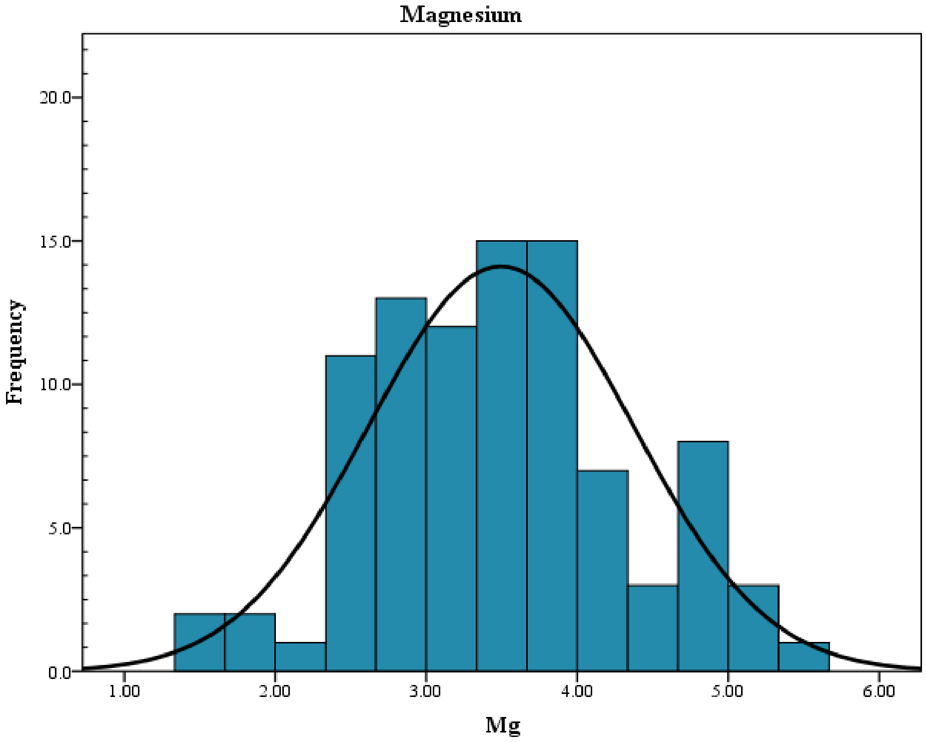

Bicarbonate
Frequency
20.0
15.0
10.0
5.0
0.0
50.00
100.00
150.00
200.00
250.00
300.00
Bicarbonate

