Motion charts for visualising long-term water quality in South African rivers

MICHAEL SILBERBAUER
Resource Quality Services, Department of Water Affairs, Private Bag X313, Pretoria 0001, South Africa
silberbauerm@dwa.gov.za

BACKGROUND

Low rainfall and runoff, together with rapid urban, mining and industrial expansion during the past 120 years, have caused many South African watercourses to become virtual conduits for waste during periods of low flow (Turton, 2008). The poor quality of water affects downstream users and the aquatic environment. Working within budgetary constraints, water resource managers are compelled to make optimal use of whatever water quality data is available for managing this difficult situation.

QUICK TREND ANALYSIS

The most rapid way for evaluating the information contained in long-term data sets is visualisation, commonly in the form of time-series graphs. Many new and more powerful representations of quantitative data have become available during the past two decades, with the advent of cost-effective computer platforms and a better understanding of visual cognition (e.g. Card *et al.*, 1999; Tufte, 2001; Ware, 2008). The ubiquitous worldwide web has placed free analytical tools such as the R statistics package (R Development Core Team, 2009) and Google Docs (docs.google.com) on the desktop of scientists everywhere where Internet connectivity is available. This contribution illustrates the application of one such tool, the Google Motion Chart or Gapminder Trendalyzer (Rosling, 2007), for investigating ionic ratios of solutes in water samples at long-term monitoring sites in South Africa.

Data preparation involved the extraction of chemical results from the Department of Water Affairs main databases for about 25 sites known to have reasonably consistent records (Van Niekerk *et al.*, 2008), then calculation of annual medians and uploading to a Google Internet spreadsheet (DWA, 2010). Google Docs has various so-called "Gadgets" available for data manipulation, notably the Motion Chart animated bubble-plot visualisation tool. Activating the Motion Chart in the spreadsheet and selecting appropriate variables, chart type and symbolisation yields an X–Y chart with a "Play" button for time-series animation (Fig. 1).

Printed representations obviously limit graphic displays to static images, so this poster consists of snapshots of the recent and historical water quality at selected sites on major rivers across South Africa. For example, when we compare sulphate with chloride, the long-term tendency is for the ratio between the two ions to remain constant for most sites, but to exhibit abnormal trends along the sulphate axis for sites affected by acid mine drainage, varying with episodic dilution capacity. The phenomenon is evident in a site on the Olifants River downstream of a coal mining region (Fig. 1).

KEY LESSONS

The advantages conferred by this approach to data visualisation are the ability to interactively choose variables on the two axes, tag sites of interest, move the time locator backwards and forwards and switch between linear and logarithmic representations. The user can quickly discern any gross patterns in the data set. Difficulties with the tool in its present state are: (1) the need to have a stable and reasonably quick Internet connection, which is not always possible in Africa;

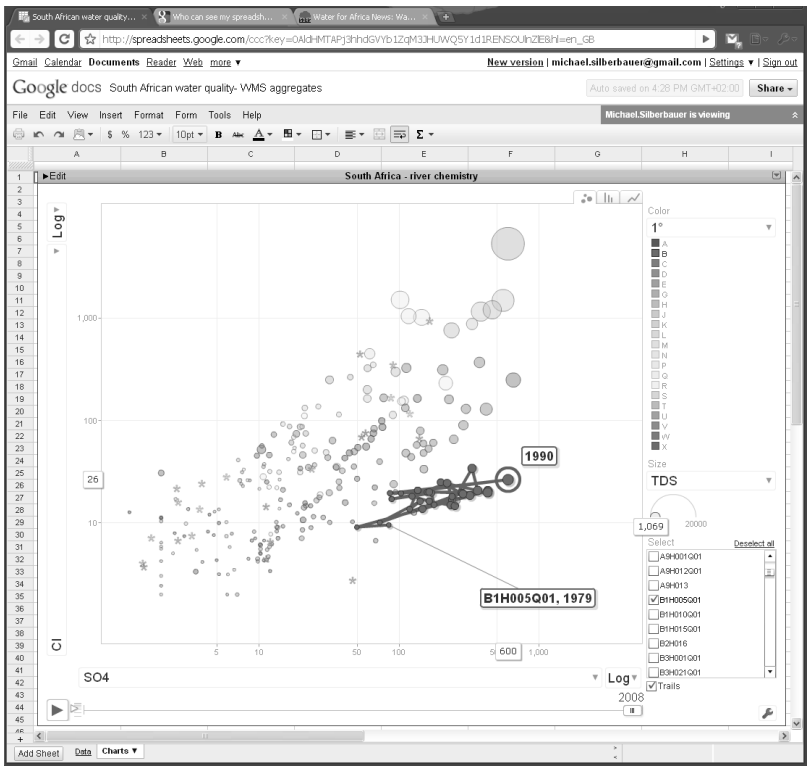

Fig. 1 Snapshot of a Google Motion Chart of chloride against sulphate, highlighting the difference between a site affected by mine drainage (B1H005Q01, where a trail connects the complete time series) against the background of 251 other sites.

(2) the inability to handle really large data sets; and (3) the complexities of customising the tool to introduce more complex visualisations such as the multivariate Maucha diagram.

This type of online, interactive tool is of great use in the visual extraction of information for problem-solving from long-term data sets, particularly in that the user is free of the burden of licence fees or in-house software development. The main drawback for its application is the need for a reliable Internet connection, and this consideration will probably encourage the development of applications useable on other very "thin-client" platforms, notably the cellular telephone network (Loudon, 2007). Certainly, in-depth analysis of problems, once identified, will require the big guns such as catchment modelling. However, simple techniques like those described here will at least give some idea of where to aim.

REFERENCES

Card, S. K., Mackinley, J. D. & Shneiderman, B. (1999) *Readings in Information Visualization: Using Vision to Think*. Academic Press, San Diego, USA.

DWA (2010) Salinity long term Gapminder graphics (http://www.dwaf.gov.za/iwqs/water_quality/NCMP/nwrqsr.asp).

Loudon, M. (2007) Data management and reporting for drinking water quality monitoring in community managed supplies. MSc Thesis, Department of Civil Engineering, University of Cape Town, South Africa.

R Development Core Team (2009) R: A Language and Environment for Statistical Computing. Vienna, Austria (http://www.R-project.org).

Rosling, H. (2007) Video: New insights on poverty and life around the world. (http://www.ted.com/talks/hans_rosling_reveals_new_insights_on_poverty.html).

Tufte, E. R. (2001) *The Visual Display of Quantitative Information*, 2nd edn. Graphics Press, USA.

Turton, A. (2008) Three strategic water quality challenges that decision-makers need to know about and how the CSIR should respond – Keynote address: A clean South Africa. In: *Science Real and Relevant*. CSIR, Pretoria (http://www.environment.co.za/documents/water/KeynoteAddressCSIR2008.pdf).

van Niekerk, H., Silberbauer, M. & Hohls, B. (2008) Monitoring programme revision highlights long-term salinity changes in selected South African rivers and the value of comprehensive long-term data sets. *Environ. Monitoring and Assessment* **154**(1), 401–411 (http://dx.doi.org/10.1007/s10661-008-0407-2).

Ware, C. (2008) *Visual Thinking: for Design*. Morgan Kaufmann, California, USA.