Streamflow generation from snowmelt in semi-arid, seasonally snow-covered, forested catchments, Valles Caldera, New Mexico

Fengjing Liu, Roger C. Bales, Martha H. Conklin, and Mark E. Conrad

Received 3 December 2007; revised 7 September 2008; accepted 22 October 2008; published 30 December 2008.

[1] Streamflow generation in the semi-arid, seasonally snow-covered, and forested mountain catchments of the Valles Caldera, New Mexico, was investigated using chemical tracers. Samples were collected from snow, subsurface flow from hillslopes, and streamflow at Redondo and La Jara Creeks from December 2004 to July 2005. A new modeling procedure was developed by combining diagnostic tools of mixing models and end-member mixing analysis to evaluate the assumptions of mixing models. This procedure was successfully used to determine conservative chemical tracers, identify the number of end-members that contribute to streamflow, and evaluate eligibility of end-members. The results show that streamflow at Redondo Creek was generated from two end-members: lateral subsurface flow (~80%) and thermal meteoric water (~20%). Streamflow at La Jara Creek was primarily from lateral subsurface flow alone. Overland flow of snowmelt was not a significant contributor to streamflow in either catchment. Lateral subsurface flow is an important process of streamflow generation in semiarid environments in the southwest United States and should play a critical role in regulating biogeochemical cycles.


1. Introduction

[2] Information on source waters and pathways of streamflow is critical in understanding hydrologic and biogeochemical processes and contaminant transport within a catchment [Cirmo and McDonnell, 1997; Holko and Lepisto, 1997; Perakis, 2002]. Mixing models using isotopic and chemical tracers (hereinafter mixing model when used in general context) have been widely and successfully applied to identify streamflow sources and catchment flow paths [e.g., Sklash et al., 1976; Hooper and Shoemaker, 1986; McHale et al., 2002; Uhlenbrook and Hoeg, 2003]. However, most of these studies have been focused on forested, humid and rainfall-dominated catchments. New studies in catchments with different climatic and human disturbance regimes may provide hydrologists with new insights into hydrologic processes [Burns, 2002]. Mechanisms of runoff generation in semiarid, mountain forested catchments where annual runoff is dominated by snowmelt, a multi-week event, may be fundamentally different from that in humid regions where frequent short-term rainfall events dominate [Wilcox et al., 1997]. In humid, forested catchments such as found in the eastern United States, runoff is generated primarily from old waters through macropore flow [e.g., McDonnell, 1990; Hooper and Shoemaker, 1986]. A few studies from an 870-m² ponderosa pine hillside at Los Alamos have indicated that both lateral subsurface flow and overland flow are important flow processes that control snowmelt runoff at hillslope scales in semiarid environments [Wilcox et al., 1997; Newman et al., 1998, 2004]. However, the importance of these processes at catchment scales in semiarid regions with a seasonal snow cover has received little attention [McNamara et al., 2005].

[3] Mixing models have also been recently challenged because the results are sometime not physically meaningful [Burns, 2002]. Mixing model results may be very different if different pairs of tracers are used for three-component hydrograph separations [Rice and Hornberger, 1998]. More importantly, chemical values in streamflow may gradually evolve through very long residence time in the subsurface, rather than a result of mixing of several end-members contributed from different geologic and hydrogeologic units with distinct isotopic and chemical signatures. Recent continuous long-term observations of streamflow chemistry in rainwater-dominated catchments in Plynlimon, Wales show fractal behavior across time scales of decades [Kirchner et al., 2000]. Solutes traveling through watersheds may initially be flushed out rapidly but leave behind a persistent low-level tail. Watershed hydrologists must therefore address an outstanding question: Are mixing models still an effective tool in modeling streamflow chemistry and contaminant transport?

[4] In addition, understanding of hydrologic and biogeochemical processes is sometime hindered by a lack of sufficient data in mountain catchments, such as streamflow discharge in ungauged catchments and chemical and isotopic data due to infrequent sampling [Bales et al., 2006].
Access into mountains in the West during snow accumulation in winter and early spring for stream sampling is challenging and sometimes dangerous. Without streamflow measurements and with only limited chemical data, can mixing models still help improve our understanding of source waters and flow paths? Answers to this question certainly help watershed hydrologists to map their research plans and strategies.

The primary aims of the research reported here were to understand how snowmelt becomes runoff in semiarid, seasonally snow-covered, forested catchments in the Valles Caldera, New Mexico using chemical tracers and to answer the two questions posed above through developing a new modeling procedure by combining diagnostic tools of mixing models developed by Hooper [2003] and end-member mixing analysis (a principal component analysis based technique for conducting mixing models as opposed to mixing models in general context) [Christophersen and Hooper, 1992].

