

# Sequential Analysis of Hydrochemical Data for Watershed Characterization

by Geoffrey Thyne<sup>1</sup>, Cüneyt Güler<sup>2</sup>, and Eileen Poeter<sup>1</sup>

---

## Abstract

A methodology for characterizing the hydrogeology of watersheds using hydrochemical data that combine statistical, geochemical, and spatial techniques is presented. Surface water and ground water base flow and spring runoff samples (180 total) from a single watershed are first classified using hierarchical cluster analysis. The statistical clusters are analyzed for spatial coherence confirming that the clusters have a geological basis corresponding to topographic flowpaths and showing that the fractured rock aquifer behaves as an equivalent porous medium on the watershed scale. Then principal component analysis (PCA) is used to determine the sources of variation between parameters. PCA analysis shows that the variations within the dataset are related to variations in calcium, magnesium,  $\text{SO}_4$ , and  $\text{HCO}_3$ , which are derived from natural weathering reactions, and pH,  $\text{NO}_3$ , and chlorine, which indicate anthropogenic impact. PHREEQC modeling is used to quantitatively describe the natural hydrochemical evolution for the watershed and aid in discrimination of samples that have an anthropogenic component. Finally, the seasonal changes in the water chemistry of individual sites were analyzed to better characterize the spatial variability of vertical hydraulic conductivity. The integrated result provides a method to characterize the hydrogeology of the watershed that fully utilizes traditional data.

---

## Introduction

The objective of this paper is to describe a methodology that uses hydrochemical data to improve characterization of watershed hydrology. The methodology is robust and objective, and uses standard statistical, spatial analysis and inverse geochemical modeling in mutually supportive sequential fashion unlike previous studies. This technique was developed to define major trends in the hydrochemical evolution of a regional-scale system of watersheds (Güler and Thyne 2004). The purpose of this paper is to describe that methodology in detail and test the procedure on a smaller watershed. In doing so, we found that the technique provided improved understanding of a study area's watershed hydrology, as well as the sources and factors control-

ling ground water and surface water quality including discriminating natural background and anthropogenic impact.

Excluding anthropogenic impacts, the chemical composition of surface water and ground water is controlled by many factors that include the composition of precipitation, mineralogy of the watershed and aquifers, climate, and topography. These factors can combine to create diverse water types that change in composition spatially and temporally. The use of major ions as natural tracers (Back 1966) has become a common method to delineate flowpaths in aquifers. Generally, the approach is to divide the samples into hydrochemical facies (water types), that is, groups of samples with similar chemical characteristics that can then be correlated with location. The spatial variability observed can provide insight into aquifer heterogeneity and connectivity as well as the physical and chemical processes controlling water chemistry.

Underlying this approach are a number of assumptions. (1) Natural water chemistry is a result of rock-water reactions such as dissolution/precipitation, reactions on aquifer surfaces, and biological reactions. (2) Distinctive chemical signatures are related to specific sets of reactions. (3) Dissolved concentrations generally increase along the

---

<sup>1</sup>Colorado School of Mines, Department of Geology and Geological Engineering, Golden, CO 80401

<sup>2</sup>Mersin Üniversitesi, Çiftlikköy Kampüsü, Jeoloji Mühendisliği Bölümü, 33343 Mersin, Turkey

Received July 2003, accepted December 2003.

Copyright © 2004 by the National Ground Water Association.

subsurface flowpath until major components reach maximum values dictated by mineral equilibrium. (4) Hydrochemical facies are directly related to the dominant processes.

In recent years, multivariate statistical methods have been employed to extract critical information from hydrochemical datasets in complex systems. These techniques can help resolve hydrological factors such as aquifer boundaries, ground water flowpaths, or hydrochemical components (Seyhan et al. 1985; Usunoff and Guzman-Guzman 1989; Razack and Dazy 1990; Join et al. 1997; Ochsekuehn et al. 1997; Liedholz and Schafmeister 1998; Suk and Lee 1999; Wang et al. 2001; Locsey and Cox 2003), identify geochemical controls on composition (Adams et al. 2001; Alberto et al. 2001; Lopez-Chicano et al. 2001; Reeve et al. 1996), and separate anomalies such as anthropogenic impacts from the background (Hernandez et al. 1991; Birke and Raush 1993; Helena et al. 2000; Pereira et al. 2003) on a variety of scales (Briz-Kishore and Murali 1992). These studies often use either or both R-mode (source of variation in a dataset from parameters) or Q-mode (variation between samples) multivariate techniques (Dalton and Upchurch 1978; Turk 1979; Upchurch 1979), but to date none has fully integrated the results of various analyses in a sequential fashion to better constrain the interpretation and include inverse geochemical modeling in the process.

## Methods

### Database Quality Assurance and Quality Control Measures

The first step is to verify the quality of the data. In this case, the water chemistry data used were collected by the U.S. Geological Survey (USGS) in cooperation with Jefferson County (U.S. Geological Survey 2001) from 74 individual domestic water-supply wells (IDWs) and 16 surface water sites along Turkey Creek Basin (TCB). Located in the mountainous part of Jefferson County ~32 km west of Denver, Colorado, this area encompasses ~116.5 km<sup>2</sup> (Figure 1). Data and data collection techniques are described in Morgan (2000). The database consists of 180 samples, collected from two separate sampling campaigns (each with 90 samples from identical locations shown in Figure 2) representing spring runoff (June 14–29, 1999) and fall base flow (October 1–November 3, 1999) conditions. Of the 38 hydrochemical variables measured (consisting of physical properties, major ions, minor ions, and trace elements), 12 parameters, specific conductance (SC), pH, dissolved oxygen (DO), calcium, magnesium, sodium, potassium, chlorine, SO<sub>4</sub>, HCO<sub>3</sub>, fluorine, and NO<sub>3</sub>+NO<sub>2</sub> (total as nitrogen) were utilized in the statistical analyses. The remaining 26 variables are not used because they contain a high percentage (40% to 100%) of censored values (i.e., concentration values reported as less than) that are not appropriate for many multivariate statistical techniques (Farnham et al. 2002; Güler et al. 2002).

Charge balance errors (%CBE =  $[\sum \text{cations} - \sum \text{anions}] / [\sum \text{cations} + \sum \text{anions}] \times 100$ ) average -1.1 for the runoff dataset and -0.2 for the base flow dataset, with standard deviations of 4.1% and 2.9%, respectively. More than 72%

of the runoff samples have a CBE within  $\pm 5\%$ , whereas this percentage is much higher for the base flow dataset (92%) and errors are evenly distributed between positive and negative values. No samples in the database have a CBE  $> \pm 10\%$ . ArcView GIS 3.2a (Environmental Systems Research Institute 1996) is used for data integration, retrieval, and visualization. Raster and vector data layers from the USGS include 1:24,000 scale digital elevation models, roads, surface water network, address locations, and basin/administrative boundaries.

### Sequence of Analysis

The advantage of this technique is the ability to use standard methods in a sequential fashion where each step builds on the prior analysis, providing increasing confidence and greater insight into the hydrochemical evolution of the watershed. Figure 2 is a flowchart of the methodology. The procedure begins with hierarchical cluster analysis (HCA) of the data to cluster similar samples. The statistical clusters are examined for spatial coherence to verify that the clusters have a physical basis. Next, the data are analysed by principal component analysis (PCA) to determine what factors (groups of parameters) account for the numerical variation of the clusters. This step serves to provide insight into the hydrochemical processes and sources of solutes. Next, a preliminary hydrochemical evolution model is formulated based on the clusters. This preliminary model is confirmed with inverse geochemical modeling that is constrained by aquifer mineralogy and the mineral saturation indices. This process produced a set of hydrochemical facies that defined flowpaths and separated natural background from anthropogenic impacts, while the spatial distribution of seasonal variation in facies provided information about local aquifer properties.

### Statistical Analyses

STATISTICA<sup>®</sup> Release 5.5 (StatSoft Inc., Tulsa, Oklahoma) is used to analyze the hydrochemical data. The statistical analysis of the runoff and base flow datasets reveals that the variables SC, calcium, magnesium, sodium, potassium, chlorine, SO<sub>4</sub>, fluorine, and NO<sub>3</sub>+NO<sub>2</sub> are skewed positively; that is, the data contain a few high values. This is common for most naturally occurring element distributions (Miesch 1976). These nine variables are log-transformed to more closely correspond to normally distributed data. Variables pH, DO, and HCO<sub>3</sub> follow a normal distribution, so are not log-transformed.

