Groundwater

Technical Commentary/

A Background Color Scheme for Piper Plots to Spatially Visualize Hydrochemical Patterns

by Luk Peeters

Abstract

The combination of ternary diagrams of cations and anions with a central diamond graph make the Piper plot very useful in visualizing groundwater chemistry datasets. One of the major drawbacks is that it is hard to link spatial attributes of the dataset to the plot. In this study, we propose a background color scheme of the Piper plot so that spatial representations of these data can be colored according to their location in the Piper plot. The color scheme is chosen to have maximum resolution while still being perceptually uniform. The linking between Piper plot and maps through this color scheme allows the interpretation of the trends and processes deduced from the Piper plot in terms of the location in the aquifer, the geology, and the groundwater flow dynamics. The colored Piper plot is applied to a groundwater quality dataset of the Condamine Alluvium in Queensland, Australia.

Introduction

The Piper plot is one of the most widely used visualization techniques for groundwater chemistry data. The combination of a ternary diagram of major cation composition and a ternary diagram of major anion composition into a central diamond plot makes it an invaluable tool to identify groups and deduce the main hydrochemical processes affecting the dataset (Piper 1944).

A major drawback of the Piper plot is that it is often not straightforward to link the position of a chemical sample on the plot to its sampling location. A priori grouping of hydrochemical data, such as according to aquifer sampled, can be visualized through different combinations of symbols and colors. Such a priori grouping is often not possible and a variety of clustering and dimensionality reduction techniques such as principal component analysis, K-means clustering and neural networks are used for grouping. An overview of such techniques is provided in Güler et al. (2002). These data-based groupings can subsequently be mapped in order to relate the groupings to the geology, groundwater flow patterns and land use (e.g., Woocay and Walton 2008; Kim et al. 2009; Vandenbohede and Lebbe 2012). Although more chemical variables than just the major anions and cations can be included in the analysis, the interpretation is hampered as the direct link to the Piper plot and the underlying geochemical processes is lost.

In addition, most hydrogeochemical processes, such as mixing of waters with different chemical composition and cation exchange, lead to gradual and continuous changes in chemical composition (Appelo and Postma 2005). While the Piper plot is ideally suited to accommodate continuous variations, in clustering techniques this will lead to ill-defined groups and samples that can belong to more than one group. Fuzzy or probabilistic clustering techniques can mitigate this effect (Güler et al. 2012), although these generally make the interpretation more challenging.

In this study a continuous color scheme for the ternary diagrams and central diamond of the Piper plot is proposed. By assigning these colors to the sampling locations on maps and cross-sections, a direct link between the sampling location and the Piper plot is established to facilitate interpretation of hydrochemical data. A similar continuous color scheme has been applied in visualizing the results of a Self-Organizing Map based exploratory data analysis of a groundwater quality dataset by Peeters et al. (2007). Nicolet and Erdi-Krausz (2003) recommend

CSIRO Land and Water, Groundwater Hydrology Program, PO Box 2, Glen Osmond, SA 5064, Australia; luk.peeters@csiro.au
Received March 2013, accepted August 2013.

doi: 10.1111/gwat.12118

a colored ternary diagram to spatially visualize the radium, thorium, and potassium concentrations measured from remote sensing. A similar approach is adopted in this study and extended so that it can be applied to the central diamond of the Piper plot as well.

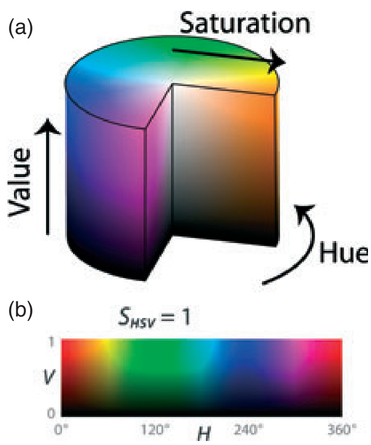

Figure 1. Hue, Saturation, Value (HSV) color scheme. (a) Cut-away 3D visualization of HSV scheme and (b) outside shell of 3D cylinder showing variation in color with hue and value for a constant saturation of 1 (adapted from Rus 2010).