2. Research Site

Valles Caldera (106°33’23”W, 35°52’19”N) is located about 15 km northeast of Jemez Springs and 40 km west of Los Alamos in the northern volcanic Jemez Mountains of northern New Mexico. Elevations of Valles Caldera range from 2300 m at Redondo Meadow to 3432 m at Redondo Peak (Figure 1). Since the early 1980s, mean annual temperature has been 9°C and precipitation averaged 476 mm, as measured in a neighboring site at an elevation...
of about 2200 m in Los Alamos [Bowen, 1996]. Mean precipitation from October to April accounts for 40% of the annual total and falls primarily as snow. Snow depth was 60–120 cm in March, 2005 in the saddle just below Redondo Peak (N. Molotch, personal communication, 2005). Mean precipitation during the monsoon season in July and August represents approximately 50% of annual precipitation [Bowen, 1996]. Streams originating from Redondo Peak include Redondo Creek, Jalamillo Creek and La Jara Creek, from west to east clockwise (Figure 1). Redondo and La Jara Creeks drain an area of 13.4 and 3.7 km², respectively.

Streams in the caldera region of the caldera, where it discharges from ring fractures and from the contacts of rhyolite with underlying volcanic and Paleozoic units (Figure 1). Deep geothermal fluids are localized along faults and fractures of the keystone graben in silicified ignimbrites [Dondanville, 1978]. The mean depth of production for these waters is about 1500 m at temperatures of 260–300°C. The deep fluid is chemically distinct from thermal meteoric water, having significant concentrations of Na⁺ and Cl⁻ (Table 1).

3. Methods

3.1. End-Member Mixing Analysis

[9] End-member mixing analysis (EMMA) was used to determine end-members and proportions of end-members contributing to streamflow, following the procedures by Christophersen and Hooper [1992] and Christophersen et al. [1990]. A principal component analysis (PCA) was performed to extract eigenvectors using a correlation matrix of conservative tracers determined using diagnostic tools of mixing models described in next section. Ionic concentrations in streamflow data were orthogonally projected using the eigenvectors by

\[ U = X^p V^T \]

where \( U \) is the orthogonally projected data matrix \((n \times m)\), in which \( n \) represents the number of samples and \( m \) one less than the number of end-members. \( X^p \) has a dimension of \( n \times p \), where \( p \) is the number of conservative tracers used to extract the eigenvectors \( V \). \( V \) has a dimension of

Table 1. Mean (Standard Deviation) Isotopic and Chemical Compositions in Stream Water, Snow, Subsurface Flow, Thermal Meteoric Water, and Deep Geothermal Fluid

<table>
<thead>
<tr>
<th>Sampling Location (Number of Samples)</th>
<th>pH (µS cm⁻¹)</th>
<th>Cond (µeq L⁻¹)</th>
<th>ANC (µeq L⁻¹)</th>
<th>Ca²⁺ (µeq L⁻¹)</th>
<th>Mg²⁺ (µeq L⁻¹)</th>
<th>Na⁺ (µeq L⁻¹)</th>
<th>K⁺ (µeq L⁻¹)</th>
<th>Cl⁻ (µeq L⁻¹)</th>
<th>NO₃⁻ (µeq L⁻¹)</th>
<th>SO₄²⁻ (µeq L⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Redondo Creek (11)</td>
<td>7.34</td>
<td>74</td>
<td>602</td>
<td>423</td>
<td>54</td>
<td>196</td>
<td>41</td>
<td>64</td>
<td>0.63</td>
<td>168</td>
</tr>
<tr>
<td></td>
<td>(0.15)</td>
<td>(15)</td>
<td>(74)</td>
<td>(75)</td>
<td>(15)</td>
<td>(38)</td>
<td>(16)</td>
<td>(36)</td>
<td>(0.04)</td>
<td>(92)</td>
</tr>
<tr>
<td>La Jara Creek (12)</td>
<td>7.14</td>
<td>42</td>
<td>385</td>
<td>266</td>
<td>36</td>
<td>97</td>
<td>21</td>
<td>18</td>
<td>0.73</td>
<td>82</td>
</tr>
<tr>
<td></td>
<td>(0.18)</td>
<td>(2)</td>
<td>(56)</td>
<td>(16)</td>
<td>(2)</td>
<td>(9)</td>
<td>(2)</td>
<td>(4)</td>
<td>(0.1)</td>
<td>(13)</td>
</tr>
<tr>
<td>Snow (1)</td>
<td>5.53</td>
<td>2.3</td>
<td>6.8</td>
<td>0.4</td>
<td>0.4</td>
<td>150</td>
<td>14</td>
<td>14</td>
<td>0.58</td>
<td>75</td>
</tr>
<tr>
<td>Subsurface flow (5)</td>
<td>7.30</td>
<td>60</td>
<td>605</td>
<td>413</td>
<td>49</td>
<td>150</td>
<td>14</td>
<td>14</td>
<td>0.58</td>
<td>75</td>
</tr>
<tr>
<td>Thermal meteoric water at Spencea(4)</td>
<td>7.85</td>
<td>268</td>
<td>2207</td>
<td>304</td>
<td>136</td>
<td>2175</td>
<td>29</td>
<td>212</td>
<td>0.81</td>
<td>342</td>
</tr>
<tr>
<td>at San Antonioö (4)</td>
<td>8.12</td>
<td>199</td>
<td>857</td>
<td>147</td>
<td>33</td>
<td>957</td>
<td>40</td>
<td>92</td>
<td>0.81</td>
<td>159</td>
</tr>
<tr>
<td>Sulfur springö (2)</td>
<td>1.69</td>
<td>18,765</td>
<td>2320</td>
<td>926</td>
<td>361</td>
<td>833</td>
<td>656</td>
<td>11,8704</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deep geothermal fluid at Baca 13ö(11)</td>
<td>7.21</td>
<td>2294</td>
<td>265</td>
<td>17</td>
<td>50,158</td>
<td>6014</td>
<td>54,512</td>
<td>895</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


Data of Rzonca and Schulze-Makuch [2003].