Kolmogorov-Smirnov normality test results confirm that the variables SC, calcium, magnesium, sodium, potassium, chlorine, SO<sub>4</sub>, fluorine, and NO<sub>3</sub>+NO<sub>2</sub> are lognormally distributed ( $p$  values  $< 0.05$ ), while the variables pH, DO, and HCO<sub>3</sub> are normally distributed (their  $p$  values  $> 0.2$ ). If the Kolmogorov-Smirnov D statistic is significant (i.e.,  $p < 0.05$ ), then the hypothesis that the respective distribution is normal is rejected (Swan and Sandilands 1995; StatSoft Inc. 1997). Subsequently, all 12 variables for both datasets are standardized to their standard scores ( $z$ -scores) as described by Güler et al. (2002). Standardization scales the data to a range of  $\sim -3$  to  $+3$  standard deviations ( $\sigma$ ), centered about a mean ( $\mu$ ) of zero, giving each variable equal weight in the multivariate statistical analyses. Without

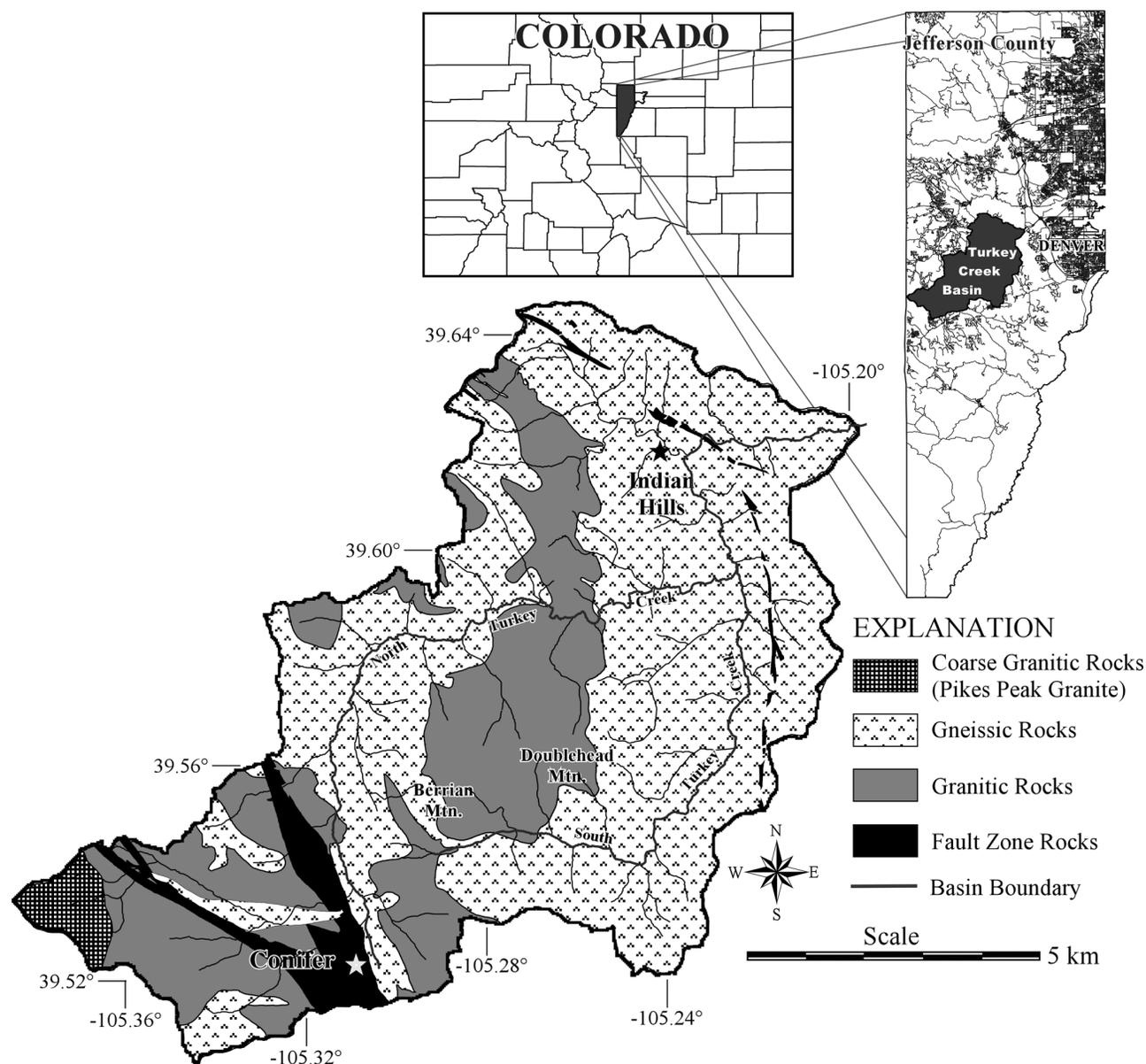


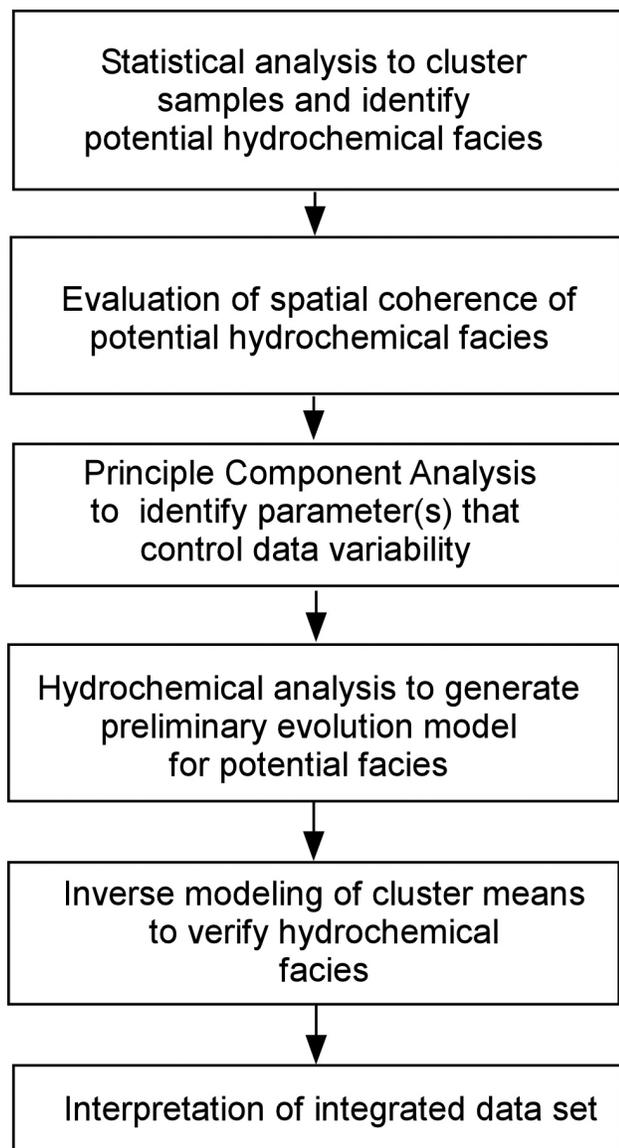
Figure 1. Location and geology of TCB. Geologic map of the basin compiled by the U.S. Geological Survey (2001) based on the map by Trimble and Machette (1979).

scaling, the results are influenced most strongly by the variable with the greatest magnitude (Judd 1980). HCA and PCA were used for the multivariate analyses. Detailed technical descriptions of HCA and PCA techniques are provided in StatSoft Inc. (1997) and are described briefly here.

HCA is a powerful tool for analyzing water chemistry data (Seyhan et al. 1985; Reeve et al. 1996; Ochsenkuhn et al. 1997) and has been used to formulate geochemical models (Meng and Maynard 2001). This method groups samples into distinct populations (clusters) that may be significant in the geologic/hydrologic context, as well as from a statistical point of view (Güler et al. 2002). Assumptions of the HCA technique include homoscedasticity (equal variance) and normal distribution of the variables (Alther 1979). To determine the relation between water samples, the standardized data matrix is imported into the statistics package. STATISTICA offers seven similarity/dissimilarity measurements and seven linkage methods.

Individual samples are compared with the specified similarity/dissimilarity, using the selected linkage method, and grouped into clusters. Ward's linkage method, which iteratively links similar samples by using the distance matrix, was used in this analysis. A classification scheme using Euclidean distance for similarity measurement, together with Ward's method for linkage, produces the most distinctive groups where each member within the group is more similar to its fellow members than to any member outside the group (Güler et al. 2002).

As a multivariate data analytic technique, PCA reduces a large number of variables to a small number of variables, without sacrificing too much of the information (Qian et al. 1994). More concisely, PCA combines two or more correlated variables into one variable. This approach has been used to extract related variables and infer the processes that control water chemistry (Helena et al. 2000; Hildago and Cruz-Sanjulian 2001). Units of measurement are very



**Figure 2.** Flowchart of sequential methodology.

important since the principal components (PCs) are meaningful only if all the variables are measured in the same units. For that reason, the standardized data matrix is used. Varimax rotation is applied to the PCs in order to find factors that can be more easily explained in terms of hydrochemical or anthropogenic processes (Helena et al. 2000). This rotation is called varimax because the goal is to maximize the variance (variability) of the new variable, while minimizing the variance around the new variable (StatSoft Inc. 1997). The number of PCs extracted (to explain the underlying data structure) is defined by using the Kaiser criterion (Kaiser 1960) where only the PCs with eigenvalues greater than unity are retained. In other words, unless a PC extracts at least as much information as the equivalent of one original variable, it is dropped (StatSoft Inc. 1997).

#### Inverse Geochemical Modeling

PHREEQC was used to calculate aqueous speciation and mineral saturation indices for samples along the topographic flowpath. Inverse modeling in PHREEQC uses the mass-balance approach to calculate all the stoichiometrically available reactions that can produce the observed

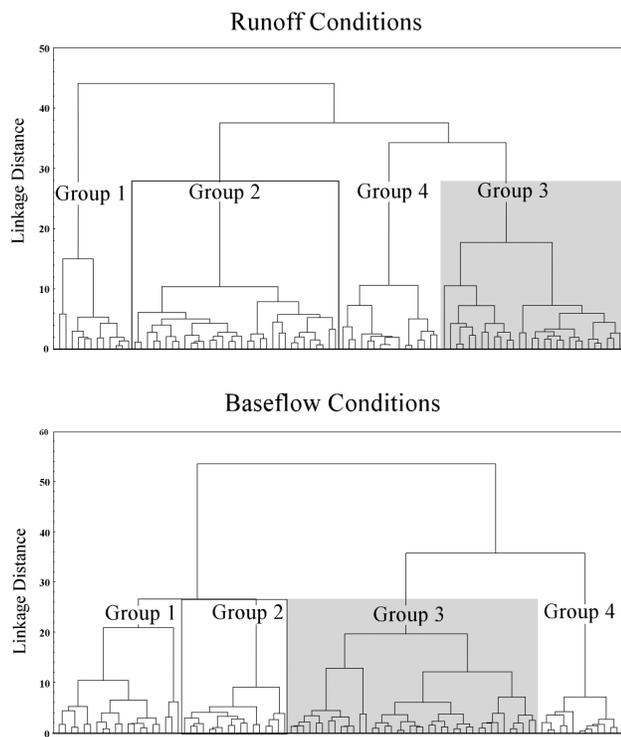
chemical changes between end-member waters (Plummer and Back 1980). This mass balance technique has been used to quantify reactions controlling water chemistry along flowpaths (Busby et al. 1991; Hidalgo and Cruz-Sanjulian 2001) and quantify mixing of end-member components in a flow system (Kuells et al. 2000). In TCB, mean values of the HCA-defined clusters provide starting and ending water compositions for the inverse models. Using the mean values of the clusters preserves information about relative differences in abundance between ions, but provides input that better utilizes the inverse modeling approach by minimizing the natural variability (numerical noise) inherent in water samples. Minerals used in the inverse geochemical models are limited to those present in the study area. Finally, the mineral reaction mode (dissolution or precipitation) is constrained by the saturation indices for each mineral, and atmospheric gases carbon dioxide and oxygen are included in the list of potential reactive components. The base flow dataset is used in the inverse modeling calculations because it exhibits the normal background chemical pattern free of the stream dilution observed during spring runoff and most closely represents a steady-state condition.

