5-1-2012

Evaluating Recharge and Dynamics of Flow in the Lower Virgin River Basin, USA: Interpretation of Hydrochemical and Stable Isotopic Data

Joseph Asante
University of Nevada, Las Vegas, asantej@unlv.nevada.edu

Repository Citation

Asante, Joseph, "Evaluating Recharge and Dynamics of Flow in the Lower Virgin River Basin, USA: Interpretation of Hydrochemical and Stable Isotopic Data" (2012). *UNLV Theses, Dissertations, Professional Papers, and Capstones*. 1533.
https://digitalscholarship.unlv.edu/thesesdissertations/1533

EVALUATING RECHARGE AND DYNAMICS OF FLOW IN THE LOWER VIRGIN RIVER BASIN, USA: INTERPRETATION OF HYDROCHEMICAL AND STABLE ISOTOPIC DATA

By

Joseph Asante

Bachelor of Science
Department of Geology, University of Ghana, Ghana
1998

Master of Science
Technische Universität Hamburg-Harburg, Germany
2003

A dissertation submitted in partial fulfillment
of the requirements for the

Doctor of Philosophy in Geoscience

Department of Geoscience
College of Science
The Graduate College

University of Nevada, Las Vegas
May 2012

THE GRADUATE COLLEGE

We recommend the dissertation prepared under our supervision by

Joseph Asante

entitled

Evaluating Recharge and Dynamics of Flow in the Lower Virgin River Basin, USA: Interpretation of Hydrochemical and Stable Isotopic Data

be accepted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Geoscience
Department of Geoscience

David Kreamer, Committee Chair

Michael Young, Committee Member

Matthew Lachniet, Committee Member

Ganqing Jiang, Committee Member

Ding Zhonghai, Graduate College Representative

Ronald Smith, Ph. D., Vice President for Research and Graduate Studies
and Dean of the Graduate College

May 2012

ABSTRACT

Evaluating Recharge and Dynamics of Flow in the Lower Virgin River Basin, USA: Interpretation of Hydrochemical and Stable Isotopic Data

By

Joseph Asante

Dr. David Kreamer, Examination Committee Chair
Professor of Hydrology
University of Nevada, Las Vegas

Because of the complex geologic setting of the Basin and Range province, groundwater flow systems of the Intermountain basins of the southwestern United States are complex and remain poorly understood. Understanding these flow systems is important for water budgeting on a regional and local scale, and development of robust numerical groundwater models for sustainable water use and protection of water-dependent ecosystems. Although for decades hydrochemistry and isotopes have been used to characterize and trace subsurface water and surface water, effectively interpreting these data are still challenging, which can be attributed to existing subjective grouping of these data and the lack of methodological framework for analyzing and interpreting the data.

In this dissertation, new analytic approaches to analyze hydrochemical and stable isotopic data are described and used to trace the sources and movement of groundwater, and better quantify Virgin River interactions with groundwater in the lower Virgin River Basin in Nevada, Arizona, and Utah. In the new approach, data analytical techniques and data interpretations are combined in a sequential and mutually supportive way to test the hypotheses of potential interbasin groundwater flow, and Virgin River interaction with groundwater in the lower Virgin River Basin. To achieve these goals of the dissertation, three studies were conducted.

The objective of the first study was to optimize the multivariate statistical grouping of hydrochemical data of groundwater. The results indicated that repeated stable cluster solutions (robust hydrochemical facies) are obtained when cluster analysis is combined with Discriminant Function Analysis and agreement measures to group hydrochemical data. Also, this integrative approach allows for a quantification of the effect of analytical errors, outliers, and data transformation on the clustering of hydrochemical data. Using this approach, an optimal number of six robust hydrochemical facies were delineated for groundwater in the lower Virgin Valley. The results indicate that inappropriate data transformation can significantly impact the delineation of robust hydrochemical facies (Cramer’s V < 0.8). In addition, the results indicate that analytical errors </= 19% and outliers </=7% can be ignored (Cramer’s V > 0.8) in clustering.

The objective of the second study was to precisely define the sources of groundwater and Virgin River interactions with groundwater in the lower Virgin River Basin through a new approach for analyzing hydrochemical data of groundwater and surface water that allows for the precise definition of sources and discharge end-members, and overlay of interpretations. The methodology developed in the first objective above was used to delineate 6 robust hydrochemical facies for all waters in the basin. Subsequently, hydrochemical end-members were identified using the Schoeller diagram and discriminant functions plot. Flowpaths were tested, based on geographic coherence and patterns of the characteristic facies. Additionally, the PHREEQC inverse modeling code was used, beyond the contemporary application for identifying processes of hydrochemical evolution, to diagnose the significance of potential recharge sources to the lower Virgin River Basin. Mineral phases and constraints selected were based on compositional plots, saturation indices calculations, and the available geologic information. PHREEQC inverse modeling indicates that at least 80% of the groundwater in the lower Virgin River Basin is derived from interbasin flow. In addition, based on the hydrochemical data, the floodplain aquifer interacts highly with the Virgin River, but both are unlikely to be hydraulically connected to the underlying Muddy Creek Formation aquifer.

Finally in the third study, linear regression, Spearman correlation tests, scatterplot, box-and-whiskers plot, and Wilcoxon Rank Sum test are unconventionally applied to glean information from δD and $\delta^{18}O$, and Na, K, SO_4, and Cl data of the hydrochemical facies delineated in Chapter 3. The δD values of the recharge end-members from adjacent basins and within the lower Virgin River Basin are high and significantly different from the low δD values of the discharge end members. Box-and-whiskers plot of $\delta^{18}O$ values, comparing to box-and-whiskers plot of δD values, indicate possible oxygen isotopic exchange between the discharge groundwater and the aquifer minerals. The isotopic exchange implies a long residence time of the groundwater which discharges in the basin. Correlation tests of δD and $\delta^{18}O$ versus Cl, and box-and-whiskers plots of δD, $\delta^{18}O$, and solute data indicate the Virgin River and floodplain aquifer are more homogeneous and evaporated than samples of the source/recharge and discharge end-members. Regionally transported deep carbonate water is invoked as possible explanation of the low δD values of the discharge waters in the lower Virgin River Basin. Mixing calculations for stable isotopes indicate that at least 50% of the

groundwater discharging in the lower Virgin River Basin derives from interbasin groundwater sources. Although this estimate is different from the estimates presented in Chapter 3, both results reinforce the significance of interbasin groundwater to the lower River Virgin Basin.

ACKNOWLEDGMENTS

This dissertation is the culmination of my course work and research at the University of Nevada Las Vegas Nevada. I gratefully acknowledge Dr. David Kreamer for his guidance and diverse support during this dissertation. The quality of this dissertation has improved, thanks to the support of my dissertation committee members and certain individuals. I thank Dr. David Kreamer, Dr. Michael Young, Dr. Matthew Lachniet, Dr. Ganqing Jiang, and Dr. Ding Zhonghai for their constructive critiquing of this dissertation. The words of encouragement, inspiration, and hope from Dr. Young, Dr. Kreamer, and Dr. Lachniet to me were really energizing.

Also I am indebted to Dr. Victor Heilweil of the United States Geological Survey (USGS) for providing me with feedback on Chapter 2 and Chapter 3 when I presented them at two meetings of the Geological Society of America (GSA). I thank Melanie Reed, a graduate student at UNLV, for helping me with a summer field sampling of wells and spring water for hydrochemical and isotopic analyses. Thanks to the UNLV LVIS Isotope Ratio Mass Spectrometer Laboratory for analyzing the isotopic values in the samples I collected. I express my sincere gratitude to the Department of Geosciences, UNLV for providing me with a teaching assistantship for five semesters and to the Urban Sustainability Initiative, UNLV for awarding me with a fellowship for two and a half years. The Virgin Valley Water District supported me for fieldwork and some summer research salary. This dissertation would not have been produced without the financial support from these organizations/departments.

Finally I thank my wife Joyce Darko and my children Nana Annor Asante and Yaa Asantewaa Asante who have supported me socially and spiritually. Your unwavering

love and care for me gives me hope and meaning to life. I dedicate this dissertation to my late Dad who had long dreamed of a PhD for me, but did not live to see this day.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

DISSERTATION OVERVIEW

1.1 Overview

Since the flow systems theory was formalized (Tóth, 1963; Freeze and Witherspoon, 1967) , basin scale hydrologic studies have been increasing, with the general objective to provide the needed scientific information for sustainable management of regional water resources and water dependent ecosystems (Eakin, 1966; Winograd and Friedman, 1972; Winograd and Thordarson, 1975; Mifflin and Hess, 1979; Johannesson et al., 1997; Davisson et al., 1999b; Thyne et al., 1999; Genereux, 2005; Hibbs and Darling, 2005; Anderson et al., 2006; Guler and Thyne, 2006; Belcher et al., 2009; Bushman et al., 2010; Hershey et al., 2010). In the Basin and Range province in the Great Basin Region of the southwestern United States, the complex geology created by Mesozoic compression and Cenozoic extensional tectonics (Page et al., 2006) complicates the flow systems of the intermountain basins. From the perspective of water supply, two aquifers are considered the most important and these are the Tertiary Muddy Creek Formation (semi-consolidated basin-fill sediment) and Paleozoic carbonate rocks. As the drilling depth of the carbonate aquifer is considerably deep (Johnson et al., 2002), the basin-fill deposits are the most important water supply source (Heath, 1984; Johnson et al., 2002). Conceptually, groundwater of the basin-fill of the intermountain basins can be replenished by infiltration of precipitation in the high topographic boundaries or in the alluvial fans of the basins, by infiltration of stream discharge on the basin floor, or by interbasin flow through the deep carbonate aquifer or permeable intervening mountains.

However, in practice, flow systems of intermountain basins remain poorly understood and controversial (Davisson et al., 1999a; Davisson et al., 1999b; Thomas, 1999; Winograd, 2001; Nelson et al., 2004; Nelson et al., 2005; Winograd et al., 2005). For example, the sources of water to the intermountain basins remain unclear.

Previous studies have examined interbasin flow on the basis of groundwater budget analysis, numerical modeling, potentiometric surface analysis, geochemical, and isotopic studies (Eakin, 1966; Winograd and Friedman, 1972; Winograd and Thordarson, 1975; Mifflin and Hess, 1979; Johannesson et al., 1997; Davisson et al., 1999b; Thyne et al., 1999; Nelson et al., 2004; Genereux, 2005; Hibbs and Darling, 2005; Anderson et al., 2006; Guler and Thyne, 2006; Belcher et al., 2009; Bushman et al., 2010; Hershey et al., 2010). The investigations are diverse due to the complexity of the problem, issues of analytical uncertainty, and quest for techniques that accurately identify flow systems of the intermountain basins. The uncertainty in quantifying precipitation, evapotranspiration, and recharge amount causes great uncertainty in the conclusions from groundwater budget and numerical modeling studies of interbasin flow. Also potentiometric surface analysis is uncertain because of lack of a sufficient number of evenly distributed wells in the deep carbonate aquifer to constrain regional flow patterns. In view of these challenges, many researchers use geochemical and isotopic data to study flow systems of the intermountain basins. Using hydrochemical data has an advantage because the groundwater quality is a manifestation of the groundwater flow patterns. Also, stable isotopes systematically vary in space or time. Because stable isotopic values of groundwater are generally conservative, they are used as tracers for sources and movement of subsurface water.

The applications of hydrochemistry and isotopes to solving hydrologic problems are myriad. Groundwater hydrochemical data have been used to define or refine conceptual models (Panno et al., 1994; Bassett et al., 2008) and to infer geological structures (Lloyd and Heathcote, 1985). Also, hydrochemical and isotopic data have been combined or used separately to trace the sources of recharge (Carrillo-Rivera et al., 1992; Ingraham et al., 2001; Kohfahl et al., 2008), evolution of water chemistry (Schwartz et al., 1981; Thomas et al., 1989; Kreamer et al., 1996), and subsurface movement and hydraulic connections (Williams, 1982; Johannesson et al., 1997; Stetzenbach et al., 2001). Hydrochemical and isotopic studies are numerous and diverse; therefore, a complete review of the hydrochemical and isotopic studies is considered unwieldy (Lloyd, 1986). The numerous studies are due to the dependence of hydrochemical and isotopic studies on research questions, geologic and hydrogeologic settings, physiographic areas, and regional or local hydrology. Although advances have been made in chemical and isotope basin hydrology, future advancement of basin hydrology research, using hydrochemistry and isotopes, demands a thorough review of past studies. In this chapter, attempt is made to review and organize the literature on the applications of hydrochemistry and isotope to flow systems using the flow system definition as a framework. Groundwater flow system is defined by the recharge (or sources), discharge, and linking flowpath (Tóth, 1963).

The first important step in effectively using hydrochemical data to characterize sources and movement of water in basins is classification of the hydrochemical data into chemically homogenous and distinct groups known as hydrochemical facies. Methods of classification are broadly classified in this research as graphical techniques (Guler et al., 2002), and multivariate statistical techniques (Ashley and Lloyd, 1978; Kreamer et al., 1996). Also, geologic, physiographic, and hydrologic characteristics have been used to classify hydrochemical and isotopic data, for example stratigraphy units, basin units, water type (thermal, groundwater, surface water/lake and spring), and depths of wells (Clark et al., 1982; Panno et al., 1994; Acheampong and Hess, 2000; Matter et al., 2006; Zhu et al., 2007; Kohfahl et al., 2008; Mahlknecht et al., 2008; Hershey et al., 2010). Graphical techniques, for example Schoeller, Piper, and Stiff diagrams of major ionic data and compositional plots of water quality parameters, are used to group hydrochemical data into hydrochemical facies. However, graphical techniques are ineffective for analyzing large datasets and limitless parameters. Also, in complex hydrogeologic settings such as the basin and range province where cross-formational flow and fracture-directed flow are important, well depth and stratigraphy alone are inappropriate for classification.

With the advent of high speed computers and statistical programs, many studies use Factor Analysis (FA), cluster analysis, and Principal Component Analysis (PCA) to group large hydrochemical datasets (Ashley and Lloyd, 1978; Dalton and Upchurch, 1978; Lawrence and Upchurch, 1982; Steinhorst and Williams, 1985; Kreamer et al., 1996; Thyne et al., 2004; Guler and Thyne, 2006; Woocay and Walton, 2008). According to a previous study (Guler et al., 2002), cluster analysis is the most efficient classification technique; however, cluster analysis as reported by previous studies is subjective. Considering that cluster analysis can produce clusters even when there are no substantive groups in the data, the need for an objective approach cannot be overstated. Also, PCA is subjective and as an ordination method sometimes groups cannot be discerned on PCA scores plots. Furthermore, hydrochemical data have been characterized using transect survey of water quality parameters along an intuitively predefined flowpath (Edmunds et al., 2002). This method, referred to as direct ordination (Shaw, 2003), is subjective, as the researcher defines the flowpath in advance. In addition, direct ordination does not allow direct overlay of future investigations.

The characterization of clusters (groups of hydrochemical data) as source and discharge end-members, based on the defining characteristics of the hydrochemical facies, is another important area of research; however, less attention has been focused in this area. End-members of small datasets have been defined using graphical techniques, compositional diagrams, and direct ordination (Panno et al., 1994; Clark et al., 2000; Matter et al., 2006; Zhu et al., 2007; Kohfahl et al., 2008; Mahlknecht et al., 2008; Hershey et al., 2010). As already mentioned, these approaches are subjective and cannot handle limitless parameters. Although cluster analysis is considered the most efficient method for classification of large datasets, it does not provide direct information on the chemistry of the statistical groups for defining end-members (Guler et al., 2002). To circumvent the problem, the class memberships of the clustering results are superimposed on the results of graphical techniques or PCA. The graphical plots are used to analyze the average of the parameters of the groups (Rosenthal et al., 1990; Thyne et al., 1999; McNeil et al., 2005; Helstrup et al., 2007; Cloutier et al., 2008). This procedure maximizes tendency for misinterpretation of end-members. As an alternative approach, the clusters can be superimposed on the results of the PCA to identify the end-members (Thyne et al., 2004; Woocay and Walton, 2008). Although this procedure has the advantage of using all cases of the dataset, cluster analysis and PCA can produce different grouping order, and that can be problematic for the interpretation of end-members. Among the commonly used graphical techniques (Piper, Stiff, and Schoeller diagrams), the Schoeller diagram is considered the most useful interpretation technique (Charron, 1969). Also end-members have been characterized using transect survey of water quality parameters along an intuitively predefined flowpath.

The end-members defined are linked by inferring flowpaths. Flowpaths are inferred by correlating end-members on the basis of their chemical characteristics. In some cases the identification of potential flowpaths has been optimized by simultaneously considering the chemical characteristics of the clusters and their areal distribution using Geographic Information Systems (GIS). To this point, I have not mentioned interpolation and geostatistical analysis of hydrochemical and isotopic data to trace flow systems. This is because interpolation and geostatistical techniques are applied to infer all the components of the flow system; therefore, it is proper to mention it at this stage. Using spatial techniques successfully is dependent on data availability and distribution. Also, flowpaths can be tested using analysis of isotopic values (Winograd and Friedman, 1972; Mifflin and Hess, 1979; Kirk and Campana, 1990; Davisson et al., 1999b; Leontiadis and Nikolaou, 1999; Ingraham et al., 2001; Smith et al., 2002; Nelson et al., 2004) and aqueous geochemical modeling techniques (Plummer and Back, 1980; Plummer et al., 1983; Plummer et al., 1990). But overlay of these additional investigations involving conservative solutes, isotopes, and aqueous geochemical modeling is efficient for unambiguously defined hydrochemical facies. The advantage of the overlay is that it will allow the synthesis of different kinds of evidence. As noted, the isotopes are most efficient when the hydrologic problem is clearly defined (Gat and Gonfiantini, 1981).

For this dissertation, a methodology of objectively classifying hydrochemical data and overlaying independent hydrochemical and isotopic evidence is presented. The goal is to precisely delineate the flow system of a basin. Herein, three independent, though related, studies are conducted to determine the sources and movement of groundwater and interactions between groundwater and surface water in the lower Virgin River Basin located in the Basin and Range province. The basin receives, at the highest elevations of its' flanking mountains, precipitation of about 600 mm. The precipitation decreases with decreasing altitude reaching a value of about 150 mm at the basin floor. Therefore, precipitation in the mountains is considered the major source of recharge to the groundwater in the basin. However, it has also been hypothesized that much of the groundwater in the basin is derived from interbasin flow through the deep carbonate rocks (Dixon and Katzer, 2002) or volcaniclastics and tuff ash rocks forming the Clover mountains at the northwest-north-northeast boundaries of the basin (CH2M HILL, 2002). In addition, faults are believed be to the main conduit of groundwater flow in the basin (Page et al., 2006).

In the first study, I innovatively combined cluster analysis, multiple discriminant analysis, and measures of association/agreement to objectively find the optimum hydrochemical groupings (or hydrochemical facies). The study results compare well with the results with test statistics, (Test statistics are a more complicated and limited approach). This study tested the hypothesis that cluster analysis combined with Multiple Discriminant Function Analysis and measures of association/agreement can be used to determine the optimum hydrochemical facies for a hydrochemical dataset. The prediction or demarcation for this hypothesis is that the optimum hydrochemical facies is stable and has the largest number of clusters (hydrochemical facies). The new approach allows for quantitative decisions to be made about the clustering technique to use for a hydrochemical dataset, number of hydrochemical facies that are significant, and effect of hydrochemical data transformation, analytical errors, and outliers on a clustering technique. In the second study, an integrated and sequential approach is used to analyze the hydrochemical facies to test the hypothesis that interbasin groundwater does not contribute to groundwater in the basin-fill aquifer of the lower Virgin River Basin and also the hypothesis that the Virgin River is hydraulically interconnected with the floodplain aquifer. The prediction is that the chemically evolved groundwaters in the lower Virgin River Basin are not linked to the interbasin groundwater hydrochemistry. Also the interconnected surface water - groundwaters are expected to have statistically the same chemistry. Firstly, the Schoeller diagram and discriminant functions plot are applied to the hydrochemical facies delineated to define source and discharge end-members. Second, the existing hydrologic and geologic information and areal distribution of the facies are used to define potential flowpaths. Third, geologic data, saturation indices, compositional diagrams are used to deduce the reactive mineral phases in the hydrogeologic environment. Finally, aqueous geochemical modeling is used to examine the hypothesis of flowpaths and mixing scenarios. In the third study, regression analysis, box-and-whisker plots, bivariate plots, Spearman correlation tests, and Wilcoxon Rank Sum Tests are used to analyze stable isotopic data to test the hypothesis that interbasin groundwater does not contribute to groundwater in the basin-fill aquifer of the lower

Virgin River Basin and also the hypothesis that the Virgin River is hydraulically interconnected with the floodplain aquifer. It was predicted that the stable isotopic values of the discharging groundwaters in the lower Virgin River Basin is statistically similar to the source groundwaters (intrabasin recharge waters). Also the interconnected surface water - groundwaters are expected to have statistically the same stable isotopic values.

The results of the hydrochemical and stable isotopic data analyzed support that at least 50% of the groundwater in the lower Virgin River Basin is derived from interbasin groundwater. The current lack of a consistent, methodological framework to guide studies makes it difficult to advance the science of basin chemical and isotope hydrology, and this review and novel methodology provides that framework. The methodology presented in this study is objective and allows overlay and synthesis of independent investigations. This review is by no means an exhaustive review of the literature on hydrochemical and isotopic studies. Nevertheless, the review provides representative studies on the trends in the research area and serves as framework for future research.

CHAPTER 2

OPTIMIZING MULTIVARIATE STATISTICAL CHARACTERIZATION OF HYDROCHEMICAL GROUNDWATER DATA: A COMPARATIVE ANALYSIS OF CLUSTERING METHODS

Abstract

Many studies have grouped hydrochemical datasets into hydrochemical facies using graphical and multivariate statistical methods. Compared to graphical methods, cluster analysis and Principal Component Analysis (PCA) can handle large datasets and limitless parameters, so are widely used by researchers. However, cluster analysis and PCA can be subjective, raising questions about the significance and confidence of the groupings. In this study, cluster analysis is combined with Multiple Discriminant Function Analysis (MDFA) and measure of association to objectively find the optimum clustering technique, number of clusters for a hydrochemical dataset, and stable clusters, and to assess the effect of hydrochemical data transformation, analytical errors, and outliers on a clustering. Using MDFA and Cohen's Kappa measure, the hydrochemical dataset is mathematically modeled into robust hydrochemical facies (or stable clusters). The cluster groupings are used as a training/test dataset for MDFA. The results indicate that within-groups linkage with squared Euclidean distance clustering method is the best method for the hydrochemical data, yielding six robust hydrochemical facies. The six stable cluster solution (optimum hydrochemical facies) is independently supported by the test statistics: explained variance, proportional reduction of error, F-Max, and Beale's 'F test'. Cramer's V coefficient between the hydrochemical facies for the log-10 and squared-root transformed datasets shows that choosing an appropriate data transformation is a key step in delineating significant hydrochemical facies (Cramer’s V < 0.8). It is demonstrated that the effect of outliers and analytical errors on clustering is insignificant (Cramer’s V > 0.8) when the dataset contains outliers </= 7% or analytical error /= 19%.

2.1. Introduction

Determining groundwater flow systems of hydrologic basins is an important and challenging research topic. The understanding of flow systems is important for accurately analyzing water budgets, developing robust numerical groundwater models, and protecting water quality. Flow systems are often inferred from numerical modeling and studying hydraulic, hydrologic, hydrochemical, lithologic, and structural data.

Because spatial groundwater quality has a causal relationship with gravity-driven groundwater flow (Tóth, 2009), hydrochemical data are used for studying and mapping basinal groundwater movement and flow rates. The goal of this study is to optimize the multivariate statistical grouping of hydrochemical data into hydrochemical facies using an approach that combines multivariate statistical techniques in a mutually supportive way, allowing for quantitative decisions in the grouping process. Hydrochemical data have been used, often together with environmental isotopes, to define or refine conceptual models (Panno et al., 1994; Bassett et al., 2008). In hydrochemical studies, data collected from hydrologic basins are used to trace sources of recharge (Carrillo-Rivera et al., 1992; Ingraham et al., 2001; Kohfahl et al., 2008), evolution of water chemistry (Schwartz et al., 1981; Thomas et al., 1989; Kreamer et al., 1996), and subsurface movement and hydraulic connections (Williams, 1982; Johannesson et al., 1997; Stetzenbach et al., 2001). Generally, hydrochemical interpretations from these studies are based on spatial mapping of hydrochemical data and temporal mapping of radioisotopic data like tritium. Identifying hydrochemical facies (Back, 1961) plays a central role in decoding the history of groundwater and in testing hypotheses about groundwater sources, movement, hydraulic connectivity, and water quality evolution.

The fundamental theory underlying the use of hydrochemistry as tracer for basin groundwater flow is the principle of groundwater chemistry evolution from rock and water interactions (Garrels and Mackenzie, 1967) and the systematic spatial and temporal variations of the chemical compositions (Chapelle, 2005). Although the theory is simple, in practice it is very difficult to map out important hydrochemical types that characterize groundwater flow in basins. Basin hydrochemical patterns are often complicated by heterogeneity and anisotropy of geological materials and the complexity of geochemical reactions. The methods for mapping hydrochemical patterns (graphical and multivariate statistics) to characterize hydrochemical data have varying weaknesses (Guler et al., 2002), which can limit their effectiveness. In general, graphical methods use restricted hydrochemical data and group samples subjectively. In some cases, hydrochemical data plotted on these diagrams cannot be visually separated, making clear demarcations of the hydrochemical facies impractical. To overcome these disadvantages, many researchers (Ashley and Lloyd, 1978; Dalton and Upchurch, 1978; Lawrence and Upchurch, 1982; Steinhorst and Williams, 1985) use multivariate statistical methods, commonly clustering techniques and Principal Component Analysis (PCA), to characterize hydrochemical data. Graphical and multivariate statistical techniques have been compared side by side and the results showed that, compared to graphical techniques, multivariate clustering techniques and PCA more effectively use large datasets and limitless parameters and are better at classifying data in the Q and R modes to determine similarities between samples and similarities between variables respectively (Ashley and Lloyd, 1978; Guler et al., 2002).

Along with the increasingly popular use of clustering techniques to characterize hydrochemical data come several challenges, including cluster analysis (i.e., choices of clustering algorithm, similarity measure, and linkage) and data preprocessing. Also, in any given multivariate technique, the number of hydrochemical facies can be considered an unresolved nuisance parameter toward the goal of using facies for hydrological interpretation. In practical applications, cluster analysis has been applied to hydrochemical data subjectively. For instance, the appeal of the cluster patterns displayed (Steinhorst and Williams, 1985) seems to be one commonly used approach for deciding about the optimum clustering solution. Also by semi-quantitatively comparing cluster analysis to graphical techniques, Guler et al. (2002) reported that Hierarchical Clustering, using Ward linkage and Euclidean distance, produce the optimum number of hydrochemical facies. This subjectivity (choice of clustering technique, no of clusters, and data preprocessing) can cause ambiguous or misleading grouping of hydrochemical data and hence incorrect interpretations of groundwater history. Evidently a need exists to optimize the cluster analysis of hydrochemical dataset using an objective approach, which can measure the stability of hydrochemical facies and quantify the effects of data preprocessing.

The proper preprocessing of data is important for recovering the ‘true’ hydrochemical facies (optimum hydrochemical facies). Data preprocessing can affect how a clustering technique groups data (Milligan, 1996) and hence recover the ‘true’ cluster structure. Recovery of this underlying cluster structure can be affected by analytical errors in data. For example, although the charge balance error of a reliable major ion analytical dataset is within ±5% (Deutsch, 1997), for practical purposes many studies have included samples with charge balance errors outside this range (Guler et al., 2002; Thyne et al., 2004). Furthermore, hydrochemical data often contain outliers, requiring a decision for data inclusion or exclusion. Also, hydrochemical data are generally positively skewed (Howarth and Earle, 1976); therefore, appropriately transforming data is often necessary before applying multivariate statistical analysis.

The basic idea underlying this study is that bias in the clustering of hydrochemical data can be reduced when stable hydrochemical facies are sought, a condition for true cluster solutions. In this study, clustering hydrochemical data is considered a modeling problem. Studies have shown that the underlying ‘true’ structure of a given dataset yields the significant number of clusters or ‘best’ clustering technique when the cluster solution is stable (Gross, 1972; McIntyre and Blashfield, 1980; Tibshirani and Walther, 2005). This process, called validation, ensures that inappropriate cluster structures are not imposed on the dataset. Also, many test statistics approaches, or ‘stopping rules’, are available for estimating the number of clusters to validate clustering solutions (Milligan and Cooper, 1985).

The objectives of this study are to innovatively combine clustering techniques with Multiple Discriminant Function Analysis (MDFA) and measures of association to find the optimum cluster solution (clustering technique and number of clusters) and to assess the effects of hydrochemical data transformation, analytical errors, and outliers on a clustering technique. The evaluative criteria are that the optimum clustering solution is a stable cluster solution (significant hydrochemical facies), has Cohen's Kappa coefficient (which measures the agreement between any cluster solution and MDFA solution) ≥ 0.8 and has the largest number of clusters (or hydrochemical facies). The n-1 cluster solution is stable if the MDFA predicted rate from jackknife classification for groups in the n-1 cluster solution (containing n-1 clusters) equals or exceeds a threshold value of 80% and the jackknife classification predicted rate for groups in the n cluster solution (containing n clusters) is less than 80%. Also cluster solutions for any two datasets (transformed data, data with analytical errors, and data with hydrochemical outliers) are considered regionally similar if Cramer's V is > 0.8. Cramer's V is a chi-square based measure of association between two independent nominal variable sets.

2.2. Study Area

The hydrochemical data for this study are collected from the lower Virgin River Basin (including Tule Desert Basin) and adjacent basin areas located in parts of Nevada, Utah, and Arizona (Figure 2. 1), hereafter called the study area. The study area has a complex stratigraphy and structure, with rocks and basin-fill sediments spatially variable and discontinuous (Maxey, 1968; Naff et al., 1974). The study area is located within the Basin and Range physiographic province of the western United States and forms part of the Colorado Basin flow system. In general, this province is comprised of valleys filled with thick alluvial sediments and mountain ranges separating the valleys (Heath, 1984). The lower Virgin River Basin has an area of about 4,463 km^2 and is bordered to the north and northeast by the Clover, Beaver Dam, and Bull Valley Mountains; to the east and south by the Virgin Mountains; and to the west and northwest by the Mormon Mesa, Mormon Mountains, and East Mormon Mountains (Dixon and Katzer, 2002). The Tule Desert basin is considered a part of the lower Virgin River Basin (though mountains separate it from the other basin areas), given that they may be hydraulically connected (Katzer et al., 2002).

Precipitation in the study area is controlled by elevation and seasons, ranging from about 609 mm/yr at the highest (2,438 m) to about 152 mm/yr on the basin floor (Dixon and Katzer, 2002). Winter precipitation dominates the study area, conversely with short duration and localized activity in the summer (Glancy and Van Denburgh, 1969; Winograd et al., 1998). Additionally, temperatures range in the summertime from highs near 40 °C, to wintertime lows below freezing. Annual potential evapotranspiration in the basin is much greater than precipitation. On the lower Virgin River Basin floor, the annual reference evapotranspiration is about 11 times greater than the annual precipitation (CEMP, 2011). Consequently, winter precipitation is considered the more important source of recharge. The lower Virgin River Basin is drained perennially by the Virgin River and Beaver Dam Wash and ephemerally by numerous washes that follow heavy precipitation. The most significant source of direct recharge to the lower Virgin River Basin is believed to be precipitation falling above the elevation of 914 m (Dixon and Katzer, 2002) of the mountains surrounding the basin.

Besides the elevation and seasons that control the amount and location of precipitation, the geology controls the location of recharge and groundwater movement. Several investigators have studied the geology of the lower Virgin River Basin (Bohannon et al., 1993; Williams, 1996; Langenheim et al., 2000; Dixon and Katzer, 2002; Page et al., 2005; Page et al., 2006). Geologic units in the lower Virgin River basin range in age from Proterozoic to Quaternary. The two major hydrostratigraphic aquifer units of water supply importance are the Paleozoic carbonate (2000 to ~4000 m thick) and the overlying Tertiary Muddy Creek Formation (~ 600 to 1,640 m thick). Much of the groundwater recharge to the lower Virgin River Basin is reported to move through the carbonate rock to the Muddy Creek Formation by upward flow (Johnson et al., 2002).

2.3. Materials and Methods

2.3.1. Sources of Data

Hydrochemical data from groundwater (wells and springs) samples collected from the study area were used in this study. The sources of the data comprised governmental and institutional reports (Glancy and Van Denburgh, 1969; Bateman, 1976; Brothers et al., 1993; Enright, 1996; Thomas et al., 2001; CH2M HILL, 2002); graduate student theses (Metcalf, 1995; Yelken, 1996), the United States Environmental Protection Agency (U.S. EPA) STORET database (U.S EPA, 2009); the United States Geological Survey (USGS) NWISWeb database (U.S Geological Survey, 2009); and sampling activities conducted during this study by the Virgin Valley Water District (VVWD), Mesquite and the University of Nevada Las Vegas (UNLV).

The database compiled contains 328 water samples and 40 attributes including water chemistry and geographic location coordinates/the U.S Public Lands Survey System. However, only 198 groundwater samples have complete data (measured/calculated and converted to consistent units) for the compositions of Ca, Mg, Na, K, HCO_3, Cl, and SO_4, total dissolved solids (TDS), and geographic location coordinates. For alkalinity values, field measurements were recorded in the database; but where unavailable, the laboratory value was used (Guler et al., 2002). Values for Si were not available for all groundwater samples in the study area; therefore, TDS was used as a surrogate variable in the multivariate analyses. Although TDS can correlate with individual ions or the major ions may be correlated with each other (Suk and Lee, 1999; Cloutier et al., 2008; Hershey et al., 2010), correlated variables in the cluster analysis violate no assumptions (Milligan, 1996). In several hydrochemical studies clustering techniques applied to TDS or electrical conductivity data in concert with major ion data have yielded useful results (Ashley and Lloyd, 1978; Williams, 1982; Seyhan et al., 1985; Riley et al., 1990; Rosenthal et al., 1990; Güler and Thyne, 2004; Thyne et al., 2004; Demlie et al., 2007; Helstrup et al., 2007). The hydrochemical data used is accessible at https://spreadsheets.google.com/pub?key=0AkihMREl9PhWdEVDMnIyakRFRUhlYmNiblJKMXZwUWc&hl=en&single=true&gid=0&output=html.

The final 198 samples were collected from 162 discrete sampling sites. A total of 16 sites had repeated analytical chemical data allowing for temporal variations to be assessed. Sample coordinates recorded in the U.S Public Land Survey System were converted to longitude and latitude coordinates using the Earth Point program (Earth Point, 2009) and Google Earth®.

2.3.2. Quality Assurance and Preprocessing of Hydrochemical Data

Although historical hydrochemical data have been shown to be useful in studying groundwater flow (Guler et al., 2002; Bushman et al., 2010), the analytical accuracy and time invariant assumption of the hydrochemical data (i.e. chemical steady state) can often be questioned. In this study the quality of all the analytical major ion composition datasets was checked using charge balance error (CBE) calculations (Feuerstein and Grimm-Strele, 1989; Hem, 1989). With the charge balance approach, analytical errors in individual cationic and anionic concentrations can compensate, but this consideration is unable to be quantified in our data (and in most field studies) and is assumed to be negligible. Also, temporal variation in the analytical data was checked using multiple samples collected at 16 discrete sampling points.