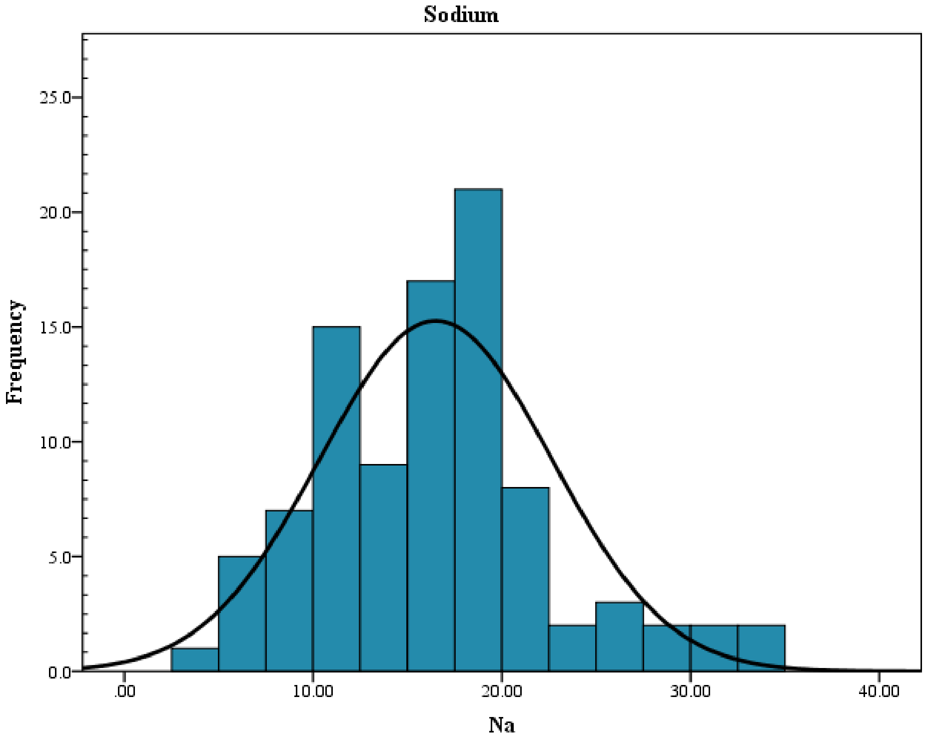

Chloride
Frequency
30.0
25.0
20.0
15.0
10.0
5.0
0.0
.00
20.00
40.00
60.00
80.00
Cl


Sulfate
Frequency
30.0
25.0
20.0
15.0
10.0
5.0
0.0
.00
10.00
20.00
30.00
40.00
50.00
SO4


Nitrate
Frequency
30.0
25.0
20.0
15.0
10.0
5.0
0.0
.00
1.00
2.00
3.00
4.00
NO3


Calcium
Frequency
20.0
15.0
10.0
5.0
0.0
25.00
50.00
75.00
100.00
125.00
Ca


Magnesium
Frequency
20.0
15.0
10.0
5.0
0.0
1.00
2.00
3.00
4.00
5.00
6.00
Mg


Sodium
Frequency
25.0
20.0
15.0
10.0
5.0
0.0
.00
10.00
20.00
30.00
40.00
Na

SAS Output – Model One

The SAS System

The DISCRIM Procedure

Total Sample Size	87	**DF Total**	86
Variables	10	**DF Within Classes**	82
Classes	5	**DF Between Classes**	4

Number of Observations Read	95
Number of Observations Used	87

Class Level Information

Basin	Variable Name	Frequency	Weight	Proportion	Prior Probability
Denny's	Denny's	18	18.0000	0.206897	0.200000
Doling	Doling	24	24.0000	0.275862	0.200000
Jones	Jones	17	17.0000	0.195402	0.200000
Sequiota	Sequiota	6	6.0000	0.068966	0.200000
Silver	Silver	22	22.0000	0.252874	0.200000

Pooled Covariance Matrix Information

Covariance Matrix Rank	Natural Log of the Determinant of the Covariance Matrix
10	14.66714

The DISCRIM Procedure

Generalized Squared Distance to Basin

From Basin	Denny’s	Doling	Jones	Sequiota	Silver
Denny’s	0	172.57213	25.23397	17.64112	32.28545
Doling	172.57213	0	139.92934	156.50550	120.57308
Jones	25.23397	139.92934	0	6.68149	7.60624
Sequiota	17.64112	156.50550	6.68149	0	9.39771
Silver	32.28545	120.57308	7.60624	9.39771	0

Linear Discriminant Function for Basin

Variable	Label	Denny’s	Doling	Jones	Sequiota	Silver
Constant		-371.13697	-503.04774	-408.37698	-384.77055	-393.57686
pH	pH	59.49625	61.20717	63.52384	59.28787	61.75245
temp	temp	10.05577	10.40467	9.69025	10.62422	9.62210
conduct	conduct	-0.71541	1.16009	-0.34251	-0.31611	0.10124
Bicarb	Bicarb	0.35691	0.29616	0.53730	0.48039	0.51751
Cl	Cl	0.73329	-0.30406	0.69253	0.67614	0.64323
SO4	SO4	-0.18962	4.24567	-0.00598	-0.11726	0.49250
NO3	NO3	31.82504	23.93725	22.71372	23.90618	16.34593
Ca	Ca	1.28813	1.23631	1.38377	1.31925	1.34341
Mg	Mg	-2.94694	-2.84759	-2.54712	-3.43439	-4.47435
Na	Na	-2.59587	-4.08804	-4.20768	-3.86075	-4.39807