#### Study Area

##### Geology and Hydrology

The fractured-crystalline rocks serve as the principal aquifer to a depth of at least 300 m. The semiarid basin has rugged topography ranging in elevation from 3000 m above mean sea level (amsl) on the southwest to 1830 m amsl near the basin outlet on the northeast (Figure 1). Rapid development, most of which occurred in the last few decades, increases the importance of quantifying and mitigating problems regarding water quantity and quality in this area. Two important issues resulting from growth are sufficient water supply and appropriate waste water disposal. At present, the ground water quality meets U.S. Environmental Protection Agency standards for drinking water, but degradation of water quality has been documented during the past 25 years (Bossong et al. 2003).

TCB is underlain by Precambrian (Proterozoic) fractured crystalline rocks, which compose the core of the Front Range of the Rocky Mountains. The locations of the major rock groups are shown in Figure 1. The Precambrian rocks in the area are an interlayered, generally conformable, sequence of gneiss, migmatite, and Boulder Creek Granite and Silver Plume Quartz monzonite intrusive igneous rocks, some of which are metamorphosed (Sims and Gable 1967). The gneisses exhibit complex structural deformation including a west-northwest folding episode associated with upper amphibolite metamorphism (Taylor 1976) and a subsequent deformational event that produced large-scale north-northeast trending folds. These rocks are intruded by numerous small dikes and irregular plutons of porphyritic igneous rocks of early Tertiary (Laramide) age and cut by abundant steep faults (Sims 1989). Unconsolidated sand and gravel locally overlay bedrock (especially along the stream channels); however,



**Figure 3. Dendograms generated from HCA of hydrochemical data from base flow and runoff seasons showing associations between samples from different parts of TCB.**

these deposits are thin, narrow, and discontinuous (Folger et al. 1996).

Rugged terrain and crystalline rocks at or near the surface make construction of reservoirs and pipelines difficult (Hofstra and Hall 1975); thus, most homes rely on IDWs drilled to depths ranging from 30 to 244 m and individual sewage disposal systems (ISDSs). Fractured-crystalline rock forms the aquifer that provides water to IDW and is the ultimate destination of processed effluent from ISDSs. Located within the TCB boundary are ~4900 homes pumping ~2.3 million m<sup>3</sup>/yr of ground water from the fractured rock aquifer (U.S. Geological Survey 2001). Many areas have very thin soil and it is sometimes necessary to import gravel to provide an effective percolation zone for ISDSs; in other cases, bedrock is excavated and the broken rock is used as percolation material.

TCB has a mean annual precipitation of ~500 mm/yr and is drained by Turkey Creek. Water level contours from the state engineer's office and U.S. Geological Survey (2001) indicate a shallow water table (averaging < 30 m below surface) that mimics topography, revealing that discharge of surface water and ground water is focused in a narrow canyon in the northeast. Streamflow volumes become relatively high in spring for 4 to 6 wk (up to 8496 L/s), then rapidly decline to low base flow volumes (8.5 to 28 L/s) for the rest of the year, with infrequent summer and fall precipitation events that produce short-term increases in streamflow. This indicates that most of the water remaining after evapotranspiration does not infiltrate the deep aquifer; rather, it flows overland or as shallow, rapid, lateral ground water flow to discharge into the streams. Evapotranspiration measurements indicate, on average, ~85% of precipitation evapotranspires (U.S. Geo-

logical Survey 2001). Of the remaining water, 75% is pumped for residential use (2.3 million m<sup>3</sup>/yr [U.S. Geological Survey 2001]). In theory, residential use is returned to the subsurface via ISDSs to supply stream base flow, ground water outflow, or recharge to storage. However, some of the ISDS water may flow laterally through the regolith, with some portion seeping into the deep fracture system and the remainder discharging to streams.

Well hydrographs exhibit an abrupt rise each spring and decline the rest of the year, indicating the major source of recharge is melting snow. Base flow is derived from extended ground water discharge of the limited, but rapid, vertical recharge during snowmelt. Most of the development in TCB occurred in the last few decades as documented by well completions and first beneficial uses reported in well records. Current development is broken down as 19% occurring before 1960, 16% during 1960–1970, 38% during 1970–1980, 11% during 1980–1990, and 16% since 1990 to the present. Limited long-term water level data exhibit an average water level decline of ~0.03 m/yr from 1973 to 1998. All wells measured during the past four years reveal an average decline of ~0.4 m/yr from 1998 to 2001. Thus, the hydrographs exhibit a long-term decline indicating the basin has not reached hydraulic equilibrium.

In spite of several recent studies, there remain questions about hydrological issues including the scale at which the fractured aquifer acts as an equivalent porous medium, the appropriate distribution of aquifer properties such as storage on the local scale within the aquifer, scale of flow-paths, the degree of hydraulic connectivity between surface water and ground water, and the sources of anthropogenic impacts (Bossong et al. 2003; Poeter et al. 2003; Vander-Beek 2003). One measure of the utility of the proposed methodology is the information it provides to address these questions.

## Results

The first step in the method is to cluster the hydrochemical data using HCA. The Q-mode analysis will group samples together based on similarity in multidimensional space. This technique is useful in providing discrimination between samples. The input data are not confined to just chemical parameters, but can include physical measurements such as temperature. In this example, we have broken the data into base flow and spring runoff before clustering in order to evaluate any temporal effects on water chemistry.

### Hierarchical Cluster Analysis

#### Spring Runoff Conditions

The spring runoff dataset were classified in 12-dimensional space and presented in a dendrogram (Figure 3). Four preliminary groups are selected based on visual examination of the dendrogram, each representing a hydrochemical facies with means for each parameter shown in Table 1. The choice of number of clusters is subjective. Choosing the optimal number of groups depends on the researcher since there is no test to determine the optimum number of

**Table 1**  
**Mean Water Chemistry of the HCA Water Groups and Local Precipitation**

Group	N <sup>a</sup>	pH	S.C.	O <sub>2</sub>	Ca	Mg	Na	K	Cl	SO <sub>4</sub>	HCO <sub>3</sub>	F	NO <sub>3</sub> +NO <sub>2</sub>	TDS <sup>b</sup>	CB <sup>c</sup>
<b>Spring Runoff</b>															
Group 1	12	6.51	108.4	5.6	6.4	1.8	16.0	1.03	6.3	7.3	52.2	0.73	0.74	76.9	-3.09
Group 2	33	7.47	299.3	2.9	35.6	9.4	13.6	1.17	8.2	14.3	166.6	1.26	0.55	176.5	-2.82
Group 3	29	6.68	478.4	4.0	52.6	12.0	24.8	1.88	53.0	30.4	142.8	0.50	4.07	256.6	2.07
Group 4	16	8.17	338.1	8.2	30.9	7.2	24.2	2.52	48.3	17.2	95.8	0.65	0.21	186.2	-1.84
<b>Base Flow</b>															
Group 1	20	7.03	169.2	6.1	13.8	3.8	17.8	0.69	5.9	5.9	92.4	0.55	0.47	94.6	-0.15
Group 2	17	7.75	286.9	2.4	31.3	10.4	15.7	1.26	4.4	9.8	175.4	1.50	0.26	162.6	-2.39
Group 3	40	6.89	503.2	3.0	59.7	13.4	21.9	2.25	50.9	34.2	177.5	0.48	3.13	272.1	0.80
Group 4	13	8.00	555.8	9.3	49.6	13.7	36.4	2.95	87.8	14.1	152.9	0.61	0.28	286.3	-0.34
Ppt <sup>d</sup>		4.92			0.15	0.02	0.03	0.03	0.06	0.42			0.89		

<sup>a</sup>Number of samples within groups  
<sup>b</sup>Total dissolved solids (mg/L); pH (standard units); specific conductance (μS/cm), mean concentrations (mg/L)  
<sup>c</sup>Percent charge-balance  
<sup>d</sup>National Atmospheric Deposition Program/National Trends Network 1999, annual precipitation-weighted mean concentrations for site CO02, 1999

groups in the dataset. This is a universal problem in all the statistical clustering schemes, sometimes called the cluster validity problem. This is why we use the criteria of spatial coherence and geochemical validity as established by inverse modeling to support the clusters chosen from HCA.