Data of Goff and Gardner [1994] and White [1986].
The orthogonal (Euclidean or U-Space) projections of streamflow samples and end-members were used to solve for proportions of end-members contributions to streamflow. This procedure is mathematically similar to a traditional mixing model, for instance, one tracer for two components as of Sklash et al. [1976], but uses the first U-space projection \( U_1 \) in a two-end-member solution.

The eligibility of end-members was evaluated using distance between original chemical compositions and U-space orthogonal projections of end-members [Christophersen and Hooper, 1992]:

\[
d_j = \| b_j - b_j^* \| \quad (2)
\]

where \( d_j \) means the Euclidean distance of end-member \( b_j \) for tracer \( j \) between original composition \( (b_j) \) and U-space projection \( (b_j^*) \) calculated by equation (3) using the eigenvector \( V \) extracted from conservative tracers. The distance was expressed as percentage in this study by dividing distance by the original chemical composition. The shorter the distance the better fit of an end-member to EMMA.

### 3.2. Diagnostic Tools of Mixing Models

Diagnostic tools of mixing models were used to determine the number of end-members and conservative tracers for EMMA using streamflow chemistry without using information of end-members following Hooper [2003]. These tools were developed to examine the appropriate rank of streamflow chemistry data set through a principal component analysis and determine the dimensions that the data span in a Euclidean U-space. Conservative mixing of end-members with distinct chemical signatures is a linear process, but chemical equilibrium of solutes with different charges is higher-order polynomials [Hooper, 2003]. This analysis is thus used to assess the “fit” of streamflow chemistry to a lower dimensional mixing subspace (Euclidean U-space). To initiate this analysis, streamflow chemistry was projected using eigenvectors extracted from streamflow chemical data:

\[
\hat{X}^* = X^* V_1^T (V_1 V_1^T)^{-1} V_1 \quad (4)
\]

where \( \hat{X}^* \) is the standardized streamflow data (mean = 0, standard deviation = 1), \( \hat{X}^* \) is the projection of the standardized data using eigenvectors of \( V_1 \), \( V_1 \) was extracted using a correlation matrix of streamflow chemical data, including all solutes to be tested (which is thus different from \( V \) in equation (1)). Superscripts \( T \) and \( -1 \) represent “transpose” and “inverse” matrix operations.

A well-posed model is indicated by a random distribution of the residuals between the measured and projected concentrations for each solute; any structure in the distribution suggests a lack of fit in the model, which can arise from the violation of any of assumptions inherent in the mixing model [Hooper, 2003].

A useful scalar measure of fit is relative root-mean-square error (RRMSE). The RRMSE for solute \( j \) is

\[
r_j = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_{ij} - \hat{x}_{ij})^2} \quad (5)
\]

where \( i \) is the \( i \)-th sample, \( x \) is the measured solute concentration, \( \hat{x} \) is the measured solute concentration, and \( \bar{x} \) is the mean of the measured solute concentration from all samples. The RRMSE provides an indication of the “thickness” of the data cloud outside the lower dimensional subspace. The dimension of the mixing subspace \( m \) should be simply the rank of \( X \) (that is, the number of dimensions that \( X \) spans), which is one less than the number of end-members [Hooper, 2003].

### 3.3. Sample Collection and Analysis

Collection of streamflow samples started on 20 December 2004 at Redondo and La Jara Creeks (Figure 1). Sampling took place approximately every two weeks during April and May 2005 and was more frequent from late May to June, with some samples less than a week apart. The last samples were collected on 19–20 July 2005 before the monsoon season. Samples were filtered in the field through pre-rinsed (300 mL), 47-mm Gelman A/E glass fiber filters with a 0.45-μm pore size and stored in polyethylene bottles. The bottles were soaked with deionized (DI) water overnight and then rinsed with DI water five times. The bottles were further rinsed three times with filtered sample water at the time of collection. The bottles for unfiltered samples were rinsed three times with sample water at the time of collection.

A depth profile of snow was sampled on 19 April 2005, approximately at the peak of snowmelt. A 60-cm snowpit was excavated at an elevation of about 3000 m on the saddle below Redondo Peak (Figure 1), where the snow depth was 60–120 cm in March 2005. A snow core was collected in plastic bags and kept frozen during shipment back to our lab at the University of California, Merced. The bag was pre-rinsed using DI water and washed with snow at least three times at the time of collection.

Subsurface flow (SSF) was sampled on 19–20 July 2005 at the foot of five hillslopes along Redondo and La Jara Creeks (numbered from SSF1 to SSF5 on Figure 1). Samples were taken where subsurface flow surfaces and starts a gully stream. Those samples were handled the same as streamflow samples.