Standard methods based on chemical relationships were used to estimate missing data and to fill in data gaps. Alkalinities recorded as $CaCO_3$ were converted to bicarbonate (HCO_3^-) concentrations (Deutsch, 1997). Total dissolved solids were either laboratory measured (residue on evaporation at 180°C) or estimated (equivalent to residue on evaporation at 180°C) by summing all major ions (mg/L) plus silica (mg/L) and subtracting half the concentration of bicarbonate (mg/L) from the total sum (Hem, 1989; Hounslow, 1995).

Hydrochemical data are usually positively skewed (Howarth and Earle, 1976) and so is often normalized for effective multivariate analysis. Thus the data were transformed before the multivariate statistical analysis. Box-Cox analysis and box-and-whiskers plot, (R Development Core Team, 2010), were used to search for the appropriate normalizing power transformations for the variables. The goal was to choose the power, from the range of power, which minimizes the number of outliers and normalizes the data. Skewness and kurtosis were calculated (SPSS, version 19) for transformed and untransformed datasets to assess effectiveness of normalization. Two datasets were prepared to evaluate how transformation can affect cluster grouping of hydrochemical data or delineation of hydrochemical facies. For data set one, the original major ions and

TDS data were square-root transformed. For data set two, major ions and TDS were subjected to a Box-Cox analysis to find the appropriate power transformation. The transformed datasets were then standardized (z-score):

$$z = \frac{x-\mu}{\sigma} \tag{1}$$

where x is the value of the variable, μ is the mean, and σ is the standard deviation. Furthermore, two other datasets were prepared to evaluate effects of outliers and CBE on classification results. The Box-Cox analysis power transformed variables were examined for outliers defined as data points exceeding $\underline{+}$ 1.5 inter-quartile ranges (IQRs) from the 1st and 3rd quartiles. Based on these tests, 14 outlier samples (7 % of the 198 hydrochemical data) were removed from the dataset and the remaining 184 samples were saved as a new, separate dataset for further evaluation. Similarly, 19 % samples with CBE outside the range of ± 5 % were removed from the 198 hydrochemical data (Box-Cox power transformed) and the remaining data (160 samples) were saved as a new dataset.

2.3.3. Multivariate Statistical Analysis of Hydrochemical Data

For the new methodology, cluster groupings were used as training dataset/test set for MDFA to mathematically group the hydrochemical data into robust hydrochemical facies. Agglomerative hierarchical cluster techniques (linkage and variance methods) or non-hierarchical K-Means clustering method were combined in a mutually supportive way with MDFA and agreement measures to find stable cluster solutions and to infer the optimal number of clusters and clustering technique for a dataset (SPSS version 19). Because clustering techniques are many, considering all of them is beyond the scope of this study; therefore, this study considered only the methods commonly used in the geophysical sciences (Gong and Richman, 1995). The clustering techniques selected for this study are the single linkage, complete linkage, between groups (average linage), within group linkage, Ward method, and K-Means method. These methods have certain characteristic tendencies that can be problematic in recovering a ‘true’ cluster solution. For example, single linkage may produce chaining clusters especially in large datasets and Ward’s method may produce small size, spherical clusters (Everitt et al., 2011). Among these techniques Ward, K-Means, and average linkage are the most commonly used (Gong and Richman, 1995). However, the other techniques not commonly used were added for performance comparison and re-evaluation using the new approach in this study. Similarly, this study considered only similarity/dissimilarity measures commonly applied in the geophysical sciences (Gong and Richman, 1995). Thus, the squared Euclidean distance, Euclidean distance, Pearson correlation, and cosine were used in this study. However, the Euclidean distance appears to be favored by researchers (Gong and Richman, 1995).

The raw datasets in n × p multivariate data matrix, X, were preprocessed and analyzed as described earlier:

$$X = \begin{pmatrix} x_{11} & \cdots & x_{1p} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{np} \end{pmatrix} \qquad (2)$$

where the rows are the water samples analyzed, and the columns are the analytical hydrochemical data. Therefore x_{np} represents the pth variable in the nth water sample. The clustering using a distance measure was performed by converting the transformed and standardized X matrix data into n x n matrix of inter-sample dissimilarity. The Euclidean distance is expressed mathematically as (Everitt et al., 2011):

$$d_{ij} = \left[\sum_{k=1}^{p}(x_{ik} - x_{jk})^2\right]^{1/2} \tag{3}$$

where d_{ij} is the distance between i and j individual samples or clusters and x_{ik} and x_{jk} are the kth analytical variable in the p variables of the multivariate data matrix. The squared Euclidean distance is simply derived by squaring the Euclidean distance in Equation 3. However, the clustering using a similarity measure was performed by converting the raw data X into n x n matrix of inter-sample similarity. This is because cosine and Pearson correlation measures are by themselves standardizing measures and are also scale independent (Le Maitre, 1982; Romesburg, 1984). The cosine similarity measure is the cosine of the angle between ith and jth samples or clusters, which is subtended by vectors of their p-dimensional variables, and is expressed mathematically (Everitt et al., 2011):

$$cos\theta_{ij} = \sum_{k=1}^{p} x_{ik}x_{jk} / \sqrt{\left(\sum_{k=1}^{p} x_{ik}^2 \sum_{k=1}^{p} x_{jk}^2\right)} \tag{4}$$

The Pearson correlation similarity is the correlation between the p-dimensional variables of the ith and jth samples or clusters and is expressed mathematically as (Everitt et al., 2011):

$$r_{ij} = \sum_{k=1}^{p}(x_{ik} - \bar{x}_i)(x_{jk} - \bar{x}_j) \, / \, \sqrt{\sum_{k=1}^{p}(x_{ik} - \bar{x}_i)^2 \sum_{k=1}^{p}(x_{jk} - \bar{x}_j)^2} \tag{5}$$

where $\bar{x}_i$ and $\bar{x}_j$ are the means of the variables of the ith and jth sample or cluster. Both Pearson correlation and cosine coefficients have limits -1 and 1; therefore values of the coefficients were transformed into the limits 0 and 1. Subsequently, clusters were formed by pairing each of the linkages with each of the similarity/dissimilarity measures. The clustering techniques and similarity measures are described in detailed in many statistical literature sources (Le Maitre, 1982; Romesburg, 1984; Everitt et al., 2011).

Determining the ‘true’ number of clusters in a dataset remains a difficult problem (Everitt et al., 2011); to circumvent this problem, this study innovatively used discriminant function analysis to find the highest number of stable clusters for corresponding clustering techniques and similarity/dissimilarity measure. The discriminant analysis was used to predict stable cluster membership using discriminant prediction equations derived from predictors (or the training dataset) (Tabachnick and Fidell, 2007). The jackknife or Leave-one-out classification, a robust cross-validation method was used to predict the stable cluster membership. The basic theory of MDFA is extraction of eigenvalues and eigenvectors and calculation of discriminant functions (Le Maitre, 1982). The predictor variables used were the log10 transformed and standardized Ca, Mg, Na, K, HCO3, SO4, Cl, and TDS.

The sequence of the procedures developed for this research is:

(1) cluster memberships are assigned to the hydrochemical samples using cluster analysis;

(2) stable MDFA model is fitted to the cluster solution by reclassifying the samples through cross-validation and measuring agreement between cluster groupings and MDFA modeled groups; each hydrochemical sample (datum) in the cross-validation process is reclassified using the MDFA functions derived from all other hydrochemical samples, excluding the datum under consideration (i.e. the “Leave-one-out” classification);

(3) Cohen's Kappa (Cohen, 1960) agreement of modeled MDFA groups, measured between the clustering technique's assigned membership and the MDFA predicted membership are used to compare performance of independent clustering techniques predicting the same number of clusters;

(4) the similarity of one modeled MDFA groups to that of another is determined from a Cramer's V agreement value measured between the two MDFA models having the same number of clusters.

Both Cohen's Kappa and Cramer's V equal 1 for a perfect agreement and zero for agreement no better than chance alone. Steps (1-2) were done for all four datasets (all data square-root transformed, all data log10 transformed, log10 without outliers, log10 excluding samples with analytical errors). Finally, the effect of outlier samples, before and after removal from the dataset, and the effect of samples containing analytical errors, before and after removal from the dataset, was also measured using step (4). Accordingly, agreement was only measured between datasets that have equal total sample sizes.

To explain the steps further, determining stable number of clusters (robust hydrochemical facies) is initiated by preselecting the number of clusters; in this case, the selected number of clusters is large, with the intent of producing an unstable cluster structure. MDFA is then used to assess the stability of the cluster structure. The n-1 cluster solution is stable if the MDFA jackknife classification predicted rate for groups in the n-1 cluster solution (containing n-1 clusters) equals or exceeds a threshold value of 80% and the jackknife classification predicted rate for groups in the n cluster solution (containing n clusters) is less than 80%. A cluster structure is declared stable if MDFA successfully predicted at least 80 % of the members in each cluster groupings. When this stability is not achieved, the number of clusters is reduced by 1 and the cluster is reassessed until stability is achieved.

2.3.4. Test Statistical Methods for Evaluating the Number of Clusters

In this study, a new approach for finding a stable cluster solution has been described. The stability idea was used to logically deduce the number of clusters fitting the dataset. A semi-check on the new approach was done using selected test statistics (‘stopping rules’) (Milligan and Cooper, 1985) to calculate the number of clusters (Milligan and Cooper, 1985; Bacher, 2001; Kahya et al., 2008). Since a single test statistic method cannot be depended upon, four methods were used (Gordon, 1999) to check the fit of the number of stable clusters to the hydrochemical data. The test statistics used are the explained variance (ETA^2_K) or R-squared (RSQ), proportional reduction of error (PRE^2_K), and F-Max statistics (Bacher, 2001), and Beale’s ‘F test’ (Everitt et al., 2011):

$$ETA_K^2 = \frac{SS_b(K)}{SS_t} \tag{6}$$

$$PRE_K^2 = 1 - \frac{SS_w(K)}{SS_w(K-1)} \tag{7}$$

$$F - MAX = \frac{SS_b(K)/(K-1)}{SS_w(K)/(n-K)} \tag{8}$$

$$F\ test = \left(\frac{SS_w(K1) - SS_w(K2)}{SS_w(K2)}\right) / \left(\frac{n-K1}{n-K2} \times \left(\frac{K2}{K2}\right)^{2/p} - 1\right) \tag{9}$$

where $SS_b(K)$, $SS_w(K)$, and SS_t are between K clusters-, within K clusters-, and total-sum of square variance respectively, K is the number of clusters, n is the number of water samples, and p is the number of variables. The test statistics, ETA_K^2, F-MAX, and F-test are among the 5 best test statistics found in a simulation study (Milligan and

Cooper, 1985). The Beale's 'F test' is used to test if a division into K2 clusters is significantly better than a division into K1 clusters (K2 > K1). All test statistics were calculated using Oneway Anova (SPSS) to find the sum of squares (between groups, within groups, and total) and Excel spreadsheet to program the equations.

Certain rules were used to determine the optimal number of clusters. For F-MAX, the optimal K coincides with the maximum F-MAX value; for ETA_K^2, and PRE_K^2, K is an optimal solution when the test value of K+1 sharply decline or makes no relative substantial gain (Bacher, 2001). Beales F test is compared with an F distribution, with the degrees of freedom p(K2 - K1) and p(n - K2) at 0.05 significant level ($p < 0.05$). A partition into K2 is optimal if the Beale's F test value for is significant K2/K1 and not significant for K3/K2. Furthermore, the stable clusters were evaluated by analyzing coherence of their areal distribution and chemical character using ArcGIS and Schoeller, scatter matrix, and compositional diagrams.

2.4. Results

2.4.1. Classification Based on Cluster Stability

Examined in this study are the stability of cluster structure of corresponding clustering techniques and effect of inappropriate data transformation, outliers, and analytical errors on clustering. Of the 198 samples, 160 samples had CBE equal to or within ± 5 %, 184 samples had -10 % </= CBE </= 10 %, 14 samples had 5 % < CBE </= 12.5 %, 13 samples showed -12.5 % </= CBE < -5 %, only 1 sample had CBE > 12.5 % and 10 samples indicated CBE < -12.5 %. For a CBE above or below ± 5 %, the analytical result of that hydrochemical sample is prone to error. The sources of error may be analytical or an indication that an important ion(s) is not analyzed (Deutsch, 1997). The Box-Cox powers estimated were closer to zero; therefore log10- transformation was used as an approximation. Standardized values of skewness and kurtosis were used to evaluate normality of the untransformed, log10-transformed, and the square-root transformed Ca, Mg, Na, K, Cl, HCO_3, SO_4, and TDS. The results showed that the skewness and kurtosis values of the log-transformed variables are closest to zero (Table 2. 1).

As anticipated, the results of the Hierarchical Agglomerative Clustering techniques and Non-Hierarchical/K-Means clustering method are wide-ranging (Table 2. 2). Several results showed no robust hydrochemical facies. In all MDFA modeled solutions, the number of samples in the smallest MDFA facies exceeded the number of variable set. Remarkably, no stable clusters (robust hydrochemical facies) were found using nearest neighbor linkage with all the similarity/dissimilarity measures (Table 2. 2). Also, no stable clusters were found using between-groups linkage with EUCLID, SEUCLID.

However, other combinations of Hierarchical Clustering linkages and distance measures produced stable clusters and high cross-validation measures (Table 2. 2). With one exception (i.e., within-groups with cosine which produced five stable clusters), cluster analysis with cosine and Pearson correlation produced the smallest numbers of stable clusters (Table 2. 2). Four clusters were delineated using furthest neighbor linkage with EUCLID and SEUCLID, and Ward linkage with EUCLID. Also, Ward linkage with SEUCLID, within-group linkage with EUCLID or with cosine, and K-Means clustering each delineated five significant hydrochemical facies. The five hydrochemical facies delineated using the Ward linkage with SEUCLID had a higher predictability (Kappa = 0.894) than the within-groups linkage with EUCLID (Kappa=0.861). However, the five hydrochemical facies delineated using the K-Means method had the highest predictability (Kappa = 0.923). Six hydrochemical facies were delineated using within-groups linkage with SEUCLID, though Kappa was lower (0.903) than for the K-means method.

Table 2. 1. Comparing different data transformations' departures from normality

	Log10-Transformed Data			Square-root Transformed Data			Raw untransformed data		
	Mean (SD)	Skewness	Kurtosis	Mean (SD)	Skewness	Kurtosis	Mean (SD)	Skewness	Kurtosis
Ca	1.953 (0.392)	0.543	-0.720	10.546 (5.260)	1.078	-.114	138.76 (142.41)	1.507	1.210
Mg	1.412 (0.553)	-1.175	3.057	6.009 (3.334)	1.353	4.056	47.17 (59.38)	4.865	40.360
Na	1.904 (0.475)	-0.212	0.606	10.362 (6.247)	3.241	22.319	146.20 (291.76)	10.232	126.011
K	0.874 (0.391)	0.264	0.986	3.044 (1.662)	3.709	27.084	12.02 (22.87)	10.411	128.957
HCO_3	2.349 (0.188)	-0.160	1.333	15.304 (3.329)	.624	.233	245.23 (109.73)	1.154	1.089
Cl	1.765 (0.553)	0.246	-0.834	9.375 (6.425)	1.594	4.201	128.96 (201.24)	5.287	46.098
SO_4	2.273 (0.686)	-0.416	-0.403	18.009 (12.994)	1.519	4.767	492.30 (802.69)	6.094	57.935
TDS	2.878 (0.376)	0.550	-0.352	30.356 (15.098)	1.941	7.538	1148.29 (1457.68)	6.066	58.268

Table 2. 2. Stable cluster solutions of the corresponding clustering methods and similarity/dissimilarity measures and K-Means method

Similarity/ Dissimilarity Measure	Linkage Methods					
	Ward			Furthest Neighbor		
	Stable Clusters	Cohen's Kappa	% XV[a]	Stable clusters	Cohen's Kappa	% XV[a]
Euclidean	4	0.933	92.4	4	0.860	90.9
Squared Euclidean	5	0.894	90.4	4	0.860	90.9
Cosine	3	0.847	89.4	3	0.921	94.4
Pearson Correlation	3	0.847	89.4	3	0.847	88.4

K-Means method produced 5 stable clusters (Kappa = 0.923)

a. %XV is the percent cross-validated grouped samples correctly classified

X denotes not applicable

Table 2. 2. Stable cluster solutions of the corresponding clustering methods and similarity/dissimilarity measures and K-Means method (continued)

Similarity/ Dissimilarity Measure	Linkage Methods								
	Nearest Neighbor			Between-Groups			Within-Groups		
	Stable clusters	Cohen's Kappa	% XV^a	Stable clusters	Cohen's Kappa	% XV^a	Stable clusters	Cohen's Kappa	% XV^a
Euclidean	X	X	X	1	X	X	5	0.861	87.9
Squared Euclidean	X	X	X	1	X	X	6	0.903	90.4
Cosine	X	X	X	3	0.932	92.9	5	0.888	88.9
Pearson Correlation	X	X	X	2	0.949	96.5	3	0.886	91.9

K-Means method produced 5 stable clusters (Kappa = 0.923)

b. %XV is the percent cross-validated grouped samples correctly classified

X denotes not applicable

Same stable cluster solutions were compared and contrasted using Cramer's V statistic. Cramer's V value indicated that the four cluster solution given by furthest neighbor with EUCLID and furthest neighbor with SEUCLID are identical (Cramer's $V=1.0$, $p < 0.001$). However, the agreement between the furthest neighbor with EUCLID and with SEUCLID solutions and the Ward with EUCLID solution was low, though still statistically significant (Cramer's $V=0.614$, $p<0.001$). A five cluster solution given by the three clustering techniques were also examined for agreement. Only the strength of the agreement between K-Means and Ward with SEUCLID exceeded a Cramer's $V > 0.8$ (Table 2. 3). In the final analysis, the largest number of clusters (hydrochemical facies), stable and significant, delineated for the hydrochemical dataset was six (the within-groups linkage with SEUCLID). Also, the agreement between hydrochemical facies with and without outliers was significant and high (Cramer's $V=0.838$, $p <0.001$), as was the agreement between facies with and without samples containing analytical errors (Cramer's $V=0.830$, $p<0.001$). In contrast, the agreement between log10-transformed and square-root transformed data was significant but low (Cramer's $V=0.655$, $p<0.001$).

2.4.2. Classification Based on Test Statistical

In this study, the cluster stability approach was used to deduce an optimal six stable clusters. The fit of the six stable cluster solution to the dataset is independently evaluated using test statistics results (Table 2. 4). Evaluation includes consideration of ETA^2_K values, PRE^2_K values, a Beale's test, and maximum F-Max values. Interestingly, three of the four test statistics results showed that the six cluster solution 'best fit' the dataset. From the ETA^2_K values, a one cluster solution explained 0 % of the variance in the dataset; a 2 cluster solution explained about 48 % and so on. The increment in the ETA^2_K remarkably dropped from cluster 3 to cluster 4 and also from cluster 6 to cluster 7. In addition, the PRE^2_K values sharply decreased from cluster 3 to cluster 4 and cluster 6 to cluster 7. The Beale's test value of a 3 cluster solution is significant over a 2 cluster solution but the preceding 4 cluster solution is not significant ($p < 0.05$). Also, the Beale's test value of a 6 cluster solution is significant over a 5 cluster solution but the preceding 7 cluster solution is not significant ($p < 0.05$). The maximum F-Max value matched a 2 cluster solution.

Table 2. 3. Comparison of the five stable cluster solutions

Clustering Techniques	Cramer's V	Approximate Significance
K-Means versus Ward with SEUCLID	0.836	0.00
K-Means versus within-group with EUCLID	0.761	0.00
Ward with SEUCLID versus within-group with EUCLID	0.769	0.00
K-Means versus within-group with cosine	0.592	0.00
Ward with SEUCLID versus within-group with cosine	0.599	0.00
Within-group with EUCLID versus within-group with cosine	0.573	0.00

Table 2. 4. Test statistics results for estimating number of clusters

No. clusters	ETA^2_K	F-MAX	PRE^2_K value	Beale's F test p-values
1	0.0000	0.000	**	**
2	0.4768	178.635	0.476866	0.000
3	0.5882	139.265	0.212825	0.014
4	0.6220	106.420	0.082233	0.347
5	0.6651	95.808	0.113776	0.038
6	0.7051	91.810	0.119515	0.008
7	0.7261	84.419	0.071299	0.090
8	0.7453	79.428	0.070052	0.054
9	0.7572	73.666	0.046567	0.199
10	0.7689	69.509	0.048383	0.125

** denotes not defined
p-value denotes the value of the significant test at $p < 0.05$

2.4.3. Classification Based on Spatial Coherence and Chemical Character of Six Stable Clusters

Although, the focus of this research is on developing methodological procedures to optimize classification of hydrochemical datasets, the areal distribution as well as chemical character and evolution of the six stable clusters are also examined, which provide further support to the methodology. The results show that in the geographic space samples having similar cluster membership are coherent and the clusters formed patterns in the basin (Figure 2. 2B). Also, temporal data from the same well or spring grouped together providing evidence for time invariant water chemistry. Additionally, in comparing the 6 stable cluster solution to the 5 stable cluster solution, it is observed that the cluster 4 of the 5 cluster solution (Figure 2. 2A) is equivalent to clusters 1 and cluster 4 of the 6 cluster solution (Figure 2. 2B). This partitioning is justified looking at the distinct chemical character of cluster 1 and cluster 4 (Figure 2. 3 and Figure 2. 5). Cluster 4 tended to have much higher Ca, Cl, and TDS compared to cluster 1 (Figure 2. 3 and Figure 2. 5). The six stable cluster solution superimposed on scatter matrices for the variables (Ca, Mg, Na, K, HCO3, Cl, SO4, and TDS) showed that the robust hydrochemical facies are controlled by field hydrogeological processes (Figure 2. 4). The chemical characters of the six water types are distinct and are described as follows (Figure 2. 5). Cluster 1: Na/Ca-HCO3, cluster 2: Ca-Mg-HCO3, cluster 3: Ca-Na-SO4, cluster 4: Na/Ca-SO4, cluster 5: Ca-Mg-SO4, cluster 6: Na-Ca-SO4. The TDS and many major ions consecutively increased from cluster 2 → cluster 1 → cluster 4→ cluster 5→ cluster 3. Cluster 6 has the lowest Mg concentration.

2.5. Discussion

The study results demonstrate the effectiveness of the new methodology to quantify the effects of outliers, charge balance errors, and inappropriate data transformation and to compare clustering techniques. A field hydrochemical dataset is scarcely normally distributed and nearly always includes outliers. The study results show that stable cluster groupings are sensitive to the type of data transformation. Both MDFA and cluster analysis (using distance measures) performances can improve by normalizing and standardizing the data as well by minimizing impact of outliers. The log10-transformed variables are closest to normality comparing to the squared-root transformed and untransformed variables. A more normalized variable set reduces the impact of outliers on the performance of the clustering technique. Similarly, the discriminant function analysis, although robust to skewness, is highly sensitive to outliers (Tabachnick and Fidell, 2007). Thus the log10 transformation is the most appropriate normalizing transformation to normalize variables in this study. Choosing log10-transformation (critically determined from Box-Cox analysis) over the unstandardized or square-root transformation depended on the skewness and kurtosis values of the variables. The Box-Cox power estimates for the variables are near zero, and therefore choosing the log10-transformation is a reasonable approximation. Kutner et al. (2005) write that such an approximation does not change the effectiveness of the transformation. However, a study by Reimann (2002) indicates that a difference in normality can result between the estimated and the approximated. How this subtle difference in normality affects the cluster groupings is not the focus of the present study. Removing multivariate outliers is a difficult process; therefore transformation may be a better way to deal with outliers. According to Dreher (2003), log transformation can result in the loss of information the outliers are carrying. Although this may be a valid argument, it is believed that transforming the data maximizes the performance of the clustering techniques and this benefit far outweighs the risk of losing information (Güler and Thyne, 2003).

The transformed variable set is standardized by the mean and the standard deviation prior to clustering (using distance measures). However, controversies exist regarding the choice of a method or use of standardization (Dreher, 2003; Dreher, 2003). According to Güler and Thyne (2003), when data are transformed and standardized, the effectiveness of the parametric method (e.g. clustering and multiple discriminant function analysis) is maximized and misclassification arising from inappropriate weighting of parameters is avoided. Reimann (2002) made a similar conclusion in factor analysis of a geochemical dataset. It has been observed that clustering of a gauged streamflow dataset standardized by range was superior to standardization by mean and standard deviation (Milligan, 1996; Kahya et al., 2008). In contrast, Reimann et al. (2008) recommend that because range transformation is vulnerable to outliers it should not be applied to standardize geochemical and environmental data. In this study, within-group linkage with SEUCLID applied to two same datasets that are differently transformed but standardized by the same method (equation 1) result in two statistically different stable cluster solutions.

The data support that the six robust hydrochemical facies solution is the optimum for the dataset. In effect within-groups with SEUCLID technique producing the largest number of stable clusters is the optimum clustering method. The six robust hydrochemical facies display spatial coherence and have distinct chemical characters as indicated by the results. In addition, three of the four test statistics results support that the six stable cluster solution fit best the dataset. Thus splitting the five cluster solution into a six cluster solution improves the classification. It was more difficult finding stable clusters when the number of clusters is large; therefore clustering techniques able to find large numbers of stable clusters are superior. In practice, choosing a stable cluster solution containing the largest number of clusters provides an opportunity to generate and test additional hypotheses that otherwise will be missed selecting the cluster solution containing the fewer number of clusters. As stated by Everitt (2011), the largest number of stable clusters is preferred unless a researcher has enough background information to know apriori the number of clusters. Clustering with Pearson correlation produced smaller numbers of stable solution because Pearson correlation ignores size displacement as is supported in the literature (Romesburg, 1984). In conclusion, Pearson correlation is an ineffective measure for clustering hydrochemical dataset for which discriminating shift in displacement is important. However, the cosine measure can be useful in clustering hydrochemical dataset as it was able to produce five stable clusters. The cosine measure's higher number of stable clusters, compared to Pearson correlation's lower number of stable clusters, is due to sensitivity of the cosine measure to additive translation (Romesburg, 1984). The distance measures are able to delineate higher numbers of clusters and this is explained by their scale dependence property (Le Maitre, 1982).

This study has shown that extrapolating cluster techniques to other field sites, as is commonly done (Thyne et al., 2004) need to be critically examined. The finding that within-groups linkage with SEUCLID best groups our dataset contrasts with a previous study showing that the Ward linkage with EUCLID is the optimum (Guler et al., 2002). The difference in the results indicates that a clustering technique appropriate for grouping a hydrochemical dataset at one study site may be inappropriate for grouping hydrochemical data at a different study site. Therefore, each dataset must be treated as unique and the procedures outlined in this study can be followed to find the optimum stable cluster solution. Similar to numerical modeling of flow, classification of hydrochemical data needs to be updated as more data becomes available.

For the hydrochemical data used in this study, the presence of outliers, analyzed after the variables set has been log10-transformed and standardized, have no significant effect on the cluster solution. Thus, there does not appear to be a need to identify and filter for outliers, when the number of outliers is relatively small (7% in this case). Similarly, we find no statistically significant difference between the cluster solution for dataset containing analytical error and the cluster solution for dataset containing no analytical errors, indicating also that these steps may not be needed for datasets containing a relatively low number of questionable samples (19 % in this case). In fact, the error percentage is even smaller (7%) if the range -10% </= CB </= 10% considered practical in hydrochemical studies (Plummer et al., 2004), is applied to this study.

2.6. Conclusions

This study demonstrates that, by using the modeling approach and the idea of stability, quantitative estimates can be made about the effects of data transformation, analytical errors, and outliers and about robustness of delineated hydrochemical facies. Cluster analysis combine with discriminant function analysis and measures of association in a mutually supportive way effectively group hydrochemical data into robust hydrochemical facies and assess the effect of outliers and charge balance errors in grouping the hydrochemical data.

The study is a significant contribution to partitioning hydrochemical dataset into robust hydrochemical facies, a critical step for using hydrochemical data to characterize basins groundwater flow and groundwater surface-water interactions. The approach in this study can be used to analyze hydrochemical data of intermountain basins to trace groundwater sources and movement and groundwater interactions with rivers. In addition, the study results show that some techniques cannot successfully generate certain numbers of clusters with respect to data being analyzed; therefore the new methodology is also a useful guide for selecting the most appropriate clustering method and similarity/dissimilarity measure assuming the number of clusters is known *a priori*. Also, data transformation and standardization is considered critical for the finding the ‘true’ cluster structure; further research in this area is recommended. Finally, clustering techniques need to be tailored to characterize different field hydrochemical datasets. We caution that, although stability is necessary for accuracy, a stable hydrochemical facies solution is not necessarily accurate. Nevertheless, stability provides a good estimate of the accuracy of cluster solutions.

2.7. Acknowledgements

Funding for this study was provided by the Virgin Valley Water District (VVWD), Mesquite Nevada, USA. We also thank the Office of Urban Sustainability, University of Nevada Las Vegas (UNLV) for the Graduate Assistantship awarded to Joseph Asante.

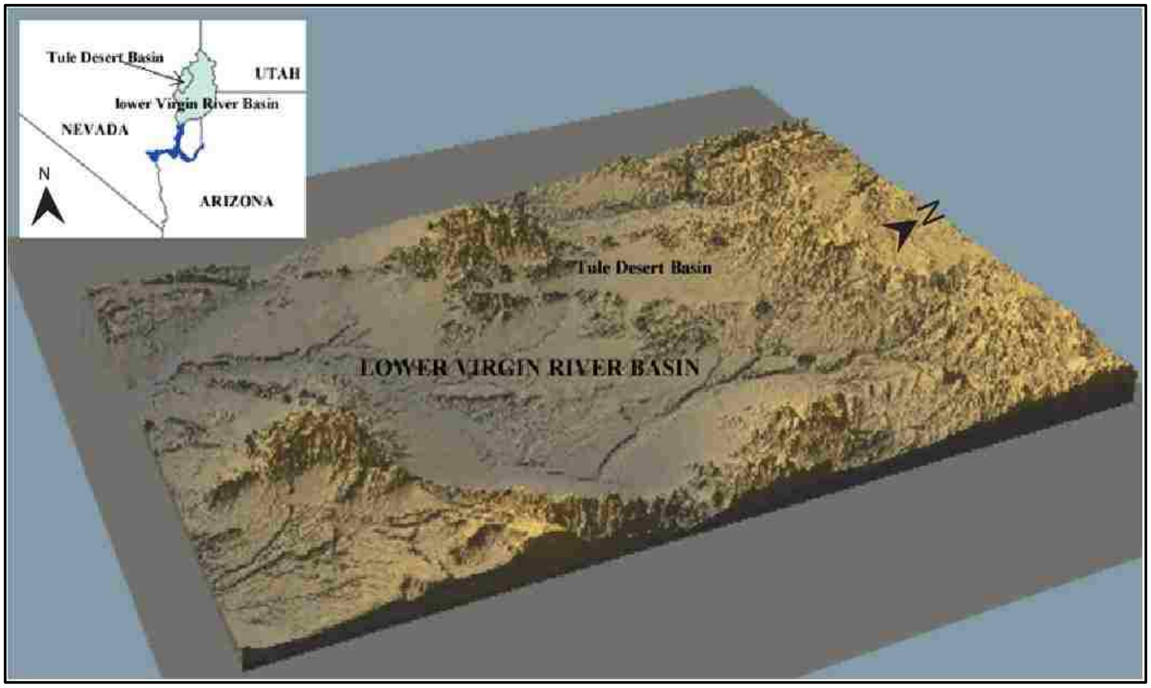

Figure 2. 1. The study area: the lower Virgin River Basin (Tule Desert Basin included) and basin adjacent areas.

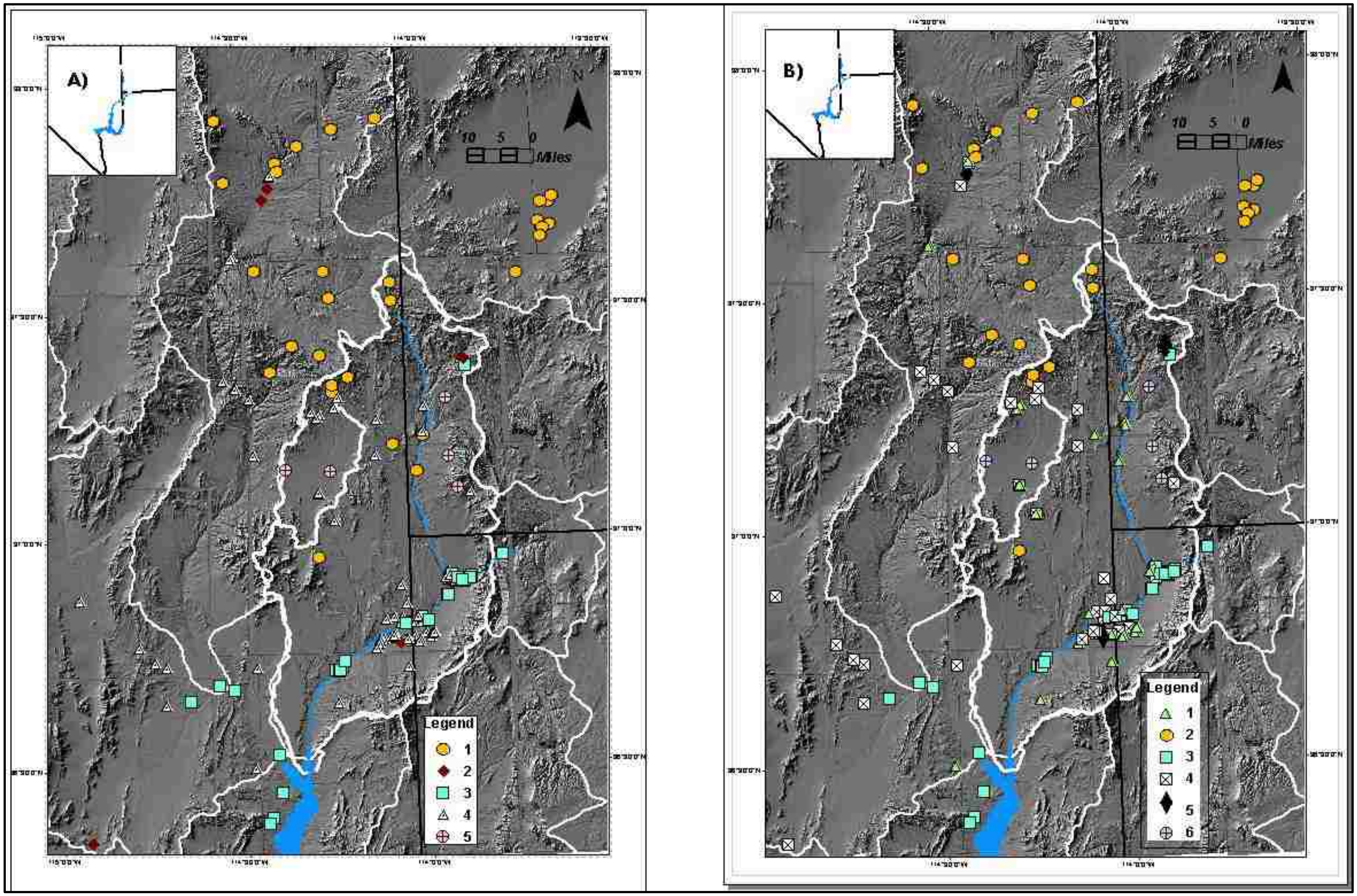

Figure 2. 2. Showing in geographic space on the background DEM are (A) the 5 Hydrochemical facies delineated using K-Means and (B) the 6 hydrochemical facies delineated using within-groups with SEUCLID.

