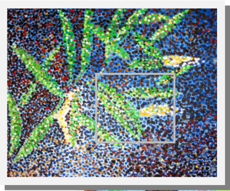
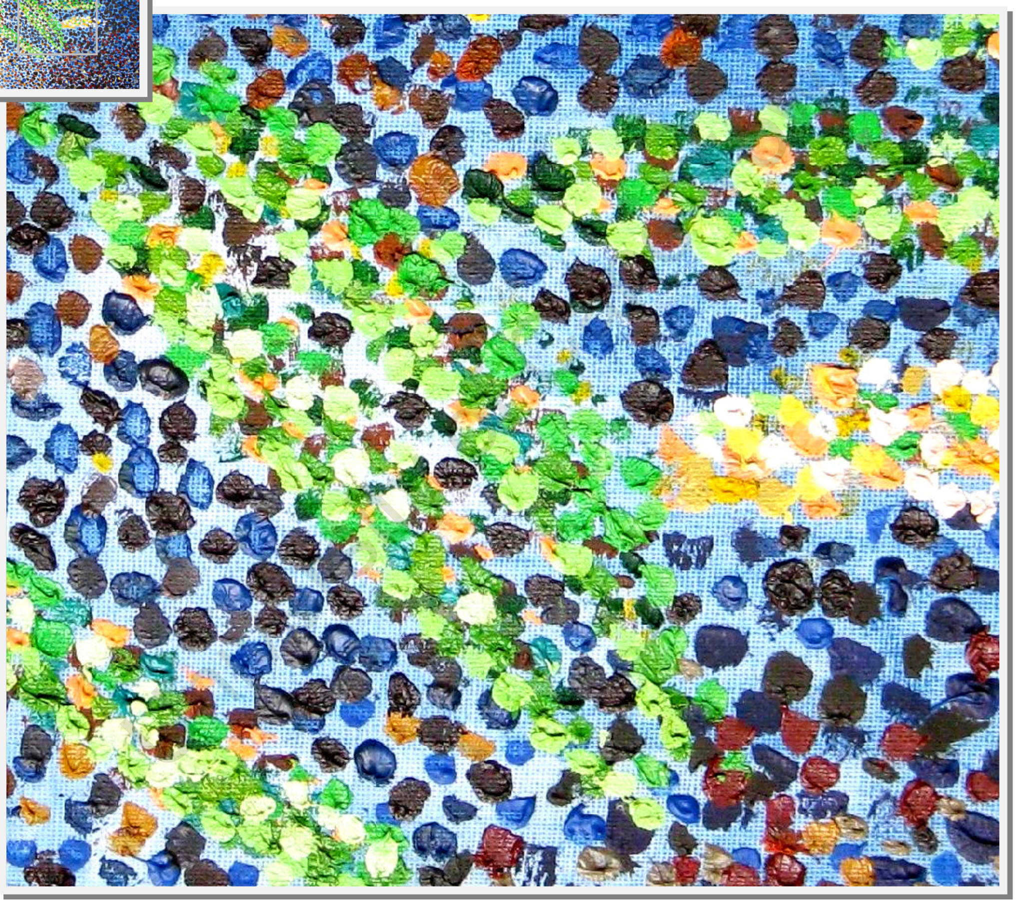
Methods for visualising complex water quality data

Michael John Silberbauer

Thesis presented for the degree of Doctor of Philosophy
in the Department of Zoology
University of Cape Town
August 2009





Cover picture / frontispiece

A graphic should not show only the leaves; it should show the branches as well as the entire tree. The eye can then go from detail to totality and discover at once the general structure and any exceptions to it (Bertin 1983).

Experimental painting of waterblommetjies (*Aponogeton distachyos*) using pointillism to demonstrate the way in which the human eye merges adjacent colour patches into their additive colours (Palmer 1999 p697). When viewed closely, the individual patches of paint are apparent. As the viewer moves away, the patches merge to produce different colours, in the same way as the dots on a colour monitor or a large advertising billboard.

Similarly, we can view water quality data as individual results or merge data from many sources to create a bigger picture.

(Painting: oil on canvas, by the author.)

Declaration

This thesis reports original research written up under the auspices of the Freshwater Research Unit, Department of Zoology, University of Cape Town, from 2006 to 2009. The work reported here took place partly at the Freshwater Research Unit and partly at the Directorate of Resource Quality Services (formerly the Institute for Water Quality Studies and the Hydrological Research Institute), between 1989 and 2009 (formal registration at the University of Cape Town was from 2006 to 2009). This work has not been submitted in part or in whole for a degree at any other university. Data and procedures presented here are the original work of the candidate unless indicated otherwise by a citation or acknowledgement.

Acknowledgements

Generations of technicians and scientists have ensured the integrity and continuity of the Department of Water Affairs and Forestry's water quality database. I am particularly indebted to Bets Davies, Marica Erasmus, Willie Geldenhuys and Kobus Myburgh for their assistance in extracting data and fixing geographical localities (Willie Geldenhuys has enthusiastically taken some of my prototypes developed here and converted them into practical Delphi applications for the national water quality database). I also appreciate the encouragement and time that managers in the Department of Water Affairs and Forestry have provided in compiling this thesis. Mbangi Nepfumbada, Quentin Espey, Axel Diefenbach and Elna Vermaak have been particularly tolerant in this regard. Many people contributed to the technical production of visual material described in Chapter 6, including Loretta Steyn (Initial State of River report and River Health Programme poster production) Tracey MacKay, Department of Education (activity books) and Marietjie Steyn (poster artist). Wilma Strydom and Liesl Hill of the CSIR and Dirk Roux of Monash South Africa also contributed to the development of Chapter 6, as explained in more detail in the introductory remarks to that chapter.

The people of the Freshwater Research Unit have been very helpful in providing advice, encouragement, discussions and office space, sometimes at short notice, on those occasions when I could escape from the government environment to the academic habitat. I am particularly grateful to my supervisor, Professor Jenny Day, for her support when I decided to embark on this process.

I thank the Institute for Geoinformatics (ifgi) of the University of Münster for the opportunity to attend their two-week Spring School for geoinformatics MSc and PhD students in 2008. This course provided new and challenging perspectives on spatial analytics, user interfaces and location-aware devices from Germany, Britain, Brazil and the USA.

Dr Gergely Maucha (Department of Remote Sensing Application in Environment, Budapest) has been most helpful in locating copies of parts of his illustrious grandfather's out-of-print publications.

Leonore Kruger (Palette Art Centre, Pretoria) provided patient and helpful guidance regarding the practice and philosophy of the art of painting.

My family has been very tolerant of my return to academia and I thank them for their support and encouragement (and for making time available on the home computer!). I am particularly grateful to my parents for many things, not least providing a place to stay during my Cape visits.

Contents

CHAPTER 1 - GENERAL INTRODUCTION 1

CHAPTER 2 - COMPACT TIME SERIES WITH THE "BARCODE" SITE INVENTORY 30

CHAPTER 3 - SYMBOLISING MULTIPLE VARIABLES AT POINTS ON A MAP 52

CHAPTER 4 - USING CUSTOMISABLE INTERNET-BASED 2D AND 3D MAPPING SYSTEMS TO PUT WATER QUALITY ON THE MAP 80

CHAPTER 5 - COMBINING CONCENTRATION AND FLOW AS LOADING ON A VISUAL CALCULATION SHEET 99

CHAPTER 6 - COMBINING VISUALISATION AND TEXT INTO A COMPREHENSIVE REPORT 109

CHAPTER 7 - GENERAL DISCUSSION 149

GENERAL CONCLUSIONS AND RECOMMENDATIONS 155

REFERENCES 157

APPENDICES 174

Abbreviations

eWQMS	South African municipal water quality database, administered by Water Services at DWAF
DWAF	Department of Water Affairs and Forestry (cabinet changes on 10 May 2009 resulted in the department losing the Forestry component and reverting to its old name of the Department of Water Affairs. DWA shares a minister with the Department of Environmental Affairs)
GIS	Geographical Information System
KML	Keyhole Markup Language, an extensible markup language (XML) for displaying user data as overlays on Google Earth
KMZ	A compressed ("zipped") KML file

Visualisation terminology

The terms "data visualisation", "information visualisation" and "scientific visualisation" have taken on specific meanings in the jargon of different disciplines (Friendly 2008). In this schema, data visualisation is the graphical display of numerical data, information visualisation is the display of non-numerical data such as document structure and scientific visualisation refers to things like engineering models of structural stresses in dams. These do not appear to be hard-and-fast categories so I have tried to avoid being dogmatic, keeping the terms neutral and using literal meanings that I hope will be clear in context.

Figure size

Many graphics are too large for legible reproduction as figures in the text, and higher-resolution copies are therefore available on the enclosed disk.

Abstract

The quality of South Africa's over-stretched water resources is a matter of concern for all who depend on them for their survival and prosperity, so access to the relevant monitoring data is essential. Visualisation is a powerful method for analysing these data and communicating the results, because it unloads complex cognitive processes from the fairly restricted human numerical processing structures onto the highly developed visual perception system. Developments in the field of visualisation during the past two decades have yielded many practical methods that are applicable to the analysis and presentation of water quality data. Judicious use of visualisation aids aquatic scientists, water resource managers and ordinary consumers in assessing the quality of their water and deciding on remedial measures. To provide some insight into the possibilities of visualisation techniques, I analyse and discuss five visual methods that I have developed or contributed to: multivariate time-series inventory plots; multivariate map symbols; spatially-referenced inventory of water quality data; mass transfer summary plots; and the use of visual methods in communicating the ecological status of rivers to a wide audience.

The first method, multivariate time-series plotting, provides the user, at a glance, sampling frequency, water quality statistics, regional locality and site description, all condensed onto a single page. The next type of visualisation comprises symbols for showing multiple variables in figures and on maps, focussing on the class of radial symbols with slices of variable radius and constant angle, which are especially useful for instant comparison between sites. A 3D public spatial inventory of water quality data takes advantage of the intuitive digital globe interface provided by Google Earth to integrate spatially-referenced access to monitoring sites with water quality plots, data and other information for several thousand surface and groundwater sites. More prosaic and pedestrian but nevertheless useful is a visual sifting tool to assess the availability of flow and water quality data for the calculation of mass transfer, presenting the information in time series plots and log-log regressions. Finally, a combination of visualisation methods supports the state of rivers reporting procedure, communicating the information from a national biomonitoring

programme to a wide audience, where sensitivity to cultural filters, background knowledge and reading ability is as important as an understanding of the neurophysical basis of visual perception.

Improvements in the cost-effectiveness of computer hardware and software have brought complex visualisation systems, formerly the privilege of a few supercomputer specialists, to millions of computer users. Scientists using visualisation techniques can improve their own understanding of water quality data and communicate their findings to others. Despite these advances, many South Africans do not have access to the Internet, and further democratisation of information about water may occur through more pervasive technology, such as the cellular telephone network. A difficulty that remains is our understanding of the human aspects of effective visual communication of water quality status, which has not kept pace with the remarkable advances in the technology, remaining largely intuitive and providing many opportunities for further research.

Chapter 1 - General Introduction

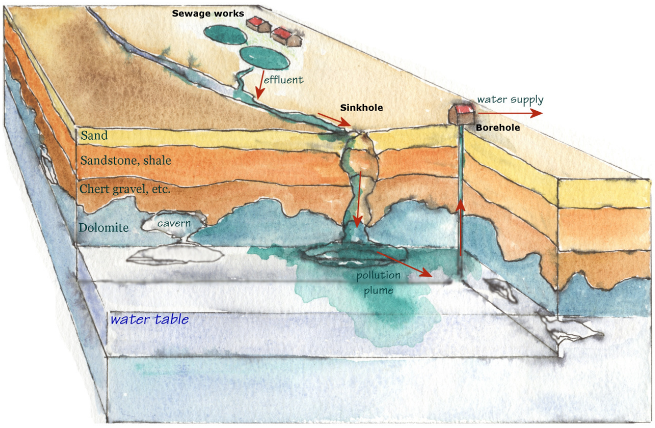


Figure 1-1. Diagram of the potential for groundwater contamination in Karst landscapes (please see Appendix 1 for a larger image and a more detailed explanation).

In 2005, an outbreak of typhoid in the small Mpumalanga town of Delmas drew attention to problems with the municipal water supply (Sidley 2005, Griesel *et al.* 2006, Pienaar and Xu 2007). Several people died and closer investigation revealed that specialists had warned of shortcomings in the water supply system as early as 1993. The municipality had focussed on other priorities in the intervening decade, and the warnings went unheeded, eventually fading from the collective memory. In short, water treatment services were inadequate because the municipality did not allocate sufficient funds (DWAF 2005, Parliamentary Monitoring Group 2007).

The deaths in 2005 prompted an intensive investigation by the South African government departments of Water Affairs and Forestry, Agriculture, Health and Social Services, Local Government and Housing and the municipality of Delmas, bringing to light a range of hazards in the water treatment, supply and disposal network (Figure 1-1). At that time, I made a schematic diagram based on the main findings of the report overlaid onto a perspective image from the newly available Google Earth 3D globe viewer (Figure 1-2). From this perhaps oversimplified elevated viewpoint, and with hindsight, the problem seems to have been an obvious disaster waiting to happen. The purpose of this thesis is to explore the potential of these kinds of visual methods for data analysis, summary and presentation to influence and facilitate decision-making regarding water resource management. It will include a review of a range of applications that I have developed before and during this study, with the aim of determining whether we can construct visualisations of water quality hazards *before* a crisis occurs, in order to effectively alert local, regional and central government officials to problem areas.

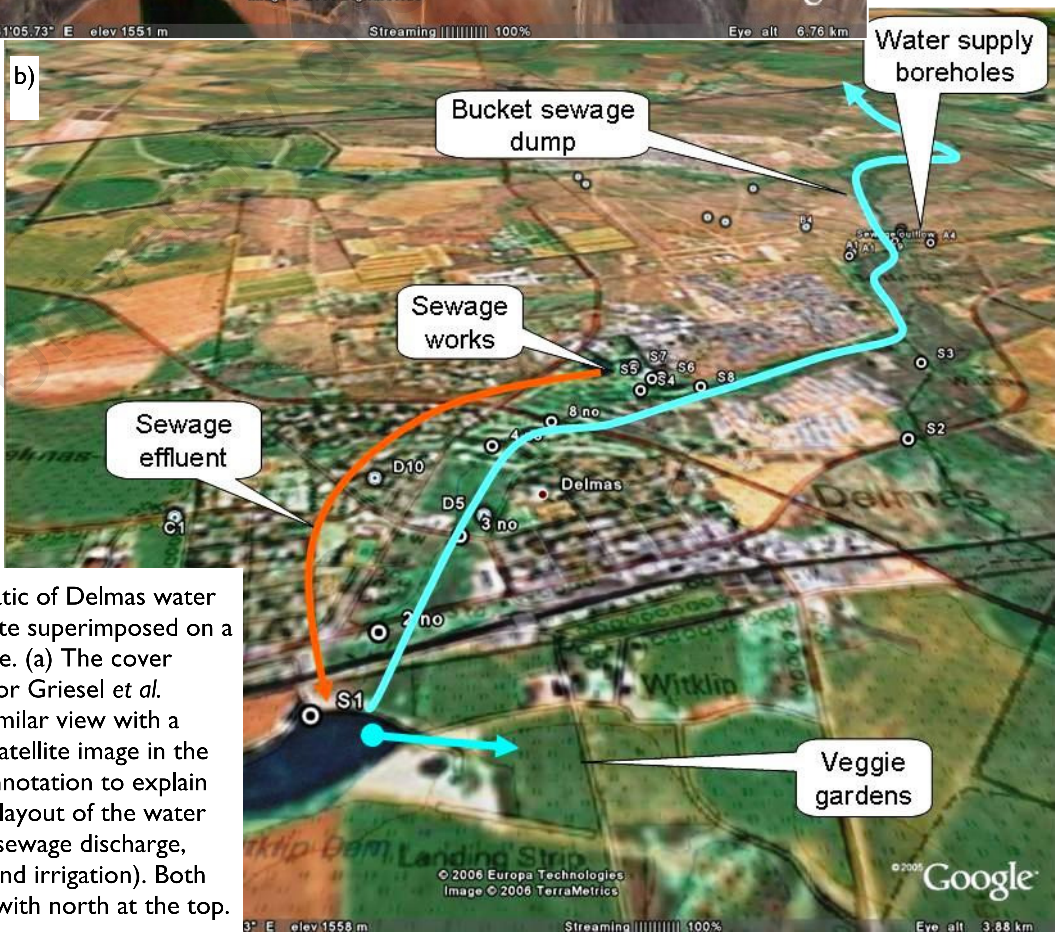


Figure 1-2. Schematic of Delmas water contamination route superimposed on a Google Earth image. (a) The cover picture prepared for Griesel *et al.* (2006) and (b) a similar view with a lower-resolution satellite image in the background and annotation to explain the hazards in the layout of the water supply (red line = sewage discharge, blue lines = river and irrigation). Both are oblique views with north at the top.

This discussion will thus have to reflect on the methods and results of past projects, while providing suggestions on how to take them forward. A prime objective is the drawing together, documentation and improvement of a number of potentially useful approaches and applications for visualising data so that others can build on these methods and, ideally, develop new and better analytical systems for creating visualisations. The review of these applications has been an opportunity to think carefully about the purpose of water quality analysis and the ways in which we report it. It also exposes what have been largely isolated, internal projects to academic (rather than managerial) scrutiny, an essential part of their development into reliable tools for decision-making. Additionally, working on a project like this gives the civil servant the opportunity return to the academic environment now and again, valuable experience in light of the government's requirement for its senior scientists to take part in what public service managers optimistically term "skills transfer". This experience is also helpful in view of the increased isolation of government scientists in small pockets in the public service, where little opportunity exists for peer review. While this closed environment encourages innovation in order to solve immediate problems in applied science, it can also lead to wasted effort because of a lack of collaboration with experts elsewhere. However, I have been privileged to present the work in Chapter 3 at an international conference in Cape Town, as referenced in the PhD, and plan to publish information from other chapters. During the past 5 years, I have had informal discussions with staff at the Global Runoff Data Centre in Germany and the GEMS water quality data centre in Canada. I have also had informal e-mail exchanges with people from around the World who have shown an interest in the Google Earth web page.

In the rest of this introduction, I will elaborate on the potential of visualisation to assist in the areas of human health, water quality and information management. I will also describe the supporting function of visualisation in previous studies and mention some aspects of physiology, culture and technology to be aware of when using visual representations.

1.1 Public water, public health

'Sir, I'll make it as plain as Peter Pasley's pike-staff--I will allow that ilk parochine, on an average, employs fifty pleughs, whilk is a great proportion in sic miserable soil as thae creatures hae to labour, and that there may be pasture eneugh for pleugh-horses, and owsen, and forty or fifty cows; now to take care o' the pleughs and cattle, we'se allow seventy-five families of six lives in ilk family, and we'se add fifty mair to make even numbers, and ye have five hundred souls, the tae half o' the population, employed and maintained in a sort o' fashion, wi' some chance of sour milk and crowdie; but I wad be glad to ken what the other five hundred are to do?'
'In the name of God!' said I, 'what *do* they do, Mr. Jarvie? It makes me shudder to think of their situation.'

--Sir Walter Scott, Rob Roy, (1817: 306)

The quotation above refers to Scotland in the early 1700s, experiencing what one would now call third-world subsistence conditions with inadequate resources, communicated in an almost incomprehensible dialect of English. This is a reminder that the developed world was not always developed, and we should take note of how they got where they are (the cynic would suggest, partly by ingenuity and partly by exporting their problems to the rest of the planet).

Life expectancy is an indicator of the presence of the essential public health services: safe water, no malaria, adequate immunisation and a reasonable diet. So, a Japanese girl born in 2004 can expect to survive until 2090 while her Zimbabwean counterpart will most likely not still be alive in 2039 (Garrett 2007). Safe water is a service that South African authorities, often rightly, pride themselves on delivering, yet this a deceptively difficult task in a dry climate with very little dilution capacity in its rivers. When runoff is low the relative contribution of effluent to streamflow is proportionally greater. Even a wealthy and well-watered country like Canada has buckled under the pressure and experienced Delmas-style incidents, for similar reasons of negligence, poor training and diversion of funding—in short, taking clean water for granted (Leveson 2002).

1.2 Running the show

I have little hope of any form of agreement in discussions. After listening to the arguments put forward during the last 2 days I feel more like entering a lunatic asylum or a nursing home than continuing with my present job. I am *absolutely* disgusted with the politicians' methods of waging a war!! Why will they imagine that they are experts at a job they know nothing about! It is lamentable to listen to them!

General Alan Brooke, later Field Marshal Lord Alanbrooke,
in Danchev and Todman (2002*:* 485), discussing Churchill and others.

Management of water quality in an arid environment would be complex and difficult enough for a wealthy and well-equipped government. In a state of extended transition, many intermediate management posts are vacant for extended periods or occupied by overworked acting officials, so information moves neither freely nor accurately up and down the hierarchy. Executive orders and actions do not take place. This is by no means to suggest that incumbents are in general insincere in their efforts, they are just too far removed from the coalface. Most people, worldwide, now grow up in an urban environment remote from the processes of primary production, acquiring their water and agricultural products with little insight into their origin (Grandin 2006: Ch 2). Likewise, in many business cultures, people no longer start at the bottom of an organisation and then spend a lifetime working their way up the ranks, obtaining valuable experience on the way. In a society in transition, even those with long years of experience find that their career ladder has taken unexpected twists on the way up, sometimes leading them to places where their accumulated knowledge is of little application.

South African legislators have made an effort to ameliorate the devastating effects of past inequalities by updating and revamping laws so that they create a legal framework that is impartial and equitable. In the National Water Act (Act 36 of 1998), this approach has resulted in a respectable structure that provides the framework for good governance and sustainable development. Underpinning the new approach to water resource legislation are almost utopian requirements for efficient management structures and comprehensive dissemination of information.

Laws make themselves felt when applied, not when enacted, though. An unfortunate by-product of any rapid change to a new political administration is a disruption in the skills transfer and mentoring processes that were inherent in a stable civil service. A new regime understandably feels the need to stamp its authority on government and this can result in an over-hasty replacement of staff. South Africa has fortunately avoided the completely disastrous revolutionary approach, so dear to humankind, which summarily purges the entire administrative structure. In Iraq, for example, this tragic approach to water management saw damage to infrastructure, loss of skills and a drop in water treatment capacity from 3 million to 1.3 million m^3/d between 2002 and 2006 (O'Hanlon and Campbell 2007). That said, South Africa has battled to extend the supply of good-quality water to the whole population, especially in rural areas, and recent reports confirm that the general state of water treatment systems is deteriorating (Cloete 2010, DWA 2010).

A complicating feature of water quality administration is that it exists in two separate compartments, resource water and supply water. Resource water is the raw substance circulating in the hydrological system, while supply water is the final treated product delivered to the consumer. Resource water is under central control and supply water is a local government responsibility. Local authorities are therefore often unwittingly the actual agents of resource management and require comprehensive and comprehensible information systems. The realisation that they cannot act without knowledge and guidance is reflected in the effort that has gone into the development of the user-friendly South African municipal water quality database eWQMS, which underlies the Blue Drop evaluation system for evaluating water works (de Souza *et al.* 2008). Staff changes and skills shortages have meant that municipalities are often not getting the right picture, particularly in areas geographically or culturally remote from research institutions and universities. Enterprises such as eWQMS improve the speed of analysis and utility of data presentation thus encouraging data capture and even (dare one hope?) data use and interpretation.

Municipal managers are now equipping themselves to ask the right questions about water supply. Who are the water users at greatest risk? Probably those who drink untreated water, are exposed to contaminated water or use contaminated water to produce food. Where are undesirable solutes, suspensoids and bacteria entering the water? Maybe from the municipal wastewater treatment works themselves. Where is water quality a threat to industry and agriculture—and consequently to the local economy? Even well-established municipal managers, to whom these are obvious questions, cannot afford to be complacent in the presence of aging and overloaded water treatment facilities.

Throughout all this, scientists need to keep a level head and achieve constructive communication of water quality results to the right audience. The Strategic Framework for National Water Resource Quality Monitoring Programmes (Grobler and Ntsaba 2004) notes the importance of accessible information systems and visualisation tools. At the Directorate of Resource Quality Services in the Department of Water Affairs and Forestry (RQS in DWAF), where much of the work described or proposed here has application, our task is to observe water quality variables in the field and laboratory, and then interpret their meaning to managers, scientists and other interested parties. This is an advisory role in that RQS does not usually intervene in the executive process, though RQS scientists need to be able to convey the urgency of important water quality findings to their senior managers. This is not always easy and specialists should be careful of clumsily trying to intervene at too high a level, unless it is their brief to provide reports at that level. The interaction between civil servant and politician is an art requiring consummate skill, as General Brooke observed in his dealings with Churchill (quotation from his diary at the head of this section). The cartoons in Appendix 2 illustrate two incidents during 2008 where South African environmental scientists, passionate about their topic, became so impatient with politicians that they could no longer present facts objectively and unemotionally. Individuals without a rare blend of skills and support usually come off second best when they take on government, corporations, organised crime—or even their peers (Berns 2008). While career martyrdom is an effective method of communication, it is a once-off event and only those with other resources to fall back on should contemplate it. A subtle approach is more likely to be sustainable.

1.3 Data collection and data management

Data collection occurs at various places in the administrative hierarchy. The Directorate of Resource Quality Services in the Department of Water Affairs and Forestry runs several national monitoring programmes that collect long-term data on surface water (some data come from other agencies such as water boards). The department's Hydrological Services directorate collects flow data at gauging weirs and groundwater quality samples from boreholes. The Water Services Regulation directorate coordinates the management and distribution of data on municipal treatment works.

The Department of Water Affairs and Forestry's Water Management System database is the repository for the RQS national programmes' water quality data, with the database design also serving as a faint echo of past decisions and operation rules. It is a complex, functional system, comprising a stable database and a set of reports and views. That said, it has fallen short of the expectations of many users (particularly regarding usability and response time). Where it has been most lacking is in providing a cost-effective data capture process for users. Part of the problem is the lag time inherent in the whole data gathering and analysis system, often creating a delay of several months between users collecting their samples and seeing the actual status of the resource at it was at the time of monitoring. This sluggishness of response is one reason why Water Services Regulation stores municipal data on the separate, leaner and more streamlined system called eWQMS (de Souza *et al.* 2008). For snap decisions, quick information with wide error margins is more help than a very accurate result several months later, although the latter may be preferred for long-term planning or legal purposes.

We have seen the importance information on water quality in a management structure. We now look at how individuals in that structure receive and interpret visual information.

1.4 Visualisation theory

"…we've been thinking about consciousness in the wrong way—as something that happens in us, like digestion—when we should be thinking about it as something we do, as a kind of living activity."

--Noë (2009) *Out of Our Heads: Why You Are Not Your Brain, and Other Lessons from the Biology of Consciousness.*

1.4.1 Perception

Two hefty publications dealing with the visual system (Palmer 2000) and information visualisation (Card et al 1999) stand as markers of the state of knowledge in these closely related fields at the end of the twentieth century. The insights of scientists with an interest in art have built on this perception of how we perceive (Tufte 2001, Ware 2004, Lanthony 2006) and a science is emerging, attuned to the human visualisation system, for designing graphics in general (Ware 2008) and maps in particular (MacEachren 2004, Slocum *et al.* 2008).

The ascendancy of the digital camera towards the end of the 1990s has made the idea of the eye as a sensor array a popular metaphor. Yet the visual system is far more complex than a pair of digital cameras sending raster images to some kind of "bioscope in the brain" (Figure 1-3). While the retina does indeed contain millions of tiny light-sensitive cells, some responding to long, medium and short wave electromagnetic radiation focussed on them by a flexible lens, they are not arranged in a regular rectangular array wired to a pixel-based image processor. Only a small area around the fovea has high enough resolution to detect fine structures, and our visual system assembles our world view in a controlled way by flicking our eyes from one point of interest to the next. Groups of cone cells sensitive to different wavelength bands work in concert with associated nerve cells to convey primitive forms to the visual system in the brain: the first stage of visual processing is automatic, where we simply know that certain objects are spots or lines (Figure 1-4).

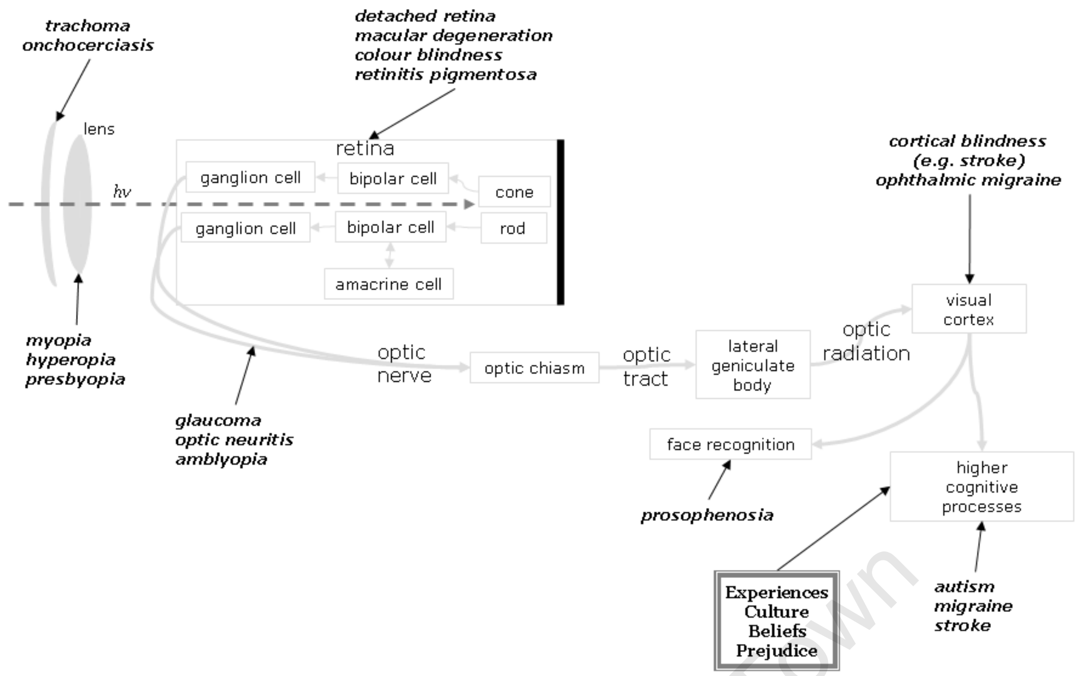


Figure 1-3. The components of the human visual system (grey boxes), information transfer (grey arrows) and various pathologies that affect the ability to see (bold italics), from the arrival of a photon (*hv*) to the triggering of an idea. The box outlined in grey suggests the level at which individual interpretation influences what a person sees. Physiological pathways compiled using information from Guyton and Hall (2000: Ch 49-51).

Further processing stages allow us to see shapes and patterns, recognise faces, turn letters into words, detect movement, predict trajectories and build a view of the world out there (Palmer 1999, Ware 2008). This is about where the theory ends: "…only one proposition about how the brain makes us conscious—how it gives rise to sensation, feeling, subjectivity—has emerged unchallenged: we don't have a clue" (Noë 2009).

The mechanics of vision are complex and at each stage of the visual system, things can go wrong. While a full description of how the human visual system works is beyond the requirements for information visualisation, we at least need to be aware of commonalities, limitations and morbidity in the visual system when preparing information for visualisation (Figure 1-3). Fortunately, amongst all the vision pathologies mentioned the most common conditions are optical (myopia, hyperopia and presbyopia, correctable with optometry) and red-green colour blindness in 8% of males (compensated for by careful choice of colour-coding) (Guyton and Hall 2000). Others are preventable by hygiene (trachoma), timely medical intervention (glaucoma, detached retina) or lifestyle changes (amblyopia). In South Africa, up to

80% of blindness, defined as visual acuity less than 3/60 (10/200 in Imperial units), is preventable or reversible (DoH 2002).

At the highest level of visual interpretation, our experience, beliefs, culture and prejudices define what we actually see when information arrives along the visual pathways (Figure 1-3). This depiction of how our minds work is naturally limited by a paradox, the "brains-in-vats" contemplating a "brain-in-a-vat", or at all events, contemplating a brain in a nuclear magnetic resonance field. Although the workings of the visual system are now understood in extraordinary detail—in terms of physics, neurophysiology and genetics—scientists and philosophers are far from understanding how we actually use this information to perceive (Noë and Thompson 2002, Noë 2009).

Returning to the mechanics of vision, during the first stages of visual perception, some reactions to light are instinctive. Focussing, pupil dilation and direction of gaze towards a visual stimulus are automatic and continue to function even in a cortically blind person. At the next level of visual processing, the primary visual cortex, we interpret certain basic visual properties of objects without needing to think: they "pop out" of their surroundings in less than a tenth of a second (Figure 1-4). The main features that pop out in this way are colour, angle, size, movement and stereoscopic depth (Ware 2008*:* 29).

Figure 1-4. The pop-out effects of angle ("quick"), colour ("brown"), size ("dog") and movement ("over", which regrettably doesn't move in print). The stereoscopic effect is represented here by the blurred word "cat" in the background. Information from Ware (2008:Ch2). The practical application of this is evident if you compare how long it takes to choose the oblique bar in the row with the time needed to choose the 12th bar.

1.4.2 Communicating without sight

I who cannot see find hundreds of things to interest me through mere touch.
--Helen Keller (Three Days to See, 1933).

When the information recipient has little or no visual ability, one must seek out other means of communication. Blind people with intact sense of hearing can use text-to-speech converters for descriptive information, but without tactile material must rely on sighted interpreters to explain graphics verbally. A Braille atlas is now available for South Africa (⠊ ⠁⠋⠗⠊⠅⠁) (Bowerman 2009), so those with no sight at all can in principle experience thematic cartography by touch. Braille publication is unfortunately very expensive because the small market does not attract advertising support (Blind SA 2009). Blind people make up about 0.75% of South Africa's population (DoH 2002).

Whatever our visual acuity, non-verbal sound can also communicate basic information: We are familiar with changes in tone and volume conveyed by lift bells and emergency vehicle sirens, while complex music conveys emotional cues, as exploited so successfully by the motion picture industry to accentuate moving images. Cellular telephone users use different ringtones to identify their devices in the rainforest cacophony of our urban environment.

While the sense of smell is less spatially oriented than sight, using smell in combination with maps ("scented cybercartography") opens up new possibilities of non-verbal communication of attribute information through association with previous olfactory-related memories (Lauriault and Lindgaard 2006). However, the concept of, say, a map of sewage works with their efficiency classified according to intensity of stench, while no doubt bringing home the point, may be a tough one to sell.

The use of multiple senses for communicating information is technologically challenging and the present cost would outweigh the benefits for scientific applications, though the entertainment business has made use of multisensory presentation for many years, for example in Disneyworld. One could imagine a complete immersive environment at some time in the future that would permit the water quality manager to stride around a virtual reality room, feeling the warmth of the Kalahari underfoot or stepping carefully to avoid the cold and spiky Drakensberg, while brightly-coloured markers indicating guideline violations bob up and down at their sites and squeak for attention.

Importantly, we should observe that Helen Keller's remark about the sense of touch at the head of this section was part of an essay lamenting the loss of vision and the need to appreciate this most important window onto the environment. Figure 1-3 lists many preventable routes to vision loss, involving water, hygiene and lifestyle choices. A public health priority must always be the avoidance of preventable blindness and management of water related diseases is an important part of this.

1.4.3 Culture and perception

Remember, with great power, comes great responsibility.

—Peter Parker's Uncle Ben (Movie: Spiderman).

We are social beings, embedded in our particular society and culture, which define the way we think about things, illustrate them and write about them (Hauser 1951), perhaps more so than we normally care to imagine. At one cultural extreme, slick urbanites expect their information pre-packaged in sound bites or shrink-wrapped PowerPoint presentations. At another point on the cultural spectrum, their rural cousins may find this attitude towards obtaining information abrupt to the point of rudeness, and prefer a more thorough process of first establishing the credentials of the messenger, seeking consensus on the message and its context and then filing it in the shared body of knowledge. Technology specialists who have grown up in the big city find to their cost that they cannot simply install a wonderful computer software solution and hope that it will take root and flourish. Often the technology disrupts the existing hierarchy in ways that make it unacceptable, and follow-up research may not reveal what went wrong because the intended users of the technology cannot bear to disappoint the researcher, so paint a respectful but misleading picture of success (Blake and Tucker 2006).

Visual communication is powerful and direct, at the most basic level triggering a defensive reaction to a perceived threat (detecting a large spotted cat in your peripheral vision triggers the response to run for your life). Visual stimuli are so direct that we react to a political symbol before we are consciously aware of noticing it (Hassin *et al.* 2007). Other parts of our visual system synthesise a picture of how the world will be a tenth of a second in the future, to compensate for the delay between a burst of photons striking the retinal pigments and a coherent picture emerging (Changizi 2008). Symbolism therefore needs to be striking while avoiding offence, balancing politeness against paranoid political correctness, and complicated by the ability of our visual system to pick out non-existent objects in such random patterns as clouds. We have to ensure that users of our information systems are as far as possible focussed on the message, not the messenger. Depending on their visual priming,, some people may take offence at the use of red and blue to denote female and male data on a graph, others may notice the cartographic symbol instead of the information when the map is sprinkled with ✞s, ☠s, ☪s, ✡s or ☹s (vertical cross, skull and crossbones, crescent moon & star, six-pointed star, grumpy face).

1.4.4 Visualising water data

"In general, if a statistical investigation of this kind is well planned and the data properly collected the interpretation will pretty well take care of itself. So-called 'high-powered,' 'refined,' or 'elaborate' statistical techniques are generally called for when the data are crude and inadequate—exactly the opposite, if I may be permitted an obiter dictum, of what crude and inadequate statisticians usually think..."

--Wallis (1949)

"Data Dumping, or the Putting Together of Statistical Compilations with a Shovel. What's important is whatever has been collected, for whatever reason, if any."

--Tufte (1977).

Sadly, these two quotations describe, perhaps a little too bluntly for comfort, some aspects of the data sets that I will be dealing with in this thesis. The planning for many water quality monitoring programmes occurred long ago, so scientists have to make the best of the available data and try to get as much information out of them as possible (but not more). One hope is that the results may help in planning and directing monitoring to obtain the most relevant data in future.

The famous map by Dr John Snow of the deaths from a cholera outbreak in London marks the beginning of modern spatial analysis of water quality (Figure 1-5, Tufte 1997, Kraak 2006). Snow plotted the number of deaths at sites on a map and deduced that a particular water source was causing users to become ill and die.

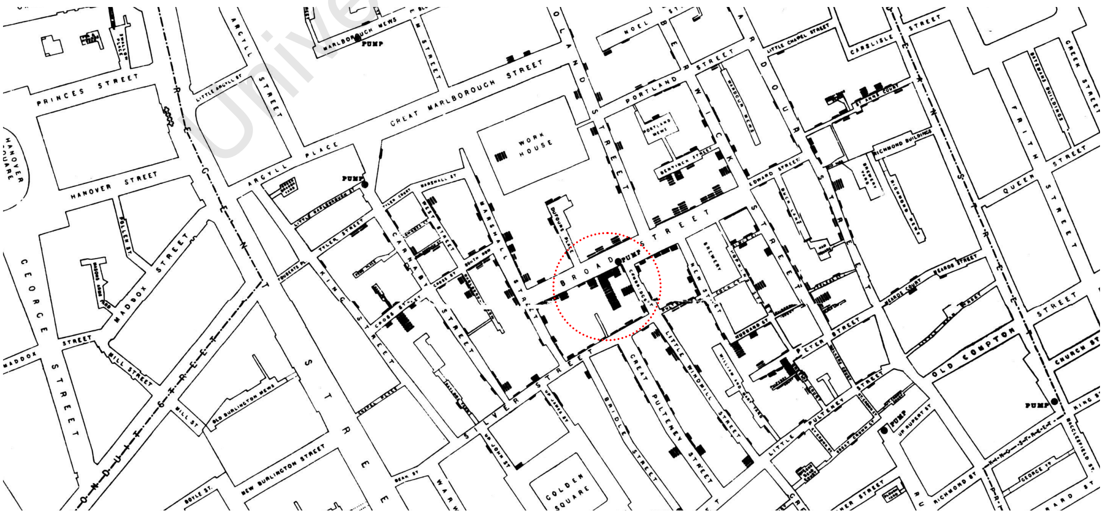
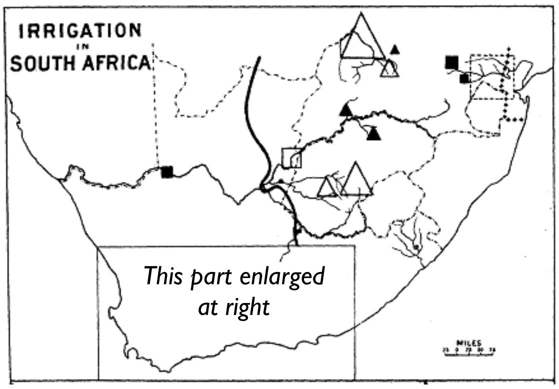


Figure 1-5. Detail of Dr John Snow's map of the Broad Street cholera outbreak in London, 1854 (Tufte 1997). The sinister black bars represent the number of deaths at each address, the largest tally being near the source of contaminated water, circled in the centre of the map.

Early examples of the use of visualisation, especially maps, for the presentation of South African water resource information, are available from at least the 1920s, for example, Forde 1925 (Figure 1-6); a 1926 hydrological chart of South Africa by the Irrigation Department (Figure 1-7); and a map of fluoride for a dental caries study (Ockerse 1946 and Figure 1-8). Map production was expensive and tedious though, so hydrologists preferred to publish their results in tabular form in large printed reports.



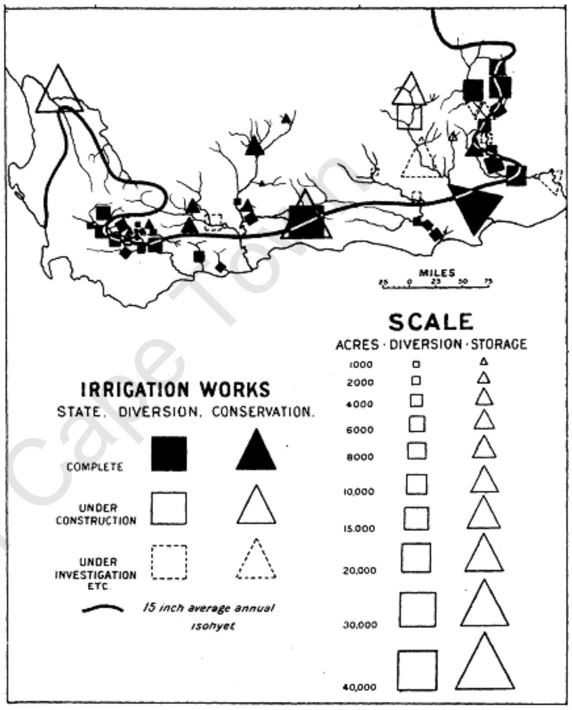


Figure 1-6. Irrigation works in South Africa in the 1920s (Forde 1925).

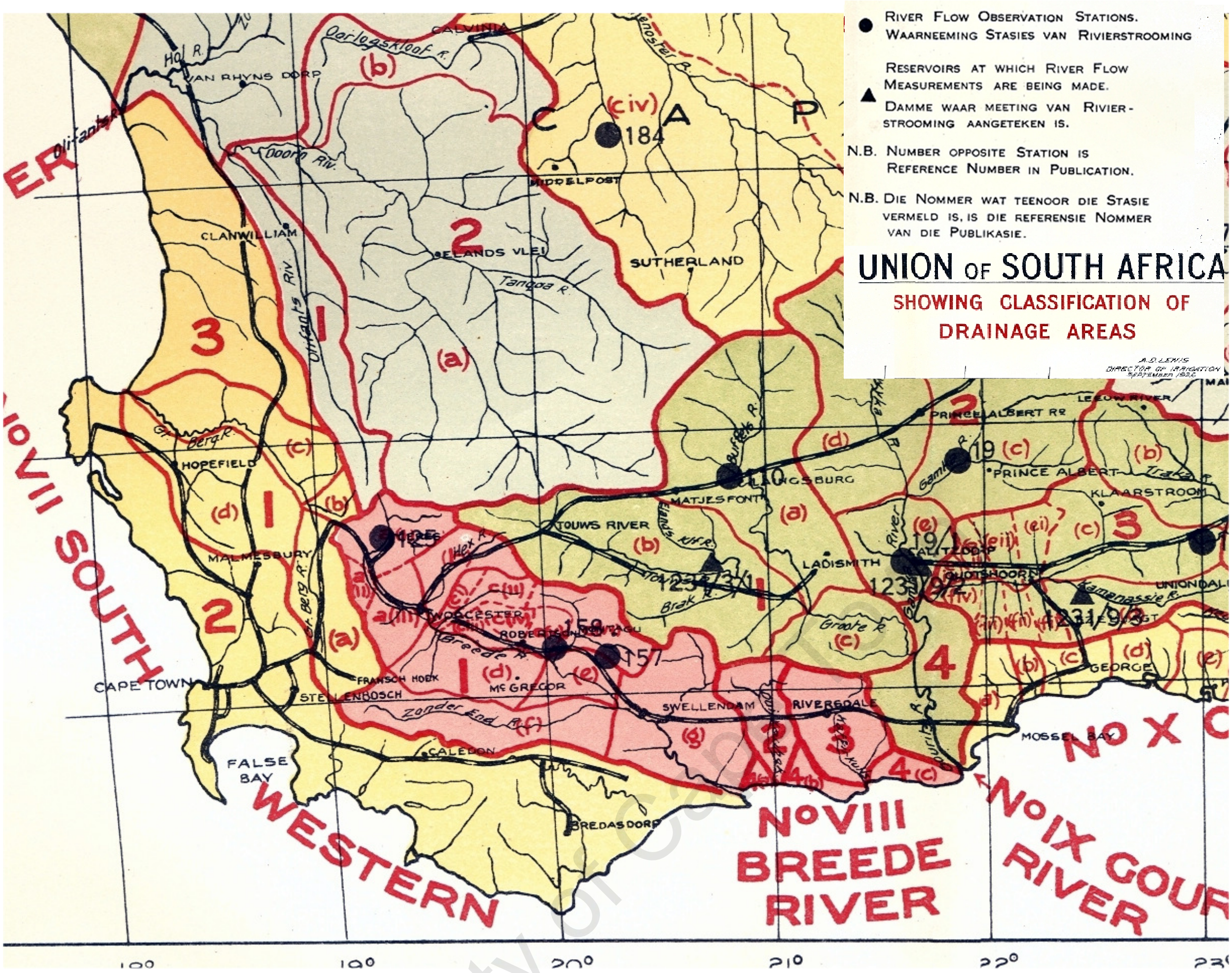


Figure 1-7. Early hydrological map of South Africa (detail of southwestern Cape Province).