Snow samples were stored frozen (–20°C) for several weeks, then placed in covered polyethylene buckets and melted at room temperature before analysis. Snow and stream samples were analyzed for pH, acid neutralizing capacity (ANC), specific conductance (Cond), and major ions. ANC was measured immediately after melting for snow, or return to the laboratory for water samples, using the Gran titration technique. As Neal [2001] noted, the Gran titration alkalinity may not be exactly equal to ANC, particularly in lower pH (<4.5). Since pH in our water is slightly greater than 7 except for snow and sulfur spring (Table 1), it is believed that the error is not significant. Cations (Ca\(^{2+}\), Mg\(^{2+}\), Na\(^+\), K\(^+\)) and anions (Cl\(^{-}\), NO\(_3^{-}\), and
were analyzed using a Dionex 2000 Ion Chromatograph. Analytical precision (1σ) for all ions was less than 3% and detection limit less than 1 µeq L⁻¹. Charge balance between cations and anions, normalized by the total charge of both cations and anions in µeq L⁻¹, was less than or near 5%.

4. Results

4.1. Chemical Composition

Concentrations of Ca²⁺, Mg²⁺, Na⁺, Cl⁻, and SO₄²⁻ in streamflow at Redondo Creek had similar temporal patterns, with a peak in early spring and lowest values in early summer (Figure 2). For example, Cl⁻ concentrations were about 77 µeq L⁻¹ in December, peaking at 135 µeq L⁻¹ in March, and then gradually decreasing to approximately 35 µeq L⁻¹ through May and June. These solutes were highly correlated, with a Pearson correlation coefficient > 0.8 (n = 11, p < 0.001). The temporal pattern of ANC was different from these solutes, with relatively constant values (602 ± 74 µeq L⁻¹) (Table 1).

Solute concentrations in La Jara Creek were much lower and less variable over time than in Redondo Creek (Figure 2). For example, the mean concentrations of Cl⁻ and SO₄²⁻ were 18 and 82 µeq L⁻¹, respectively, 30–50% of those at Redondo Creek (Table 1). The standard deviations of Cl⁻ and SO₄²⁻ at La Jara Creek were only 4 and 13 µeq L⁻¹, respectively, 10–15% of those for Redondo Creek. Nevertheless, the temporal patterns of solute concentrations were similar between the two creeks for Ca²⁺, Mg²⁺, Cl⁻, SO₄²⁻, and Cond, except that solute concentrations peaked one month later at La Jara Creek than at Redondo Creek. The Na⁺ and ANC concentrations were inversely correlated with Ca²⁺, Mg²⁺, Cl⁻, and SO₄²⁻, with Pearson correlation coefficients < -0.55 (n = 12, p < 0.05).

Figure 2. Time series of chemical compositions in streamflow and mean values in snow, subsurface flow, and thermal meteoric water, with vertical bars showing standard deviation (1σ) when more than one sample was collected. Note that the Spence Spring is out of the ordinate scale for Na⁺ and ANC because of their very high concentrations.
4.2. Diagnosis of Streamflow Chemistry for Mixing

Solute concentrations in snow were 1–10% of those in streamflow from both catchments (Table 1). Solute concentrations in subsurface flow did not vary significantly and were very close to those in streamflow at La Jara Creek for all solutes except ANC and Ca\(^{2+}\) (Table 1 and Figure 2), but were 30–50% of those in streamflow at Redondo Creek. Concentrations in thermal meteoric water at Spence were higher than streamflow for all solutes except Ca\(^{2+}\) and K\(^+\), while those of San Antonio Spring were not distinct from streamflow at Redondo Creek for all solutes except Ca\(^{2+}\), ANC, and Na\(^+\).

Solute concentrations of sulfur spring and deep geothermal fluid were very different from streamflow and thermal meteoric water, with concentrations of Na\(^+\), Cl\(^-\), and SO\(_4^{2-}\) 10–100 times those in streamflow (Table 1). Sulfur spring and deep geothermal fluid did not appear to influence the streamflow chemistry at either Redondo or La Jara Creeks.

4.2. Diagnosis of Streamflow Chemistry for Mixing

The distributions of residuals between measurements and predictions (equation (4)) of streamflow chemistry in 1- and 2-dimensional (1-D and 2-D) mixing space are shown on Figure 3. To assist examination of the residual distribution, p (probability) for slope and RRMSE (equation (5)) were used to measure the degree of randomness. Note that the slopes of the fitted residual lines are negative and equal to \(R^2\) in magnitude.

At Redondo Creek the distributions of residuals of Cond, Mg\(^{2+}\), Na\(^+\), Cl\(^-\), and SO\(_4^{2-}\) were near a random pattern in 1-D mixing space (Figure 3). The \(R^2\) values (as well as the magnitude of the slope) for these solutes were less than 0.1, with \(p > 0.4\). The RRMSE values of these solutes were less than 2% except for SO\(_4^{2-}\) at 4% (Figure 4). The \(R^2\) and p values of these solutes changed slightly in 2-D mixing space. In contrast, the residuals of ANC, K\(^+\), and NO\(_3^-\) were strongly patterned in 1-D mixing space, with \(R^2 > 0.3\) and \(p < 0.06\). Their RRMSE values were also higher than the aforementioned solutes. Therefore the streamflow chemistry was primarily controlled by conservative mixing of two end-members (1-D mixing space) for Cond, Mg\(^{2+}\), Na\(^+\), Cl\(^-\), and SO\(_4^{2-}\).