All groups have relatively low mean total dissolved solids (TDS), ranging from 77 to 257 mg/L (Table 1). Low TDS is common in granitic rock aquifers and reflects a combination of factors including slow reaction rates, short residence times, and limited reactive surface area in fractures. Low TDS in spring data is also related to dilution of base flow with recharge or overland flow, typical of many high altitude streams (Sullivan and Drever 2001). The clusters in the dendrogram are ordered from left to right in a near monotonic increase of TDS. Detailed evaluation of the data (comparing Figures 3 and 4) revealed that samples from groups 1, 2, and 3 are exclusively composed of ground water (except one surface water sample in group 3—A04). Group 4 samples consisted of surface water (except one ground water sample—N99), indicating that surface water and ground water maintain distinct chemical signatures even during spring recharge. However, the linkage distances between groups, especially groups 3 and 4 (Figure 3), suggest that all samples have more hydrochemical similarities during runoff, and the fact that group 4 surface water flow samples have lower TDS than the ground water samples from group 3 shows the dilution effect of spring runoff. This is consistent with the hydrographs that show most recharge is rapidly discharged over one to two months (Poeter et al. 2003).

In gross chemical terms, group 1 samples are characterized by low calcium, magnesium, SO<sub>4</sub>, and HCO<sub>3</sub> concentrations, while group 2 samples are characterized by lower O<sub>2</sub> (DO) and higher amounts of rock-forming components (Table 1). Group 3 samples are distinguished by significantly higher chlorine and NO<sub>3</sub>+NO<sub>2</sub> content. For instance, NO<sub>3</sub>+NO<sub>2</sub> (total as nitrogen) averages 4.07 mg/L in group 3, 7.4 to 19 times higher than the other groups. The group 4 (stream water) samples are characterized by higher

pH, DO, and potassium, and lower NO<sub>3</sub>+NO<sub>2</sub> relative to group 3. The enrichment in chlorine relative to sodium in groups 3 and 4 is evidenced in the sodium:chlorine molar ratios of 0.72 and 0.77, respectively, compared to the ratios of groups 1 and 2 (3.91 and 2.58, respectively).

#### Base Flow Conditions

The base flow data were also classified in 12-dimensional space and are presented in a dendrogram (Figure 3). Four preliminary clusters were chosen for this dataset as well. Identifying the source of the samples from groups 1, 2, and 3 in the dendrogram revealed they are composed primarily of ground water (except two surface water samples in group 1—I01 and K01, and two surface water samples in group 3—A04 and A05). Ground water sample N99 is grouped with surface water samples. TDS of the base flow samples range from 95 to 286 mg/L (Table 1). The clusters in the dendrogram increase in TDS from left to right. The chemistry of the clusters is similar to that of the clusters from the spring dataset, except now the surface water samples have the highest TDS indicating that base flow is derived from ground water discharge as expected. The cluster classifications are changed since spring sampling for some locations, as indicated by the change in the size of the cluster from spring to fall. This behavior can be interpreted as reflecting local variations in aquifer conductivity and are discussed in more detail later.

In the base flow dataset, group 1 samples are characterized by low calcium, magnesium, potassium, SO<sub>4</sub>, HCO<sub>3</sub>, and TDS concentrations (Table 1). Group 2 samples are characterized by higher amounts of rock-forming components sodium, potassium, calcium, magnesium, fluorine, higher HCO<sub>3</sub>, but lower DO, chlorine, and NO<sub>3</sub>+NO<sub>2</sub> concentrations. Group 3 samples have proportionally higher calcium, magnesium, sodium, potassium, and HCO<sub>3</sub> values, but show disproportional increases in chlorine, SO<sub>4</sub>, and NO<sub>3</sub>+NO<sub>2</sub> values, as well as lower pH values. In fact, the group 3 samples have NO<sub>3</sub>+NO<sub>2</sub> (total as nitrogen) and chlorine concentrations an order of magnitude higher than

groups 1 and 2. Group 4 (stream water) has even higher sodium, potassium, chlorine, and TDS values than group 3, but both pH and DO also have higher values consistent with atmospheric contact. Samples from groups 3 and 4 are enriched in chlorine relative to the sodium with sodium:chlorine molar ratios of 0.67 and 0.64, respectively, compared to the samples of groups 1 and 2, 4.71 and 5.53, respectively.

#### Spatial Distribution of Statistical Water Groups

The spatial distribution of the base flow data clusters is shown in Figure 4. The base flow data should most closely reflect steady-state conditions in this aquifer where the water chemistry is most closely linked to water-rock processes rather than dilution and flushing of spring runoff. Mapping the groups shows their spatial coherence. Considering the spatial distribution of samples based on their chemical similarity allows the user to combine the two types of information (hydrochemical and spatial) into one representation and directly evaluate potential hydrochemi-

cal facies in a spatial context. Spatial coherence of the statistical clusters is expected in a system where water chemistry is dominated by a few underlying hydrochemical processes. Further, any clusters defined by Q-mode analysis can reflect the spatial distribution of lithological or hydrochemical conditions that represent the underlying processes.

Ground water samples from low TDS group 1 are associated with the highest elevations in TCB (average well surface elevation of 2608 and 2515 m for runoff and base flow datasets, respectively). Samples from groups 2 and 3 are located at lower elevations than group 1 samples with average well surface elevations of 2369 m. Group 3 samples are intermingled with group 2 samples, but appear most strongly correlated to the higher population density. Of note is the fact that the surface samples in group 4 show little change in chemistry along the length of streams.

The overall pattern in the basin is low TDS in the higher elevations in the southwest and higher TDS at the lower elevations to the northeast. This pattern is consistent

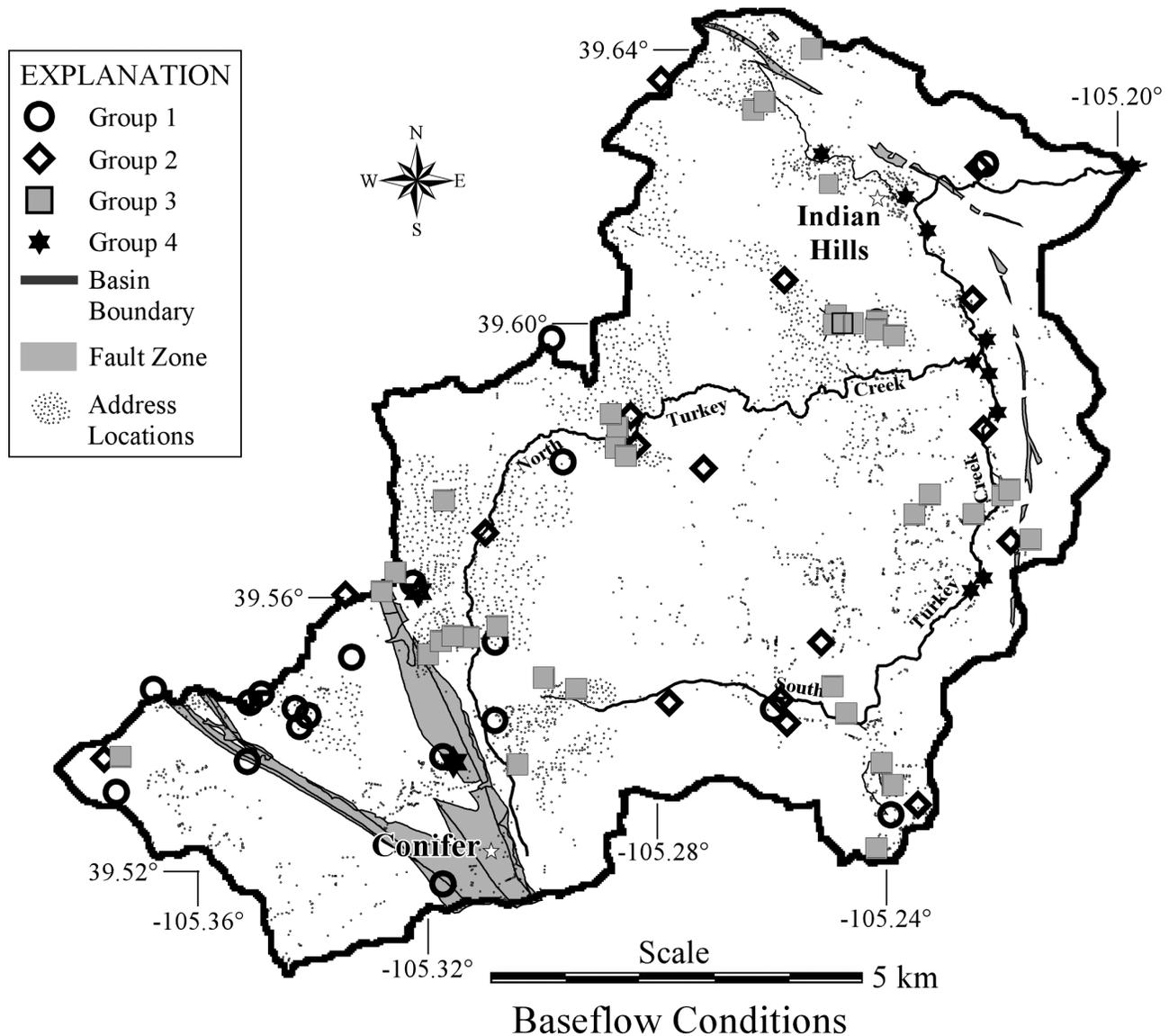


Figure 4. Map view of water groups (determined from HCA results) and population density for the base flow conditions. Estimated current households number 4900 (U.S. Geological Survey 2001).

with the expectation that greater rock-water interactions will increase TDS along topographic flowpaths. Samples that belong to the same group are usually located in close proximity to one another, suggesting similar processes and/or flowpaths. The high degree of spatial and statistical coherence suggests that the statistically derived groups have hydrochemical significance and that the changes between these hydrochemical facies represent the hydrochemical evolution of water in TCB.