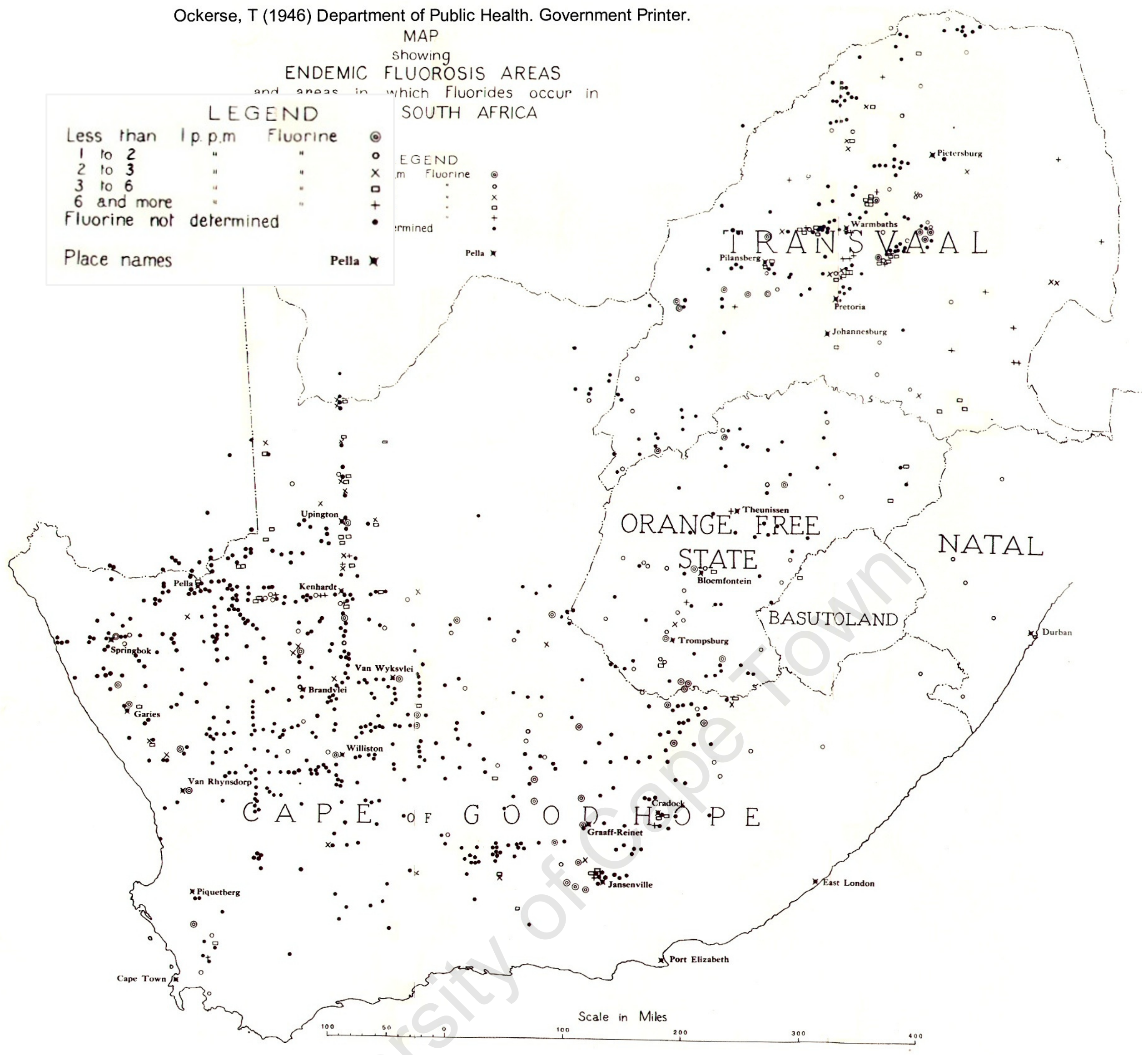


Figure 1-8. Fluoride data for South Africa in the 1940s (Ockerse 1946).

The wider availability of computer systems during the 1980s and onwards made the display of time-series data more straightforward (one method, "Barcode", is covered in detail in Chapter 2). Spatial data representation is a much greater challenge, as reflected in the higher operating cost of geographical information systems. Even when the software licence is free, users need time and training to master the system, understand the data, learn how to present them and perform successful spatial analysis.

As regards South African water quality data, the first serious departure from purely tabular listing was the use of pie symbols to represent the components of water quality at sites on a map of the middle Vaal catchment (Figure 1-9, van Vliet and Nell 1986). This involved semi-automated calculation of symbol sizes and angles, followed by the actual drafting task, performed by highly qualified drawing-office personnel. The drafting team used painstaking manual methods, including cutting shaded transfer sheets into sectors with a craft knife and pasting them in position. Within a decade, graphics software and pen plotters had become sufficiently advanced that computer output could be used directly in reports and posters. Vector output devices then gave way to machinery based on raster technology, such as ink-jet and electrostatic printers. "Cutting and pasting" now happens with no fuss and no mess.

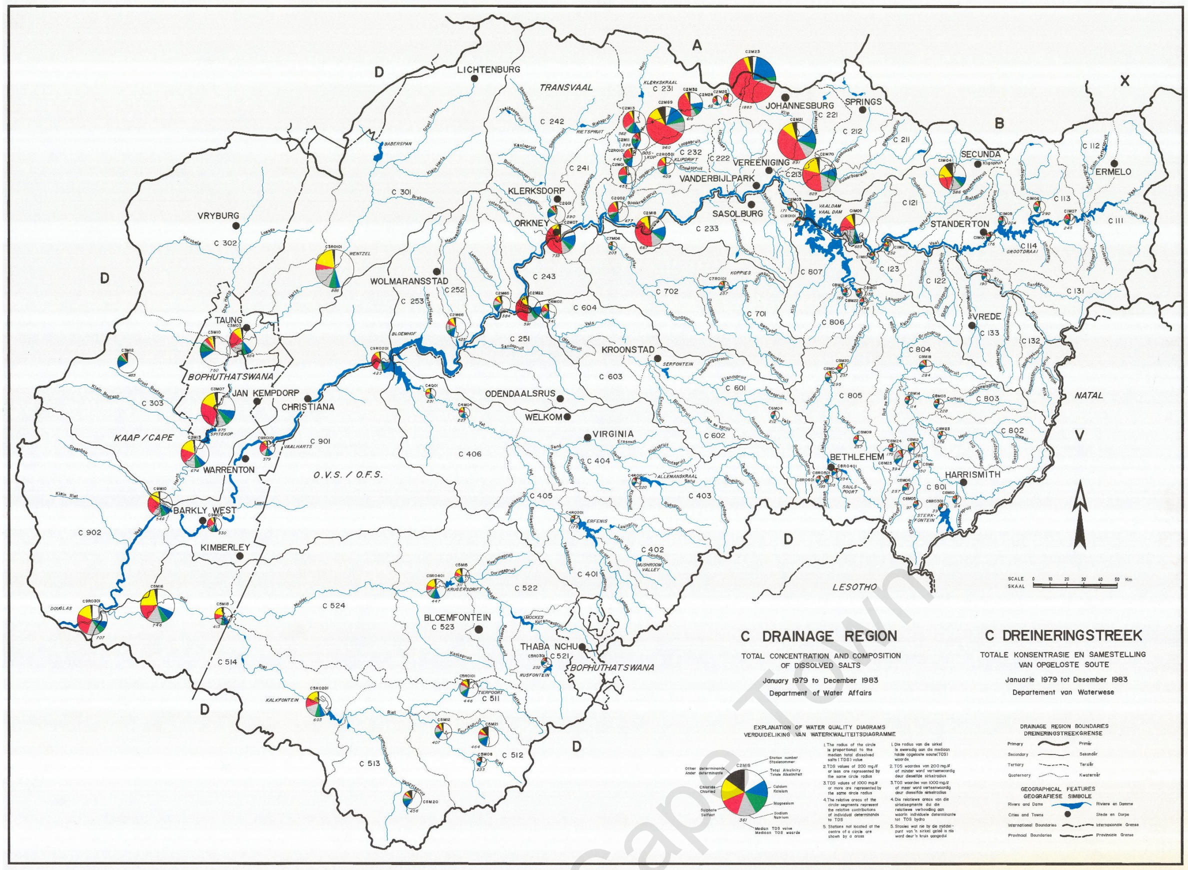


Figure 1-9. Hand-drawn water quality map of the Vaal River catchment showing chemical composition at monitoring sites as pie charts. Reproduced from van Vliet and Nell (1986).

For normal time-series and scatterplot graph plotting, scientific data visualisation has been the poor cousin of business-oriented graphics such as Excel and PowerPoint (Tufte 2003), which favour extraordinary default graph formats such as blue lines on a grey background. Point-and-click scientific graphing has remained an expensive niche market, although applications like the free R-statistics package (R Development Core Team 2008) have become very powerful for those prepared to invest the effort to learn them. Development of in-house graphics is a costly option. The Water Management System (www.dwaf.gov.za/iwqs/wms) is one example, with many graphing options for water quality data, and a separate map-plotting procedure, but needing a team of programmers for development and maintenance. The ambitious WaterMarque system (Cobban and Silberbauer 1993, Cobban 1994) is another example, consisting of a complex system of graphical user interfaces in Unix ArcInfo for viewing spatial data and methods for plotting water quality symbols and time-series graphs on a map (Figure 1-10).

Figure 1-10. The WaterMarque system, an ambitious graphical user interface to South African water quality data, built in the mid-1990s.

Representing the time dimension adds an interesting challenge to spatial data representation, and the oceanographic disciplines appear to have attracted more funding for this than the limnological ones. For example, Boyer *et al.* (2000) developed a very useful set of visualisation tools that included animation of the movement of phosphorus and chlorophyll *a* in tidal Florida Bay—however this was a once-off application for a specific, delimited study area.

Maps use many visual simplifications that we usually accept early on and then interpret automatically. An appropriate example for this discussion is the river on a topographical map. Thin blue lines trace the path that runoff follows downhill across the landscape. Usually blue and not a shade of brown as in real life, this delicate tracery represents a theoretical, unidirectional network that transports water and pollutants downstream under the effect of gravity. We can analyse the network using tools such as Strahler ordering (Horton 1945, Strahler 1957) and create tree structures like those used to show the hierarchy of directories and files on a computer disk. We can reduce the network to a node and link structure as described in Chapter 6.2.2 (railway diagrams) and illustrated in Figure 1-11.

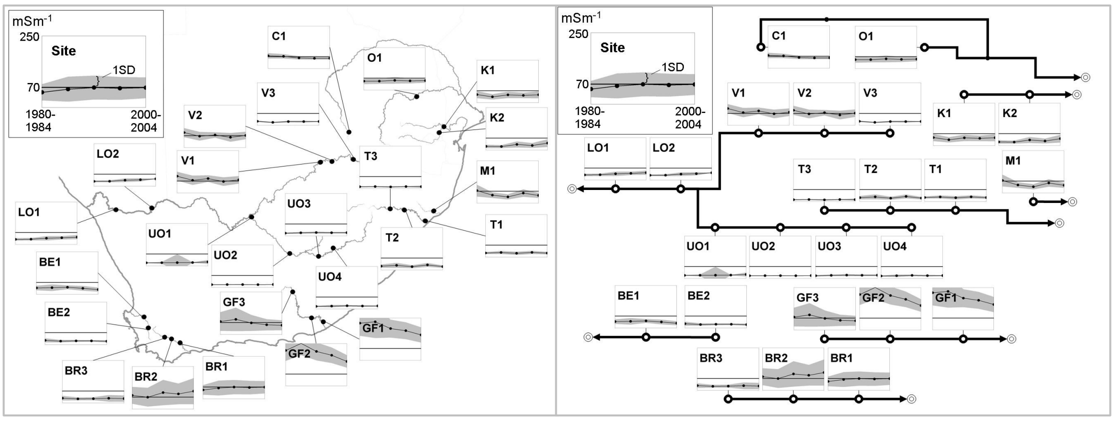


Figure 1-11. Two alternative draft maps that I prepared for publication in van Niekerk *et al.* (2008), illustrating the different types of information conveyed by geographical and "railway" depictions of a river network and monitoring sites. The geographical version was accepted for the final publication.

Simplifications of the river type may improve understanding but may also act as a cognitive block. Representing the river network as a mesh of channels does not convey to the user the interaction with other physical compartments of the hydrological cycle. The blue lines on the map do not evaporate, seep into the ground or accept water from aquifers. The challenge with map visualisations is to reduce the cognitive load just enough to make things clearer but not enough to conceal important information.

1.4.5 Computing visualisations

Now, with the computer, all manner of comparisons seem within rapid reach. The computer provides the necessary simplification, and we believe that's the end of it. *Vive l'intelligence artificielle !* No further need for thought!

--Bertin 1983

Computers, distributed digital databases and automated algorithms have augmented both the speed and the computing power of our brains, and that newfound speed and power is capable of inducing changes in our mental self-image not unlike the ones that sensorimotor technology can induce in our body image: If being deprived of one's spectacles or one's automobile feels rather like the loss of eyes or limbs, being deprived of one's computer or cell-phone feels like the loss of one's intrinsic cognitive and communicative capacity.

--Dror and Harnad 2008

Studies of human creativity suggest that we go through certain processes when producing something new and original. The first stage is preparation (working out a roadmap), followed by ideation (brainstorming), selection (sifting out the workable ideas), production (getting a workable prototype or draft) and eventually evaluation (determining advantages and flaws in the product) (de la Harpe 2007). In all these stages, computer software can accelerate or impede our progress. “The entire problem is one of augmenting this natural intelligence in the best possible way, of finding the artificial memory that best supports our natural means of perception” (Bertin 1983).

Water quality components interact with one another in complex ways. When working with water quality data, the aquatic scientist most likely has some picture in mind (“The pH seems to be going down, I’m sure the sulphate concentration is increasing in this river.”) which he or she then needs to compare with an external visualisation (“OK, so let’s plot sulphate for a couple of sites along this tributary…”) and then internalise it again (“Whoa! Why did they stop monitoring here in 1995… Hmmm, at this site sulphate is high, but it fluctuates like crazy…”) and perhaps use the external visualisation tool again (“This may make more sense if we try a seasonal plot…”). This to and fro process is an important part of conceptualising mental and mathematical models of what is going on in reality, whatever that might be (Trafton *et al.* 2005).

The development of interactive visualisation tools is an expensive and specialised task, and many visualisations that I will discuss are static graphics, designed for inclusion in reports that communicate information to peers and to middle managers (including scientist-managers) in government and research organisations. These are the people best placed to interpret the results in context and communicate them to senior managers, the public, politicians, regional administrators and local authorities. While it appears obvious that clarity and speed are important attributes and that information should be unambiguous, we are dealing with the real world with all its limitations. Nevertheless, certain principles are achievable without too much extra effort. If we settle on a standard and mostly automated reporting process that creates graphics and tables that retain their layout from one reporting period to the next, readers do not have to strain their imaginations to try to reconcile the information in different documents (Wainer 2008). Part of the reason for plotting graphs is so that one can detect anomalous (or "surprising") patterns, which are those whose frequency differs from previous experience (Keogh *et al.* 2002), a task that is much easier when units and scales remain the same.

1.4.6 State of the art

A research direction that has already become established in consumer products is location aware maps, which appear on GPS navigation devices. Instead of looking at a paper map and trying to figure out where you are on it and which direction you are going (exocentric), you now travel through the terrain with a map that continuously presents a representation of where you are and what is in front of you (egocentric) (Porathe 2008). More and more people are in constant contact using cell phones that have small, high-resolution colour screens and these devices are location-aware to a greater or lesser extent. Commercial enterprises have already exploited this for advertising services such as bank machines and retail outlets, so the potential for providing water quality data in this way is very real.

Occasionally expertise, funding, insight and talent come together and produce the kind of graphics that the theorists have been talking about: Intuitive controls, subtle colouring and animation. An example of this is the Gapminder project (www. gapminder.org) which developed the Trendalyzer software, available as the Motion Chart Gadget in Google Documents Spreadsheet (Google 2009), with some resemblance to the Spotfire business intelligence software (Card, 1999: 627). The first application of the Gapminder software was for global demographic data (Figure 1-12). The motion chart is an excellent tool for exploratory analysis of the products of data mining. For water quality applications, please see Figure 3-19 and Figure 3-20.

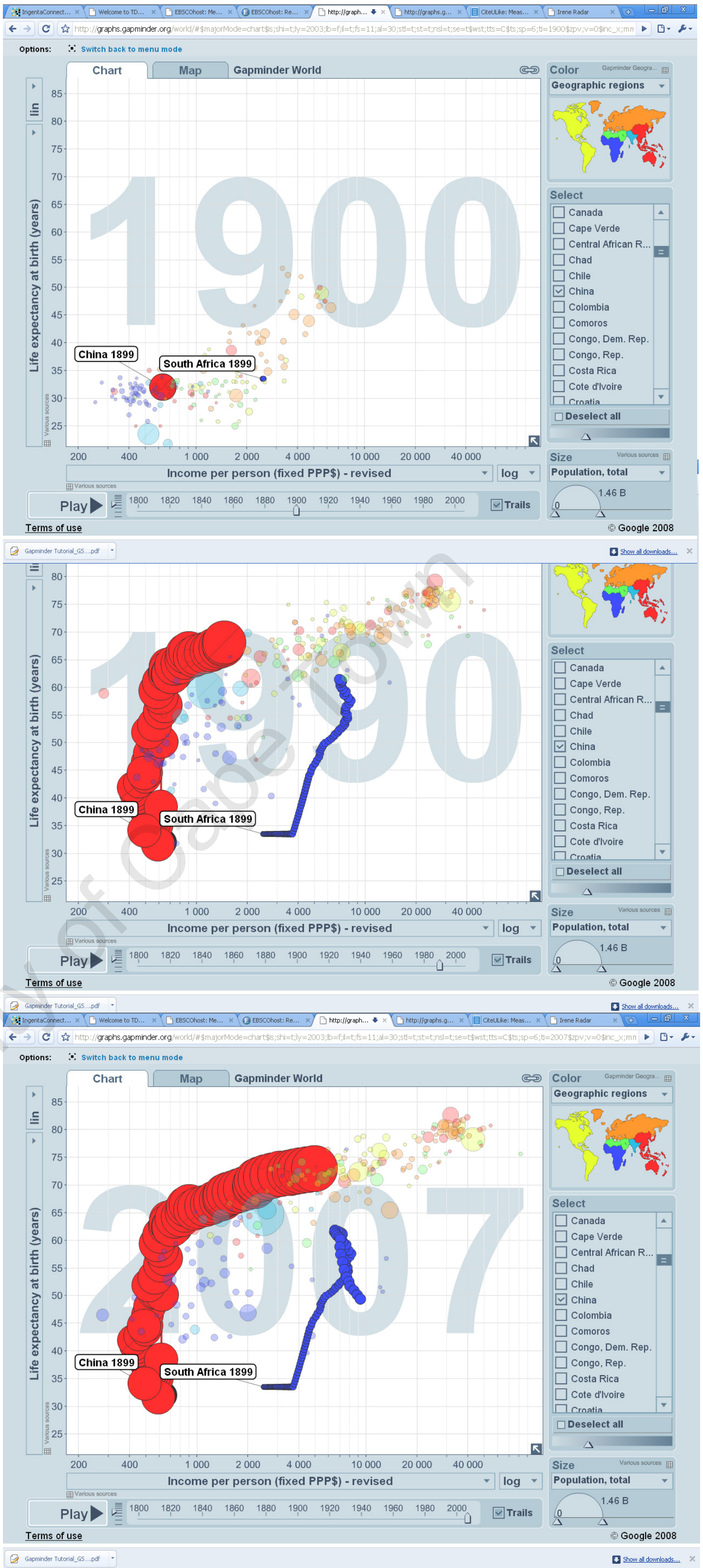


Figure 1-12. Gapminder plots of life expectancy (y) and income (x) in 1900, 1990 and 2007 for the nations of the world. The size of each dot is proportional to the population of the country. The red dots trace the progress of China and the blue dots that of South Africa. Note how life expectancy in South Africa declined after 1990. (After Rosling 2007a.)

Another type of Gapminder chart traces the massive increase in human immunodeficiency virus infections in South Africa during the 1990s (Figure 1-13). If Gapminder charts had been available to present the widely available statistics in a more compelling way, would it have persuaded politicians to react to the public health crisis more effectively? This important question will resurface later as we assess the effectiveness of data presentation methodologies in the following chapters.

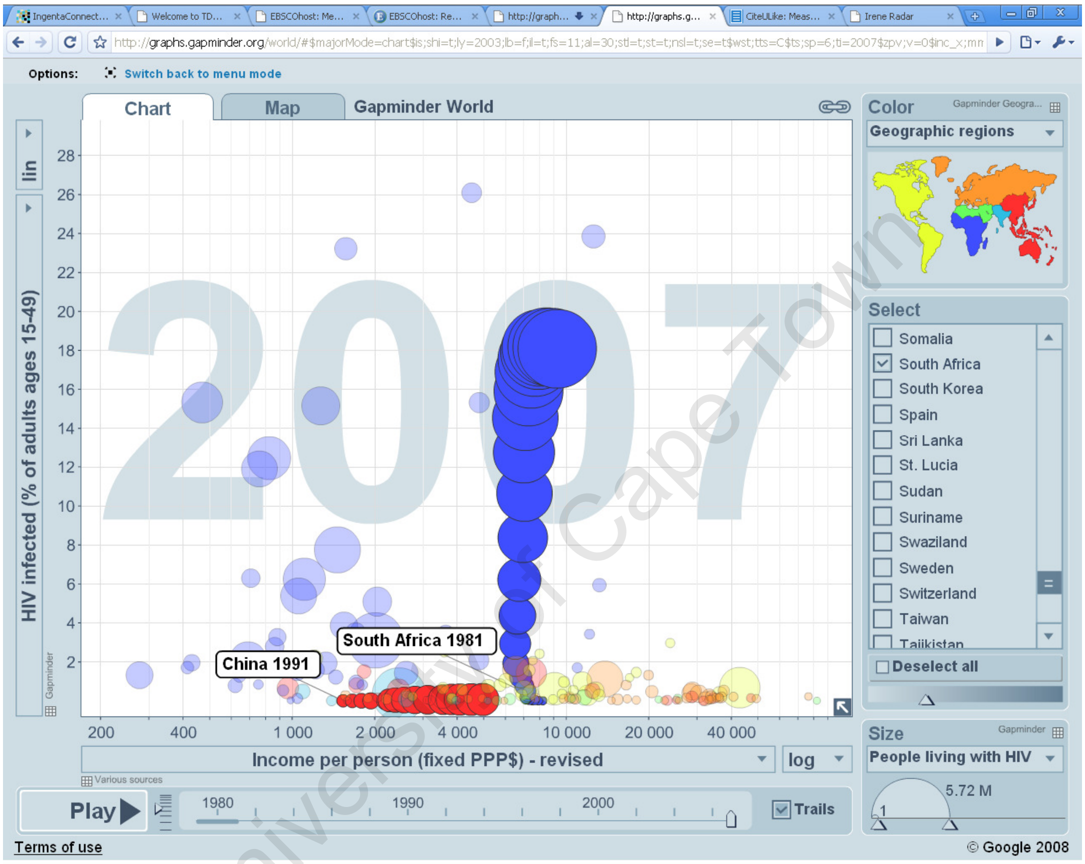


Figure 1-13. Gapminder plot of percentage infection with human immunodeficiency virus (HIV) infection against per capita income, with China in red and South Africa in blue, starting at the dates shown and ending in 2007. The size of each circle is proportional to the number of people living with HIV. (After Rosling 2007a.)

1.4.7 Cautionary notes

Before embarking on the main text of this thesis, we need to pause and consider some vagaries of technological development that constrain data storage and analysis. Although computer hardware has become cheaper to buy and own since the 1980s, the same is not true for software. First-world architecture, commercially-developed applications are expensive, difficult to customise and often fall short of expectations (Brooks 1995, Metzger 1981). Informal, "shanty" architecture developed in-house solves immediate problems but is costly in personnel time and often not transferrable even within the organisation. Free and open source software is convenient because of the lack of licensing concerns, but operational costs are the same as for other applications. Tailor-made, high-quality, interactive scientific graphics require arduous dedication, infinite patience on the part of managers and a large budget.

Computing occurs in a hostile environment. Network management is heavily centralised within organisations, and distribution of security updates and antivirus signatures is not always well coordinated. This project had two month-long delays because of preventable malware infestations. Unstable power resulting from central supply problems, building renovations and cable theft caused damage to equipment and loss of data. The worst events involved the direct theft of 55 computers from the establishment in 2007, causing further losses and delays. Fortunately, backups existed, but the replacement of equipment and restoration of systems took months.

Finally, though it pains me to say this as a graphics enthusiast, visualisation is only a small part of the reporting mechanism. If any of the other crucial components are missing—sampling, operational management, scientific expertise, report-writing skills, the political will to act on information—then the best visualisation software in the world will do little more than record the demise of aquatic ecosystem functioning.

1.4.8 Thesis layout

The layout of this thesis follows a rough progression from simple to complex methods for visual data presentation. Chapters are in an approximate chronological sequence of method development, though considerable overlap occurs because all methods are still undergoing refinement and some methods are dependent on others. In each chapter the type of data and the intended audience differ.

In Chapter 2, I introduce the difficulties of getting to grips with an increasingly unwieldy national water quality database. The solution described developed from an existing concept for representing monitoring frequency at each site, adding variable concentration to give monitoring programme managers a snapshot of water quality trends. The graphic summaries became useful to other scientists in government and elsewhere, for trawling the database in search of potentially useful information.

Chapter 3 is an introduction to the power of radial symbols for representing sequential or multivariate data. The compact format of this symbol, if carefully crafted, is ideal for use on maps. The most striking application is the representation of major ion ratios in their spatial context. These Maucha symbols are useful for ecologists, water quality managers and geochemists.

In Chapter 4, I describe the use of the Google Earth three-dimensional globe viewer to provide a fly-through interface to monitoring sites across South Africa. The application is hierarchical in that the interested user can obtain more and more detailed information about a site simply by clicking on it. This visualisation method is popular with specialists and the public alike.

Chapter 5 contains a description of how the combination of overlapping time series in the water quality and flow databases using standard load calculation methods yields the mass of solutes exported from catchments. This is a specialised application for data exploration and the advantage that it has over other load calculators is that it can run in batch mode and generate hundreds of graphical summaries for later visual comparison.

Finally, in Chapter 6, I review methods for visualising biomonitoring data for a diverse audience ranging from schoolchildren to scientists. The example is the River Health Programme state-of-rivers reporting procedure. The difficulty of communicating the correct visual message to people from vastly different backgrounds becomes evident.

Chapter 2 - Compact time series with the "Barcode" site inventory

Note on originality

Except where citations explicitly state otherwise, the procedures and graphics described in this chapter are my own work.

2.1 Introduction

By 1990, the number of active sites in the South African national surface water chemistry database had approached one thousand, and even the hydrologists closely involved in running the Department of Water Affairs and Forestry (DWAF) monitoring network and database battled to keep tabs on locations of sites, frequency of monitoring and state of water quality. At the same time, scientists in other government departments and at universities were seeing the potential of mining the database for new information about South Africa's water resources, so DWAF was keen to make the data more accessible. One way to showcase the data available was to provide a catalogue or inventory of the national monitoring sites, consisting of one summary page per site, showing the sampling frequency, length of record and other information about the site. DWAF hydrologists already had a series of printed "Green Book" inventories of flow gauging stations and reservoirs, and the chemists developed on this theme with a complementary "Yellow Book" series, the first documents to present virtually the entire DWAF water quality database graphically (TR146 series, Swart *et al.* 1991). TR146 included summaries of each station on a page, with text and a figure consisting of shaded vertical bars along a time axis to show the sampling frequency (Figure 2-1).

South African water quality data users have become accustomed to the Water Management System (WMS) Informix database with data transfer in delimited text format through E-mail and the Internet. This is a great improvement over access in the 1980s, when the Hydrological Information System was located on a Burroughs (Unisys) mainframe database. An information technology high priesthood would extract data to ASCII tables on 9-track reel-to-reel tape, for use on other mainframes. By 1990 the computing environment had become more democratic and users could receive their data on 360kb 5¼" floppy disks for use on personal computers. However, the production of the Yellow Books still required technological gymnastics that included data transfer from mainframe to personal computer and then ingenious use of the then novel personal computer graphics, mainly by DWAF chemist-programmers Ulrich Nell and Willie Geldenhuys, who saw the potential of this new DOS environment and began setting up Pascal programs and Quattro spreadsheets.

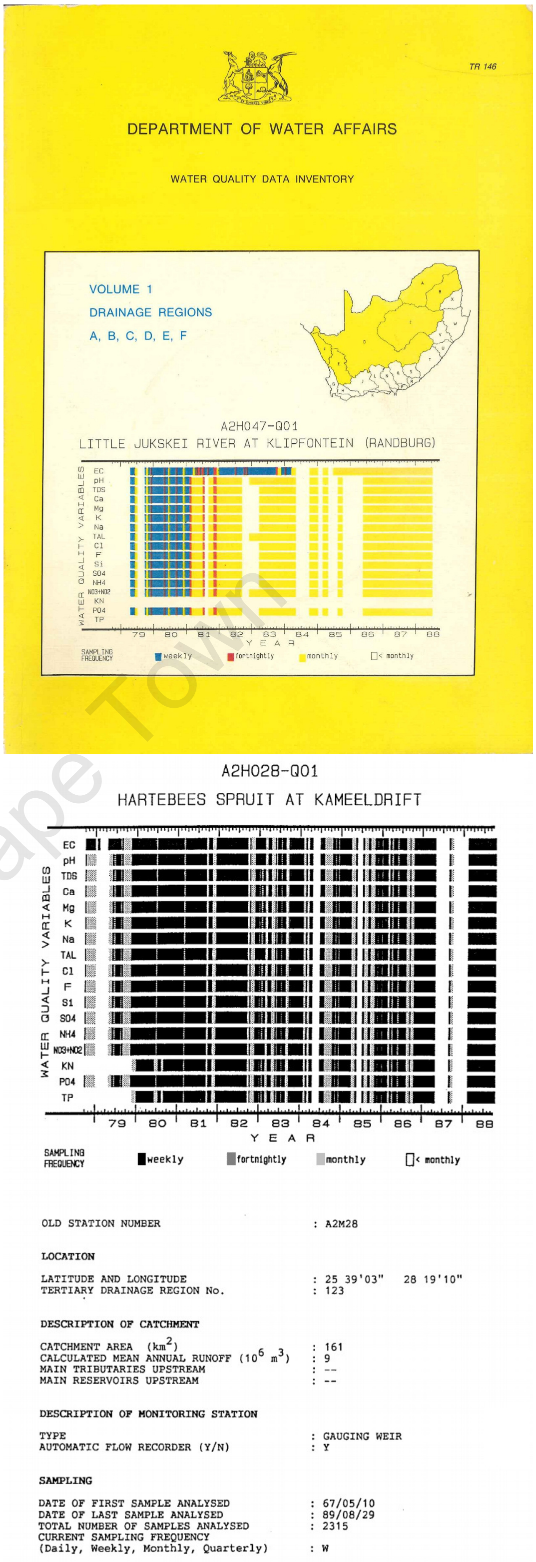


Figure 2-1. Cover and example page of the first volume in the "Yellow Book" water quality monitoring site series, Technical Report TR 146 (Swart *et al.* 1991). A larger version of this figure appears in Appendix 3.

Another report series provided water quality summaries for all stations using box plots (TR145, van Veelen *et al.* 1990). The reports also included some of the first geographical information system (GIS) maps of water quality, in the form of shaded maps of 90th percentile and 75th percentile TDS per tertiary drainage region. This type of shaded area (choropleth) map extrapolated from water quality data collected at points along a drainage system unintentionally implies that one or two sites in a 2000km^2 area represent the entire catchment. The approach is only useful very small scales, in the order of 1:10 000 000 (at which the whole map of South Africa fits onto a report page or computer screen), when the view generalises sampling localities sufficiently. (Note that a later intranet-based water quality mapping system used catchment and river shading– http://intranet.dwaf.gov.za/iwqs/waterlaw/chemical/chemclass.html generated using qatrivhtml.aml –Silberbauer and Howman 2001 and enclosed disc.)

The next step was to include time, site locality, monitoring frequency and water quality into a single summary page, and this is where the Barcode concept originated. Barcode as a visualisation tool was a logical progression from the earlier data summaries, and made use of newly available GIS tools to provide an enhanced descriptive inventory that included water quality results at a glance.

When South Africa emerged from isolation in the 1990s, DWAF investigated off-the-shelf software for data management. For flow data, Hydstra turned out to be suitable but nothing was available for the complex DWAF water quality database, which had various sample management tools built in. STORET (King and Manning 1999) had some potential but was coded for conditions in the USA and did not demonstrate any advanced graphical capability. Liechti and Jakob (1992) produced a beautiful map of Swiss water quality two decades ago but a careful search of their hydrological atlas site has not picked up an update.

(Note: the Barcode name, chosen in the early 1990s, has nothing to do with the Barcode in R-statistics [Emerson and Green 2007, R Development Core Team 2008]; they both just bear a vague resemblance to the machine-readable Universal Product Code).

2.2 Methods

2.2.1 Programming platform

Barcode, originally intended to be a test bench prototype to demonstrate a concept using available technology, has survived since the mid-1990s, when the best available graphics platform on DWAF computers happened to be the Unix-based ArcInfo (v.6, ESRI 2006), programmable in Arc Macro Language (AML), a batch scripting tool. Besides its comprehensive set of GIS functions, AML has the simple pen-plotter-based commands MOVE, DRAW, LINE, SHADE and PATCH, which permit fine control of output for precise mapping and symbolisation. AML also has basic programming elements such as lists, looping and conditional branching. Important at that time was access to and manipulation of large datasets in a relational database and a function to import from and export to text files, abilities that were limited on DOS-based PC hardware and software. Anecdotal information on Internet discussion groups suggests that AML still has a loyal following because of its ease of use compared with more recent graphical and geographical programming methods, and the steep learning curve for adapting old tricks to a new application environment.

2.2.2 Data source

The data layout is the same as that developed for the WaterMarque geospatial data visualisation system for water quality, using the Henco INFO database that forms part of ArcInfo (Cobban and Silberbauer 1993). At about six-monthly intervals, the manager of the Informix water quality database exports two ASCII (text) tables; the first is a station catalogue comprising t least 40 000 records and the second a listing of a subset of the water quality variables sorted by station and date, somewhat more than half-a-million records. This flat-file format is amenable to rapid record-by-record manipulation with simple command-line tools such as awk and sort. AML batch processes import the time-series into an INFO database (sorted on station, date, time, depth) and the station catalogue into a spatial database (coverage or shapefile). Barcode is hard-coded for two standard types of water quality, the inorganic data set and the trace element data set (Table 2-1).

Table 2-1. The two preset groups of data types in Barcode are the inorganic constituents (also known as the "macro" group) and the trace elements.

	Inorganic	**Trace**
1	Conductivity	Arsenic
2	pH	Cobalt
3	TDS	Barium
4	Calcium	Zirconium
5	Magnesium	Zinc
6	Potassium	Lead
7	Sodium	Cadmium
8	Total Alkalinity	Nickel
9	Chloride	Manganese
10	Fluoride	Iron
11	Silica	Chromium
12	Sulphate	Molybdenum
13	NH4(N)	Aluminium
14	NO3(N)	Vanadium
15	KN	Copper
16	PO4(P)	Strontium
17	TP	Beryllium
18		Boron
19		Mercury
20		Antimony
21		Titanium

2.2.3 Graphs and graphics

Time-series plots created using the AML GRAPHBAR command and a common, customised time axis command generated using day-to-number conversion routines, where day 1 = 1801-01-01. Plots are stacked one above the other with a single, common time axis and individual, range-based y axes. Statistics are calculated using built-in AML statistics and a percentile routine provided by John Carter (Hydrological Research Institute, 1994, personal communication: the algorithm does a linear interpolation between percentiles of sorted variables). The minima for variables are generally zero and the maxima are either the actual maximum or a user-selected percentile, except for the special case of pH, which is plotted around the neutral value of 7.0 with a fixed maximum and minimum.

Barcode includes a Maucha diagram (Maucha 1932 and Chapter 3) if the required ionic data are present (potassium, calcium, magnesium, sulphate, chloride and total alkalinity). On the assumption that ionic ratios remain relatively constant with time, the Maucha routine makes use of median values for each ion. For more in-depth studies, time-series Maucha diagrams are available (Chapter 3).

2.2.4 Operation

The user starts Barcode in an ArcInfo Unix or DOS session using an instruction on the command line or in a batch file (Figure 2-2). For example, to plot all the inorganic data and flow records for site A2H028Q01 from 1972 to 2008, the user simply enters:

```
&run barcode # # # a2h028q01 macro flow 1972 2008 print
```

To create a postscript file for all trace element data in secondary drainage region A2, the user enters:

```
&run barcode a 2 # # trace # 1990 2000 file
```

```
2009-04-24 - 12:53:58  by michael in workspace /rpd2/spek/prjws8/users/michael/aml
Arc: &r barcode
******************** Command-line interface! ********************
SunOS operating system using directory /hri
Running /rpd2/spek/prjws8/users/michael/aml/barcode.aml v8.03 on 2009-04-24 -
12:54:02
Usage: barcode CatPri CatSec CatTer Station VarType Flow Year1 Year2 Display
     Pcntl BigFont StnType noZ PctMax GLV inFile Type DebugD
CatPri    limits the search to one primary catchment, (e.g. C, or # to skip)
CatSec    narrows the search to a secondary catchment, if CatPri is set
CatTer    further narrows the search if CatPri and CatSec are set
Station   selects only one station, or CatQat
VarType   analysis group (only macro or trace currently available)
Flow      to plot flow or dam level (where available)
Year1     the first year to plot, e.g. 1983
Year2     the last  year to plot, e.g. 1994
Display   p=print, f=PDF, m=PDF + maucha.jpg file, n=no (or s=Sun screen)
Pcntl     percentile (= 90 percent by default)
BigFont   for setting large fonts for posters and presentations
StnType   R or H (default is B for both)
noZ       Z or noZ (to exclude Z stations)
minNsmp   minimum number of samples to accept (default 1)
PctMax    Y = use the selected percentile as the graph max (default No)
GLV       guideline DHC,ALW,AGI,AAQ,ICn,RFC,RIC,AES (default none)
inFile    input data file (ignore)
Type      sun1 sun2 sun3 sun4 tek1040 tek4107 (ignore)
DebugD    used by programmers to set debugging mode (ignore)
>
```

Figure 2-2. The startup prompt of Barcode, showing the selection criteria. Understandably, users usually find a command line input that works and enter it into a batch job, to save on typing.

2.2.5 Metadata

In this context, the term "metadata" refers to the descriptive information about a monitoring site and the parameter settings for generation of the summary page. It includes the site code, alternative codes, long description, geographic coordinates, input data files, user name, and job run information (date and time of the summary and the version of Barcode). Considerable work was required to develop routines to improve readability of metadata, by transforming the geographical coordinates into degrees, minutes and seconds format and converting descriptive text from ALL CAPITALS (a relic of mainframe card punch days) to a more readable Mixed Case, retaining capitals in special cases such as site codes (e.g. A2H028Q01). See McDonald (1989) and Appendix 6 for an explanation of site codes.

2.2.6 Output

Output can go to the default printer, a graphics monitor or a Postscript file for conversion to PDF, a convenient format for distribution by E-mail or on the Internet (See Chapter 4). An option is available to produce Barcode reports with minimal text in a large font, for presentations.

2.3 Results

The Barcode water quality representation provides a compact summary of the monitoring and analysis of chemical water quality at a site, with time series, monitoring frequency, general statistics, salinity and locality. Below is an example of the Barcode summary page, with data from the Hartbeesspruit upstream of Roodeplaat Dam (Figure 2-3). Monitoring frequencies are apparent, including a complete absence of monitoring halfway through the period shown. Monitoring of total phosphorus and Kjeldahl nitrogen only began five years after the start of the sampling programme. The time-series charts also show seasonal cycles in solute concentrations.

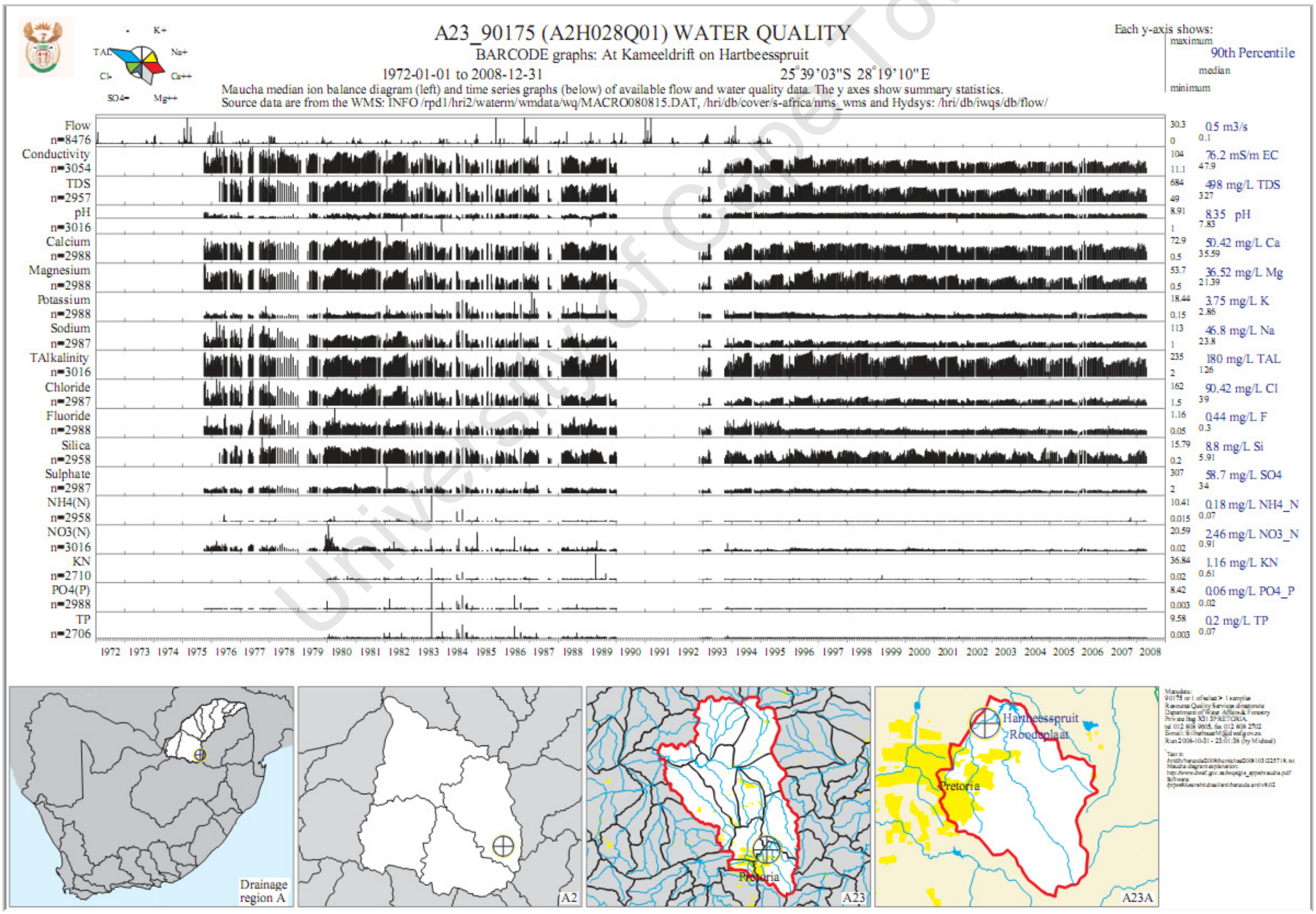
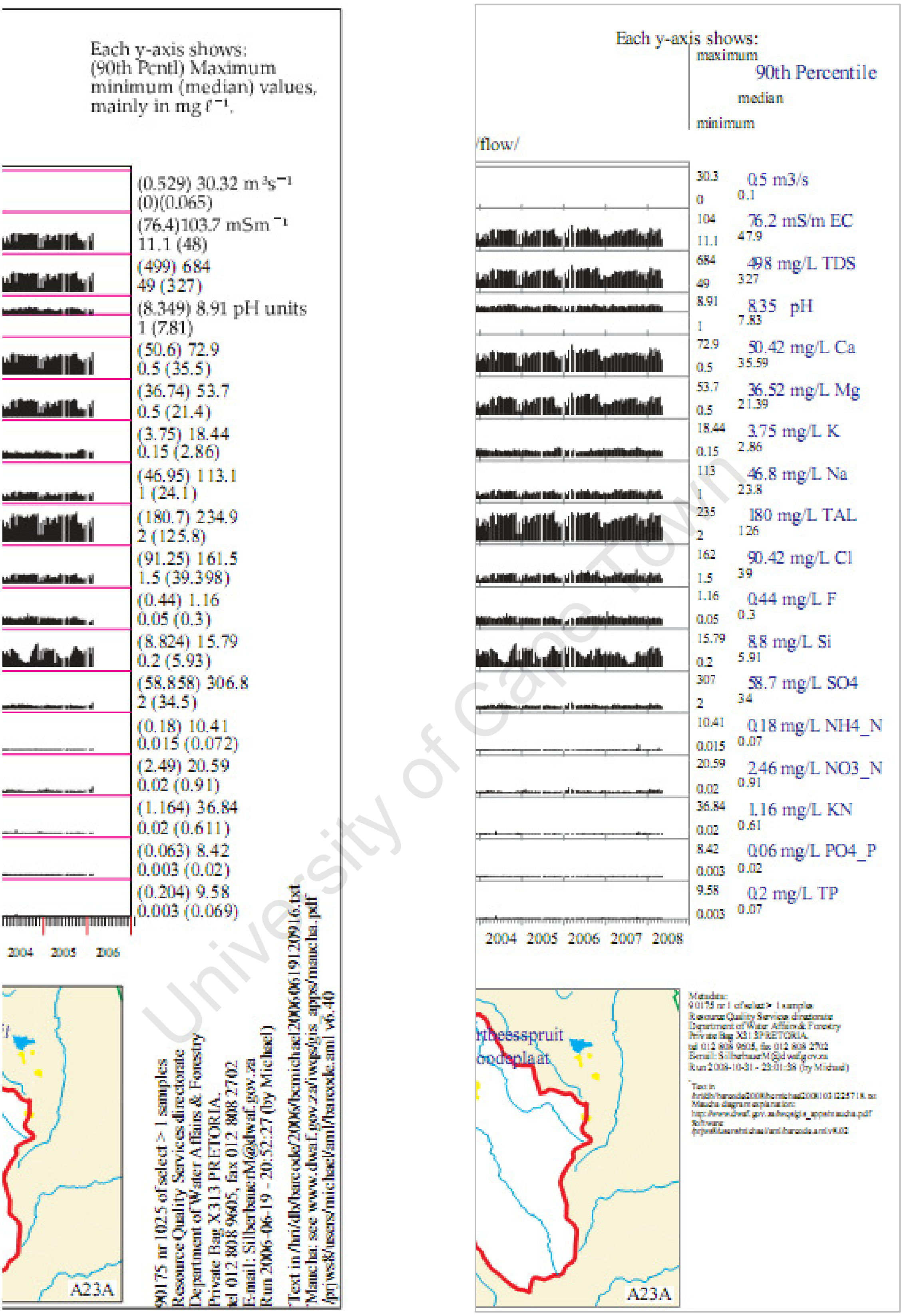


Figure 2-3. Barcode plots for station A2H028Q01 on the Hartbeesspruit, an inflow to Roodeplaat Dam in Gauteng. For larger examples, please see the Appendix.