The SAS System

The DISCRIM Procedure
Classification Results for Calibration Data: WORK.MODELONE80
Resubstitution Results using Linear Discriminant Function

Posterior Probability of Membership in Basin

Obs	From Basin	Classified into Basin		Denny's	Doling	Jones	Sequiota	Silver
41	**Silver**	**Jones**	*	0.0000	0.0000	0.7416	0.0068	0.2516
42	**Silver**	**Jones**	*	0.0000	0.0000	0.6475	0.0834	0.2691
54	**Jones**	**Sequiota**	*	0.0032	0.0000	0.1608	0.7993	0.0368
62	**Jones**	**Sequiota**	*	0.0000	0.0000	0.3704	0.6276	0.0019

*** Misclassified observation**

```
                         The DISCRIM Procedure
        Classification Summary for Calibration Data: WORK.MODELONE80
         Resubstitution Summary using Linear Discriminant Function

       Number of Observations and Percent Classified into Basin

From Basin  Denny's   Doling    Jones  Sequiota   Silver    Total

Denny's          18        0        0         0        0       18
             100.00     0.00     0.00      0.00     0.00   100.00

Doling            0       24        0         0        0       24
               0.00   100.00     0.00      0.00     0.00   100.00

Jones             0        0       15         2        0       17
               0.00     0.00    88.24     11.76     0.00   100.00

Sequiota          0        0        0         6        0        6
               0.00     0.00     0.00    100.00     0.00   100.00

Silver            0        0        2         0       20       22
               0.00     0.00     9.09      0.00    90.91   100.00

Total            18       24       17         8       20       87
              20.69    27.59    19.54      9.20    22.99   100.00

Priors          0.2      0.2      0.2       0.2      0.2


                    Error Count Estimates for Basin

           Denny's  Doling   Jones  Sequiota  Silver   Total

Rate        0.0000  0.0000  0.1176    0.0000  0.0909  0.0417

Priors      0.2000  0.2000  0.2000    0.2000  0.2000
```

SAS Output – Factor Analysis

```
                              The SAS System

                          The FACTOR Procedure
                 Input Data Type               Raw Data
                 Number of Records Read              95
                 Number of Records Used              87
                 N for Significance Tests            87



              Means and Standard Deviations from
                       87 Observations

              Variable           Mean         Std Dev
              pH              6.64023        0.565488
              temp           17.47701        2.546484
              conduct        56.19195       10.566800
              Bicarb        217.95402       38.686849
              Cl             31.76013       14.015504
              SO4            18.34250       13.826425
              NO3             2.23073        0.901694
              Ca             90.84368       21.038264
              Mg              3.56736        0.840681
              Na             16.99540        5.965844
```

Correlation

		pH	**temp**	**conduct**	**Bicarb**	**Cl**	**SO4**	**NO3**	**Ca**	**Mg**	**Na**
pH	**pH**	1.00000	0.28614	-0.45894	-0.49242	-0.16996	-0.17431	-0.55151	-0.53494	-0.22070	-0.10764
temp	**temp**	0.28614	1.00000	-0.45370	-0.48248	0.11147	-0.43236	-0.72474	-0.57594	-0.04705	0.02677
conduct	**conduct**	-0.45894	-0.45370	1.00000	0.83621	0.44526	0.74288	0.80297	0.87923	0.54600	0.60070
Bicarb	**Bicarb**	-0.49242	-0.48248	0.83621	1.00000	0.01765	0.71407	0.74643	0.88390	0.56155	0.19199
Cl	**Cl**	-0.16996	0.11147	0.44526	0.01765	1.00000	-0.06062	0.18617	0.06245	0.10452	0.90608
SO4	**SO4**	-0.17431	-0.43236	0.74288	0.71407	-0.06062	1.00000	0.57973	0.75442	0.50479	0.23652
NO3	**NO3**	-0.55151	-0.72474	0.80297	0.74643	0.18617	0.57973	1.00000	0.81161	0.39944	0.25855
Ca	**Ca**	-0.53494	-0.57594	0.87923	0.88390	0.06245	0.75442	0.81161	1.00000	0.47191	0.25044
Mg	**Mg**	-0.22070	-0.04705	0.54600	0.56155	0.10452	0.50479	0.39944	0.47191	1.00000	0.22125
Na	**Na**	-0.10764	0.02677	0.60070	0.19199	0.90608	0.23652	0.25855	0.25044	0.22125	1.00000