At La Jara Creek the residuals of all solutes were highly structured against measured concentrations in streamflow in 1-D mixing space, with \(R^2\) usually higher than 0.3 and \(p < 0.07\) (Figure 4). The degree of randomness significantly increased in 2-D mixing space only for Cond, Ca\(^{2+}\), and Mg\(^{2+}\). The residuals in 2-D mixing space were still highly correlated with solute concentrations for most solutes, with \(R^2\) near or > 0.1. The RRMSE values also did not significantly change from 1-D to 2-D mixing space for most solutes (Figure 4). It appears that the streamflow chemistry at La Jara Creek was not controlled by mixing of different end-members.

4.3. Determination of End-Members

The tracers Cond, Mg\(^{2+}\), Na\(^+\), Cl\(^-\), and SO\(_4^{2-}\), which are conservative upon mixing, were used in EMMA for Redondo Creek. Mg\(^{2+}\) and Cl\(^-\), which had the highest p values for the residual distributions, were used to develop an additional EMMA model (Figure 5). U-space projections (PCA scores) were made using eigenvectors extracted from these conservative tracers using a correlation matrix, following equation (1). Temporal variation of the first U-space projection (\(U_1\)) was used to screen end-members and examine if the projections of end-members were distinct from streamflow (Figure 5). Subsurface flow samples collected at each site were treated as potential end-members to account for their spatial variation, following the protocol of Liu et al. [2004] for talus water samples. This analysis was not conducted for La Jara Creek because streamflow at La Jara Creek was not a mixture of end-members, based on the diagnostic tools of mixing models.

End-members were further evaluated based on the distance between their original compositions and U-space projections calculated using equations (2) and (3) (Table 2). For the five-tracer model, the distance of subsurface flow at SSF1 and SSF2 was shorter for all tracers than all other subsurface flow samples. In particular, the distance of Cl\(^-\) was only 7% of the original concentration for SSF2, much shorter than for SSF1 (23%). The subsurface flow at SSF2 appeared to be superior to SSF1 in parameterizing subsurface flow contributing to streamflow. The distance of snow was higher than 50% of the original concentration for all tracers and snow did not appear to be a significant contributing end-member. The distance of thermal meteoric water at Spence and San Antonio was between 30% and 63% for all tracers except Cond, indicating that neither of them was an ideal end-member candidate. However, the distance of thermal meteoric water at Spence was much shorter than San Antonio for Mg\(^{2+}\) and Cl\(^-\). Thus thermal meteoric water at Spence was selected to represent the thermal meteoric water that contributes to streamflow in the five-tracer model.

Compared to the five-tracer model, the distance was significantly shortened for all tracers in the two-tracer model, particularly for SSF2 and thermal meteoric water at Spence (Table 2). The distance of SSF2 was only 1% and 6% for Mg\(^{2+}\) and Cl\(^-\), respectively. The distance of thermal meteoric water at Spence was 4% and 6% for Mg\(^{2+}\) and Cl\(^-\), respectively. The same as for the five-tracer model, SSF2 and thermal meteoric water at Spence were selected as two contributing end-members in the two-tracer model.

4.4. Quantification of End-Member Contributions

The calculated end-member contributions were very different for the five-tracer versus two-tracer EMMA models, even though the same end-members were used (Figure 6). Subsurface flow and thermal meteoric water contributed 91% and 9% on average, respectively, for the five-tracer model compared to 77% and 23% for the two-tracer model. However, the temporal pattern of end-member contributions was identical for the two model results (\(R^2 = 1.0\)). The percent of subsurface flow in the total discharge was lowest in early March at the peak of snow accumulation, as shown
Figure 3. Distribution of residuals against measured solute concentrations in streamflow under 1-D and 2-D mixing space for Redondo Creek and La Jara Creek. $R^2$ and $p$ value for slope were shown for fitted lines; the slope was the same as $R^2$ with a negative sign.
by SWE and SCA data (Figure 6), and gradually increased through late May at the end of snowmelt.

Streamflow chemistry was re-created using the EMMA results and chemical compositions in end-members (Figure 7). Regression lines between measurements and predictions from EMMA results were forced to have zero intercept so that the goodness of prediction was evaluated on the basis of 1:1 line ($y = x$). For the five-tracer model, concentrations of Cl$^-$, Cond, and Na$^+$ were only fairly well predicted since either their slope or $R^2$ was much lower than 1. Prediction of Mg$^{2+}$ was poor, with $R^2$ of only 0.23. Prediction of SO$_4^{2-}$ was apparently not successful since its intercept could not be set as zero in order to get a realistic $R^2$ value. Overall, the five-tracer model was not successful.

In contrast, the two-tracer model was reasonably successful. Concentrations of Mg$^{2+}$ and Cl$^-$ in streamflow were reproduced with both slope and $R^2$ close to 1 (Figure 7). Concentrations of Cond, Na$^+$ and SO$_4^{2-}$ were also reproduced reasonably well, particularly for SO$_4^{2-}$, which yielded a slope of 0.71 and a $R^2$ of 0.56, though these tracers were not used in the model.