The software went through a number of incremental improvements during its development, for example changes to the y-axis to improve readability (Figure 2-4).



a) b)

Figure 2-4. The y-axes of Barcode plots for version 6.40 (a) and 8.02 (b) of the software at site A2H028Q01 on the Hartbeesspruit in 2006 and 2008, showing changes to the statistics key designed to improve readability.

The Barcode example shows the sampling frequency (compare with the Yellow Pages example, Figure 2-1), each sample appearing as a vertical hairline bar. Descriptive information from and about the database appears at the top, and the data plots appear as a set of stacked graphs with a common time axis and automatically varying Y axes annotated with descriptive statistics. A Maucha diagram at the top left indicates that this is calcium bicarbonate water contaminated with sodium. Four automatically generated and annotated maps at the bottom progressively zoom into the site locality at the primary, secondary, tertiary and quaternary scale. The block of text at the lower right includes information such as run time and program version for debugging purposes. As far as they could be implemented in a rigid, automated system, the design principles of Tufte (2001) influenced the later development of the software.

The gap in the data from 1990 to 1992 represents a misunderstanding between different sections of the Department of Water Affairs and Forestry about who was actually responsible for monitoring this site. A smaller data gap during 1987 is also apparent in the yellow pages diagram (Figure 2-1). The digital hydrological flow record for this station terminates in 1995.

Barcode has an option to plot data against the water quality guideline for a particular use (DWAF 1998). The unacceptable range appears as a shaded area behind the bars (Figure 2-5 and Appendix 7). Table 2-2 lists the eleven types of guideline currently available with Barcode.

Table 2-2. The guidelines from DWAF (1996) available in Barcode.

Code	Guideline
DHC	Domestic human consumption
DHF	Domestic human consumption - F (fluoride range included)
ALW	Agriculture livestock watering
AGI	Agriculture irrigation
AAQ	Agriculture aquaculture
IC1	Industry category 1
IC2	Industry category 2
IC3	Industry category 3
IC4	Industry category 4
RFC	Recreation full contact
RIC	Recreation intermediate contact
AES	Aquatic ecosystems

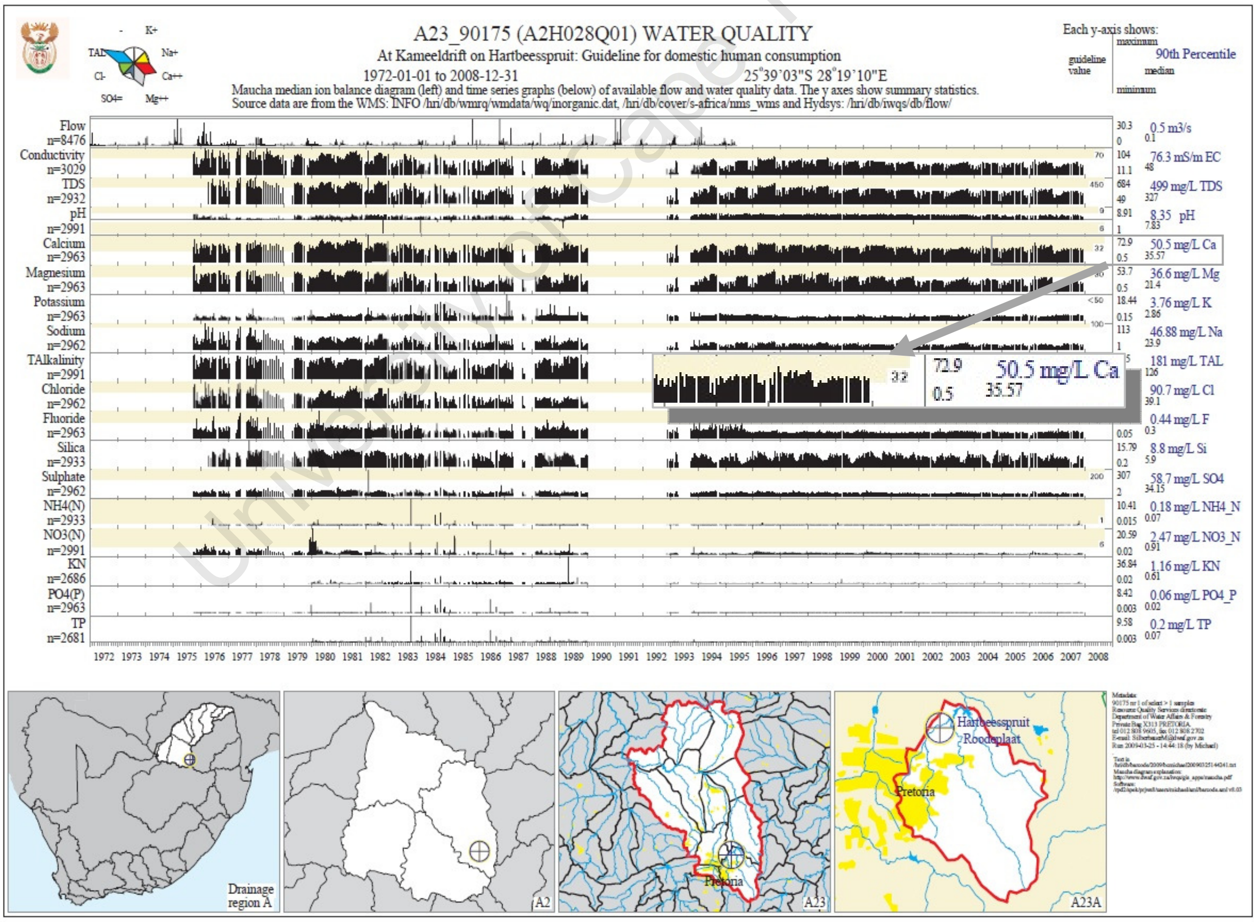


Figure 2-5. Examples of Barcode output with DWAF (1998) "domestic: human consumption" guidelines as shaded rectangles behind the time series plots (a section of the calcium plot is enlarged to show how the concentration just exceeds the 32 mg/L ideal guideline on occasion). The actual guidelines for domestic use include quality indicators not plotted by Barcode, so this type of display highlights problems, not suitability for use. A larger image and examples of other guidelines are in the Appendix.

2.4 Discussion

Barcode has shown its worth as a workhorse for bulk visualisation. The batch process system under which it runs provides a means of maximising use of expensive spare computer capacity to pre-process large data sets for later inspection. However, times are changing and the diminishing reliability of hardware platforms combined with the expense of maintaining the ArcInfo Unix system will probably drive a change to another method for mass generation of visual summaries, as I will discuss in section 2.4.3.

2.4.1 Visualisation

Just as concepts merely refer to more abstract representations in the mind, data are a proxy for the phenomena they are intended to measure.

--MacEachren *et al.* (2004)

The principles of clear no-nonsense graphics articulated so convincingly by Tufte (1997, 2001) were not locally well-known during the early development of Barcode, but had a great influence on the later development of the Barcode graphical summary sheet and related work. For example, Tufte's ideas influenced changes, seemingly obvious in retrospect, such as a pH graph showing the deviation around 7.0, the inclusion of date tics on all x-axes, clearer y-axis descriptors, the option to show guidelines and proper casing of the descriptive text. (Sadly, database managers, steeped in the tradition of card-based data entry, have consistently resisted changing the convention of ENTERING DESCRIPTIONS IN CAPITALS.) The vision research of Ware (2004) and Lanthony (2006) became available even later in the project. Like Tufte, these authors combine an interest in art with their professional skills in psychology, ophthalmology or statistics, to shed light on design principles for graphics that communicate efficiently with the visual system of *Homo sapiens*.

The thin vertical bars that represent solute values in the Barcode application are better suited than line or area plots for showing where sampling has occurred or, looked at another way, not hiding where samples are missing. The gap tooth effect where sampling has not occurred alerts programme managers to problems in the sampling procedure (e.g. Figure 2-3). Line and area plots bridge the gaps and give an illusion of completeness, even with markers at data points. In an ideal world with a fast turnaround time for the entire sample collection-analysis-reporting chain, the programme manager would have the opportunity to take quick corrective action based on this kind of information. In reality, the site presented in Figure 2-3 went unmonitored for nearly three years before complaints from data users and test runs of the prototype Barcode alerted us to the problem. Furthermore, anomalies in the data, such as the high nutrient concentrations and pH values of 1.0 in 1983, ought, if detected, to have triggered an immediate reaction to check whether these were real events, or analysis errors for exclusion from the database.

In each row of Barcode, maxima may thus represent extreme pollution events, misfiled samples or errors in analysis. The option to use the 90th or 95th percentile as a maximum on the y-axis is useful for visual analysis because excluding outliers expands the scale for the remaining data and reveals other patterns. Likewise, the median is a more representative statistic than the average here, as a measure of the general concentration at a site, as it is less susceptible to bias from extreme outliers (Dixon 1953). Examples of possible outliers in the lower three time-series plots (Kjeldahl nitrogen, phosphate and total phosphorus) are visible in Figure 2-5.

Barcode reveals other characteristics of the data that are not obvious from lists or line graphs. For example, where data came from two laboratories, Roodeplaat Dam and Umgeni Water, slight systematic differences in sample handling or analytical methods are evident because the two time series remain quite distinct even where they overlap (Figure 2-6).

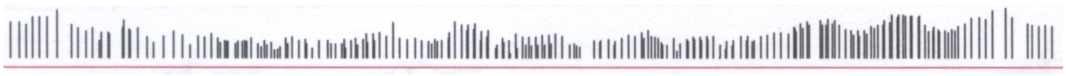


Figure 2-6. Enlargement of a portion of the conductivity graph on a Barcode summary for Mzinto Dam from 1994 to 2006: the maximum conductivity is about $40mS.m^{-1}$. Between 1995 and 2005, the Department of Water Affairs and Forestry and Umgeni Water were both regularly monitoring the same site (on different days). The two interleaved data sets are clearly visible, but the duplication only became apparent when the Umgeni Water data were imported into DWAF's database in 2006.

The common time axis allows a user to compare seasonal and episodic changes in variables with one another, for example flow events with decreases in salinity. However, an area where Barcode's visual comparison is less successful is between the chemical variables and flow. Long-term hydrographs of widely fluctuating stream flow are often imperceptible on a small, linear y-axis, while the picture presented by a logarithmic y scale is difficult to interpret. Flow-versus-concentration comparisons work better in a dedicated load-calculation environment (Chapter 5).

The background shading of guideline values complicates the picture slightly but provides an important benchmark, especially for those not familiar with water quality variables and their meanings. Confusion arises where the whole data set is below the guideline (nothing shaded) or substantially above the guideline (extent of exceedance not clear), although the advantage of this approach over shading everything below the guideline is that shading only appears in cases where a problem exists. An alternative would be to have the part of a bar that exceeds the guideline in a different colour, though the thinness of each bar does not present a very large surface area for colour detection by the eye. Another method is to show deviation from the guideline, similarly to the way that the pH chart shows deviation from pH 7 (Figure 2-3).

The concept of water quality "conforming to a guideline" is in itself unclear, as experts need to evaluate a number of different chemical, physical and biological variables and their interactions before giving guidance on the suitability of water for a particular use and the extent of the risk of using it (Dallas and Day 2004). We cannot completely replace the responsibilities and intuitive problem-solving methods of technicians and managers with a mechanical "expert" system (Leveson 1994), and simply presenting the data with the guidelines and no supporting information may cause unwarranted alarm or complacency. In many ways, Barcode is a system for water quality experts to perform data exploration and analysis, rather than to communicate their final results.

2.4.2 User interface and batch processing

The user interface is plain and to the point, while providing for a fixed range of user-defined options. Usability testing and customer acceptance a decade ago went something like this: "Here it is." "OK, thanks." Any requests for improvements proceeded in a similarly informal way.

Developed at a time when graphical user interfaces such as DWAF's WaterMarque (Cobban and Silberbauer 1993) and ESRI's first version of the PC-based ArcView were becoming available, Barcode's development went against the trend for interactive user interfaces, with a very rigid command-line option and an emphasis on mass data processing rather than user friendliness. This was a conscious decision taken for several reasons that may not be obvious, so I will elaborate on them here.

Computer power was limited and licences for commercial software expensive, so unattended overnight and weekend data processing made economic sense. Using all available CPU power for data manipulation rather than driving elaborate graphical interfaces made maximum use of expensive computing hardware. The tools available in the mid-1990s for developing user interfaces were primitive and time-consuming to implement. In the case of WaterMarque, perhaps 90% of development time and expenditure went into the user interface and training and only 10% into useful analysis of data. Formal user testing of user interfaces did not occur, which added to the expense of training. X-terminal interfaces to the Sun Unix GIS system were few and overpriced, so a command-line input was convenient, working from any text terminal emulator and providing a platform-independent output.

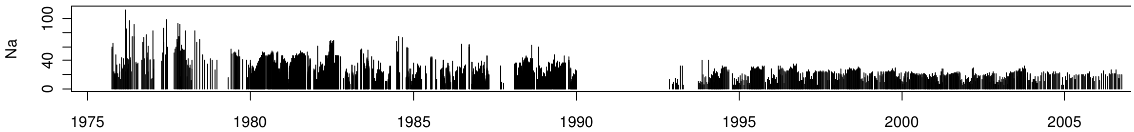
Furthermore, a batch job environment sets bounds on the number of user options and does not need user intervention between runs (users can set up a run for an entire region and leave it to finish over a weekend or longer). Batch processing conveniently leaves a text record of command executions that allows users to repeat routine queries without repeatedly having to step through interminable menus. Finally, even now, adjustments to style in Barcode are tedious, but once implemented the style is identical for all output (no user interference, one size fits all). Batch jobs have a down side, which relates to technological progress not always being uniformly beneficial, as discussed on page 49 (Technical snags).

2.4.3 Future development

While the ArcInfo GIS AML programming environment is stable and powerful, even enthusiasts admit that it has become very much a niche area for development. Few GIS specialists can still code in AML and young programmers understandably prefer to develop marketable skills in more widely applicable languages like Visual Basic or Delphi.

The Google Earth interface described in Chapter 4 takes care of the interactive geographic requirements of an inventory, being superior to the thumbnail maps on the Barcode data summary, meaning that the system does not actually need to be on a GIS at all. A much more widely usable platform than AML for plotting data and statistics is the free R-statistics package (R Development Core Team 2008). It produces a wide range of powerful graphics (see examples in Figure 2-7 and Appendix 4) although it may not be fast enough for interactive applications on large datasets. A possible web application could combine R and Google Maps. R also has its own routines to draw simple maps using digital spatial layers such as shapefiles.

```
# Change all variants of -9.000, -9.0 to -9 in a2h028q01.txt
# In R:
chem <- read.csv ("a2h028q01.txt",na.strings = "-9")
dates <- as.Date (chem$DATE, "%d-%b-%Y") # read 02-FEB-1981 (mil) format
Na<-chem$SODIUM
plot (dates, Na, type="h")
# File-> Save as-> Metafile...
```




```
library(YaleToolkit) # Load package YaleToolkit
sparkline(Na, times=chem$NDATE, xaxis=TRUE, yaxis=TRUE, margins =
unit(c(1,1,1,1), 'cm'), ptopts=('min.max'), IQR = gpar(fill = 'grey',
col = 'grey'))
# File-> Save as-> Metafile...
```

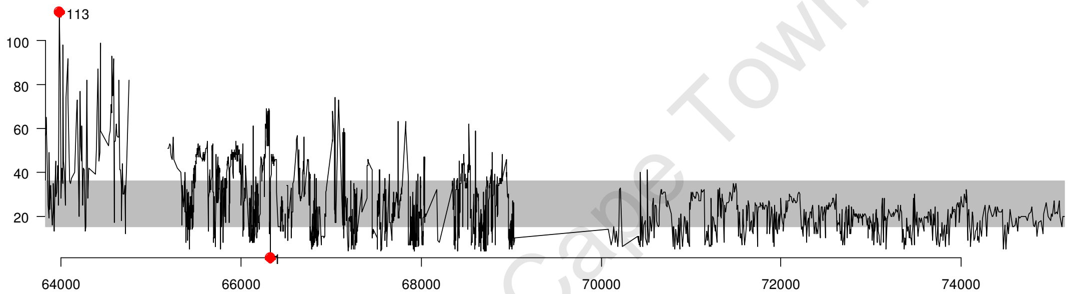


Figure 2-7. Time-series plots of water quality data in bar and sparkline formats using R-statistics (R Development Core Team 2008). The input file is in the comma-delimited form used in Silberbauer and Geldenhuys (2008) and Chapter 3. The first example requires three lines of code and the second example a further two lines. Note that this sparkline example uses day numbers instead of dates on the x-axis, and has not handled missing values consistently.

Far more ubiquitous than R is the Microsoft Corporation's Excel spreadsheet software. As noted in the general introduction, this is not an ideal platform for scientific applications, and its default graphics settings are abysmal. Excel 2003 has 14 chart types of which only two, XY (Scatter) and Bubble, use Cartesian coordinates. All charts default to a grey background. However, with some tweaking, it can be persuaded to do useful things, for example an inventory of water quality data using bubble symbols (Figure 2-8). This type of display is quite tedious to apply in a spreadsheet environment, because of the amount of manual intervention required to get data in the right format and transfer it to the application, but it provides a useful test bed for ideas that a programmer can implement later if the concept is acceptable to potential users.

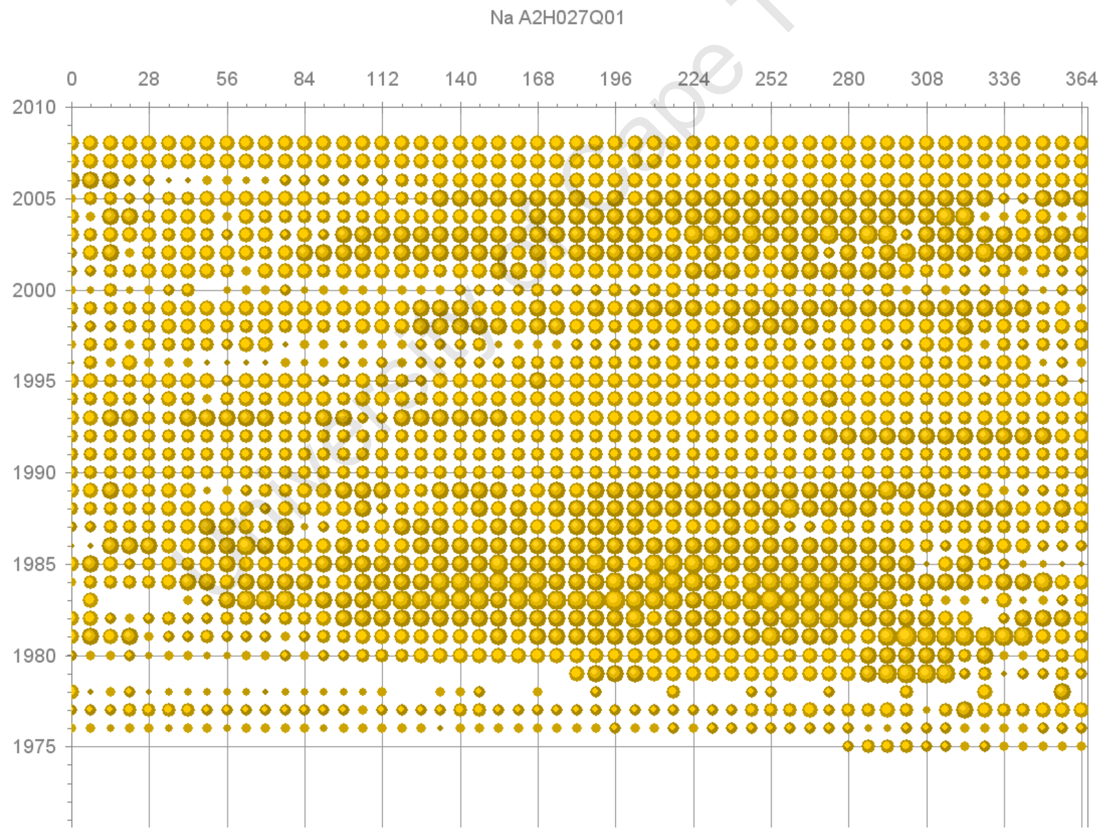


Figure 2-8. Demonstration of bubble plots in Excel to show sampling frequency and sodium concentration at the Pienaars River monitoring site upstream of Roodeplaat Dam. Refinements that a purpose-written program should include are a scale and colour-coding to show exceedance.

Whatever platform we use, assembling time-series data tables from a relational database is an inefficient process. In our work environment, we overcome this by pre-extracting commonly used data into a time-series flat tabular format (often in scheduled overnight batch jobs) rather than bombarding the live relational database with complex queries in real time. This way of doing things also avoids slowing down manual interactive data entry processes during the day.

2.4.4 Technical snags

Technical problems have a way of diverting effort away from the actual task of providing information about water quality monitoring. An early example is the monitoring site A2H027Q01 that I originally wanted to use as an example in this chapter because it has the most data. However the entire site is missing from TR146 (Swart *et al.* 1991), apparently because the data set exceeded the size of a floppy disk (360 kilobytes) and the transfer process from the mainframe failed. Detection of the error occurred too late for inclusion of the site in the final technical report.

A difficulty with maintaining input routines that are reliant on data on another system is that arbitrary changes to the other system can require substantial modifications to input and data formats on your own system. An example is the inconsistent use of site codes. The water quality database previously used the Department of Water Affairs and Forestry standard monitoring site code, for example A2H028Q01 (McDonald 1989) which contains embedded information about the primary and secondary drainage regions and whether this is a river or reservoir site. The water quality site code then changed to an arbitrary number, e.g. 90175, although the standard site code was retained in a separate field and in the hydrological database. The database administrators then began using the alphanumeric site code in non-standard ways, for example truncating the last three characters. They also began adding intelligence to the numeric site code by allocating blocks of number sequences to regions, nullifying the arbitrary nature of the code. For a time I maintained backward compatibility with the original nine-character data field with an elaborate process for converting the new ten-digit code into a shorter hexadecimal number, but this proved far messier than the eventual substantial changes adopted to accommodate the 10-digit code and the non-standard use of the old alphanumeric code.

Computer system management has also affected the data processing environment. Until 2006, an on-site Unix administrator ran our GIS server. The uninterruptible power supply was more or less uninterruptible. Barcode on Unix could process tens of thousands of graphs in one continuous five-week run. However, the IT section decided to concentrate their efforts mainly on Microsoft and its “office” applications. The Microsoft operating system has proved less satisfactory than Unix for running batch processes on personal computers even after a decade or two of development. It is interrupted by virus infections, automatic reboots for operating system upgrades and, in the case of portable systems, the need to shut down in transit. In some ways, administrators and system managers appear to dismiss computers as little more than office appliances for writing memos and wasting time on the Internet.

2.5 Conclusion

The Barcode summary page has proved to be a useful tool for data exploration, especially for seeing at a glance whether a site is active and which water quality variables are present. Resource Quality Services scientists and technicians have generated many thousands of Barcode graphs to help with data exploration, especially as a service to outside users unfamiliar with the WMS database. Chapter 4 describes the generation of about 40 000 Barcode summaries to provide a spatially accessible inventory in conjunction with Google Earth.

Software and hardware platform changes mean that it is time to take this development to another level, using free and powerful tools such as R-statistics. This will require the cooperation of water quality database managers to provide simple, comma-delimited text files of time series data, generated overnight when the system is quiet. The benefit could be faster detection of anomalies in the sampling and analysis procedures, but only if it is part of a wider process of speeding up all the links in the monitoring and analysis network so that we provide information that is both useful and rapidly available.

Chapter 3 - Symbolising multiple variables at points on a map

Note on originality

Except where citations explicitly state otherwise, the procedures and graphics described in this chapter are my own work.

3.1 Introduction

An extremely useful application of spatial analysis systems is the display of information about conditions at points dispersed across a region of interest, for example water quality monitoring sites in a management area. To show simple information about a single variable, we need nothing more elaborate than a digit, symbol or coloured dot, analogous to the pushpins used in paper wall charts. When the information that we wish to convey becomes more complex, for example multiple water quality variables, time series, or exceedance of limits, the choice of a suitable visualisation method demands more care. Circular, radial and star symbols are of great value in certain of these multivariate applications. This chapter will focus on a particular radial symbol, the Maucha ionic diagram, for mapping water quality, while many technical principles described here are also more widely applicable to other types of radial diagrams for maps.

Pie ("camembert") charts are the most commonly available circular symbols provided with commercial software. The pie consists of sectors, one for each variable, with a constant radius and variable angle. The total radius, or pie size, can vary to show differences in total quantity between sites. Whether pies are more useful than bar charts or even tables depends on the data set. Generally, pies are better at comparing proportions of small numbers of variables (Bertin 1967*:* 105, Spence and Lewandowsky 1991, Spence 2005). More complex data sets may demand a different approach: Wainer and Thissen (1981*:* 211) include a pie chart cited by

Bertin (1980) as "an example of poor practice" and observe that sorted rows of bar charts gave a far clearer picture of the data on meat production in Europe.

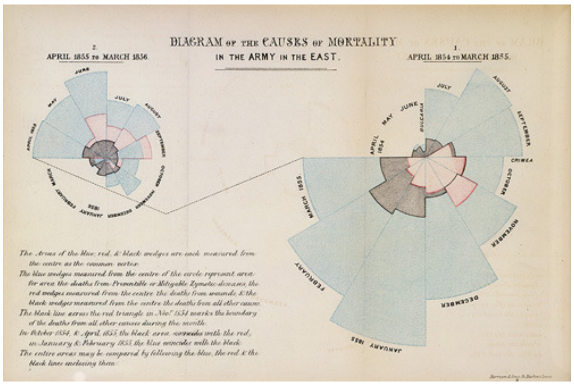


Figure 3-1. Nightingale's 1858 diagram of the causes of mortality in the British Army in Crimea. The largest areas represent deaths from preventable diseases.

3.1.1 Radial diagrams

Florence Nightingale famously used radial diagrams ("polar-area plots", "Nightingale rose diagrams") to plot mortality in the Crimea (Figure 3-1, Nightingale 1858, Small 1998, Wainer 2007, Friendly 2008). A similar diagram to map the suitability of water for use as related to six domestic consumption guidelines appears in Hohls *et al.* (2002) and Silberbauer and Hohls (2002). In rose diagrams, also known as "papillons" or "segment diagrams", the sector representing each variable has a constant angle and a radius or area that is proportional to the variable value (technically it should be the *area*—Wainer 2007). Shading the sectors further emphasises differences.

Another important variety of rose diagram is Feitelson's (2003) *spie* chart (Figure 3-2, which has variable sector angles like a normal pie chart, with the addition of a variable radius to show changes from the initial pie. The spie chart makes use of the human visual system's ability to see quite small displacements in overlapping images much better than even substantial differences between images side-by-side (Palmer 1999: 538).

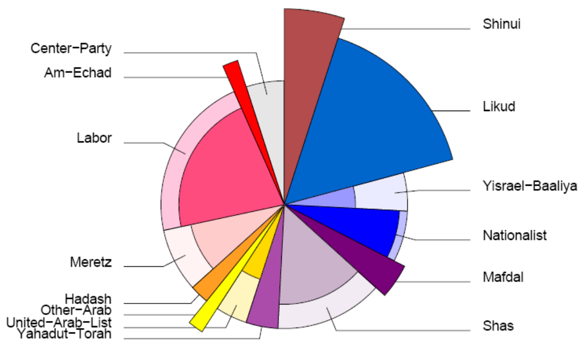


Figure 3-2. Feitelson's 2003 spie diagram of party representation in the Israeli parliament. The circle denotes the starting condition, and the wedges fan out or recede to show changes.

Wells (1999) published a QBASIC system for evaluating a variety of roses, angular histograms, coronas, kites and other radial diagrams for geoscientific applications

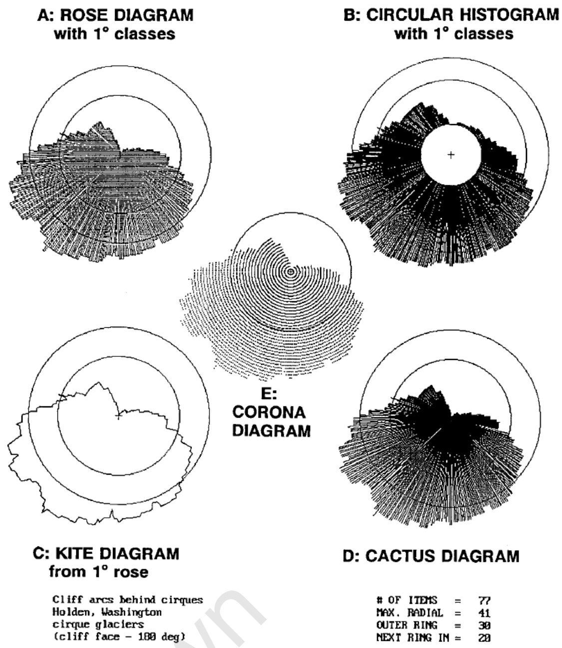


Figure 3-3. Wells' 1999 rose, circular histogram, kite, cactus and corona for aspect data.

requiring the representation of angular information (Figure 3-3). Baas (2000) developed a similar process involving script generation in Excel, execution in Corel Draw, and manual editing to produce publication quality diagrams. The R-statistics package (Dalgaard 2002, Everitt 2005), Statgraphics and Microsoft Excel also provide various radar or spider symbols, which represent each variable by a radial spike, possibly with a polygon envelope joining the tips of the spikes. This type of representation is suitable only for certain applications. "The STAR glyph cannot handle many variables, for even with seven it is hard to remember which is which and thus be able to interpret why a particular state looks as it does" (Wainer and Thissen 1981*:* 226-228), although with careful design to avoid confusion, "multivariate ray-glyphs" and "whisker" or "fan" plots (star plots without a polygon envelope) are ideal for showing multiple variables at points on a map (Ware 2004, Slocum *et al.* 2005).

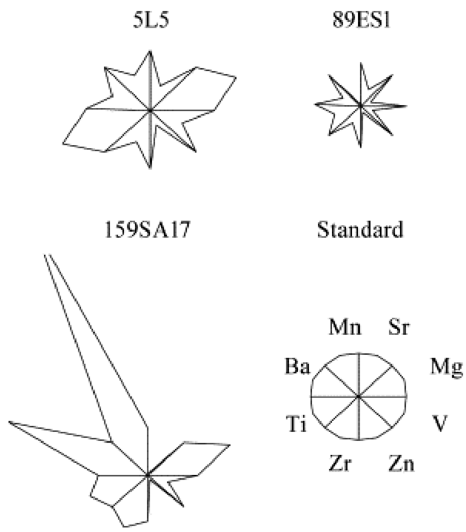


Figure 3-4. Tamás and Abonyi's 2002 fingerprinting symbols for clinker.

Radial diagrams with constant angles are easy to compare and are suitable for visual fingerprinting of the chemical nature of samples, for example trace elements in clinker from different parts of the world (Figure 3-4, Tamás and Abonyi 2002).

The examples above have been for single sites without explicit location on a map. However, radial plots can effectively illustrate spatial and temporal trends for several organic and inorganic compounds on a site map, for example to trace variation in water quality and redox state in contaminated groundwater plumes (Figure 3-5 and Carey 2003).

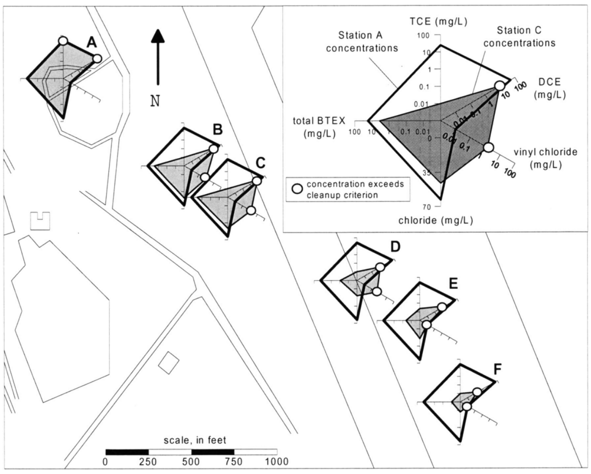


Figure 3-5. Radial diagrams showing changes in organic and inorganic groundwater contaminants in a pollution plume (B-F), relative to the source site A (Carey 2003).

A real difficulty in implementing radial diagrams in spatial analysis systems is that the only standard radial symbol for variable representation on a map in most geographical information systems, including ESRI's ArcGIS and various commercial and open source systems, is the humble pie. As a workaround, one can create radial diagrams as a composite of special font symbols representing, for example, each constituent of a water fitness-for-use evaluation, though the limited availability of standard sector symbol fonts restricts this to approach to two or four variables. This method also requires the plotting of multiple copies of the locality points, one for each variable, which creates confusing interleaving effects when symbols overlap. I have developed an ArcGIS VBA script (based on the Maucha diagram discussed later; Figure 3-10) to produce a less confusing result than that obtained with fonts but at the expense of considerably greater effort (Figure 3-6).

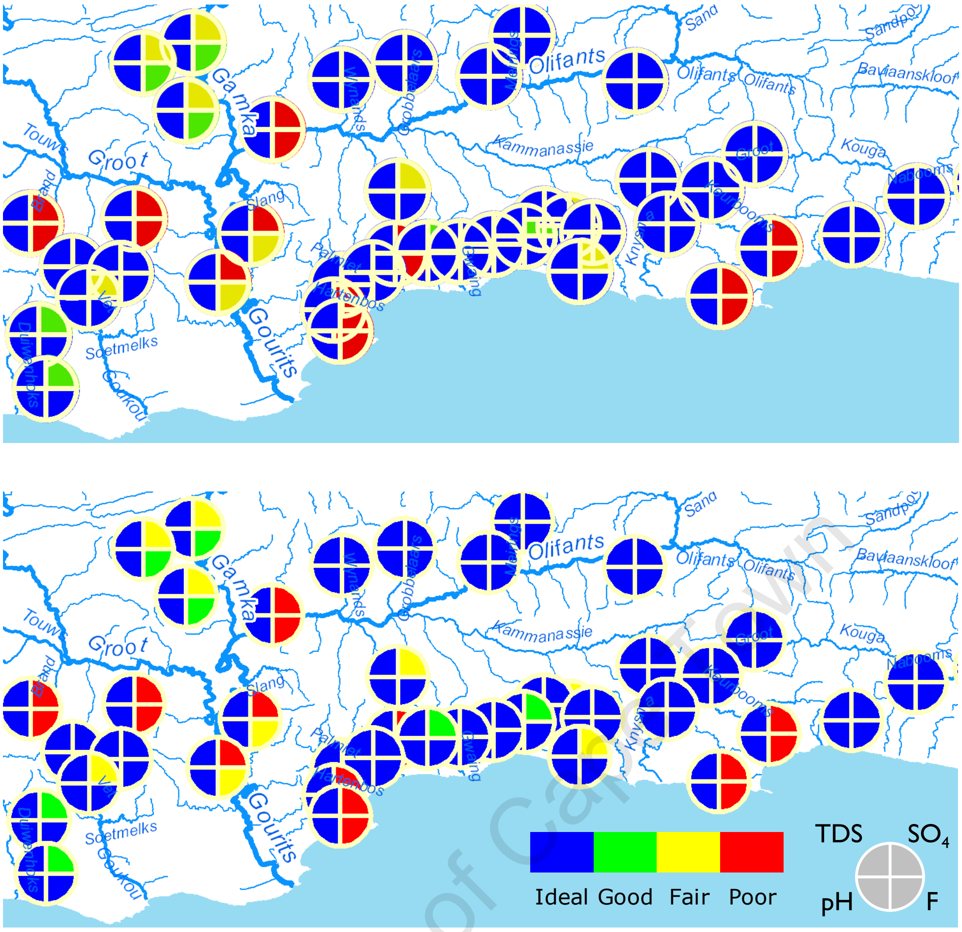


Figure 3-6. Upper: The use of markers (ESRI IGL Font22 122-125 and ESRI default 62) compared with Lower: the use of a VBA script to draw sector symbols. Note that while the VBA-generated symbols do not interleave, overlap is still a problem. The symbols show the class of water use for each quaternary drainage region: this map is only for illustrating the use of symbols and should not be used for planning.

3.1.2 Representing salinity

The salinity of an aquatic habitat and the proportion of the major ions comprising the salinity broadly define the type of aquatic life that the habitat will support. In South African uncontaminated surface waters the dominant ionic types are, broadly, sodium chloride in the southwest or calcium bicarbonate in the northeast (e.g. Figure 3-13). Deviations from the norm may indicate gross pollution, for example with sulphate (from mining) or sodium (from sewage) and, by implication, associated contaminants that are harder to detect. Ionic composition is also a "fingerprint", similar to the clinker and groundwater examples above (Figure 3-4 and Figure 3-5), allowing us to trace the movement of water of a certain chemical type. A useful aspect from the point of view of the analytical chemist is that the cations and anions in a water sample need to balance, and any major discrepancy implies an error in analysis or an unusually high concentration of a normally minor solute not represented in the Maucha diagram.

In comprehensive evaluations of graphical and statistical methods for classifying sites or regions according to their water chemistry, Hem (1985) and Güler *et al.* (2002) have tested the Collins Bar, Pie, Stiff and semi-logarithmic Schoeller plots. Güler *et al.* (2002) evaluated cluster analyses of their data and concluded that

"…the efficiency and semi-objective nature of the [numerical] statistical techniques makes these techniques superior to the graphical methods in order to group samples based on water chemistry data. However, the graphical methods provide valuable information about the chemical nature of the groups. By combining the two techniques we can gain additional information that neither technique by itself can offer."

Several authors have used the Maucha diagram for inter-site comparison (e.g. Maucha 1932, Hassel and Martin 1995, Hassel *et al.* 1997). An unusual variant is a six-sided star diagram with the anions at the bottom and the cations at the top (sodium and potassium combined) used for comparing data in Yugoslavian saline lakes with data worldwide (Petrovic 1981 and Figure 3-7). Some have taken a step further and used the symbols on maps to show regional distribution (Silberbauer and King 1991, King *et al.* 1992, Blinn 1993, Day and King 1995).

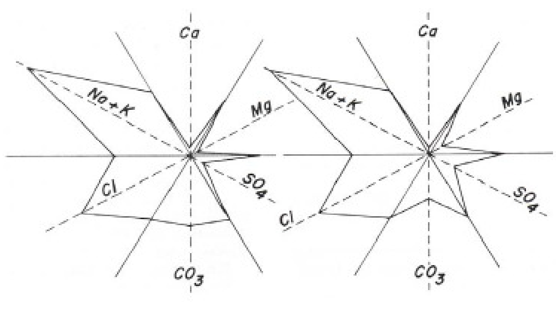


Figure 3-7. The six-sided ionic diagrams of Petrovic (1981).

In order to compare a selection of the graphical techniques for water quality data, I have prepared an Excel spreadsheet where one can enter the solute concentrations in mg/L and immediately see how well the different representations work (Figure 3-8 and Figure 3-16).

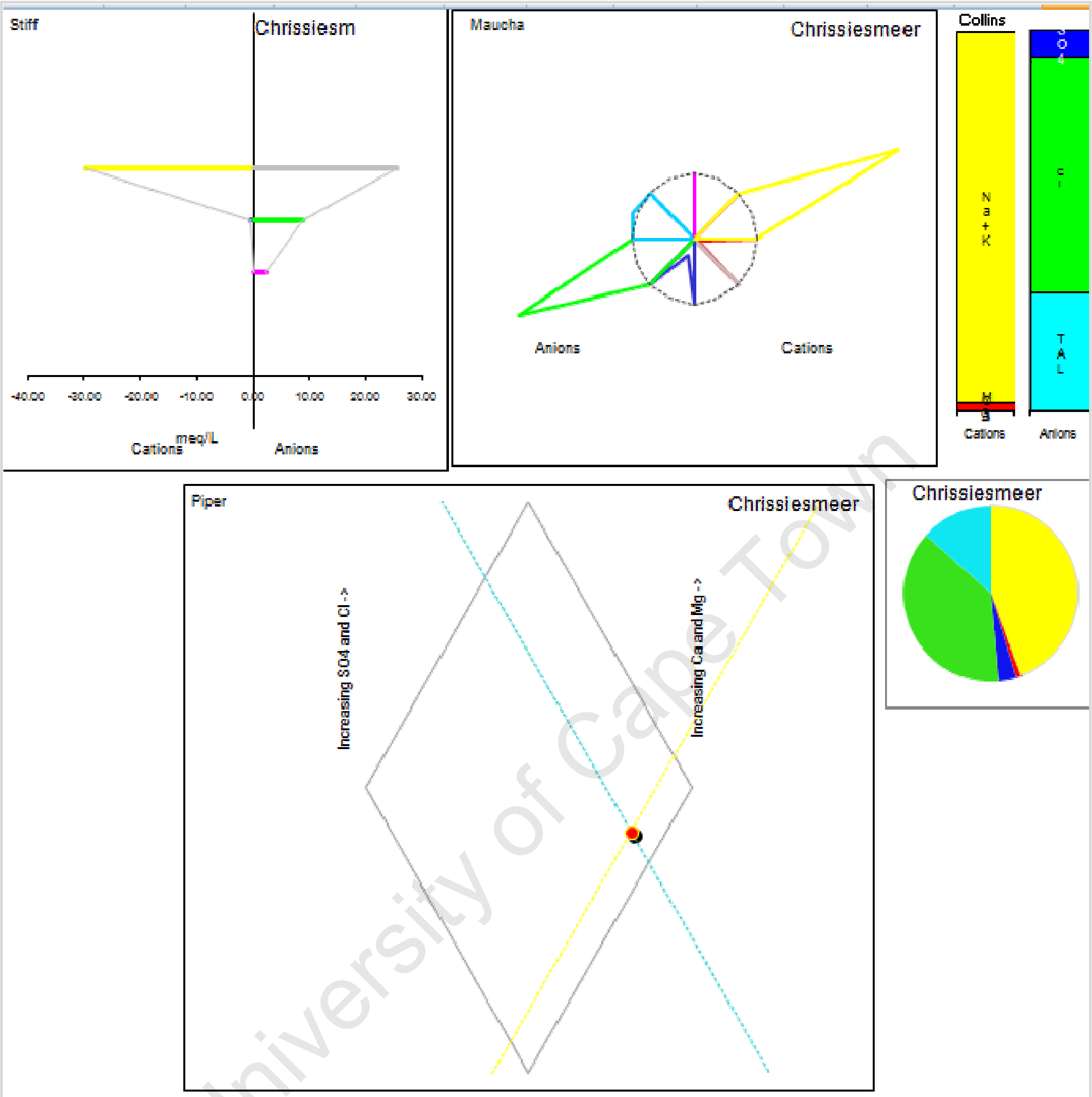


Figure 3-8. A portion of an Excel spreadsheet to compare the representation of the ionic composition of a water sample using Stiff, Maucha, Collins, Piper and pie diagrams. The sample was collected from Chrissiesmeer, a shallow saline lake in Mpumalanga, in 1969. The spreadsheet also contains a filled Maucha graphic (Figure 3-16). The spreadsheet file is available on the enclosed DVD and at http://www.dwaf.gov.za/iwqs/water_quality/NCMP/NWRQSR.htm. Note that all calculations in this spreadsheet are performed with standard Excel in-line formulae and graphics, not Visual Basic for Applications. Colour coding is yellow for sodium, red for calcium, grey for magnesium, blue for sulphate, green for chloride and cyan for total alkalinity. These colours are not completely arbitrary, having some relation to the colours for salts and metals that people might come across in a basic chemistry course.

3.2 Methods

3.2.1 Single diagram

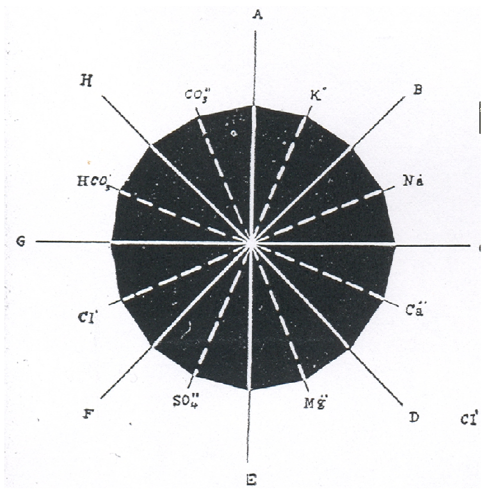


Figure 3-9. Construction of the Maucha diagram (Maucha1932). The outline of the polygon defines a theoretical but unlikely state of equinormality of all eight solutes.

The Maucha diagram is an ionic balance diagram, where the area of each spike is proportional to the ionic concentration in milliequivalents per litre (Figure 3-9 and Maucha 1932). Broch and Yake (1969) introduced the idea of scaling the symbols to represent total ionic concentration in combination with ionic composition. For many years, scientists wishing to use Maucha diagrams had to do the milliequivalent conversion, trigonometry and drafting by hand. The simple but tedious calculation sets a limit to the number of symbols that a person can realistically produce manually.

```
for Spine := 1 to 8 do
begin
    Angle       := Spine - 1;
    Equivalents := ChemInput [Spine];
    IonArea     := (Equivalents / EquivSum) * TotalArea;
    IonRadius   := IonArea / (TotalRadius * Sin(r(22.5)));
    StartAngle  := Angle * 45;
    IonAngle    := StartAngle + 22.5;
    EndAngle    := (Spine) * 45;
    Vertex[1,1] := ScaleX (XCentre);
    Vertex[1,2] := ScaleY (YCentre);
    Vertex[2,1] := ScaleX (TotalRadius * Cos(r(StartAngle)));
    Vertex[2,2] := ScaleY (TotalRadius * Sin(r(StartAngle)));
    Vertex[3,1] := ScaleX (IonRadius   * Cos(r(IonAngle  )));
    Vertex[3,2] := ScaleY (IonRadius   * Sin(r(IonAngle  )));
    Vertex[4,1] := ScaleX (TotalRadius * Cos(r(EndAngle  )));
    Vertex[4,2] := ScaleY (TotalRadius * Sin(r(EndAngle  )));
    Vertex[5,1] := ScaleX (XCentre);
    Vertex[5,2] := ScaleY (YCentre);
    SeedX       := ScaleX (IonRadius/2.0 * Cos(r(IonAngle  )));
    SeedY       := ScaleY (IonRadius/2.0 * Sin(r(IonAngle  )));
    SetFillStyle (SolidFill, Angle+2);
    FillPoly (5, Vertex);
    SetColor (0);
    SetLineStyle (SolidLn, 0, ThickWidth);
    Line
    (Vertex[1,1], Vertex[1,2], Vertex[2,1], Vertex[2,2]);
    Line
    (Vertex[4,1], Vertex[4,2], Vertex[5,1], Vertex[5,2]);
    SetLineStyle (SolidLn, 0, NormWidth);
    SetColor (Angle+2);
end;
```

Figure 3-10. Pascal code fragment for drawing the Maucha symbols in Silberbauer and King (1991). See also Appendix 5.