The FACTOR Procedure
Initial Factor Method: Principal Components

Prior Communality Estimates: ONE

Eigenvalues of the Correlation Matrix: Total = 10 Average = 1

	Eigenvalue	Difference	Proportion	Cumulative
1	5.25843213	3.28858743	0.5258	0.5258
2	1.96984470	0.91051706	0.1970	0.7228
3	1.05932764	0.24549787	0.1059	0.8288
4	0.81382977	0.40248796	0.0814	0.9101
5	0.41134182	0.19467717	0.0411	0.9513
6	0.21666465	0.08801152	0.0217	0.9729
7	0.12865312	0.03763133	0.0129	0.9858
8	0.09102179	0.05026306	0.0091	0.9949
9	0.04075873	0.03063308	0.0041	0.9990
10	0.01012565		0.0010	1.0000

		Factor1	**Factor2**	**Factor3**
pH	**pH**	-0.56315	0.03819	0.44051
temp	**temp**	-0.59578	0.40777	0.45955
conduct	**conduct**	0.95885	0.22247	0.05092
Bicarb	**Bicarb**	0.89876	-0.19652	0.13074
Cl	**Cl**	0.25686	0.93549	-0.19347
SO4	**SO4**	0.77763	-0.18681	0.36772
NO3	**NO3**	0.88882	-0.11712	-0.26925
Ca	**Ca**	0.93650	-0.17346	-0.00405
Mg	**Mg**	0.58487	0.05246	0.62387
Na	**Na**	0.42354	0.87029	-0.00590

Variance Explained by Each Factor

Factor1	**Factor2**	**Factor3**
5.2584321	1.9698447	1.0593276

Final Communality Estimates: Total = 8.287604

pH	**temp**	**conduct**	**Bicarb**	**Cl**	**SO4**	**NO3**	**Ca**	**Mg**	**Na**
0.51264630	0.73241328	0.97148719	0.86349353	0.97854120	0.77482331	0.87621458	0.90713140	0.73403120	0.93682247

SAS Output – Model Two

```
                              The SAS System

                          The DISCRIM Procedure
             Total Sample Size  88  DF Total               87

             Variables           5  DF Within Classes      83

             Classes             5  DF Between Classes      4

                   Number of Observations Read  95

                   Number of Observations Used  88

                         Class Level Information

     Basin     Variable  Frequency  Weight   Proportion        Prior
               Name                                      Probability

     Denny's   Denny's          19  19.0000    0.215909     0.200000

     Doling    Doling           24  24.0000    0.272727     0.200000

     Jones     Jones            17  17.0000    0.193182     0.200000

     Sequiota  Sequiota          6   6.0000    0.068182     0.200000

     Silver    Silver           22  22.0000    0.250000     0.200000

                      Pooled Covariance Matrix
                             Information

                    Covariance  Natural Log of the
                   Matrix Rank  Determinant of the
                                 Covariance Matrix

                             5            12.53399
```

The DISCRIM Procedure

Generalized Squared Distance to Basin

From Basin	Denny's	Doling	Jones	Sequiota	Silver
Denny's	0	42.50905	16.68908	9.59263	11.03283
Doling	42.50905	0	17.51778	27.50600	21.41770
Jones	16.68908	17.51778	0	5.42490	1.50046
Sequiota	9.59263	27.50600	5.42490	0	2.68067
Silver	11.03283	21.41770	1.50046	2.68067	0