### PCA

PCA was used to reduce the number of variables, investigate the degree of continuity or clustering of the samples, and identify the variables most important to separating the groups; in effect, extracting the factors that control the chemical variability. In this analysis, the axes (PCs) may represent the dominant underlying processes and should help constrain any process-based models of hydrochemical evolution. In this situation, we anticipate that the chemistry of surface water and ground water are derived from typical rock-water interaction (weathering) as precipitation reacts with aquifer minerals during flow. In addition, there may be contributions from anthropogenic sources that can produce distinct chemical differences compared to the natural background. This technique can highlight those outliers or groups of samples that are controlled by such factors from the more pervasive natural background. Detailed analysis showed that three PCs are the most meaningful choice for the system. We used both Kaiser's criterion and Scree plots to choose the number of PCs.

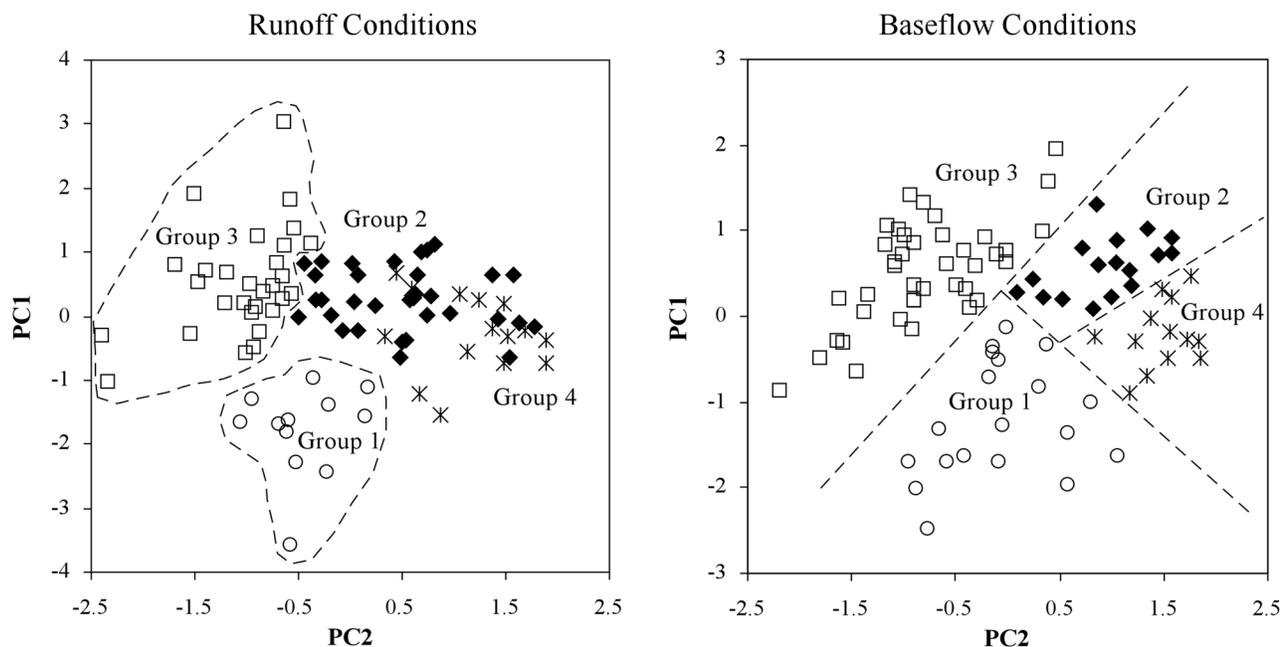
#### Spring Runoff Conditions

For the spring dataset, three significant PCs explain 72.8% of the variance of the original dataset. Most of the variance is contained in PC1 (40.6%), which is associated with the variables SC, calcium, magnesium,  $\text{SO}_4$ , and

$\text{HCO}_3$ . PC2 explains 16.3% of the variance and is mainly related to pH and  $\text{NO}_3+\text{NO}_2$ . The variables DO and chlorine contribute most strongly to the third component (PC3) that explains 15.9% of the total variance. PC1 contains classical hydrochemical variables originating from weathering processes, whereas PC2 and PC3 are related to nitrate and chlorine. High concentrations of nitrate and chlorine are generally attributed to anthropogenic sources. Figure 5 shows the projection of the first two PC scores (PC1 and PC2) in a scatter-plot. The distribution suggests a more continuous variation of the chemical and physical properties of some of the samples. Samples from groups 1 and 3 are more compactly clustered, whereas samples from groups 2 and 4 mingle in PC space. Compact PC distributions suggest that all the water samples in that group have similar chemistries, hence similar flowpaths or sources. If distribution of the samples in the PC space is broad, it may indicate changes in the water chemistry due to processes such as a source of contamination, dilution, or abrupt changes in vertical-horizontal connectivity of the aquifer.

#### Base Flow Conditions

Three PCs explain 71.9% of the variance of the base flow dataset. Most of the variance is contained in PC1 (43.5%), which is associated with the variables SC, calcium, magnesium,  $\text{SO}_4$ , and  $\text{HCO}_3$ . PC2 explains 16% of the variance and is related to pH and  $\text{NO}_3+\text{NO}_2$ . The variables DO and chlorine contribute most strongly to the third component (PC3), which explains 12.4% of the total variance. This pattern is exactly the same as the one seen in the runoff dataset and the information obtained from PCA is fully consistent with that provided by HCA. The scores for the first two principal components are plotted in Figure 5, which shows four well-separated groups defined in PC space that correspond to the results obtained from HCA. It appears that during low flow conditions, the clustering is



**Figure 5.** Plot of the PCA results showing the distribution of HCA-derived classification of samples for the runoff and base flow seasons.

more distinct without the overlap between ground water (group 2) and surface water samples (group 4) seen in the spring data. This is consistent with the hydrographs that show most spring recharge is rapidly discharged over one to two months (Poeter et al. 2003).

### Hydrochemical Evolution and Inverse Geochemical Modeling

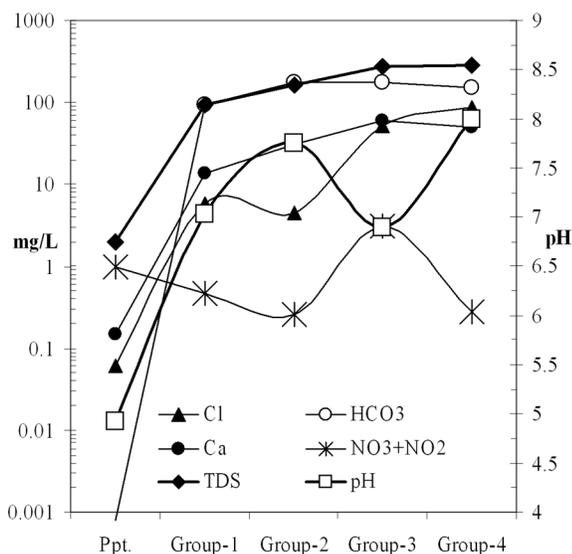
We will use the base flow dataset for further analysis. Ideally, the data for inverse modeling are from a steady-state condition, but, in fact, few natural hydrologic systems ever reach steady state and this procedure has been widely applied with useful results. Determination of the watershed's hydrochemical evolution requires identification of the dominant processes that control water chemistry, which include natural background and anthropogenic influences (e.g., road salt, septic tank discharge). The preliminary hydrochemical evolution is based on the behavior of individual components within the context of PCA. The hydrochemical evolution for selected components from the base flow data is presented in Figure 6. The fundamental premise is that TDS increases as water moves from the recharge to the discharge areas, or precipitation evolving sequentially to group 4. The change in TDS is significant as snow melts and reacts with the more reactive soil layer composed of aquifer material and vegetative debris (precipitation to group 1). The difference in water chemistry between samples from groups 1 and 2 is relatively smaller, indicating the limited water-rock interaction as the precipitation loses contact with the atmosphere (decreasing  $O_2$ ) and the entrained  $CO_2$  is converted to  $HCO_3$  (Table 1). The process of systematic increases in most parameters continues until either an upper limit is reached due to mineral equilibrium (e.g., calcite saturation) or the water exits the basin. This conceptual model is supported by the PCA results that showed the majority of variation in the dataset was related to components associated with dissolution of

major aquifer minerals (sodium, potassium, magnesium, sulfur, and calcium). However, an abrupt and nonsystematic change occurs between groups 2 and 3. The second PCA axis contains two of four parameters—chlorine,  $SO_4$ ,  $NO_3$ , and pH—that change nonsystematically between groups 2 and 3. Finally, the discharge samples, group 4, have distinctly higher chlorine and  $O_2$  parameters that compose the third axis, while calcium,  $SO_4$ ,  $HCO_3$ , and  $NO_3$  decrease.

Thus, the preliminary inverse model should mimic a process such as rock weathering reactions of the granitic and gneissic rocks, which contain potassium feldspar, plagioclase, hornblende, biotite, potassium mica, calcite, pyrite, and quartz to supply the majority of the solutes calcium, magnesium, sodium, potassium, and  $SO_4$  with  $HCO_3$  derived from consumption of atmospheric  $CO_2$ . Some of the sodium, calcium, chlorine,  $CO_3$ , and  $SO_4$  may be supplied by dry deposition of calcite, halite, and gypsum dust (Morgan 2000). The available mineral phases in the PHREEQC database were supplemented with appropriate data from the WATEQ4F and MINTEQ databases. Another process or component is required for the lowered pH and elevated  $NO_3$  found in group 3 samples. Finally, another process, or processes, control the elevated chlorine and  $O_2$  of group 4 chemistry. This conceptual model is now tested with inverse modeling.