When personal computers with graphics output became widely available in the 1980s, they provided an ideal platform for automatically generating large numbers of Maucha diagrams for sample quality control and for comparing water quality between sites. Silberbauer and King (1991) used a Pascal program (Figure 3-10, DWAF-RQS 2009), later adapted for Quattro and then Excel spreadsheets (Figure 3-8). Others used BASIC to construct the symbols, for example, Hassel and Martin (1995) and Hassel *et al.* (1997) who compared various water types in Florida.

While computer technology was advancing, our routine chemistry sampling and analysis took a step backwards in respect of discriminating between the carbonate, bicarbonate and aqueous carbon dioxide fractions in water samples. They are lumped into a single entity, "total alkalinity", in the water quality database so only seven of Maucha's original eight rays or petals have data. The eighth ray is retained to balance the symbol shape. (I attempted to use pH to estimate the split between carbonate and bicarbonate, but the uncertainty in the pH of samples that have been in storage for several months after collection means that the results would be misleading.)

3.2.2 Geographical

The Department of Water Affairs and Forestry's first multivariate water quality map of the upper Vaal River catchment consisted of manually drawn pie charts on a hand-drafted map (Figure 1-9 and van Vliet and Nell 1986). Silberbauer and King (1991) published maps of the south western Cape with scaled Maucha diagrams created using the previously-mentioned Pascal program (DWAF-RQS 2009) to generate individual symbols which they then cut and pasted (in the analogue sense of paper, scissors and glue) onto the map.

The advent of ESRI's ArcInfo, programmable with the Arc Macro Language (AML), enabled the construction of maps and geometric symbols on the same platform, using the SHADE command (links to code examples are on the Internet at http://www.dwaf.gov.za/iwqs/water_quality/NCMP/NWRQSR.htm and on the enclosed disk). ArcView 3 later became available, with the Avenue scripting language for customisation. While I found the procedure for creating Maucha symbols technically more difficult in Avenue than in AML, the ability to programmatically group the polygons making up the symbol gives the user the option to manually shift overlapping symbols away from each other or map features, when necessary for clarity. A log scale of concentration, rather than a linear or square root scale, keeps the range of symbol sizes within reasonable bounds, so that samples with low concentrations do not shrink to invisibility and those with high concentrations do not engulf the whole map. The use of log scaling sacrifices qualitative visual interpretation of salinity in the interests of cartographic legibility.

To further improve legibility, preparation of water quality data for Maucha maps includes sorting in descending order by salinity so that the highest salinity symbols are drawn first and do not obscure smaller symbols. A final map of surface water salinity represented by Maucha diagrams at sites across South Africa appears in Figure 3-13 and Silberbauer and Hohls (2002).

ArcGIS 9 with its VBA scripting language added a further degree of complexity to the process of designing custom symbols for maps. After a long battle to understand the different way of coding, and much searching on the Internet, I have written a procedure that emulates the earlier Avenue script, up to a certain point. The most serious problem that remains is that the user can only shift the grouped Maucha symbols around by means of the arrow keys. Any attempt to move a grouped symbol with the mouse causes ArcGIS to crash instantly, with the loss of all edits, causing hostility to the method in all but the most persistent user. The script is nevertheless available on the DVD and at http://www.dwaf.gov.za/iwqs/water_quality/NCMP/NWRQSR.htm.

3.2.3 Time series

One approach to showing the change in ionic composition with time is the use of multiple static images of a single site (Silberbauer 1995), known as "small multiples" in the terminology of Tufte (2001). This "cartoon strip" approach uses a series of Maucha diagrams plotted on a time line, possibly with ancillary information such as river flow rate. Ion ratio changes stand out more clearly than in a multivariate time series plot.

Our visual system is very good at detecting movement, so another way to represent time series is to animate the diagram. I developed an animation method for Maucha diagrams on a map using a batch program in the ArcMacro Language that runs for several days and generates hundreds of image files comprising Maucha diagrams distributed on a simplified map, with the associated hydrographs. Where flow data are available, river width also fluctuates with flow and river colour with water quality. The program includes code to fade images so that they do not abruptly disappear during periods with no data, which would result in difficulty in following the action, a phenomenon related to change blindness (Harrower and Fabrikant 2008). A series of thumbnail images (Figure 3-11) illustrates just a few days of a 29-year record of chemical and hydrological data for the Crocodile River catchment upstream and downstream of Hartbeespoort Dam. The software used to combine the images into movie files was Animation Shop 3 (JASC Software, now owned by Corel and not being developed further, so it will not run under Microsoft Windows Vista) or mpeg_encode.aml (DeMerchant 1997), simply because they happened to be available. One could quite easily use other processors, such as Windows MovieMaker (Microsoft 2004) or Flash (Adobe software). Note that generating polygons within an animation package such as Flash, instead of exporting images and splicing them together later, would give a more polished effect to the final product. The vector method would, however, require new development within the Flash environment. A difficulty with the sequential image method was that computer memory limitations restricted any splicing operation to about 1000 images at a time, resulting in a series of "sub-scenes" which were then further spliced together.

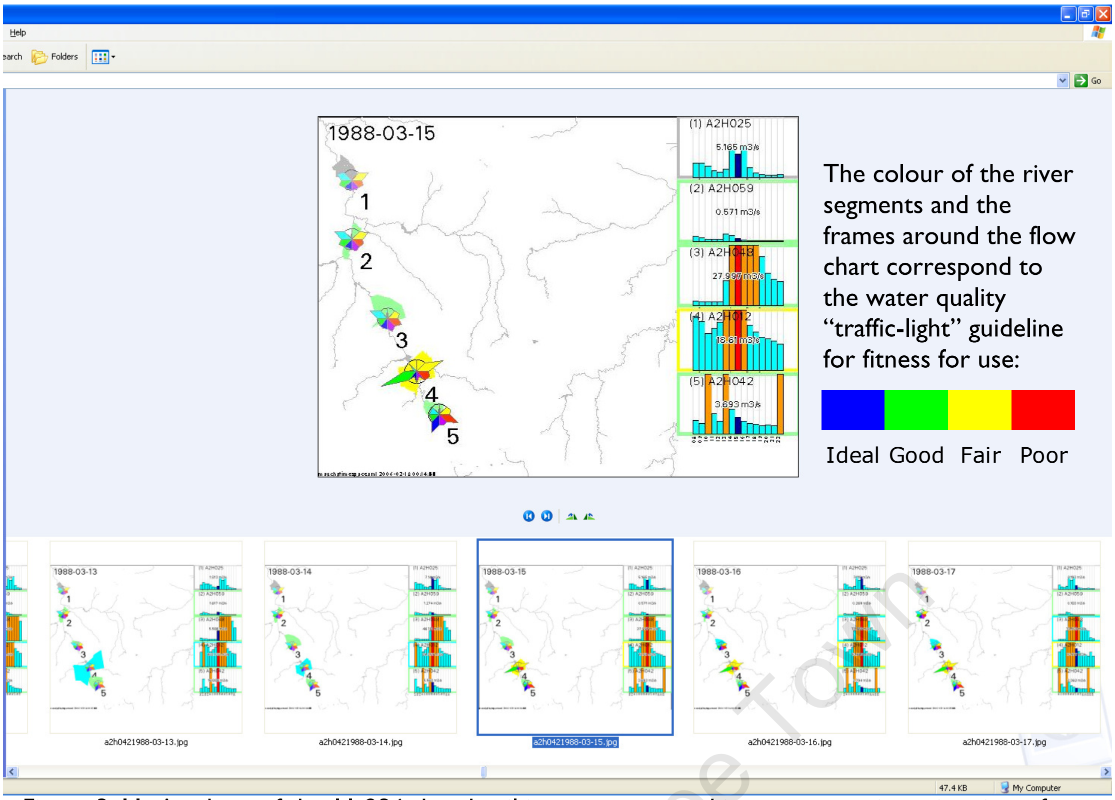


Figure 3-11. A subset of the 11 236 thumbnail images converted into a composite animation of water chemistry and flow in the Crocodile River catchment north of Johannesburg, from 1975-01-01 to 2005-10-05. The splashes of colour behind the Maucha symbols are an artefact of ArcInfo's line width function: river width is proportional to flow rate, also shown as time series plots.

3.2.4 Legend

While users are becoming familiar with a map symbol, an explanatory legend is essential, until they interpret the encoding automatically. However, the complex nature of the Maucha symbol makes construction of a scaled legend difficult. Aquatic scientists are generally more familiar with total dissolved salt concentrations measured in milligrams per litre than the Maucha symbol's milliequivalents per litre. One solution is to use a range of examples at different milligram per litre concentrations for a qualitative comparison (Figure 3-12, Hohls *et al.* 2002, Silberbauer and Hohls 2002).

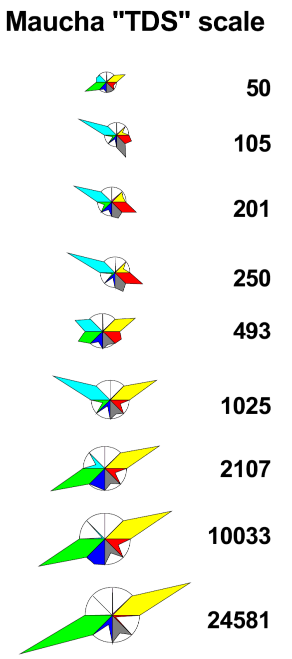


Figure 3-12. A scaled Maucha legend comprising a range of site symbols of varying concentration. Labels are total dissolved salts in *milli*grams per litre.

3.3 Results

3.3.1 Maucha on the map

Maucha diagrams created by the Pascal program and manually pasted onto a map successfully illustrated the distribution of water types in wetlands in the southwestern Cape (Silberbauer and King 1991) and ionic types across southern Africa (Day and King 1995). The methods that I developed in ArcView to produce ionic diagrams directly on maps have been applied in many internal reports and in a published report of water quality in South Africa (Figure 3-13, Hohls *et al.* 2002, Silberbauer and Hohls 2002). Similar maps produced with ArcGIS in VBA are still at the prototype stage (Figure 3-14).

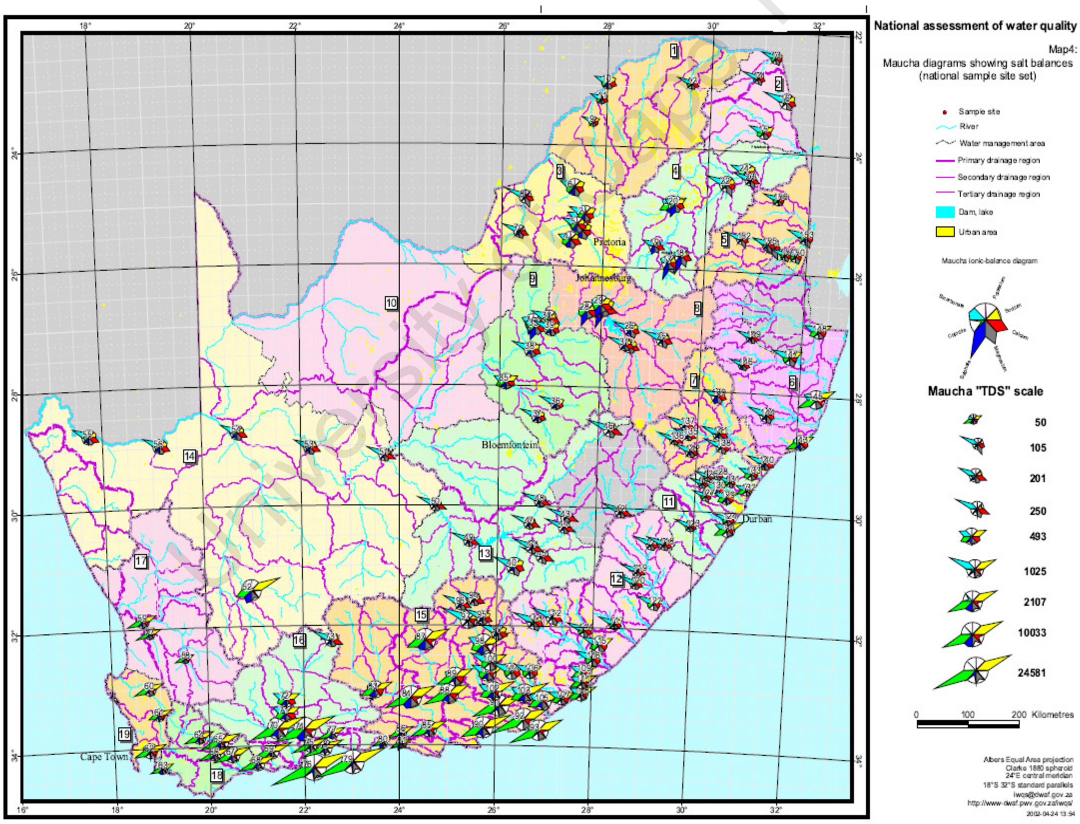


Figure 3-13. Maucha map of aquatic salinity types across South Africa created with an Avenue script in ArcView 3.2 (Hohls et al. 2002).

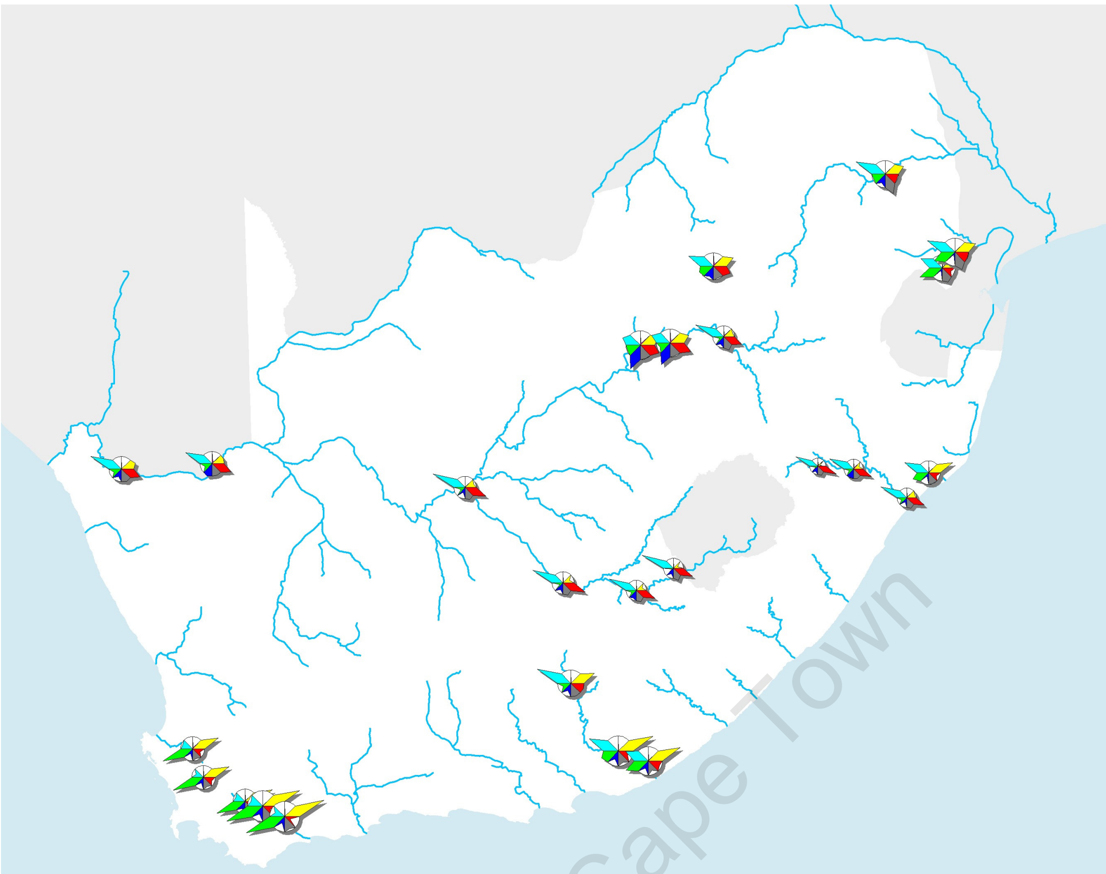


Figure 3-14. Maucha map of aquatic salinity types across South Africa created with a VBA script in ArcGIS 9.2 (code fragment shown below). The complete code is on the accompanying DVD.

```
RayNumber = 0
For Each io In Ion
  Equivalents = Ion(RayNumber)
  angle = RayNumber
  RayNumber = RayNumber + 1
  IonArea = Equivalents / EquivSum * TotalArea
  IonRadius = IonArea / (TotalRadius * Sin (angrad *
22.5))
  StartAngle = angle * 45
  IonAngle = StartAngle + 22.5
  EndAngle = RayNumber * 45
  x(0) = 0
  y(0) = 0
  x(1) = TotalRadius * Cos(angrad * StartAngle)
  y(1) = TotalRadius * Sin(angrad * StartAngle)
  x(2) = IonRadius * Cos(angrad * IonAngle)
  y(2) = IonRadius * Sin(angrad * IonAngle)
  x(3) = TotalRadius * Cos(angrad * EndAngle)
  y(3) = TotalRadius * Sin(angrad * EndAngle)
  x(4) = x(0)
  y(4) = y(0)
  xyv = 0
  Set pPointColl = New Polygon
  Do While xyv < 5
    Set pPoint = New Point
    pPoint.SetEmpty
    pX = (x(xyv) * SymbolScale) + xCentre
    pY = (y(xyv) * SymbolScale) + yCentre
    pPoint.PutCoords pX, pY
    pPointColl.AddPoint pPoint
    Set pPoint = Nothing
    xyv = xyv + 1
  Loop
  Set pFillSymbol = New SimpleFillSymbol
  With pFillSymbol
    .Style = esriSFSSolid
    .Color = pRgb_ray(RayNumber - 1)
    .Outline = pOutline
  End With
  Set pPolygonElement = New PolygonElement
  Set pElement = pPolygonElement
  Set pPolygon = New Polygon
  Set pPolygon = pPointColl
  pElement.geometry = pPolygon
  Set pIFillShapeElement = pElement
  pIFillShapeElement.Symbol = pFillSymbol
  Set pElement = pIFillShapeElement
  pGroupElement.AddElement pElement
Next io
pGraphicsContainer.AddElement pGroupElement, 2
```

3.3.2 Time series

The cartoon strip approach uses a series of Maucha diagrams plotted on a time line, with the option of ancillary information such as river flow rate (Figure 3-15).

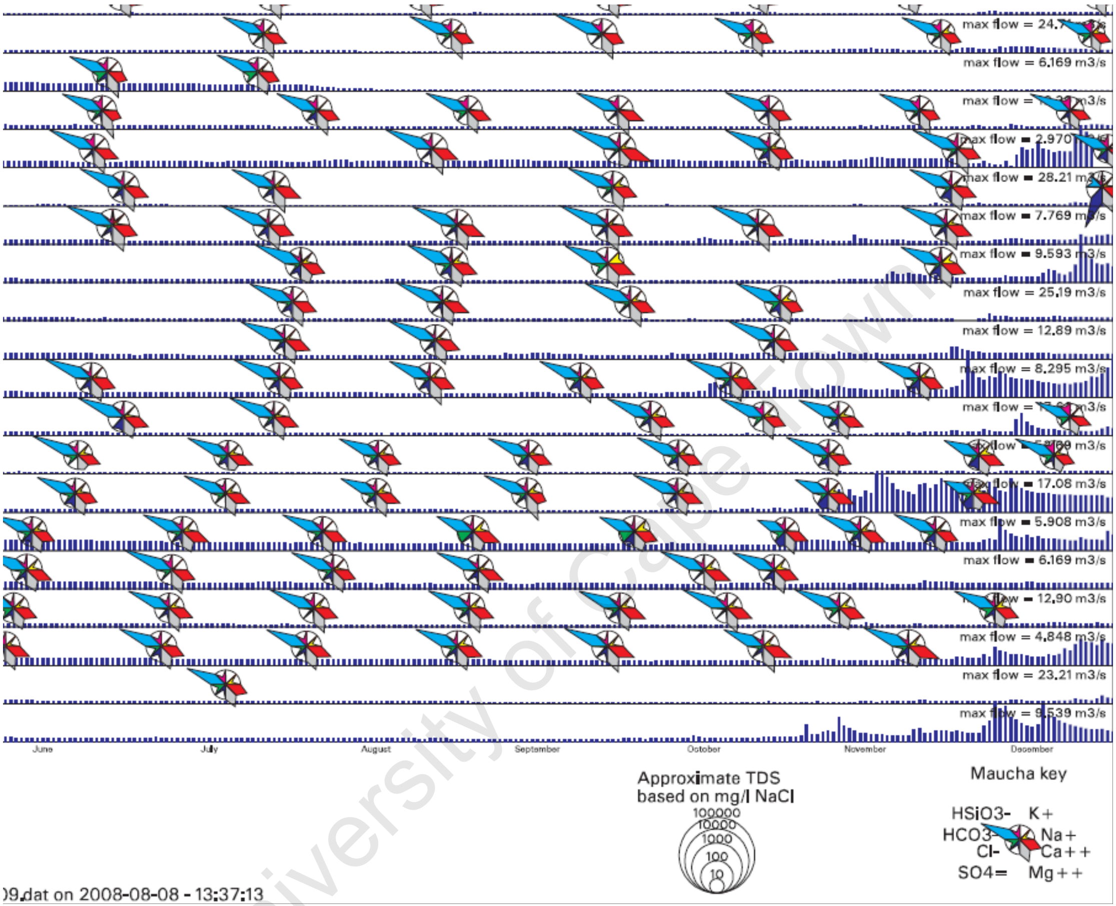


Figure 3-15. A series of annual Maucha time-series diagrams showing the relative constancy of the ionic ratios at a site, even during periods of high flow. The high-sulphate sample at the right, fifth row down, stands out clearly because of its large blue spike.

Animation is obviously difficult to present in print, so sample animations are included on the enclosed disk. The image in Figure 3-21 in the Discussion section is one frame from a movie of the water quality and flow for a set of monitoring sites across South Africa for the period 1980-01-01 to 2006-12-27, compressing 17 years into 329 seconds (a time scale of 1: 1 631 358).

3.4 Discussion

3.4.1 The features that make the Maucha diagram work

The Maucha diagram is a compact and comparative data summary, which works well for the applications described here because of the specific aquatic environmental purpose for which Maucha (1932) designed it. It represents data that must add up to a constant sum on the left and right, so asymmetry and ion imbalance become immediately apparent. In the trained user the expectation of ionic balance and the knowledge that only a few common water types exist, restricts the number of possible shapes to far fewer than for a random eight-pointed star. According to visual theory, his or her low-level (preattentive) visual system is thus able to instantly decode the shape of individual symbols and the patterns of groups of symbols, causing them to seem to pop out with no effort (Ware 2004).

While the scaled Maucha diagram of Broch and Yake (1969) is technically a quantitative representation, the ability of the human visual system to perceive differences in areas of map symbols is qualitative or at best comparative (Slocum *et al.* 2005, 2008). The viewer thus needs to have access to further information for quantitative evaluation as, for example, in the Google Earth interface described in Chapter 4 and Silberbauer and Geldenhuys (2008), where clicking on a Maucha symbol at a monitoring site brings up additional graphs and tables. Maucha symbols have a constant variable angular alignment, an improvement over pie symbols, which have sectors whose angular size and position varies with quantity. One pie on a slide presentation works for a small number of variables, side-by-side comparisons of pies are also useful for small datasets, but dozens of pies on a map display are confusing, because the angular position of each sector shifts with changes in the variable values. However, map designers can improve pie pattern detection by using prominent colouring for the most important sectors. The consistent shape of the Maucha symbol and constant angular position of salient data avoid this limitation and helps us more easily to do visual grouping on a map, because we are good at detecting the angle of objects (MacEachren 1994: Ch 3). When less-regular shapes are necessary for data representation, the inclusion of a reference polygon greatly assists in interpretation. The spie chart is a good example of this, with its start and end conditions overlain for ease of comparison (Figure 3-2 and Feitelson 2003), as is Carey's redox symbol with the desired polygon shape overlain on each map symbol (Figure 3-5 and Carey *et al.* 2003).

The clearly defined rays of a Maucha diagram are easier to discern than the amorphous polygons of a star symbol (Figure 3-16). The segments of the Maucha symbols also lend themselves to shading or colour-coding to simplify comparison. A frequently-used alternative polygon diagram is that of Stiff (1951). Although quantitatively useful, it expands in the horizontal direction only, obscuring map features, although some users have bravely attempted to make it scalable and map-friendly (Tonjes *et al.* 1995 and Figure Figure 3-17). It has the same failing of spider and radar plots in the lack of easily-defined segments to simplify object recognition.

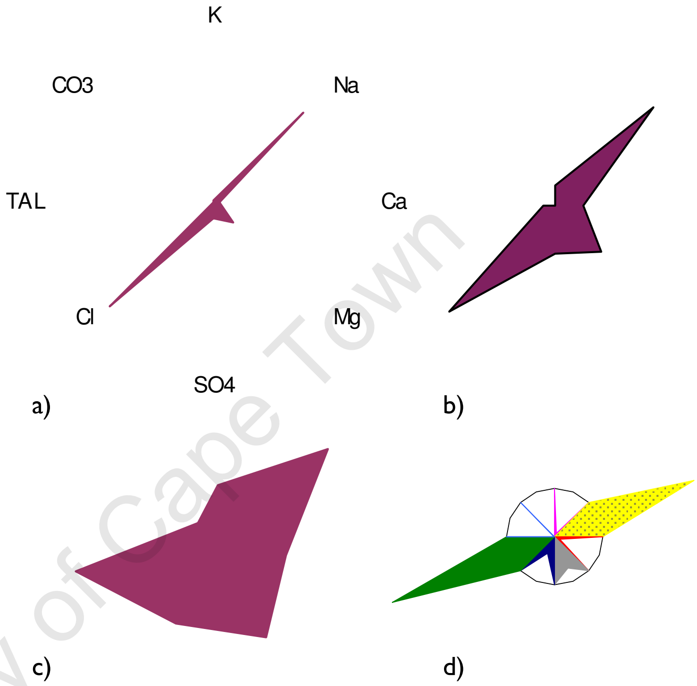


Figure 3-16. Radial diagrams (a to c) of ion ratios produced in an Excel spreadsheet (see Figure 3-8) compared with a Maucha diagram (d), all for standard seawater. Data transformations are linear (a), square root (b) and logarithmic (c).

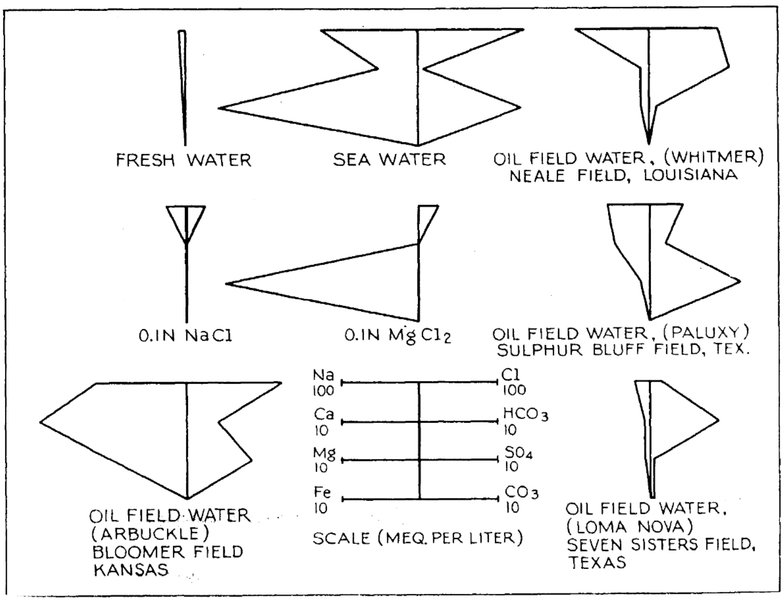


Figure 3-17. Examples of the Stiff diagram by Tonjes *et al.* (1995).

Another important characteristic that sets the Maucha diagram apart is the circle (actually a 16-sided polygon) of equinormality, which provides a reference point for the rays: the displacement of each ray from the reference circle shows the deviation from the norm. The importance of this circle can be seen in the six-sided star diagrams of Petrovic (1981) mentioned above (Figure 3-7), where it is absent, making the results difficult to compare. To sum up in vision theory terminology, the Maucha diagram forms a pattern that has *gute Gestalt* or figural goodness (Palmer 1999).

For completeness, we should note here the "other" less-successful Maucha diagram, which has six linear vectors of ionic concentrations in milliequivalents per litre instead of the more striking radial petals (Figure 3-18). It is quantitatively simpler to interpret but lacks the compact form factor of the area-based Maucha diagram (Maucha 1949 in Hem 1985). For some reason, and unfortunately for the wider adoption of the 1932 version, this is the only Maucha diagram mentioned in the classic American reference on reporting natural water quality (Hem 1985: 175).

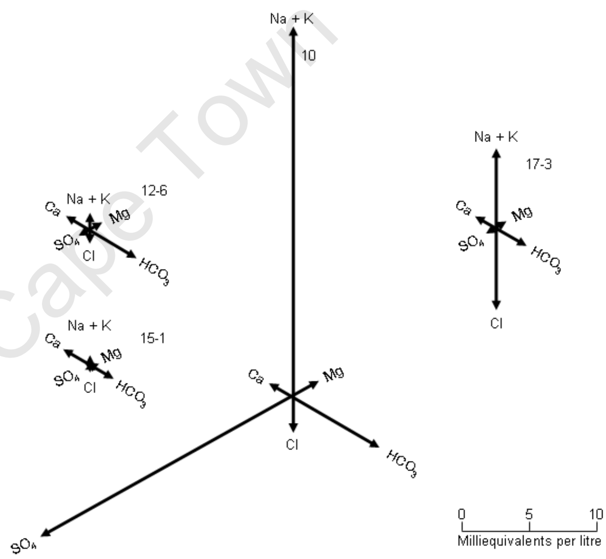


Figure 3-18. The "other" Maucha diagram, using six linear radial vectors (Maucha 1949). Examples of symbols for four water quality samples, redrawn from Hem (1985 Figure 31).

In the end, much depends on the user's familiarity with the symbol. Looking at the problem mechanistically, repeatedly used symbols no longer require cognitive interpretation; we cease to refer to the legend to understand the symbol every time we use it, and it becomes part of our precognitive visual system, where we automatically interpret the meaning without being aware of doing it (MacEachran 1994: 45). The Maucha symbol has several features that correspond with aspects of preattentive processing: size, orientation, colour and spatial grouping (Ware 2004).

When implementing the Maucha diagram as a map symbol (sometimes called a map glyph), one has to strike a balance between these powerful characteristics. If the symbols are too large, they obscure the sampling locality and any other symbols that constitute a group. If they are too small the orientation and colour coding of the rays is indistinct. Not much of this theory was known in the 1930s. Maucha (1932) thus appears to have intuitively hit upon a symbol that allows us to *see* the differences between aquatic salinity types (especially calcium bicarbonate versus sodium chloride) without consciously thinking about them.

3.4.2 Time series

A series of Maucha diagrams plotted along a time axis is an example of the application of “small multiples" (Tufte 2001). Small multiples allow the viewer to make instant comparisons of the system state at different times, something that is difficult with a video because of the limited capacity of the human visual system for remembering temporal patterns (Ware 2008: 174).

Animation allows us perceive other types of information, such as the movement of clusters across a map (Griffin *et al.* 2006). Animation also helps the viewer to keep track of the transition of an object between one state and the next (Heer and Palmer 1999, Robertson 2007) and animated cartography elevates the map from a description of state to a description of processes (Harrower and Fabrikant 2008). For effective use of a complex animation, fine control of the action replay process is essential, otherwise the action flashes by too quickly for comprehension (Slocum *et al.* 2000). This tight control of frame rate, frame stepping and reverse stepping is unfortunately not a characteristic of most off-the-shelf free digital movie viewers. However, the Gapminder tool used by Rosling (2007a) is now available as part of the Google Desktop suite of Internet software (see also Figure 1-12 in the general Introduction). Figure 3-19 and Figure 3-20 show a sample application of this tool that I have set up to track the mean annual salinity, sulphate and chloride at twenty-five monitoring sites in South Africa.

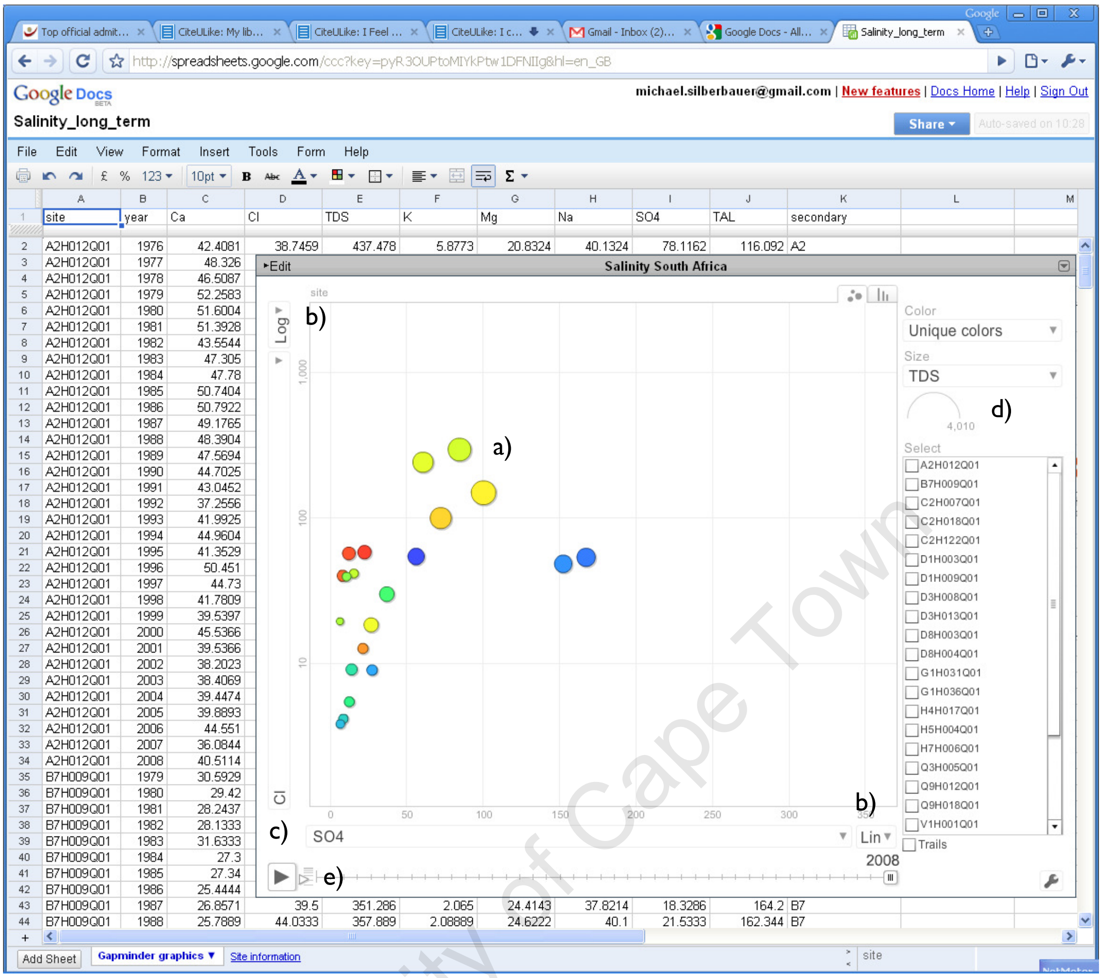


Figure 3-19. Google Docs Motion Chart Gadget for observing the changes with time of any two variables. The example here uses the sites selected for a study of long-term salinity trends collected by the Department of Water Affairs and Forestry (van Niekerk et al. 2008). The user has control over symbol coding (a), labelling (Figure 3-20), axis type (b—linear or logarithmic) the choice of variables (c), the scaling variable (d) and the rate of change of the time slider (e). The application is currently available on-line at a link on this page: http://www.dwaf.gov.za/iwqs/water_quality/NCMP/NWRQSR.htm

Data preparation consisted of calculating a simple annual mean for the available salinity data at each site and exporting the results in the tabular form required by the application (site, year, mean1, mean2, mean3…).

The Motion Chart application gives the user much greater control of the display and is therefore preferable to the pre-packaged animation method. Users can also write their own code to produce applications (or "Gadgets") specific to their own needs. Configuration of the software to produce gadgets for dynamic radial or even Maucha diagrams is therefore quite feasible. This comes close to fulfilling the dream of the prescient statistician Tukey, "for when one watches a movie of the progress of such an iteration—something the computer and the microfilm plotter can easily produce—one may truly see the points flowing back and forth" (Tukey 1965).

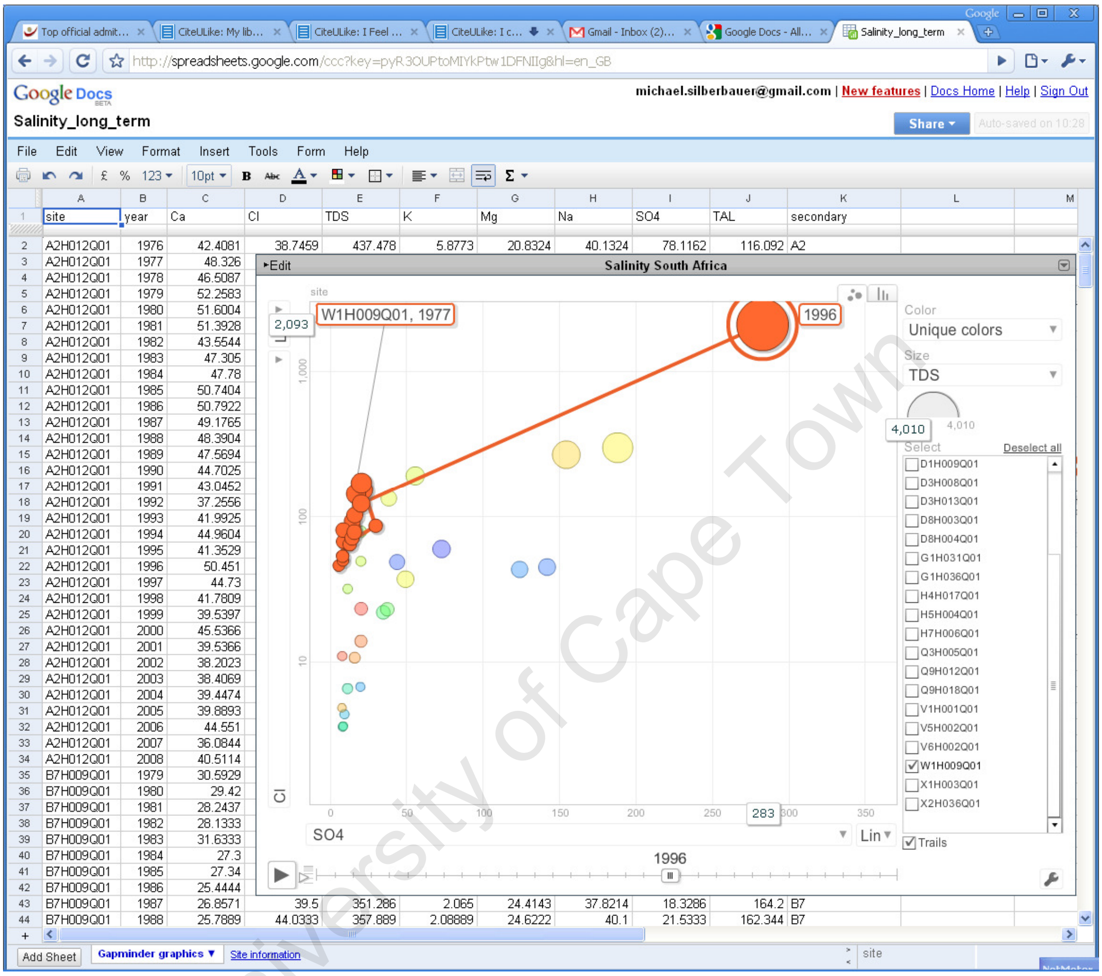


Figure 3-20. The data from Figure 3-19 in the year 1996, showing unusually high mean annual salinity at a site on the Mhlatuze River. Note the facility to label a site and track its progress with a trail, and to display the year and the sulphate, chloride and TDS concentrations (milligrams per litre) when the cursor hovers over a site symbol.

A difficulty with the non-interactive variety of animation (and PowerPoint slide shows suffer the same drawback) is an association with going to the movies, which encourages the viewer to slide into a passive "entertain-me!" frame of mind. With no control of the display or interaction with the developer, he or she may quickly enter that state of impaired cognition, boredom. At the other extreme, unless provided with careful cueing and user interface design, animated displays run the risk of inducing cognitive overload in the viewer. Pre-packaged animations should therefore be short and to the point, rather like a television advertisement, and have a temporal legend where appropriate, such as a digital clock or graphical time indicator (Harrower and Fabrikant 2008). Keeping track of time indicator, legend and changing map symbols is a complex multitasking process that has the undesired effect of dividing the viewer's attention. Possible solutions to cope with this confusion include sound cues or temporal legends embedded in the map, although widely accepted guidelines for the time annotation of cartographic animation do not yet exist (Harrower and Fabrikant 2008). In an example developed during this study, a simplification of the layout that would have helped associate flow with locality (Figure 3-21) was not feasible because the home-built animation process did not lend itself to easy modification—the first test had fortuitously used a set of monitoring sites stacked one above the other on a map (Figure 3-11). Optimal layout is clearly an area for further development.

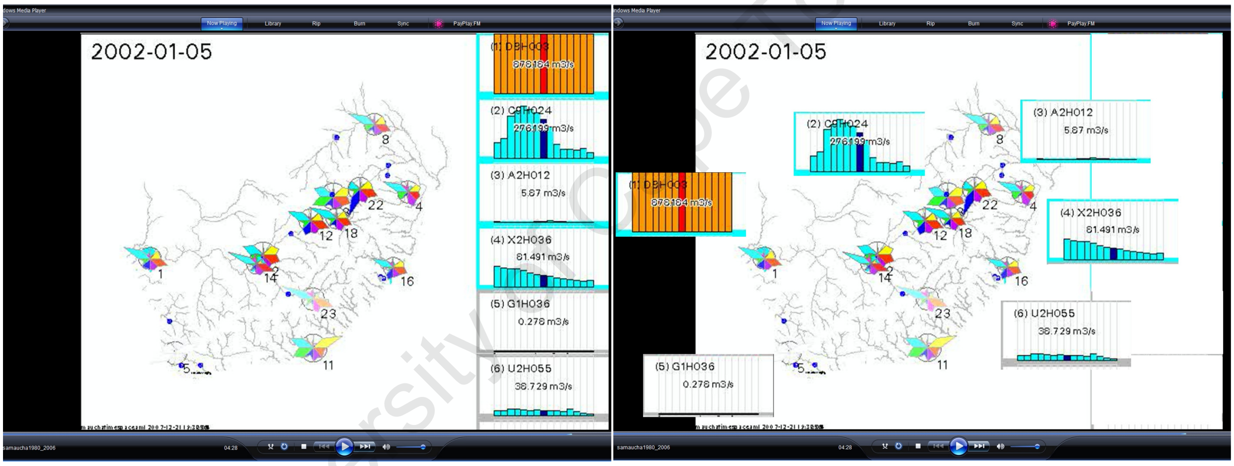


Figure 3-21. Proposed change to the animated map layout to reduce the visual and cognitive load of following an animation of water quality Maucha diagrams and flow hydrographs (samaucha1980_2006.mpg on enclosed disk).

For change monitoring, whether using animations or small multiples, the Maucha diagram's advantage over the pie diagram is that transitions for each variable are linear along the radii, rather than circular around the centre point (in other words, the symbol pulses in and out with changes in variables, rather than wobbling back and forth about its centre point). This property is the inherent advantage of all "rose" diagrams over the more common pie (Feitelson 2003, Wainer 2007).

3.4.3 Taking the Maucha diagram further

Minor tweaking includes the possibility of using the spare "$CO_3^{=}$" petal at the top left of the Maucha diagram to represent compounds like NO_3^- that are sometimes important components of TDS in certain studies. SiO_2 is also a candidate, but it contributes to the TDS in a more complicated way (Hem 1985) that might make its inclusion as an anion problematic. And although the Maucha diagram has already proved its worth as a compact multivariate water chemistry marker in Google Earth (Chapter 4, Silberbauer and Geldenhuys 2008, 2009), further refinements are required to give it a more finished appearance in the slick Google environment (for example, having a transparent symbol background to avoid the white block around each point marker).

The discussion has focussed on placing information in geographic and temporal space, but there is nothing to stop us representing other spaces, for example ionic space. Firstly, some general observations about the fundamental difference between the ways surface water and groundwater specialists regard their data. The surface water scientist works on the assumption that water quality data at sites on the river network represent water at that site and perhaps a little way up and down the network from that point. The geohydrologist views underground water as large, continuous bodies influenced by rock chemistry, the chemical quality of which he or she can legitimately interpolate in two dimensions (and perhaps three).

Following this train of thought, we can map the surface water Maucha salinity symbols to the groundwater Piper diagram of Vegter (1995), in effect creating a Rosetta stone to help scientists make the cognitive leap to and fro between the ternary and radial representations. The visualisation in Figure 3-22 illustrates this with a map using two independently perceived data symbols. Shading gives the "continuous" groundwater data boundaries and point symbols show the river monitoring site data. The viewer (after thinking carefully about the symbols) can perceive two sets of data simultaneously with less confusion than if both were point symbols or both shaded (Ware 2004). Even a highly cluttered image (Figure 3-23) provides useful information.