5. Discussion
5.1. Evaluation of End-Member Mixing Analysis Results

Spatial and temporal variations in the chemical composition of end-members may result in significant uncertainty in hydrograph separations [Genereux, 1998]. In the current study the uncertainty caused by spatial variability of chemical composition in subsurface flow appears to be not significant. Except for the subsurface flow at SSF4, which drains to a stream that is disconnected from Redondo Creek (Figure 1), the $U$-space projections of subsurface flow at all sites collected at the same time were very close to each other (Figure 5). Subsurface flow was generated from the same soil in the two catchments, the Redondo rubble land association (Figure 1), which may explain why the chemical compositions in subsurface flow were relatively invariant across different locations.

Though subsurface flow was only sampled once in this study, its temporal variability may be seen from streamflow at La Jara Creek (Figure 2), which was primarily composed of subsurface flow alone as discussed in next section. Similar to the outflow of a ponderosa pine hillslope near Los Alamos [Newman et al., 1998], the solute concentrations in streamflow at La Jara Creek varied over time (Figure 2). A sensitivity test of EMMA was conducted by substituting the time series of chemical compositions in streamflow at La Jara Creek for the single subsurface flow value represented by SSF2 in the two-tracer model for Redondo Creek. The EMMA results (data not presented) using constant and varying values for subsurface flow were almost identical, with a $R^2$ of 0.99 ($n = 11$, $p < 0.001$). The estimated mean contribution of subsurface flow to streamflow at Redondo Creek increased only 2%, from 77% to 79%, when time series of chemical compositions in La Jara Creek was used for subsurface flow. It is because of the significant difference of chemical compositions between

Figure 4. Relative root mean square error (RRMSE) under 1-D and 2-D mixing space for (top) Redondo Creek and (bottom) La Jara Creek.

Figure 5. Temporal variation of the orthogonal projections ($U_1$) by the first PCA eigenvectors extracted from (a) five tracers and (b) two tracers.
thermal meteoric water and subsurface flow that it did not cause a significant error in EMMA when the variability of chemical compositions in subsurface flow was ignored or not known.

[35] The very different EMMA results using all five tracers (Cond, Mg$^{2+}$, Na$^+$, Cl$^-$, and SO$_4^{2-}$/C$_0$) versus just two (Mg$^{2+}$ and Cl$^-$/C$_0$) (Figure 6) appear to be primarily caused by the chemical composition of thermal meteoric water at Spence. The chemistry of this thermal water doesn’t fit well the mixing space determined using all five tracers at Redondo Creek, though it is better than thermal meteoric water at San Antonio (Table 2). If only Mg$^{2+}$ and Cl$^-$/C$_0$ are used, however, it does fit the mixing space very well. Thus thermal meteoric water that discharges to Redondo Creek may be very different from thermal meteoric water at Spence for Na$^+$ and SO$_4^{2-}$/C$_0$ concentrations, but similar for Mg$^{2+}$ and Cl$^-$/C$_0$ concentrations. In fact, the SO$_4^{2-}$ concentration was lower in the Spence thermal water than in streamflow on 19 April 2004 (Figure 2), suggesting that the thermal meteoric water contributing to Redondo Creek may have a much higher SO$_4^{2-}$ concentration.

5.2. How Important is Overland Flow Versus Subsurface Flow and Groundwater?

[36] EMMA results indicate that subsurface flow from hillslopes exerts the major control on Redondo Creek, consistent with a study at an 870-m$^2$ ponderosa pine hillslope near Los Alamos [Wilcox et al., 1997]. Streamflow chemistry at La Jara Creek is primarily controlled by a single flow component (Figure 3). Since chemical compositions at La Jara Creek were very close to those of subsurface flow (Figure 2), this single flow component is deemed to be subsurface flow. Soils in the La Jara Creek catchment are primarily composed of well-drained soils, Redondo rubble land associations (Figure 1), which facilitate snowmelt infiltration. Consistent with Newman et al. [1997, 1998, 2004], subsurface flow at both catchments appears to be generated from a saturated zone in the lower soil horizons above the bedrock in hillslopes, essentially lateral subsurface flow. Lateral subsurface flow from hillslopes is not commonly considered an important agent of runoff generation in semiarid environments [Wilcox et al., 1997], though some previous studies have found pedogenic evidence that it does occur [Thorns, 1983]. This result highlights the importance of lateral subsurface flow in hydrologic and biogeochemical studies in semiarid environments.

[37] EMMA results indicate that overland flow (either infiltration- or saturation-excess) did not exert a significant control on streamflow during the snowmelt period at either catchment, in agreement with past findings that overland flow rarely occurs in ponderosa pine forests [Dunford, 1954; Heede, 1984; Williams and Buckhouse, 1993].
observations also do not support overland flow even during the peak of snowmelt at both catchments. However, Wilcox et al. [1997] found that infiltration-excess overland flow was the second largest flow component (after lateral subsurface flow) at their site, occurring primarily in winter, particularly if soils are frozen before a snowpack develops. Soils in our catchments are primarily composed of well-drained rubble land associations (Figure 1), implying that infiltration rate of snowmelt is high. It is unclear whether or not the soils in our catchments were frozen during WY2005, but catchment size may play a significant role in the difference between the results. Our catchment area is at least 4000 times that of Wilcox et al. [1997]. Infiltration-excess overland flow may be generated from one area of a catchment rather than the entire catchment.