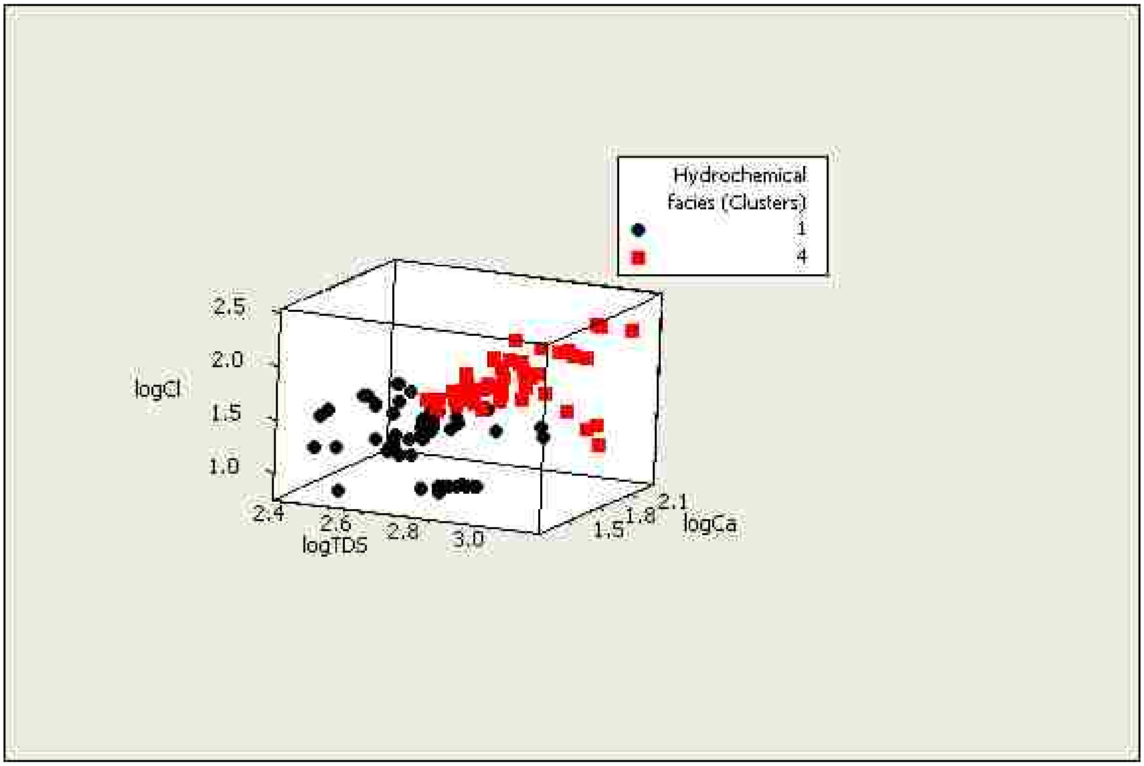

Figure 2. 3. Ca-Cl-TDS composition 3D-diagram comparing the hydrochemical facies 1 and 4 in the within-group hydrochemical facies.

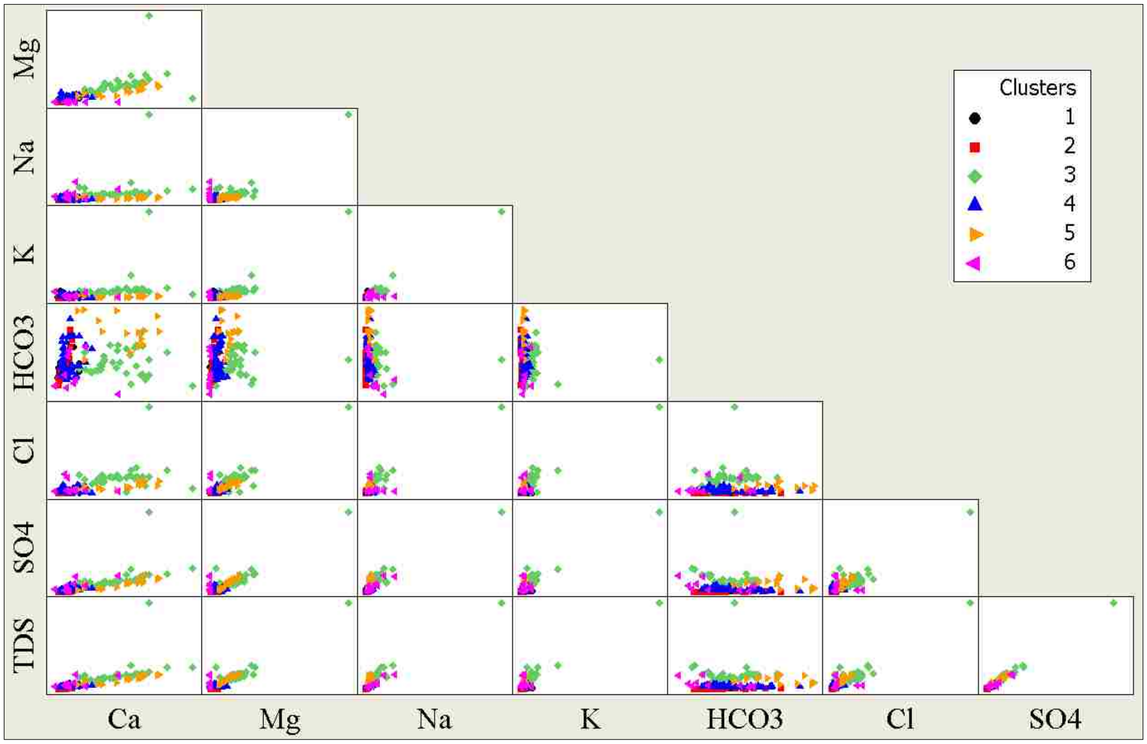

Figure 2. 4. Hydrochemical facies superimposed on scatter matrix of the hydrochemical variable.

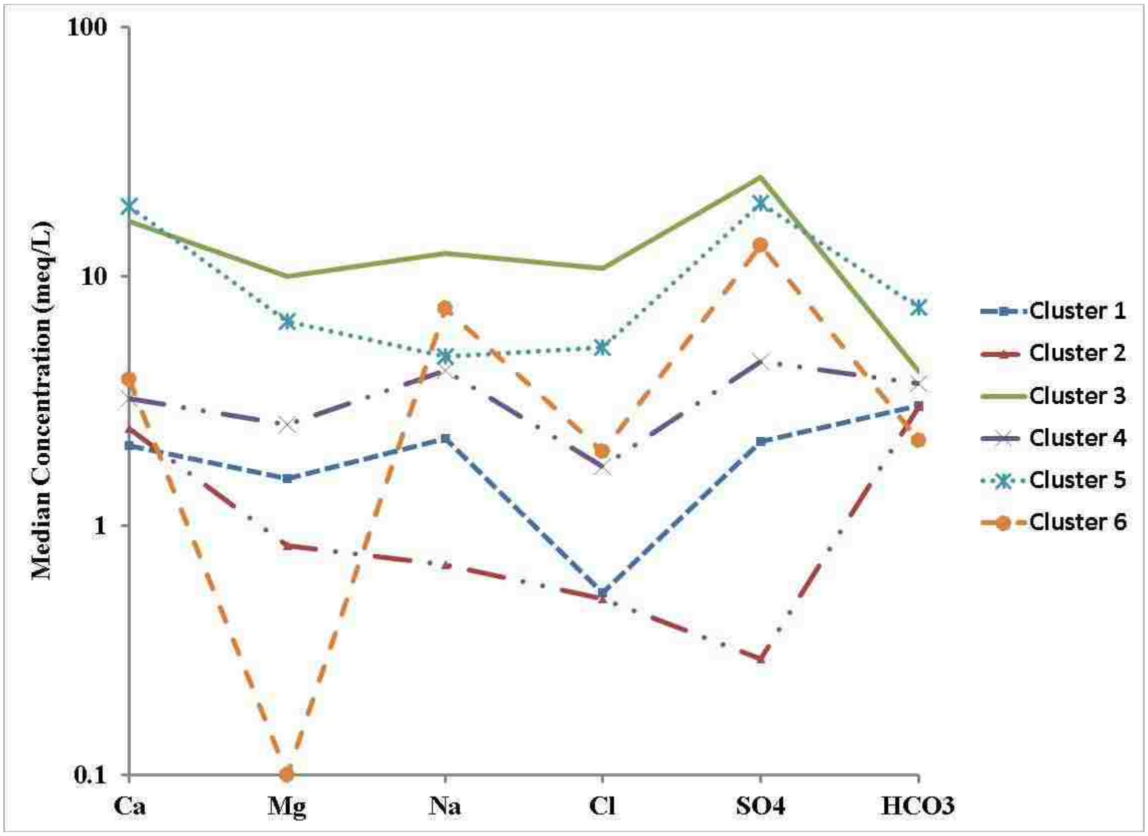

Figure 2. 5. Characteristics groundwater types of the median compositions of the clusters.

CHAPTER 3

BASIN HYDROCHEMISTRY AND RELATED FLOW SYSTEMS OF INTERMOUNTAIN BASINS: INTEGRATED AND SEQUENTIAL DATA ANALYTIC APPROACH

Abstract

Interbasin and intrabasin recharge sources and surface water interactions with groundwater are examined in the lower Virgin River Basin (Nevada, Arizona and Utah) using hydrochemical data. In this study a new approach that uses mutually supportive methods and applies them in sequence to the hydrochemical data is demonstrated. Using this approach the characteristics of the flow systems are more precisely defined to evaluate the significance of interbasin and intrabasin recharge to the basin's water budget and of the interactions among the floodplain aquifer, the Muddy Creek Formation aquifer, and the Virgin River.

Six robust hydrochemical facies were defined using cluster analysis, Discriminant Function Analysis, and measures of association (Cohen's Kappa and Kramer's V), which allows for repeatable facies definition. Two source and four discharge end-members of the hydrochemical evolution are defined using Schoeller diagram and discriminant functions plot. Interbasin and intrabasin flowpaths are inferred based on the geographic coherence of the end-members, characteristic facies, and available information on geology and hydrology. In addition, PHREEQC inverse modeling code was used beyond the contemporary application for identifying processes of hydrochemical evolution to further evaluate the inferred flowpaths. Mineral phases and constraints are selected based

on compositional plots, saturation indices calculations, and the available geologic information. Modeling results indicate that at least 80% of the groundwater in the lower Virgin River Basin is derived from interbasin flow. In addition, the floodplain aquifer was shown to interact highly with the Virgin River, but both are unlikely to be hydraulically connected to the underlying Muddy Creek Formation aquifer.

3.1. Introduction

Understanding basin scale groundwater flow dynamics is critical for water resources modeling and budget calculations and effective groundwater monitoring, all of which are necessary for a sustainable water management (Panno et al., 1994; Stuyfzand, 1999; Bassett et al., 2008). Groundwater hydrochemical data have been used to define or refine conceptual models (Panno et al., 1994; Bassett et al., 2008) and to infer geological structures (Lloyd and Heathcote, 1985). Also hydrochemical and stable isotopic data have been combined or used separately to trace the sources of recharge (Carrillo-Rivera et al., 1992; Ingraham et al., 2001; Kohfahl et al., 2008), evolution of water chemistry (Schwartz et al., 1981; Thomas et al., 1989; Kreamer et al., 1996), and subsurface movement and hydraulic connections (Williams, 1982; Johannesson et al., 1997; Stetzenbach et al., 2001). In addition, many studies have used aqueous geochemical modeling codes to explain the chemical processes of hydrochemical evolution (Thomas et al., 1989; Acheampong and Hess, 1998; Thyne et al., 2004; Helstrup et al., 2007).

Groundwater quality is one of a myriad of phenomena generated by gravity-driven groundwater flow (Tóth, 2009). The quality of groundwater is controlled by rock-water interactions and geochemical processes (Feth et al., 1964; Garrels and Mackenzie,

1967) and transport and distribution of dissolved minerals by moving groundwater, which groundwater creates a dynamic equilibrium sustaining many geochemical processes that are normally self-limiting (Tóth, 2009). The combined processes create hydrochemical patterns that can reflect groundwater flow patterns and groundwater-surface water interactions.

Analyzing hydrochemical data of the desert intermountain basins can be challenging due in part to the complex geology of the basins. This difficulty has encouraged the use of stable isotopes and radioisotopes independently or combined with major ion geochemistry to trace the hydrogeology of the intermountain basins. However, compared to water quality parameters (major ionic chemistry, temperature, and pH), isotopes are more expensive to collect and analyze. Also isotopes require detailed understanding of the background hydrochemistry and chemical behavior of the tracer for correctly interpreting the data (Guler and Thyne, 2006). In contrast, major ions, temperature, and pH data are routinely analyzed, and offer a less expensive approach to interpreting hydrologic systems. The analysis of hydrochemical data into hydrochemical facies can guide an isotopic research study. Clearly, the methods for interpreting hydrochemical data are a critical research area.

Using hydrochemical data effectively to trace sources and movement of subsurface water requires objectively classifying the data into chemically homogenous and distinct groups known as hydrochemical facies (Back, 1961) and more accurately defining end-members and linking flowpaths. Broadly, hydrochemical data can be characterized using graphical methods and multivariate statistical analysis. For large datasets and limitless parameters, graphical methods are less effective than multivariate techniques. Graphical methods used commonly are Piper, Schoeller, Stiff, and compositional diagrams, and direct ordination where compositions are plotted along a transect survey. Also multivariate statistical methods used commonly are Principal Component Analysis (PCA), Factor Analysis, and Cluster Analysis (Ashley and Lloyd, 1978; Dalton and Upchurch, 1978; Lawrence and Upchurch, 1982; Steinhorst and Williams, 1985; Kreamer et al., 1996; Thyne et al., 2004; Guler and Thyne, 2006; Woocay and Walton, 2008). Although cluster analysis is considered the most effective technique for classification of hydrochemical data (Guler et al., 2002), the method is fraught with subjectivity. The fact that clustering will produce clusters even when no substantive clusters exist is worrisome. Therefore a new approach, which allows for repeat delineation of hydrochemical facies, was demonstrated in Chapter 2 and the approach is used in the present study. Contemporary, to define end-members, the memberships of the clusters are superimposed on the graphical (Rosenthal et al., 1990; Thyne et al., 1999; McNeil et al., 2005; Helstrup et al., 2007; Cloutier et al., 2008) or PCA analysis (Thyne et al., 2004; Woocay and Walton, 2008). However, end-members defined using graphical methods are subjective and do not use the whole dataset. Although PCA can use the whole dataset, overlaying results from cluster analysis on PCA analysis can be problematic because the two results can differ.

Sequentially combining mutually supportive methods to analyze hydrochemical data can be a very powerful tool to study basin hydrochemistry and flow systems. This idea of formalizing the analysis of hydrochemical data to enhance interpretation has been expressed also by Thyne et al. (2004). By formalizing the analysis, a reduction can be achieved in ambiguous interpretations. In this study, major ions, temperature, silica, and pH data are analyzed, integrating and sequentially applying multivariate statistical techniques, Geographic Information Systems (GIS), Schoeller and major ion compositional diagrams, and PHREEQC geochemical code. Robust clusters are delineated by combining cluster analysis, discriminant function analysis, and measures of association, an approach described in Chapter 2. The end-members are defined using the Schoeller diagram and a discriminant functions plot, and the flowpaths are inferred with physiographic, geologic, and hydrogeologic information, and chemical characteristics and areal distribution of the facies. Further evaluations of the flowpaths are conducted using PHREEQC aqueous geochemical modeling. This sequential approach is used to test the hypotheses that groundwater of the lower Virgin River basin and the surrounding basins are hydraulically continuous, and that the lower Virgin River interacts with the Muddy Creek formation aquifer. Specifically, I test the hypothesis that interbasin groundwater does not contribute to groundwater in the basin-fill aquifer of the lower Virgin River Basin and also the hypothesis that the Virgin River is hydraulically interconnected with the floodplain aquifer. The prediction is that the chemically evolved groundwaters in the lower Virgin River Basin are not linked to the interbasin groundwater hydrochemistry. Also the interconnected surface water - groundwaters are expected to have statistically the same chemistry. Also the hypotheses are tested using the basic principles that the direction of groundwater flow is the direction of hydrochemical evolution and plausible geochemical reactions.

3.2. Geology and Hydrology of the Study Area

The lower Virgin River Basin (Figure 3.1.B) has an area of about 4,463 km^2 and is separated from the neighboring basins by the Clover Mountains, Bull Valley Mountains, Beaver Dam Mountains, Virgin Mountains, and Mormon Mountains (Figure 3.1.A). The lower Virgin River Basin is a tri-state basin located in parts of Utah, Nevada, and Arizona (Figure 3.1.B). The highest elevations of the Virgin Mountains and Beaver Dam Mountains exceed 2438 m. The highest elevations of the Clover Mountains and Mormon Mountains are about 2134 m and 1829 m respectively. The highest elevations in the Tule Spring hills and the East Mormon Mountains elevations are about 1219 m. The lowest elevation in the basin is about 366 m at the Lake Mead area, which is the outlet for the Virgin River.

Precipitation in the study area is controlled by elevation and seasons. In the lower Virgin River Basin precipitation ranges from about 609 mm/yr at the highest elevation (2,438 m) to about 152 mm/yr on the basin floor (Dixon and Katzer, 2002). Winter and summer are the two major rainfall seasons. Winter precipitation dominates the study area, with short duration and localized convective activity in the summer (Glancy and Van Denburgh, 1969; Winograd et al., 1998); winter precipitation is considered the more important source of recharge to the groundwater in the study area (Winograd and Friedman, 1972). The most significant source of direct recharge is believed to be precipitation in the mountains, above an elevation of 914 m (Dixon and Katzer, 2002). The lower Virgin River Basin is drained by two perennial rivers, the Virgin River, flowing northeast to southwest, and the Beaver Dam Wash, flowing nearly southwards, and by numerous ephemerally washes that follow heavy precipitation. Temperatures range in the summertime from highs near 40 °C, to wintertime lows below freezing. On the lower Virgin River Basin floor, the annual reference evapotranspiration is about 11 times greater than the annual precipitation (CEMP, 2011).

Due to Mesozoic compression and Cenozoic extensional orogenic activities, the geology of the study area is complex. A simplified geology (GBCGE, 2005) of the study area is shown in Figure 3. 2C. Details of the stratigraphy, lithology, structure, and aquifer potential have been reported by several investigators (Bohannon et al., 1993; Williams, 1996; Langenheim et al., 2000; Dixon and Katzer, 2002; Page et al., 2005; Page et al., 2006). Rocks of the mountain ranges are mainly consolidated metamorphic gneisses and schist, sedimentary clastics and carbonate, and volcanic ash flow tuffs. The Clover Mountain rocks are highly deformed and can transmit groundwater from its north to the south (CH2M HILL, 2002). The basin-fill sediment is thick and consists of deposits and sediments of limestone, conglomerates, sandstone, gypsum, ash flow tuffs, alluvium, colluviums, playa, and eolian. The Paleozoic carbonate rock and the Tertiary sedimentary rock (Muddy Creek Formation) are the most prolific aquifers of intermountain basins. As shown in Figure 3. 2, the rock units are cut through by numerous Cenozoic faulting. Therefore it is believed that groundwater transmits through fractures formed by solution cavities and Cenozoic faulting (Page et al., 2006).

3.3. Methodology

The hydrochemical data (spring, well, and surface water samples) used is accessible at https://spreadsheets.google.com/spreadsheet/ccc?key=0AkihMREl9PhWdER3aGxKc3lrRElacWZ5LXNiS0hoaHc&hl=en_US. The database comprised data from governmental and institutional reports (Glancy and Van Denburgh, 1969; Bateman, 1976; Brothers et al., 1993; Enright, 1996; Thomas et al., 2001; CH2M HILL, 2002); graduate student theses (Metcalf, 1995; Yelken, 1996); the U.S. EPA STORET database (U.S EPA, 2009); the U.S. Geological Survey NWISWeb database (U.S Geological Survey, 2009); and sampling activities conducted during this study by the Virgin Valley Water District (VVWD), Mesquite and the University of Nevada Las Vegas (UNLV). Although the database contained 328 samples and 40 variables, only 222 samples and 11 variables (Ca, Mg, Na, K, Cl, HCO_3, SO_4, total dissolved solids (TDS), Silica (SiO_2), temperature, and pH) were used in the study. The variables, Ca, Mg, Na, K, Cl, HCO_3, SO_4, and TDS were complete for the 222 samples, but results for SiO_2, temperature, and pH were incomplete. The 222 samples were distributed between 198 groundwater samples (162 discrete sampling sites and repeated data) and 24 surface water samples (21 lower Virgin River and 3 Beaver Dam Wash). To evaluate groundwater flow from the carbonate aquifer to the lower Virgin River Basin, the water quality parameters (hydrochemical and field parameters) from a USAF TEST WELL (730 m) tapping the deep carbonate aquifer was used (Bunch and Harrill, 1984). The well is located in the Dry Lake Valley north of lower Virgin River Basin, which is a possible upgradient regional groundwater flow to the lower Virgin River Basin. The major ions, temperature, silica, and pH data were analyzed applying integrative and sequentially using multivariate statistics, GIS, Schoeller and major ion compositional diagrams, and the PHREEQC geochemical modeling code.

3.3.1. Multivariate Statistical Classification of Hydrochemical data

Preceding the statistical analysis, Box-Cox analysis and box-and-whiskers plots were used to select the most appropriate normalizing transformation for each of the complete variables. Proper transformations are vital to improving cluster analysis results (Chapter 2); therefore in this study Box-Cox estimated power or rounded power was used (Fox et al., 2011) depending on the best normalization and reduced outliers as revealed by box-and-whiskers plots. The transformed variables were standardized by their mean and standard deviation. Hydrochemical facies were optimally defined by combining clustering techniques with Multiple Discriminant Function Analysis (MDFA) and measures of associations. The techniques and procedures used to delineate the optimal hydrochemical facies are extensively described in Chapter 2. The clustering technique producing the largest number of clusters and significant Cohen's Kappa value measured between cluster assigned membership and MDFA predicted membership is the optimum technique. The MDFA modeled cluster groups is the optimum hydrochemical facies.

As described in Chapter 2, the robustness of the optimum clustering technique to the effects of outliers and analytical errors were assessed by independently removing the outliers and analytical errors from the data and then cluster analyzing the resultant data. The hydrochemical facies defined using the data without outliers were compared with the hydrochemical facies defined using the data with outliers. The comparison was done using Cramer's V coefficient that measures the agreement between two nominal variable sets. The outliers were defined as those data points exceeding ±1.5 inter-quartile ranges (IQRs) from the 1st and 3rd quartiles. Based on these tests, 4 outlier samples (1.8% of the 222 hydrochemical data) were found and removed from the dataset to determine their effect on the cluster analysis results. Also, only 18% of the samples (41 out of 222 samples) had charge balance errors (CBE) outside the range of ± 5 % and were removed from the dataset to determine their effect on the cluster analysis results.

3.3.2. Discriminant Functions Plot and Schoeller Diagram

For the hydrochemical facies delineated, the chemical characteristics and chemical evolution of end-members were defined using discriminant functions plot and Schoeller diagram. The means (back transformed of mean of transformed data) of the major ionic values of the facies were plotted on the Schoeller diagram and together with the class memberships plotted on the discriminant functions plot, the end-members were defined. Because the major ionic data have been transformed, the mean values and standard deviations of the major ions were calculated from the transformed variables and then back transformed. Mean calculated from transformed data and then back transformed is similar to the Median because by this approach the leverage of maximum values is less. Median values were calculated for the pH, Temperature, and silica concentrations. Only subsets of the facies were used where data analyses involved incomplete variables.

3.3.3. Geographic Information Systems, Compositional Plots, Saturation Indices, and Aqueous Geochemical Modeling

Based on the chemical character and areal distribution of the end-members and the available geology and hydrology information, interbasin and intrabasin flowpaths and groundwater-surface water interactions were inferred. The mineralogical phases interacting with the water were deduced using compositional plots, saturation indices calculations, and the available geologic information. In addition, aqueous geochemical modeling was used to determine whether the flowpaths inferred are consistent with designated source and discharge end-members and selected minerals phases and constraints.

The PHREEQC geochemical code (Parkhurst and Appelo, 1999) was used to perform inverse geochemical modeling. Inverse geochemical modeling uses the mass balance concept (Garrels and Mackenzie, 1967) to identify and to quantify net geochemical reactions accounting for the evolution of groundwater along flowpaths. PHREEQC solves the mass balance equations for element or element valence state, electrons, alkalinity, charges, and water (Parkhurst, 1997). The general and simplest form of the equations is (Plummer and Back, 1980):

$$\sum_{p=1}^{j} \alpha_p \, \beta_{p,i} = \Delta m_i \qquad (10)$$

where j = the number of phases, p = reactants and phases used in the mass balance model, α_p = mass transfer coefficient of any phase p in moles, $\beta_{p,i}$ = stoichiometric coefficient of element i in phase p, and m_i = change in concentration of element i between the initial and final water along groundwater flowpaths. Besides solving for the unknown mole transfers of the phases α_p, the program was used to calculate mixing fractions for two waters. An inherent weakness of the modeling code is that mole balance models calculated using an inverse geochemical modeling are non-unique. Therefore data on mineral phases, thermodynamic feasibility of reactions, and saturation indices were used to constrain the PHREEQC models calculated.

The saturation index (SI) of the rock minerals is calculated using the speciation code of PHREEQC. Mathematically the saturation index is expressed as:

$$\mathrm{SI} = \log\left[\frac{\mathrm{IAP}}{\mathrm{K_{sp}}}\right] \tag{11}$$

where IAP = ion-activity product and K_{sp} = solubility product constant at equilibrium. If the solution is at equilibrium with the phase SI = 0; when the solution is undersaturated with the phase SI < 0; for conditions of supersaturation, SI >0. The saturation indices calculated were used to infer plausible mineral phases for the inverse modeling and to evaluate thermodynamic feasibility of precipitation or dissolution of minerals. The phases were optimally selected by combining inferences from the saturation indices plots, geologic information, and compositional plots of the major ionic data, TDS, and silica.

The conceptualized groundwater flowpaths were evaluated using the inverse geochemical modeling code. The model input data were hydrochemical data for the source and discharge end-members along inferred flowpaths, mineral phases and phase constraints in the aquifer. For hydraulically connected end-members, the PHREEQC code produces models consistent with the given set of phases, constraints for the phases, and reaction thermodynamic feasibility of the phases. No models are produced for hydraulically disconnected end-members. Between any two models the one with the smallest sum of residual is considered the best since it is the model most consistent ("best fit model") with the modeling data. In this study the mean (back transformed) Ca, Mg, Na, K, SO_4, Cl, HCO_3 and the median pH, temperature, and SiO_2 data of the hydrochemical facies inferred to be connected along flowpaths are used in the inverse modeling calculations to test the hypotheses about interbasin and intrabasin flowpaths for the lower Virgin River Basin. To account for spatial uncertainty, 5% uncertainty was assigned to all analytical data used in modeling calculations (Parkhurst, 1997). Also for each of the conceptualized flowpaths "minimal" models (Parkhurst, 1997) were determined.

3.4. Results

3.4.1. Hydrochemical Facies

Based on Box-Cox analysis and box-and-whiskers plots, the ions Ca, Na, Cl, HCO_3, and TDS are log10-transformed, Mg is square root transformed, and K and SO_4 are transformed by the powers 0.1285 and 0.1473 respectively. The transformation normalized the variables as well as reduced the number of outliers (Figure 3. 3 and Figure 3. 4). Among all the clustering techniques applied to the hydrochemical datasets, the Ward linkage with Euclid produced the optimum cluster solution consisting of six stable clusters (Table 3.1.). Outliers' effects on the optimal clustering method and the grouping of the hydrochemical data was acceptable (Cramer's V= 0.887, p<0.000). Similarly, the effect of data containing analytical errors (CBE</>±5) on the optimal method and the groupings was acceptable (Cramer's V=0.809, p<0.000). Thus the six hydrochemical groups as modeled by MDFA are the best possible hydrochemical facies for the study area.

Table 3. 1. Stable cluster solutions obtained using the Hierarchical clustering techniques and K-Means method

Similarity/ Dissimilarity Measure	Linkage Methods								
	Ward			Furthest Neighbor			Nearest Neighbor		
	Stable Clusters	Cohen’s Kappa	% XV	Stable clusters	Cohen’s Kappa	% XV	Stable clusters	Cohen’s Kappa	% XV
Euclidean	6	0.91	89.6	X	X	X	X	X	X
Squared Euclidean	3	0.903	92.3	X	X	X	X	X	X
Cosine	2	0.881	93.2	3	0.946	96.4	X	X	X
Pearson Correlation	2	0.881	93.2	2	0.881	93.2	X	X	X

Similarity/ Dissimilarity Measure	Linkage Methods					
	Between-Groups			Within-Groups		
	Stable clusters	Cohen’s Kappa	% XV	Stable clusters	Cohen’s Kappa	% XV
Euclidean	X	X	X	5	0.906	89.2
Squared Euclidean	X	X	X	5	0.9	89.6
Cosine	3	0.947	94.6	3	0.868	90.5
Pearson Correlation	2	0.963	97.7	2	0.881	93.2

%XV denotes the percent cross-validated grouped samples correctly classified.
X denotes 1 cluster solution or no cluster solution found.
The K-Means method produced 3 stable clusters (Kappa = 0.966, %XV=97.3%)

3.4.2. Areal Distribution and Chemical Characteristics of Clusters, End-Members and Flowpaths

The clusters have distinct chemical and spatial character. These are revealed by a spatial GIS map, Schoeller diagram and discriminant functions plot. The hydrochemical facies displayed geo-spatial coherence and patterns (Figure 3. 5). It is observed that the lateral variation of the clusters is more important than the temporal and vertical variations. Cluster 1 is located in the north basin-fill of the Virgin River Basin, near the Clover Mountain fronts. Cluster 2 is found in the Clover Valley and the Escalante Desert Valley areas. Cluster 3 plotted in the floodplain aquifer and along the Virgin River in the lower Virgin River Basin. Cluster 4 and cluster 5 is southward of clusters 1 and 2. The clusters form a sequence pattern from cluster 2 to cluster 1, then to clusters 4 or cluster 5 or cluster 6. Cluster 6 samples are localized at the Tenneco Gold Mine Wells located around the Bull Valley Mountains in Utah.

From the discriminant functions loadings (eigenvector loadings), important trends are extracted (Figure 3. 6). The groundwater clusters ordered 2, 5, 4, and 3 are distinguished by a successively increasing Cl concentration (Figure 3. 6). The clusters 2, 3, 4, and 5 have higher Na and K and lower HCO_3 and TDS compared to that of the clusters 1 and 6 (Figure 3. 6). These distinguishing features are also observed on the Schoeller diagram (Figure 3. 7). In clusters 2, 3, 4, and 5, the TDS sequentially is increased from cluster 2 → cluster 5→ cluster 4 → cluster 3 (Figure 3. 7). Groups 4 and 5 have the same characteristic pattern (same water type) but the chemical composition ions of the former are systematically higher than the latter. Unless specifically stated cluster 3 consist of floodplain groundwater and Virgin River samples. Cluster 3 (only the groundwater samples) and samples collected from the Virgin River (VR) have similar pattern and nearly the same composition (Figure

3. 7). The major water types (Figure 3. 7) are Ca-Mg-HCO_3 (clusters 1), Ca-Na-Mg-HCO_3 (cluster 2), Na-Ca-SO_4-HCO_3 (clusters 4 and 5), Ca-Na-Mg-SO_4-Cl (cluster 3 and Virgin River), and Ca-Mg-SO_4-HCO_3 (cluster 6). A groundwater sample collected from the deep carbonate aquifer in the Dry Lake Valley, located about north of the lower Virgin River Basin, is also plotted on the Schoeller diagram to compare with the groundwater clusters and to be used in searching for sources of groundwater in the lower Virgin River Basin. The deep carbonate aquifer is Ca-Mg-HCO_3 in chemical character and has a chemical pattern similar to the pattern of clusters 1.

3.4.3. Comparing Discriminant Functions Plot to Principal Component Analysis Plot

PCA is an ordination method used for data reduction through a linear combination of variables. The theory and procedures of PCA can be found elsewhere (Le Maitre, 1982; Kreamer et al., 1996) and are not described here. Although a PCA plot, similar to a discriminant functions plot, has the capability for extracting trends and explaining geochemical evolution, the ordered samples by PCA may differ from the cluster grouping. To demonstrate this shortcoming, the hydrochemical data has been analyzed using PCA with correlation matrix and Varimax rotation. The Principal Component (PC) scores of the first and second PC coordinates are plotted and the cluster groups are superimposed on the PCA ordered samples. The results show groups that are well separated on the discriminant functions plot compared to the PCA plot (Figure 3. 6 and Figure 3. 8).

3.4.4. Deduction of Rock and Water Interactions

The history of rock-water interactions in the study area is gleaned from scattergrams of composition of the major ions and SiO_2. Na/Na+Cl ratios of clusters

2, 4, 5 exceeded the halite dissolution line (gradient = 0.5), and that of Virgin River and cluster 3 (only groundwater samples) almost equaled the halite dissolution line. Clusters 1 and 6 plotted below the dissolution line (Figure 3. 9). Clusters 1and 2 plotted above Ca/Ca+SO4 ratio line (gradient =0.5), clusters 3 and 4 plotted below the line, cluster 5 bisected the line and cluster 6 plotted nearly on the line (Figure 3. 10). HCO_3/SiO_2 ratios of clusters 2 and some of cluster 5 were <5 meq/mmol, and that of clusters 1, 3, and 4 were >5 meq/mmol (Figure 3. 11). The HCO_3/SiO_2 ratios for cluster 6 are not calculated due to the lack of SiO_2 measurements. Mg/Mg+Ca ratios of all the clusters were <0.5 (Figure 3. 12).

The saturation indices of calcite, dolomite, gypsum, anhydrite, and halite for the clusters are plotted as a function of log10(Cl) ions (Figure 3. 13). An uncertainty of ±0.5 is assumed for the SI calculations (Deutsch, 1997). Within this uncertainty, most of the groundwater and surface water are saturated and a few are oversaturated with calcite and dolomite (Figure 3. 13). All the groundwaters and surface waters are undersaturated with the minerals gypsum, halite, and anhydrite. The undersaturation increased towards saturation progressively from cluster2→ cluster5→ cluster4→ clusters3 and 6 (Figure 3. 13).

3.4.5. Hydrochemical Evolution Modeling

Three flowpaths are formulated and evaluated using the PHREEQC inverse geochemical modeling code: (1) interbasin groundwater flow from the deep carbonate aquifer to the basin-fill aquifer—groundwaters of cluster 4 and 5; (2) interbasin groundwater flow from Clover Valley/Escalante Desert Basin—cluster 2 to the basin-fill aquifer; (3) intrabasin groundwater (cluster 1) flow to cluster 4 or 5. In addition to the three inferred flowpaths stated, the sources of water to cluster 3 (located in the floodplain of the Virgin River) are investigated. Model calculations are performed using representative chemical values of the hydrochemical facies (Table 3.2).