Results of inverse modeling are presented in Table 2. Local rain water chemistry was the starting water composition for the first inverse model that calculates the mass of minerals and gases required to produce the mean group 1 composition from precipitation. Potential reactive components are limited to minerals identified in the watershed and the reaction mode (dissolution or precipitation) set based on saturation indices calculated with PHREEQC for each group's mean chemical composition. We have found that using the mean compositions of each group as model inputs facilitates rapid convergence and produces more robust models by minimizing the inherent variability between individual samples. Using this approach, the changes in water chemistry from precipitation to group 1, and group 1 to group 2, are easily modeled as a result of normal rock-water interactions (natural background).

However, no inverse models using the available minerals can account for the observed changes from group 2 to group 3 or group 3 to group 4 samples. The changes in  $NO_3$ , chlorine, and pH between groups 2 and 3 are unusual as water-rock reactions usually increase rather than decrease pH values, and the local minerals cannot provide the large amounts of chlorine and  $NO_3$  required. The most likely explanation is mixing of group 2 samples with an anthropogenic component of lower pH and higher chlorine,  $SO_4$ , and  $NO_3$  content. Bossong et al. (2003) noted that the chloride mass balances required an additional component other than road salt to account for the elevated chloride in the ground water. Hofstra and Hall (1975) proposed that septic tank effluent (STE) could be a significant source of elevated chloride. Figure 7 displays the major ion composition of local STE (Dano et al. 2003) and the more evolved group 2 ground water. The figure shows that STE has a chemical composition that when mixed with unaltered ground water could account for the impacted samples



**Figure 6.** Plot of pH, calcium, chlorine,  $NO_3$ ,  $HCO_3$ , and TDS values for the statistically defined base flow sample groups showing the hydrochemical evolution of the surface water and ground water in TCB.

**Table 2**  
**Results of Inverse Modeling Using the Means of Each Statistical Group as Input**

Plagioclase	Hornblende	Biotite	K-mica	K-spar	Calcite	CO <sub>2</sub> (g)	Gypsum	Halite	Kaol.	SiO <sub>2</sub> (a)
<b>Precipitation to Group 1</b>										
8.18 × 10 <sup>-04</sup>	2.12 × 10 <sup>-05</sup>	1.69 × 10 <sup>-05</sup>	0.00 × 10 <sup>+00</sup>	0.00 × 10 <sup>+00</sup>	0.00 × 10 <sup>+00</sup>	1.48 × 10 <sup>-03</sup>	5.66 × 10 <sup>-05</sup>	1.64 × 10 <sup>-04</sup>	-5.73 × 10 <sup>-04</sup>	-1.03 × 10 <sup>-03</sup>
7.65 × 10 <sup>-04</sup>	3.14 × 10 <sup>-05</sup>	0.00 × 10 <sup>+00</sup>	0.00 × 10 <sup>+00</sup>	1.69 × 10 <sup>-05</sup>	0.00 × 10 <sup>+00</sup>	1.44 × 10 <sup>-03</sup>	5.66 × 10 <sup>-05</sup>	1.64 × 10 <sup>-04</sup>	-5.36 × 10 <sup>-04</sup>	-1.05 × 10 <sup>-03</sup>
7.65 × 10 <sup>-04</sup>	3.14 × 10 <sup>-05</sup>	0.00 × 10 <sup>+00</sup>	1.69 × 10 <sup>-05</sup>	0.00 × 10 <sup>+00</sup>	0.00 × 10 <sup>+00</sup>	1.44 × 10 <sup>-03</sup>	5.66 × 10 <sup>-05</sup>	1.64 × 10 <sup>-04</sup>	-5.53 × 10 <sup>-04</sup>	-1.01 × 10 <sup>-03</sup>
<b>Group 1 to Group 2</b>										
2.51 × 10 <sup>-04</sup>	6.94 × 10 <sup>-05</sup>	0.00 × 10 <sup>+00</sup>	0.00 × 10 <sup>+00</sup>	1.46 × 10 <sup>-05</sup>	0.00 × 10 <sup>+00</sup>	1.11 × 10 <sup>-03</sup>	4.07 × 10 <sup>-05</sup>	0.00 × 10 <sup>+00</sup>	-1.81 × 10 <sup>-04</sup>	-9.23 × 10 <sup>-04</sup>
2.51 × 10 <sup>-04</sup>	6.94 × 10 <sup>-05</sup>	0.00 × 10 <sup>+00</sup>	1.46 × 10 <sup>-05</sup>	0.00 × 10 <sup>+00</sup>	0.00 × 10 <sup>+00</sup>	1.11 × 10 <sup>-03</sup>	4.07 × 10 <sup>-05</sup>	0.00 × 10 <sup>+00</sup>	-1.95 × 10 <sup>-04</sup>	-8.94 × 10 <sup>-04</sup>
2.77 × 10 <sup>-04</sup>	6.06 × 10 <sup>-05</sup>	1.46 × 10 <sup>-05</sup>	0.00 × 10 <sup>+00</sup>	0.00 × 10 <sup>+00</sup>	0.00 × 10 <sup>+00</sup>	1.11 × 10 <sup>-03</sup>	4.86 × 10 <sup>-05</sup>	0.00 × 10 <sup>+00</sup>	-1.98 × 10 <sup>-04</sup>	-8.84 × 10 <sup>-04</sup>

Three possible models were obtained for precipitation to group 1 and group 1 to 2. Values are moles added or subtracted to obtain mass balance. Positive values are dissolved and negative are precipitated.

(group 3 pattern is intermediate between end members). Mixing calculations using most conservative parameters, sodium and chlorine, suggest STE fractions of 8% and 12%, respectively, could be added to group 2 water and produce group 3 samples.

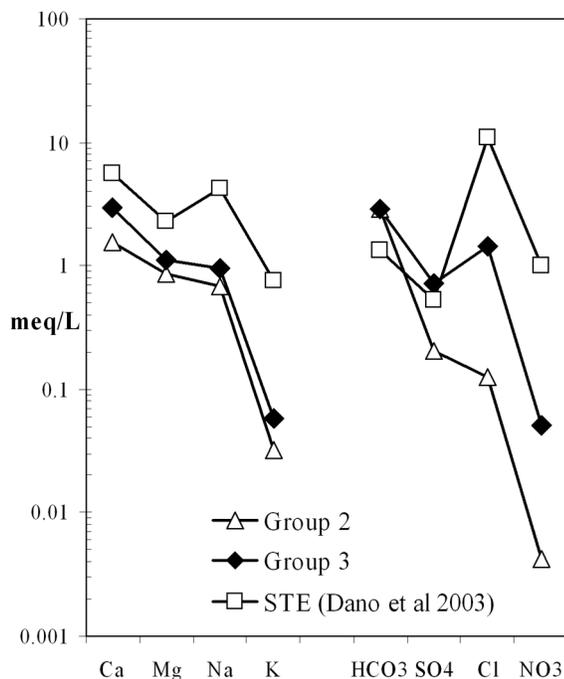
Group 4 samples are primarily surface water from the north and south forks of Turkey Creek. During base flow, creek water is assumed to be a result of aquifer discharge. The samples have much higher O<sub>2</sub> content consistent with atmospheric contact. The chemistry of group 4 samples appears to be a mixture ground water discharging into the stream, itself a mixture of waters from groups 2 and 3, but with an additional NaCl-rich component and processes to remove SO<sub>4</sub>, HCO<sub>3</sub>, and NO<sub>3</sub>. These results indicate the hydrochemical evolution model for group 4 is not very well constrained.

## Discussion

The cluster interpretations in this study are general and made on the watershed scale. There may be better interpretations within the context of a more localized system. However, in the context of a watershed scale, we can use the hydrochemical evolution model to explore questions about the watershed. For instance, the source of the anthropogenic impact is an important concern. The anthropogenic component was clearly distinguished in group 3 samples and we can show that the STE composition reported by Dano et al. (2003), mixed with group 2 water, can explain the impacted group 3 samples. The identification of this component was facilitated by the multivariate analysis since no single parameter had a bimodal distribution, meaning univariate analysis would have been problematic. The high degree of correspondence between population density and group 3 samples supports this interpretation, as does the fact that ISDS effluents are a primary source of nitrate in ground water and surface water in the area, where other sources for nitrate such as large-scale fertilizer applications and animal operations are lacking (Bossong et al. 2003). Therefore, the relative proportions of ISDS effluent and fresh recharge water will influence the ground water chemistry, and we formulate the working hypothesis that wells with an anthropogenic component are locations where STE volumes are high relative to normal recharge and/or vertical conductivity is good.

Based on limited STE composition data, the chemical mixing calculations suggest the wells with group 3 water chemistry have a fraction of STE between 8% and 12%. If the hydrologic estimates of effluent volumes vs. total recharge are accurate, then TCB residents generate a volume of STE up to 75% of present recharge. Given the apparent limited anthropogenic impact on ground water, it appears we have only begun to see the impact of ISDS on ground water quality or much of STE is discharging directly to surface runoff.

Another question in TCB is whether the system behaves as an equivalent porous medium (EPM). While there is clear evidence that the aquifer is not an EPM on a well scale (VanderBeek 2003), the hydrochemical data show the EPM assumption appears to be valid for the



**Figure 7. Scholler plot showing the mean chemical compositions of groups 2 and 3, and the composition of local STE.**

watershed scale. In an EPM watershed, we expect weathering reactions to produce systematic changes in water chemistry along the flowpath. In contrast, spatially abrupt changes would indicate behavior that cannot be represented with EPM assumptions. The PHREEQC modeling demonstrated that groups 1 and 2 represent natural background and those samples generally show a systematic spatial distribution correlated with topography with only a few abrupt changes in water chemistry along the topographic flowpaths. The higher TDS group 3 samples are excluded from this analysis since the higher TDS of this group is not solely a result of weathering reactions. This overall pattern suggests that TCB can be represented as an equivalent porous media at the watershed scale, but not on the smaller scale as evidenced by the adjacent locations with water chemistry from different clusters.