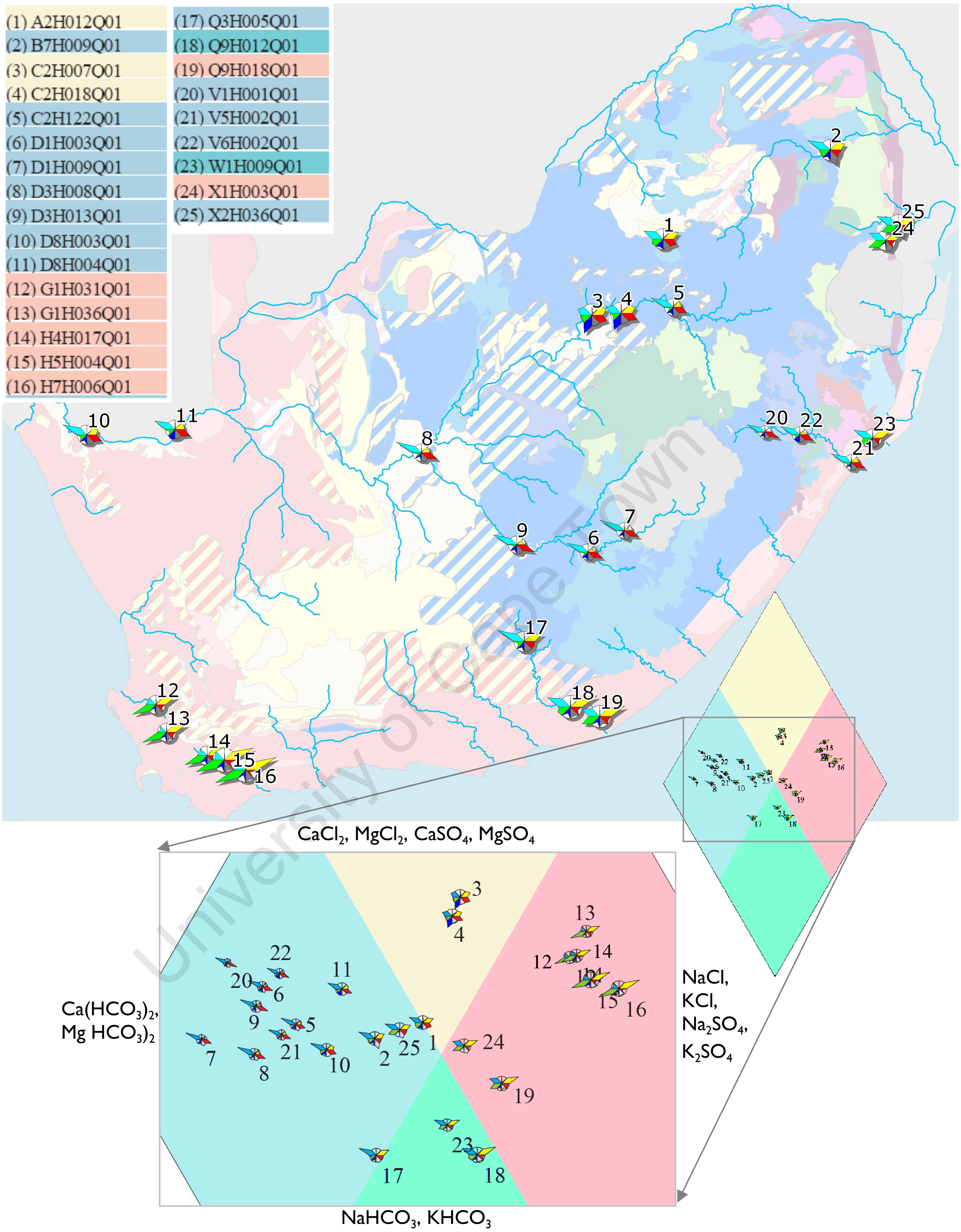
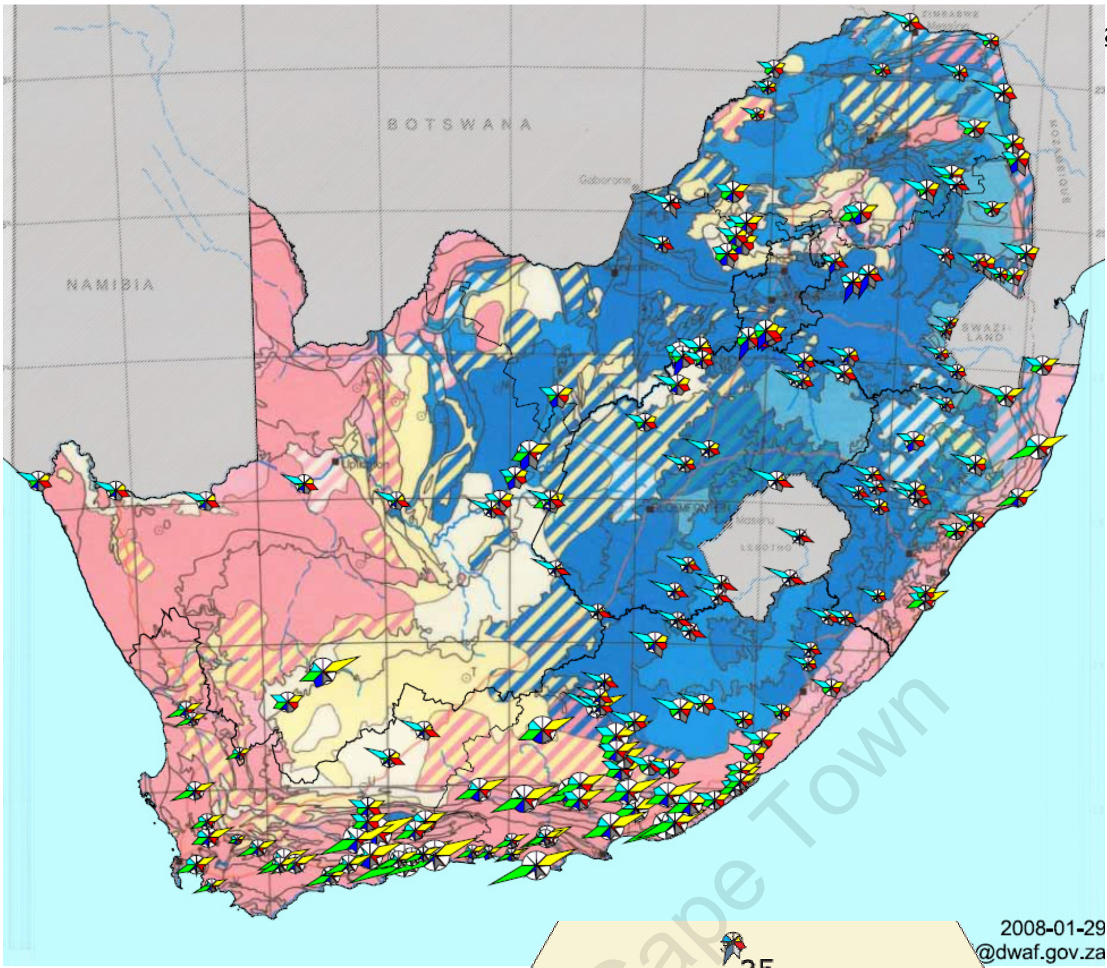
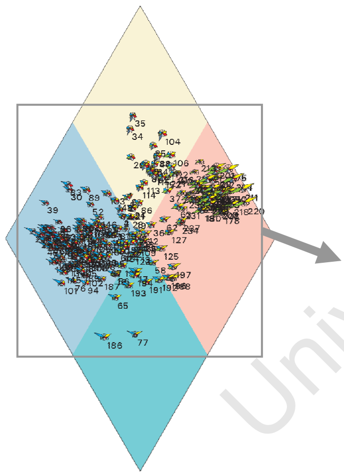


Figure 3-22. Mapping of Maucha (1932) radial symbols to the Piper ternary graph, showing clustering of Maucha symbol shapes on a Piper diagram according to the ionic makeup of each water quality sample. The colours of the shaded lozenges represent the shaded groundwater regions on the map. Map and Piper diagram after Vegter (1995), but generated independently using ArcGIS VBA for the map and ArcInfo AML for the piper diagram. Colour differences and overlapping labels are limitations inherent in the process of synthesising images from different hardware and software platforms. I have made manual adjustments to colour and label position for clarity.





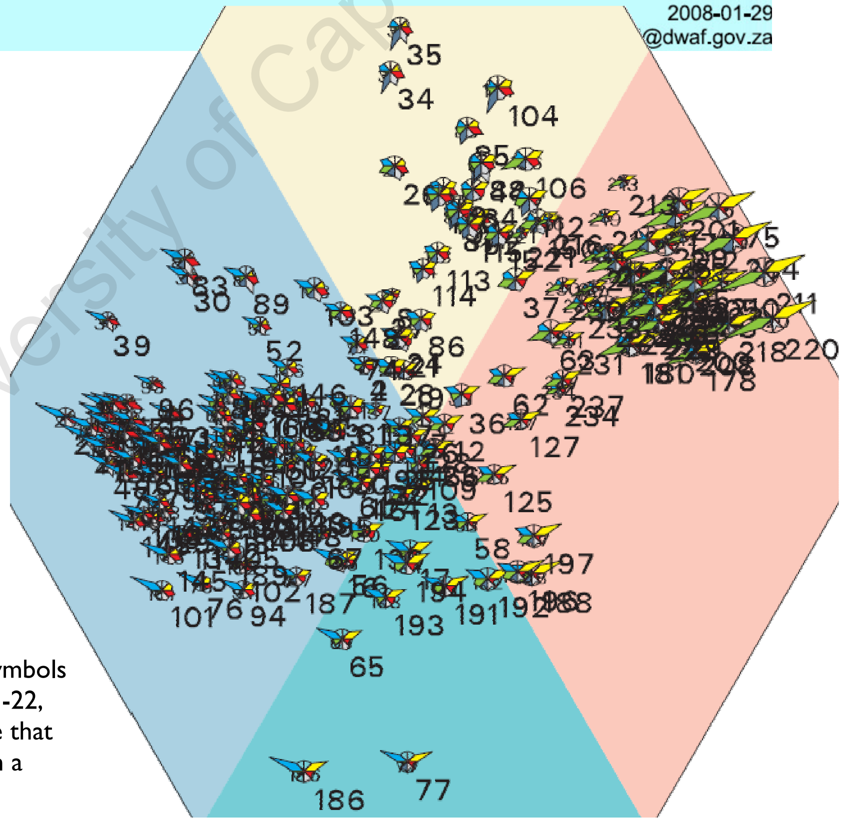


Figure 3-23. Mapping of Maucha symbols to the Piper diagram as in Figure 3-22, using a larger data set, to illustrate that general patterns are still evident in a cluttered image.

3.5 Conclusion

The construction of the Maucha (1932) diagram is more difficult than that of the pie diagram, which is ubiquitous in off-the-shelf software, but the effort is worthwhile if it provides improved at-a-glance understanding of salinity variables. While also technically more difficult to construct than the Stiff (1951) diagram, the Maucha diagram has the advantage of radial compactness. Furthermore, Stiff developed his symbol for petrochemical groundwater applications while Maucha's aim was closer to our applications, in representing aquatic habitat suitability. The Maucha diagram is very specific for the ionic ratio application, and is simpler to interpret than other radial diagrams in this context.

A similar approach to the programming methods for constructing a Maucha diagram can also be applied to generate constant-angle pies with as few as four or even two segments, to avoid the overlap problems caused by using multiple copies of a point layer to achieve the same effect. In fact, the built-in methods available for point symbolisation with commercial and open source geographical information systems are surprisingly limited. At the very least, a star or radar symbol option should be standard, for everything from seasonal wedge charts to wind rose diagrams. The symbols should also obey the rules for labelling that automatically arrange symbols to avoid overlaps while inserting leader lines to the point from which they moved. These automated symbol manipulation abilities would, if implemented, allow specialists to concentrate on data presentation rather than obscure trigonometric programming.

The method for drawing the Maucha diagram may have application to symbolic representations of other features, for example algal species distribution expressed as a fraction, or trace elements in clinker samples (Tamás and Abonyi 2002). Theory would suggest that 12 rays are about the maximum that we can discriminate (assuming that we can distinguish a 30° difference between rays—Ware 2004) so symbols looking like sea urchins or cacti (e.g. Figure 3-3) need to supply other cues, such as colour-coding or clustering of results, to provide useful information.

In the case of Maucha diagram time series, the ability to view the complete trend in a set of small multiples provides different information to the recognition of moving patterns in an animation. The user is able to view the state of the system at different intervals without scrolling back and forth. In the next chapter, we will look at ways of allowing the user to explore the spatial dimension of the water quality data set.

Chapter 4 - Using customisable Internet-based 2D and 3D mapping systems to put water quality on the map

Note on originality

This chapter is based on a paper by Silberbauer and Geldenhuys (2008, 2009), omitting the work of Geldenhuys, who used the Delphi programming environment to embed keyhole markup code generation into the Water Management System. Reference to his work occurs at the appropriate places.

4.1 Introduction

To appreciate the importance of South Africa's water-quality monitoring network, and the need for a spatial inventory, some background information on the region's climate, hydrology and economic development is necessary.

Rainfall over South Africa is in the low to moderate range with an average of about 480 mm per year (Figure 4-1a). Annual evaporation rates increase from the southeast to northwest and mostly exceed 2000 mm (Figure 4-1b). The runoff generated from rainfall is generally less than 50 mm per year, representing a yield of 9% against the world average of 35% (Allanson 2004; Schultze 2006; Schultze and Lynch 2006). Rainfall is also very seasonal and river flow is often low or non-existent in the dry season, even in major rivers such as the Limpopo.

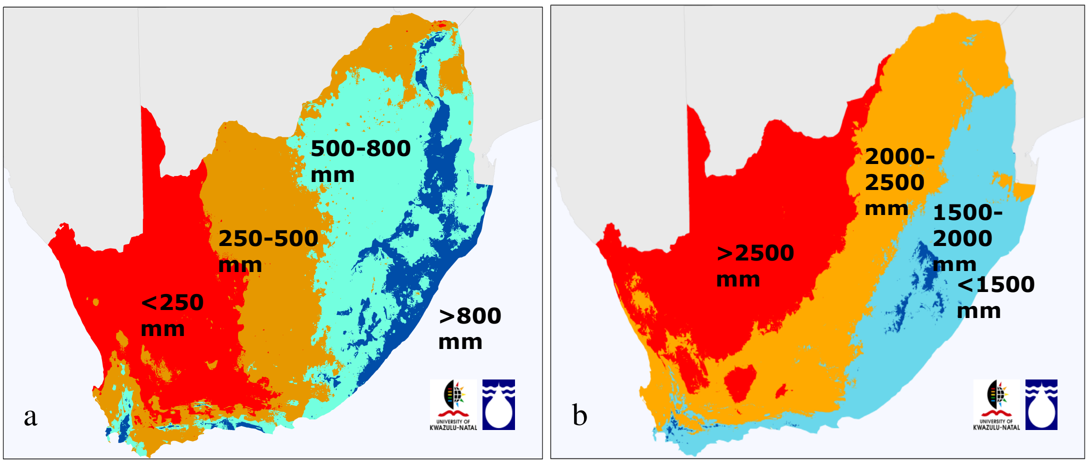


Figure 4-1 a) Mean annual rainfall for South Africa (data from Schulze and Lynch 2006); b) annual evaporation for South Africa (data from Schulze and Maharaj 2006).

This hydrological regime severely limits the capacity of South African surface waters to absorb the pollutants that are an inevitable consequence the country's tremendous growth in agriculture, mining, industry and urban population during the past century. Scientists sounded a warning during the 1950s that the quality of the water in streams draining Johannesburg was already poor and that the main impoundment to the north of the city, Hartbeespoort Dam, had become eutrophic because of high nutrient inputs (Cholnoky 1958; Allanson 1961; Allanson and Gieskes 1961).

In response to concerns about water quality, the government department that administers national water resources (Department of Water Affairs and Forestry, DWAF) stepped up the national sampling effort during the 1960s. The number of samples and the types of analyses soon became too much for a conventional wet chemistry laboratory, and DWAF established a partially automated laboratory in the mid-1970s, with a laboratory information management system feeding data to the national hydrological database. Procedures have since undergone many updates and the data now reside on the Informix-based Water Management System (WMS). The WMS includes tools for database administration, monitoring programme management and report generation (in the form of tables, flat files and maps created in ArcView 3.2 with Avenue). The national network now consists of many thousands of sites, some sampled routinely and others in response to events. DWAF's

Resource Quality Services directorate administers the network under a number of national programmes dealing with water chemistry, eutrophication, environmental health and microbiology.

As the database grew, DWAF water quality managers saw the need for an easily accessible inventory of monitoring sites and data, together with a facility to produce tabular summaries, maps and data sets for water quality modelling. A team of aquatic scientists and technicians hand-drafted a comprehensive map of the Vaal River basin water quality from 1979-1983 (van Vliet and Nell 1986) and a massive paper inventory of 1652 sites using a combination of mainframe database and personal computer Pascal (Swart *et al.* 1991). A parallel report on water quality at 683 sites also included the first GIS water quality maps (van Veelen *et al.* 1990). During the 1990s, rapid developments in GIS, remote sensing, desktop computing power and data distribution via the Internet enabled the development of graphical interfaces to water quality data, such as the ArcInfo-based WaterMarque (Cobban and Silberbauer 1993). DWAF staff also took the first steps in publishing data on the web.

By 2005, the maturing and convergence of the technologies for GIS, image processing and the Internet had resulted in a number of global earth viewers becoming widely available, notably the remarkable Keyhole Earth Viewer system released as Google Earth (Crampton 2008). One of the developers has described its unique image algorithms as "feeding an entire planet piecewise through a straw" (Bar-Ze'ev 2007). The convincing perspective relief view of satellite imagery draped over a digital elevation model of the surface of the Earth, animated using motion parallax, resembled a computer-game rather than a GIS and attracted the attention of a much wider audience than previously exposed to GIS, especially as it works on fairly ordinary consumer hardware. The facility for customising the interface with Keyhole Markup Language (KML) made the system useful to GIS specialists as well.

Soon after Google released Google Earth, I came across the Keyhole Markup Language (KML) tutorial (Google Code 2008) and knocked together a rapid prototype, which convinced me that the platform would provide an astonishingly fresh perspective on water quality monitoring networks and data access. Management information such as relationships between sources of pollution, water users and catchment topography were so clear that I decided to develop a fully-fledged system as described in the next section. The application solved a number of GIS problems at once: how to create perspective views, drape satellite images over topography in real time, investigate temporal changes in monitoring patterns and provide a geographically based inventory to internal and external clients. Best of all, no tedious motivations were required to obtain the software—one simply downloaded it.

4.2 Methods

This section describes the general processes followed in using Google Earth's extensible markup language, KML, for the construction of a visual monitoring site inventory with metadata pertaining to location, data available and period of activity. The version of Google Earth at the time of testing was 4.3.7191.6508 (beta). For details of scripts and code, please see Figure 4-3 and the enclosed disk.

4.2.1 Data extraction and preparation

The inventory described here is on a static web site, updated annually, where users can display available sites for regions of interest and bring up metadata, canned graphs and associated information for each site. Note that an alternative method, part of the WMS, gives users the option to visualise selected sites in Google Earth from within the WMS database (Silberbauer and Geldenhuys 2008, 2009).

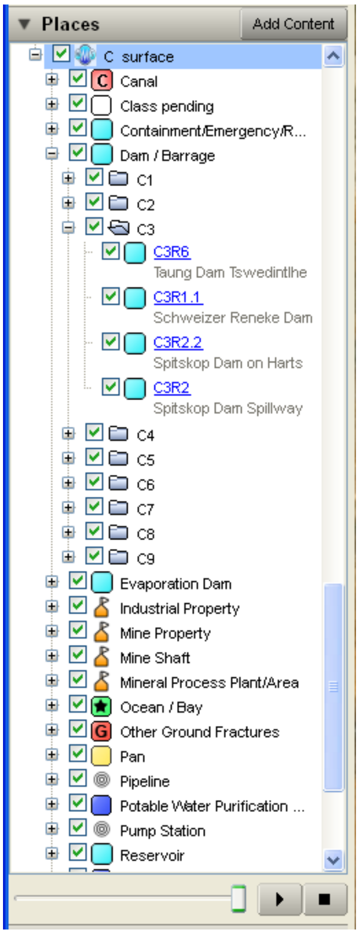
The inventory requires two flat files exported from the WMS: the first is a list of the more than 40 000 active sites with identifier, coordinates and description and the second is a list of the 500 000 sample results (one set of results for each date, time and depth). Utilities adapted from the WaterMarque project mentioned above import the sites into a point layer and the data into an INFO database.

4.2.2 Generating the monitoring site KML files

With no experience of KML, a useful approach to get started was to set up a layout manually with a few sites in Google Earth itself, setting the required symbols, folder structure and descriptions. When saved as a KML file, the code gives useful insights into the required syntax, enabling a more rapid learning process than starting from scratch with the on-line manuals and tutorials (e.g. Google Code 2008). Note that

KMZ files such as those on the DWAF web site are simply KML files zipped for faster download. I used SecureZIP do convert KML to KMZ (PKWARE 2007) but any zip programme with a batch facility would work.

Using the gawk implementation of the awk scripting tool (Gnu 2008), and a sorted inventory list as input, I generated KML "point and balloon" code for more than 41 000 sites within South Africa's 22 primary drainage regions, structured in a logical hierarchy of site type, secondary drainage region and site code (see Figure 4-2a for an example using the C1 secondary drainage region). An alternative hierarchy has the 19 administrative Water Management Areas at the top level with drainage regions below. Standard map symbols for water use types are still under development, so I arbitrarily chose the most suitable icons from the extensive set supplied by Google (Figure 4-2b). Representative listings of KML code are in Figure 4-3, and the full listings are on the enclosed disc.



Site class	Symbol
normal	paddle/wht-circle-lv.png
unknown	paddle/wht-blank-lv.png
canal	pal5/icon50l.png
potable	paddle/blu-blank-lv.png
industry	pal2/icon10.png
hydro	paddle/blu-circle-lv.png
lake	paddle/ltblu-blank-lv.png
marine	paddle/grn-stars-lv.png
municipal	pal3/icon21.png
pipeline	shapes/target.png
sewage	paddle/ylw-blank.png
tunnel	shapes/donut.png
wetland	paddle/grn-blank-lv.png
pan	paddle/ylw-blank-lv.png
ground	paddle/g-lv.png
borehole	paddle/B-lv.png
borehole00	paddle/0-lv.png
borehole01	paddle/1-lv.png
borehole02	paddle/2-lv.png
borehole03	paddle/3-lv.png
borehole04	paddle/4-lv.png
borehole05	paddle/5-lv.png
borehole06	paddle/6-lv.png
borehole07	paddle/7-lv.png
borehole08	paddle/8-lv.png
borehole09	paddle/9-lv.png
borehole10	paddle/10-lv.png

Figure 4-2 a) Folder structure in Google Earth for DWAF's "C" drainage region, showing site types, secondary drainage regions ("C1", "C2", etc.) and site names; b) Complete list of icons used: most boreholes had 10 or fewer data records, so the number of records could be included in the icon.

In parallel with the KML production process, the awk script outputs a set of HTML site tables with data links. This is to provide a workaround for downloading .zip files and to give access to water quality information for users who do not have Google Earth. (A security feature prevents the direct download of .zip files from a link within a Google Earth balloon).

The Delphi process in the WMS water quality database generates similar KML files to those discussed above (Silberbauer and Geldenhuys 2008, 2009). The main difference in operation is that a user working in the WMS environment selects sites of interest and then activates the process to generate a KML document for those sites.

Users have occasionally requested pre-packaged fly-throughs of study areas for presentations. I have used FRAPS (Beepa 2008) to screen-capture short movie sequences from Google Earth and Windows Movie Maker (Microsoft 2004) to splice and edit them into a longer fly-through (example at DWAF-RQS 2008 and on the enclosed disk).

I used the "Export to KML" extension for ArcGIS 9.1 (Martin 2007) to create KMZ files of the hydrologically important rivers and drainage basin boundaries to serve as background information for South African users. The source file for rivers was the South African 1: 500 000 layer (Silberbauer 2006) and I subdivided the KMZ river files by stream order to reduce their size. An option exists to show river names only, because the 10 000 river arcs slow down the performance of Google Earth considerably (and the rivers are already visible in the background images anyway). For the same reason, drainage regions created from the DWAF layers are available in each of the four hierarchy levels separately, with the quaternary drainage region labels in a file on their own.

4.2.3 Water quality data summaries

While the inventory is useful for planning, what users ultimately need is the water quality information at each site. With limited skills available in Internet database access programming, the most cost-effective approach was to dust off two existing tools from the old WaterMarque project, Barcode (Chapter 2) and Maucha (Chapter 3) and create a set of static data summaries, text files and salinity symbols.

Data summaries are in the form of Barcode graphs, which are one-pagers of the major water quality indicators for a site (see Chapter 2). Sixteen bar graphs of major water quality indicators and a bar graph of flow data (where available) are stacked above one another with a common time scale (Silberbauer 1997). Summary statistics (Minimum, median, 90th percentile and maximum) for each variable appear on the right-hand y-axis and locality maps of monitoring sites at primary, secondary, tertiary and quaternary scale ensure that the plot serves as a stand-alone data summary. The application also writes the data to a text file that users can download and plot with their own spreadsheet or graphing software. Metadata about the source data of the plot appear in the heading and a text box. The Barcode macro runs in ESRI's Arc Macro Language (AML) using data in the INFO database and takes about one minute per monitoring site in batch mode on a Sun Fire V490 server (i.e. at least 28 days to populate the whole web site, barring server and power failures).

Maucha symbols (discussed in more detail in Chapter 3) portray the major dissolved ions in a water sample, showing at a glance the ionic balance and dominant ions (Maucha 1932, Silberbauer and King 1991, Day and King 1995). Designed to indicate habitat type, as discussed in Chapter 3, their radial format and scalability makes them a better choice for mapping than the more commonly used Stiff diagram, developed for groundwater studies by the petrochemical industry (Stiff 1951).

The first lines of a KML set the environment and top level folder name:

```
<?xml version="1.0" encoding="UTF-8"?>
<kml xmlns="http://earth.google.com/kml/2.2">
<Document>
<name>S  surface</name>
```

Marker and balloon styles are at the head of the script:

```
<Style id="hydroPlacemark">
  <IconStyle><scale>0.4</scale>
    <Icon>
      <href>http://maps.google.com/
mapfiles/kml/paddle/blu-circle-lv.png</href>
    </Icon>
  </IconStyle>
  <BalloonStyle>
  <text>$[description]</text>
  <bgColor>ffffffff</bgColor>
  </BalloonStyle>
</Style>
```

Geographic coordinates (longitude, then latitude):

```
<Point>
  <coordinates>28.015556, -32.515278,0
</coordinates>
</Point>
```

<Timespan> links the visibility of each site to the time slider:

```
<TimeSpan id="ID">
  <begin>1990-10-17</begin>
  <end>2007-12-11</end>
</TimeSpan>
```

Defining a site type folder (Rivers):

```
</Folder></Folder>
<Folder>
  <name>Rivers</name>
  <styleUrl>#hydroPlacemarkL</styleUrl>
  <open>0</open>
<Folder>
  <name>S1</name>
  <open>0</open>
```

A lowest-level folder entry, for a monitoring site:

```
<Placemark>
<name>S7H4</name>
<Snippet maxLines="1">At Area 8 Springs B
on Groot-Keirivier</Snippet>
```

The details within a popup balloon are coded in HTML, between <![CDATA [**and**]]>**:**

```
<description><![CDATA[
<img src="http://www.dwaf.gov.za/iwqs/wms/data/s70/s70_102568.jpg" align="right" />
<b>WMS S70_102568</b><br />
S7H004<br />
At Area 8 Springs B on Groot-Keirivier<br />
<i>Rivers</i> samples: <b>325</b><br />
1990-10-17 to 2007-12-11<br />
<a href="http://www.dwaf.gov.za/iwqs/wms/data/s70/s70_102568.pdf">graph</a>|
<a href="http://www.dwaf.gov.za/iwqs/wms/data/S_reg_WMS_nobor.htm#102568">data</a>|
<a href="http://www.dwaf.gov.za/hydrology/cgi-bin/his/cgihis.exe/StationInfo?Station=S7H004">flow</a>|
<a href="http://www.dwaf.gov.za/iwqs/wms/data/000key.htm">home</a>|
<a href="http://www.dwaf.gov.za/iwqs/gis_apps/maucha.pdf">Maucha key</a>|
]]></description> <styleUrl>#hydroPlacemark</styleUrl>
```


Figure 4-3. Examples of KML code performing the main functions of the Google Earth water quality inventory. The interested reader can extract full code listings by unzipping KMZ files on <http://www.dwaf.gov.za/iwqs/wms/data/000key.asp>.

4.3 Results

On the Department of Water Affairs and Forestry web site, internal and external users can now access a 3D globe interface to the WMS water quality inventory grouped by hydrological or management region, and can view flow gauging and dam level monitoring sites (Figure 4-4). Depending on their area of interest, users can decide whether to show surface water sites or groundwater sites and whether to symbolise sites by monitoring type icon or by Maucha salinity symbol.

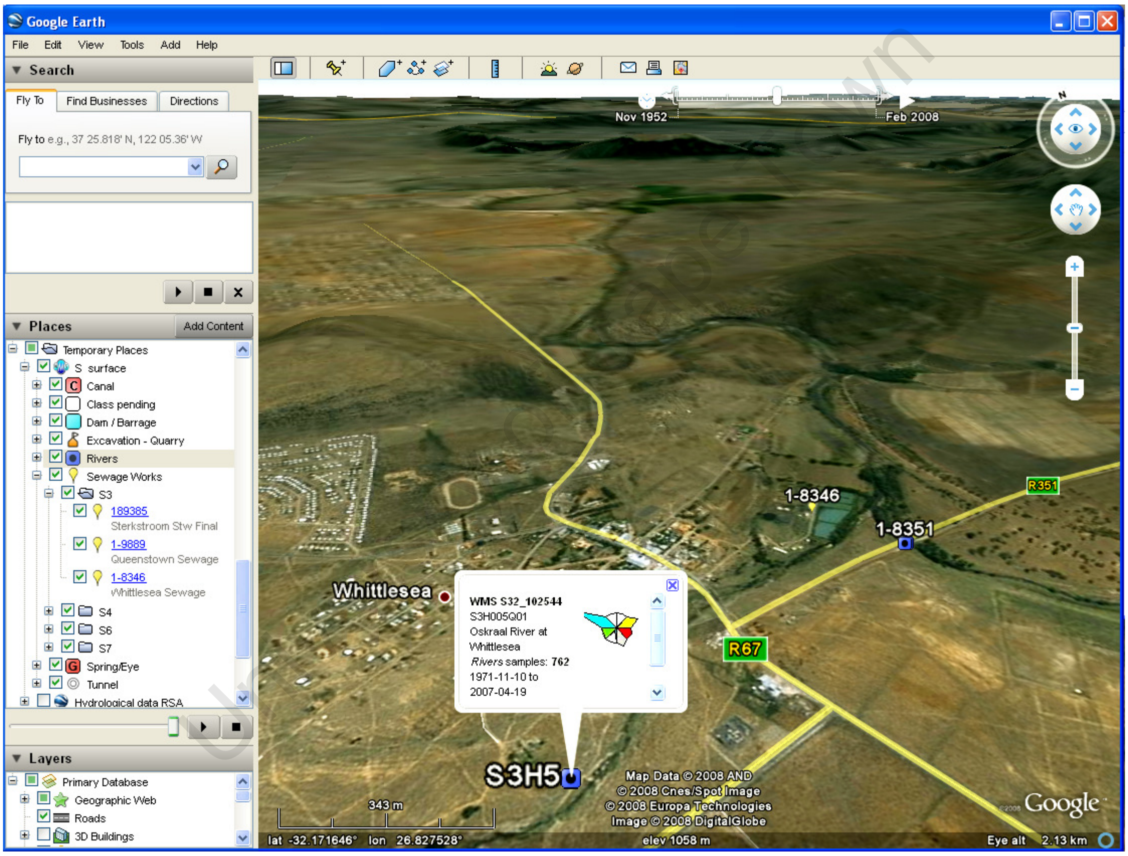


Figure 4-4 Water quality monitoring sites viewed in Google Earth, with a balloon at a river site including the Maucha symbol (main ions Na^{+}, Mg^{++} and HCO_3^{-}) and links to further information. URL <http://www.dwaf.gov.za/iwqs/wms/data/000key.asp>.

At each site, the user can click on the symbol and display a balloon with metadata (site description, number of results, time period sampled, ionic diagram) and links to further information (time-series graphs of major ions, raw data and flow data). Within the national water quality database, the WMS, users can quickly visualise the localities of a site or group of sites in Google Earth. With the tools developed for monitoring sites, one can rapidly develop specific views of sites on request, such as topographical map overlays, estuaries or environmental monitoring sites (an example of the application of these methods to estuaries is in Figure 4-5).

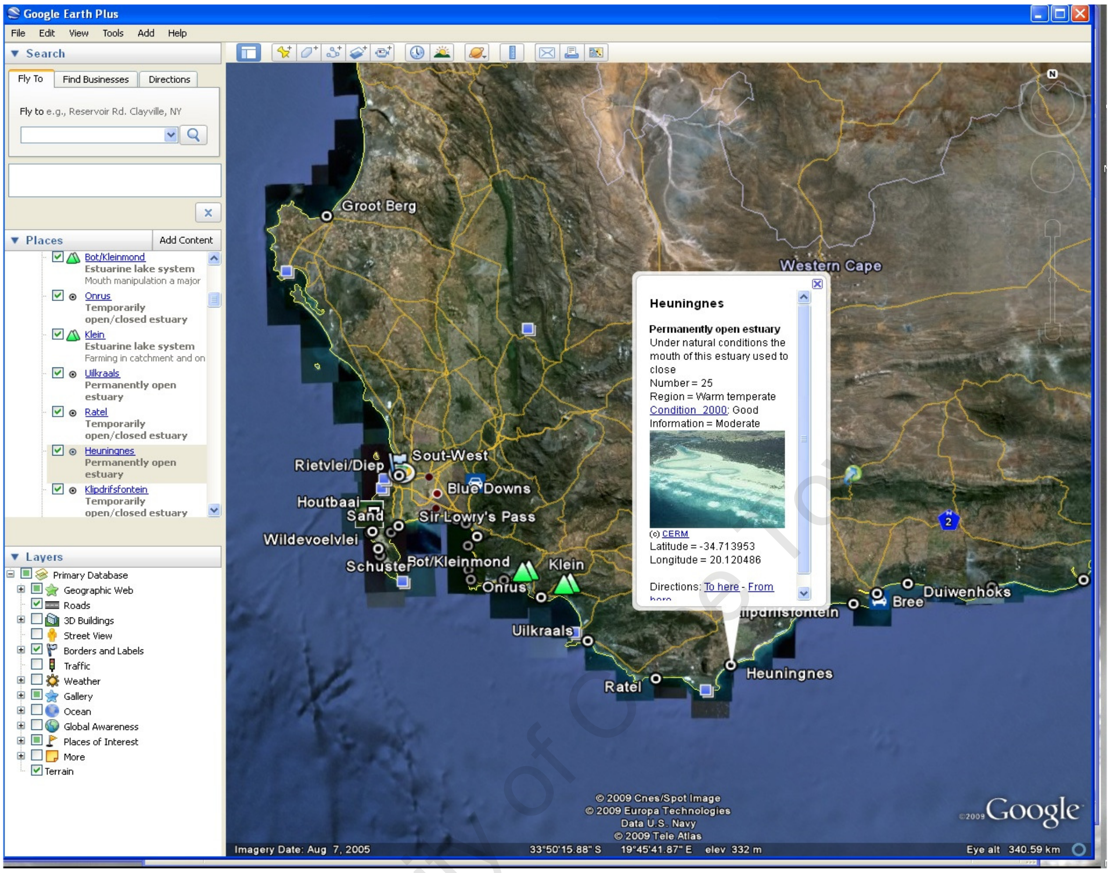


Figure 4-5. Inventory of estuaries (Whitfield 2003) converted to a KML file for use in Google Earth. URL < http://www.dwaf.gov.za/iwqs/gis_data/river/Estuaries.kmz>

The web site has proved to be quite popular, with at least 1000 unique visitors to the base page during its first year of operation (Figure 4-6). Most visits were from South Africa, mainly within the Department of Water Affairs and Forestry, though the system recorded 95 unique page views from the Americas, 82 from Europe and 39 from the east (Asia and Oceania).

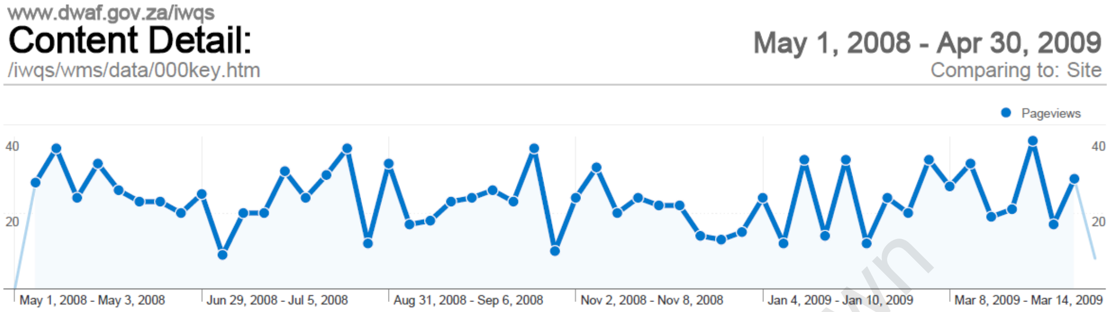


Figure 4-6. Number of visits per week to the inventory site for the year 2008-2009 ("new visitor" statistics from Google Analytics). These statistics give only a quantitative "feel" for the activity on the site, because most people actually using the KML files would download them to their computers and carry on working off-line.

4.4 Discussion

4.4.1 Customising Google Earth

Customisation of Google Earth is not difficult and the choice of KML generation method is immaterial. For this application, I simply used the scripting tools left over from past projects on my digital “workbench”. User requests are straightforward to implement, for example the requirement for a hierarchical folder structure and the separate display of borehole and hydrological sites from general water quality sites. The ability to create hyperlinks in HTML balloons allows for integration with many other data sources. However, certain links such as <file:///*.zip> and <mailto:> did not function as expected inside balloons, possibly for reasons of system security. The static data summaries (Barcode PDF plots) on the web site worked well, though users have to download the data from a separate HTML table if they wish to plot the data themselves with different date ranges or to show guideline levels.

4.4.2 Using Google Earth for a spatial inventory

The Google Earth interface is easy to use and allows rapid changes of scale from global to local and back. The water quality inventory interface has been extremely useful for finding available sites in a study area, for seeing where sites are in relation to one another, and for detecting sites where coordinates are in error. Usage tracking (Figure 4-6) has been difficult to interpret because internal testing and internal usage have the same network location, but personal communications suggest that users within DWAF and outside the organisation are satisfied with the way the inventory works. The effectiveness of Google Earth is evident in the number of developments of this type of application in other countries. Not surprisingly, other scientific organisations saw the advantages of customising Google Earth for their own data dissemination. GEMS/Water in Canada (www.gemstat.org) developed an interface to their international database and after submission of the first draft of this chapter as an abstract to the FOSS4G 2008 conference organisers, the Global Runoff

Data Centre in Germany began using Google Earth to map hydrological monitoring sites worldwide (grdc.bafg.de). Each group customised the Google Earth application in its own way, for example GEMS/Water and GRDC provide less information in the pop-up balloon at each site, rather taking the user to an interactive time-series graph composer. In France, the Institute de Recherche pour le Développement has also evaluated Google Earth for use in West African hydrological applications (Valero 2009) and the World Health Organization proposed the use of Google Earth for mapping tropical diseases in developing countries (Lozano-Fuentes 2008).

Google's massive centralised spatial database keeps the background data up to date with no effort on the part of the user: dam construction sites, pumps, sewage works are all there on the satellite images. The vector road network for South Africa displayed by Google Earth, in conjunction with the rivers and drainage regions from DWAF, ensure that hydrologists have no difficulty navigating through their study areas. Since 2005, the percentage of Africa covered by high-resolution images in Google Earth has increased greatly, giving an armchair view of land cover and surface structures not previously available so conveniently and on such a wide scale.

The built in functions and data are very welcome in a spatial analysis development environment where our department has insufficient skills and resources to develop geoportal systems from the ground up. For example, something as deceptively simple as the Google Earth "spread" function that displays overlapping points by offsetting them with leader lines, is extremely useful in identifying lazy coordinate capture, such as where all the sites for an impoundment are incorrectly located at one place. This operation is not easy to do on a conventional GIS.

4.4.3 Other ways to create KML files

For smaller, once-off data sets, off-the-shelf KML generation methods are more practical than the development of specialised scripts. One application that I have found to be useful for scientists without GIS training who need to view their own sets of coordinates in Google Earth is CSV2KML (CSV2KML 2007), an efficient, functional and free converter using spreadsheet data as input. Google Earth also provides KML creation tools on their site. Users find Google Earth less intimidating and simpler to configure than GIS, and after a brief introduction are soon able to continue on their own. Those who are more comfortable in the GIS environment can create geographic layers and export them as KML files (e.g. Martin 2007, for ArcGIS).

4.4.4 Other KML viewers

Many software developers, including the traditional GIS companies, have developed 3D earth viewers. ArcGIS Explorer (build 480) reads KML and KMZ files but opens the HTML text of a pop-up balloon in the default browser instead of inside the balloon. It does not spread overlapping symbols, scale labels, or interpret time data and has a more rigid GIS feel to it than Google Earth. World Wind 1.4 is supposed to have a KML interpreter but I could not get it to work with my KML files. The Microsoft Live Maps system automatically picks up and displays KMZ files, including those on the DWAF website, on the Internet under the Live Search mapping service.

4.4.5 Problems

One limiting factor in deploying this type of application more widely is the digital communications infrastructure in Africa. Although very ingenious algorithms optimise Google Earth's download rate, the slow bit rate is still frustrating for the dial-up Internet user. In 2002, Africans made up about 1% of the world's Internet users (Oyelaran-Oyeyinka and Lal 2005). By 2006 this fraction had reached 2.3% (Fuchs and Horak 2008), and although South Africans were relatively better off, fewer than one in eight had any access to the Internet at all. By 2007 less than 5% of South Africans had access to "broadband" connections (Miniwatts 2008). While I have verified that Google Earth is able to operate over a dialup link, thanks to its advanced caching ability, some of the immediacy that makes it so effective is lost when the data download rate is too slow.

Users not skilled in spatial analysis understandably have difficulty in distinguishing between Google Earth and the KML water quality database front-end, delivering unwarranted praise or derision depending on how they experienced the system. Some concerns include the perceived "age" and "fuzziness" of the images, the occasional mis-registration of a scene and illogical system errors (for example, failure by the operating system to recognise KMZ files after certain system updates).

4.4.6 Future developments

Eventually, our group at DWAF hopes to develop a complete and stable set of customised water use icons for use in the WMS, GIS and KML environments. For future applications, an enterprise-based system might be more versatile for internal users, enabling live database queries from links in the popup balloons. The time slider is suitable not only for showing when stations become commissioned and decommissioned, but also for following changes in water quality status from year to year. Figure 4-7 shows snapshots from an animation of mercury water quality classification changes with time, using the Google Earth time slider.

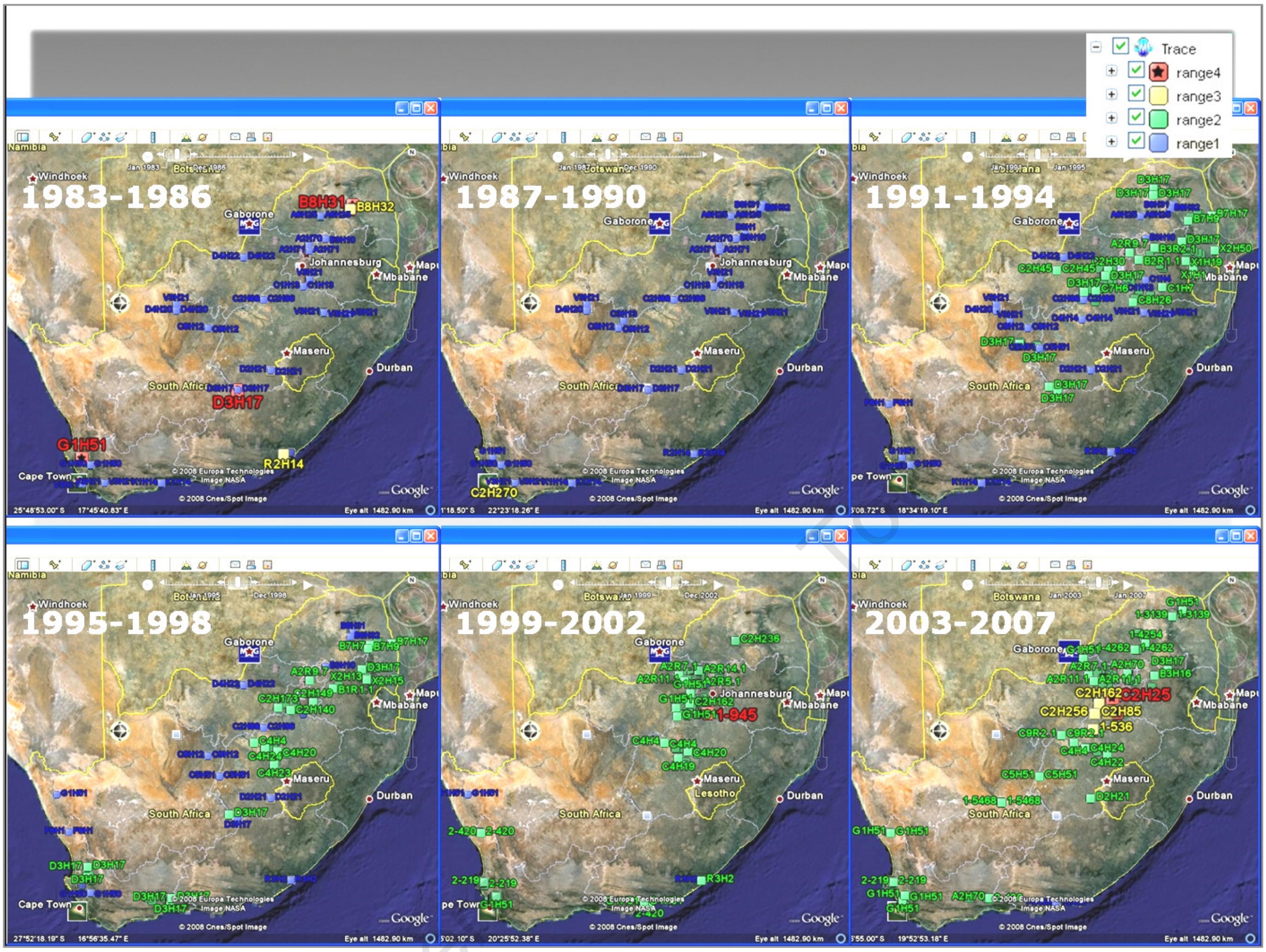


Figure 4-7. Snapshots of mercury water quality classification changes with time, using the Google Earth time slider (data from DWAF).

4.5 Conclusion

Google Earth provides a remarkable platform for a practical visualisation system for monitoring sites, creating a convincing, immersive view of the 3D surface hydrology network for many types of data users, including aquatic scientists, anglers and students. The background satellite and road data sets are far more up to date than any small organisation with a limited budget could hope to achieve. Properties of the sampling network that have become lost as the corporate memory fades, such as gaps or duplicate monitoring, become obvious when presented in a scalable interactive system. Google Earth epitomises the designer's information-seeking mantra:

> Overview first, zoom and filter, then details on demand
> Overview first, zoom and filter, then details on demand
> Overview first, zoom and filter, then details on demand
>
> -Card *et al.* 1999*:* 625.

At a more fundamental level, this kind of convincing graphical representation of a visual environment helps to convey information directly to our well-developed perception systems. The system creates "an actively constructed, meaningful model of the environment that allows perceivers to predict what will happen in the future so that they can take appropriate action and thereby increase their chances of survival" (Palmer 1999).

This chapter has hinted at the division of hydrological discharge data and water quality data into two completely different data sets. The next chapter will return to specifics and discuss how to combine these datasets to extract further information about the mass of solutes moving down river networks.