The next section introduces the color scheme and is followed in the results and discussion section by an application of the technique to a dataset from the Condamine River Alluvium in Queensland, Australia.

Methods

Choosing a color scheme for a triangular- or diamond-shaped diagram is not trivial. Firstly, the color scheme needs to be perceptually uniform, that is, the perceived difference between colors assigned to data points needs to correspond to the difference in chemistry between the data points. Secondly, the color scheme has to provide sufficient resolution, that is, variation in color, such that even subtle differences in groundwater chemistry can be perceived. Lastly, the color scheme must be able to be represented in most commonly used scripting and plotting software.

The Hue, Saturation, Value (HSV) color scheme is a widely used color scheme and provides an acceptable compromise between these three objectives (Schwarz et al. 1987, Figure 1). In the HSV scheme, hue is specified as an angle between 0° and 360° with 0° corresponding to red. Saturation and value are both defined between 0 and 1, where a saturation of 1 is the fully saturated color and a saturation of 0 is fully unsaturated or white (Figure 1a).

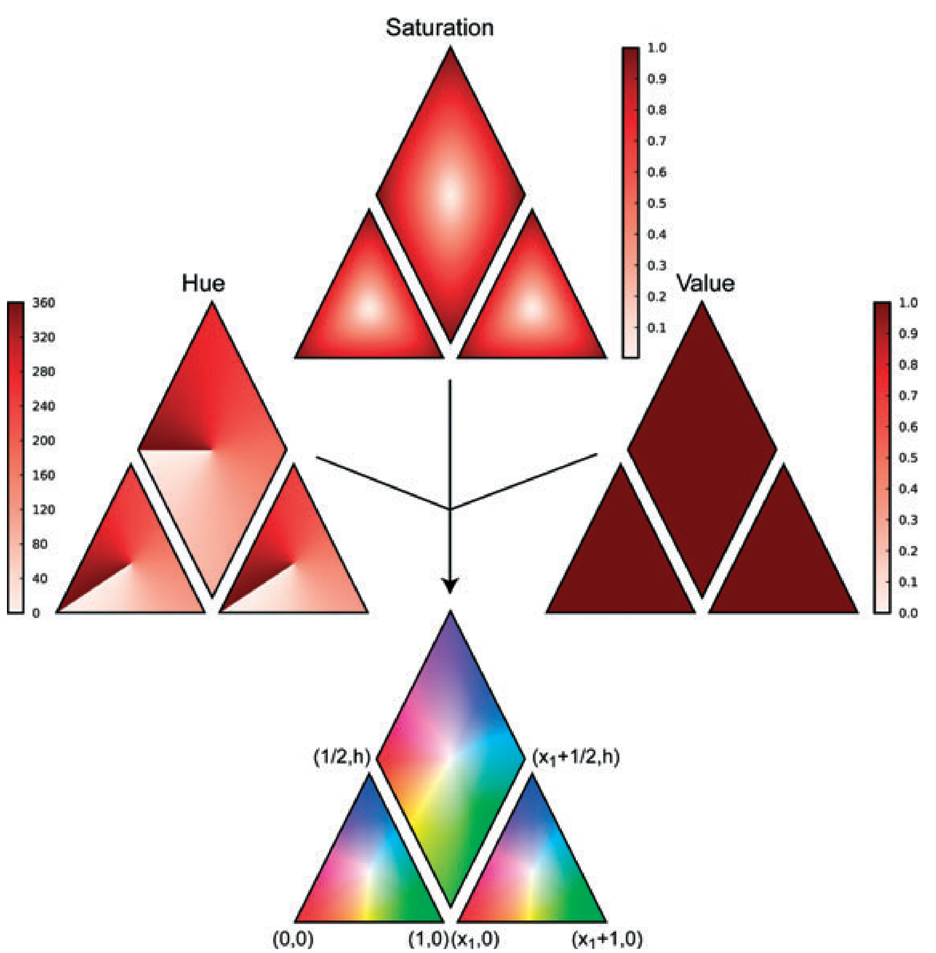

Figure 2. Components of the HSV scheme in the Piper plot and the resulting color-coded Piper plot.

The value is a measure of the brightness of the color, with 0 equal to black.