Table 3. 2. Mean chemical values (unless otherwise stated) of the clusters (water types) in mg/l used in inverse modeling and saturation indices calculations

Facies	Ca	Mg	Na	K	Cl	SO_4	HCO_3	SiO_2*	pH*	Temp*
cluster1	84.87	26.50	18.59	2.43	23.33	62.65	288.56	36.00	7.55	18.70
cluster2	39.64	8.06	20.94	4.01	17.63	13.84	170.03	49.00	7.77	21.60
cluster3	311.74	118.93	318.77	26.94	324.01	1258.97	243.82	19.80	7.20	23.00
cluster4	67.40	23.59	115.08	8.49	76.44	264.84	204.35	27.50	7.81	24.30
cluster5	42.61	17.52	62.15	5.84	19.11	119.54	182.77	31.00	7.81	26.18
VR	268.62	99.93	298.90	27.71	411.02	925.57	283.44	18.65	7.94	23.25
Carbonate	76.00	30.00	18.00	6.50	5.00	20.00	404.00	24.00	7.30	27.00

Temp denotes temperature (°C).

VR denotes the Virgin River samples.

Cluster 3 is groundwater samples only and does not include the Virgin River samples

Carbonate is the carbonate rock well sample (station name USAF TEST WELL) collected from the carbonate aquifer (730 m deep) underlying Dry Lake Valley located north of lower Virgin River Basin (Bunch and Harrill, 1984).

* denotes the median value.

Mineral saturation indices are calculated using the representative chemical values of the hydrochemical facies (clusters) in Table 3.2. Results show that all the hydrochemical facies and the carbonate group are saturated to supersaturated with calcite, aragonite, chalcedony, dolomite, and quartz (Table 3.3). However, all the clusters are undersaturated with anhydrite, gypsum, halite, magnesite, and sepiolite. The partial pressure of CO_2 (g) in the carbonate and cluster 1 appears to be higher than the atmospheric value of $10^{-3.5}$ atm. Mineral phases and constraints are selected and used in the inverse modeling calculations (Table 3.4).

Table 3. 3. Selected saturation indices of mineral phases with respect to the representative chemistry of the clusters

Phases	cluster1	cluster2	cluster3	cluster4	cluster5	VR	Carbonate
Anhydrite	-2.0	-2.8	-0.6	-1.5	-2.0	-0.7	-2.5
Aragonite	0.2	0.0	0.1	0.2	0.1	0.9	0.2
Calcite	0.3	0.1	0.2	0.3	0.2	1.0	0.3
Chalcedony	0.4	0.5	0.1	0.2	0.2	0.1	0.1
CO2(g)	-2.1	-2.6	-1.9	-2.5	-2.6	-2.6	-1.7
Gypsum	-1.8	-2.6	-0.4	-1.3	-1.7	-0.5	-2.3
Halite	-7.9	-8.0	-5.7	-6.7	-7.5	-5.6	-8.6
Magnesite	-0.5	-0.8	-0.5	-0.4	-0.4	0.3	-0.3
Quartz	0.9	0.9	0.5	0.6	0.7	0.5	0.6
Sepiolite	-1.8	-1.3	-3.1	-1.2	-1.1	-0.2	-2.9
Dolomite	0.4	-0.1	0.3	0.6	0.4	1.9	0.6

VR represents the Virgin River samples.

Carbonate is the well water sample (station name USAF TEST WELL) collected from the carbonate aquifer (730 m deep) underlying Dry Lake Valley located north of lower Virgin River Basin (Bunch and Harrill, 1984).

Table 3. 4. Selected phases for inverse geochemical modeling calculations

Phase	Composition	Constraint
Dolomite	$CaMg(CO_3)_2$	none
Calcite	$CaCO_3$	none
Halite	NaCl	dissolve
Gypsum	$CaSO_4$:2H2O	dissolve
Albite	$NaAlSi_3O_8$	dissolve
Aragonite	$CaCO_3$	dissolve
Quartz	SiO_2	precipitate
Anhydrite	$CaSO_4$	none
K-feldspar	$KAlSi_3O_8$	dissolve
Kaolinite	$Al_2Si_2O_5(OH)_4$	none
NaX	NaX	none
CaX2	CaX_2	none

The potential sources of recharge to the basin-fill aquifer (cluster 4 and 5) are cluster 1, cluster 2, and the carbonate aquifer designated cluster 7. The mole balance calculations for the flowpaths inferred to link the source and discharge end-members are shown in Table 3.5. Considering 100% of a source flowing to the basin-fill groundwater of cluster 4, the flow from the carbonate aquifer groundwater of cluster 7 to the basin-fill groundwater of cluster 4 had the smallest sum of residual. Also considering that mixed groundwater sources recharge the basin-fill aquifer groundwater of cluster 4, 87% Clover Valley groundwater of cluster 2 mixed with 13% carbonate aquifer groundwater had the smallest sum of residuals. No models are found for 100% of cluster 2 flowing to cluster 4. Among the models calculated when considering a single recharge source, the model for groundwater of cluster 1 flowing to groundwater of cluster 4 had the largest sum of residual.

Similarly, considering 100% of a source flowing to the basin-fill groundwater of cluster 5, the flow from cluster 2 to cluster 5 produced the only model and with a large sum of residual. The modeling showed that neither the carbonate groundwater nor groundwater of cluster 1 recharges the groundwater of cluster 5. Also, considering mixed groundwater sources to recharge the basin-fill groundwater of cluster 5, 84% carbonate aquifer groundwater and 16% groundwater of cluster 1 flowing to cluster 5 had the smallest sum of residuals. No models are found when mixtures of cluster 7 and cluster 2 and mixtures of cluster 1 and 2 are considered to flow to the basin-fill groundwater of cluster 5. Interestingly, considering groundwater of cluster 2 or groundwater of carbonate aquifer flow to groundwater of cluster 1 yielded no models. Only consistent models that had calcite, dolomite and quartz are selected; this is because these minerals are highly reactive phases (Table 3.3). Finally sources of water to cluster 3, groundwater collected from wells located on the floodplain, were also considered. No models were found when cluster 4 and 5 were considered as sources recharging cluster 3. However, a consistent model was found between cluster 3 and VR (Virgin River water) for cluster 3 flow to the Virgin River and not vice versa.

Table 3. 5. Models calculated using the PHREEQC inverse geochemical modeling code for the conceptualized groundwater flowpaths in the study area. Units are in moles per liter (mol/l).

Phase mole transfers	cluster 7 to cluster 4	cluster 1 to cluster 4	40% cluster7 + 60% cluster1 to cluster4	13% cluster7 + 87% cluster2 to cluster4	26% cluster1 + 74% cluster2 to cluster4	cluster 2 to clsuter 5	84% cluster7 + 16% cluster1 to cluster5
Dolomite	-2.64E-04	-6.50E-05	-1.38E-04	5.20E-04	4.58E-04	3.53E-04	-4.82E-04
Calcite	-3.29E-03	-1.70E-03	-2.28E-03	-1.07E-03	-9.77E-04	-6.99E-04	-2.86E-03
Halite	2.02E-03	1.50E-03	1.69E-03	1.71E-03	1.62E-03		3.15E-04
Albite	4.32E-03	1.89E-03	2.77E-03				4.32E-03
Quartz	-8.68E-03	-4.22E-03	-5.85E-03	-5.16E-04	-5.52E-04	-3.83E-04	-8.56E-03
Gypsum	2.41E-03	1.95E-03	2.14E-03	2.47E-03	2.34E-03	1.16E-03	1.02E-03
K-feldspar	5.09E-05	1.55E-04	1.17E-04	1.06E-04	1.25E-04	4.18E-05	
Kaolinite	-2.18E-03	-1.02E-03	-1.45E-03	-5.32E-05	-6.26E-05	-2.09E-05	-2.16E-03
NaX	-1.89E-03	1.03E-03		2.62E-03	2.72E-03	1.66E-03	-2.85E-03
CaX2	9.45E-04	-5.16E-04		-1.31E-03	-1.36E-03	-8.29E-04	1.43E-03
Sum of Residuals	3.0	6.4	4.8	2.6	3.7	11.0	4.3

A negative sign denotes precipitation and a positive sign denotes dissolution.
Carbonate denotes carbonate aquifer groundwater.
VR denotes Virgin River water
Cluster 3 is the groundwater samples only and does not include the Virgin River samples

3.5. Discussion and Conclusions

In this study, groundwater sources and movement in an intermountain basin is examined using an integrated and sequential approach to analyze the hydrochemistry of groundwater and surface water. A methodology of objective classification and sequencing of mutually supportive methods applied to the hydrochemical data of the lower Virgin River Basin allows overlay of evidences, which helps to precisely diagnose groundwater sources, movement, connectivity, and discharge in complex intermountain basins. Results of this study show the importance of objectively identifying hydrochemical facies as an essential step in the hydrochemical study of a basin hydrology. Additionally, this study confirms the conclusion in a previous study that hydrochemical data is objectively classified by combining cluster analysis with discriminant function analysis, and association measures.

The results show that the chemical character of the groundwater facies of the lower Virgin River Basin progressively evolves from Ca-HCO_3 type in the Clover Mountain front areas of the lower Virgin River Basin or in the outside basin areas north of the lower Virgin River basin to Na/Ca-SO_4 types in the middle and south basin-fill sediments of the lower Virgin River Basin, and along the Virgin River. The TDS of Na/Ca-SO_4 waters, which is higher than the TDS of Ca-HCO_3 type supports this interpretation. Again the TDS of the Na/Ca-SO_4 groundwaters supports that there are 3 groundwater sub-classes: the poorest water quality, localized and found in Bull Valley Mountains in Utah (cluster 6); the medium water quality, the commonest and wide spread water in the basin (cluster 4), and the best water quality found mostly along the Virgin River (cluster 5). Considering all the groundwater types and the Virgin River, the groundwater of cluster 3 and Virgin River have the poorest water quality with highest major ions and TDS. The progressive change in chloride or sodium ionic concentrations distinguishes and shows evolution of the hydrochemical facies. Results of Schoeller and discriminant functions plot indicate the groundwaters of clusters 1, 2, and carbonate aquifer are youngest, but the groundwaters of clusters 4, 5, and 6 are oldest. These conclusions are supported by the lower TDS, chloride and major ionic concentrations of clusters 1, 2, and the carbonate aquifer, and the higher TDS, chloride, and most major ions of clusters 4 and 5. In addition, the spatial pattern of the clusters reflects the variation in precipitation. In conclusion the Schoeller diagram and a discriminant functions (discriminant scores) plot are effective for defining source and discharge end-members.

The approach used to collect data on mineral phases interacting with groundwater and surface water is rapid and cost effective and is can be useful in studies where the mineralogy of the aquifer is not fully accessible. Synthesizing geologic information, saturation indices calculations, and major ion compositional plots permit gleaning of the mineral phases in the groundwater environment interacting with the groundwater. This approach is particularly important in the study area considering the complexity of the hydrogeology of the Basin and Range Province. To infer reactive minerals, all major ions used in compositional plots are in meq/l. Using meq/l provides clearer evidence of the relationship of the composition of water to the composition of solid minerals (Hem, 1989). Carbonate dissolution is dominant when HCO_3/SiO_2 ratio is >5, and for values <5 silicate weathering dominate is dominant. Therefore groundwater of cluster 2 and some of cluster 5 indicate strong silicate weathering, and the remaining clusters indicate strong carbonate weathering. Because of the low Mg/Mg+Ca ratios (<0.5) of groundwater of cluster 2 and 5, granitic weathering is possible. The Tertiary ashflow and volcaniclastic rocks predominant in the groundwater zone in the Clover Mountains and Escalante Desert Areas corroborate this inference. The high Na/Na+Cl ratios of groundwater of cluster 2 and some of 5 further suggest albite mineral phase reacting in addition to halite. Expectedly, some carbonate dissolves as Gypsum dissolution alone cannot account for the amount of Ca for the groundwater of cluster 2 and some of cluster 5. The water chemistry of cluster 3 (groundwater and Virgin River water) and cluster 6 (Figure 3. 9 to Figure 3. 12) indicate the dissolution of halite, gypsum, and limestone-dolomite or precipitation of calcite. The groundwater quality of cluster 1 potentially evolves from dissolution of halite, gypsum, and carbonate-dolomite or precipitation of calcite. Possible evolutions of the groundwater of cluster 4 are reaction with halite and gypsum, dissolution/precipitation of carbonate, and ion exchange. In addition to evidence of phases from the compositional plots and geologic information, saturation indices data also provide evidence that calcite, dolomite, and quartz are reactive phases in the groundwater and surface water environments.

Models calculated using an inverse geochemical code strongly supports interbasin groundwater flow from the Clover Valley Basin or the deep carbonate aquifer to the basin-fill aquifer of the lower Virgin River Basin. Considering a single replenishing source to the basin-fill aquifer, the inverse geochemical calculations indicate that the deep carbonate or the Clover Mountains groundwaters are more probable sources than the intrabasin groundwater of cluster 1. Considering mixing of 2 sources, the inverse modeling calculations indicates that at least 80% of the groundwaters in the basin-fill aquifer (clusters 4 or 5) are from interbasin groundwater flow. Thus mixing 13% cluster 7 and 87% cluster 2 and mixing 84% cluster 7 and 16% cluster 1 explain the groundwater of the basin-fill aquifer. Although the models do not provide conclusive answers, undeniably greater insights have been gained about the significance of the potential recharge sources to the lower Virgin River Basin.

Also the evolution of groundwater quality in the lower Virgin River Basin is explainable using the two best fit models found with the mole balance calculation. In the first model, 13% cluster 7 mixes with 87% cluster 2, and with dissolution of dolomite, halite, gypsum, and K-feldspar, and Na/Ca ion exchange. The net result of this reaction is the loss of Ca ions from the solution by ion exchange, addition of Na ions in solution by ion exchange, and the precipitation of calcite, quartz, and kaolinite. The dissolution of gypsum and dolomite while calcite precipitates suggests a dedolomitization process. In the second model, 87% cluster 7 mixes with 16% cluster 1, and with dissolves halite, albite, and gypsum, and Ca/Na ions exchange. The net result of this reaction is the loss of Na ions from the solution, addition of Ca ions to the solution, and the precipitation of dolomite, calcite, quartz, and kaolinite. Regarding groundwater interactions with the Virgin River, the model calculation of groundwater of cluster 3 (only groundwater samples in the floodplain aquifer), flow to the Virgin River indicates the dissolution of calcite and halite, precipitation of dolomite and anhydrite, and the loss of Na from solution through Na/Ca ion exchange. The precipitation of dolomite and the corresponding increase in the Mg/Ca ratio of the solution suggest a dolomitization.

In conclusion a new approach for analyzing hydrochemical data to test hypotheses about sources and movement of water in basins has been presented. In this approach, discriminant function analysis was used to model hydrochemical data, using cluster delineated groups as a training dataset. This approach is objective, allowing groups to be repeatedly defined. Furthermore the discriminant functions scores were plotted and used to identify hydrochemical end-members. This approach removes the problem of incongruence that can arise when PCA is used to identify end-members for cluster delineated groups. In addition, this methodology allows additional independent investigations to be overlaid for a more precisely testing of hypotheses. For example an aqueous geochemical modeling was applied to the hydrochemical facies datasets. We observed that using Box-Cox and box-and-whiskers plots help in finding the best power transformation values for normalizing data and minimizing the outliers in data for a better classification result. From the analysis of the current data collected we conclude that interbasin groundwater significant source of recharge to the lower Virgin River Basin accounting for at least 80% of the basin’s water budget. The methodology described can be applied to study flow systems in other intermountain basins.

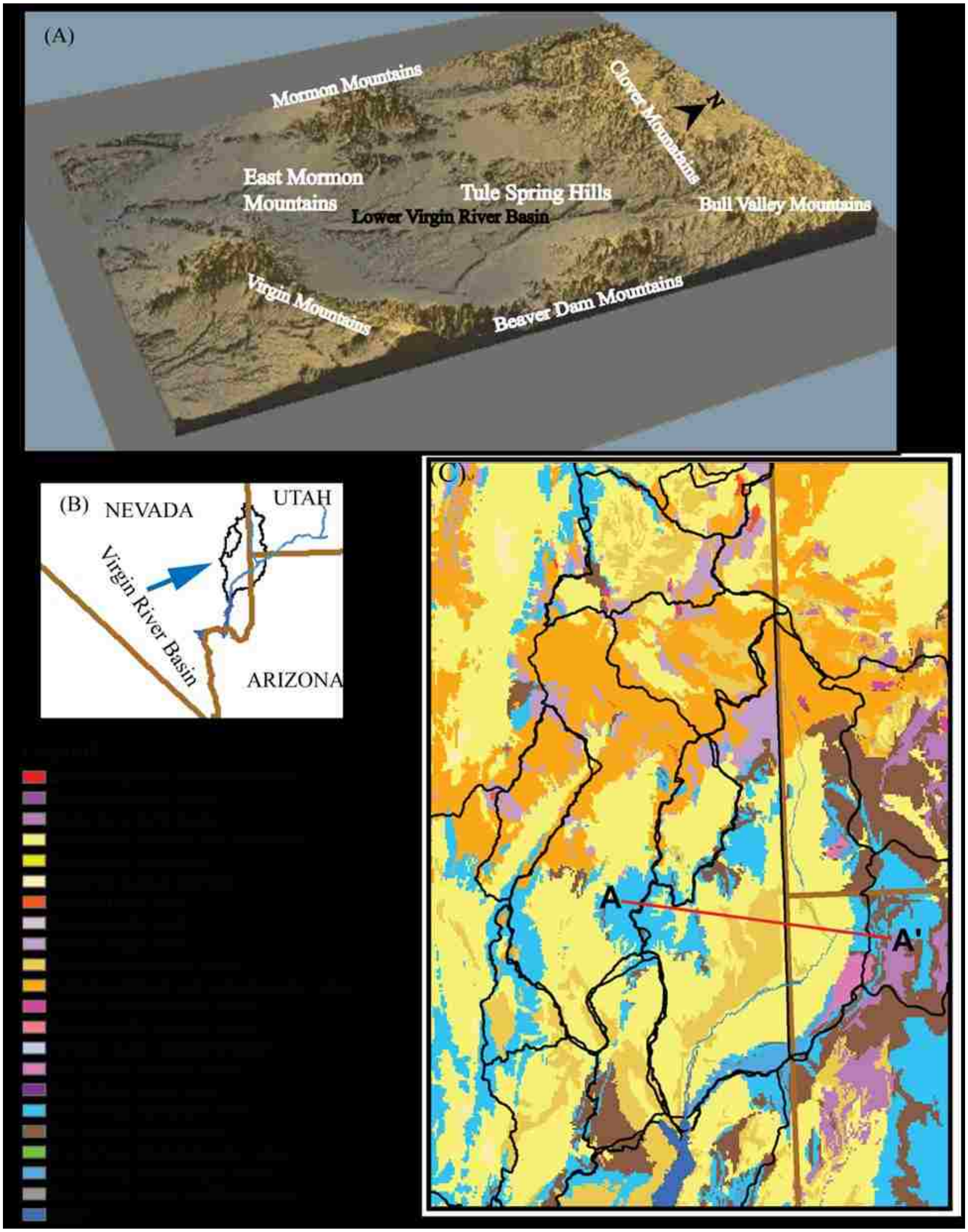

Figure 3. 1.Physiography (A), Location (B), and simplified geology map (C) of the lower Virgin River Basin and the adjacent basin areas. The geologic map is modified after GBCGE (2005)

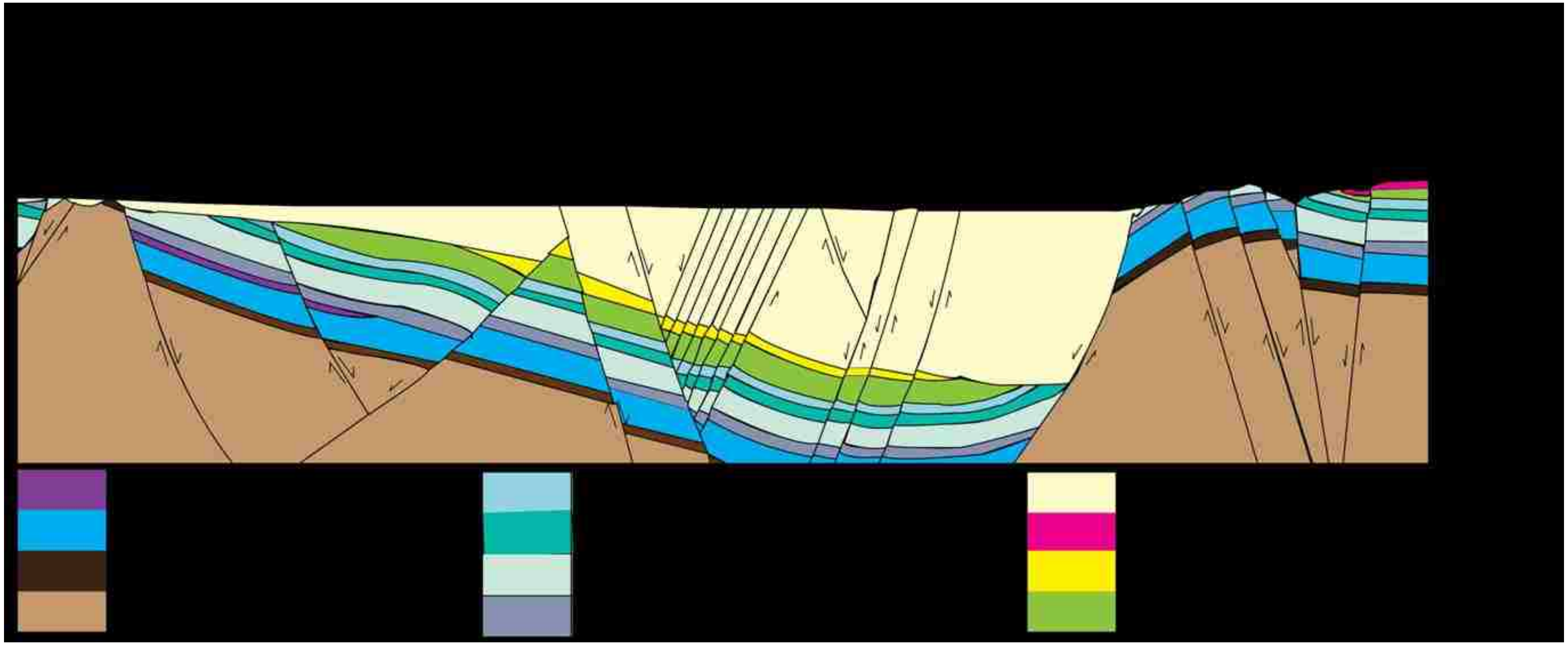

Figure 3. 2. Geologic cross-section along the line A—A' on the geology map. Modified after Page et al. (2006)

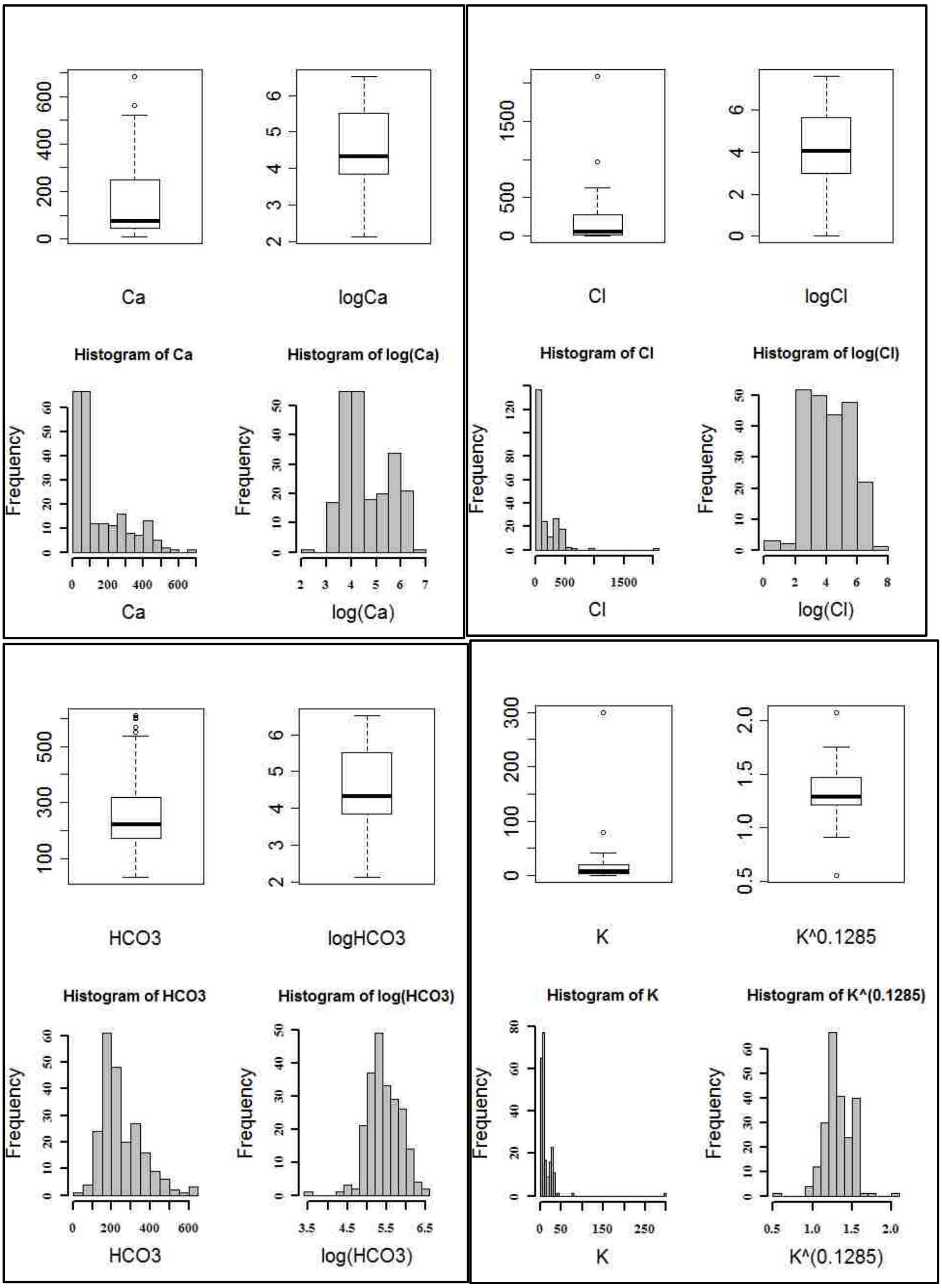

Figure 3. 3. Statistical distribution of the raw and transformed variables in mg/L

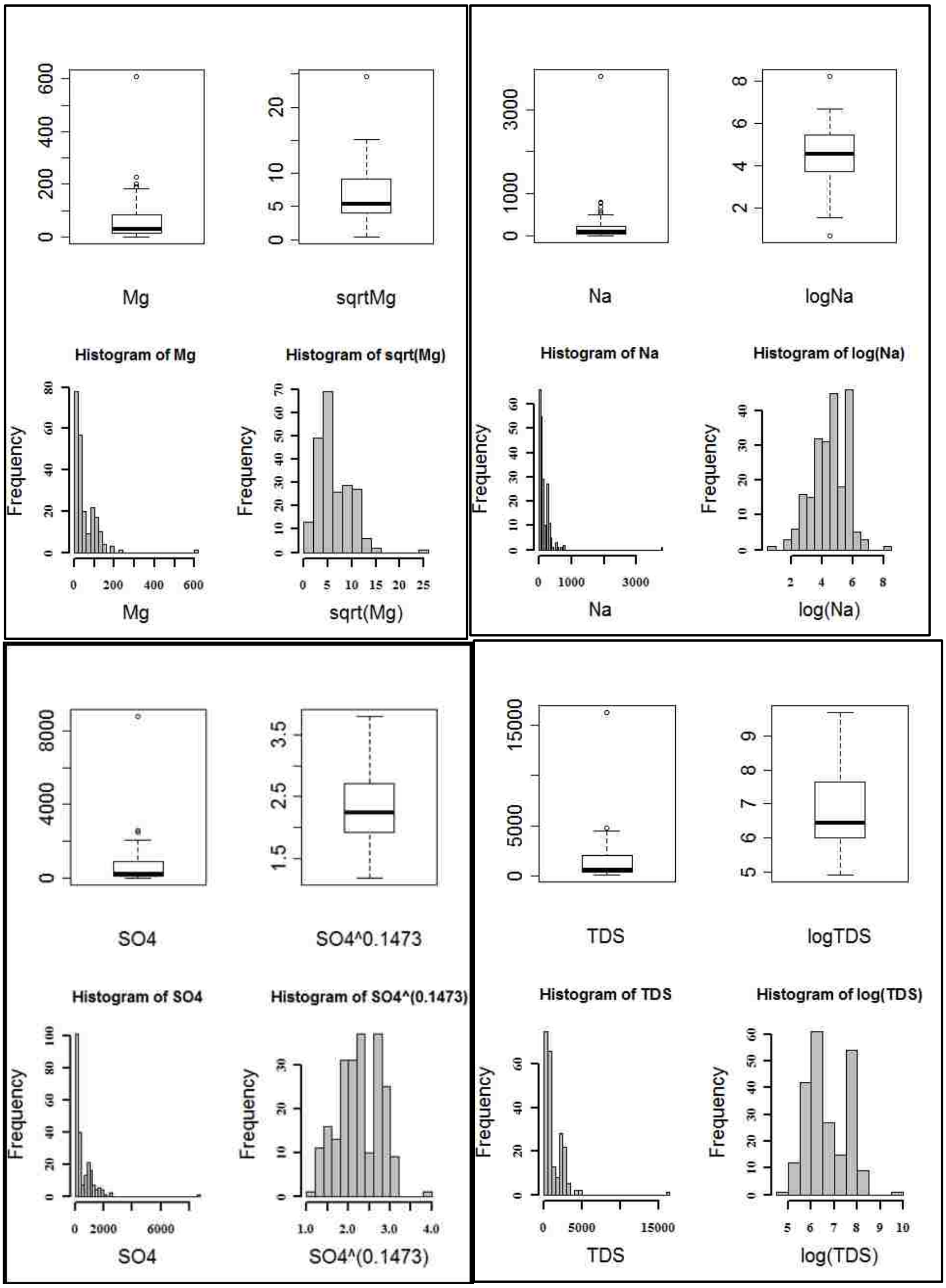

Figure 3. 4. Statistical distribution of the raw and transformed variables (mg/L)

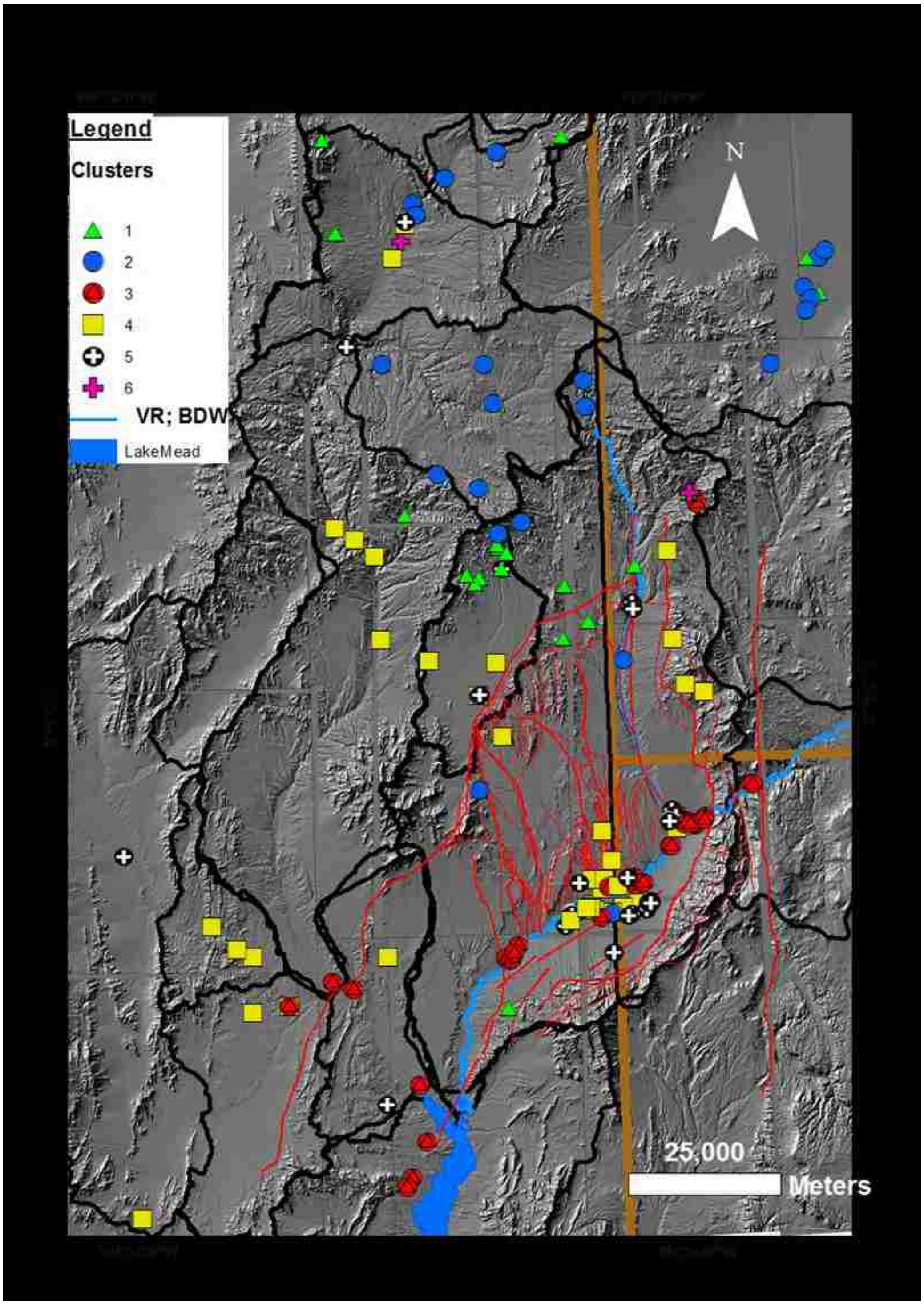

Figure 3. 5. Spatially distributed hydrochemical facies of the lower Virgin River Basin and adjacent basins. The lines running roughly north-south are fault lines (Dixon and Katzer 2002; Page et al., 2005). Also Virgin River (VR) and Beaver Dam Wash (BDW) shown on the map.