Chapter 5 - Combining concentration and flow as loading on a visual calculation sheet

Note on originality

Except where citations explicitly state otherwise, the procedures and graphics described in this chapter are my own work.

5.1 Introduction

The discussion so far has centred on representing the concentration of solutes in rivers and water bodies. An estimate of concentration is sufficient for some resource management applications such as determination of the ecological reserve for a river in terms of the National Water Act (e.g. Malan *et al.* 2003), because we are dealing with a continuous body of water. For other applications, such as dam eutrophication control (e.g. Enongene and Rossouw 2007), the actual mass of solutes and suspensoids is important because the river is in effect a conduit for reagents to a large mixed reactor. The mass transfer or load of a substance is thus a measure of the potential effect that it may have on other water bodies. For example, the load of phosphorus to a reservoir will affect the severity of eutrophication. Heavy loads of sodium chloride from a pollution source to a river can contribute to agricultural and municipal water supply problems for users many kilometres downstream.

Load has units of rate of transfer, for example tons per year. Some specialists draw a distinction between "flux" as the loading rate (mass per time) and load as the mass alone (Richards, 1998) but to avoid a proliferation of terms I will simply refer to load. To simplify matters further, I will limit the discussion to surface stream flow while acknowledging that groundwater and atmospheric transfer may be important components in some areas.

Two main groups of methods for estimating the rate of mass transfer are the direct approach using river data and the indirect approach using catchment land-use information. The direct approach consists of some form of multiplication of concentration (mass per volume) by flow (volume per time) to yield mass per time. The indirect approach is the use of a balance sheet of land use surface areas in a catchment to estimate the export of various solutes using predetermined coefficients related to factors such as population density and type of agriculture. A third type of estimation method is available for the special case of suspended clay particles, involving the direct measurement of the amount of sediment accumulated in reservoirs (Rooseboom and Lotriet 1992).

The direct mass-times-volume method is the subject of this chapter. Its application in well-gauged catchments is a common way of estimating the coefficients for the indirect method (Grobler and Silberbauer 1985, Enongene and Rossouw 2007). Even a large number of gauging stations cannot adequately cover a country the size of South Africa, as has been emphasised with the Google Earth tools in Chapter 4. We will frequently have to estimate loads from the best available data.

Direct load-calculation requires a continuous flow record and regular water quality monitoring, especially during high flows. The Department of Water Affairs and Forestry has long-term river flow and water quality data for several hundred monitoring sites across South Africa. The Hydstra database stores flow data and the Water Management System stores water quality data; both databases have functions that enable the export of their data in tabular format suitable for use in other applications.

Direct load calculation usually involves the combination of a continuous flow-record with water quality data collected at discrete intervals. In this chapter, I describe a method for visualising intermediate data and results of load calculation on a single information sheet, so that the assumptions and simplifications in the process are easier to see. The first prototype dates back to the mid-1990s but development stopped when the Department of Water Affairs and Forestry purchased the Hydstra system, which has algorithms to calculate various loadings such as the United States Environmental Protection Agency's Total Maximum Daily Load (TMDL) specifications. However, the database administrators eventually decided that importing water quality data into Hydstra would introduce unnecessary complexities into an already stretched system. I have therefore dusted off the program again with the aim of using it to check the availability of the necessary data, test the method and generate a set of results for all sites with suitable data. It serves as a sifting tool to identify sites and catchments for more detailed study with advanced tools such as FLUX (Walker 1999) and various models, such as the load module of MIKE BASIN (DHI 2009). Note that the US Army Corps of Engineers has lately begun a project to convert FLUX to operate in the Windows environment as Flux32 (Dr Dave Soballe, pers. comm).

Detailed load calculations for a site on the Pienaars River near Roodeplaat Dam provide a useful example (Figure 5-1). The figure plots the raw flow data for the whole of 1979 and a detailed view for February of that year, illustrating the diurnal sewage discharge flow that follows a sinusoidal curve. The spreadsheet (Figure 5-1c) shows 40 hours of data collected during the storm event, using phosphorus concentrations to illustrate the behaviour of a non-conservative solute, plotted in Figure 5-1d. Multiplication of concentration by flow yields the rate of mass transfer. Integrating under the curve of total phosphorus (TP) yields the load. Figure 5-1f shows the general relationship between $\log_{10}$ concentration and $\log_{10}$ flow. Selection of low-flow data or high-flow data yields markedly different ratios, a fact that can be exploited to stratify the data and improve the results.

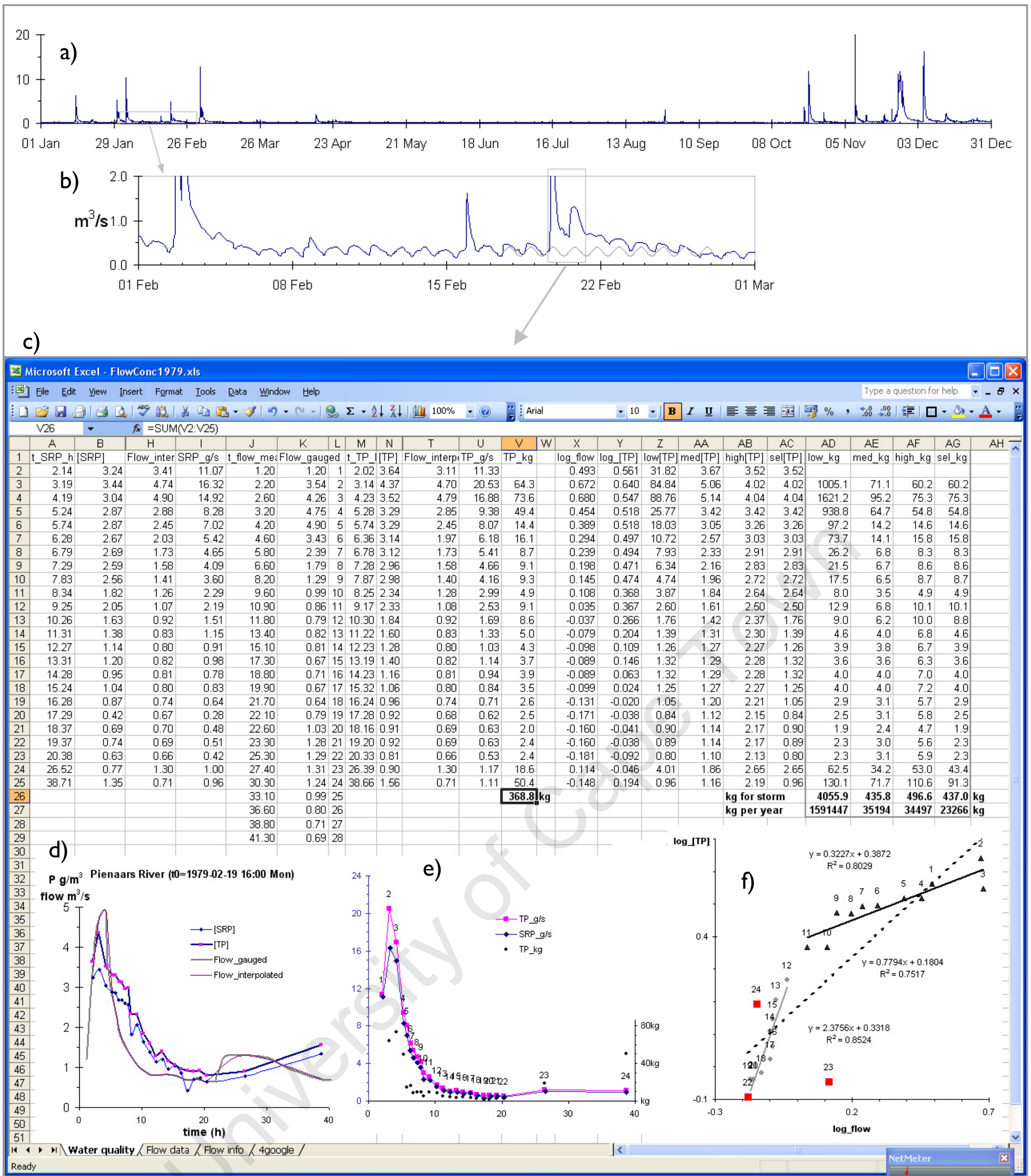


Figure 5-1. A detailed study of flow, concentration and load at site A2H027Q01 on the Pienaars River between Baviaanspoort Sewage Works and Roodeplaat Dam. Instantaneous flow data are from the Department of Water Affairs and Forestry Hydstra database and water quality data from an unpublished gauging of a 1979 storm event (Silberbauer 1979). The aspects to note are: a) the "inflection-point" (or "primary") flow data for the whole of 1979; b) a detailed view of the month of February, with the flow axis exaggerated to show the diurnal sewage discharge flow (simulated with a sinusoidal curve under the storm event. The spreadsheet (c) shows 40h of data for the storm event, with phosphorus data to illustrate the behaviour of a non-conservative solute (d). Multiplication of concentration by flow yields the rate of mass transfer (e) and integrating under the curve of total phosphorus (TP) yields the load (represented as dots in e). Plotting $\log_{10}$ concentration against $\log_{10}$ flow (f) yields a low-flow regression (grey line), a general linear regression (dotted line) and high-flow regression (black line). The three red markers are the three samples collected after the storm event, excluded for the regression analyses. The last four columns of the spreadsheet show the results of using each regression model to generate loads: low-flow, general, high flow and stratified (flow less than 1m^3/s uses the low-flow regression and flow more than 1m^3/s uses the high-flow regression).

5.2 Methods

The development platform was ESRI's Arc Macro Language, again for the reasons given in previous chapters: the availability of licences for the software, simple graphics commands, processing of much larger data sets than DOS-based computers could handle in the mid 1990s and the stability of the Unix platform. The data sets were the INFO database of chemical results described in Chapter 2 and flat ASCII files of mean daily flow rates extracted from Hydstra, the Department of Water Affairs and Forestry's flow data base. The program, load.aml, performs the operations listed here:

For the selected monitoring site, solute and time period, create a file containing all flow and concentration data matching the selection criteria.

For each concentration result, calculate the load as flow x concentration, and write the flow, concentration, load, $\log_{10}$ concentration and $\log_{10}$ flow.

For each year, perform a linear regression of $\log_{10}$ concentration against $\log_{10}$ flow and plot the result.

Plot the concentration as a time series (highlight interpolated data using lighter intensity, with lightness increasing in proportion to the number of days to the nearest actual measurement).

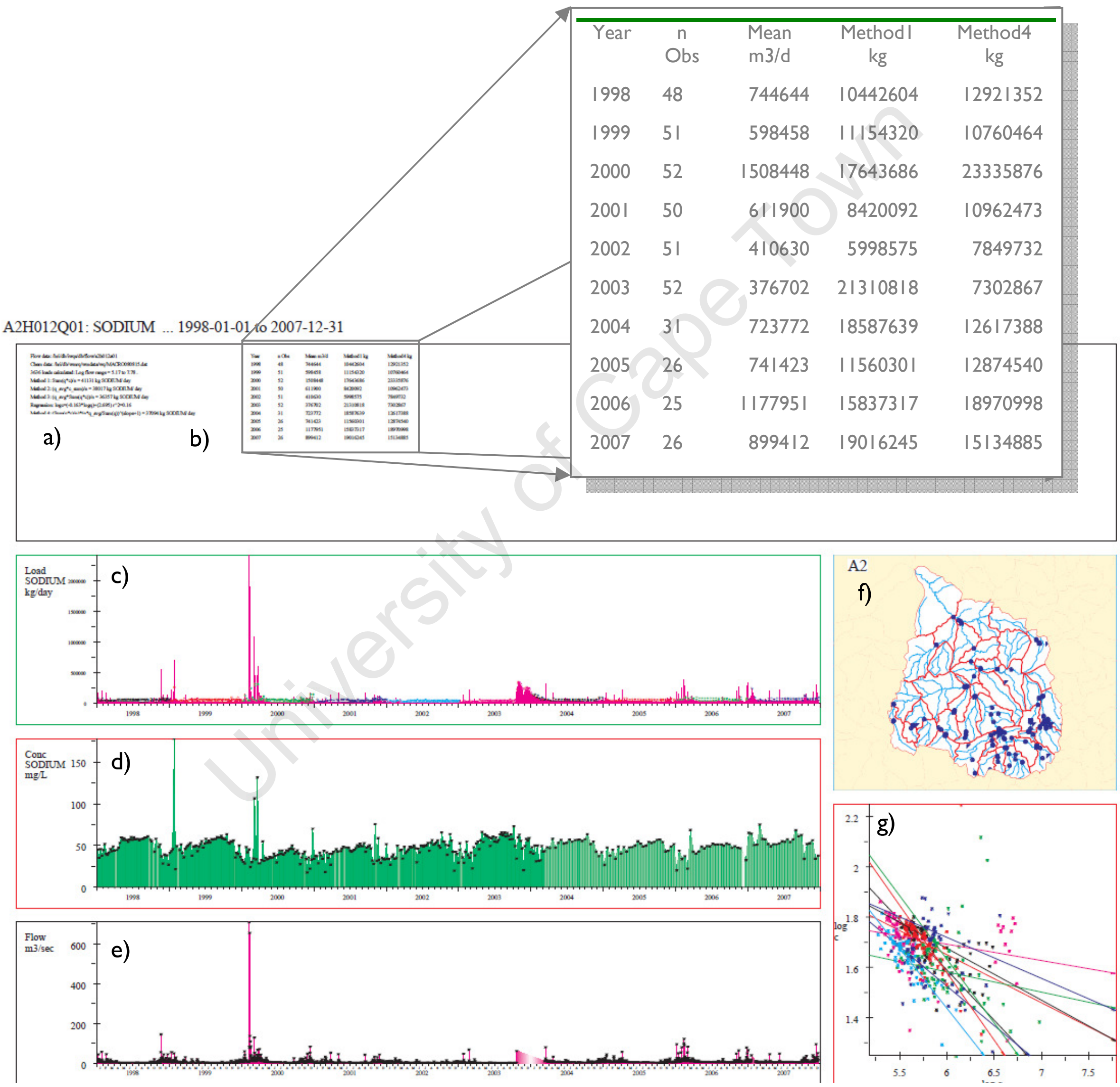
Plot the flow as a time series.

Plot the load as a time series.

Write a table of annual flows and loads.

5.3 Results

Figure 5-2 is an example of the layout of a sample load plot. In Figure 5-3 a number of plot images are shown as file thumbnails to illustrate how the program output enables the user to scan a large number of sites for a suitable candidate for more detailed investigation. Each ten-year analysis and plot took about ten minutes on a Sun Unix server.



Year	n Obs	Mean m3/d	Method1 kg	Method4 kg
1998	48	744644	10442604	12921352
1999	51	598458	11154320	10760464
2000	52	1508448	17643686	23335876
2001	50	611900	8420092	10962473
2002	51	410630	5998575	7849732
2003	52	376702	21310818	7302867
2004	31	723772	18587639	12617388
2005	26	741423	11560301	12874540
2006	25	1177951	15837317	18970998
2007	26	899412	19016245	15134885

Figure 5-2. Example of a load page for the monitoring site on the Crocodile River before it enters Hartbeespoort Dam. The regions are a) list of different calculation methods; b) table of annual loads (enlarged for clarity); time-series plots of c) load, d) concentration and e) flow; f) site locality and g) annual regression plots of log concentration against log flow, colour-coded to match (c). Method 1 is the flow x volume calculation and Method 4 uses regression. (A larger version of image appears in Appendix.) Note that the purpose of this summary page is to act as part of a data sifting process, and the data analyst would investigate the data in (g) in detail using other procedures.

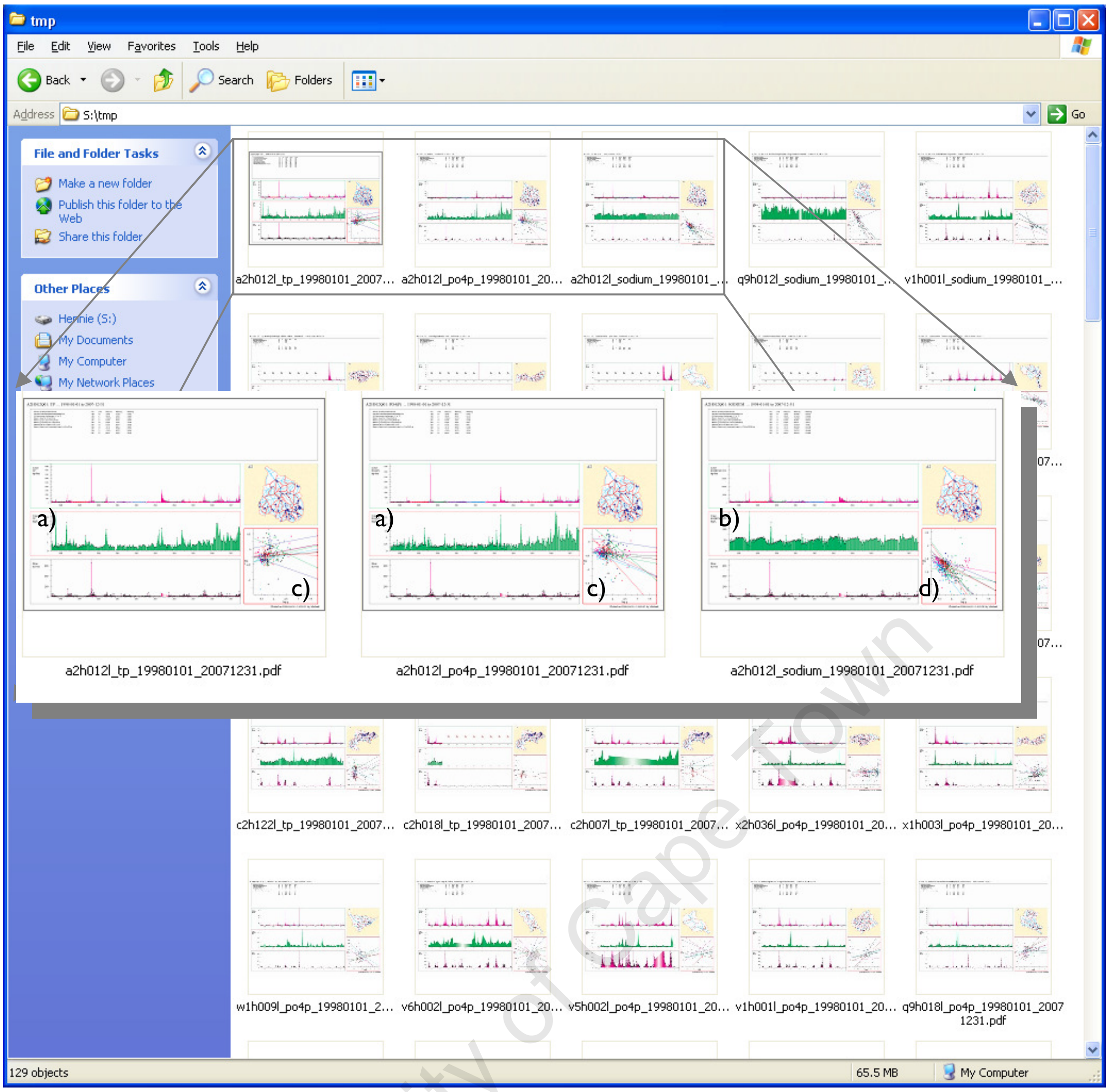


Figure 5-3. Sample thumbnail images of output from the load calculation script for sites across South Africa. The three enlarged images show how, with some practice, the user can pick up patterns in the data before examining each image in detail. a) The total and dissolved phosphate concentrations follow a long-term upward trend while b) the sodium concentration has an annual cycle. c) The total and dissolved phosphate concentrations show a variable relationship to flow, while d) high flows consistently dilute sodium as seen in Figure 5-2.

5.4 Discussion

What is striking about the flow-to-concentration ratio is how it varies between sites, varies for different solutes and changes from year to year. In an environment where large volumes of wastewater discharge to a drainage network with limited and unreliable capacity for dilution, catchment managers need to understand, monitor and manage the interaction between stream flow and solute concentrations. The most effective way is by direct monitoring, but this is also the most expensive in skills and technology. As a stopgap measure, experts in the field of load calculation can maximise the value of national investments in flow and quality monitoring by using the available data to produce export coefficients that catchment managers can plug into models for estimating the consequences of different management approaches.

The original aim of the load visualisation method described here was to provide a global, automatic load calculation method. However, it has fallen short of this ideal because of the uncertainties inherent in the behaviour of solutes in overland runoff and stream flow. This load method must therefore assume a more modest role as a visual method for screening large data sets in order to find patterns indicating sites with useful information for further investigation. For example, the type of slope in the log-log plot of concentration against flow (e.g. Figure 5-3c and d) can show whether point sources are an important factor (Walker 1999).

The load visualisation method suffers from the drawback mentioned in previous chapters for static graphical presentations, in that the user cannot intervene to investigate specific changes and nuances. In this regard, the approach used by the Gapminder software (Chapter 2 and Chapter 3) provides some inspiration as to how an interactive data exploration tool might look. A snapshot of the Gapminder-style motion chart with data from the sites from van Niekerk *et al.* (2008) appears in Figure 5-4, representing a very simplified data set consisting of annual means of solutes plotted against total annual discharge. The chart includes only about 700 records in all, because larger data sets do not work too well on Internet-based software at low data speeds. The coarse scale hides the subtle interactions of the type seen in Figure 5-1.

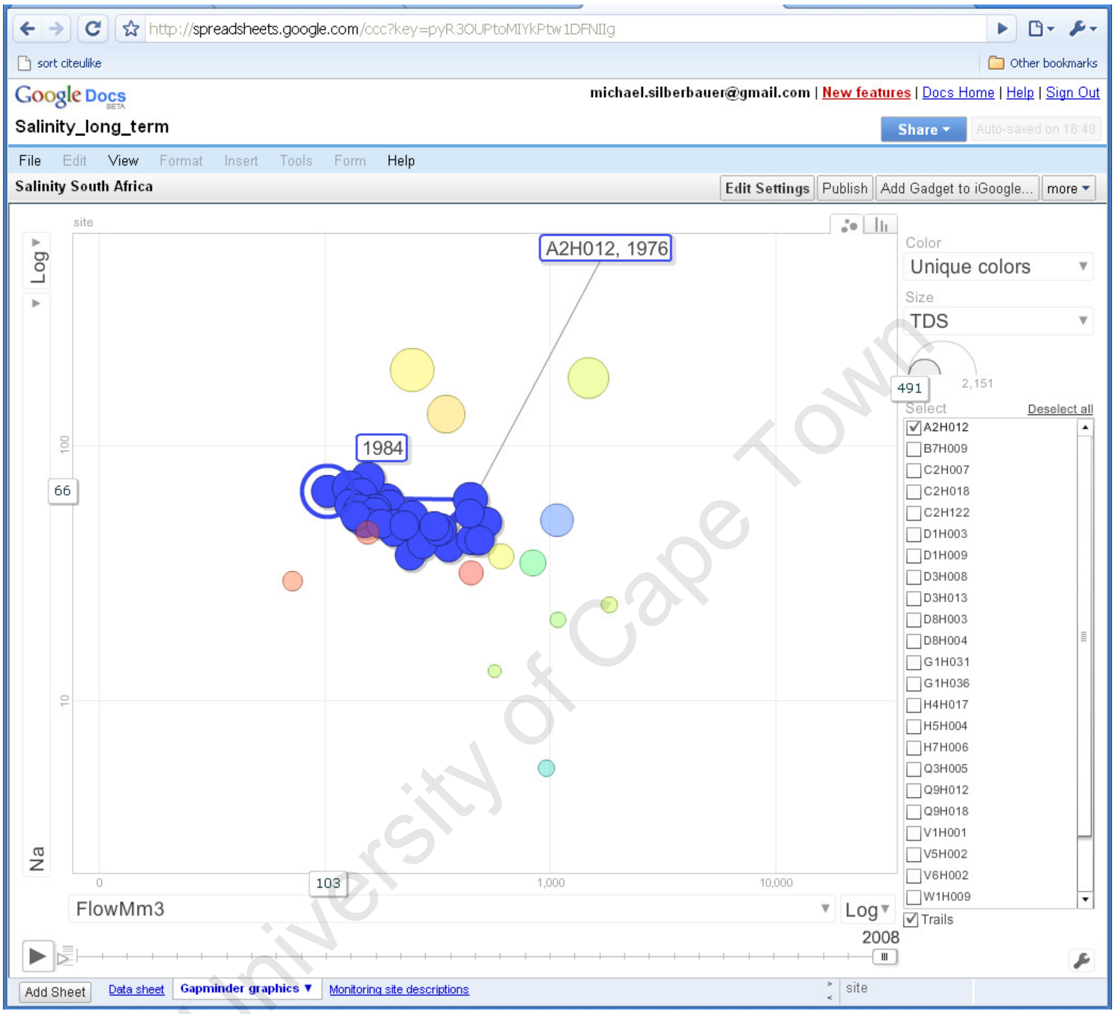


Figure 5-4. Google Documents spreadsheet Motion Chart with a log-log plot of mean annual sodium concentration in milligrams per litre against total annual flow in millions of cubic metres, highlighting the site on the Crocodile River upstream of Hartbeespoort Dam. The dots track the annual changes in the ratio of sodium to flow—at this site the sodium concentration varies little.

In the end, the final concentration in the “reactor vessel”, the dam, is what drives water quality problems such as eutrophication, purification costs and crop damage. If the catchment manager already has concentration data, the empirical answer is there, though with no predictive ability.

5.5 Conclusion

The load program generates a one-page summary of the flow, concentration, flow-concentration relationship, load and locality for a single solute at a single monitoring site. While the original intention was that the one-page report should be used in isolation, it has turned out to be more applicable as a data overview and sifting device when large numbers of pre-generated pages are displayed as thumbnail images. This batch processing ability that gives it the edge over interactive Windows software can save many hours of examining data in tabular form or generating graphs manually in an interactive system.

Load information is particularly important as a measure of the effect of land use changes on water quality. The combined analysis of flow and concentration data presents technical problems in the combination of two time series with different intervals from separate databases, with the likelihood that gaps will exist in both series. This load visualisation tool has only gone part of the way to assisting data users, and future developments should concentrate on giving users control over how they view changes in the concentration versus flow relationship through an interactive visualisation tool.

Chapter 6 - Combining visualisation and text into a comprehensive report

Note on originality

While I have developed the arguments of the discussion and analysis of the State of Rivers reporting process presented here, this chapter arises from a collaborative effort that aimed to produce a joint article in a scientific journal. Contributions by colleagues are indicated in the text. Wilma Strydom (CSIR) in particular, who led the production of many documents described here, has provided valuable criticism of my interpretation. I am also indebted to Dirk Roux (Monash South Africa) and Liesl Hill (CSIR) for their comments and discussion.

6.1 Introduction

So far, we have been dealing with scientists' internal and external visual models developed to help themselves and their colleagues understand the meaning of data sets. The focus now moves away from these individual, specialised visual presentation techniques to more ambitious attempts to create visual models of the aquatic environment in the minds of a wider audience. The example here is the Department of Water Affairs' National Aquatic Ecosystem Health Monitoring Programme, in particular the River Health Programme. This biomonitoring programme illustrates the practical problems involved in applying the theory of visual communication in a real-life, multi-disciplinary document-production environment. The programme's State of Rivers reports combine water resource visualisations of various types with text commentary to communicate the results of each regional River Health Programme study to a wide audience.

6.1.1 The River Health Programme

The River Health Programme is the most extensive of the South African National Aquatic Ecosystem Health Monitoring programmes, and has been functioning since 1994. The national Department of Water Affairs and Forestry (DWAF) administers the River Health Programme in collaboration with other organisations, notably the CSIR, Water Research Commission and provincial environmental departments (Strydom *et al.* 2006, Roux *et al.* 2008).

During their monitoring work, scientists involved in the River Health Programme have accumulated a wide range of data, from which the River Health Programme team as a whole can extract material and repackage it for a variety of audiences including water resource managers, politicians and the general public (Strydom *et al.* 2007). Several participants have reviewed the effectiveness of the presentation and distribution of River Health Programme information and have found that it is has been less successful than expected in influencing opinion and decision-making across the whole target audience (Strydom *et al.* 2002, Strydom *et al.* 2007, Strydom 2009a,b). In this chapter, I concentrate on the visual presentation methods developed for the most widely distributed products of the River Health Programme, especially the State of River reporting materials and comment on the effectiveness of this means of communication. The purpose of this study is to evaluate the visual communication process and suggest improvements where appropriate.

Common reasons for scientists to publish their work in technical reports and formal journal publications are to broaden the base of scientific knowledge and to subject their work to the scrutiny of their peers. Beyond these intellectual pursuits, scientists also have an obligation to communicate their research results to the wider audience of non-scientists in what ideally becomes an ongoing interactive process of "social learning" (Weber and Word 2001, Ison *et al.* 2004). The difficulty with this process is that the broad umbrella of "non-scientist" covers such disparate entities as the subsistence farmer, the municipal administrator and the corporate manager. Managers in business and government rarely possess qualifications in the natural sciences, yet their decisions can have far-reaching effects on the state of the environment. Furthermore, in a democracy public opinion carries much weight in deciding whether to implement scientific recommendations. Public participation is only fair, because individual and corporate taxes fund government research, and decisions around environmental management affect everyone living in that environment. Reporting effectively to the wide range of audiences involved is vital for informed decision-making at all levels (Strydom *et al.* 2007).

Preparing information for the public can be challenging for scientists, because they are used to communicating mainly amongst fellow specialists, and are immersed in the imagery and background of their own field (Nowotny 2003). Scientists may easily overlook how much additional information they need to provide if they are to engage the minds of non-scientists, or even specialists from a different field. The process becomes further complicated when the expected readership includes the previously-mentioned mix of water resource managers, strategic planners, policy makers and interested members of the public (Van Wyk *et al.* 2008). The value of the communication effort is only realised if the intended audience actually reads the reports, and has sufficient confidence in the credibility of the contents to use them, perhaps in public debate of environmental problems (Denisov and Christoffersen 2001). The information contained in State of Rivers reports must compete in an environment overshadowed by other more powerful sources, such as advertising, political propaganda and deeply-held beliefs. If the reports provide suitable information for helping policy-makers, managers and landowners to take measures to improve the state of the aquatic environment, the reporting process will have attained an important objective. The next step of translating knowledge into action is far more subtle, and has to take place in the minds of individuals living in the catchments (Strydom *et al.* 2007).

The task of shepherding information from the philosophical arena to the pragmatic takes skill and practice, and treading the fine line between patronising and incomprehensible prose requires care (and practice). "To be of true service to humanity, science must be an exquisite blend of data, theory and narrative" (Schermer 2007, a writer for *Scientific American*). The choice of analytical and visualisation methods contributes both to credibility (Tufte 2001) and palatability (Denisov and Christoffersen 2001).

"First, I am going to show you something very beautiful [shows music score on screen]: Nocturne, of Chopin. Isn't it beautiful music? But so very few people will appreciate it if I only show them the notes. Only a very few musicians can read the notes and say, 'Oh, this is beautiful music!' And I think this is often how we are, we who love statistics, who work with statistics; we often show the notes, we don't play the music."

--Rosling (2007b).

Guidelines for report production, from data collation and identification of target audience, to evaluation of the usefulness of the final product—drawing on the experience gained during the first decade of the River Health Programme—are available in the River Health Programme State of River Reporting Manual (Strydom 2003). The manual touches only briefly on the symbolic representation of habitat integrity, and in this chapter I deal in greater depth with the visual methods of the River Health Programme, noting its successes and shortcomings. Graphical representation of information is effective for conveying information and it is essential for the many readers of State of River reports who only have time to skim through the figures and tables, or who have limited reading skills (Strydom, 2003).

6.1.2 The philosophy of ecosystem status reports

The term "River Health" reflects the nature of the River Health Programme as an integrative approach to river ecosystem science (Saner 1999). In this process, aquatic scientists treat the river system as a virtual patient whose vital signs they can monitor, applying triage (classification) to decide whether the system's condition might improve with intervention, or if it is already too far gone. The "vital signs" in this context are validated ecological indices, feasible to measure at regular intervals and responsive to changes caused by human activities in the environment. River Health indices include the South African Scoring System for invertebrates (SASS – Chutter 1998), Macroinvertebrate Response Assessment Index (MIRAI – Thirion 2007) the Fish Response Assessment Index (FRAI – Kleynhans 2007), the Riparian Vegetation Response Assessment Index (VEGRAI – Kleynhans *et al.* 2007) and the Index of Habitat Integrity (IHI – Kleynhans *et al.* 2008). The River Health Programme team has continuously refined these methods while investigating additional indicators, for example geomorphology, water quality and diatoms (Strydom *et al.* 2006).

6.2 Preparing biomonitoring reports

This section describes the process of putting together a typical State of River report, and other products such as posters and brochures. It also notes changes in methods that have occurred since the reporting process began in the mid-1990s.

6.2.1 Colour-coded system status icons

The River Health Programme team experimented with a range of icons or pictograms to represent the ecological state of river systems. We eventually created a tailor-made set of symbols, because the icons readily available in publications and printer fonts were for foreign biological zones, with misleading visual references to exotic species such as conifers, salmon, marsupials and deer. The Results section (below) summarises the symbols currently used and some of their predecessors.

Most people rapidly distinguish and assess graphical symbols based on colour (Lewandowsky and Spence 1989). This is part of our low-level feature analytical system and is so hard-wired that we do not think about it (Ware 2008: 25). In order to be consistent, the River Health Programme team adopted the "traffic-light" colour scheme used by DWAF for classifying water quality according to suitability for use (DWAF 1998). This colour symbolism, while quickly interpreted, is not necessarily self-evident, nor universal, so every report includes a detailed key. Furthermore, colour vision varies between individuals, causing ambiguity in the interpretation of colour codes. (See section 1.4.1 for more on colour-blindness, lens aging and cataracts.) To reduce misinterpretation of River Health Programme data resulting from subtle colour perception differences, colour-coded icons in River Health Programme reports include confirmatory text (“natural”, “good”, “fair” or “poor”) and each individual icon contains the letter “N”, “G”, “F” or “P”. Table 6-1 has further details.

6.2.2 Railway diagrams of river networks

An idealised river differs from most ecosystems in that it is a linearly constrained aquatic habitat with a constant downhill movement of the habitat matrix. In river systems in the real world, though, water is not always present above ground and does not always flow, especially in an arid climate (Uys and O'Keeffe 1997). The linear model is nevertheless useful for simplifying the interrelationship between points in the catchment and on the river network. We therefore borrowed the approach used in railway diagrams in order to illustrate how disturbance and response in the river habitat is dependent on where they lie in the network. Here, the term "railway diagram" refers to the simplified city transport map, a form of linear cartogram where nodes and links take precedence over geographic accuracy, to help commuters find routes through complex urban transport systems such as the London Underground (Ware 2008: 15). In a similar way, the river network diagram emphasises the route followed by water under the influence of gravity, and shows possible links for migration of aquatic species.

6.2.3 Cartography

The River Health Programme team has been developing methods for displaying habitat integrity as symbols on maps since at least 1996, when I used a script in the Arc Macro Language (AML) to plot topographical maps with ecological status icons for the Crocodile River system in Mpumalanga (Figure 6-1 and DWAF-RQS 2009). The difficulties in applying geographical information systems (GIS) at that time included automatic placement of icons so that they did not overlap and the complexities of constructing vector marker symbols (Cobban and Silberbauer 1993). State of River report editors eventually resorted to manual desktop publishing methods for placing text and icons on maps, using GIS for the base layer only, because this approach gave much finer control over the quality of the final product.

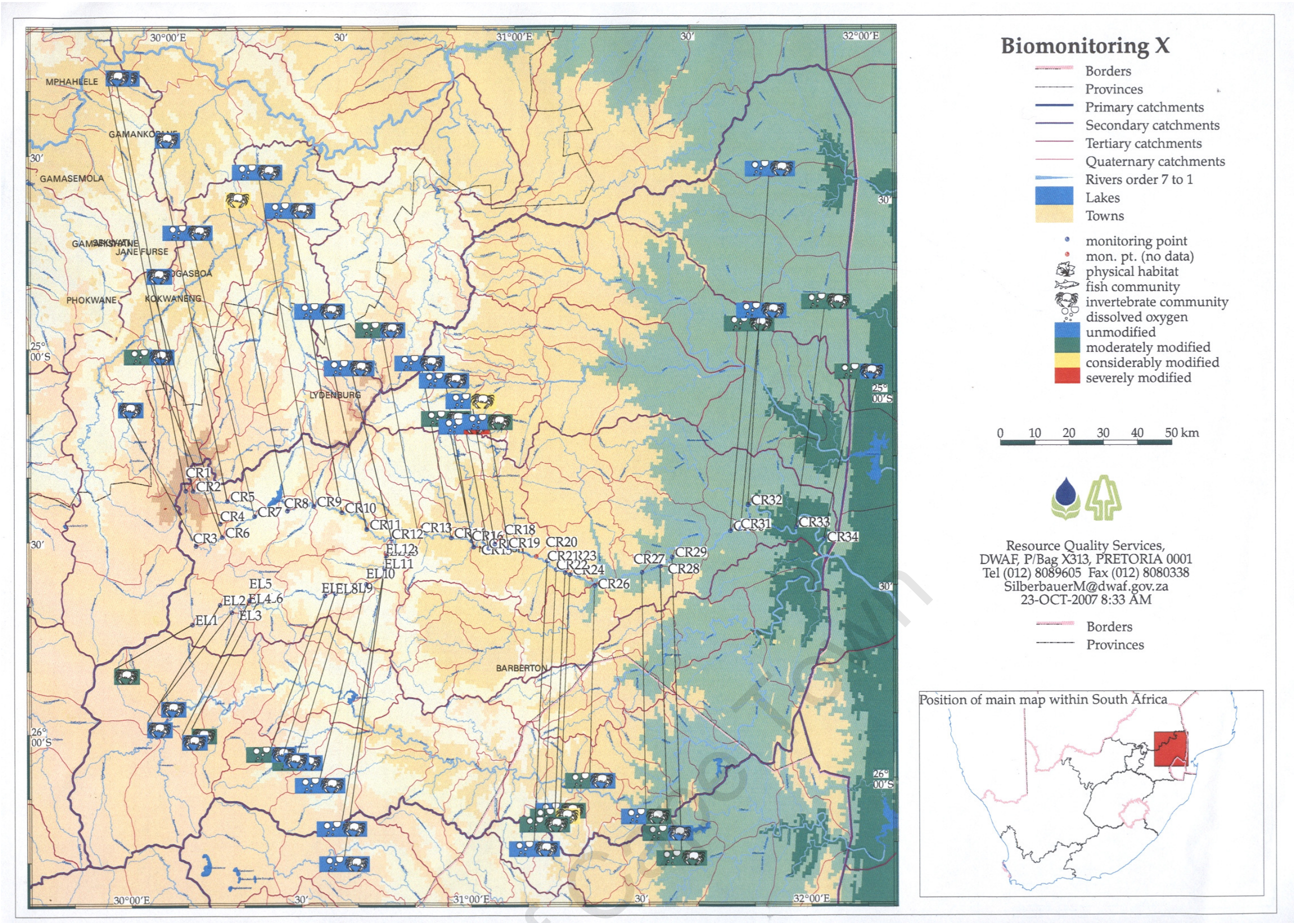


Figure 6-1. A method to automate the process of mapping biomonitoring data and avoid overlapping symbols, with moderate success (Silberbauer 1996). Commercial GIS developers have still not satisfactorily resolved this problem.

As a rule, GIS specialists prepared digital base maps with the required information, for example ecoregions, topography, hypsographic tinting, hillshading and drainage, at a suitable resolution for printing (600 to 1200 dpi) and the report editor completed the map using page layout software to overlay text and symbols. This two-step process was necessary because the available GIS software packages did not include all the tools for creating a product of publication quality. The maps in State of River reports and posters provide the reader with the locality of monitoring sites in relation to factors that might affect the aquatic environment, such as land use, climate and topography. For applications requiring more precision, such as navigation, site localities are stored in the Rivers Database (Dallas *et al.*, 2007).

Ecoregions described in Kleynhans *et al.* (2005) and Moolman (2007) set the broad framework for ecological classification of rivers. Ecoregion boundaries are under constant revision, so slight differences occur between reports.

6.2.4 Photography

Photographic images provide a convenient and detailed record of sites, activities, environmental threats and key organisms. Besides technical documentation, images also served as friezes and backdrops in State of River reports, to convey a general idea of catchment type, inhabitants and environmental factors.

The technology of photography changed substantially during the first decade of the River Health Programme, spanning the silver-to-silicon revolution where negative film and transparencies rapidly gave way to ever-higher resolution digital image recording. The advantages of the new medium include instant replay, simple image storage, cataloguing and processing, no development costs and easy indexing. Digital photography has also opened the way to new opportunities for recording aquatic animal and plant species and for compiling identification guides.

6.2.5 Compiling State of River reports

The compilation of State of River reports began informally but very soon had to adopt disciplined procedures for meeting quality criteria and deadlines. The production of a report goes through the following stages (Strydom 2003): The editor or project leader produces a document structure, in collaboration with local experts, using publication layout software (standard word processing packages are inadequate to the task of correctly placing text and graphics). The document then goes through a tightly controlled iterative process with strict deadlines to refine the content, style and layout. Individual specialists write technical pieces, more or less to specification, using their own observations and data from the River Health

Programme Rivers Database (Dallas *et al.* 2007) to define the present state of the river systems and to suggest desired future states. The next stage is to obtain uniformity of style, aiming at the informed member of public with a secondary school education. It includes text boxes to highlight specialist material or items of importance to help the reader find relevant information quickly. A consistent style is difficult to sustain in a document with many authors, and the editor starts the process by assembling the text blocks from individual sources into a draft structure. A team member who has not been closely involved in the production phases then rewrites and ruthlessly prunes the draft, aiming for clarity of expression and elimination of “technobabble”, striving to be informative but not patronising.

After returning the draft layout to the technical writers for verification, the editor incorporates any changes and combines photographs, maps, revised text boxes and diagrams into a final draft, which is distributed to the development team for content verification. The document then undergoes management and political approval and a final layout check before printing and distribution, sometimes with an official launching function.

6.2.6 Poster production

Poster preparation is a less formal process than State of River report production, although it draws from the same basic State of River material. The first River Health Programme poster was a general overview of the river health monitoring process, containing a short overview of the project, contact information and an example map with symbols showing the status of plants, invertebrates and fish. I created the technical draft using Microsoft PowerPoint, after which a professional graphic designer took the concept and changed it into a print-ready format (Figure 6-2).

Following on the technical poster, team members agreed to the need for a special poster that would generically portray both adverse impacts on rivers and good management practices (Free State poster, RHP 2004d). Such a poster would not only be suitable for teaching but also for display in public and for communication with people lacking adequate reading skills. The River Health Programme team in the Eastern Cape needed a more catchment-specific poster, enabling their audience to identify specific landmarks such as places of residence or work. This accommodation of regional requirements was essential for the adoption and buy-in of River Health Programme concepts.

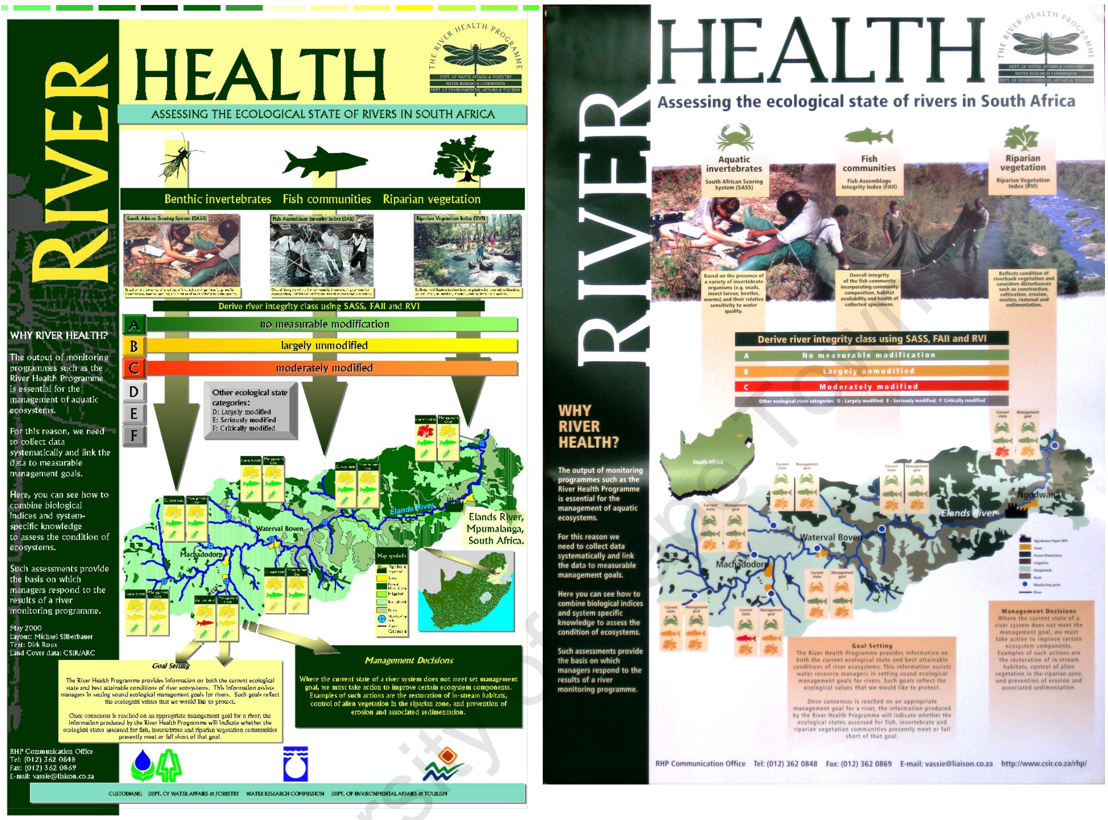


Figure 6-2. Original draft of the first River Health poster (left) and the final product (right). Fundamental changes by the layout artist include the invertebrate and fish species for the icons, the use of more subtle colours and the removal of the heavy arrows.

6.2.7 Compiling activity books

Materials developed for the receptive minds of younger children could fit in well with South Africa's outcomes based education (OBE) system. If teachers were to use these materials within the curriculum's water theme, they would have the advantage of materials aligned to the local environment. The River Health Programme team therefore collaborated with the Department of Education, building on the basic OBE curriculum themes of water use and water cycle, and expanding the activity books

(Table 6-5) to include information on river ecosystems, human impacts and mitigation. With the emphasis on "what can I do?" and "what is good and what is bad", children can begin to understand the importance of responsible water and river use with the help of games and activities. Inclusion of the basics of eye-hand coordination and counting encourages the teaching facilitators to use the activity books (Strydom 2009a,b).

6.2.8 From English to other official languages

While the language of the main State of River reports is English, the use of English for documents aimed at people in rural areas would introduce a serious barrier to understanding. River Health Programme team members with the appropriate language skills therefore undertook translation into different official languages. Where the Programme made use of commercial translation services, the local River Health teams would carefully evaluate the accuracy of the product because of the specialised technical nature of the text and the non-equivalence of many technical terms in indigenous languages. The idiom and dialects of regions also differ, and the local teams contributed advice and translation options to ensure correct interpretation in their area (W. F. Strydom, personal communication). Three languages for which successful translations of some River Health documents have taken place are Afrikaans, isiXhosa and Setswana.