The HSV color scheme spans a wide variety of colors, providing ample resolution, while the HSV values can easily be transformed into red, green, and blue values, which are used in most scripting and plotting software to visualize colors. The scheme is not entirely perceptually uniform; for example, there is visually more difference over the same hue angle in the yellow section than there is in the green section of the color scheme (Figure 1b). In the following implementation of the HSV on the triangular- and diamond-shaped diagram, effort is made to minimize this effect.

The application of the color scheme to the Piper plot is carried out in three steps; firstly, a Cartesian grid is defined over the Piper plot and the data points are expressed in this coordinate system. Secondly, the variation in hue, saturation, and value is defined in function of the x and y coordinates. Finally, these functions are evaluated over the grid spanning the Piper plot and for each data point. The HSV values are then transformed to RGB for ease of plotting.

The cation ternary diagram plots the proportion of Ca^{2+}, Mg^{2+}, and Na^{+}+K^{+} in meq/L, with Ca^{2+} in the lower left corner, Mg^{2+} in the top corner, and Na^{+}+K^{+} in the right corner. Let the origin of the Cartesian coordinate system coincide with the left corner of the cation diagram and let the coordinates of the right corner be (1,0). As the ternary diagram is an equilateral triangle, the coordinates of the top corner are $\left(\frac{1}{2}, h\right)$ with $h = \frac{1}{2}\tan\left(\frac{\pi}{3}\right)$.

As the proportions of cations need to sum to 1, only two components are needed for plotting. The Cartesian coordinates for the cations of an arbitrary sample can be written as (Mertie 1964):

$$x_{\text{cat}} = \left[\text{Na}^{+} + \text{K}^{+}\right] + \frac{\left[\text{Mg}^{2+}\right]}{2}$$

$$y_{\text{cat}} = h\left[\text{Mg}^{2+}\right]$$

The anion ternary diagram has endmembers HCO_3^{-}+CO_3^{2-}, SO_4^{2-}, and Cl^{-}. The coordinates for the anions can be written as:

$$x_{\text{an}} = x_1 + \left[\text{Cl}^{-}\right] + \frac{\left[\text{SO}_4^{2-}\right]}{2}$$

$$y_{\text{an}} = h\left[\text{SO}_4^{2-}\right]$$

with x_1 as the x-coordinate of the left corner of the anion diagram.

The location of a sample in the central diamond is the intersection of the projection of the cation position along the Ca^{2+}-Mg^{2+}-axis and the projection of the anion position along the Cl^{-}-SO_4^{2-}-axis. This can be expressed as:

$$x_{\text{d}} = \frac{1}{4h}\left(y_{\text{an}} - y_{\text{cat}}\right) + \frac{1}{2}\left(x_{\text{an}} + x_{\text{cat}}\right)$$

$$y_{\text{d}} = \frac{1}{2}\left(y_{\text{an}} + y_{\text{cat}}\right) + h\left(x_{\text{an}} + x_{\text{cat}}\right)$$

The second step is formulating a function to map the color scheme to the Cartesian grid for both triangular diagrams and the central diamond diagram. As hue is specified as an angle, it is chosen to be equal to the angle between the line spanned between the left most corner and the center of the diagram and the line spanned between an arbitrary (x,y) location and the center of the diagram (Figure 2). The angle is measured counter-clockwise from the line between the left corner and the center. The resulting color scheme assigns the primary colors, red, green, and blue to the endmembers of the composition diagram.

Saturation is specified to vary from 1 at the edges of the diagram to 0 in the center. This ensures that all colors blend to white in the center of the diagram. A linear interpolation, however, would result in a color scheme with a jagged appearance which compromises the condition for a perceptually uniform color scheme. To alleviate this, Clough-Tocher interpolation is used (Alfeld 1984) to create a smoothly varying saturation surface on the triangles and diamond. This interpolation leads to a large area with white-dominated colors centrally in the diagram, which decreases the resolution of the visualization. This is overcome by taking the square root of the saturation values after interpolation (Figure 2).

Value is set uniform at 1 to provide colors with maximal brightness. This implies that from the three variables available in the HSV scheme only two are used in defining the color scheme.