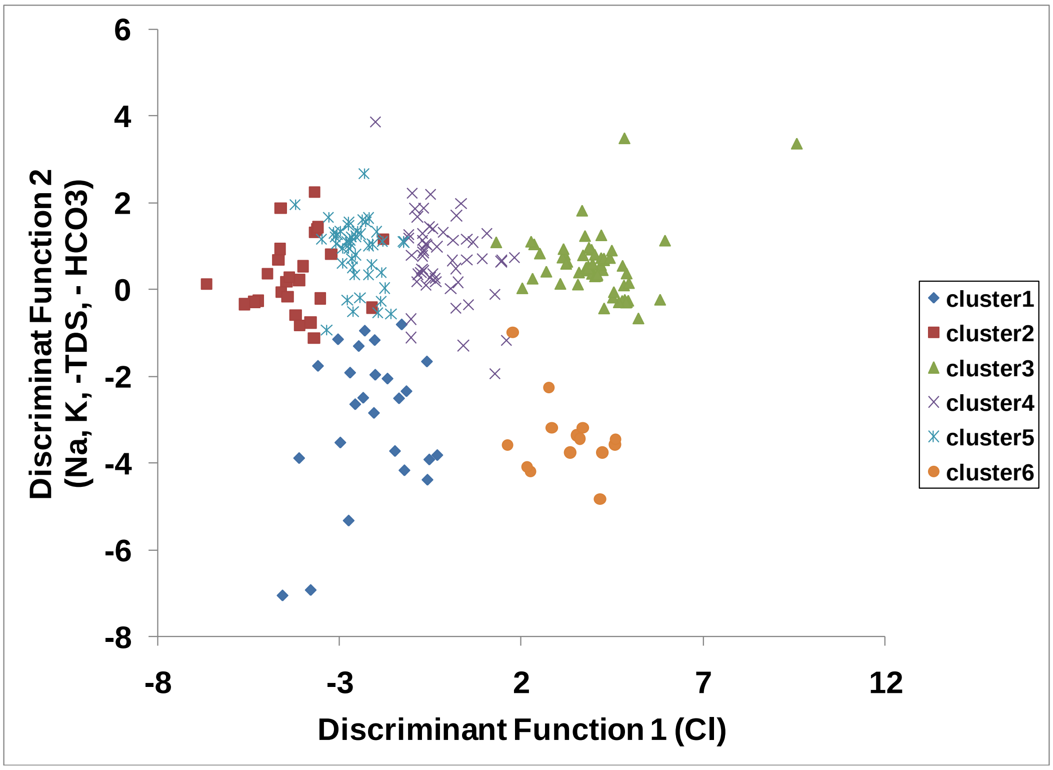

Figure 3. 6. Discriminant functions plot of the discriminant function analysis modeled hydrochemical facies

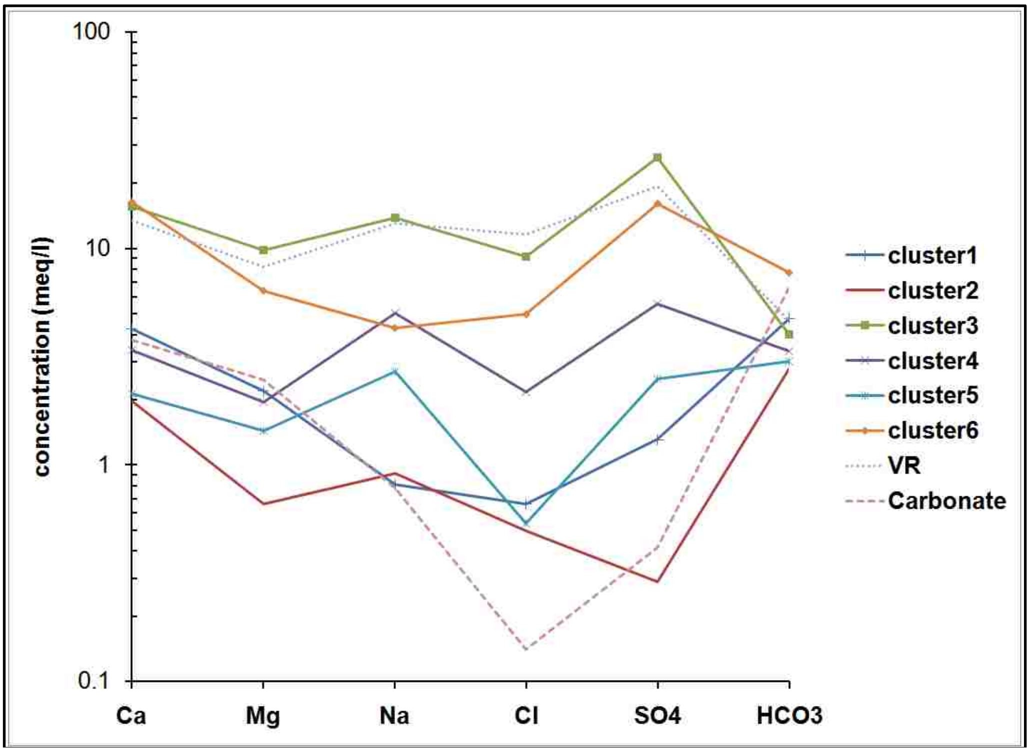

Figure 3. 7. Schoeller diagram of mean (back transformed mean of transformed data) major ionic values of the hydrochemical facies modeled. Cluster 3 is only the groundwater samples. Carbonate is the carbonate rock well sample (station name USAF TEST WELL) collected from the carbonate aquifer (730 m deep) underlying Dry Lake Valley located north of lower Virgin River Basin (Bunch and Harrill, 1984).

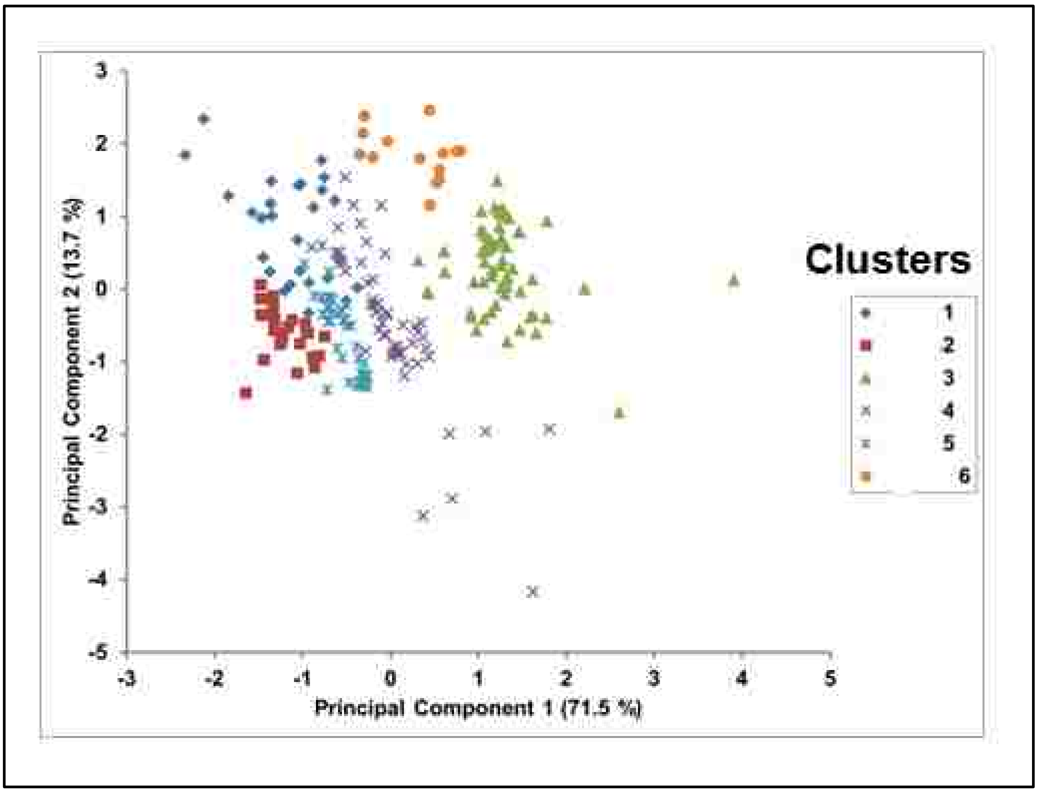

Figure 3. 8. Principal Components plot with superimposed cluster groups

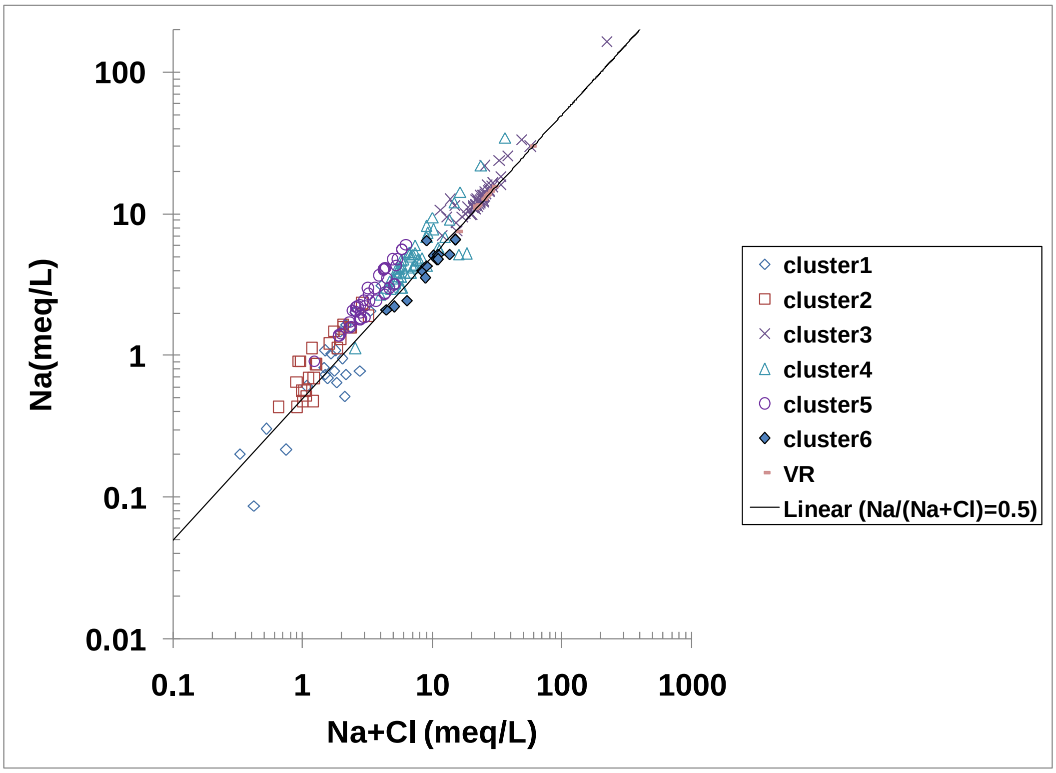

Figure 3. 9. Na/Na+ Cl ratio compositional diagram of clusters.

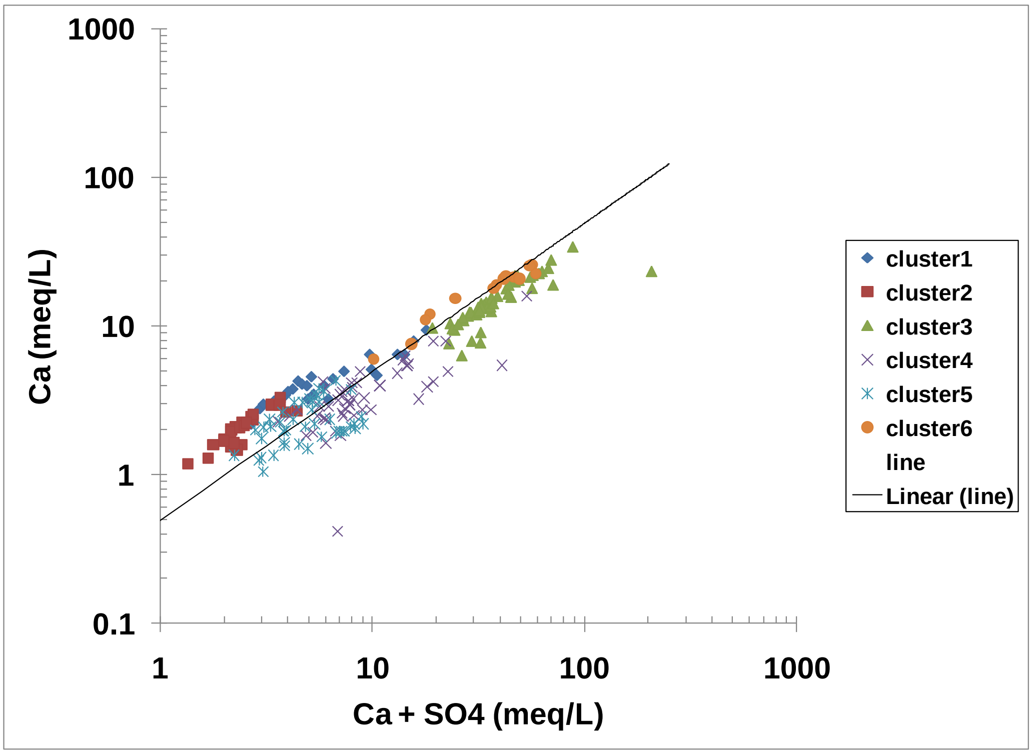

Figure 3. 10. Ca/Ca+SO4 ratio compositional diagram of clusters. Gradient of the line =0.5

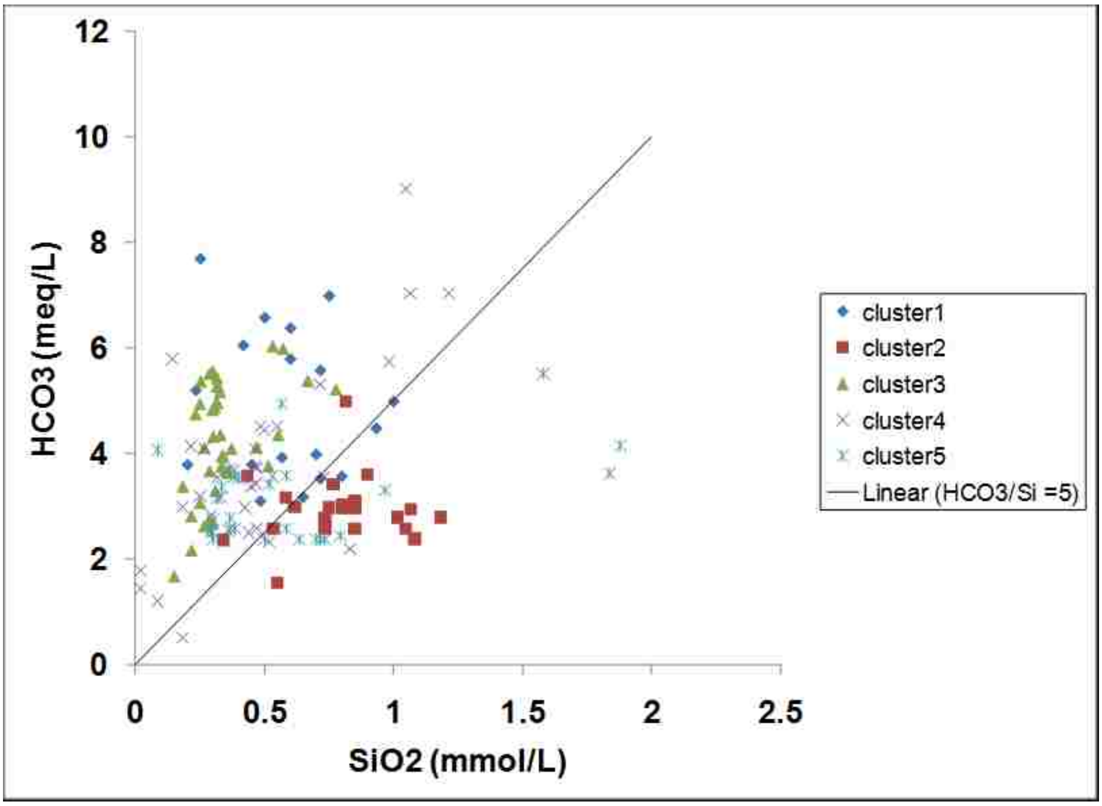

Figure 3. 11. HCO3/SiO2 ratio compositional diagram of clusters

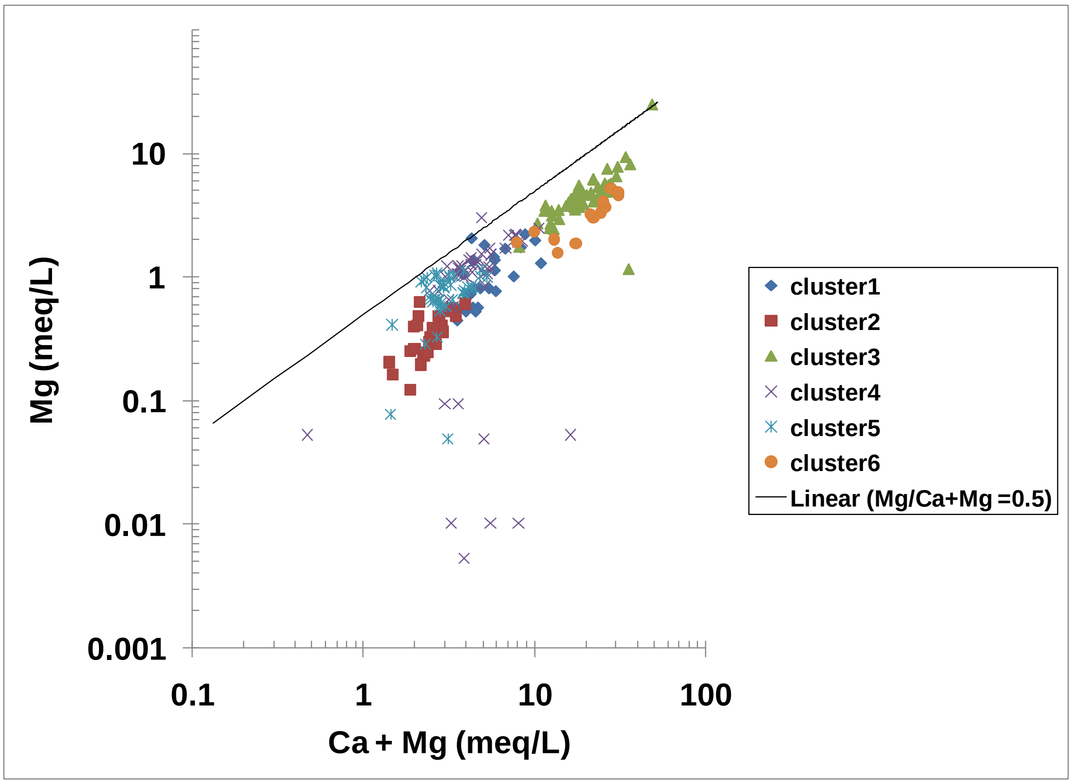

Figure 3. 12. Mg/Ca ratio compositional diagram of clusters

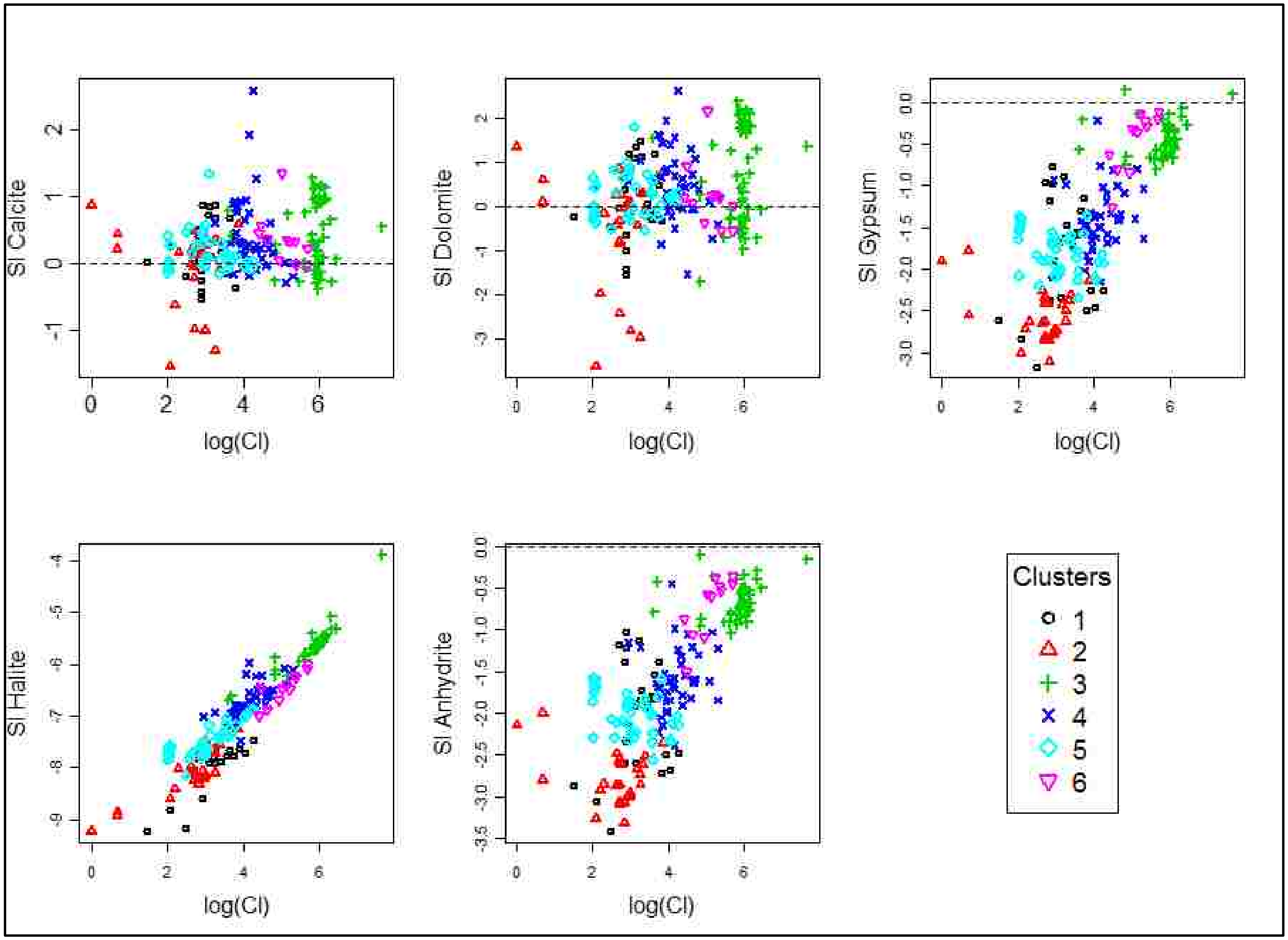

Figure 3. 13. Comparison of saturation indices (SI) of calcite, dolomite, gypsum, anhydrite, and halite, and comparison of clusters based on saturation index versus log chloride concentration.

CHAPTER 4

EVALUATING INTERBASIN FLOW AND GROUNDWATER-SURFACE WATER INTERACTIONS: UNUSUAL STATISTICAL AND GRAPHICAL ANALYSIS OF STABLE ISOTOPIC AND SOLUTES DATA

Abstract

Rocks and basin-fill sediments of the intermountain basins in the Great Basin region are complexly folded and faulted and, as a result, understanding the flow systems is challenging. In this study, stable isotopic (δD and $\delta^{18}O$) and selected chemical solutes (Na, K, SO_4, and Cl) data from waters in the lower Virgin River Basin and adjacent basin areas are used to test the hypotheses of interbasin groundwater flow and the Virgin River interactions with groundwater. Unconventionally, linear regression, Spearman correlation test, scatterplot, box-and-whisker plot, and Wilcoxon Rank Sum test are applied to glean hydrogeologic information from the stable isotopic and solutes data of the hydrochemical facies for the study area. For the hydrochemical facies, δD is regressed over $\delta^{18}O$, the correlation of δD and $\delta^{18}O$ versus Cl is tested, and the distribution of stable isotopic and solutes data is analyzed using box-and-whiskers plots and Wilcoxon Rank Sum Test. The results indicate the Virgin River with floodplain aquifer is more homogeneous and evaporated than groundwater samples designated source and discharge end-members. Regional springs representative of the regional carbonate groundwater, a potential replenishing source to the basin-fill aquifer of the lower Virgin River Basin, have a median δD of -109.0‰. Possible sources besides the carbonate water are intrabasin and adjacent extra-basin waters with median δD of -87.0‰ and -93.0‰ respectively. The

Virgin River with floodplain aquifer has median δD of -93.0‰, and of the discharging groundwaters have median δD of -98.25‰ and -101.50‰. Mixing calculations show that mixing of 51-66% carbonate aquifer groundwater and 34-49% intrabasin recharge account for the discharging waters. Again mixing calculations show that mixing of 33-53% carbonate groundwater and 47-67% interbasin groundwater from the Clover or Escalante Desert Valley account for the discharging waters. In conclusion, at least 50% of the groundwater in the basin-fill of the lower Virgin River Basin is from interbasin sources.

4.1. Introduction

Parts of the southwestern United States of America have limited surface water resources; therefore groundwater is a vital resource. A large part of the Southwest is within the Great Basin region. The general physiographic features of the Great Basin are basin-fill sediments and intervening mountain ranges. Generally, the basin-fill deposits are the most important water supply source (Heath, 1984; Johnson et al., 2002). The approximate area of the Alluvial Basins region is 1,025,000 km^2, stretching from the Puget Sound-Williamette Valley, Washington and Oregon, to West Texas (Heath, 1984). The climate in parts of the Great Basin ranges from semi-arid to arid (Heath, 1984) and the natural recharge is low (U.S. Geological Survey, 2000). Understanding the flow systems of these basins is critical for sustainably managing these vital groundwater and surface water resources.

Interbasin flow is an important potential source of recharge to the basin-fill aquifers of the intermontane basins. This is a concept by which groundwater from one basin moves through permeable intervening mountains to another basin, or by which regional water in the deep underlying Paleozoic carbonate rock moves to the basin-fill aquifers. Additional sources of recharge are local infiltrating precipitation in the mountain blocks or mountain front adjacent to the alluvial valleys and infiltrating streams/rivers originating in the adjacent mountains. Studying the sources, movement, and quantity of recharge to the basin-fill aquifers is a prolific research area (Eakin, 1966; Winograd and Friedman, 1972; Winograd and Thordarson, 1975; Mifflin and Hess, 1979; Johannesson et al., 1997; Davisson et al., 1999b; Thyne et al., 1999; Genereux, 2005; Hibbs and Darling, 2005; Anderson et al., 2006; Guler and Thyne, 2006; Belcher et al., 2009; Bushman et al., 2010; Hershey et al., 2010). The abundant research has been propelled by the need to sustainably manage groundwater of the Great Basin region to prevent causing unacceptable environmental, economic, and social consequences (Alley et al., 1999).

Testing the hypothesis of interbasin flow is a challenge for hydrologists (Davisson et al., 1999a; Davisson et al., 1999b; Thomas, 1999; Winograd, 2001). Interbasin flow has been tested in many ways, including using hydrochemical, isotopic, and hydrogeologic data, and water budget analysis. Basins hydrologic budgets and spring discharge rates higher than can be supported by local recharge alone have also been used to support the idea of interbasin groundwater flow occurring (Eakin, 1966). However, in arid regions, the water budget approach is useful only for preliminary studies because the empirical approaches for calculating recharge are uncertain, and as a result, recharge calculations become questionable.

Using limited water level data, hydraulic gradients of potentiometric surface of the deep carbonate groundwater have been used to interpret regional movement of groundwater in the Paleozoic carbonate aquifer in the southern Great Basin (Winograd and Thordarson, 1975). Additionally, the regional carbonate groundwater has been characterized using physical and chemical quality and radio-isotopic compositions of water discharged by major springs (Mifflin and Hess, 1979; Hershey et al., 2010). Previous studies reported that for the Great Basin region, groundwater concentrations of Na, K, SO_4, and Cl increase with groundwater travel distance (Mifflin and Hess, 1979; Hershey et al., 2010). This conclusion forms part of the reason for selecting these solutes for the current study. Mifflin (1979) used bomb-pulse tritium to draw the conclusion; however, because of the short half-life for tritium, concentrations in the field are now very low for hydrological studies. Conventionally, stable isotopic testing of interbasin flow has been performed by comparing the stable isotopic values (δD, $\delta^{18}O$) of precipitation in a basin to the stable isotopic values of groundwater in the basin, correlating contours of stable isotopic values of groundwater, or comparing the position of water groups and their relationships to the global meteoric waterline (GMWL) on a δD-$\delta^{18}O$ plot (Davisson et al., 1999b; Smith et al., 2002; Anderson et al., 2006; Guler and Thyne, 2006; Belcher et al., 2009; Bushman et al., 2010; Hershey et al., 2010). Less commonly, for a more preliminary approach, deuterium has been used to calibrate conceptual groundwater flow model covering 13 topographic basins of the White River Flow Systems (Thomas et al., 2001). Also correlation of the distribution of deuterium in geographic areas and mixing calculations have been to support the hypothesis that the

Ash Meadow springs are fed by recharge from the Spring Mountains-Sheep Range and underflow from Pahranagat Valley (Winograd and Friedman, 1972).

Watersheds in the Great Basin have diverse characteristic flow systems, which demand evaluating individual basins. In the Moapa Valley in southern Nevada, the Muddy River Springs discharge through the basin-fill aquifer has been interpreted as interbasin groundwater from the lower Meadow Valley (Kirk and Campana, 1990) or the Spring Mountains (Winograd and Friedman, 1972). Also, the Smith Creek Valley in central Nevada is believed to be a closed basin (Thomas et al., 1989). Conversely, for the Ash Meadows Valley in southern Nevada, the alluvial aquifer is believed to be hydraulically connected to the carbonate aquifer (Mifflin and Hess, 1979; Bushman et al., 2010).

Similar to other intermountain basins, a number of sources of recharge have been proposed for the basin-fill aquifer in the lower Virgin River Basin. Firstly, groundwater flow conditions at the boundaries north of the lower Virgin River Basin are considered as no flow, Neumann’s-type boundary condition (Glancy and Van Denburgh, 1969; Baron, 2007). Consequently, significant groundwater recharge to the basin-fill aquifer of the lower Virgin River Basin is assumed to be supplied only from precipitation in the adjacent mountains, which percolates the mountain block or the mountain front. Second, it is posited that groundwater transmits through the Clover, Beaver Dam, and Bull Valley Mountains bordering north and northeast of the lower Virgin River Basin to the lower Virgin River Basin (CH2M HILL, 2002). The interbasin transfer of water may be facilitated by faults and fractures in the volcanic rock of the Clover Mountains and permeability of the carbonate rocks lying deeply beneath. Much of the groundwater in the basin-fill aquifer of the lower Virgin River Basin is believed to be supplied from the deep carbonate aquifer beneath or the adjacent basins north (Dixon and Katzer, 2002; Johnson et al., 2002).

Although stable isotopic data can provide good evidence of flow systems, due to intermingling of the hydrochemical facies in the Euclidean (metric) space at times, it is difficult to extract flow dynamics information using the conventional approach of analyzing these data. For this study, an unusual and robust approach is used to evaluate the interbasin flow and groundwater-surface water interactions in the lower Virgin River Basin. Interbasin flow and groundwater-surface water interactions are tested by analyzing and interpreting the stable isotopic and solutes data (δD, $\delta^{18}O$, Na, K, Cl, and SO_4). Specifically I tested the hypothesis that interbasin groundwater does not contribute to groundwater in the basin-fill aquifer of the lower Virgin River Basin and also the hypothesis that the Virgin River is hydraulically interconnected with the floodplain aquifer. It was predicted that the stable isotopic values of the discharging groundwaters in the lower Virgin River Basin is statistically similar to the source groundwaters (intrabasin recharge waters). Also, I predicted that the Virgin River and the floodplain aquifer have statistically the same stable isotopic values. Unconventionally, pieces of information of hydrologic processes were gleaned from the data using linear regression, scatterplot, box-and-whiskers plot, Spearman correlation test, and non-parametric Wilcoxon Rank Sum test. We assume the stable isotopic compositions are conservative and chemical reactions that can sink Na, K, Cl, and SO_4 are insignificant. The interpretations were enhanced by the available hydrologic and geologic information.

4.2. Description of the Study Area

The lower Virgin River Basin (Figure 4. 1B), covering an area of about 4,463 km^2, is located in the Basin and Range physiographic province of the southwestern United States. As typical of the intermountain basins, mountains created by east-west extension during the Cenozoic (Page et al., 2006) separate the lower Virgin River Basin (Figure 4. 1A) from the adjacent basins.

The mountain ranges and basement rocks consist of pre-Tertiary rocks and are overlain with Tertiary and Quaternary rocks. Figure 4. 1C is a simplified geology map (GBCGE, 2005) showing the spatial distribution of rocks in the study area. In cross-section, Figure 4. 2 illustrates the lithologic units of the study area, ranging from Early Proterozoic to Quaternary. The following discussion of the lithologic units and their hydrogeological importance are a summary from Page et al. (2005) and Dixon and Katzer (2002). Early Proterozoic crystalline basement rocks consisting of gneiss, schist, and granite have low permeability and are considered a groundwater no flow boundary. Late Cambrian and Late Proterozoic rocks are mostly low permeability quartzite, conglomerate, sandstone, siltstone, and shale. Middle Cambrian to lower Permian rocks consisting mostly of limestone and dolostone, are considered a regional aquifer. Mesozoic rocks, mostly consisting of low permeable conglomerate, sandstone, siltstone, mudstone, shale, and gypsum, are considered as confining units. Quaternary to Tertiary rocks are alluvium and colluvium sediments, and semi-consolidated to consolidated sedimentary rocks. Tertiary volcanic rocks are ash-flow tuffs from the Caliente caldera in the north-west area of the lower Virgin River Basin. The Tertiary rocks of water resource importance are named the Muddy Creek Formation. The Muddy Creek Formation is mostly conglomerate, tuffaceous sandstone, gypsum, halite, and mudstone.

The climate in the study area is semi-arid to arid. The topography, seasons, and geology of the study area influence the distribution of precipitation and recharge in the lower Virgin River Basin. Rainfall distribution is highly orographic - the highest elevations in the basin at about 2438 m receive annual precipitation of about 609 mm; in contrast the basin floor receives precipitation of about 152 mm (Dixon and Katzer, 2002). The annual reference evapotranspiration on the basin floor is about 11 times greater than the annual precipitation (CEMP, 2011). Winter precipitation dominates the study area; in contrast, summer precipitation is short duration and localized (Glancy and Van Denburgh, 1969; Winograd et al., 1998). Additionally, temperatures range in the summertime from highs near 40 °C, to wintertime lows below freezing. Therefore winter precipitation is considered the most important source of recharge (Friedman et al., 2002). Two perennial rivers, the Beaver Dam Wash and the Virgin River, drain the lower Virgin River Basin. Potential groundwater recharge sources in the lower Virgin River Basin are precipitation in the higher elevations surrounding the basin, interbasin groundwater, and stream/river infiltration (Dixon and Katzer, 2002). As shown in Figure 4. 2, the rock units are cut through by abundant Cenozoic faulting. Cenozoic faults are believed to be the main conduit for movement of groundwater in the study area (Page et al., 2006).

4.3. Materials and Methods

This study examined interbasin flow and groundwater-surface water interactions by analyzing the chemical solutes (Na, K, SO4, and Cl) and stable isotopic values (δD, and δ^{18}O) for six hydrochemical facies of the lower Virgin River Basin (Chapter 3). The chemical solutes and stable isotopic data used in this study are accessible at https://spreadsheets.google.com/pub?key=0AkihMREl9PhWdEVDMnIyakRFRUhlYmNiblJKMXZwUWc&hl=en&single=true&gid=0&output=html. The database comprised data from governmental and institutional reports (Glancy and Van Denburgh, 1969; Bateman, 1976; Brothers et al., 1993; Enright, 1996; Thomas et al., 2001; CH2M HILL, 2002); graduate student theses (Metcalf, 1995; Yelken, 1996); the U.S. EPA database STORET (U.S EPA, 2009); the U.S. Geological Survey database NWISWeb (U.S Geological Survey, 2009); and sampling activities conducted during this study by the Virgin Valley Water District (VVWD), Mesquite and the University of Nevada Las Vegas (UNLV). The database details including the sources of data and evaluation of the analytical data are described in Chapter 2. The database contains 328 samples and 40 variables. The standard methods for collecting the water samples and analyzing the hydrochemical and stable isotopic compositions in the water are described in several sources in the hydrochemistry literature (Clark and Fritz, 1997; American Public Health Association, 2005) and are not described here.

