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

C. Güler · G. D. Thyne

We thank the author for his comments that raise several important points (Dreher 2003). Our goal in this first paper was to offer a comparison of different graphical and statistical methodologies available for hydrogeologists in the context of a typical regional dataset and to point out the limitations of the purely graphical approach in defining hydrochemical facies. Dr. Dreher notes that consideration of the geology and chemistry of the study are minimal in our paper. We acknowledge the comment, but the article was constrained by the normal editorial procedures as to length and content. A more complete discussion of those aspects can be found in the dissertation of Güler (2002), and a series of papers that are in review that explicitly consider the effects of both chemical and physical factors.

The fundamental concept underlying the idea of hydrochemical facies is that there is a direct relationship between water chemistry and the lithological properties and hydrochemical processes that control that chemistry. The methodology of first differentiating the samples with a statistical technique (Q-mode cluster analysis) offers an efficient and objective method to evaluate the similarity between different samples. Considering the spatial distribution of samples based on their similarity allows the user to combine the two types of information (hydrochemical and spatial) into one representation and directly evaluate hydrochemical facies in a spatial context. For instance, the spatial coherence of the statistical clusters in Fig. 4 is

exactly what you would expect in a system where water chemistry is dominated by a few underlying hydrochemical processes. Furthermore, the spatial clusters defined by Q-mode analysis help define the spatial distribution of lithological or hydrochemical conditions that represent the underlying processes. This approach seems to function well on a range of spatial scales from our regional scale down to the local scale (Suk and Lee 1999). This information combined with the results of R-mode analysis, which extracts the specific factors that control the chemical variability, better constrain any geochemical models of hydrochemical evolution. It also highlights outliers or groups of samples that are controlled by additional factors from the more pervasive background.

A portion of the comments is directed at the specific statistical procedures that we used. One comment suggested that we could have used non-parametric statistical methods for data with non-normal distributions. That option was considered, but we felt that most hydrogeologists are not familiar with these techniques. Instead we chose to use parametric methods that would have some familiarity to most users. We also found, while avoiding the complication of applying the non-parametric techniques to the parameters that did have normal distributions, that the parametric techniques produced very good results.

The author notes that log transformation of skewed or bimodal chemical datasets may not be appropriate as it causes information to become “lost,” specifically information regarding anomalies (we had no bimodal distributions in our dataset), and that standardization of the data shifts relationships between the ions and neglects information derived from relative abundances. The reason we used these particular procedures (transformation and standardization) was twofold. We used transformation to maximize the effectiveness of the parametric methods and standardization to avoid misclassifications arising from the inappropriate weighting of parameters with large magnitudes. In addition, our experience with multivariate analysis has been that as long as more than eight parameters are used, delineation of the outliers remains clear.

Regarding the comment that information from outlying points was lost in the transformation, anomalies are

Received: 4 February 2003 / Accepted: 3 July 2003
Published online: 2 September 2003

© Springer-Verlag 2003

C. Güler (✉)
Çiftlikköy Kampüsü,
Mersin Üniversitesi,
Jeoloji Mühendisliği Bölümü, 33343 Mersin, Turkey
e-mail: cguler@mersin.edu.tr
Tel.: +90-324-3610001 ext. 605
Fax: +90-324-3610032

G. D. Thyne
Department of Geology and Geological Engineering,
Colorado School of Mines,
1500 Illinois Street, Golden, CO, 80401, USA

preserved as outlying points that form the edges of the transformed distributions and that variance is still incorporated during the clustering process. The loss of the information about relative differences in abundance between ions during standardization was still preserved in the means of raw data for each cluster. We agree that you can get different results using non-transformed datasets, but in our study area those clustering results were not as useful as the transformed data. Ultimately, we would encourage users to try several variations (raw, transformed, transformed and standardized) and compare the results.

So while we acknowledge the limitations of defining groups or subgroups using only statistical analysis without considering the geological and hydrogeological settings (aquifer layers, catchment areas, etc.), we would hope our contribution would be considered in the larger context; that using statistical clustering to objectively separate samples into groups (clusters) followed by plotting those clusters in a spatial context actually allows for a more robust interpretation and leads to a better hydrochemical evolution model that can account for all the samples. For our field area, this approach was much more useful in defining the hydrochemical evolution than traditional methods used in the past (Whelan et al. 1989; Houghton 1994; Berenbrock and Schroeder 1994; Thyne

et al. 1999) and removed much of the subjective bias inherent in traditional methods.

References

- Berenbrock C, Schroeder RA (1994) Ground water flow and quality, and geochemical processes in Indian Wells Valley, Kern, Inyo and San Bernardino Counties, California, 1987–88. USGS Water-Res Inv Rep, pp 93–4003
- Dreher T (2003) Comment on: Güler C, Thyne GD, McCray JE, Turner AK (2002): evaluation of graphical and multivariate statistical methods for classification of water chemistry data. *Hydrogeol J* 10:455–474
- Güler C (2002) Hydrogeochemical evaluation of the groundwater resources of Indian Wells–Owens Valley area, southeastern California. PhD Thesis, Colorado School of Mines, Golden, 262 pp
- Houghton BD (1994) Ground water geochemistry of the Indian Wells Valley. Master's Thesis, California State University, Bakersfield, 85 pp
- Suk H, Lee K-K (1999) Characterization of a ground water hydrochemical system through multivariate analysis: clustering into ground water zones. *Groundwater* 37:358–366
- Thyne GD, Gillespie JM, Ostidick JR (1999) Evidence for interbasin flow through bedrock in the southeastern Sierra Nevada. *Geol Soc Am Bull* 111:1600–1616
- Whelan JA, Baskin R, Katzenstein AM (1989) A water geochemistry study of Indian Wells Valley, Inyo and Kern Counties, California. Naval Weapons Center Tech Publ 7019, Geochemistry Study and Appendix A