Afrikaans

Even when technical Afrikaans language skills are already available in a project team, the document editor requests the University of the Free State Centre for Environmental Management to review the final text for correctness. A problem that Afrikaans has in common with any local language used closely with an international language like English, is that Afrikaans speakers unconsciously adopt many anglicised terms where "correct" Afrikaans versions exist.

isiXhosa

The translation into isiXhosa was an important step that required the combined skills of professional translators and local River Health Programme teams in the Eastern Cape. IsiXhosa does not have an extensive scientific vocabulary so the elimination of ambiguity is an important consideration (Strydom 2009a,b). The task was more complex than expected and disputes over dialectal differences delayed the publication of an activity book and a poster by several months.

Setswana

Some material developed by the North West Provincial team was in Setswana, for example the Crocodile (West) and Marico poster (RHP 2005c).

6.2.9 Illiteracy and visual impairment

Some posters are available in a purely graphical format for the audience whose reading skills are poor or nonexistent. However, the most challenging translation of all is into a format accessible to the visually-impaired and the River Health Programme has unfortunately not yet embarked on this type of publication. Although South Africa has the technology to produce text and maps in Braille (Bowerman 2009), the material would require extensive translation from a visual format to a descriptive format.

6.2.10 Evaluation of State of Rivers reports

Members of the River Health Programme team have formally investigated the effectiveness of the programme from the philosophy to the practical application. Strydom *et al.* (2002) conducted interviews with scientists and managers, and distributed questionnaires, to gauge readers' opinions about two reports, one on the Crocodile, Sabie-Sand & Olifants River Systems (RHP 2001a) and the other on the Letaba and Luvuvhu River Systems (RHP 2001b). Roux *et al.* (2008) took a broader and more theoretical view of the quality and relevance of the research carried out under the programme by evaluating the publications that came out of it and by studying its wider influence on the management of resources. Strydom (2009a,b) used questionnaires and visits to the catchments to find out what the people in the Buffalo, and Hartenbos and Klein Brak River catchments had learned, or could learn, from the River Health Programme, and how far the information was distributed. Van Wyk *et al.* (2008) and Funke *et al.* (2008) have examined how differences in "knowledge culture" influence the way in which scientists and managers receive and interpret scientific information, with special attention to South African resource management. Scientists and managers need to build and maintain lines of communication for an extended period before they can understand one another's cultural and institutional viewpoints well enough for meaningful interchange of information and support (van Wyk et al. 2008). Bridging the communication gap between scientists and politicians is particularly difficult because of the different scales at which they view the environment, people and the country (Funke et al. 2009).

6.3 Products of the reporting process

6.3.1 Icons for classifying suitability for use

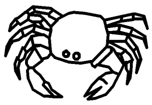
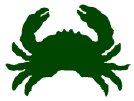
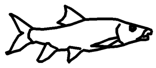
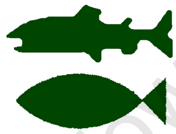
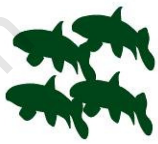
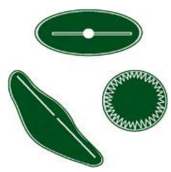
The colour codes used for habitat status appear in Table 6-1 and the prototype and final iconic symbols used to show habitat status at sites on a map are listed in Table 6-2. When used in a report, each symbol takes on the appropriate habitat status colour from Table 6-1.

Table 6-1. Colour scale for habitat status classification.

Colour	Ecological meaning	Management meaning (Strydom, 2006).	Comment
blue	natural / N	No or negligible modification of in-stream and riparian habitats and biota	Protected rivers; relatively untouched by human hands; no discharges or impoundments allowed
green	good / G	Ecosystems essentially in good state; biodiversity largely intact	Some human-related disturbance but mostly of low impact potential.
yellow[1]	fair / F	A few sensitive species may be lost; lower abundances of biological populations are likely to occur, or sometimes higher abundances of tolerant or opportunistic species occur.	Multiple disturbances associated with need for socio-economic development, e.g. impoundment, habitat modification and water quality degradation.
red	poor / P	Habitat diversity and availability have declined; mostly [...] tolerant species present; species present are often diseased; population dynamics have been disrupted (e.g. biota can no longer reproduce or alien species have invaded the ecosystem).	Often characterised by high human densities or extensive resource exploitation. Management intervention is needed to improve river health--e.g. to restore flow patterns, river habitats or water quality.
purple or grey	Artificial / A [or "Un-acceptable" or "seriously modified "]	Transformed to such an extent that their habitat types, biological communities and ecosystem processes bear little or no resemblance to those that would occur under natural conditions.	Modified beyond rehabilitation to anything approaching a natural condition. Example- canalised rivers in urban environments.
black	critically modified [rarely used]	Critically or extremely modified.	[definition uncertain: "worse than purple"]

[1] Note that in practice "yellow" requires the addition of a small amount of red to improve the contrast with "white" paper, which is usually slightly cream-coloured. The result varies widely on different printers and papers.

Table 6-2. Symbols developed for the River Health Programme to map the health at monitoring sites, showing the prototype and currently used versions.

Indicator	Symbol description	Prototype symbols	Current symbol
Index of habitat integrity	Riparian habitat combined with instream habitat as a stylised landscape with the river colour-coded differently from the banks.		
South African scoring system invertebrates	Crabs, notonectid bug, crab again...  		
Fish assemblage integrity index	More or less identifiable indigenous *Labeo*  species or a simplified 'alpha'-symbol. Eventually, a school emphasising a fish *community*: the number of fish (*Pseudobarbus afer*) is greatest for the natural habitat status.		
Riparian vegetation index	Simplified *Ficus* sp., simplified *Acacia* spp., generic tree.  		
Geomorphological index	Stylised river in landscape.		
Water quality	Glass (drinking water); later used water drop to emphasise that the River Health Programme cannot evaluate suitability of water for drinking.		
Diatoms	Frustules of generalised diatoms		

6.3.2 Railway diagrams of rivers

All State of River reports carried conventional locality maps of river networks, mostly based on the 1: 500 000 scale vector dataset of southern African rivers (Silberbauer 2006). From 2001 some reports included a simplified river network in railway diagram format (Figure 6-3 and Figure 6-4) with the colour-coded icons at relative positions along the river.

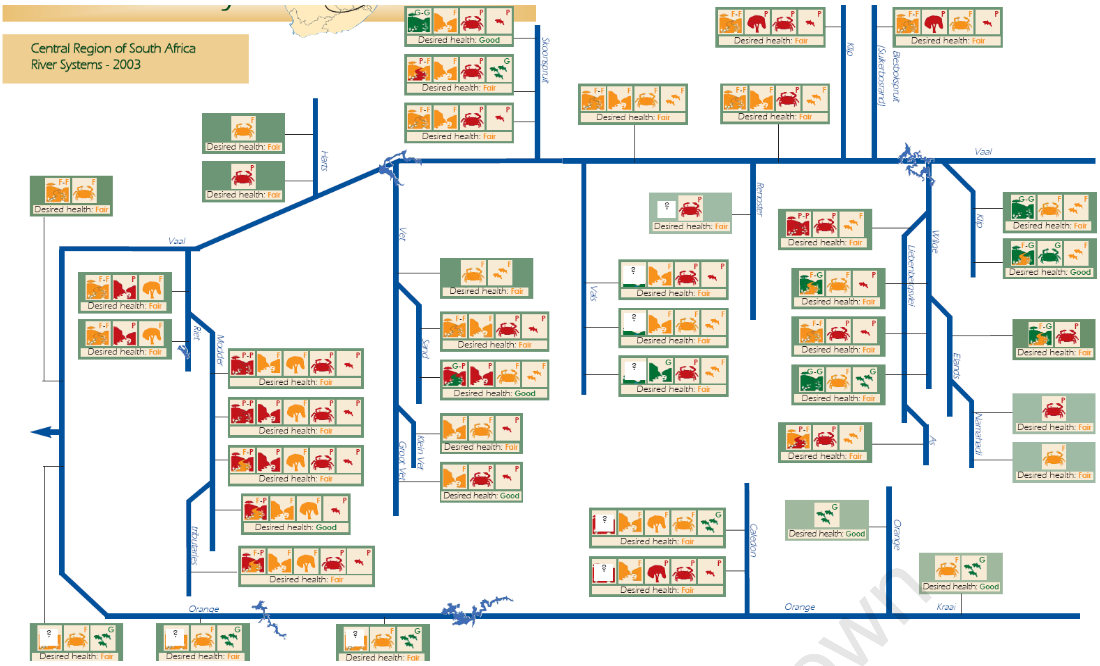


Figure 6-3. Railway diagram of river networks in the Free State Province and adjacent catchments.

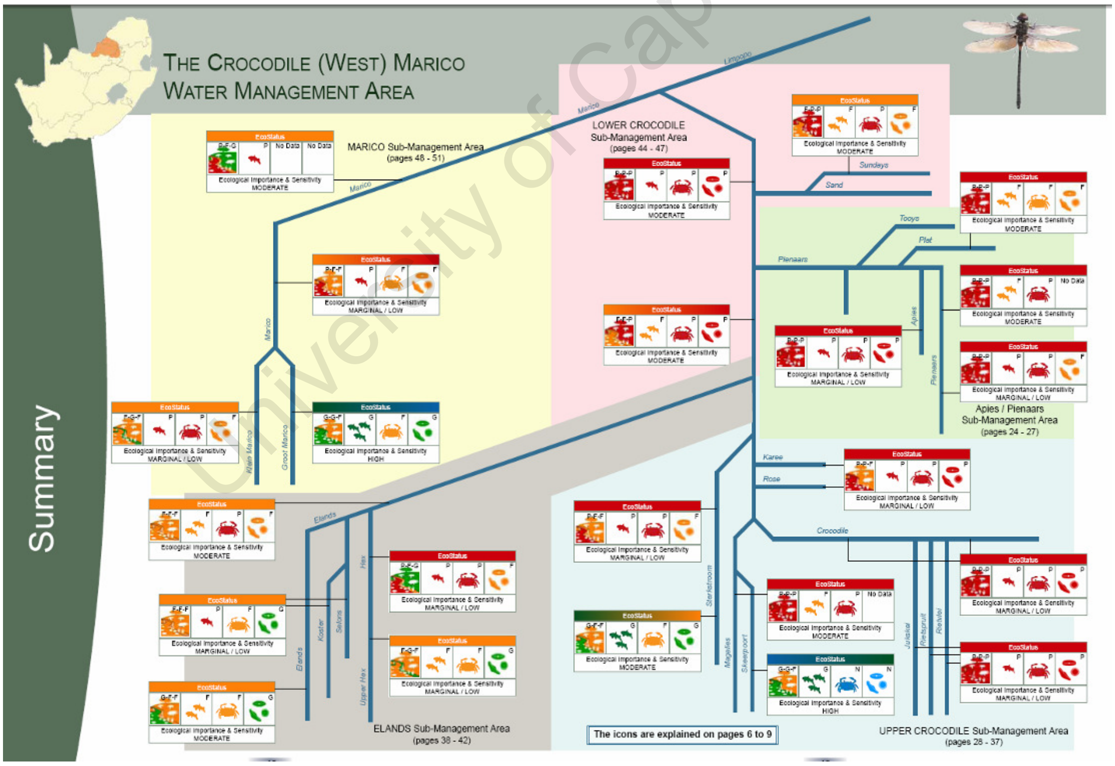


Figure 6-4. Railway diagram of the Crocodile-Marico river network, with background shading to show the main drainage basins.

6.3.3 Reports

A full account of the reports and findings of the River Health Programme from 1994 to 2004 appears in Strydom *et al.* (2006). Table 6-3 lists all the reports in the series.

Table 6-3. The River Health Programme State of River reports from 1998 to 2007.

Title	**Reference**
State of the Crocodile River	RHP 1998
Crocodile, Sabie-Sand and Olifants Rivers	RHP 2001a
Letaba and Luvuvhu Rivers	RHP 2001b
uMngeni River	RHP 2002
Hartenbos and Klein Brak Rivers	RHP 2003a
Diep, Hout Bay, Lourens and Palmiet River Systems	RHP 2003b
Free State Region State of River report	RHP 2003c
Berg River System	RHP 2004a
Buffalo River System	RHP 2004b
Crocodile(West) Marico WMA	RHP 2005a
Greater Cape Town's Rivers	RHP 2005b
Olifants/Doring and Sandveld Rivers	RHP 2006a
Mokolo River	RHP 2006b
Rivers of the Gouritz Water Management Area	RHP 2007a

The covers of the first and second State of River reports (RHP 1998, RHP 2001a) differ markedly as a result of rapid changes in approach in the early years of the River Health Programme (Figure 6-5). The vertical green bar remained a characteristic of the report covers but the landscape page layout, although ideal for maps, was inconvenient for shelving. The dragonfly logo had not yet emerged when the first report came out.

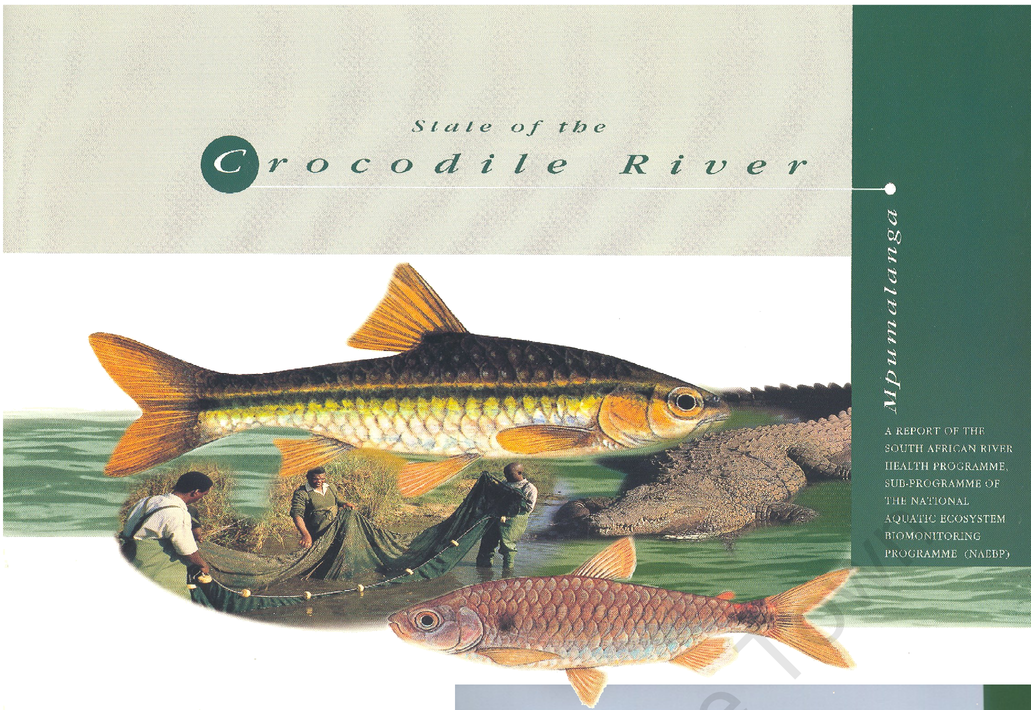


Figure 6-5. The first State of River reports, on the Crocodile River in Mpumalanga (RHP 1998, RHP 2001a). The emphasis shifted from appearance to content in the transition from outsourced publication to in-house document preparation by scientists in the River Health team.

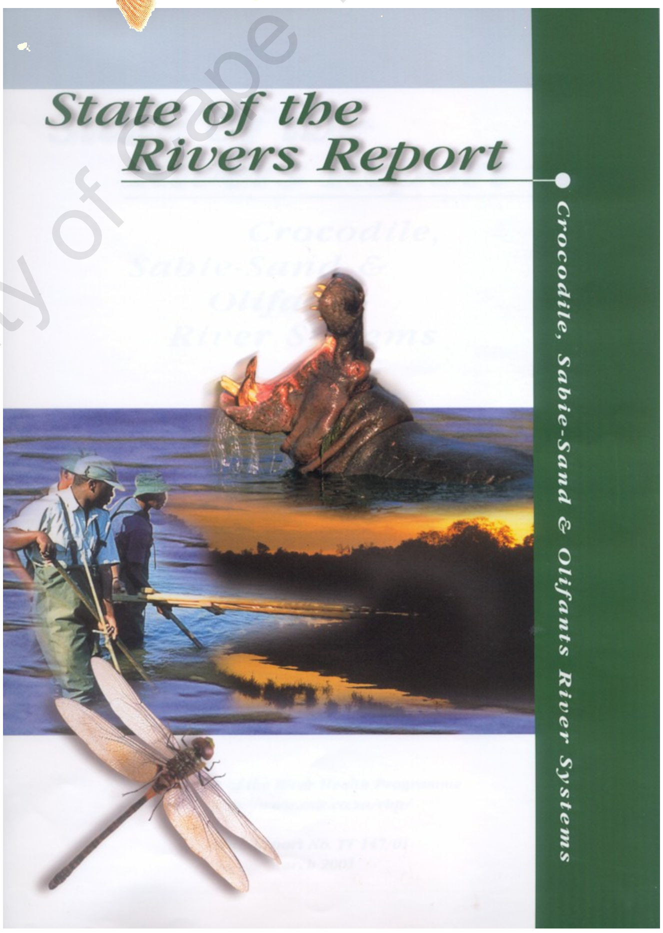


Figure 6-6. Montage of river reports and posters from Strydom et al. (2006).

6.3.4 State of Rivers posters

Various teams working on the River Health Programme have produced about ten posters covering topics from the specific conditions within a catchment to the general principles of habitat integrity (Table 6-4). An example of a detailed State of River poster appears in Figure 6-7.

Table 6-4. Posters produced by various teams for the River Health Programme.

Subject	Languages	Reference
Crocodile River	English	RHP 2000a
Modder River	English	RHP 2001c
Hartenbos & Klein Brak (A3 fold)	English	RHP 2003d
Southern Gauteng Rivers March	English	RHP 2003f
Free State Region State of River poster	English, Afrikaans, seSotho	RHP 2004c
Free State Region fun poster	English	RHP 2004d
Crocodile(West) Marico poster	English & Afrikaans, English & Setswana	RHP 2005c
Cape Town's Rivers poster	English, isiXhosa	RHP 2005d
Berg River poster	English, Afrikaans	RHP 2005e
Goukou and Duiwenhoks Rivers	English	RHP 2006c
Rivers of the Overberg Region	English	RHP 2006d
Selected rivers of the North West Province	English	RHP 2007b

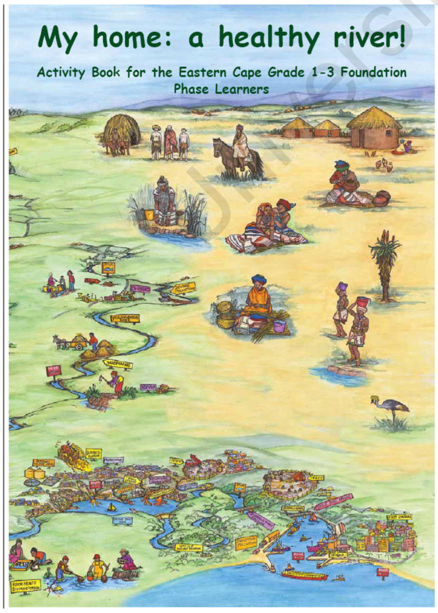
Figure 6-7. Detail of a poster (isiXhosa edition): Ecological state of Cape Town's rivers (RHP 2005d).

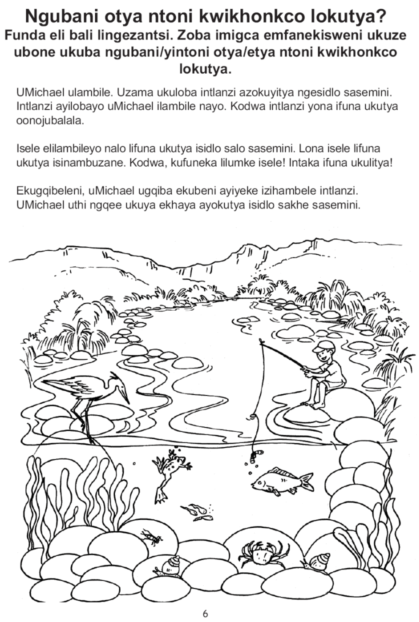
River Health Programme activity books

Various teams have produced colouring-in books aimed at young schoolchildren. A selection of the titles appears in Table 6-5.

Table 6-5. Activity books prepared as part of the River Health Programme for use in schools.

Book	Language	Reference
	English, isiZulu & Sesotho.	RHP 2000b
My home: a healthy river! Colouring-in Book (Finny Fish)	English, Afrikaans & isiXhosa	RHP 2001d
	English, Afrikaans & Sesotho	RHP 2003e
My home: a healthy river! Activity Book for the Eastern Cape Grade 1-3 Foundation Phase Learners	English	RHP 2005f
Water-wise with our rivers. Activity Book for the Free State Grade R-3 Foundation Learners	English & Afrikaans	RHP 2005g
Rivers are fun! Activity Book for the Western Cape Foundation Phase Learners (grade 1-3)	English	RHP 2006e
Rivierpret! Aktiwiteitsgids vir die Wes-Kaapse Grondslagfase Leerders (Graad 1-3)	Afrikaans	RHP 2006f
Yonwabele imilambo yetho! Incwadana yemisebenzi yabaFundi besiga esisisiseko baseNtshona Koloni (ibanga loku-1 ukuya kwelesi-3)	isiXhosa	RHP 2007c




Ngubani otya ntoni kwikhonkco lokutya?
Funda eli bali lingezantsi. Zoba imigca emfanekisweni ukuze ubone ukuba ngubani/yintoni otya/etya ntoni kwikhonkco lokutya.

UMichael ulambile. Uzama ukuloba intlanzi azokuyitya ngesidlo sasemini. Intlanzi ayilobayo uMichael ilambile nayo. Kodwa intlanzi yona ifuna ukutya oonojubalala.

Isele elilambileyo nalo lifuna ukutya isidlo salo sasemini. Lona isele lifuna ukutya isinambuzane. Kodwa, kufuneka lilumke isele! Intaka ifuna ukulitya!

Ekugqibeleni, uMichael ugqiba ekubeni ayiyeke izihambele intlanzi. UMichael uthi ngqee ukuya ekhaya ayokutya isidlo sakhe sasemini.

6

Figure 6-8. The cover of an activity book (RHP 2005f) and a sample page (RHP 2007c) produced as part of the River Health Programme in the Eastern Cape Province.

6.3.5 Evaluation of river health reporting in South Africa

British state of environment reports: "Anthropological evidence" from a researcher:

I have been looking out for physical evidence of use of the reports. For example, I happened to be using one of the Heads of Function's offices. Not only was their copy of the Coastal Report in pristine condition, but I could not get it out of the magazine holder it was in! This strongly suggests it does not get used. I have noticed that people with well-thumbed copies tend to keep them on the shelf just above their desk rather than out of reach in a cabinet. In the interviews several people mentioned the condition of their reports:
"It looks a bit tatty!" [Freshwater Report] "I've used little sticky labels" [Coastal Report]
While these data are by no means comprehensive I have certainly noticed several 'perfect' reports shelved out of reach in private offices and I do wonder whether some of these could be put to better use.

--Wolfenden, 1999

The evaluation of the influence of information dissemination projects such as those used in the River Health Programme and its reports has been the subject of various reports. Some findings relevant to the present discussion are summarised in Table 6-6 and are dealt with in more detail in the next section.

Table 6-6. A summary of conclusions reached by several authors investigating the effectiveness of South African water management communications. (See Discussion section for more details.)

Summary of findings relevant to this study	**Authors**
One cannot assume that the target audience for a state of rivers report will actually read it and assimilate the information. Follow-up procedures are necessary.	Strydom (2003)
If scientists and managers are to cooperate constructively in the study and management of natural resources, they need to understand their cultural differences.	Van Wyk *et al.* (2008)
Although the River Health Programme has steered South African aquatic resource managers out of the rut of monitoring for the sake of monitoring, they face the risk of being stuck in a new rut of monitoring for the sake of reporting. The link between generating reports and bringing about improvements in the aquatic environment needs strengthening.	Roux *et al.* (2008)
Communication efforts can trip up on the failure of such seemingly obvious factors as ensuring the distribution of reports to the right people and ensuring that they know how to interpret the information. Interaction by resource managers and scientists with children and parents in the school environment is a practical way to evaluate the effectiveness of high-level management and information programmes in a catchment.	Strydom (2009a,b)

6.4 Discussion

Eye-tracking studies have shown that carefully structured and well-cued illustrated documents reduce the cognitive load on the reader trying to make sense of the content (Holsanova *et al.* 2008). Combining information from disparate sources and presenting graphical information alongside text involves considerable effort, including the application of powerful and frustrating software. Even leaders in the field of graphics have encountered problems applying their own precepts when publishing. Ware (2008: Preface) battled with his publisher to achieve the required layout of figures and text, and was clearly disappointed with the final product. Tufte (2001) gave up trying to use publishing houses and undertook the risky and expensive route of self-publishing. The River Health Programme also followed the course of laying out the State of River reports in-house. This process requires much effort on the part of scientists in the field of design, which is unfamiliar to many, and commits managers to expenditure on what may appear to be cosmetic products. All this is justified if the reports show themselves to be an effective medium for advising those who make decisions regarding the environment.

The results of various evaluations of scientific reporting (Table 6-6) have demonstrated forcefully that, despite our best efforts in design and production, we cannot assume that we simply need to launch our reports like best-sellers to a grateful public, who will promptly assimilate them and apply the knowledge therein. This is not a particularly local phenomenon: similar problems of lack of use occurred with British state of environment reports, despite a publication price of £35 to £50 per copy (Wolfenden 1999). Many managers feel that they do not have the time to apply the information contained in this type of report (Wolfenden 1999, Strydom *et al.* 2002 and van Wyk *et al.* 2008). This could be a result of work stress, where individuals struggle to deal with multiple and conflicting responsibilities, or the consequence of scientists aiming environmental status reporting at too wide an audience and not meeting specific requirements for detail, depth, frequency or interpretation in the context of each reader's own field. Since the reader might be a biologist, engineer, lawyer, hydrologist or school pupil, a lack of applicable information for any specific subject area would not be surprising.

6.4.1 Compilation of reports

Notwithstanding the difficulty of gauging the usefulness of State of River reports to readers, the compilation process has doubtless benefited the writers and contributors in their understanding of resource function and management requirements. The editing process compels contributors to communicate with one another about data presentation and to meet regularly, exchange ideas and get their results and interpretations on paper in a logical way that is at least intelligible to their fellow-scientists. In essence, the report-writing process encouraged participants with a wide range of experience and expertise to interact and to articulate their research assumptions and conclusions in terms understandable to others—and in the process probably improved their own understanding of interactions between aquatic ecosystem components. The compilation of reports was educational for scientists at all stages of their careers, providing them with a form of "external cognition" for sharing ideas and brainpower.

Despite these efforts, a gulf still exists between the gatherers of knowledge and those who need to apply it in running a country (Roux *et al.* 2006, Van Wyk, *et al.* 2008). Some scientists have expressed personal reservations about the suitability of state of environment reports for bridging the gap, and suggest that the considerable resources that go into this process might be better applied directly to scientific or management activities. For the moment, we will put these iconoclastic thoughts to one side and concentrate on the visualisation aspects of River Health reporting.

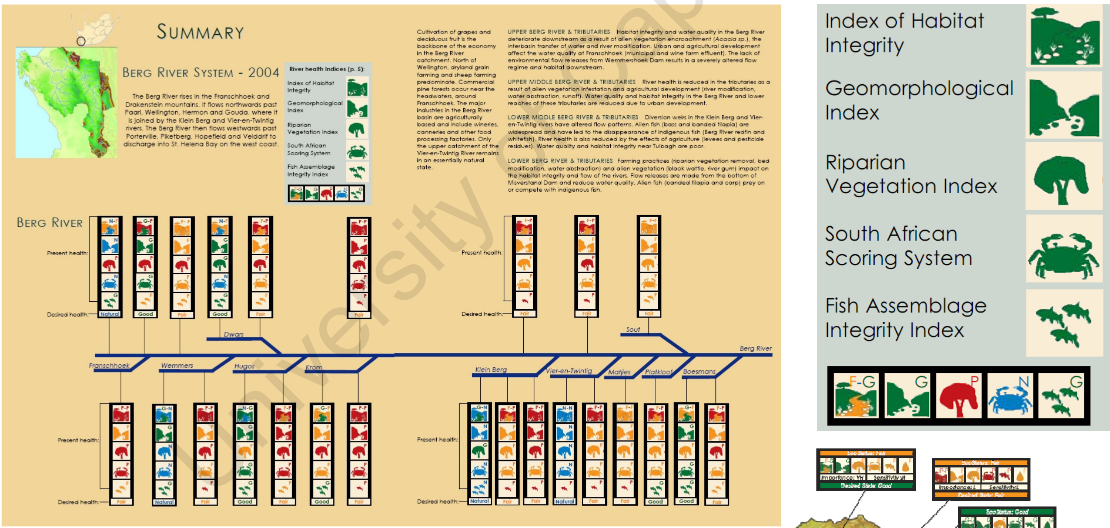
6.4.2 Visual communication with icons

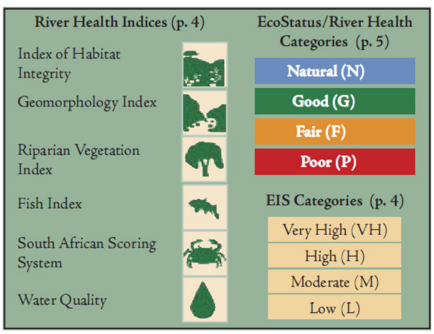
Simple and striking icons draw one's attention to important facts, and are particularly effective for marking points on a map without obscuring geographical information. Icon design at first appeared trivial, but as the River Health Programme developed, we became aware that communication by symbolic images has a direct and subliminal nature that demands great care on the part of the developer. Subliminal exposure to political symbols can undermine subjectivity (Hassin *et al.* 2007) and extremely brief exposure to images, in the order of a tenth of a second, can trigger emotions without conscious awareness (Ruys and Stapel 2008). The need for caution in a nation with diverse cultural imagery is evident. Consider, for example, the care with which South African heraldic specialists have designed emotive symbols of power such as the national flag by F. Brownell (e.g. Burgers 2005) and the military crests for the first warships purchased under the new government (Bennet 2007). Scientists can often ignore cultural sensitivities in the formal framework of a journal paper, but when communicating with a wider audience such as the readers of a State of River report, unintended subconscious cues could divert the reader from the essential message of a map or table. Depending on the circumstances, naive use of a cross, skull or crescent moon in a document might cause a susceptible reader to lose the thread of an argument completely.

The development of the State of River reporting procedure was incremental, and the authors of the reports learned new methods as they went along. This approach promoted innovation at the expense of a uniform "corporate image". One case is the 2001 SoR Report for the Crocodile, Sabie-Sand and Olifants River Systems, whose authors for unknown reasons replaced the system of icons for site status with simple dot markers on the map and general classifications of status per ecoregion in the document text (RHP 2001a). This sort of lack of uniformity could cause difficulties later, when comparing between river systems or presenting trends in a single system. River Health icons have never undergone formal user testing and a structured evaluation experiment might make up a useful honours or masters project.

6.4.3 The effectiveness of railway diagrams

Compare the Berg River system diagram near the beginning of the Berg River 2004 report with the complex realistic map in the Olifants/Doring and Sandveld Rivers 2006 (RHP 2004a, RHP 2006a, Figure 6-9). Each is a different representation of reality: both are extreme simplifications, using one or a few measurements at widely spaced sites to pin down the functioning of many square kilometres of ecosystem. The realistic map helps to locate sites within a familiar geographic framework whereas the railway diagram emphasises relationships between sites and possible transport routes for solutes, suspensoids and organisms, while distorting scale and direction. The constraints of the printed page limit both methods, and perhaps the ideal solution is an interactive digital map where the user can turn layers on and off and zoom in and out to examine details (see Chapter 4 for the use of Google Earth to provide this kind of environment for water quality data).





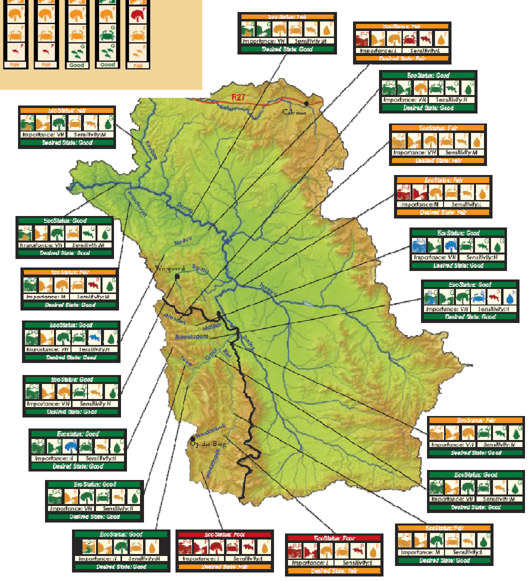


Figure 6-9. The Berg River "railway" network diagram (top, RHP 2004a) and the Doring River "realistic" river map (right, RHP 2006a). Both show the habitat status at monitoring sites, and the Doring map includes a desired future state.

Conventional mapping

GIS specialists have a formidable array of presentation tools at their disposal. They can use software-generated hypsographic tinting for contextual information about terrain. Overlaying political outlines, natural drainage and man-made structures provides reference points. Furthermore, accentuating or fading different layers with judicious choice of colour is technically simple. Subtle hillshading provides cues about slope and aspect, as illustrated in Figure 6-10, where relief shading accentuates the flatter parts of the Free State which contains many endorheic pans, while drainage lines with few tributaries cut right across the pan region. For purposes other than topography, background colouring and hillshading easily confuses the map viewer, and Figure 6-10 is unusually complex.

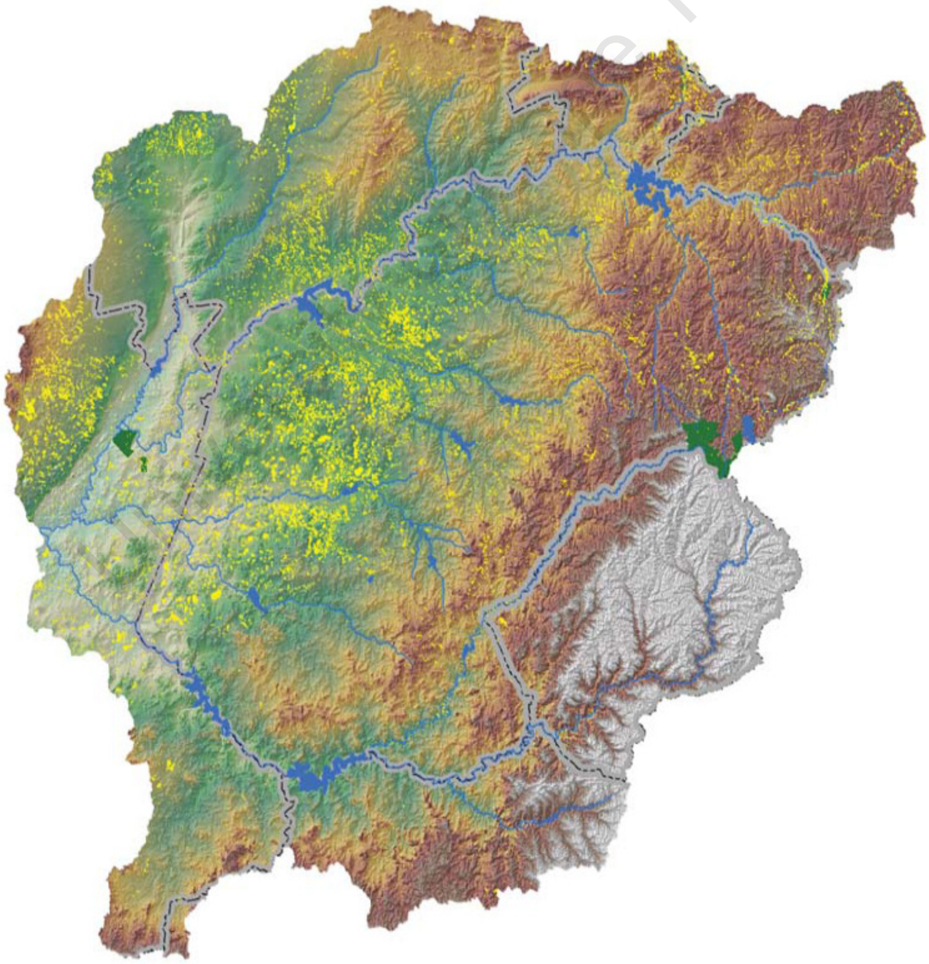


Figure 6-10. The topography of the Free State Province and surrounding areas, with endorheic pans highlighted in yellow. The purpose of including the pans was to illustrate the unusual geomorphology of the area and how it relates to drainage patterns (RHP 2003c: 37).

6.4.5 The photographic record

Figure 6-11. Photograph at site G1BOES-BANGH, illustrating the upper reaches of the Berg River Boesmans tributary near Piketberg (3 November 2003, by the author).

Despite advances in camera technology, photography remains largely an art, and many images submitted for inclusion in reports were technically unacceptable (badly exposed, out of focus, motion blurred) or aesthetically inadequate (poor composition, lack of scale indicators). Photography is nevertheless a powerful medium for communicating the appearance of a monitoring site and specific problems in a river and its catchment. Again, caution is necessary, because the message communicated depends not only on the content of the image but also on the background knowledge of the reader of the report. For example, Figure 6-11 contains a large amount of information, and with a little thought, one can imagine various interpretations by different viewers. Three items provide examples of potential alternative understandings of what the image represents.

Brownish river water: To the aquatic scientist the position of this site, apparently in a sandstone mountain catchment, suggests that the water is probably of good quality. However, the colour from humic acids originating in the Fynbos vegetation might mislead a casual observer into thinking that the water is “dirty”.

Person carrying an object: The human figure provides a useful scale for the specialist and non-specialist alike, and helps the viewer to perceive the image as a real environment that they could imagine themselves interacting with. People familiar with aquatic science will further recognise the activity as some sort of sampling procedure by a scientist (and perhaps even recognise who it is). Others might dismiss it as part of a picnic scene and skip to the next photo.

Green undergrowth and trees in the background: To the casual observer, the background vegetation is a row of healthy green trees setting off an idyllic rural scene. Scientists specialising in climatology might perceive a patch of forest helping to offset the CO_2 deficit. Plant specialists and local environmentalists would be less accommodating and might want to alert the authorities to the encroachment of an invasive alien wattle.

These observations suggest that field photography for this kind of report goes beyond the level of point-and-click or even composition for aesthetic appeal, because the image must convey habitat quality, scale, environmental factors and activities that might affect the river. This requires training, skill, experience and insight.

A completely different class of photograph is the historical record, for example family photographs and old news pictures, often not originally taken with the intention of documenting the environment. Note that archival photos and text may unintentionally introduce cultural bias because examples may be freely available from a well-documented culture and hardly any from one based on oral tradition.

6.4.6 Language matters

Text may seem to lie outside a discussion of visualisation, yet the written word also has a fundamental visual aspect in the human visual system. The visual cortex passes groups of letters or letter patterns to the language comprehension system, where images become descriptive information (Guyton and Hall, 2000). Presentation of text should therefore make these processes as efficient as possible, to avoid confusion and wasted effort in comprehension.

Readers have difficulty wading through large intimidating blocks of text, and searching for linked graphical material. Yet standard text-writing software makes the juxtaposition of words and pictures difficult, leading to confusion about the message (Ware 2008: Preface). The document you are reading is no exception, and follows the conventional way of cross-referencing figures and tables because preparing it in a proper layout form would have taken many months longer. Ideally, an entity on the page should be self-contained, so that a figure or table includes enough information for a reader to understand it. Maps should be self-explanatory, without the need to switch back and forth between the image and a complex legend (Tufte 2001). Text boxes also help to provide structure to a page and are useful as containers for ancillary information or places to summarise the contents of the page for a reader who is just skimming through.

Style is very personal and can sow disharmony between report collaborators. Consider, for example, the choice of font or typeface. Font styles are deeply embedded in our culture (Whalley 1980), to such an extent that we react to them subconsciously. Each of us in our lifetimes will have a slightly different exposure to texts that have entertained or annoyed, with unconscious links to the typeface in which we read the material, and our reading ranges from sublime literature to ridiculous advertising copy. Individuals differ in their awareness of subtle stylistic details such as correct use of opening and closing apostrophes or the inclusion of serifs at the top and bottom of letters. Content, style and layout should be the driving forces in document preparation, and writers should select a font type and size that is easy on the eye, of the appropriate formality and readable in the medium of publication Poole (2005). The use of a clearly readable font may introduce ambiguities in scientific text, for example confusion in this document's font between 1 (one), l (ell) and I (capital i): (New Century Schoolbook 1lI, Times New Roman 1lI, Arial 1lI, Verdana 1lI). In short, the aim in presenting information should always be to keep the reader's attention on the content and not the medium, with the aim of avoiding spurious associations where possible.

Writing style can be in a range of keys spanning prosaic to poetic. Consider these two comments on the conservation of an area in Scotland, one appealing to the intellect and the other to the emotions:

To help achieve sustainable development, it is desirable that people can appreciate both the intrinsic value and diversity of the natural heritage and the economic benefits which it provides.

(Scottish Natural Heritage 2002)

What would the world be, once bereft
Of wet and wildness? Let them be left,
O let them be left, wildness and wet;
Long live the weeds and the wilderness yet.

(Gerard Manley Hopkins, Inversnaid.)

The State of River style gives writers the freedom to move away from the formal scientific prose of a technical report, so contributors have adopted a narrative tone for explaining historical changes. Perhaps wisely, they have so far refrained from expressing themselves in verse. The efforts of even accomplished scientific poets to describe the environment may jar, for example the use of "vegetable" in this poem may have been solemn in the 18th century but now conjures up an absurd image far removed from the royal vision of the poet.

"So sits enthroned in vegetable pride
Imperial Kew, by Thames' glittering side;"

--Erasmus Darwin, grandfather of Charles, 1784:
The Botanical Garden (Underwood 1899).

The River Health Programme's collaborative approach to report writing aims at sustaining a consistent style across reports, without stifling regional "accents" or individual creativity in different centres. Having a distinct identity fosters a sense of ownership and not only helps to motivate the provincial teams but is also important in ensuring continued support by local administrations for monitoring and assessment.

Where possible, report editors encouraged contributors to shun the bureaucratic passive voice and be specific about the organisations responsible for doing things, especially in sections on management recommendations. Thus, a writer would have to rephrase a statement of the form "action X needs to be taken" as "organisation Y should perform action X". Identifying the grammatical subject of such a sentence is not always trivial and sometimes requires a certain amount of research, perhaps boldness. The identification of the person or organisation responsible for an action is also linked to the position of the government logo on the document and the rank of the signatory in the preface. A government minister has the authority defined in his or her department's defining Act of Parliament to insist that a group take action to resolve a problem under the department's jurisdiction. Other organisations such as the Water Research Commission, CSIR, consultants and NGOs regrettably have no such authority and can only make recommendations.

6.4.7 Lost in translation

Direct and literal translation of technical material from English into other official languages such as isiZulu, Sesotho and isiXhosa may not be as effective as anticipated, because those in the intended audience who have a technical background will very likely have received their training with English terminology and would prefer to use the English text. Those without technical training will need more background information than supplied in the original text, so translation develops into the more complex process of converting technical data into the information that users need (Denisov and Christoffersen 2001: 21). Furthermore, certain topics may already be ingrained in the tradition and knowledge of the proposed audience, requiring insight on the part of the translators in order to tailor of the text appropriately. Where the information provided by external sources conflicts with local expert knowledge, local experience will likely carry more weight than the "scientific" input (Weber and Word 2001: 488). Identifying local experts and merging their knowledge with the scientific evidence is a complex, two-way process (Chalmers and Fabricius 2007).

Inter-cultural blunders are a hazard in the production of documents by and about people from divergent backgrounds. A case that illustrates this point was the inclusion in a colouring book of a picture and item about the importance of water in the initiation of young Xhosa men. Scientists had viewed this from the distance as an interesting anthropological event, but were unsure whether its inclusion was insensitive. A principle of the River Health Programme is to build up regional teams, not only to collect data but also to prepare reports. Furthermore, local teams can better deal with thorny regional cultural matters.

6.4.8 Poster communication

Posters are excellent documents for summarising the geographical location of problems and spatial interactions, showing cause and effect on a single sheet of paper. Graphic illustrations of water problems can transcend language barriers, though their immediacy requires more care in preparation than do textual reports, to avoid misinterpretation resulting from cultural filtering. In parallel with the cumbersome and expensive poster format, short take-away documents or flyers are useful for distribution to schools and in libraries and other public buildings.

Poster styles are naturally visual with less emphasis on text. Concise text boxes focus the reader's attention on one topic at a time and pictures are more closely integrated with the text. The layout of the posters aims to guide the reader from the background information through the current state to what is wrong (pressures and impacts) and what can be done (management actions and recommendations) in a visually pleasing way. Scientists, myself included (Figure 6-2), tend to fuss over detail and clutter posters with information, trying to communicate too much and not leaving enough blank space between concepts. The State of River reporting manual provides helpful guidelines for bringing clarity to a crowded page (Strydom 2003).

6.4.9 In the classroom

Strydom (2009a,b) made a detailed study of River Health materials used in the classroom in the Buffalo, Hartenbos and Klein Brak River catchments. She found that the "official" distribution route did not function too well, so she did her own distribution. School materials in the "activity book" style aim to open up a new world to learners, especially those in rural and underprivileged areas without access to books, magazines or the Internet, and to shape their future perceptions. In line with the curriculum for foundation phase learners, activity books give them the opportunity to draw, colour in and to add water uses. The books provide material for eye-hand coordination, language, reasoning, life orientation and encouraging curiosity about natural phenomena. Strydom (2009a,b) found that, through activities, the learners' knowledge broadened with regard to water uses and the benefits to people of clean water. The activity books tried to extend the teachers' knowledge and perceptions as well.

6.4.10 Reporting trends

Each State of River report is a snapshot of the state of a river at a particular time, sometimes with information as to how it arrived in that state. To paint a more complete picture of the effects of external drivers and possible remedial measures, the River Health Programme needs to produce follow-up reports at regular intervals Lack of continuity appears to be a problem in administering and carrying out the work required (Roux, *et al.* 2008). When the data become available, the River Health Programme administrators need to facilitate the publication, at regular intervals, of shorter follow-up documents to State of River reports, consisting of updated key pages in the same format as the initial report. Distribution could be in both print and electronic format. A summary of public comment on the content and applicability of previous versions of the report would add great value, as would descriptions of management actions taken in response to challenges (cf. Weber and Word 2001).