Six hydrochemical facies (clusters) are analyzed in this study. The details of the procedures on how the six hydrochemical facies are delineated are found in Chapter 2. The class sizes of the clusters 1 through 6 are 25, 26, 59, 53, 45, and 14 respectively. Two hundred and twenty two groundwater (wells and springs) and surface water samples were partitioned into the six hydrochemical facies. There were 198 groundwater samples and 24 surface water samples. All 222 samples have measurements for the chemical solutes (Na, K, SO4, and Cl), but only 123 samples have measurements for δD and δ^{18}O pair.

The spatial distribution and the chemical characteristics of the hydrochemical facies, and geology and hydrology information were used to help in interpreting the stable isotopic values of the facies.

The chemical solutes were selected for tracing interbasin flow and interactions between groundwater and surface water because, as mentioned previously, in the Great Basin region, the concentrations of these solutes increase with longer flowpath and residence time (Mifflin and Hess, 1979; Hershey et al., 2010). In this study, it is assumed chemical reactions that can sink Na, K, Cl, and SO_4 are insignificant. Generally, chloride accumulates and remains in solution, as it is not affected by many of the chemical and biological reactions that affect other major ions. But sulfate reduction can reduce the concentration of sulfate with corresponding increase in the concentration of bicarbonate. This reaction is not yet known to occur in the study area. Similarly, potassium may be taken up by plants but for this study it is assumed that is insignificant.

The stable isotopic value measured is the enrichment composition of the isotopic ratios ($^2H/^1H$ and O^{18}/O^{16}) in water relative to the composition of the isotopic ratio in the Vienna Standard Mean Ocean Water (VSMOW). The stable isotopic value is expressed by the δ notation and in unit of permil (‰) (Clark and Fritz, 1997):

$$\delta = \left(\frac{R_{sample}}{R_{s\tan dard}} - 1 \right) \mathrm{X}100 \qquad (1)$$

where R_{sample} and $R_{standard}$ are the isotopic ratios of the water sample and VSMOW respectively.

The stable isotopic values in the hydrological cycle exhibit distinctive spatial and temporal variation, which account for their use is tracing the hydrological cycle. The variation of the isotopic values in precipitation is due to fractionation accompanying the phase changes in the water cycle (Gibson et al., 2005). Fractionation occurs in all condensation processes and in the evaporation of well mixed liquid water in the hydrological cycle because the volatility of the lighter isotopes (^{1}H and ^{16}O) is much higher than that of the heavy isotopes (^{18}O and ^{2}H) (Dansgaard, 1964). In general, fractionation of the isotopic composition of precipitation is controlled by a number of processes including temperature (altitude, latitude), seasonal variation, distance from vapor source and evaporation (Dansgaard, 1964). But in the groundwater saturated zone fractionation of the stable isotopes of water generally does not occur. Therefore δ^{18}O and δD of groundwater commonly behave conservatively in the saturated zone. In rare cases the stable isotopes of oxygen and hydrogen may exchange with rocks (particularly carbonate) and gases (H_2S, CO_2) in the saturated zone and the isotopic exchange rate increases in geothermal environments. Also in shallow groundwater conditions, evaporation can fractionate the stable isotopes of water. In the study area, the desert region of the southwestern United States, high stable isotope values empirically correlates with warmer lower elevation precipitation, but low stable isotopic values empirically correlates with cooler high elevation precipitation (Blasch and Bryson, 2007). In this study, the stable isotopic values of δD and δ^{18}O of are used to trace interbasin groundwater flow to the lower Virgin River Basin and interactions between groundwater and the Virgin River.

4.3.1. Statistical and graphical analyses of stable isotopic and solutes data

The stable isotopic data of the hydrochemical facies (six clusters) are analyzed using histograms, bivariate and linear/non-linear regression analysis, Spearman correlation coefficient, box-and-whisker plots, and Wilcoxon Rank Sum test. The solutes (Na, K, SO_4 and Cl) are analyzed using box-and-whisker plot. Using these analytical techniques the hydrogeologic environment and the recharge, mixing, and residence time information are extracted from the hydrochemical facies. The results are interpreted to support or reject the hypotheses of interbasin groundwater flow and the interactions between the groundwater and surface water.

All the statistical analyses are performed using the statistical program R (R Development Core Team, 2010). A histogram is used to analyze the distribution of the δD and $\delta^{18}O$ data to study the hydrogeologic conditions in the study area. For each cluster, the δD is plotted against $\delta^{18}O$, and δD regressed over $\delta^{18}O$ to characterize the recharge and mixing conditions of the distinct clusters. The method of least-squares is used to fit the best regression line to the sub-datasets and to test the hypotheses about the parameters of the linear regression model fitted. The hypotheses of the parameters are tested at the 95% confidence level. Because of the sensitivity of the regression model to outliers, outliers and outlying points influencing the regression were checked using leverage (hat) and Cook's distance values. A point with hat values greater than $2 * p / n$ ($p = 2$ (number of variables) and n = number of samples/cases) was considered a potential outlier and removed if Cook's distance value is greater than or equal 1 (Logan and Logan, 2010). The gradient and the intercept parameters of the regression equation are estimated such that the sum of residuals is the minimum. The box-and-whiskers plots of the solutes (Na, K, SO_4, and Cl) and the stable isotopic values ($\delta^{18}O$ and δD) are used to trace the movement of subsurface water to and within the lower Virgin River Basin. The Wilcoxon Rank Sum test (also called Mann-Whitney test) is used to test whether two hydrochemical facies proposed to be hydraulically connected are statistically ‘truly’ connected by flowpath.

The Wilcoxon Rank Sum test is non-parametric and so does not require the data of the two hydrochemical facies to be normally distributed. The Wilcoxon test examines whether the two hydrochemical facies come from the same population. Two hydrochemical facies are from different population if the resulting p-value for the Wilcoxon test is less than 0.05 (Reimann et al., 2008). Alternatively, the two hydrochemical facies are from the same population if the resulting p-value for the Wilcoxon test is equal to or greater than 0.05 (Reimann et al., 2008).

4.4. Results

4.4.1. Hydrochemical facies

As shown in Figure 4. 3, the study area is characterized, on the basis of the major ion chemistry, into six hydrochemical facies (Chapter 3). The hydrochemical facies superimposed on a digital elevation model (DEM) reveal that the facies are coherent in the geographic space. On the basis of the chemical character and spatial distribution of the hydrochemical facies (clusters) as well as the information about the distribution and amount of rainfall received in the study area, the facies are characterized as source or discharge end member waters (Chapter 3). Facies 1 and 2 are characterized as source end-member, facies 4, 5, and 6 are characterized as discharge end member, and facies 3 consist of the floodplain groundwater discharge and the Virgin River water. The chemical character of the major groundwater types are described as Ca-Mg-HCO_3 (clusters 1), Ca-

$Na-Mg-HCO_3$ (cluster 2), $Ca-Na-Mg-SO_4-Cl$ (cluster 3 and Virgin River), $Na-Ca-SO_4-HCO_3$ (clusters 4 and 5), and $Ca-Mg-SO_4-HCO_3$ (cluster6).

4.4.2. Stable Isotopic Data

Distribution of isotopic values of the study area is non-normal, unimodal, and has a large range. For the pairs of δD and $\delta^{18}O$ data of the groundwater and surface water (n =123), the δD values range from -109‰ to -70‰ and the $\delta^{18}O$ values range from -14.6‰ to -6.7‰ (Figure 4. 4). The wide variation of the stable isotopic values suggests fracture controlled groundwater flow with limited mixing. Also the mean δD distribution is influenced by a few very low δD values (median = -94.00, mean = -94.49) and mean $\delta^{18}O$ distribution is influenced a few very high $\delta^{18}O$ values (median = -12.80, mean = -12.49).

The bivariate δD - $\delta^{18}O$ plot (Figure 4. 5) shows that generally groundwaters of clusters 1 and 2 plot along the Craig's meteoric waterline (Craig, 1961) commonly called the global meteoric waterline (GMWL); however, groundwaters of clusters 3, 4, 5, and 6 plot to the right of the meteoric line (Figure 4. 5). On the δD - $\delta^{18}O$ plot, the clusters are intermingled so that it becomes difficult to differentiate the clusters. However, differing patterns emerge when the δD is regressed on $\delta^{18}O$ for each cluster (Figure 4. 5). Note that in cluster 6, only one sample had stable isotopic values; therefore cluster 6 has no regression line. The regression coefficients with their respective errors and statistically significant values for the individual hydrochemical groups are shown in Table 4.1 . Coefficients estimated for the models are reliable as the sub-datasets of the clusters contain no influential data points (Cook's distances < 1). The gradients of the regressions of all the clusters are lower than the gradient of the meteoric water line (mathematically:

$\delta D = \delta^{18}O + 10$). Clusters 2 and 3 have the largest and the lowest gradients respectively. Statistically the gradients of groundwaters of clusters 4 and 5 are the same, but different from the gradients of groundwaters of clusters 1, 2, and 3. Clusters 2, 4, and 5 appear diverging from a common point on the GMWL. The regression lines cross the GMWL at different points and in studies elsewhere the δD and $\delta^{18}O$ values of the intersection point has been interpreted as the average of the isotopic values of the precipitation recharging the groundwater (Clark et al., 2000). Cluster 1 and the GMWL meet at δD = -91.63 ‰ and $\delta^{18}O$ = -12.70 ‰. Cluster 2 intersects the GMWL at δD = - 108.67 ‰ and $\delta^{18}O$ = - 14.83 ‰. Cluster 3 crossed the GMWL at δD = -99.13 ‰ and $\delta^{18}O$ = -13.64 ‰ and cluster 4 crossed the GMWL at δD = -108.86‰ and $\delta^{18}O$ = -14.86 ‰. The cluster 5 intersects the GMWL at δD = -110‰ and $\delta^{18}O$ = -15‰. The point where the cluster 1 intersects the GMWL has the most enriched stable isotopic values. Progressively, the isotopic values of the intersected points decrease in the order of cluster 1, cluster 3, cluster 2, cluster 4, and cluster 5. Thus the intersection point of cluster 5 has the most depleted isotopic values.

The relationships between δD versus Cl and $\delta^{18}O$ versus Cl for the clusters were analyzed using the scatterplot function in the 'car' package and the Spearman rho (ρ) correlation test function (significant level $p < 0.05$). From Figure 4. 6, there is no significant trend between δD and Cl for cluster 1 ($\rho = 0.245$, $p = 0.379$), clusters 2 ($\rho = -0.069$, $p = 0.778$), cluster 4 ($\rho = -0.076$, $p = 0.67$), and cluster 5 ($\rho = 0.380$, $p = 0.081$). However for cluster 3, increasing concentration of Cl is weakly but significantly related with increasing δD values ($\rho = 0.408$, $p = 0.020$). Also, Figure 4. 6 shows three groundwater groups: low salinity and high δD water (cluster 1 and 2); low-medium salinity and low δD water (cluster 4 and 5); high salinity and high δD water (cluster 3). In contrast to the δD values, the $\delta^{18}O$ values do not differentiate the clusters (Figure 4. 7). As shown in Figure 4. 7, no significant trends is observed for the plot $\delta^{18}O$ versus Cl for the clusters except for cluster 5 where a significant but a weak relationship is found (ρ = 0.580, p = 0.005) and cluster 3 (ρ = 0.578, p = 0.000). The $\delta^{18}O$ versus Cl correlation for cluster 5 is accidental as will be explained later. Remarkably for cluster 3, both δD versus Cl and $\delta^{18}O$ versus Cl are significantly related.

Figure 4. 8 shows remarkable differences in the δD values of the clusters. Comparing the median δD values among the hydrochemical facies, clusters 4 and 5 have contrastingly low δD values. With the exception of cluster 6, the median values of cluster 1, 2, and 3 are above the upper quartile (75th percentile) of clusters 4 and 5 signaling that the δD values of clusters 4 and 5 may be significantly different from clusters 1, 2, and 3 (Figure 4. 8). The δD value of cluster 6 lies within the Interquartile Range (IQR) of the boxes for clusters 4 and 5. Wilcoxon Rank Sum test shows significant difference between the median δD value of the discharge end member waters (clusters 4 and 5) and the recharge end member waters—cluster1 (p = 0.00) and cluster 2 (p = 0.01). Also Wilcoxon test shows significant difference between the median δD value of the floodplain-Virgin River group (cluster 3) and cluster 4 and 5 with which cluster 3 likely interact with (p = 0.00). No significant difference is found between the median δD values of the two recharge end member waters (p = 0.18).

Table 4. 1. Linear regression coefficients of stable isotopic values characteristic of clusters

Clusters	Slope ± error	Intercept ± error	Linear model R-Squared Correlation
1	5.3 ± 0.5	-24.3 ± 6.6	0.881
2	7.4 ± 0.8	1.1 ± 10.5 *	0.999
3	4.1 ± 0.3	-43.2 ± 3.4	0.880
4	5.9 ± 0.4	-21.2 ± 5.7	0.848
5	6.2 ± 0.4	-17.0 ± 5.4	0.919

Coefficients are significant at $P<0.05$.
Coefficient marked * is not significant
The linear model R-squared correlation is interpreted as the variance of δD explained by $\delta^{18}O$

4.4.3. Hydrochemical solutes (Na, K, SO4, and Cl)

Box-and-whiskers plot and Wilcoxon Rank Sum Test are used to compare the clusters characterized as recharge and discharge end member waters and the surface water-floodplain group (cluster 3). Median concentrations of the solutes Na, K, SO4, and Cl are relatively larger for the discharge end member waters than concentrations for the recharge end member waters. Recharge end member waters have similar concentrations of Na (p = 0.96), but significantly different concentrations of K (p = 0.00), SO4 (p = 0.00) and Cl (p = 0.01). Groundwater of cluster 1 has relatively higher SO_4 and Cl concentrations than groundwater of cluster 2. Cluster 2 has higher K concentrations than the concentrations of cluster 1. Among the designated discharge end-member, cluster 6 has the largest concentrations of all the selected ions except for potassium. Also concentrations of the solutes of cluster 4 are relatively larger than and significantly different (p = 0.00) from the concentrations of cluster 5. For the selected ion concentrations, cluster 3 have the lowest variance comparing to the other clusters. The relatively wide variation of concentration for the groundwater groups is evidence of fracture controlled groundwater flow.

4.5. Discussion

Generally, modern precipitation data plot on the global meteoric water line (Craig, 1961). However in arid regions, the stable isotopic compositions of precipitation often plot to the right of the global meteoric water line (GMWL) as a result of a modification of the stable isotopic compositions by local climatic conditions (Gat, 1971). Furthermore, recharge processes can modify the stable isotopic composition of precipitation so that the isotopic compositions of groundwater become different from those of the source precipitation. The results of this study indicate that partially evaporated precipitation recharges groundwater in the lower Virgin River Basin. Evaporation prior to recharge is inferred from the plotting of most of the hydrochemical samples to the right of the GMWL. However, other processes and conditions including rock and water interactions and subsurface high temperature can affect the stable isotopic compositions of the groundwater as are discussed below. The use of the GMWL in this study is justified because winter precipitation, the dominant source of recharge (Winograd et al., 1998; Lachniet et al., 2011) in the Great Basin (southern Nevada), plot close to the GMWL (Friedman et al., 2002). Also remarkably stable isotopic values of nine drip water of Pinnacle Cave in southern Nevada fall on the GMWL (Lachniet et al., 2011).

The results of this study provide a basis for diagnosing interbasin groundwater flow to the lower Virgin River Basin. One interesting finding, based on the gradients of the regression lines of the clusters, is that evaporation has affected the different clusters to different extents. Comparing the discharge end member waters (cluster 4 and 5) to the recharge end member waters (cluster 1 and 2), cluster 1 is more affected by evaporation and cluster 2 is less affected by evaporation prior to infiltration and percolation. This finding does not support cluster 1 or 2 as the dominant sources supplying water to clusters 4 and 5. The interpretation is based on the assumption that evaporation is the only dominant process affecting the stable isotopic compositions and no subsurface exchange reactions are occurring. It is noted that other processes may complicate this interpretation. Despite the differences in evaporation, the intersection points of the regression line of clusters 2 is closest to the intersection points of the clusters 4 and 5, indicating the approximate oxygen and hydrogen isotopic compositions of the precipitation recharging these three groups are similar.

Additional information to diagnose interbasin flow is provided by results from the analysis of the chemical solutes data and further analysis of the stable isotopic data. The relative differences in the selected ion compositions of the hydrochemical facies indicate that the recharge end member waters are younger than the discharge end member waters. All other factors being equal, a longer flowpath has relatively higher concentrations of the selected major ions than a shorter flowpath. The solute data support groundwater movement from cluster 1 or 2 to cluster 4 or 5; however the hydraulic connection between these recharge end member waters and the discharge end member waters cannot be supported by their stable isotopic data. The box-and-whiskers plot analyses indicate clusters 1 and 2 are not the dominant source of water for clusters 4 and 5. The deuterium isotopic values of the source waters are high and significantly different from the low isotopic values of the discharge waters. This difference indicates that interbasin movement of groundwater from Clover Valley or Escalante Valley to the Virgin River

Basin, or the intrabasin recharge in the lower Virgin River Basin cannot alone account for discharge groundwater in the lower Virgin River Basin. The significant difference between the recharge and discharge waters indicate that, even in a mixing case, the contribution from the recharge end member waters are smaller.

In contrast to the significant differences in deuterium isotopic values observed, the oxygen isotopic values of the source waters do not differ from the oxygen isotopic values of the discharge waters. This may indicate that the oxygen isotopic values of the discharge waters have shifted to more positive values approaching those of the discharge end member waters. This observation points to rock and water interactions leading to exchange of oxygen isotopic values of the groundwater with the isotopic values of the rocks. Groundwater oxygen isotopic values exchanges with the oxygen isotopic values of the aquifer media when the temperature in the hydrogeologic environment is higher and/ when the residence time of the water is longer (Clayton et al., 1966; Gat, 1971; Pearson et al., 1991). This explanation is possible for water flowing though the deep carbonate aquifer. Because the deep carbonate aquifer flow is regional, the residence time of the groundwater is longer, hence the potential for isotopic exchange reaction exist. Previous studies also recognize the possibility of $\delta^{18}O$ exchange between the groundwater and the deep carbonate aquifer (Winograd and Friedman, 1972; Thomas et al., 2001; Smith et al., 2002). Studies using $\delta^{18}O$ data alone have to be re-evaluated (Davisson et al., 1999b). An alternative explanation for the lack statistically significant difference between the oxygen isotopic values of the clusters is that, the variation of oxygen isotopic composition in natural water is small and is explained by the small mass difference between ^{18}O and ^{16}O isotopes (Ferronsky and Polyakov, 1982). In contrast, deuterium varies widely in natural water because mass the difference between deuterium and protium is large; therefore in natural processes, the separation of deuterium and protium is more effective.

A previous study has identified the Ash, Crystal, and Hiko Springs as discharging from the deep carbonate aquifer (Thomas et al., 1996). These springs are located in the northwest of the lower Virgin River Basin and in the Colorado flow systems, of which the lower Virgin River Basin is part. Therefore δD values of these springs (Hershey et al., 2010) are used in this study to represent the deep carbonate aquifer source to the lower Virgin River Basin. The springs have median δD value equal to -109.0‰. The median value of cluster 1 is -87.0‰, cluster 2 is -93.0‰, cluster 3 is -93.0‰, cluster 4 is -98.25‰, and cluster 5 is -101.50‰. It is evident that the discharge in the lower Virgin River Basin is mixing between intrabasin and interbasin sources. Mixing calculations results show that 51-66% deep carbonate aquifer source mixing with 34-49% intrabasin source account for the discharging cluster 4 or 5. Also mixing calculations results show that 33-53% deep carbonate aquifer source mixing with 47-67% interbasin source from the Clover or Escalante Desert Valley account for the discharging cluster 4 or 5.

Information extracted from hydrochemical and stable isotopic data provide evidence to support interactions between the floodplain aquifer and the Virgin River. However the interaction of the Virgin River and Muddy Creek Formation aquifer is not supported. The Virgin River samples group with the groundwater samples located along its' course (floodplain aquifer) and together they formed cluster 3. The groundwater in the study area has very diverse stable isotopic values. The variability in stable isotopic values is explainable by altitude, latitude, evaporation, isotopic exchange reaction, or mixing and fracture controlled groundwater flow. However, cluster 3 consisting of the floodplain water and Virgin River water has a relatively small spread. This means that the Virgin River with floodplain aquifer water is well mixed compared to the Muddy Creek aquifer groundwater. Also the Virgin River with floodplain aquifer water samples are significantly evaporated compared to the groundwater samples. The evaporation of surface waters in arid regions is a well known process. However for this study, the groundwater chloride concentration increases due to dissolution reactions, hence the lack of correlation between isotopic compositions and chloride concentration. When the regression lines are extrapolated to the GMWL the average isotopic values of recharge water for the different hydrochemical facies are inferred. However, the isotopic values of the discharge waters are lower than the actual values due to the oxygen isotopic exchange reaction in the groundwater aquifer system.

4.6. Conclusions

Data analyzed in this study support that at least 50% of water in the lower Virgin River Basin is derived from interbasinal sources. Groundwater in the lower Virgin River Basin is a mixture of groundwater of the deep carbonate aquifer or the Clover and Escalante Valleys and recharge derived from precipitation in the lower Virgin River Basin. Low stable isotopic values and high solutes concentrations of the discharge waters indicate contribution of the regional groundwater source. Also, if one has to analyze only a single isotopic parameter, δD parameter should be preferred to $\delta^{18}O$.

This study did not answer the hypothesis that the discharge water is paleowater. Also the carbonate-water $\delta^{18}O$ exchange reaction has been suggested in this study and needs to be further investigated. The carbonate-water $\delta^{18}O$ exchange reaction inferred in this study has a serious implication for studies using $\delta^{18}O$ to interpret flow systems. We conclude that to use $\delta^{18}O$ data for studying flow systems, $\delta^{18}O$ values are to be adjusted for the carbonate-water $\delta^{18}O$ exchange reaction. Finally, the procedures described in this study can be used to investigate flow systems in other intermountain basins.

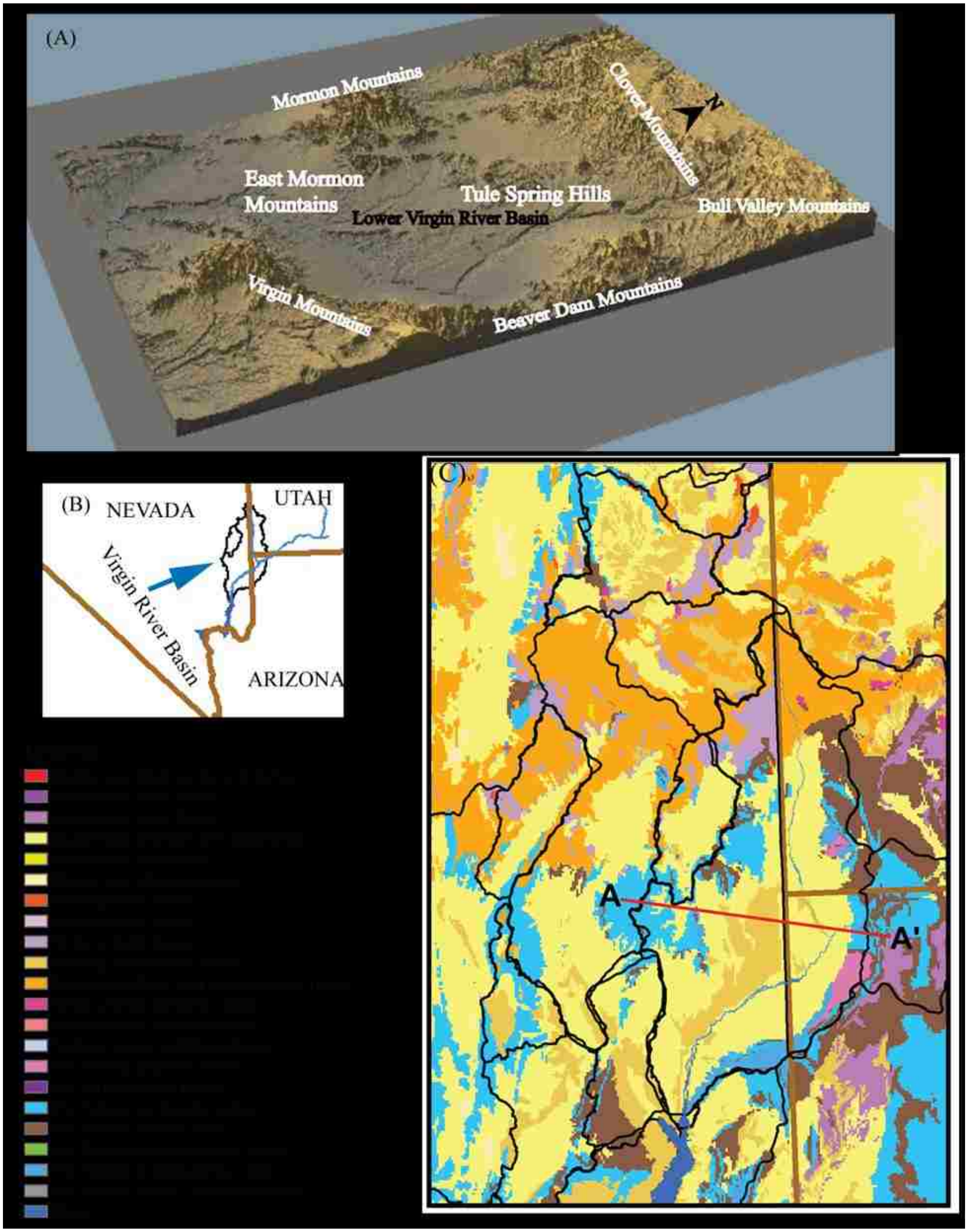

Figure 4. 1. Physiography (A), Location (B), and simplified geology map (C) of the lower Virgin River Basin and the adjacent basin areas. The geologic map is modified after GBCGE (2005)

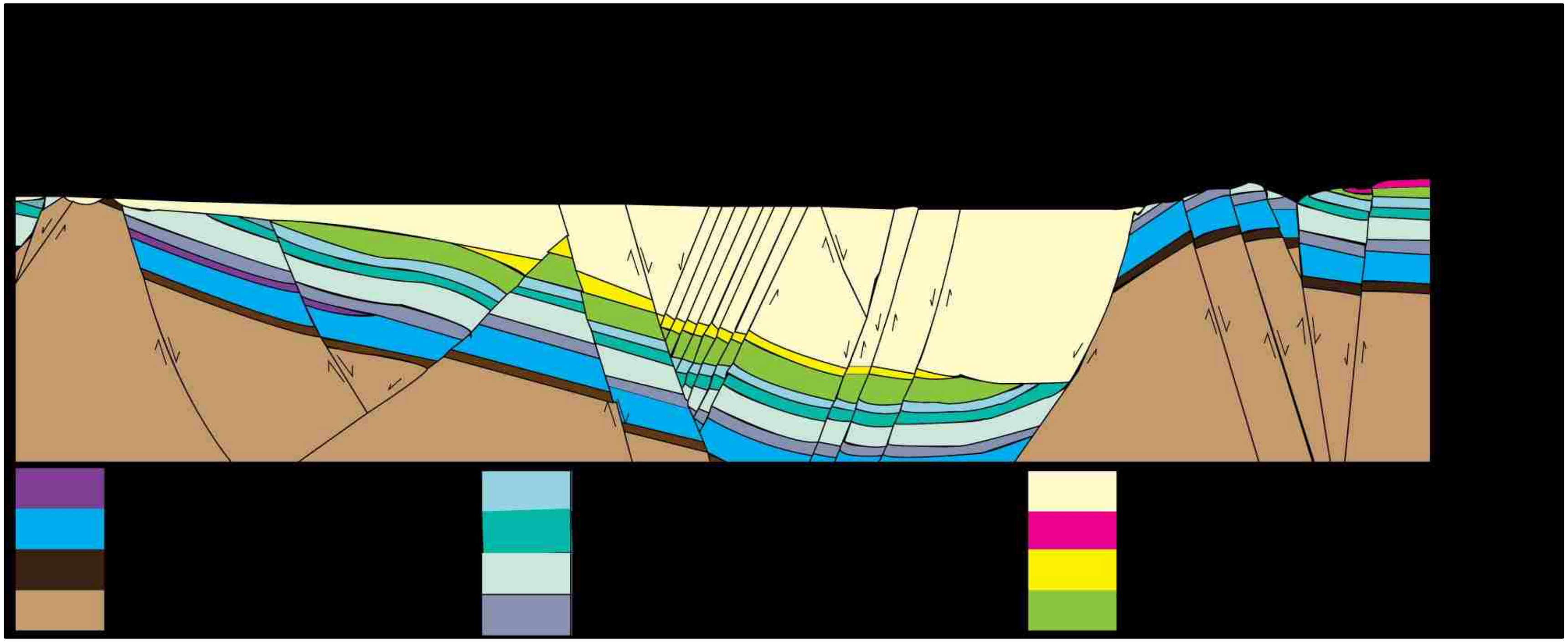

Figure 4. 2. Geologic cross-section along the line A—A' on the geology map. Modified after Page et al. (2006).

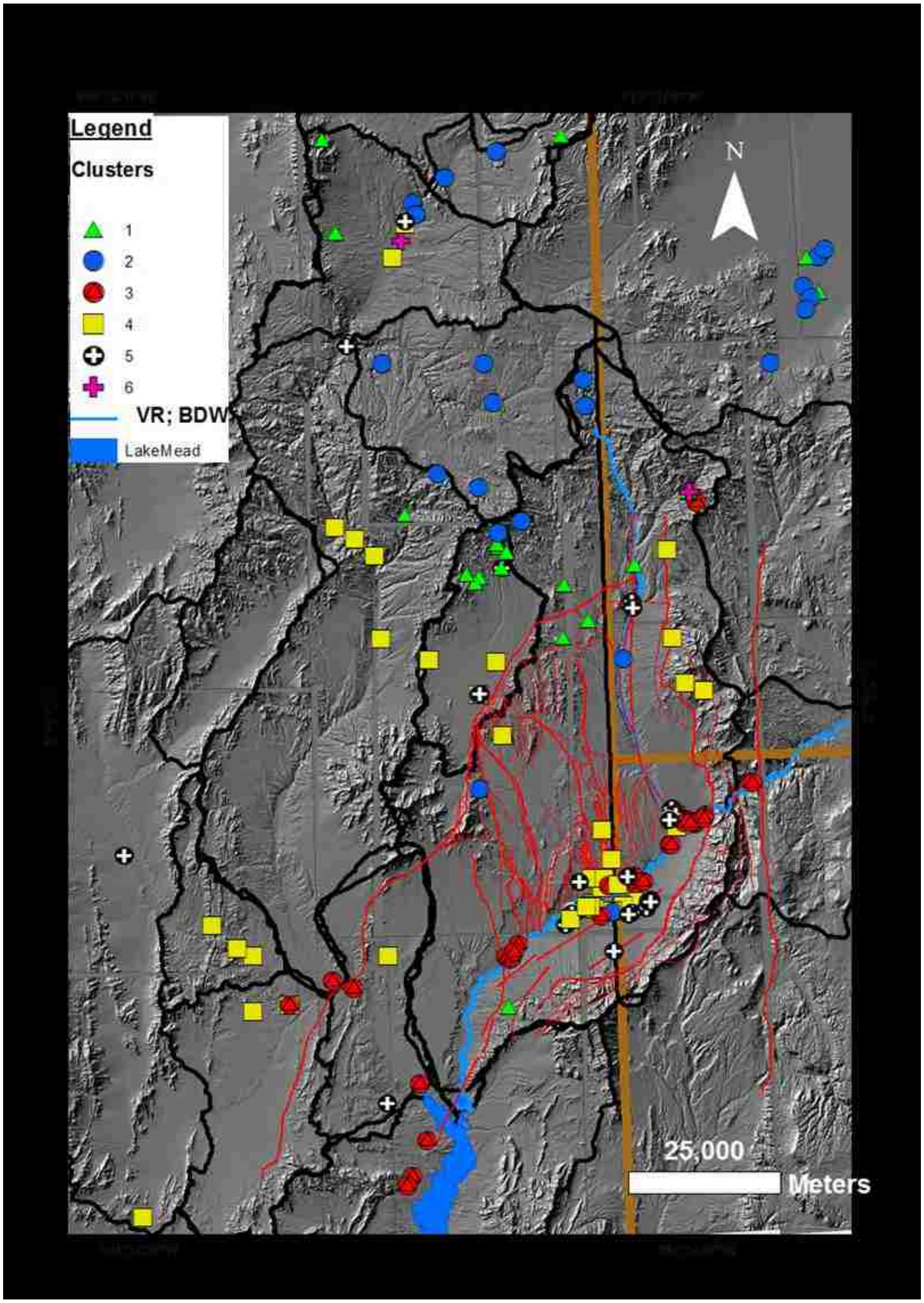

Figure 4. 3. Hydrochemical facies in geographic space (DEM) for the lower Virgin River Basin (Chapter 3). Note, there are coinciding sampling points due to the small scale of the map. The red lines are faults (Page et al., 2005; Dixon and Katzer, 2002).