[38] The contribution of thermal meteoric water to streamflow at Redondo Creek was responsive to snowmelt (Figure 6). Thermal meteoric water circulates only in the upper 500 m of the moat zone of the caldera [Goff and Grigsby, 1982]. The release of thermal meteoric water was

**Figure 7.** Prediction of conservative tracer concentrations at Redondo Creek using the results of five-tracer and two-tracer based EMMA. The intercept of the fitted line was set to be zero to compare with 1:1 line ($y = x$). Note that a zero-intercept cannot be fitted for SO$_4^{2-}$. 

Using 5 tracers: Cond, Mg$^{2+}$, Na$^+$, Cl$^-$, and SO$_4^{2-}$

Using 2 tracers: Mg$^{2+}$ and Cl$^-$
probably driven by snowmelt infiltration. The infiltration of snowmelt creates a higher hydraulic gradient, displacing thermal meteoric water stored in the fractured rocks, akin to the piston flow mechanism as described by Buttle [1994]. In agreement with thermal meteoric water, the discharge measured from a sulfur spring at Sulfur Springs also varied over time, peaked in early summer [Goff and Grigsby, 1982]. The contribution of thermal meteoric water, which has higher chemical concentrations than lateral subsurface flow, explains why the chemical concentrations in streamflow at Redondo Creek were higher than at La Jara Creek (Figure 2).

5.3. Are Mixing Models Still an Effective Tool in Modeling Streamflow Chemistry and Flow Paths?

[39] Mixing models were often used in a very simple way, e.g., one tracer for two components and two tracers for three components based on mass balance equations of tracers and water, and expressed as linear equations as follows, using two tracers for three components as an example [e.g., Hooper and Shoemaker, 1986].

\[ A_f 1 + A_f 2 + A_f 3 = A_s \] (6)

\[ B_f 1 + B_f 2 + B_f 3 = B_s \] (7)

\[ f_1 + f_2 + f_3 = 1 \] (8)

where \( f \) is the fraction of total streamflow discharge due to an end-member; \( A \) and \( B \) represent compositions of tracers \( A \) and \( B \); subscripts \( 1, 2, 3 \), and \( s \) represent end-members \( 1, 2, 3 \), and streamflow. Two key assumptions usually made are that tracers are conservative and the number of end-members is known [e.g., Buttle, 1994]. However, conservative tracers and number of end-members are usually not known \textit{a priori}. Hydrologists have relied on analysis of watershed hydrology and geology to identify conservative tracers and to acquire the number of end-members. Because of a lack of sufficient information on hydrology and geology in some, if not most, catchments, conservative tracers and the number of end-members are sometime misinterpreted, resulting in significant errors or even an unrealistic presentation of the conceptual model of streamflow generation for a watershed, as illustrated by Burns et al. [2001]. This problem is the major reason why the use of mixing models has diminished [Burns, 2002].

[40] Taking the reverse procedure from the above mixing model (equations (6) and (7)), diagnostic tools of mixing models were used to determine the rank of streamflow chemical data, or in other words, to decompose the data set into a combination of linear equations using eigenvectors of streamflow chemical data. Note, however, that diagnostic tools of mixing models are not used to exactly reconstruct equations (6) and (7). Instead, the tools are used to test if such linear equations can be established (through distribution of residuals between measurements and projections) and if so, how many equations (the number of eigenvectors, which is one less than number of end-members) are needed and which solutes can be expressed linearly (therefore conservative tracers).

[41] If streamflow chemistry is dominated by chemical interactions between streamwater and rocks and soils, streamflow chemistry cannot be decomposed as linear equations because multivalent ions involve polynomial processes (number of charges serves as power of ionic concentrations in equilibrium constant equation). If streamflow is dominated by subsurface flow from myriad sources with varying residence times due to the heterogeneity of subsurface media, streamflow chemistry follows a power law of fractal contributions of source waters [Kirchner et al., 2000]. In these cases, the residuals of streamflow chemistry between measurements and projections by eigenvectors of PCA should show a structured pattern with pertinent measurements.

[42] Mixing and fractal analysis of streamflow chemistry complement each other in understanding controls of water quality and (bio)geochemical processes in watershed hydrology. Mixing models apply to watersheds where flows from different hydrologic units have distinct chemical signatures or periods when multiple pathways are activated. In snowmelt-dominated catchments such as Redondo Creek of this study (Figure 2) and Green Lakes Valley in Colorado [Liu et al., 2004], chemical compositions in overland flow or lateral subsurface flow are distinct from groundwater, even though variable over time. In rainfall-dominated catchments such as those in Phynlimon, Wales [Kirchner et al., 2000], there are no flow components that can be defined as end-members. Diagnostic tools of mixing models provide a means to help hydrologists evaluate if end-members can be defined and thus determine if end-member mixing analysis is potentially a useful tool in a study. Note that end-member mixing for conservative tracers depends on there being no chemical reactions once source waters from different hydrologic units converge in streams, and does not explicitly account for geochemical weathering and chemical interactions between water and rocks and soils within each end-member reservoir or hydrologic unit.