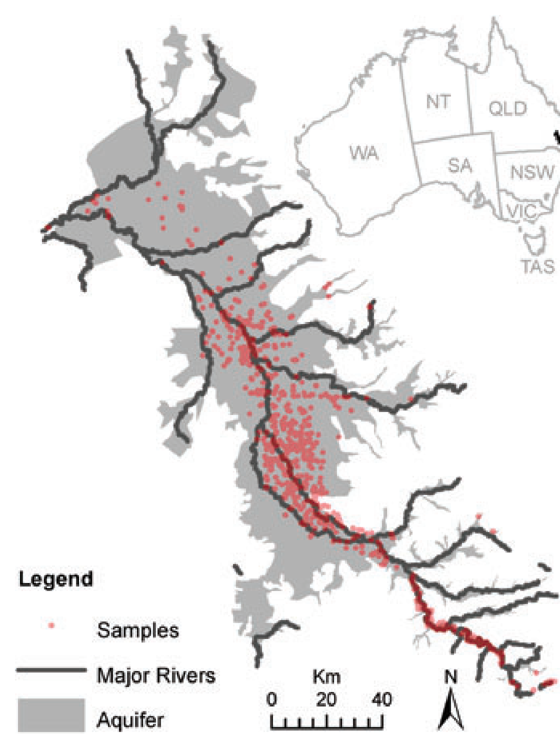

Figure 3. Condamine River alluvium in Queensland, Australia, with the location of the 752 sampling locations with cation and anion data.

In the final step, the resulting HSV values are converted to RGB values to facilitate plotting. As the color scheme is expressed in function of the (x,y) coordinates of the Piper plot, the HSV or RGB triples for a hydrochemical sample can be calculated directly and subsequently used in spatial plotting of these data.

The plotting algorithm is implemented in Python 2.7.2 using the libraries NumPy 1.6.1 (Oliphant 2006) and Matplotlib 1.1.0 (Hunter 2007) and is provided in Appendix S1.

Study Area

A hydrochemical dataset of the Condamine River Alluvium in Queensland, Australia is chosen as test case for this methodology (Figure 3). The predominantly coarse sand and gravel aquifer is formed by the quaternary alluvial deposits of the Condamine River centrally in the valley and colluvial sheetwash deposits near its margins (CSIRO 2008). The aquifer is hydraulically connected to Tertiary basalt aquifers in the east and to the underlying Jurassic and Cretaceous sand and mudstones from the Great Artesian Basin. The alluvial sediments are covered with a thick, clayey soil that limits diffuse recharge. Most of the recharge is from the Condamine River, which drains the Paleozoic basement in the south of the aquifer. While the river is considered losing in the entire study area, the northern anabranch in the central region is maximally losing or disconnected because of intensive pumping for irrigated agriculture (CSIRO 2008).

The dataset is obtained from the Queensland Government and consists of 752 sampling locations where major ion chemistry is analyzed (Figure 3). Details on the monitoring network and sampling protocol can be found through the resource center of the Queensland Government (http://www.derm.qld.gov.au/services_resources/index.php).

Results and Discussion

Figure 4 shows the Piper plot with the background color scheme and the corresponding color-coded maps. From the Piper plot, it is clear that the cations cluster along a line with equal proportions of Ca^{2+} and Mg^{2+} with increasing $Na^{+}+K^{+}$ proportions. Similarly, the anions