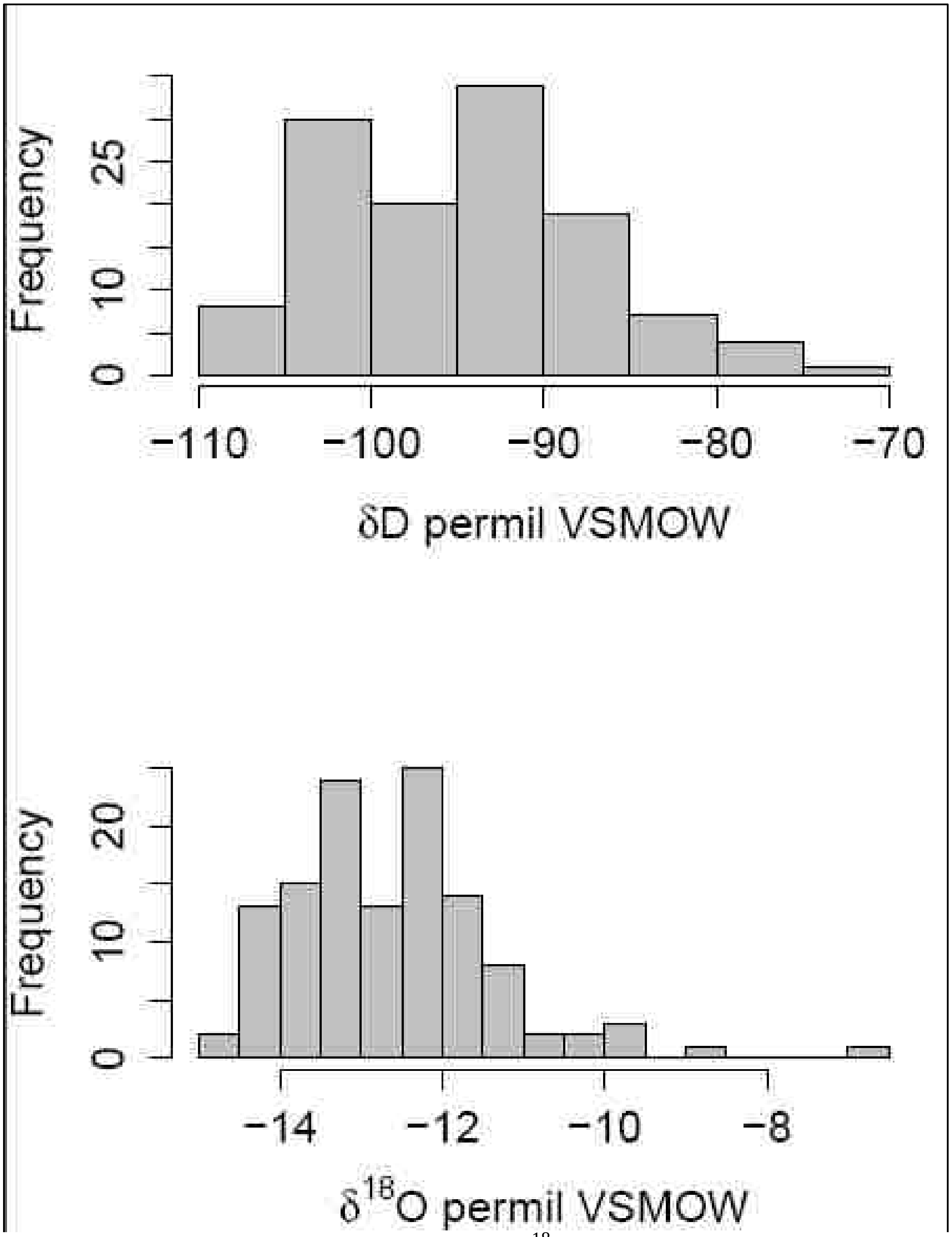

Figure 4. 4. Distribution the available δD and δ^{18}O data in the study area

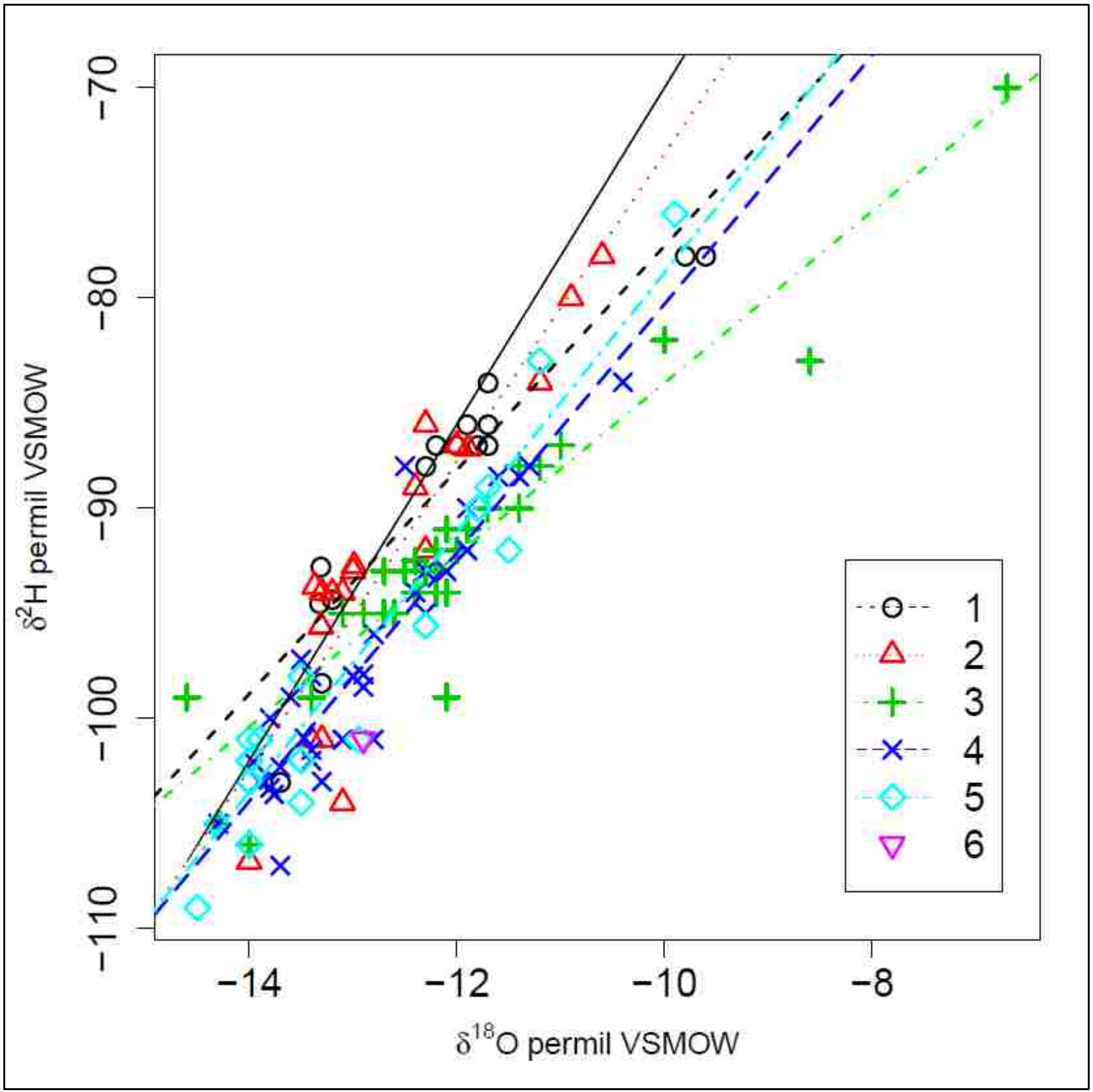

Figure 4. 5. δD and δ^{18}O of the hydrochemical facies (clusters 1 to 5). Lines are regressions fitted to facies 1 to 5 (dashed, point, point dash, double dash, and two dash respectively). The only inverted triangle is stable isotopic values of cluster 6. Black line (continuous) is the GMWL (Craig, 1961).

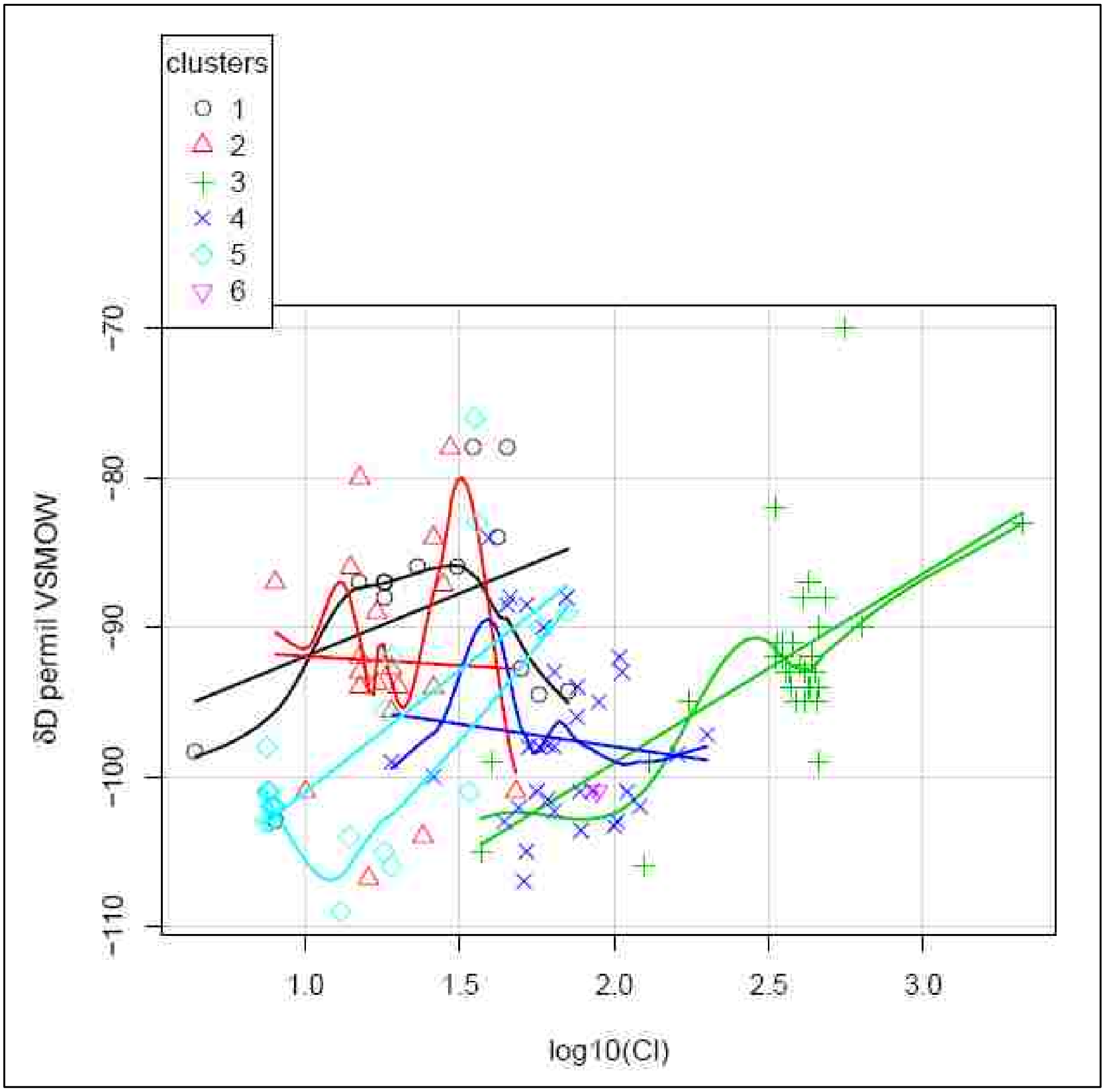

Figure 4. 6. δD and logCl relationship (linear and non-parametric regressions) of the hydrochemical facies (clusters)

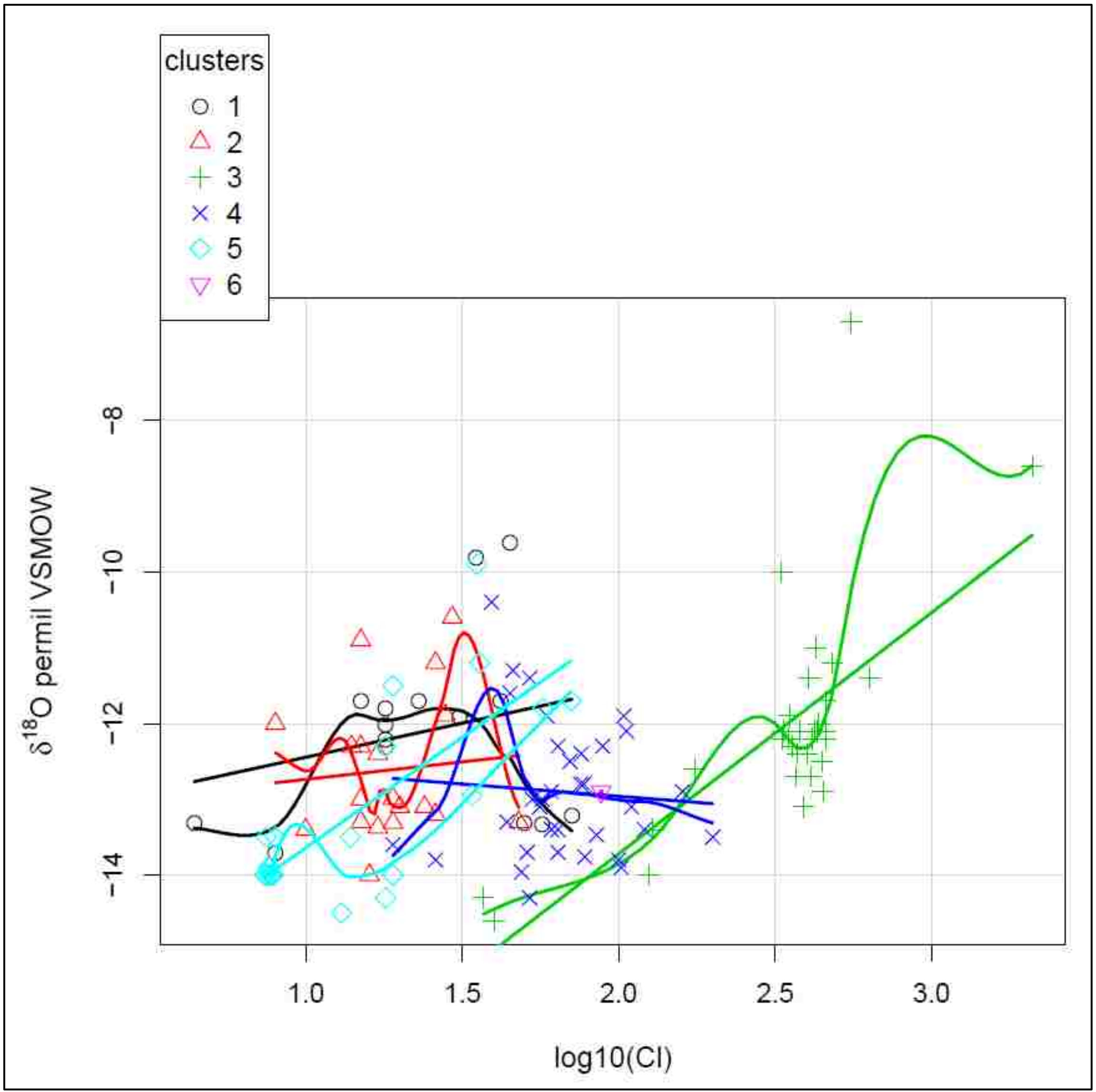

Figure 4. 7. $\delta^{18}O$ and logCl relationship (linear and non-parametric regressions) of the hydrochemical facies (clusters)

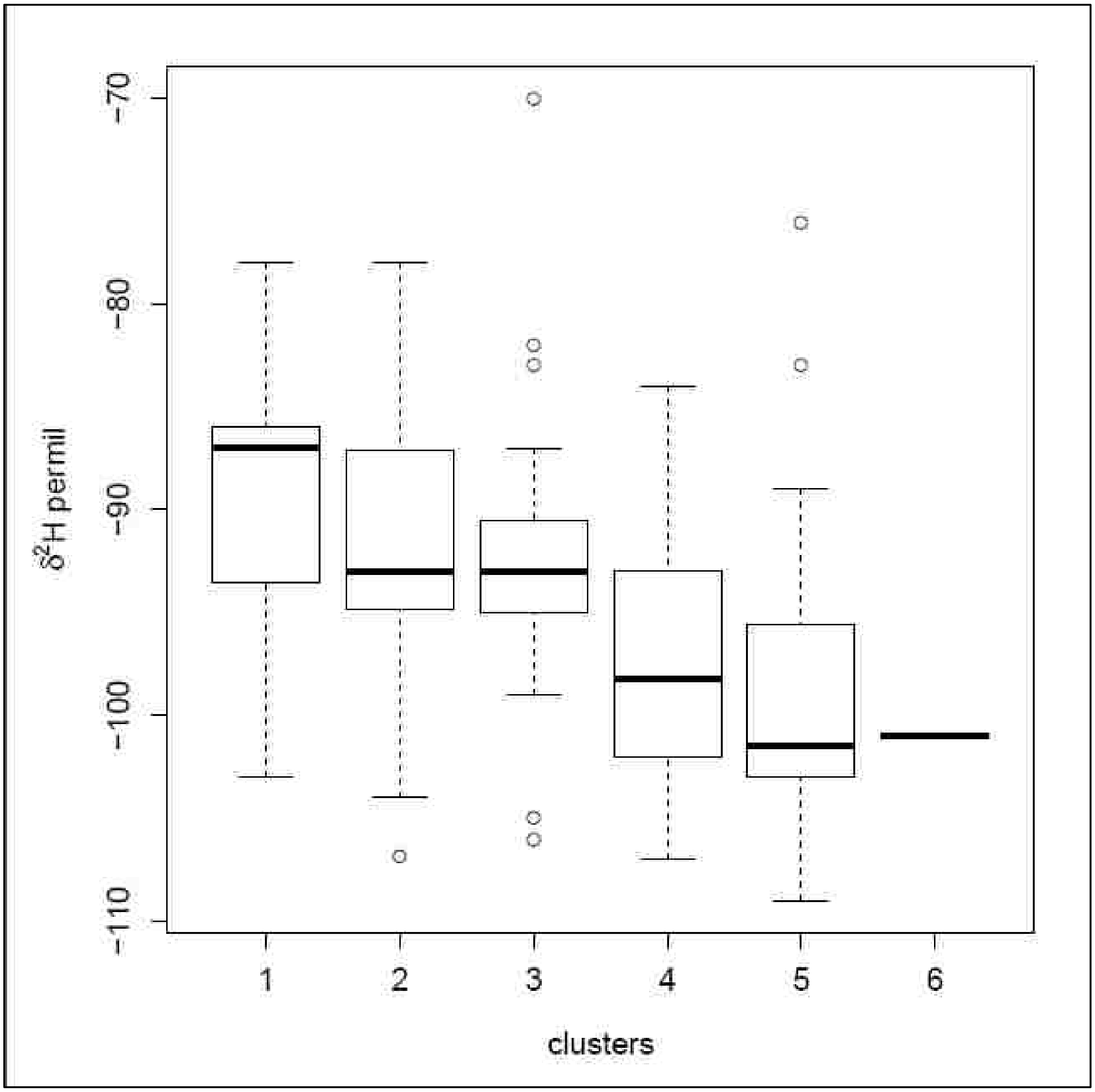

Figure 4. 8. Box-and whiskers plot of δD values of the hydrochemical facies

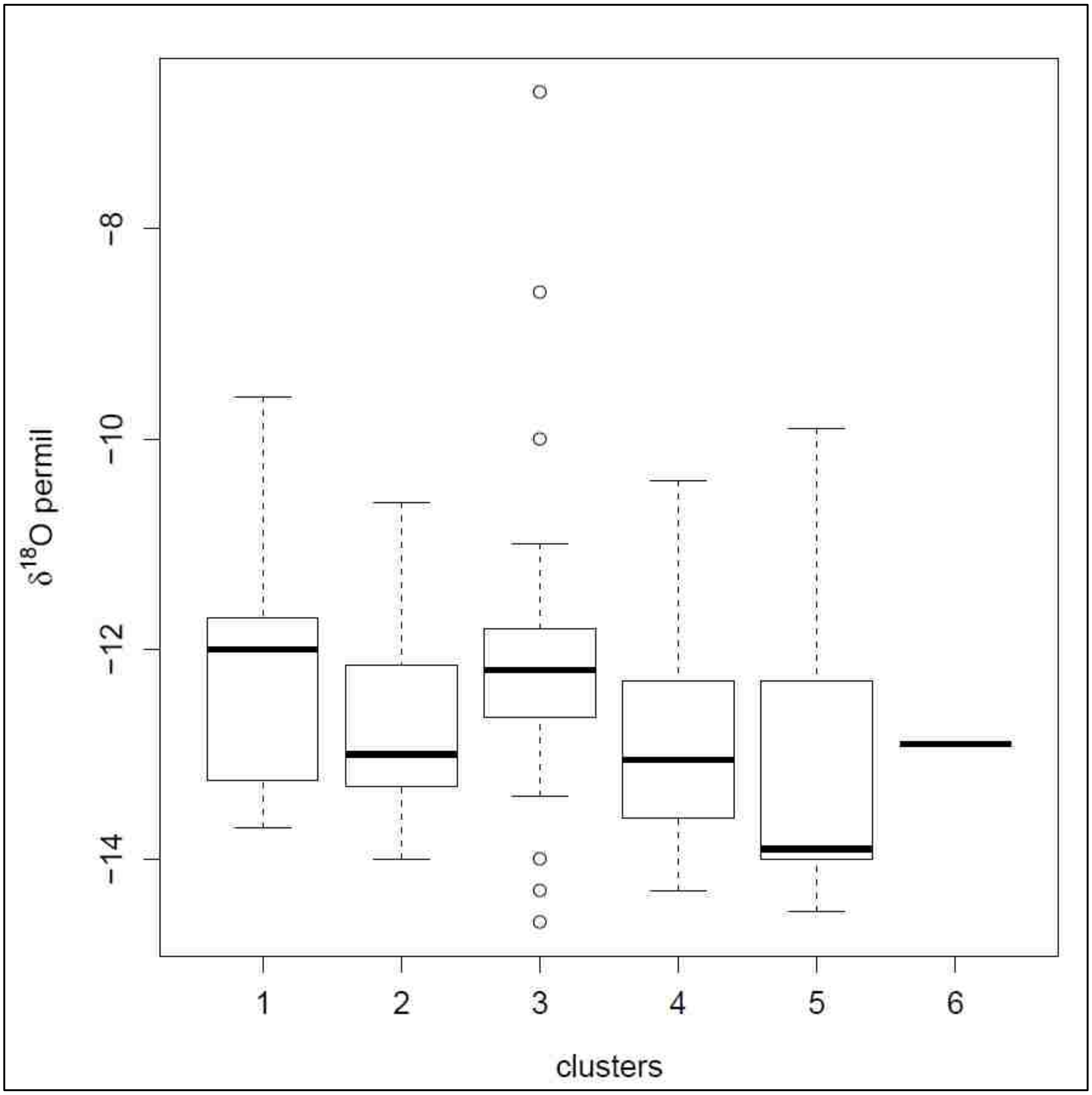

Figure 4. 9. Box-and whiskers plot of δD values of the hydrochemical facies

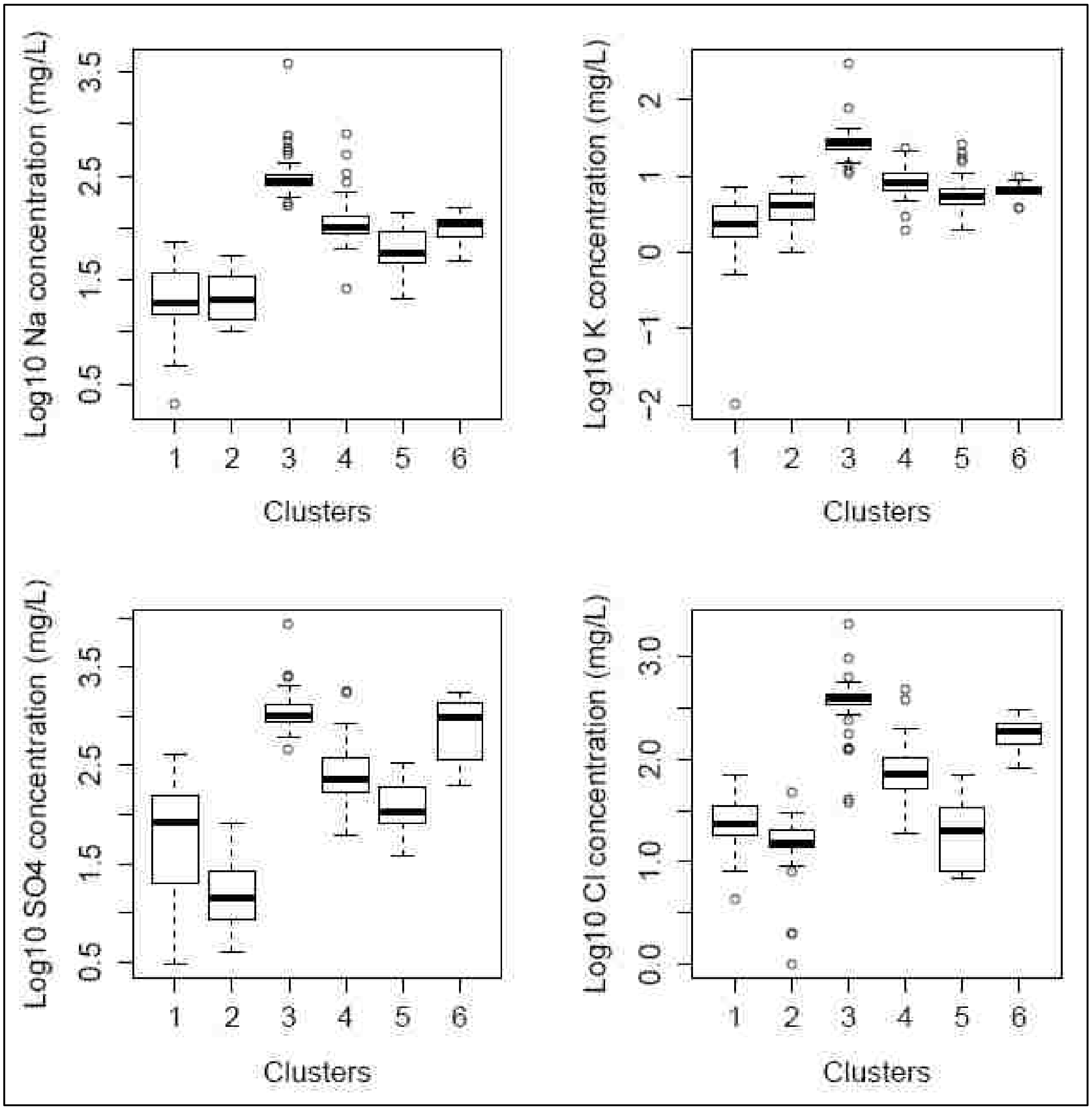

Figure 4. 10. Relative concentrations of selected solutes in the hydrochemical clusters

CHAPTER 5

DISSERTATION SUMMARY

5.1. Summary

The following is a summary of this dissertation. It covers the aspects of describing the major accomplishments of each section of the work, and continues on to describe the hydrogeologic ramifications of the findings.

The first section of this work introduced integrated methodology for establishing optimum hydrochemical facies in a groundwater basin, combining clustering techniques with Multiple Discriminant Function Analysis and measures of association. The method also allows quantitative analysis of the effect of data transformation, analytical error, and data outliers. The method was applied to the lower Virgin River Basin, which lies on the border of the States of Nevada, Arizona, and Utah in the United States.

In the second study, the clusters were characterized as source or discharge end-members based on their chemical character. Aqueous geochemical modeling was used to simulate the progression of chemical evolution in the lower Virgin River Basin. The data used for the modeling included the source end-members located near the Clover Mountains front, in the deep carbonate aquifer, and in the Clover/Escalante Valley. Also mineralogical phases were deduced from geologic information and compositional plots of major ions.

In the third study regression analysis, histograms, box-and-whisker plots, Pearson correlation. and Wilcoxon Rank Sum Test were used to decipher the physical recharge processes from stable isotopic values for source and discharge end-members, which were then correlated to test hypotheses of interbasin groundwater flow to the lower Virgin River Basin, and interactions between groundwater and the Virgin River.

Discernible relationships exist between the hydrochemical facies (clusters 1 through 6) and the topography, geology, climate, and basin characteristics (Figure 3.1, Figure 3.2, and Figure 3.5). An optimal number of six hydrochemical facies were identified and a general progression of facies from north to south, from source to discharge, were consistent with topographic, basin characteristics (depth of basin-fill sediments and length), climatic, stratigraphic, and structural elements of the study area. The general trend is that Cluster 1 (Mountain Front – e.g. upper Beaver Dam Wash, upper Tule Basin, Meadow Valley basin) and Cluster 2 (silicic ashflow and volcanic clastics – e.g. northern uplands) are considered recharge/ source areas, in addition to any potential upwelling from the deep carbonate aquifer. Discharge is represented by Clusters 4, 5, and 6 (basin-fill sediments and/or fault controlled flow – i.e. dominantly central Beaver Dam wash, Tule Basin, Muddy river drainage), Detailed geologic structures and basin characteristics information available for this study are for only the lower Virgin River Basin, the focus of the study. Therefore, interpretation of structural control on groundwater flow for other basins is based on extrapolation from that in the lower Virgin River Basin. The correlation of the patterns of the hydrochemical facies to the elements listed above provides information for selecting suitable sites for water development, protection of water resources and ecosystem, and urban/regional planning. The hydrochemical facies display spatial coherence which indicates the physical and chemical state of the water as influenced by these elements of the hydrogeologic environment.

The lower Virgin River Basin receives different amounts of precipitation, (and correspondingly different amounts of potential recharge), which is generally correlated to topographic elevation and/ or shows a latitudinal distribution. The north boundary of the lower Virgin River Basin and beyond receives more precipitation than the locations below this boundary. Generally, in the Great Basin region average annual precipitation in the north Valley locations is about 241 mm, but the southern locations receive precipitation of about 101 mm (Hershey et al, 2010). These climatic differences influence the spatial distribution of the source end-member and discharge end-member waters (Figure 3.5). The areas of higher elevation and high latitude receive more precipitation than the basin floor and low latitude areas. While this is a general trend, there are exceptions. For example, in the lower Virgin River Basin two samples from a source end-member (cluster 2) are located in the basin floor. The occurrence of these two source end-member water samples in the basin floor/low latitude of the lower Virgin River Basin, can be interpreted as resulting from their location in an area of highly faulted major washes (Figure 3.5) (Dixon and Katzer 2002; Page et al., 2005) that can transmit groundwater to the water table. In spite of some exceptions, general trends throughout the basin, from recharge to outlet, can be recognized.

The topography, depth of basin-fill sediments, and length of basin control the distribution of the source and discharge end-member waters. Generally, groundwaters at the higher elevations north of the lower Virgin River Basin are interpreted as source end-member waters and that at the lower elevations or basin floors can be considered discharge end-member waters. Also, groundwater samples from locations near the lower Virgin River and lower Muddy River to the east, show a distinct hydrochemical signal (cluster 3), which is interpreted as part of a floodplain aquifer, and have hydrochemistry similar to river water quality.

The size and depth of basin fill sediments influences flowpath, water quality and consequently the way facies are distributed. This can be seen in comparing the basin characteristics of the large and deep lower Virgin River Basin to other, smaller basins. For example, topographic influence on the pattern of hydrochemical facies generated is greater in basins with considerable length and basin-fill sediments. Within basins, particularly large basins, sub-patterns can exist. In the lower Virgin River Basin, the source end-member, intrabasin water (cluster 1) is located at or near the mountain front. Basins with considerable basin-fill sediments exhibit the interbasin discharge end-member water (clusters 4, 5, and 6). This observation is in line with the concept of regional groundwater flow (Toth, 2009).

The patterns of the hydrochemical facies correlate with the stratigraphy and fault systems. On the basis of the stratigraphy, the two source end-members waters (cluster 1 and 2) can be differentiated. Cluster 1 tends to be located in areas of carbonate rocks or alluvium close to mountains, and cluster 2 is located in the silicic ash flow and volcaniclastics rocks, as mentioned earlier. The discharge source end-member waters are located in the basin-fill sediments and on or closer to fractures (Figure 3.2 and Figure 3.5). For example in the lower Virgin River Basin four samples of the discharge end-member water (cluster 4) are located in the vicinity of a major fault known as the piedmont fault (Figure 3.2 and Figure 3.5). Similar major faults are observed between the East Mormon Mountain and Tule Spring Hills where some samples of the end-member discharge waters are located. Also, within major faults and dense fracture zones of smaller faults north of the Virgin River, abundant samples of the discharge end-member waters comingle. These observations are supporting evidence that the fractures are the main conduit for discharging regional groundwater in the basins.

Also these observations underpin the complexity of the flow systems. Groundwater in lower Virgin River Basin more likely occurs from mixing lateral flowing intrabasin water and upwelling interbasin water of significant proportion > 50%, as described in previous chapters. This explains why the same hydrochemical facies are located over considerable lateral distance in the basin floor in the lower Virgin River Basin.

The chemistry of the hydrochemical facies is the result of subsurface water interactions with reactive mineral phases of aquifer materials. In the study area, three potential sources of water replenishing the groundwater in the lower Virgin River Basin have been identified. Two of these have been identified as interbasin source (cluster 2 and the deep carbonate aquifer) and the other is identified as intrabasin water (cluster 1).

The source end-members are bicarbonate waters with relatively low TDS, Cl, and Na/Ca ratio. The three potential source waters are described as Ca-Mg-HCO_3 water (cluster 1), Ca-Na-HCO_3 (cluster 2) and Ca-Mg-HCO_3 (Carbonate aquifer). The sodic and calcic characters are interpreted in terms of geologic formations where infiltration and recharge took place. Compositional plots indicate that the intrabasin source (cluster 1) dissolves halite (sodium chloride), gypsum, calcite, and dolomite. The interbasin source in the Clover Valley/Escalante Valley (Cluster 2) dissolves albite in the silicic ash flow and volcaniclastics, and sodium chloride and gypsum in the soil cover during infiltration. In the carbonate aquifer, the regional interbasin flow dissolves calcite, dolomite, and minor minerals like gypsum and halite.

Although four springs in Cluster 2 are located in the lower Virgin River Basin at the Clover Mountains, the ensemble group of Cluster 2 wells and springs can still be considered an interbasin source. These four exceptions commonly have small flow rates (Hershey et al., 2010) and so are not considered an important source of recharge to the lower Virgin River Basin. Therefore, it is logical to label the source end-member water (cluster 2) primarily as interbasin source recharged in the silicic ash flow and volcaniclastics rocks in the Clover Valley or beyond.

The aquifer along the Virgin River, which in this study is designated as the floodplain aquifer, is interpreted to be in strong hydraulic interconnection with the Virgin River as indicated by the similarity of the chemistry and TDS of the two waters. Based on the results of the hydrochemical data analyses no link is recognized between the Virgin River/Floodplain aquifer and the Basin-fill Aquifer.

The discharge end-members (cluster 4, 5, and 6) have relatively higher TDS and Cl concentrations than source waters. The combined higher TDS and Cl are interpreted as resulting from a long flowpath through the deep subsurface. Discharge end-members are described as $Na\text{-}Ca\text{-}SO_4\text{-}HCO_3$ (cluster 4), $Na\text{-}Ca\text{-}HCO_3\text{-}SO_4$ (clusters 5), and $Ca\text{-}SO_4$ (cluster 6). The Na/Ca ratio is relatively high (Figure 3.7) for all discharge waters (except cluster 6) and is interpreted as resulting from increasing flowpath and subsequent calcite saturation. The discharge end-members evolution follows two pathways. In one pathway, carbonate groundwater mixes with groundwater in the Clover/Escalante Valley as they both flow southwards towards the Virgin River. For the other pathway, carbonate mixes with intrabasin source end-member water (cluster 1). The reactions that occur are dissolution of dolomite, halite, gypsum, and k-feldspar, Na/Ca ion exchange, and precipitation of calcite, quartz, and kaolinite. Reactions for the second pathway are dissolution of halite, albite, and gypsum, Ca/Na ion exchange, and precipitation of dolomite, calcite, quartz, and kaolinite.

The findings of this study have implications for water resources development, planning, and management for the region, including

1. the study supports the possibility of groundwater upwelling,
2. the research indicates enhanced flow in faults/fractures,
3. the findings lend evidence that basin depth is important to regional flow,
4. this work finds mineral supersaturation of groundwater is prevalent and may affect the long term permeability of fractures, and
5. the results indicate that cluster 6 is different than other discharge waters (specifically with regard to TDS concentration and dissolved major ions) and perhaps could be influenced by anthropogenic activities.