Linear Discriminant Function for Basin

Variable	Label	Denny's	Doling	Jones	Sequiota	Silver
Constant		-288.32862	-400.80056	-331.16541	-309.76269	-329.10366
pH	pH	42.96479	48.31261	46.23670	42.73703	46.02084
temp	temp	6.06615	7.06519	6.33222	6.92507	6.58455
conduct	conduct	4.47962	6.41771	4.62557	4.56772	4.63529
Bicarb	Bicarb	0.18627	0.13457	0.31891	0.26791	0.28256
Cl	Cl	-2.04101	-3.25664	-2.54398	-2.39799	-2.46566

The DISCRIM Procedure
Classification Results for Calibration Data: WORK.MODELTWO80
Resubstitution Results using Linear Discriminant Function

Posterior Probability of Membership in Basin

Obs	From Basin	Classified into Basin		Denny’s	Doling	Jones	Sequiota	Silver
34	**Silver**	**Sequiota**	*	0.0020	0.0000	0.0147	0.7722	0.2111
36	**Silver**	**Sequiota**	*	0.0020	0.0000	0.0587	0.7942	0.1451
37	**Silver**	**Jones**	*	0.0013	0.0000	0.5051	0.0586	0.4349
38	**Silver**	**Jones**	*	0.0000	0.0001	0.7369	0.0204	0.2425
39	**Silver**	**Jones**	*	0.0000	0.0379	0.7383	0.0101	0.2136
40	**Silver**	**Jones**	*	0.0002	0.0034	0.5876	0.0652	0.3435
41	**Silver**	**Jones**	*	0.0000	0.0001	0.7087	0.0170	0.2742
42	**Silver**	**Jones**	*	0.0009	0.0001	0.5596	0.0932	0.3463
43	**Silver**	**Jones**	*	0.0025	0.0001	0.4793	0.1822	0.3359
44	**Silver**	**Sequiota**	*	0.0124	0.0000	0.2509	0.4543	0.2823
45	**Silver**	**Sequiota**	*	0.0026	0.0000	0.2369	0.5508	0.2096
46	**Silver**	**Jones**	*	0.0005	0.0000	0.4433	0.3731	0.1831
47	**Silver**	**Sequiota**	*	0.0021	0.0000	0.3402	0.4612	0.1966
48	**Jones**	**Silver**	*	0.0355	0.0000	0.4077	0.1017	0.4551
54	**Jones**	**Sequiota**	*	0.0110	0.0000	0.2747	0.4596	0.2547
61	**Jones**	**Silver**	*	0.0000	0.0005	0.3244	0.1588	0.5163
62	**Jones**	**Silver**	*	0.0001	0.0005	0.1629	0.2997	0.5368
67	**Denny’s**	**Sequiota**	*	0.0358	0.0000	0.0001	0.9385	0.0255
80	**Denny’s**	**Silver**	*	0.1875	0.0004	0.1694	0.1540	0.4885

*** Misclassified observation**

The SAS System

The DISCRIM Procedure
Classification Summary for Calibration Data: WORK.MODELTWO80
Resubstitution Summary using Linear Discriminant Function

Number of Observations and Percent Classified into Basin

From Basin	Denny’s	Doling	Jones	Sequiota	Silver	Total
Denny’s	17	0	0	1	1	19
	89.47	0.00	0.00	5.26	5.26	100.00
Doling	0	24	0	0	0	24
	0.00	100.00	0.00	0.00	0.00	100.00
Jones	0	0	13	1	3	17
	0.00	0.00	76.47	5.88	17.65	100.00
Sequiota	0	0	0	6	0	6
	0.00	0.00	0.00	100.00	0.00	100.00
Silver	0	0	8	5	9	22
	0.00	0.00	36.36	22.73	40.91	100.00
Total	17	24	21	13	13	88
	19.32	27.27	23.86	14.77	14.77	100.00
Priors	0.2	0.2	0.2	0.2	0.2	

Error Count Estimates for Basin

	Denny’s	Doling	Jones	Sequiota	Silver	Total
Rate	0.1053	0.0000	0.2353	0.0000	0.5909	0.1863
Priors	0.2000	0.2000	0.2000	0.2000	0.2000	