[43] EMMA used in this study is not simply a replacement of traditional mixing models (e.g., two tracers for three components) as a mathematical manipulation, but is part of the process to test a conceptual model of streamflow generation established using chemical tracers. The projection of chemical compositions in end-members using eigenvectors extracted from streamflow chemistry provides an effective means to evaluate if chemical compositions in end-members fit in mixing spaces of streamflow chemistry and thus determine eligibility of end-members (e.g., Spence Spring versus San Antonio Spring for thermal meteoric water at Redondo Creek). Re-creation of streamflow chemistry for conservative tracers using the results of EMMA and chemical compositions in end-member provides a quantitative assessment of model results and ensures that the results are physically meaningful. This re-creation is not recurring because correlations of solutes instead of concentrations are used in PCA to solve for contributions of end-members to streamflow. Particularly, if some tracers are not used in EMMA but used in predictions, it may enhance the evaluations such as the predictions of Cond in the two-tracer model at Redondo Creek (Figure 7).

[44] The combination of diagnostic tools of mixing models and end-member mixing analysis not only helps to enforce the assumptions of mixing models but also enhances
results of mixing models, particularly when chemical data are limited. The combination of these tools reduces uncertainties, particularly in choosing the number of end-members, or conservative tracers, i.e., the so-called “model uncertainty” [Joerin et al., 2002]. Also, EMMA uses correlations between tracers other than concentrations of tracers. Ionic concentrations in streamflow significantly varied from snowmelt to summer dry seasons (Figure 2). With several samples collected during snowmelt, however, the data are sufficiently distributed over the year to capture the seasonal variability of ionic concentrations at Redondo Creek (Figure 2). It is assumed that the correlation coefficients between tracers will not be significantly changed if more samples during snowmelt season are added. However, it merits further work in future to evaluate the sensitivity of correlation coefficients to the number of samples, particularly during snowmelt season.

6. Conclusions

[45] Streamflow in the semiarid, seasonally snow-covered, forested catchments in Valles Caldera, New Mexico is controlled primarily by lateral subsurface flow. Overland flow is not a significant contributor to streamflow. Because of similar climate, vegetation and soils, the dominance of lateral subsurface flow in regulating streamflow should apply to the many similar catchments elsewhere in the Southwest United States. The results of this study improve our understanding of streamflow generation in these catchments and thus should help understand biogeochemical cycling and ecohydrology.

[46] Mixing models based on natural tracers remain effective tools in watershed hydrology, particularly during snowmelt period in snowmelt-dominated catchments. Our results show that a combination of diagnostic tools of mixing models and end-member mixing analysis helps determine conservative tracers, identify end-members and validate quantitative results of mixing models; this combination enables renewed use of mixing models and builds confidence in the results. Diagnostic tools of mixing models are useful in determining the dominant processes of streamflow chemistry, end-member mixing versus fractal and non-conservative behavior over the time scales investigated using streamflow chemistry data alone. One can learn from the results if end-members can be defined and thus greatly help map research plans and strategies. The combination of diagnostic tools of mixing models and end-member mixing analysis also reduces modeling uncertainties through determination of the number of end-members and conservative tracers and provides a conceptual understanding of streamflow generation, even if only limited chemical data are available. It is recommended that diagnostic tools of mixing models be applied before EMMA or any tracer-based hydrograph separations are performed to assure the quality of hydrograph separations.

[47] Acknowledgments. The authors thank J. Kostrzewski, R. Parmenter, S. Gifford, J. Gamlin, and M. Frisbee for their assistance in stream sampling; K. Anne, S. May, and G. Shaw for their assistance in sample analysis; and Tom Painter for processing of the remote sensing data. Funding was provided by National Science Foundation’s Center for the Sustainability of Semiarid Hydrology and Riparian Areas (SAHRA) (NSF EAR9876800). Research at Lawrence Berkeley National Laboratory was supported by the Director, Office of Energy Research, Basic Energy Sciences, Chemical Sciences Division of the U.S. Department of Energy under Contract DE-AC02-05CH11231.

References


Muldavin, E., and P. Tonne (2003), A Vegetation Survey and Ecological Assessment of Valles Caldera National Preserve, New Mexico, 118 pp., Natural History, Albuquerque, N. M.


Rzonca, B., and D. Schulze-Makuch (2003), Correlation between microbiological and chemical parameters of some hydrothermal springs in New Mexico, USA, *J. Hydrol.*, 280, 272–284.


R. C. Bales, M. H. Conklin, and F. Liu, Sierra Nevada Research Institute and School of Engineering, University of California, P.O. Box 2039, 5200 N. Lake Road, Merced, CA 95344, USA. (fliu@ucmerced.edu)

M. E. Conrad, Lawrence Berkeley National Laboratory, Building 70A, Room 4418, Berkeley, CA 94720, USA.