This study provides supporting evidence that a significant proportion (over 50%) of the groundwater in the basin-fill aquifer of the lower Virgin River Basin is either from interbasin water upwelling through the deep carbonate aquifer, and/or from interbasin groundwater recharged in the Clover Valley transmitted through faults in the Clover Mountains. Two elements tend to control the discharge of deep interbasin groundwater - fractures and deep basin depth as clearly observed in the lower Virgin River Basin. Groundwater in the basins becomes quickly saturated to supersaturated with calcite and dolomite during recharge. A supersaturation state has implications for aquifer permeability because precipitation of calcite or dolomite reduces the fracture space for groundwater movement. The discharge water (cluster 6) is chemically unrelated to the other discharge groundwaters (clusters 4 and 5). Anthropogenic activities such as mining or farming may be influencing this water quality, but presently there is no ground verification for this assertion.

What is clear is that much of the recharge to the lower Virgin River Basin occurs at the lower Virgin River Basin north bordering mountains and Valleys. For the study area, groundwater recharge tends to be strongly controlled by climate and topography, and groundwater discharge tend to be controlled by the depth of basin-fill sediments and fractures. In future studies, radioisotopic analysis, trace element hydrochemistry, and examination of chemicals related to anthropogenic activities can further test interpretations made herein. With increasing population and prosperity in the region, and a commensurately greater demand for water, an increasing clear understanding of groundwater sources and movement will be critical for water planning and management.

REFERENCES CITED

Acheampong, S.Y., Hess, J.W., 1998. Hydrogeologic and hydrochemical framework of the shallow groundwater system in the southern Voltaian Sedimentary Basin, Ghana. Hydrogeology Journal 6 (4), 527-537.

Acheampong, S.Y., Hess, J.W., 2000. Origin of the shallow groundwater system in the southern Voltaian Sedimentary Basin of Ghana: an isotopic approach. Journal of Hydrology 233 (1), 37-53.

Alley, W.M., Reilly, T.E. and Franke, O.L., 1999. Sustainability of ground-water resources. 1186, U.S. Department of the Interior, U.S. Geological Survey, Denver, Colorado, 79p.

American Public Health Association, 2005. Standard methods for the examination of water and wastewater :including bottom sediments and sludges, 21st ed. American Public Health Association, New York.

Anderson, K., Nelson, S., Mayo, A., Tingey, D., 2006. Interbasin flow revisited: The contribution of local recharge to high-discharge springs, Death Valley, CA. Journal of Hydrology 323 (1-4), 276-302.

Ashley, R.P., Lloyd, J.W., 1978. An example of the use of factor analysis and cluster analysis in groundwater chemistry interpretation. Journal of Hydrology 39 (3-4), 355-364.

Bacher, J., 2001. Teststatistiken zur Bestimmung der Clusterzahl für Quick Cluster. ZA-Information / Zentralarchiv für Empirische Sozialforschung 4871-4897.

Back, W., 1961. Techniques for mapping of hydrochemical facies. Hydrogeology P 0424-D, U. S. Geological Survey, Reston, VA, United States (USA). pp. D380-D382.

Baron, A.G., 2007. Interaction of surface and subsurface hydrological processes in the lower Virgin Valley. M.Sc. Thesis University of Nevada, Las Vegas, Nevada.

Bassett, R., Steinwand, A., Jorat, S., Petersen, C., Jackson, R., 2008. Forensic Isotope Analysis to Refine a Hydrologic Conceptual Model. Ground Water 46 (3), 372-383.

Bateman, R.L., 1976. Inventory and chemical quality of ground water in the White River-Muddy River-Meadow Valley wash area, southern Nevada. Project No.40, Water resources Center Desert Research Insitute University of Nevada Systems, Nevada.

Belcher, W.R., Bedinger, M.S., Back, J.T., Sweetkind, D.S., 2009. Interbasin flow in the Great Basin with special reference to the southern Funeral Mountains and the source of Furnace Creek springs, Death Valley, California, U.S. Journal of Hydrology 369 (1-2), 30-43.

Blasch, K.W., Bryson, J.R., 2007. Distinguishing Sources of Ground Water Recharge by Using δ^2H and δ^{18}O. Ground Water 45 (3), 294-308.

Bohannon, R.G., Grow, J.A., Miller, J.J., Blank, R.H., JR., 1993. Seismic stratigraphy and tectonic development of Virgin River depression and associated basins, southeastern Nevada and northwestern Arizona. Geological Society of America Bulletin 105 (4), 501-520.

Brothers, K., Katzer, T., Mojib, R.M., Grinnell, G., Bernholtz, A. and Johnson, M., 1993. Addendum to hydrology and interactive computer modeling of ground and surface water in the lower Virgin River Valley, primarily in the Clark County Nevada. 1-A, Las Vegas Valley Water District, Cooperative Water Project, Las Vegas.

Bunch, R.L., Harrill, J.R., 1984. Compilation of selected hydrologic data from the MX missile-siting investigation, east-central Nevada and western Utah. Open-File Report84-702, U.S Geological Survey, pp. 1-123.

Bushman, M., Nelson, S.T., Tingey, D., Eggett, D., 2010. Regional groundwater flow in structurally-complex extended terranes: An evaluation of the sources of discharge at Ash Meadows, Nevada. Journal of Hydrology 386 (1-4), 118-129.

Carrillo-Rivera, J.J., Clark, I.D., Fritz, P., 1992. Investigating recharge of shallow and paleo-groundwaters in the Villa de Reyes Basin, SLP, Mexico, with environmental isotopes. Applied Hydrogeology 1 (4), 35-48.

CEMP, 2011. Community Environmental Monitoring Program (CEMP). A joint program between the Department of Energy's (DOE) National Nuclear Security Administration Nevada Site Office (NNSA/NSO), and the Desert Research Institute (DRI) of the Nevada System of Higher Education. Available at: http://www.cemp.dri.edu/cemp/. Accessed 4/21, 2011.

CH2M HILL, 2002. Focused hydrological assessment of the Tule Desert hydrographic area including relevant aspects of the Virgin River hydrographic area in Southern Nevada. Toqoup Energy, Inc., Nevada.

Chapelle, F.H., 2005. Geochemistry of groundwater, in: Holland, H.D., Turekian, K.K. (Eds.), Surface and Ground Water, Weathering, and Soils (eds. Drever, J.I.) Vol. 5 Treatise on geochemistry. Elsevier Pergamon, Oxford, UK; San Diego, CA, USA, pp. 425-449 .

Charron, J.E., 1969. Hydrochemical interpretation of groundwater movement in the Red River Valley, Manitoba. Inland Waters Branch, Dept. of Energy, Mines and Resources, Ottawa.

Clark, I.D., Douglas, M., Raven, K., Bottomley, D., 2000. Recharge and Preservation of Laurentide Glacial Melt Water in the Canadian Shield. Ground Water 38 (5), 735-742.

Clark, I.D., Fritz, P., 1997. Environmental isotopes in hydrogeology. CRC Press, Lewis Publishers, Boca Raton, FL, United States.

Clark, I.D., Fritz, P., Michael, F.A., 1982. Isotope hydrogeology and geothermometry of the Mount Meager geothermal area. Canadian Journal of Earth Sciences 19 (7), 1454-1473.

Clayton, R.N., Friedman, I., Graf, D.L., Mayeda, T.K., Meents, W.F., Shimp, N.F., 1966. The origin of saline formation waters 1. Isotopic composition. Journal of Geophysical Research 71 (16), 3869-3882.

Cloutier, V., Lefebvre, R., Therrien, R., Savard, M.M., 2008. Multivariate statistical analysis of geochemical data as indicative of the hydrogeochemical evolution of groundwater in a sedimentary rock aquifer system. Journal of Hydrology 353 (3-4), 294-313.

Cohen, J., 1960. A coefficient of agreement for nominal scales. Educational and Psychological Measurement 20 (1), 37-46.

Craig, H., 1961. Isotopic variations in meteoric waters. Science 133 (3465), 1702-1703.

Dalton, M.G., Upchurch, S.B., 1978. Interpretation of hydrochemical facies by factor analysis. Ground Water 16 (4), 228-233.

Dansgaard, W., 1964. Stable isotopes in precipitation. Tellus 16 (4), 436-468.

Davisson, M.L., Rose, T.P., Smith, D.K., Kenneally, J., 1999a. Reply [to “Comment on ‘Isotope hydrology of southern Nevada groundwater: Stable isotopes and radiocarbon’ by M. L. Davisson et al.”]. Water Resources Research 35 (11), 3581-3583.

Davisson, M.L., Smith, D.K., Kenneally, J., Rose, T.P., 1999b. Isotope hydrology of southern Nevada groundwater: Stable isotopes and radiocarbon. Water Resources Research 35 (1), 279-294.

Demlie, M., Wohnlich, S., Wisotzky, F., Gizaw, B., 2007. Groundwater recharge, flow and hydrogeochemical evolution in a complex volcanic aquifer system, central Ethiopia. Hydrogeology Journal 15 (6), 1169-1181.

Deutsch, W.J., 1997. Groundwater geochemistry :fundamentals and applications to contamination. Lewis Publishers, Boca Raton.

Dixon, G.L., Katzer, T., 2002. Geology and hydrology of the lower Virgin River Valley in Nevada, Arizona, and Utah. VVWD-01, Virgin Valley Water District, Nevada.

Dreher, T., 2003. Comment on Güler C, Thyne GD, McCray JE, Turner AK (2002): Evaluation of graphical and multivariate statistical methods for classification of water chemistry data (Hydrogeology Journal 10:455–474). Hydrogeology Journal 11 (5), 605-606.

Eakin, T.E., 1966. A regional interbasin groundwater system in the White River Area, southeastern Nevada. Water Resources Research 2 (2), 251-271.

Earth Point, 2009. Township and Range – Public land survey system on Google Earth. Available at: http://www.earthpoint.us/Townships.aspx. Accessed 10/20, 2009.

Edmunds, W.M., Carrillo-Rivera, J., Cardona, A., 2002. Geochemical evolution of groundwater beneath Mexico City. Journal of Hydrology 258 (1), 1-24.

Enright, M., 1996. Selected hydrologic data for the Beaver Dam Wash area, Washington County, Arizona, 1991-95. Open-File 96-493, United States Geological Survey, 36p.

Everitt, B.S., Landau, S., Leese, M., Stahl, D., 2011. Cluster analysis, 5th ed. John Wiley & Sons, Ltd, Great Britain.

Ferronsky, V.I., Polyakov, V.A., 1982. Environmental isotopes in the hydrosphere, Translated by Ferrosnky, S.V. ed. John Wiley & Sons, Chichester Sussex; New York.

Feth, J.H., Robertson C.E., Polzer W.L., 1964. Sources of mineral constituents in water from granitic rocks, Sierra Nevada, California and Nevada. USGS Numbered Series Water Supply Paper (1535-I), 1-76.

Feuerstein, W., Grimm-Strele, J., 1989. Plausibilitätstests für eine routinemäßige Erfassung von Grundwasserbeschaffenheitsdaten. Vom Wasser 73, 375-398.

Fox, J., Weisberg, S., Fox, J., 2011. An R companion to applied regression, 2nd ed. SAGE Publications, Thousand Oaks, California.

Freeze, R.A., Witherspoon, P.A., 1967. Theoretical analysis of regional groundwater flow: 2. Effect of water-table configuration and subsurface permeability variation. Water Resources Research 3 (2), 623-634.

Friedman, I., Smith, G.I., Johnson, C.A., Moscati, R.J., 2002. Stable isotope compositions of waters in the Great Basin, United States 2. Modern precipitation. Journal of Geophysical Research 107 (D19), 4401-44022.

Garrels, R.M., Mackenzie, F.T., 1967. Origin of the chemical compositions of some springs and lakes, Chapter 10; Equilibrium concepts in natural water systems--Am. Chem. Soc., 151st Ann. Mtg., Pittsburgh, 1966, Div. Water, Air and Waste Chemistry Symposium. 222-242.

Gat J., Gonfiantini R. editors, 1981. Stable isotope hydrology: Deterium and oxygen-18 in the water cycle. Austria: International Atomic Energy Agency Vienna, Technical Reports Series No. 210.

Gat, J., 1971. Comments on the stable isotope method in regional groundwater investigations. Water Resources Research 7 (4), 980-993.

GBCGE, 2005. GIS regional data download. Available at: http://www.unr.edu/geothermal/datalist.html. Accessed June/30, 2011.

Genereux, D.P., 2005. A paired-watershed budget study to quantify interbasin groundwater flow in a lowland rain forest, Costa Rica. Water Resources Research 41 (4), 1-17.

Gibson, J.J., Edwards, T.W.D., Birks, S.J., St Amour, N.A., Buhay, W.M., McEachern, P., Wolfe, B.B., Peters, D.L., 2005. Progress in isotope tracer hydrology in Canada. Hydrological Process 19 (1), 303-327.

Glancy, P.A., Van Denburgh, A.S., 1969. Water-resources appraisal of the lower Virgin River Valley area, Nevada, Arizona, and Utah. State of Nevada, Division of Water Resources, Carson City.

Gong, X., Richman, M.B., 1995. On the Application of Cluster Analysis to Growing Season Precipitation Data in North America East of the Rockies. Journal of Climate 8 (4), 897-931.

Gordon, A.D., 1999. Classification, 2nd ed. Chapman & Hall/CRC, Boca Raton.

Gross, A.L., 1972. A Monte Carlo study of the accuracy of a hierarchical grouping procedure. Multivariate Behavioral Research 7 (3), 379-389.

Güler, C., Thyne, G.D., 2003. Reply to comment by T. Dreher to Güler C, Thyne GD, McCray JE, Turner AK (2002): Evaluation of graphical and multivariate statistical methods for classification of water chemistry data (Hydrogeology Journal 10:455–474). Hydrogeology Journal 11 (5), 607-608.

Güler, C., Thyne, G.D., 2004. Hydrologic and geologic factors controlling surface and groundwater chemistry in Indian Wells-Owens Valley area, southeastern California, USA. Journal of Hydrology 285 (1-4), 177-198.

Guler, C., Thyne, G.D., 2006. Statistical clustering of major solutes; use as a tracer for evaluating interbasin groundwater flow into Indian Wells Valley, California. Environmental & Engineering Geoscience 12 (1), 53-65.

Guler, C., Thyne, G.D., McCray, J.E., Turner, A.K., 2002. Evaluation of graphical and multivariate statistical methods for classification of water chemistry data. Hydrogeology Journal 10 (4), 455-474.

Heath, R.C., 1984. Ground-water regions of the United States. U.S. Dept. of the Interior, Geological Survey, Washington, D.C., 78p.

Helstrup, T., Jørgensen, N., Banoeng-Yakubo, B., 2007. Investigation of hydrochemical characteristics of groundwater from the Cretaceous-Eocene limestone aquifer in southern Ghana and southern Togo using hierarchical cluster analysis. Hydrogeology Journal 15 (5), 977-89.

Helstrup, T., Jørgensen, N., Banoeng-Yakubo, B., 2007. Investigation of hydrochemical characteristics of groundwater from the Cretaceous-Eocene limestone aquifer in southern Ghana and southern Togo using hierarchical cluster analysis
. Hydrogeology Journal 15 (5), 977-989.

Hem, J.D., 1989. Study and interpretation of the chemical characteristics of naatural water. Water-Supply Paper2254, United States Geological Survey, United States Government Printing Office, Washington, D.C. pp. 1-263.

Hershey, R.L., Mizell, S.A., Earman, S., 2010. Chemical and physical characteristics of springs discharging from regional flow systems of the carbonate-rock province of the Great Basin, western United States. Hydrogeology Journal 18 (4), 1007-1026.

Hibbs, B.J., Darling, B.K., 2005. Revisiting a Classification Scheme for U.S.-Mexico Alluvial Basin-Fill Aquifers. Ground Water 43 (5), 270-263.

Hounslow, A., 1995. Water quality data :analysis and interpretation. Lewis Publishers, Boca Raton.

Howarth, R., Earle, S., 1976. Application of a generalized power transformation to geochemical data. Mathematical Geology 11 (1), 45-62.

Ingraham, N.L., Zukosky, K., Kreamer, D.K., 2001. Application of stable isotopes to identify problems in large-scale water transfer in Grand Canyon national park. Environmental Science and Technology 35 (1), 299-302.

Johannesson, K.H., Stetzenbach, K.J., Hodge, V.F., Kreamer, D.K., Zhou, X., 1997. Delineation of ground-water flow systems in southern Great Basin using aqueous rare earth element distributions. Ground Water 35 (5), 807-819.

Johnson, M., Dixon G.L., Rowley P.D., Katzer T.C., Winters M., 2002. Hydrology and groundwater conditions of the Tertiary Muddy Creek Formation in the lower Virgin River Basin of southeastern Nevada and Adjacent Arizona and Utah. Geological Society of America 2002 Rocky Mountain Section, Annual Meeting, Cedar City, Utah Field Trip Manual 285-315.

Kahya, E., Demirel, M.C., Bég, O.A., 2008. Hydrologic homogeneous regions using monthly streamflow in Yurkey. Earth Sciences Research Journal 12 (2), 181-193.

Katzer, T., Dixon, G.L. and Johnson, M., 2002. Impact of water-resource development in Tule Desert, Lincoln County, Nevada on the water resources of the lower Virgin River Valley, Clark Country, Nevada. VVWD-02, VVWD, Nevada.

Kirk, S.T., Campana, M.E., 1990. A deuterium-calibrated groundwater flow model of a regional carbonate-alluvial system. Journal of Hydrology 119 (1-4), 357-388.

Kohfahl, C., Sprenger, C., Herrera, J.B., Meyer, H., Chacon, F.F., Pekdeger, A., 2008. Recharge sources and hydrogeochemical evolution of ground water in semiarid and karstic environments; a field study in the Granada Basin (southern Spain). Applied Geochemistry 23 (4), 846-862.

Kreamer, D.K., Hodge, V.F., Rabinowitz, I., Johannesson, K.H., Stetzenbach, K.J., 1996. Trace element geochemistry in water from selected springs in Death Valley National Park, California. Ground Water 34 (1), 95-103.

Kutner, M.H., Nachtsheim, C.J., Neter, J., Li, W., 2005. Applied linear statistical models, 5th ed. McGraw-Hill Irwin, New York.

Lachniet, M.S., Asmerom, Y., Polyak, V., 2011. Deglacial paleoclimate in the southwestern United States: an abrupt 18.6 ka cold event and evidence for a North

Atlantic forcing of Termination I. Quaternary Science Reviews 30 (27–28), 3803-3811.

Langenheim, V.E., Glen, J.M., Jachens, R.C., Dixon, G.L., Katzer, T. and Morin, R.L., 2000. Geophysical constraints on the Virgin River Depression, Nevada, Utah, and Arizona. 00-407, U.S. Dept. of the Interior, U.S. Geological Survey, Menlo Park, California.

Lawrence, F.W., Upchurch, S.B., 1982. Identification of recharge areas using geochemical factor analysis. Ground Water 20 (6), 680-687.

Le Maitre, R.W., 1982. Numerical Petrology: Statistical Interpretation of Geochemical Data. Elsevier Scientific Publishing Company, New York.

Leontiadis, I.L., Nikolaou, E., 1999. Environmental isotopes in determining groundwater flow systems, northern part of Epirus, Greece. Hydrogeology Journal 7 (2), 219-226.

Lloyd, J.W., 1986. A review of aridity and groundwater. Hydrological Processes 1 (1), 63-78.

Lloyd, J.W., Heathcote, J.A., 1985. Natural inorganic hydrochemistry in relation to groundwater :an introduction. Clarendon Press, Oxford, Oxfordshire.

Logan, M., Logan, M., 2010. Biostatistical design and analysis using R :a practical guide. Wiley-Blackwell; Wiley-Blackwell, Chichester, UK; Hoboken, NJ; Chichester, UK; Hoboken, NJ.

Mahlknecht, J., Horst, A., Hernández-Limón, G., Aravena, R., 2008. Groundwater geochemistry of the Chihuahua City region in the Rio Conchos Basin (northern Mexico) and implications for water resources management. Hydrological Processes 22 (24), 4736-4751.

Matter, J., Waber, H.N., Loew, S., Matter, A., 2006. Recharge areas and geochemical evolution of groundwater in an alluvial aquifer system in the Sultanate of Oman. Hydrogeology Journal 14 (1-2), 203-224.

Maxey, G.B., 1968. Hydrogeology of desert basins. Ground Water 6 (5), 10-22.

McIntyre, R.M., Blashfield, R.K., 1980. A Nearest-Centroid Technique for Evaluating the Minimum-Variance Clustering Procedure. Multivariate Behavioral Research 15 (2), 225-238.

McNeil, V.H., Cox, M.E., Preda, M., 2005. Assessment of chemical water types and their spatial variation using multi-stage cluster analysis, Queensland, Australia. Journal of Hydrology 310 (1–4), 181-200.

Metcalf, L., 1995. Groundwater-surface water interactions in the lower Virgin River area, Arizona and Nevada. MSc Thesis, University of Nevada Las Vegas, Las Vegas, USA. 181p.

Mifflin, M.D., Hess, J.W., 1979. Regional carbonate flow systems in Nevada. Journal of Hydrology 43 (1-4), 217-237.

Milligan, G.W., 1996. Clustering validation: results and implications for applied analyses, in: Arabie, P., Hubert, L.J., Soete, G.D. (Eds.), Clustering and classification. World Scientific, Singapore, pp. 341-374 .

Milligan, G.W., Cooper, M.C., 1985. An examination of procedures for determining the number of clusters in a data set. Psychometrika 50 (2), 159-179.

Naff, R.L., Maxey, G.B. and Kaufmann, R.F., 1974. Interbasin ground-water flow in southern Nevada : prepared in conjunction with Geological Society of America

Cordilleran Section Meeting, Las Vegas, Nev., March 1974. 20, Mackay School of Mines, University of Nevada, Reno, 28 p.

Nelson, S., Anderson, N.K., Mayo, A., 2005. Reply to [“Comment on “Testing the interbasin flow hypothesis at Death Valley, California’”] Winograd et al. Eos Transactions American Geophysical Union 86 (32), 296-296.

Nelson, S.T., Anderson, K., Mayo, A.L., 2004. Testing the interbasin flow hypothesis at Death Valley, California. Eos Transactions American Geophysical Union 85 (37), 349-349.

Page, W.R., Dixon G.L., Rowley P.D., Brickey D.W., 2005. Geologic map of parts of the Colorado, White River, and Death Valley groundwater flow systems, Nevada, Utah, and Arizona. 1:250,000ma.

Page, W.R., Scheirer, D.S. and Langenheim, V.E., 2006. Geologic cross sections of parts of the Colorado, White River, and Death Valley regional ground-water flow systems, Nevada, Utah, and Arizona. 2006-1040, U.S. Geological Survey, Reston.

Panno, S.V., Hackley, K.C., Cartwright, K., Liu, C.L., 1994. Hydrochemistry of the Mahomet Bedrock Valley Aquifer, East-Central Illinois: Indicators of Recharge and Ground-Water Flow. Ground Water 32 (4), 591-604.

Parkhurst, D.L., 1997. Geochemical mole-balance modeling with uncertain data. Water Resources Research 33 (8), 1957-1970.

Parkhurst, D.L., Appelo, C.A.J., 1999. User's guide to PHREEQC, Version 2: a computer program for speciation, batch-reaction, one-dimensional transport, and inverse geochemical calculations. U.S. Geological Survey: Earth Science Information Center, Open-File Reports Section distributor, Denver.

Pearson, J.F.J., Balderer, W., Loosli, H.H., Lehmann, B.E., Matter, A., Peters, T., et al, 1991. Applied isotope hydrogeology: A case study in Northern Switzerland. Elsevier Science Publishers B.V., Amsterdam, Netherlands.

Plummer, L.N., Back, W., 1980. The mass balance approach: Application to interpreting the chemical evolution of hydrologic systems. American Journal of Science 280, 130-142.

Plummer, L.N., Bexfield, L.M., Anderholm, S.K., Sanford, W.E. and Busenberg, E., 2004. Geochemical characterization of ground-water flow in the Santa Fe Group aquifer system, Middle Rio Grande basin, New Mexico. Water-Resources Investigation Report03-4131, U.S. Geological Survey, Reston, 395p.

Plummer, L.N., Parkhurst, D.L., Thorstenson, D.C., 1983. Development of reaction models for ground-water systems. Geochimica et Cosmochimica Acta 47 (4), 665-686.

Plummer, N.L., Busby, J.F., Lee, R.W., Hanshaw, B.B., 1990. Geochemical modeling of the Madison aquifer in parts of Montana, Wyoming, and south Dakota. Water Resources Research 26 (9), 1981-2014.

R Development Core Team, 2010. R: A Language and environment for statistical computing.

Reimann, C., Filzmoser, P., Garrett, R., 2002. Factor analysis applied to regional geochemical data: problems and possibilities. Applied Geochemistry 17 (3), 185-206.

Reimann, C., Filzmoser, P., Garrett, R., Dutter, R., 2008. Statistical data analysis explained :applied environmental statistics with R. John Wiley & Sons, Chichester, England; Hoboken, New Jersey.

Riley, J.A., Steinhorst, R.K., Winter, G.V., Williams, R.E., 1990. Statistical analysis of the hydrochemistry of ground waters in Columbia River basalts. Journal of Hydrology 119 (1-4), 245-262.

Romesburg, H.C., 1984. Cluster analysis for researchers. Lifetime Learning Publications, Belmont, California.

Rosenthal, E., Adar, E., Issar, A.S., Batelaan, O., 1990. Definition of groundwater flow patterns by environmental tracers in the multiple aquifer system of southern Arava Valley, Israel. Journal of Hydrology 117 (1-4), 339-368.

Schwartz, F.W., Muehlenbachs, K., Chorley, D.W., 1981. Flow-system controls of the chemical evolution of groundwater; Symposium on geochemistry of groundwater. Journal of Hydrology 54 (1-3), 225-243.

Seyhan, E., Van De Griend, A.A., Engelen, G.B., 1985. Multivariate analysis and interpretation of the hydrochemistry of a dolomitic reef aquifer, northern Italy. Water Resources Research 21 (7), 1010-1024.

Shaw, P., J.A., 2003. Multivariate Statistics for the environmental sciences, First ed. Hodder Arnold, a member of the Hodder Headline Group, London, UK.

Smith, G.I., Friedman, I., Veronda, G., Johnson, C.A., 2002. Stable isotope compositions of waters in the Great Basin, United States 3. Comparison of groundwaters with modern precipitation. Journal of Geophysical Research 107 (D19), 4402-4416.

Steinhorst, R.K., Williams, R.E., 1985. Discrimination of groundwater sources using cluster analysis MANOVA, canonical analysis and discriminant analysis. Water Resources Research 21 (8), 1149-1156.

Stetzenbach, K.J., Hodge, V.F., Guo, C., Farnham, I.M., Johannesson, K.H., 2001. Geochemical and statistical evidence of deep carbonate groundwater within overlying volcanic rock aquifers/aquitards of southern Nevada, USA. Journal of Hydrology 243 (3–4), 254-271.

Stuyfzand, P.J., 1999. Patterns in groundwater chemistry resulting from groundwater flow. Hydrogeology Journal 7 (1), 15-27.

Suk, H., Lee, K., 1999. Characterization of a ground water hydrochemical system through multivariate analysis; clustering into ground water zones. Ground Water 37 (3), 358-366.

Tabachnick, B.G., Fidell, L.S., 2007. Using multivariate statistics. Pearson and Allyn & Bacon, Boston.

Thomas, J.M., 1999. Comment on 'Isotope hydrology of southern Nevada groundwater: Stable isotopes and radiocarbon' by M. L. Davisson et al. Water Resources Research 35 (11), 3577-3583.

Thomas, J.M., Calhoun, S.C. and Apambire, W.B., 2001. A deuterium Mass-balance interpretation of groundwater sources and flows in southern Nevada. Publication No. 41169, Desert Research Institute, Nevada, 46p.

Thomas, J.M., Welch, A.H. and Dettinger, M.D., 1996. Geochemistry and Isotope hydrology of representative aquifers in the Great Basin Region of Nevada, Utah, and

Adjacent States. Professional Paper1409-C, U.S Geological Survey, Washington, DC, 100p.

Thomas, J.M., Welch, A.H., Preissler, A.M., 1989. Geochemical evolution of ground water in Smith Creek Valley—a hydrologically closed basin in central Nevada, U.S.A. Applied Geochemistry 4 (5), 493-510.

Thyne, G., Guler, C., Poeter, E., 2004. Sequential analysis of hydrochemical data for watershed characterization. Ground Water 42 (5), 711-723.

Thyne, G.D., Gillespie, J.M., Ostdick, J.R., 1999. Evidence for interbasin flow through bedrock in the southeastern Sierra Nevada. Geological Society of America Bulletin 111 (11), 1600-1616.

Tibshirani, R., Walther, G., 2005. Cluster Validation by Prediction Strength. Journal of Computational and Graphical Statistics 14 (3), 511-528.

Tóth, J., 1963. A theoretical analysis of groundwater flow in small drainage basins. Journal of Geophysical Research 68 (16), 795-812.

Tóth, J., 2009. Gravitational systems of groundwater flow : theory, evaluation, utilization. Cambridge University Press, Cambridge; New York.

U.S EPA , 2009. STORET Data Warehouse. Available at: http://www.epa.gov/storet/dw_home.html. Accessed 04/05, 2009.

U.S Geological Survey, 2009. *National Water Information System (NWISWeb) data*. Available at: http://nwis.waterdata.usgs.gov/usa/nwis/qwdata. Accessed 06/10, 2009.

U.S. Geological Survey, 2000. Desert basins of the southwest. USGS Fact Sheet 086-00, U.S. Dept. of the Interior, U.S. Geological Survey, Reston, 4p.

Williams, R.E., 1982. Statistical identification of hydraulic connections between the surface of a mountain and internal mineralized sources. Ground Water 20 (4), 466-478.

Williams, V.S., 1996. Preliminary geologic map of the Mesquite quadrangle, Clark and Lincoln counties, Nevada, and Mohave County, Arizona. 1:24 000.

Winograd, I.J., 2001. Interbasin groundwater flow in south central Nevada: A further comment on the discussion between *Davisson et. al.*. [1999a, 1999b] and *Thomas* [1999]. Water Resources Research 37 (2), 431-433.

Winograd, I.J., Fridrich, C.J., Sweetkind, D., Belcher, W.R., Thomas, J.M., 2005. Comment on "Testing the Interbasin Flow Hypothesis at Death Valley, California". Eos Transactions American Geophysical Union 86 (32), 295-296.

Winograd, I.J., Friedman, I., 1972. Deuterium as a tracer of regional ground-water flow, Southern Great Basin, Nevada and California. Geological Society of America Bulletin 83 (12), 3691-3708.

Winograd, I.J., Riggs, A.C., Coplen, T.B., 1998. The relative contributions of summer and cool-season precipitation to groundwater recharge, Spring Mountains, Nevada, USA; Special issue in honor of Eugene S. Simpson, founding editor. Hydrogeology Journal 6 (1), 77-93.

Winograd, I.J., Thordarson, W., 1975. Hydrogeologic and hydrochemical framework, South-Central Great Basin, Nevada-California; with special reference to the Nevada test site. U.S Geological Survey Professional Paper 712-C, 126p.

Woocay, A., Walton, J., 2008. Multivariate Analyses of Water Chemistry: Surface and Ground Water Interactions. Ground Water 46 (3), 437-449.

Yelken, M.A., 1996. Trace element analysis of selected springs in the Virgin River Basin. MSc Thesis, University of Nevada Las Vegas, Las Vegas, Nevada, USA, 128p.

Zhu, G.F., Li, Z.Z., Su, Y.H., Ma, J.Z., Zhang, Y.Y., 2007. Hydrogeochemical and isotope evidence of groundwater evolution and recharge in Minqin Basin, Northwest China. Journal of Hydrology 333 (2), 239-251.

VITA

Graduate College
University of Nevada, Las Vegas

Joseph Asante

Home Address:
4386 Escondido Street
Las Vegas NV 89119

Degrees:
Bachelor of Science, Geology, 1998
University of Ghana Legon, Ghana

Master of Science, Environmental Engineering, 2003
Technische Universität Hamburg-Harburg, Germany

Awards and Scholarships:

- Urban Sustainability Initiative Graduate Fellowship, UNLV, Fall 2009-Spring 2012
- Teaching Assistantship, Department of Geoscience, UNLV, Spring 2007-Spring 2009
- Geoscience Geosymposium most outstanding poster award, UNLV, April, 2009
- International Graduate Access Scholarship, UNLV, Spring 2008, 2011, & 2012
- Department of Geoscience Graduate Scholarship, UNLV,2008 & 2009
- Geological Society of America Rocky Mountain (63rd Annual) and Cordilleran (107th Annual) Joint Meeting (18–20 May 2011) travel grant recipient, May, 2011
- Geological Society of America 2011 Annual Meeting (9–12 October 2011) travel grant recipient, October, 2011

Publications:

Asante, J., Kreamer, D.K., Cross, C.L, and Y. Michael, (in review), Optimizing Multivariate Statistical Characterization of Hydrochemical Groundwater Data: A Comparative Analysis of Clustering Methods

Asante, J., Kreamer, D.K., and Y. Michael, (to be submitted), Basin Hydrochemistry and Related Flow Systems: Integrated Data Analytic Methods.

Asante, J., Kreamer, D.K., and Y. Michael, (to be submitted), Evaluating Interbasin Flow and Groundwater-Surface Water Interactions, lower Virgin River Basin USA: Unusual Statistical and Graphical Analysis of Stable Isotopes, and Solutes

Abstracts:

Asante, J., and D.K., Kreamer, Basin Hydrochemistry and Related Flow Systems: Integrated Data Analytic Methods. GSA Annual Meeting & Exposition, October, 9-12, 2011 (abstract and oral presentation)

Asante, J., D.K., Kreamer, and C.L Cross, Enhancing Multivariate Statistical Characterization of Hydrochemical Groundwater Data: A Comparative Analysis of Clustering Methods. GSA Joint Rocky Mountain/Cordilleran Conference, May 17-20, 2011 Logan, Utah (abstract and oral presentation)

Asante, J. and D.K., Kreamer, Enhancing the Definition of Hydrochemical Facies to Evaluate Recharge Sources and Hydrodynamics of the Lower Virgin Basin, USA. Nevada Water

Resources Association Annual Conference, March 2-4, 2010, Las Vegas, NV (abstract and poster presentation)

Asante, J. and D.K., Kreamer, Evaluating Interbasin Flow and Groundwater-Surface Water Interactions, lower Virgin River Basin USA: Unusual Statistical and Graphical Analysis of Stable Isotopes, and Solutes. NGWA Ground Water Summit, May 2012 Garden Grove, CA (abstract accepted for oral presentation in May, 2012)

Dissertation Title: Evaluating Recharge and Dynamics of Flow in the Lower Virgin River Basin, USA: Interpretation of Hydrochemical and Stable Isotopic Data

Dissertation Examination Committee:
Chairperson, Dr. David Kreamer, Ph.D.
Committee Member, Dr. Michael Young, Ph.D.
Committee Member, Dr. Matthew Lachniet, Ph.D.
Committee Member, Dr. Ganqing Jiang, Ph.D.
Graduate Faculty Representative, Dr. Ding Zhonghai, Ph.D.