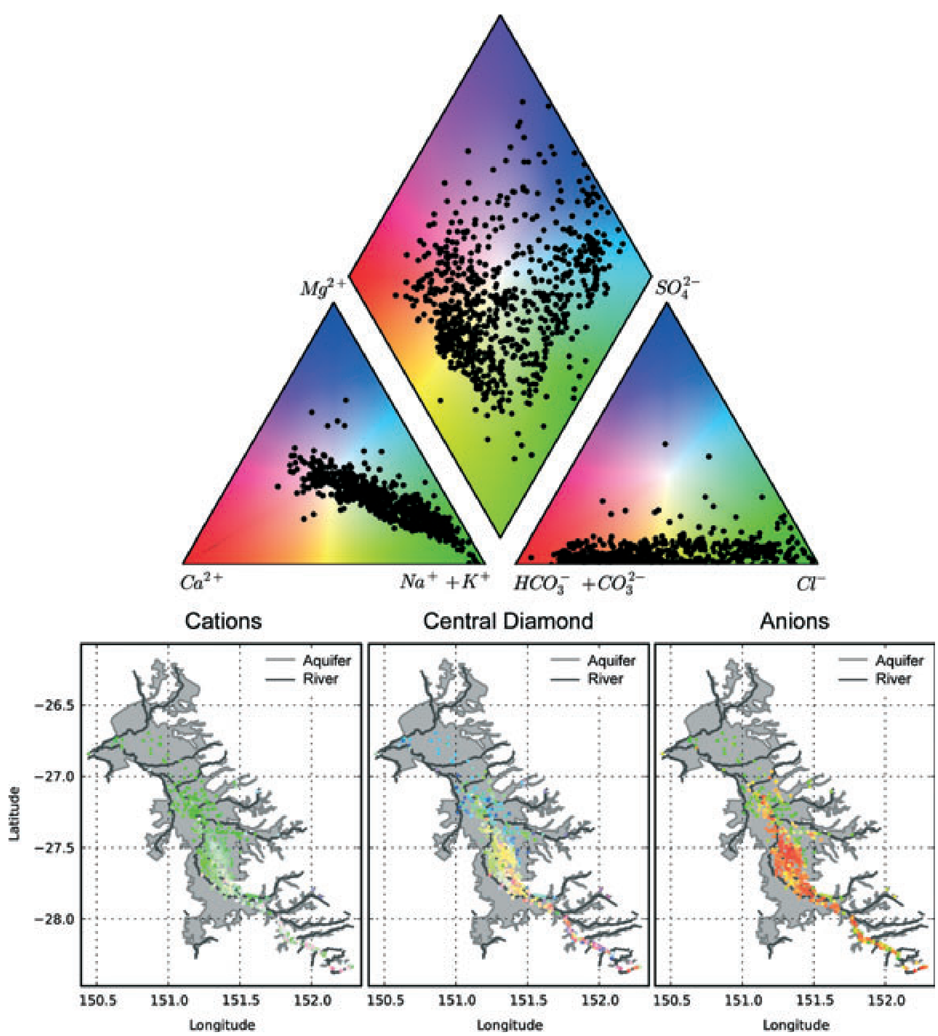

Figure 4. Color-coded Piper plot with corresponding color-coded maps for the Condamine Alluvium dataset.

cluster along the line between alkalinity and chloride. The cation and anion maps show that the chemical composition moves toward the Na^+ and Cl^- endmembers in a downstream direction and away from the river. These trends are also apparent in the plot and map of the central diamond, where water gradually changes from $CaMgHCO_3$-type in the red and purple colors to $NaHCO_3$-type in green and yellow and finally to NaCl-type in the blue colors.

This pattern of groundwater chemistry agrees with low diffuse recharge and dominant recharge from the river (Huxley 1982). Low diffuse recharge due to low rainfall and high evapotranspiration results in elevated salinity of the recharging water and thus the aquifer. Recharge from the river introduces water with lower salinity into the aquifer. The NaCl-type waters in the north and east of the study area are linked to lower transmissivity of the aquifer, and hence longer residence times, and interaction with deeper bedrock formations. Besides the general trends, this methodology also allows to visualize subtle anomalies, such as a zone with slightly elevated Cl^- concentrations around coordinates (151.4,27.75).

Conclusion

The color-coded Piper plot and corresponding maps provide an intuitive visualization of complex, multidimensional hydrochemical data, without the need for advanced clustering algorithms. The case study in the Condamine Alluvium shows that the colored Piper plot allows visualization of the general trends in the dataset as well as subtle deviations from these trends.

While the analysis of major ions through Piper plots alone cannot be considered a complete investigation of the groundwater chemistry of an aquifer, it certainly serves as an exploratory data analysis tool to identify major patterns and processes in the aquifer and as basis to direct further research, especially when combined with other spatial information such as geology or piezometry.