However, the degree of heterogeneity of hydraulic conductivity is not well known; it would be an important parameter in any quantitative flow model. Hydrographs from wells in TCB show that some wells respond to spring recharge in days to weeks, while others do not show water

level increases for several months (Bossong et al. 2003). Assuming that local hydraulic conductivity will be reflected in the connection between the ground water and surface water, we can use the response of an individual well's water chemistry to spring recharge to evaluate spatial trends. In TCB, the assumption is that rapid, vertical recharge during snowmelt and precipitation events supply the base flow. Figure 8 shows the locations where surface and ground water quality did or did not change between spring runoff and fall base flow conditions. There are several locations where spring samples showed normal background chemistry (groups 1 and 2), but, by fall, the water quality had degraded due to anthropogenic impact (group 3). At locations where ground water quality during base flow conditions is not as good as during spring recharge, it can be inferred that there is good hydraulic connection between the surface and aquifer. Locations with no significant change in water quality between sample periods have poor vertical hydraulic connection. Based on the distribution of such locations, we can see that the scale of heterogeneity is small since adjacent wells such as F8 and F14 have one well that

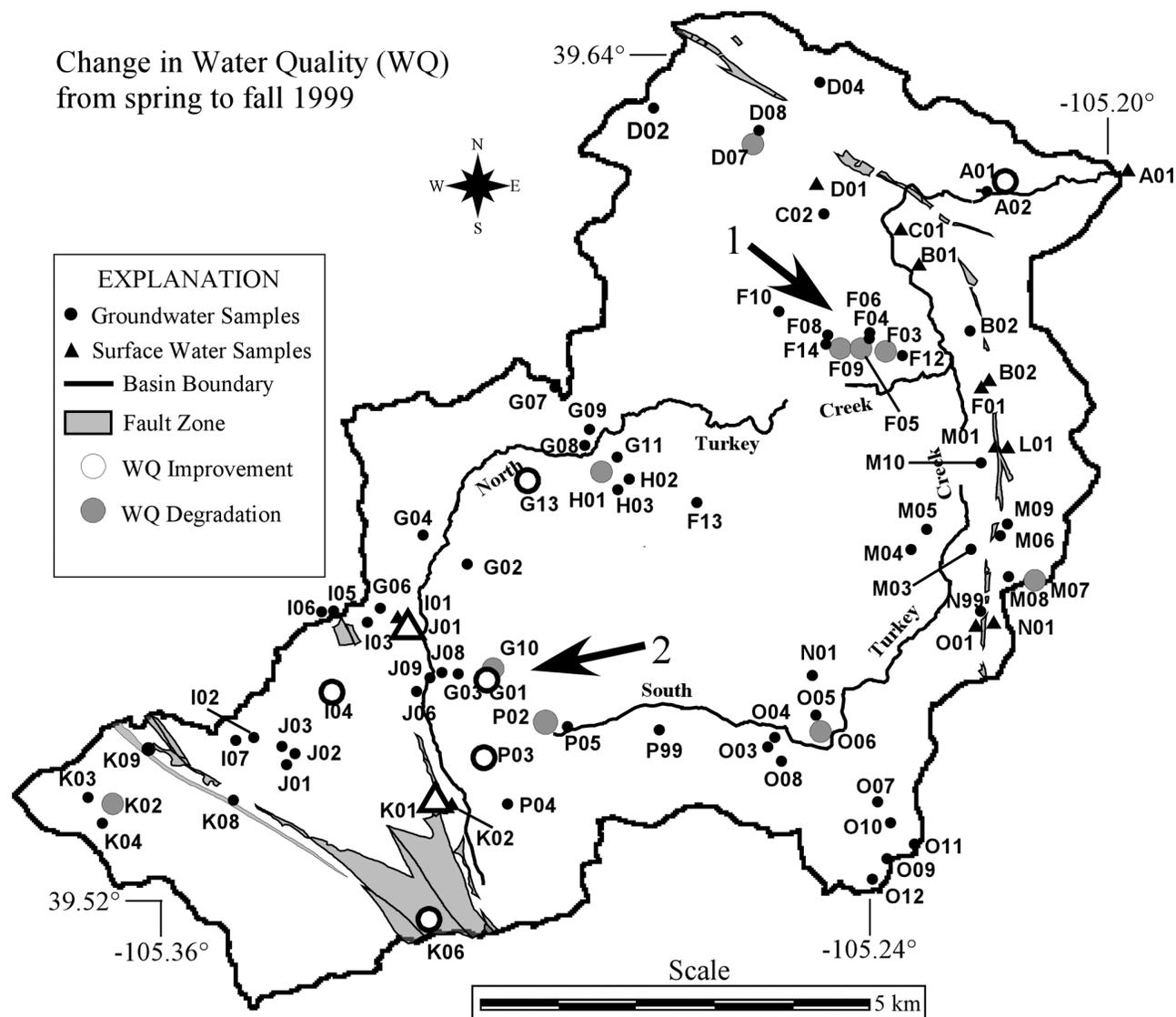


Figure 8. Map view of Turkey Creek showing the change in water quality between runoff and base flow. Arrows 1 and 2 show the locations of specific samples referred to in the text.

shows the spring dilution effect, while the adjacent well does not. Even more extreme variation is seen at the well pair G10 and G01.

Since the samples were collected over a period of a month, it is possible that the lack of dilution response may be due to the timing of sample collection where the recharge was pushing older, more saline ground water toward the well and later samples would have shown the response (Nagorski et al. 2003). Correlation of the chemical behavior with the well hydrographs response time (from days to months) to recharge events would allow confirmation of this hypothesis.

## Conclusions

The integrated statistical/spatial/geochemical analysis showed that some locations (groups 1 and 2) have water chemistry due to natural water-rock interactions, while other locations (group 3) were impacted by an anthropogenic source or sources. In this case, the source of degradation of water quality is strongly associated with increasing populations that employ ISDS. Mountain watersheds, which serve as the principal source areas for the majority of recharge in the western United States, are particularly vulnerable because low aquifer storage and rapid flow rates produce limited opportunity for natural processes that attenuate anthropogenic pollution. The water chemistry also showed that the surface flow is derived from ground water, while spring recharge simply dilutes the base flow chemistry. Further comparison of the temporal changes in water chemistry in wells helped infer local aquifer properties in a complex fractured rock aquifer.

It appears that the methodology illustrated in this paper allows investigators to incorporate typical hydrochemical information into the analysis of a watershed-scale aquifer system. The procedure is aimed at the broader aspects of the hydrochemical evolution and may not be as effective on small spatial scales or in situations with little chemical variability. However, the methodology can be useful in many typical watersheds where basic hydrochemical and hydrogeological data are available. The methodology takes advantage of strengths of standard statistical, geochemical, and spatial analysis tools, and allows sequential integration of the information from each technique producing a more robust interpretation. In this example, we were able to identify the major processes controlling hydrochemical variations in TCB (natural water/rock interaction and an anthropogenic component), determine the location and chemical signature of anthropogenic impact, and provide information about the aquifer properties (scale of EPM behavior and heterogeneity of the hydraulic conductivity). The results support the previous hydrological interpretations of the basin as a single watershed with limited storage, and suggest that housing density can be a problem due to the widespread use of ISDS.

## Acknowledgments

We would like to thank the Ministry of Education of Turkey and the National Institute of Water Resources for their financial support. Discussions with the Turkey Creek

Group at Colorado School of Mines were very helpful in developing our concepts and keeping us grounded in hydrologic reality. The final manuscript was greatly improved based on suggestions by Edward Mehnert and two anonymous reviewers.

**Editor's Note:** The use of brand names in peer-reviewed papers is for identification purposes only and does not constitute endorsement by the authors, their employers, or the National Ground Water Association.

## References

- Adams, S., R. Titus, K. Pietersen, G. Tredoux, and C. Harris. 2001. Hydrochemical characteristics of aquifers near Sutherland in the western Karoo, South Africa. *Journal of Hydrology* 241, no. 1–2: 91–103.
- Alberto, W.D., D.M. del Pilar, A.M. Valeria, P.S. Fabiana, H.A. Cecilia, and B.M. de los Angeles. 2001. Pattern recognition techniques for the evaluation of spatial and temporal variations in water quality, a case study: Suquia River Basin (Cordoba-Argentina). *Water Resources* 35, 2881–2894.
- Alther, G.A. 1979. A simplified statistical sequence applied to routine water quality analysis: A case history. *Ground Water* 17, 556–561.
- Back, W. 1966. Hydrochemical facies and ground water flow patterns in the northern part of the Atlantic Coastal Plain. U.S. Geological Survey Professional Paper 498-A.
- Birke, M., and U. Rauch. 1993. Environmental aspects of the regional geochemical survey in the southern part of East Germany. *Journal of Geochemical Exploration* 49, no. 1–2: 35–61.
- Bosson, C.R., J.S. Caine, D.I. Stannard, J.L. Flynn, M.R. Stevens, and J.S. Heiny-Dash. 2003. Hydrologic conditions and assessment of water resources in the Turkey Creek Watershed, Jefferson County, Colorado, 1998–2000. U.S. Geological Survey Water Resources Investigations Report 03–4034.
- Briz-Kishore, B.H., and G. Murali. 1992. Factor analysis of revealing hydrochemical characteristics of a watershed. *Environmental Geology and Water Sciences* 19, no. 1: 3–9.
- Busby, J.F., L.N. Plummer, R.W. Lee, and B.B. Hanshaw. 1991. Geochemical evolution of water in the Madison Aquifer in parts of Montana, South Dakota, and Wyoming. U.S. Geological Survey Open File Report: F1-F89.
- Dalton, M.G., and S.B. Upchurch. 1978. Interpretation of hydrochemical facies by factor analysis. *Ground Water* 16, no. 4: 228–233.
- Dano, K., E. Poeter, and G. Thyne. 2003. Geochemical and geophysical determination of the fate of septic tank effluent in Turkey Creek Basin, Colorado. In *Abstracts with Programs of the Annual Meeting of the Geological Society of America*, November 2–5, Seattle, Washington, 34, no. 7: 564–565.
- Environmental Systems Research Institute. 1996. *Arc/View GIS Manual*. Redlands, California: ESRI.
- Farnham, I.M., K.J. Stetzenbach, A.K. Singh, and K.H. Johannesson. 2002. Treatment of nondetects in multivariate analysis of ground water geochemistry data. *Chemometrics and Intelligent Laboratory Systems* 60, 265–281.
- Folger, P.F., E. Poeter, R.B. Wanty, D. Frishman, and W. Day. 1996. Controls on <sup>222</sup>Rn variations in a fractured crystalline rock aquifer evaluated using aquifer tests and geophysical logging. *Ground Water* 34, no. 2: 250–261.
- Güler, C., and G.D. Thyne. 2004. Hydrologic and geologic factors controlling surface and ground water chemistry in Indian Wells–Owens Valley area and surrounding ranges, California, USA. *Journal of Hydrology* 285, 177–198.
- Güler, C., G. Thyne, J.E. McCray, and A.K. Turner. 2002. Evaluation of graphical and multivariate statistical methods for classification of water chemistry data. *Hydrogeology Journal* 10, 455–474.