Acknowledgment

The author would like to thank the Queensland government for providing the dataset, Cui Tao for compiling the dataset, and the constructive comments of the anonymous reviewers.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Appendix S1 Python implementation of the plotting algorithm and dataset for the case study

References

Alfeld, P. 1984. A trivariate Clough-Tocher scheme for tetrahedral data. *Computer Aided Geometric Design* 1, no. 2: 169–181.

Appelo, C., and D. Postma. 2005. *Geochemistry, Groundwater and Pollution*. Amsterdam: AA Balkema Publishers.

CSIRO (2008). Water availability in the Condamine-Balonne. A report to the Australian Government from the CSIRO Murray-Darling Basin Sustainable Yields Project. Canberra, Australia: CSIRO.

Güler, C., M.A. Kurt, M. Alpaslan, and C. Akbulut. 2012. Assessment of the impact of anthropogenic activities on the groundwater hydrology and chemistry in Tarsus coastal plain (Mersin, SE Turkey) using fuzzy clustering, multivariate statistics and GIS techniques. *Journal of Hydrology* 414, no. 1: 435–451. DOI:10.1016/j.jhydrol.2011.11.021

Güler, C., G.D. Thyne, and J.E. McCray. 2002. Evaluation of graphical and multivariate statistical methods for classification of water chemistry data. *Hydrogeology Journal* 10, no. 4: 455–474. DOI:10.1007/s10040-002-0196-6

Hunter, J.D. 2007. Matplotlib: a 2D graphics environment. *IEEE Computer Society* 9, no. 3: 90–95. DOI:10.1109/MCSE.2007.55

Huxley, J.W. 1982. The hydrogeology, hydrology and hydrochemistry of the Condamine River Valley alluvium. Masters by Research thesis, Queensland University of Technology, Brisbane. http://eprints.qut.edu.au/35951/

Kim, K.-H., S.-T. Yun, B.-Y. Choi, G.-T. Chae, Y. Joo, K. Kim, and H.-S. Kim. 2009. Hydrochemical and multivariate statistical interpretations of spatial controls of nitrate concentrations in a shallow alluvial aquifer around oxbow lakes (Osong area, central Korea). *Journal of Contaminant Hydrology* 107, no. 3-4: 114–127. DOI:10.1016/j.jconhyd.2009.04.007

Mertie, J.B. 1964. Transformation of trilinear and quadriplanar coordinates to and from Cartesian coordinates. *The American Mineralogist* 49, no. 4: 926–936.

Nicolet, J. P. and Erdi-Krausz, G. 2003. Guidelines for radioelement mapping using gamma ray spectrometry data. International Atomic Energy Agency Technical Documents (IAEA-TECDOCs) no. 1363.

Oliphant, T. E. 2006. Guide to NumPy. http://www.tramy.us/numpybook.pdf (accessed August 26, 2013).

Peeters, L., F. Baçao, V. Lobo, and A. Dassargues. 2007. Exploratory data analysis and clustering of multivariate spatial hydrogeological data by means of GEO3DSOM, a variant of Kohonen's self-organizing map. *Hydrology and Earth System Sciences* 11: 1309–1321.

Piper, A.M. 1944. A graphic procedure in the geochemical interpretation of water-analyses. *Transactions-American Geophysical Union* 25, no. 6: 914–923

Rus, J. 2010. HSL and HSV models.svg. http://en.wikipedia.org/wiki/HSL_and_HSV (accessed August 26, 2013).

Schwarz, M.W., W.B. Cowan, and J.C. Beatty. 1987. An experimental comparison of RGB, YIQ, LAB, HSV, and opponent color models. *ACM Transactions on Graphics* 6, no. 2: 123–158. DOI:10.1145/31336.31338

Vandenbohede, A., and L. Lebbe. 2012. Groundwater chemistry patterns in the phreatic aquifer of the central Belgian coastal plain. *Applied Geochemistry* 27, no. 1: 22–36. DOI:10.1016/j.apgeochem.2011.08.012

Woocay, A., and J. Walton. 2008. Multivariate analyses of water chemistry: surface and ground water interactions. *Ground Water* 46, no. 3: 437–449. DOI:10.1111/j.1745-6584.2007.00404.x