- Helena, B., R. Pardo, M. Vega, E. Barrado, J.M. Fernandez, and L. Fernandez. 2000. Temporal evolution of ground water composition in an alluvial aquifer (Pisuerga River, Spain) by principal component analysis. *Water Research* 34, 807–816.
- Hernandez, M.A., N. Gonzalez, and M. Levin. 1991. Multivariate analysis of a coastal phreatic aquifer using hydrochemical and isotopic indicators, Buenos Aires, Argentina. In *Proceedings of the International Association on Water Pollution Research and Control's International Seminar on Pollution, Protection and Control of Ground Water*, as published in *Water Science & Technology* 24, 139–146.
- Hidalgo, M.C., and J. Cruz-Sanjulian. 2001. Ground water composition, hydrochemical evolution and mass transfer in a regional detrital aquifer (Baza Basin, southern Spain). *Applied Geochemistry* 16, 745–758.
- Hofstra, W.E., and D.C. Hall. 1975. Geologic control of supply and quality of water in the mountainous part of Jefferson County, Colorado. Colorado Geological Survey Bulletin 36.
- Join, J.L., J. Coudray, and K. Longworth. 1997. Using principal components analysis and Na/Cl ratios to trace ground water circulation in a volcanic island: The example of Reunion. *Journal of Hydrology* 190, no. 1–2: 1–18.
- Judd, A.G. 1980. The use of cluster analysis in the derivation of geotechnical classifications. *Bulletin of the Association of Engineering Geology* 17, 193–211.
- Kaiser, H.F. 1960. The application of electronic computers to factor analysis. *Educational and Psychological Measurement* 20, 141–151.
- Kuells, C., E.M. Adar, and P. Udluft. 2000. Resolving patterns of ground water flow by inverse hydrochemical modeling in a semiarid Kalahari basin. *Tracers and Modeling in Hydrogeology* 262, 447–451.
- Liedholz, T., and M.T. Schafmeister. 1998. Mapping of hydrochemical ground water regimes by means of multivariate-statistical analyses. In *Proceedings of the Fourth Annual Conference of the International Association for Mathematical Geology*, October 5–9, Ischia, Italy, ed. A. Buccianti, G. Nardi, and R. Potenza, 298–303. Kingston, Ontario, Canada: International Association for Mathematical Geology.
- Locsey, K.L., and M.E. Cox. 2003. Statistical and hydrochemical methods to compare basalt- and basement rock-hosted ground waters: Atherton Tablelands, north-eastern Australia. *Environmental Geology* 43, no. 6: 698–713.
- Lopez-Chicano, M., M. Bouamama, A. Vallejos, and B.A. Pulido. 2001. Factors which determine the hydrogeochemical behaviour of karstic springs: A case study from the Betic Cordilleras, Spain. *Applied Geochemistry* 16, no. 9–10: 1179–1192.
- Meng, S.X., and J.B. Maynard. 2001. Use of statistical analysis to formulate conceptual models of geochemical behavior: Water chemical data from the Botucata Aquifer in Sao Paulo state, Brazil. *Journal of Hydrology* 250, 78–97.
- Miesch, A.T. 1976. Geochemical survey of Missouri—Methods of sampling, laboratory analysis and statistical reduction of data. U.S. Geological Society Professional Paper 954-A.
- Morgan, K. 2000. Spatial analysis and modeling of geochemical distribution to assess fracture flow in Turkey Creek Basin, Jefferson County, Colorado. M.S. thesis, Department of Geology and Geological Engineering, Colorado School of Mines.
- Nagorski, S.A., J.N. Moore, and T.E. McKinnon. 2003. Geochemical response to variable streamflow conditions in contaminated and uncontaminated streams. *Water Resources Research* 39, 1044–1058.
- Ochsenkuehn, K.M., J. Kontoyannakos, and P.M. Ochsenkuehn. 1997. A new approach to a hydrochemical study of ground water flow. *Journal of Hydrology* 194, no. 1–4: 64–75.
- Pereira, H.G., S. Renca, and J. Sataiva. 2003. A case study on geochemical anomaly identification through principal components analysis supplementary projection. *Applied Geochemistry* 18, 37–44.
- Plummer, L.N., and W.W. Back. 1980. The mass balance approach—Application to interpreting the chemical evolution of hydrological systems. *American Journal of Science* 280, 130–142.
- Poeter, E., G. Thyne, G. VanderBeek, and C. Guler. 2003. Ground water in Turkey Creek Basin of the Rocky Mountain Front Range in Colorado. In *Engineering Geology in Colorado—Contributions, Trends, and Case Histories*. Denver, Colorado: Association of Engineering Geologists.
- Qian, G., G. Gabor, and R.P. Gupta. 1994. Principal components selection by the criterion of the minimum mean difference of complexity. *Journal of Multivariate Analysis* 49, 55–75.
- Razack, M., and J. Dazy. 1990. Hydrochemical characterization of ground water mixing in sedimentary and metamorphic reservoirs with combined use of Piper's principle and factor analysis. *Journal of Hydrology* 114, no. 3–4: 371–393.
- Reeve, A.S., D.I. Siegel, and P.H. Glaser. 1996. Geochemical controls on peatland pore water from the Hudson Bay Lowland: A multivariate statistical approach. *Journal of Hydrology* 181, no. 1–4: 285–304.
- Seyhan, E., A.A. van-de-Griend, and G.B. Engelen. 1985. Multivariate analysis and interpretation of the hydrochemistry of a dolomitic reef aquifer, northern Italy. *Water Resources Research* 21, no. 7: 1010–1024.
- Sims, P.K. 1989. Central City and Idaho Springs, Front Range, Colorado. In *Mineral Deposits and Geology of Central Colorado*, ed. B. Bryant and D.W. Beatty. 28th International Geological Congress Field Trip Guidebook T129. Washington, D.C.: American Geophysical Union.
- Sims, P.K., and D.J. Gable. 1967. Petrology and structure of Precambrian rocks, Central City quadrangle, Colorado. U.S. Geological Society Professional Paper 554-E.
- StatSoft Inc. 1997. *Electronic Statistics Textbook*. Tulsa, Oklahoma: StatSoft Inc., <http://www.statsoft.com/textbook/stathome.html>.
- Suk, H., and K.-K. Lee. 1999. Characterization of a ground water hydrochemical system through multivariate analysis: Clustering into ground water zones. *Ground Water* 37, no. 3: 358–366.
- Sullivan, A.B., and J.I. Drever. 2001. Spatiotemporal variability in stream chemistry in a high-elevation catchment affected by mine drainage. *Journal of Hydrology* 252, 237–250.
- Swan, A.R.H., and M. Sandilands. 1995. *Introduction to Geological Data Analysis*. Malden, Massachusetts: Blackwell Science Ltd.
- Taylor, R.B. 1976. Geologic map of the Black Hawk quadrangle, Gilpin, Jefferson, and Clear Creek counties, Colorado. U.S. Geological Society Quadrangle Map GQ-1248, scale: 1:24,000.
- Trimble, D.E., and M.N. Machette. 1979. Geologic map of the Colorado Springs–Castle Rock area, Front Range urban corridor, Colorado. U.S. Geological Survey Miscellaneous Investigations Series.
- Turk, G. 1979. Discussion of “Interpretation of hydrochemical facies by factor analysis.” *Ground Water* 17, no. 2: 212–213.
- Upchurch, S.B. 1979. Reply to discussion of “Interpretation of hydrochemical facies by factor analysis.” *Ground Water* 17, no. 2: 213–215.
- U.S. Geological Survey. 2001. Mountain ground water resource study phase I report summary: Water resources assessment of the Turkey Creek Watershed, 1998 to 2000. Prepared for Jefferson County Planning and Zoning Department, by the U.S. Geological Survey Colorado District.
- Usunoff, E.J., and A. Guzman-Guzman. 1989. Multivariate analysis in hydrochemistry: An example of the use of factor and correspondence analysis. *Ground Water* 27, 27–34.
- VanderBeek, G. 2003. Estimating recharge and storage coefficient in a fractured rock aquifer, Turkey Creek Basin, Jefferson County, Colorado. M.S. thesis, Department of Geology and Geological Engineering, Colorado School of Mines.
- Wang, Y., T. Ma, and Z. Luo. 2001. Geostatistical and geochemical analysis of surface water leakage into ground water on a regional scale: A case study in the Liulin karst system, north-western China. *Journal of Hydrology* 246, no. 1–4: 223–234.