6.4.11 Developments in design

Parts of this discussion seem "wise after the event", but we should remember that the work described here took place in parallel with rapid development in the fields of information visualisation (Card *et al.* 1999), the physiology of interpreting colour, pattern and movement (Palmer 1999) and the optimisation of information presentation to make use of these insights (Ware 2008). The field of design is experiencing a transition from art and craft based purely on the on the intuitive knowledge of talented and skilled illustrators to a much wider field transcending many disciplines. This multi-disciplinary approach to studying communication and cognition has also filtered through to popular culture, as evident in the success of films such as *The Matrix* (1999), which explores consciousness, perception and the use of digital methods to feed different views of reality direct to the sensory regions of the brain (Grau 2005).

6.5 Conclusion

Communication of technical information to a broad, non-specialist audience requires constant vigilance to ensure correct interpretation of the intended message. The River Health Programme has developed a comprehensive and stable reporting method and symbolism for communicating the results of aquatic ecosystem health monitoring. The successful State of River report format combined maps, photographs, icons and text into integrated double-page spreads, where the reader could comprehend a large amount of information at a glance without constantly paging back and forth to find the links. This type of presentation was also the most challenging and time-consuming to construct.

The River Health Programme participants have strived to avoid crossing the fine line between information and propaganda. When we distil information to its essence before serving it to the user, the message is a powerful brew, all the more reason for us to ensure that it is also correct. Caution is important in this regard, because the State-of-River reports in South Africa are in some respects an impassioned attempt by scientists to alert managers and politicians to the importance of river systems, rather than a simple presentation of information as in other countries. The subsequent release of the Green Drop report (DWA 2010) and Cloete (2010) have eclipsed this effort to some extent.

We know or suspect the success of River Health Programme methods at certain levels, such as schools and the ecosystem health scientific community, but at others such as the public and natural resource management, the communication value for the audience is only dimly understood. The diverse cultures existing side-by-side in South Africa provide many opportunities for further research in these areas.

Future communication methods, customisable "landscape fly-through models" such as Google Earth and virtual reality devices, provide a powerful medium for communicating data and metadata—we can not only show the status at various sites, but also how widely-spread these sites are and what assumptions we are making, sometimes on more tenuous grounds than we like to admit. Distributing information in this way is cheaper and easier for the distributor (we have tested the concept already, as in Chapter 4) but puts a great reliance on an Internet-enabled and technically savvy audience, something which is not yet fully in place in 2009.

South African river health reporting, while not yet routine nor sufficiently funded for continuity, is in a much better state than on the rest of the continent, where intensive localised investigations are more likely, e.g. Chakona (2009). First-world administrations, such as the European Union, have a more structured approach than ours, with a large staff complement (water.europa.eu/en/welcome). With structure comes rigidity, and we appear to have more leeway in South Africa to experiment with different monitoring methods and to interact informally with other agencies. In some respects, South Africa is ahead of more wealthy nations, and the Swiss Federation is only now implementing a regular national fish biomonitoring programme (Appendix 11) although their methods are well established (Schager 2004). One could speculate that this reflects the complexities of the federal system compared with the provincial system.

The River Health Programme has served as a valuable college of learning by doing, for all participants. Those who were experts in one area learned new skills in others. Many novices became experts. Catchment management is best learned experientially, where even senior managers occasionally leave their suits and panelled offices behind and get their feet wet (Ison *et al.* 2004).

Chapter 7 - General discussion

"Water resources management is not about engineering and science, it is about people, and their relationship with one another and with the resource on which they depend.
"The ideal situation is a catchment where all the users appreciate the resource and realise the importance of taking care of the resource collectively. In Inkomati, we are all children of the same river."

—Sizile Ndlovu, Chief Executive Officer of Inkomati Catchment Management Agency in van Vuuren (2007)

In this thesis I have traced a more or less logical progression of the graphical representation of water quality, from multivariate time series graphs, through multivariate map symbols, a spatial interface to a monitoring network, visualisation of the mass transfer calculations in catchments, to methods for visually conveying scientific findings about aquatic habitats to a broad audience. Beginning with mechanistic aspects of visualisation, we have seen how the use of motion parallax and perspective views of draped digital elevation models can produce a convincing illusion of the planet-in-a-box. This we can constructively exploit to allow people to forage for data in a landscape that they can traverse at lightning speed, comparing sites and the interaction of human and hydrological features. Using Google Earth is considerably more cost-effective—and less terrifying for all concerned—than taking managers on helicopter trips over their areas of jurisdiction. Google Earth also makes these virtual helicopter rides accessible to millions of people who would not otherwise have the opportunity of getting a bird's-eye view of the area where they live.

At a less dramatic level, we have seen how we can test the over-simplified but convenient model of flow and concentration for many hundreds of sites by generating large arrays of graphs. Here, the data forager is searching for islands of order in a sea of chaos.

Having built up confidence in graphics for our own scientific analyses, we embark on the treacherous process of presenting aquatic ecosystems to others not familiar with our reasoning, such as it is. Here we need to be very clear in our own minds about what we know or can surmise about a system, before imposing this viewpoint on others. We also have to be sure that the necessary foundations and points of reference are there for others to follow our arguments. The "others" may also have their own observations and beliefs about how things work and may regard them as more reliable than ours. We therefore also need to be open to discovery and learning from our audience.

The methods described in this thesis had their origin at a time when Lewandowsky and Spence (1989) could assert, “there have been only a few experimental studies of how people use graphs, and most of these are not known to the average practitioner”. The science of visualisation has moved forward since then, and we are now at the stage where specialists have written journal articles and books on the mechanism and psychology of vision, and on designing graphical data presentations to make best use of this knowledge.

We can accept that most people have a similar optical input to their visual system, with known pathologies and shortcomings, and that the first stage of visual interpretation of certain shapes is involuntary and happens very rapidly (<0.1s). Subsequent processing steps are progressively more dependent on our experiences, upbringing and culture. For example, the geometric, red shape of a stop sign is more likely to make sense universally than a political cartoon depicting a particular time and country.

We have also seen that people use visualisations in two broad and widely different ways. The one is an analytical tool that is an extension to our own cognitive processes, and the other is a means of expressing our internal thought processes in the minds of others—as a means of communication. This is where a difficulty lies. We know our own minds reasonably well, having spent a lifetime in them, but we can only surmise the workings of other minds. Some gifted individuals have an innate skill at “reading” minds, while others struggle because of upbringing or pathology (Grandin 2006). This is a much more complex area than it first appears, and will no doubt provide material for many research projects.

Visionaries like Tukey (1965) and Bertin (1983) foresaw the power of computer graphics for improving our understanding of large and complex datasets. Commercial statistical analysis systems have become commonplace, though expensive, and are complemented by an extraordinary array of freely distributed graphics software, notably the eclectic statistical package, R (R Development Core Team 2008). Tukey moderated his enthusiasm for the coming computer age by saying that "Since the computer is a sharp enough tool to be really useful, you can cut yourself on it", though he implied that this was not a bad thing, and that the statistically unskilled should not shy away for fear of making the odd mistake here and there (Tukey 1965).

The potential for obscuring information by over-use of unnecessary graphical techniques is an ever-present risk, and Tufte (1987) has railed against "chartjunk", those embellishments that make a graphical representation harder to interpret, although Lewandowsky and Spence (1989) countered with a plea for moderation in minimalism. The next great leap in graphical representation is epitomised by Google Earth, which solved at a stroke the problem of combining digital terrains, fly-through animations and point data in a single, democratic application. The KML language indeed allows us to "stand on the shoulders of giants" in achieving our visualisation goals.

In this thesis, I have argued that graphic representations are an important component of understanding and communicating water quality information. I have also cautioned that users of graphical methods should understand that judicious use of certain shapes helps the viewer make more sense of the data; while conversely, a confusing visualisation can cloud the information and make the meaning more obscure.

We have also seen how important it is to be aware of whether we are using particular visualisation methods for helping us in our own analytical work or for communicating our findings with others. An understanding of the audience is vital if we are to convey information successfully using visual methods, even if the audience is ourselves. I hope that this study has shown that great opportunities exist for collaborative research in sociology, psychology, physiology and visual analytics to investigate new methods of data presentation and interpretation.

In developing future visual methods for exploring data, we should not lose sight of the investigative work of Trafton *et al.* (2005) and Trickett and Trafton (2006). The research by these authors into the way scientists apply visualisations has contributed valuable insights into the cognitive processes of visual data interpretation.

For me, a disappointing aspect of this study has been the use of animation. In order to avoid confusing the viewer, animations must provide supplementary information in the form of instructive narrative, subtitles and a time legend, or user controls for stepping back and forth, pausing and jumping to a specific event. My efforts have so far fallen short of these ideals. Fortunately, tools such as the Gapminder Google Motion Chart provide a new platform on which to build completely different ways of viewing time-series data. The Motion Chart's ability to create animations on the fly without the need for pre-generation provides many user options for intervention and many developer options for presentation ("playing the music, not just showing the notes"—Rosling 2007b).

The financial burden of maintaining a computer with Internet, network and E-mail services is another consideration that we cannot ignore. The cost of information and communications technology has come down, but it is still expensive, so we are obliged make the most effective use of it, as with any limited resource. While a technically savvy person can save on licence fees by configuring a computer to run a huge range of free and useful software, this approach is often too confusing, time-consuming and tedious for the average user, and is not yet widely supported by conventional IT staff. When working in isolation from technical support, for example in a rural setting, the user needs an appliance that operates with no fuss. SMS and other cellular communication methods have great potential, and many application developers have begun exploiting them (e.g. Loudon *et al.* 2006, Loudon 2007). With a cellular network, management is highly centralised and the user is able to distribute costs over a long period at an affordable rate. The advantages of piggy-backing onto existing mass-market platforms include lower price of technology and general user familiarity with the interface. The costs of developing a user-friendly interface architecture are considerable (Reiner and Schaper 2010) and we can avoid them by using standard consumer apparatus. Wilson et al. (2007) emphasised this principle in constructing supercomputers from games machines.

Before and during the development work described in this thesis, the author was privileged to work on many programming and visualisation platforms, from a VAX 11/750 running International Imaging Systems software to contemporary desktop personal computers with fast graphics cards, running ESRI's ArcGIS package. The organisation bore the costs of hardware and proprietary software, allowing for a reasonably unbiased comparison of the AML/Avenue/VBA environment with free environments such as the R statistical programming language. In retrospect, what seemed to make the most difference was having a team working on the same platform. From a high point of ten spatial analysts in the late 1980s our group has dwindled to just one in 2010. The synergy that comes from having different specialists working on complementary facets of the same software environment has gone. Only a shadow of this mutual support process exists in online discussion groups.

During the development of the software described here, I took great pains not to reinvent various wheels but rather to build on the experience of others. At a low level, for example the drawing of polygon symbols on a map, some prior development was publicly available, but surprisingly I could not locate anything freely available on a grander scale. A closed system like AquaChem has a slick interface but the USD1500 price tag precludes trying it out in case it works. The most important criterion would be the ability to run in unattended batch mode.

As noted in Chapter 2, the elegant Swiss water quality map produced by Liechti and Jakob (1992) with its four-petal radial symbols seems to have been a once-off production. The ideas for a general process for map symbol construction published by Schnabel (2009) from his thesis (Schnabel 2007) became available only after submission of this thesis for examination. When I suggested that ESRI should evaluate the concept, I received the rebuff reproduced in Appendix 12. It would seem that this is an area in dire need of attention, and the next logical step would be to encourage developments in the open source arena.

An undisputable advantage of open source software is that one can download it, get updates and deploy it anywhere without the red tape of motivations, applying for licences for new users and justifying the licence fee to unsympathetic financial officers. However, open access to specialised services such as the Internet brings its own complications, with some users who have too much time on their hands or too little supervision and self-control, whittling away at what we have noted remains a very expensive commodity. At an even more serious level, protecting a network, its hardware and software from criminal activities such as electronic and physical intrusion adds further unnecessary costs to data processing and reporting. At the Institute where I work, the cost of recovering from a series of thefts by gangs during 2007 totalled millions of rand, including the replacement and configuration of 55 computers, the construction of a kilometre of perimeter fencing and the installation of security equipment. The event not only diverted funds from research and development, but also had unpredicted and severe psychological effects on staff members, including aggression, mistrust and fear.

General conclusions and recommendations

Visualisation methods elucidate the quality of water and reservoirs by creating a world-view that may help us survive in our increasingly complex and crowded environment. The methods described here are less about providing answers directly, than about providing a framework within which water quality data users can discover which questions to ask, and methods by which they can sift the water-related databases for clues as to how to go about finding the answers.

Visualisations are often exciting in that they help us to comprehend how our data relate to the greater scheme of things, but newer visualisations like Google Earth can also have a sobering effect in that they emphasise just how superficial are our attempts to measure the workings of the hydrological cycle. Visualisations can sometimes help us quickly make the connection between problem and solution, for example when viewing pollution sources and water users in a 3D fly-through view of a catchment.

An occupational hazard with the development of complex graphical methods is that they easily become an end in themselves rather than a means to further discovery. Graphics specialists always need to be part of a team that includes subject specialists, who can evaluate the visualisation tools, apply the results and formulate sensible reactions to them. To use the survival analogy, the visual system and the tooth and claw interacting with the external world need to be part of the same animal.

Complex software systems need to run on sufficiently powerful computers and networks, or we cannot fully exploit the benefits of visualisation, depending as it does on a short response time to input. Given a fast enough network, the processing power can reside in a central location and deliver images to less powerful peripheral devices, such as low-cost computers and cellular telephones.

I set out to examine the idea that scientific visualisation methods can increase the value of water quality data to water resource managers. The answer has turned out to be "indirectly, yes". We are far from being able to set up a slick water control bunker with huge monitors showing a live display of the state of the nation's water resources, and I'm not sure that we need to. What we do have are visualisation tools that help specialists explore their data in different ways, and offload some of the burden of cognition onto mechanical scratchpads where they can look at them in a different light and from new perspectives.

A field that has great promise for further development is the Motion Chart discussed in Chapter 3. Two enhancements that would improve the chart's applicability to exploring water quality data would be multivariate symbols and range indicators. Dynamic Maucha diagrams would be a good starting point for multivariate symbols and would enable the user to follow changes in ionic composition, for example sulphate dominance. Range indicators could be limits of statistical uncertainty or guidelines for water usability.

Have we succeeded in developing visual methods for communicating the state of South Africa's aquatic ecosystems to a wide audience in such a way that they can see the problems and the potential solutions? This thesis suggests some progress along these lines, using advanced visualisations to the best of our abilities and creating a great deal of understanding within the scientific community. To make better use of these methods, we will need more studies, such as the work of Bielak (2009) or Strydom (2009a,b), of actual rather than hoped-for reactions of people directly affected by and reliant on water resources, before we can safely say that our investments are bearing fruit. The further the visual message travels from our comfortable circle of mutual understanding, the more we need to reinforce the message by other means, such as personal dialogue with the people living out there.

References

Allanson, B. R. 1961. Investigations into the ecology of polluted inland waters in the Transvaal, part I: The physical, chemical and biological conditions in the Jukskei-Crocodile river system. Hydrobiologia 18:2-76.

Allanson, B. R. 2004. Limnology in South Africa: Past and Present Status and Future Needs. In B. Gopal & R.W. Wetzel (eds.), Limnology in Developing Countries, vol. 4, pp. 1-116.

Allanson, B. R. and Gieskes, J. M. T. M. 1961. Investigations into the ecology of polluted inland waters in the Transvaal, Part II: An introduction to the limnology of Hartbeespoort Dam with special reference to the effect of industrial and domestic pollution. Hydrobiologia 18:77-94.

Baas, J. H. 2000. EZ-ROSE: a computer program for equal-area circular histograms and statistical analysis of two-dimensional vectorial data. Computers and Geosciences 26:153-166.

Bar-Ze'ev, A. 2007. How Google Earth really works. [online] < http://www.realityprime.com/articles/how-google-earth-really-works> viewed 2009-05-11.

Beepa 2008. FRAPS real-time video capture and benchmarking. [online] <http://www.fraps.com> viewed 2008-08-10.

Bennett, C. H. 2007. Names and ships' badges of the Valour class frigates. Naval Digest No 12. Just Done Productions, Durban, South Africa. 76 pp and 4 plates.

Berns, G. 2008. Iconoclast: A Neuroscientist Reveals How to Think Differently. Harvard Business Publishing.

Bertin, J. 1967. Graphics and Graphic Information Processing. Walter de Gruyter & Co, Berlin.

Bertin, J. 1980. Graphics and the graphical analysis of data. Walter de Gruyter & Co, Berlin.

Bertin, J. 1983. Semiology of graphics. University of Wisconsin Press, Madison, Wis.

Blake, E. H. and Tucker, W. D. 2006. User interfaces for communication bridges across the digital divide. AI Soc. 20:232-242.

Blind SA 2009. [online] < http://www.blindsa.org.za/> viewed 2009-03-22.

Blinn, D. W. 1993. Diatom community structure along physicochemical gradients in saline lakes. Ecology 74:1246-1263.

Bowerman, P. 2009. SA Braille atlas launched. PositionIT p. 51. [online] <http://www.eepublishers.co.za/view.php?sid=15603> viewed 2009-05-08.

Boyer, J., Sterling, P., and Jones, R. 2000. Maximizing information from a water quality monitoring network through visualization techniques. Estuarine, Coastal and Shelf Science 50:39-48.

Broch, E. S. and Yake, W. 1969. A modification of Maucha's ionic diagram to include ionic concentrations. Limnology and Oceanography 14:933-935.

Brooks, F. P. 1995. The Mythical Man-Month: Essays on Software Engineering, Anniversary Edition. 2nd ed. Addison-Wesley Professional.

Burgers, A. P. 2005. The South African Flag Book. Protea Book House, South Africa.

Card, S. K., Mackinley, J. D., and Shneiderman, B. 1999. Readings in Information Visualization: Using vision to think. Academic Press, San Diego.

Carey, G. R., Geel, P. J., Wiedemeier, T. H., and McBean, E. A. 2003. A modified radial diagram approach for evaluating natural attenuation trends for chlorinated solvents and inorganic redox indicators. Ground Water Monitoring & Remediation 23:75-84.

Chakona, A., Phiri, C., Chinamaringa, T. and Muller, N. 2009. Changes in biota along a dry-land river in northwestern Zimbabwe: declines and improvements in river health related to land use. Aquatic Ecology. 43(4):1095-1106.

Chalmers, N. and Fabricius, C. 2007. Expert and generalist local knowledge about land-cover change on South Africa's Wild Coast: can local ecological knowledge add value to science? Ecology and Society 12. [online] <http://www.ecologyandsociety.org/vol12/iss1/art10/> viewed 2009-04-05.

Changizi, M. 2008. Harnessing vision for computation. Perception 37:1131-1134.

Cholnoky, B. J. 1958. Hydrobiologische Untersuchungen in Transvaal II. Selbstreinigung im Jukskei-Crocodile Flusssystem. Hydrobiologia 11:205-266.

Chutter, F.M. 1998. Research on the rapid biological assessment of water quality impacts in streams and rivers. Report No. 422/1/98. Technical report, Water Research Commission, Pretoria.

Cloete, D., de Villiers, R., Bernstein, A., Segal, N. and Altbeker A. 2010. Water: A looming crisis ? Round Table Number 14. Transmedit Place, 5 Eton Road, Parktown, Johannesburg 2193, South Africa: The Centre for Development and Enterprise.

Cobban, D. 1994. Assessment of water quality through a menu-driven user interface on a geographic information system platform. Computer Graphics '94 6-9Sep1994.

Cobban, D. A. and Silberbauer, M. J. 1993. Water quality decision-making facilitated through the development of an interface between a geographic information system and a water quality database. In S. A. Lorenz, S. W. Kienzle, and M. C. Dent (eds.), Hydrology in Developing Regions... the Road Ahead. Proceedings of the Sixth South African National Hydrological Symposium, volume 2, pp. 523-530, Pietermaritzburg.

Crampton, J. W. 2008. Keyhole, Google Earth, and 3D Worlds: An Interview with Avi Bar-Zeev. Cartographica 43:85-93.

CSV2KML 2007. csv2kml 1.0.2 compiled in February 2008. [online] <http://www.tchartdev.com/csv2kml.htm> viewed 2008-08-09.

Dalgaard, P. 2002. Introductory statistics with R. Springer-Verlag, New York.

Dallas, H. F. and Day, J. A. 2004. The effect of water quality variables on aquatic ecosystems: A review. Report Number TT 224/04. Technical report, Water Research Commission, Pretoria.

Dallas, H., Molteno, A., Ewart-Smith, J., and Janssens, P. 2007. Rivers database version 3, a user manual. River Health Programme. Technical report, Department of Water Affairs and Forestry, Pretoria.

Danchev, A. and Todman, D. 2002. War diaries 1939-1945: Field Marshal Lord Alanbrooke. Phoenix Press, London.

Day, J. A. and King, J. M. 1995. Geographical patterns, and their origins, in the dominance of major ions in South African rivers. South African Journal of Science 91:299-306.

de la Harpe, H. 2007. Creative intelligence: the designer's creativity handbook. Platinum Press, Potchefstroom.

de Souza, P., Wensley, A., Manus, L., Delport, E., Stevens, F., Mackintosh, G., and Ramba, M. 2008. Development and implementation of an open-source based internet accessible water quality management system for improving the quality of water services in south africa. In FOSS4G 2008 Free and Open Source Software for Geospatial. OSGeo, GISSA. [online] <http://conference.osgeo.org/index.php/foss4g/2008/paper/view/74/28> viewed 2009-05-08.

DeMerchant, I. 1997. ENCODER.AML, a script that encodes animation frames into an MPEG animation. Code supplied with ESRI ArcInfo.

Denisov, N. and Christoffersen, L. 2001. Impact of environmental information on decision-making processes and the environment. Technical report, UNEP/GRID-Arendal.

DHI 2009. MIKE BASIN – River basin management in GIS. [online] <http://www.dhigroup.com/Software/WaterResources/MIKEBASIN.aspx> viewed 2009-05-23.

Dixon, W. J. 1953. Processing data for outliers. Biometrics 9:74-89.

DoH 2002. National guideline: prevention of blindness in South Africa. Technical report, Directorate of chronic diseases, disabilities and geriatrics, Department of Health, South Africa.

Dror, I. and Harnad, S. 2008. Offloading Cognition onto Cognitive Technology. John Benjamins, Amsterdam.

DWA 2010. Green drop report 2009 version 1: South African waste water quality management performance. South Africa: Department of Water Affairs. [online] <http://www.dwaf.gov.za/Documents/GreenDropReport.pdf> viewed 2010-05-16.

DWAF 1998. Quality of domestic water supplies – Volume I: Assessment Guide. Water Research Commission report number TT 101/98. ISBN 1 86845 416 9. 114 pp. [online] <http://www.dwaf.gov.za/iwqs/AssessmentGuides/AssessmentGuide/AssessmentGuide.pdf> viewed 2009-04-05.

DWAF 2005. Internal Question Paper No 33. [online] <http://www.dwaf.gov.za/communications/Q&A/2005/1567NA14Oct05.doc> viewed 2009-04-05.

DWAF-RQS 2008. Resource Quality Services water quality data exploration supplementary information. [online] <http://www.dwaf.gov.za/iwqs/wms/data/000key2.htm> viewed 2008-08-10.

DWAF-RQS 2009. NCMP Information processing. [online] <http://www.dwaf.gov.za/iwqs/water_quality/NCMP/NWRQSR.htm> viewed 2009-03-26..

Emerson, J. W. and Green, W. 2007. YaleToolkit: Data exploration tools from Yale University. [online] <http://cran.nedmirror.nl/web/packages/YaleToolkit/index.html> viewed 2009-02-19.

Enongene, G. N. and Rossouw, J. N. 2007. Collation and development of nutrient export coefficients for South Africa - P14/12/16/2. Technical report, Directorate: Water Resources Planning Systems, Department of Water Affairs and Forestry, Pretoria.

ESRI. 2006. GIS dictionary. [online] <http://support.esri.com/index.cfm?fa=knowledgebase.gisDictionary.search&searchTerm=AML> viewed 2009-04-05.

Everitt, B. S. 2005. An R and S-PLUS companion to multivariate analysis. Springer.

Feitelson, D. G. 2003. Comparing partitions with spie charts. Technical report, School of Computer Science and Engineering, The Hebrew University of Jerusalem.

Forde, C. D. 1925. Irrigation in South Africa. The Geographical Journal 65:342-349.

Friendly, M. 2008. Milestones in the history of thematic cartography, statistical graphics, and data visualization. *In* Seeing Science: Today. American Association for the Advancement of Science. [online] <http://www.math.yorku.ca/SCS/Gallery/milestone/milestone.pdf> viewed 2009-05-08.

Fuchs, C. and Horak, E. 2008. Africa and the digital divide. Telematics and Informatics 25:99-116.

Funke, N. S., Shaxson, L., and Bielak, A. T. 2009. Evidence-based policy for environmental sustainability: a path forward for South Africa. In Collaborative Workshop on Evidence-based Policy-making, pp. 1-40. Council for Scientific and Industrial Research.

Garrett, L. 2007. The challenge of global health - beware what you wish for. Foreign Affairs 86. [online] <http://www.foreignaffairs.com/articles/62268/laurie-garrett/the-challenge-of-global-health> viewed 2009-06-07.

Gnu 2008. Gawk. [online] <http://www.gnu.org/software/gawk/> viewed 2008-08-10.

Google 2009. Motion chart. [online] < http://www.google.com/ig/directory?url=www.google.com/ig/modules/motionchart.xml> viewed 2009-05-08.

Google Code 2008. KML Reference. [online] <http://code.google.com/apis/kml/documentation/kml_tags_21.html> viewed 2008-08-10.

Grandin, T. 2006. Animals in Translation: The Woman Who Thinks Like a Cow. Bloomsbury Publishing PLC, new edition.

Grau, C. (ed.) 2005. Philosophers Explore The Matrix. Oxford University Press, USA.

Griesel, M., Kühn, A., Kempster, P., Mamabolo, M., and Silberbauer, M. J. 2006. Report on an integrated water quality programme conducted in the town of Delmas, November 2005 to June 2006 (Delmas water quality report). Technical report, Resource Quality Services, Department of Water Affairs and Forestry, Pretoria, South Africa.

Griffin, A. L., MacEachren, A., Hardisty, F., Steiner, E., and Li, B. 2006. A comparison of animated with static small multiple maps for visually identifying space-time clusters. Annals of the Association of American Geographers 96:740-753.

Grobler, D. C. and Ntsaba, M. 2004. Strategic framework for national water resource quality monitoring programmes-DWAF Report No. N/0000/REQ0204. Technical report, Resource Quality Services, Department of Water Affairs and Forestry, Pretoria, South Africa.

Grobler, D. C. and Silberbauer, M. J. 1985. The combined effect of geology, phosphate sources and runoff on phosphate export from drainage basins. Water Research 19:975-981.

Güler, C., Thyne, G., McCray, J., and Turner, K. 2002. Evaluation of graphical and multivariate statistical methods for classification of water chemistry data. Hydrogeology Journal 10:455-474.

Guyton, A. C. and Hall, J. E. 2000. Textbook of medical physiology. W. B. Saunders, Philadelphia Pa., tenth edition.

Harrower, M. and Fabrikant, S. 2008. The Role of Map Animation for Geographic Visualization. Wiley and Sons.

Hassel, A. and Martin, D. F. 1995. A computer-generated (Q-BASIC) program to construct Maucha diagrams. Florida Scientist 58:300-312.

Hassel, A. L., Dooris, P. M., and Martin, D. F. 1997. Maucha diagrams and chemical analyses to diagnose changes in lake chemistry. Florida Scientist 60:75-80.

Hassin, R. R., Ferguson, M. J., Shidlovski, D., and Gross, T. 2007. Subliminal exposure to national flags affects political thought and behaviour. Proceedings of the National Academy of Sciences of the USA 104:19757-19761.

Hauser, A. 1951. The social history of art from prehistoric times to the middle ages. In The social history of art, volume 1. Routledge & Kegan Paul, London and Henley.

Heer, J. and Robertson, G. G. 2007. Animated transitions in statistical data graphics. IEEE Information Visualization (InfoVis) Includes movie AnimatedTransitions.mov. 13:1240-1247.

Hem, J. D. 1985. Study and interpretation of the chemical characteristics of natural water. Water-Supply Paper 2254. US Geological Survey, United States Government Printing Office, third edition.

Hohls, B. C., Silberbauer, M. J., Kühn, A. L., Kempster, P. L., and van Ginkel, C. E. 2002. National water resource quality status report: Inorganic chemical water quality of surface water resources in SA; the big picture. Report no. N/0000/REQ0801. Technical report, Pretoria, South Africa.

Holsanova, J., Holmberg, N., and Holmqvist, K. 2008. Reading information graphics: The role of spatial contiguity and dual attentional guidance. Applied Cognitive Psychology. [online] DOI: 10.1002/acp.1525, viewed 2009-05-17.

Horton, R. E. 1945. Erosional development of streams and their drainage basins; hydrophysical approach to quantitative morphology. Geological Society of America Bulletin. 56(3):275-370.

Ison, R. L., Steyaert, P., Roggero, P. P., Hubert, B., and Jiggins, J. 2004. The SLIM (social learning for the integrated management and sustainable use of water at catchment scale) final report. Technical report, European Commission. [online] URL: <http://slim.open.ac.uk/objects/Outcomes/SLIM%20Final%20Report.pdf> viewed 2009-04-05.

Keller, H. 1933. Three days to see. The Atlantic Monthly 151:35-42. [online] <http://www.theatlantic.com/issues/33jan/keller.htm>, viewed 2009-03-22

Keogh, E., Lonardi, S., and Chiu, B. . Y.-c. 2002. Finding surprising patterns in a time series database in linear time and space. In KDD '02: Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 550-556, New York, NY, USA. ACM.

King, R. and Manning, L. 1999. Modernized STORET. In: Symposium on Water Resources and the World Wide Web. American Water Resources Association. [online] <http://www.awra.org/proceedings/www99/w50/index.htm> viewed 2010-05-16.

King, J. M., de Moor, F. C., and Chutter, F. M. 1992. Alternative ways of classifying rivers in South Africa. River Conservation and Management Ch 14:213-228.

Kleynhans, C. J. 2007. Module D: Fish response assessment index in river ecoclassification: Manual for ecostatus determination (version 2). Technical report, Water Research Commission and Department of Water Affairs and Forestry. [online] <http://www.dwaf.gov.za/iwqs/rhp/eco/EcoStatus/ModuleD_FRAI/Module%20D_FRAI.pdf> viewed 2009-06-13.

Kleynhans, C. J., Thirion, C., and Moolman, J. 2005. A Level I River Ecoregion classification System for South Africa, Lesotho and Swaziland. Report No. N/0000/00/REQ0104. Technical report, Resource Quality Services, Department of Water Affairs and Forestry, Pretoria, South Africa.

Kleynhans, C. J., Mackenzie, J., and Louw, M. D. 2007. EcoStatus determination (Version 2) Module F: Riparian Vegetation Response Assessment Index (VEGRAI). Report No: TT 333. Technical report, Water Research Commission and Department of Water Affairs and Forestry. [online] <http://www.wrc.org.za/downloads/report%20lists/web%20rpts/ecosyst/TT%20333-CONSERVATION.pdf> viewed 2009-06-17.

Kleynhans, C. J., Louw, M. D., and Graham, M. 2008. River ecoclassification manual for ecostatus determination (Version 2) Module G: Index of habit integrity Section 1: Technical Manual. Report No: TT378-09. Technical report, Water Research Commission and Department of Water Affairs and Forestry. [online] <http://www.wrc.org.za/downloads/report%20lists/web%20rpts/ecosyst/TT378-09%20Conservation%20of%20water%20ecosystems.pdf> viewed 2009-06-17.

Kraak, M. J. 2006. Why maps matter in GIScience. The Cartographic Journal 43:82-89.

Lanthony, P. 2006. Des yeux pour peindre / An eye for painting. Réunion des musées nationaux, France.

Lauriault, T. P. and Lindgaard, G. 2006. Scented cybercartography: Exploring possibilities. Cartographica, Special issue on Cybercartography 41:73-90.

Leveson, N. G. 1994. High-pressure steam engines and computer software. Computer 27:65-73.

Leveson, N. G. 2002. Model-based analysis of socio-technical risk. Technical report, Engineering Systems Division, Massachusetts Institute of Technology.

Lewandowsky, S. and Spence, I. 1989. The perception of statistical graphs. Sociological Methods and Research 18:200-242.

Liechti, P and Jakob, A. 1992. Hydrologischer Atlas der Schweiz (Hydrological Atlas of Switzerland) - Plate 7.2, Average concentration of selected chemical parameters in surface waters. Switzerland. [online] <http://www.hydrologie.unibe.ch/hades/hades/tafel_72_en.html> viewed on 2010-05-16.

Loudon, M. 2007. Data management and reporting for drinking water quality monitoring in community managed supplies. Master's thesis, Department of Civil Engineering, University of Cape Town.

Loudon, M., Rivett, U., and Richards, T. 2006. Open source GIS and mobile devices for water demand management. In Proceedings of the 2nd Community Informatics For Developing Countries Conference.

Lozano-Fuentes, S., Elizondo-Quiroga, D., Farfan-Ale, J. A., Loroño-Pino, M. A., Garcia-Rejon, J., Gomez-Carro, S., Lira-Zumbardo, V., Najera-Vazquez, R., Fernandez-Salas, I., Calderon-Martinez, J., Dominguez-Galera, M., Mis-Avila, P., Morris, N., Coleman, M., Moore, C. G., Beaty, B. J., and Eisen, L. 2008. Use of Google Earth to strengthen public health capacity and facilitate management of vector-borne diseases in resource-poor environments. Bulletin of the World Health Organization 86:718-725.

MacEachren, A. 1994. Some Truth With Maps: A Primer on Symbolization and Design (Resource Publications in Geography). Association of American Geographers.

MacEachren, A. M. 2004. How Maps Work: Representation, Visualization, and Design. The Guilford Press.

MacEachren, A., Gahegan, M., and Pike, W. 2004. Visualization for constructing and sharing geo-scientific concepts. Proceedings of the National Academy of Sciences of the United States of America 101 Suppl 1:5279-5286.

Malan, H., Bath, A., Day, J., and Joubert, A. 2003. A simple flow-concentration modelling method for integrating water quality and water quantity in rivers. Water SA 29:305-311.

Martin, K 2007, Export to KML Extension Version 2.4 documentation, Bureau of Planning, City of Portland. [online] <http://outreach.cast.uark.edu/east/east/resources/Export_to_KML_documentation.pdf> viewed 2008-08-09.

Maucha, R. 1932. Hydrochemische Methoden in der Limnologie. Die Binnengewässer 12:173.

Maucha, R. 1949. The graphical symbolization of the chemical composition of natural waters. Hidrológiai Közlöny 13:117-118.

McDonald, R. D. 1989. New numbers for gauging stations, monitoring points and systems as well as for data and information TR 141 (in Afrikaans). Technical report, Department of Water Affairs, Pretoria.

Metzger, P. W. 1981. Managing a Programming Project. 2nd ed. Prentice Hall.

Microsoft 2004. Windows Movie Maker 2.1. [online] <http://www.microsoft.com/windowsxp/downloads/updates/moviemaker2.mspx> viewed 2008-08-10.

Miniwatts Marketing Group 2008. Africa Internet usage and population statistics. [online] <http://www.internetworldstats.com/africa.htm> viewed 2008-08-09.

Moolman J. 2007. Aquatic ecoregions of South Africa. Poster. Resource Quality Services, Department of Water Affairs and Forestry.

Nightingale, F. 1858. Notes on matters affecting the health, efficiency, and hospital administration of the British Army. Founded chiefly on the experience of the late war. Presented by request to the secretary of state for war. Privately printed. Original not seen, only copies of parts, e.g. URL <http://upload.wikimedia.org/wikipedia/commons/1/17/Nightingale-mortality.jpg>.

Noë, A. 2009. Out of Our Heads: Why You Are Not Your Brain, and Other Lessons from the Biology of Consciousness. Hill and Wang, 1st edition.

Noë, A. and Thompson, E. T. 2002. Vision and Mind: Selected Readings in the Philosophy of Perception. The MIT Press, first edition.

Nowotny, H. 2003. Democratising expertise and socially robust knowledge. Science and Public Policy 30:151-156.

O'Hanlon, M. and Campbell, J. 2007. Iraq index - tracking variables of reconstruction & security in post-Saddam Iraq. [online] <http://www.brookings.edu/fp/saban/iraq/index20070329.pdf> viewed 2009-05-08.

Ockerse, T. 1946. Endemic fluorosis in South Africa. PhD thesis, University of the Witwatersrand.

Oyelaran-Oyeyinka, B. and Lal, K. 2005. Internet diffusion in sub-Saharan Africa: A cross-country analysis. Telecommunications Policy 29:507-527.

Palmer, S.E. 1999. Vision science—photons to phenomenology, The MIT Press, Cambridge, Massachusetts.

Parliamentary Monitoring Group 2007. Delmas Diarrhea Outbreak 2007: Department briefing. [online] <http://www.pmg.org.za/print/9833> viewed 2009-04-10.

Petrovic, G. 1981. On the chemistry of some salt lakes and pools in Yugoslavia. Hydrobiologia 81:195-200.

Pienaar, H. and Xu, Y. 2007. Protecting what's underneath the tap. The Water Wheel pp. 24-27. [online] <http://www.wrc.org.za/downloads/waterwheel/jul-aug%2007/Gr%20wat%20p%2024-27.pdf> viewed 2009-06-16.

PKWARE 2007. Data security software. [online] <http://www.pkware.com/> viewed 2009-04-10.

Poole, A. 2005. Literature Review - Which Are More Legible: Serif or Sans Serif Typefaces? [online] <http://www.alexpoole.info/academic/literaturereview.html> viewed 2009-04-05.

Porathe, T. 2008. Measuring effective map design for route guidance - an experiment comparing electronic map display principles. Information Design Journal 16:190-201.

R Development Core Team (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, [online] <http://www.R-project.org> viewed 2009-06-16.

Reiner J. and Schaper, M. 2010. Tackling IT complexity in product design. McKinsey Quarterly. 2010 March: 1-9.

RHP 1998. State of River report: State of the Crocodile River [online] <http://www.dwaf.gov.za/iwqs/rhp/state_of_rivers.html> viewed 2009-04-05.

RHP 2000a. Poster: The Ecological State of the Crocodile River (Mpumalanga). [online] <http://www.dwaf.gov.za/iwqs/rhp/state_of_rivers.html> viewed 2009-04-05.

RHP 2000b. My home: a healthy river! Colouring-in Book (Finny Fish) in English, isiZulu and Sesotho. Department of Water Affairs and Forestry, Pretoria.

RHP 2001a. State of River report: Crocodile Sabie-Sand and Olifants Rivers [online] <http://www.dwaf.gov.za/iwqs/rhp/state_of_rivers.html> viewed 2009-04-05.

RHP 2001b. State of River report: Letaba and Luvuvhu Rivers [online] <http://www.dwaf.gov.za/iwqs/rhp/state_of_rivers.html> viewed 2009-04-05.

RHP 2001c. Poster: The Ecological State of the Modder River (Free State). [online] <http://www.dwaf.gov.za/iwqs/rhp/state_of_rivers.html> viewed 2009-04-05.

RHP 2001d. My home: a healthy river! Colouring-in Book (Finny Fish) in English, Afrikaans and isiXhosa. Department of Water Affairs and Forestry, Pretoria.

RHP 2002. State of River report: uMngeni River [online] <http://www.dwaf.gov.za/iwqs/rhp/state_of_rivers.html> viewed 2009-04-05.

RHP 2003a. State of River report: Hartenbos and Klein Brak Rivers [online] <http://www.dwaf.gov.za/iwqs/rhp/state_of_rivers.html> viewed 2009-04-05.

RHP 2003b. State of River report: Diep, Hout Bay, Lourens and Palmiet River Systems [online] <http://www.dwaf.gov.za/iwqs/rhp/state_of_rivers.html> viewed 2009-04-05.

RHP 2003c. State of River report: Free State Region [online] <http://www.dwaf.gov.za/iwqs/rhp/state_of_rivers.html> viewed 2009-04-05.

RHP 2003d. Poster: Hartenbos & Klein Brak (A3 Z-fold) [online] <http://www.dwaf.gov.za/iwqs/rhp/state_of_rivers.html> viewed 2009-04-05.

RHP 2003e. My home: a healthy river! Colouring-in Book (Finny Fish) in English, Afrikaans and Sesotho. Department of Water Affairs and Forestry, Pretoria.

RHP 2003f. Poster: The Ecological State of the Southern Gauteng Rivers [online] <http://www.dwaf.gov.za/iwqs/rhp/state_of_rivers.html> viewed 2009-04-05.

RHP 2004a. State of River report: Berg River System [online] <http://www.dwaf.gov.za/iwqs/rhp/state_of_rivers.html> viewed 2009-04-05.

RHP 2004b. State of River report: Buffalo River System [online] <http://www.dwaf.gov.za/iwqs/rhp/state_of_rivers.html> viewed 2009-04-05.

RHP 2004c. Poster: The Ecological State of the rivers of the Free State Region [online] <http://www.dwaf.gov.za/iwqs/rhp/state_of_rivers.html> viewed 2009-04-05.

RHP 2004d. Poster: Free State Region Fun Poster [online] <http://www.dwaf.gov.za/iwqs/rhp/state_of_rivers.html> viewed 2009-04-05.

RHP 2005a. State of River report: Crocodile(West) Marico WMA [online] <http://www.dwaf.gov.za/iwqs/rhp/state_of_rivers.html> viewed 2009-04-05.

RHP 2005b. State of River report: Greater Cape Town's Rivers [online] <http://www.dwaf.gov.za/iwqs/rhp/state_of_rivers.html> viewed 2009-04-05.

RHP 2005c. Poster: The Ecological State of the Crocodile (West) Marico (English and Afrikaans; English and seTswana). [online] <http://www.dwaf.gov.za/iwqs/rhp/state_of_rivers.html> viewed 2009-04-05.

RHP 2005d. Poster: The Ecological State of Cape Town's Rivers (English and isiXhosa). [online] <http://www.dwaf.gov.za/iwqs/rhp/state_of_rivers.html> viewed 2009-04-05.

RHP 2005e. Poster: The Ecological State of the Berg River (English and Afrikaans). [online] <http://www.dwaf.gov.za/iwqs/rhp/state_of_rivers.html> viewed 2009-04-05.

RHP 2005f. My home: a healthy river! Activity Book for the Eastern Cape Grade 1-3 Foundation Phase Learners. Department of Water Affairs and Forestry, Pretoria.

RHP 2005g. Water-wise with our rivers. Activity Book for the Free State Grade R-3 Foundation Learners in English and Afrikaans. Department of Water Affairs and Forestry, Pretoria.

RHP 2006a. State of River report: Olifants/Doring and Sandveld Rivers [online] <http://www.dwaf.gov.za/iwqs/rhp/state_of_rivers.html> viewed 2009-04-05.

RHP 2006b. State of River report: Mokolo River [online] <http://www.dwaf.gov.za/iwqs/rhp/state_of_rivers.html> viewed 2009-04-05.

RHP 2006c. Poster: The Ecological State of the Goukou and Duiwenhoks Rivers [online] <http://www.dwaf.gov.za/iwqs/rhp/state_of_rivers.html> viewed 2009-04-05.

RHP 2006d. Poster: The Ecological State of the Rivers of the Overberg Region. [online] <http://www.dwaf.gov.za/iwqs/rhp/state_of_rivers.html> viewed 2009-04-05.

RHP 2006e. Rivers are fun! Activity Book for the Western Cape Foundation Phase Learners (grade 1-3). Department of Water Affairs and Forestry, Pretoria. [online] <http://www.dwaf.gov.za/iwqs/rhp/state_of_rivers.html> viewed 2009-04-05.

RHP 2006f. Rivierpret! Aktiwiteitsgids vir die Wes-Kaapse Grondslagfase Leerders (Graad 1-3). Department of Water Affairs and Forestry, Pretoria. [online] <http://www.dwaf.gov.za/iwqs/rhp/state_of_rivers.html> viewed 2009-04-05.

RHP 2007a. State of River report: Rivers of the Gouritz Water Management Area. [online] <http://www.dwaf.gov.za/iwqs/rhp/state_of_rivers.html> viewed 2009-04-05.

RHP 2007b. Poster: Habitat Integrity of Selected Rivers of the North West Province: Major impacts and management actions. [online] <http://www.dwaf.gov.za/iwqs/rhp/state_of_rivers.html> viewed 2009-04-05.

RHP 2007c. Yonwabele imilambo yetho! Incwadana yemisebenzi yabaFundi besiga esisisiseko baseNtshona Koloni (ibanga loku-1 ukuya kwelesi – 3). CSIR, Pretoria. [online] <http://www.dwaf.gov.za/iwqs/rhp/state_of_rivers.html> viewed 2009-04-05.

Richards, P. R. 1998. Estimation of pollutant loads in rivers and streams - a guidance document for NPS programs - grant x998397-01-0. Technical report, U.S. Environmental Protection Agency, Region VIII, Water Quality Laboratory, Tiffin, Ohio.

Rooseboom, A. and Lotriet, H. H. 1992. The new sediment yield map for southern Africa. In Erosion and Sediment Transport Monitoring Programmes in River Basins, pp. 527-538. IAHS Publ. no. 210.

Rosling, H. 2007a. Video: New insights on poverty and life around the world. [online] < http://www.ted.com/index.php/talks/hans_rosling_reveals_new_insights_on_poverty.html> viewed 2009-05-08.

Rosling, H. 2007b. Video: Unveil the beauty of statistics. [online] < http://radar.oreilly.com/archives/2007/07/oecd-world-forum.html> viewed 2009-05-11.

Roux, D. J., Rogers, K. H., Biggs, H. C., Ashton, P. J., and Sergeant, A. 2006. Bridging the science–management divide: moving from unidirectional knowledge transfer to knowledge interfacing and sharing. Ecology and Society 11:4+.

Roux, D. J., Hill, L., and Strydom, W. 2008. Assessing the impact of research funded by the Water Research Commission in support of the River Health Programme. WRC Report No. TT 360/08. Technical report, Water Research Commission, Pretoria.

Ruys, K. I. and Stapel, D. A. 2008. The secret life of emotions. Psychological Science 19:385-391.

Saner, M. A. 1999. Two cultures: not unique to ecology. Conservation Ecology 3(1):r2. [online] <http://www.consecol.org/vol3/iss1/resp2/> visited 2009-06-06.

Schager, E. and Peter, A. 2004. Méthodes d'analyse et d'appréciation des cours d'eau en Suisse: Poissons niveau R (région) MGS-44-F.

Schermer, M. 2007. The really hard science. Scientific American 297(4):23-24.

Schnabel, O. 2007. Benutzerdefinierte Diagramm-signaturen in Karten: Konzepte, Formalisierung und Implementation. Eidgenössischen Technischen Hochschule Zürich. [online] <http://e-collection.ethbib.ethz.ch/eserv/eth:29352/eth-29352-01.pdf> 2010-05-16

Schnabel, O. and Hurni, L. 2009. Primitive-based Construction Theory for Diagrams in Thematic Maps. Cartographic Journal. 46(2):136-145.

Schulze, R. E. 2006. Water Resources Complexities and Hydrological Needs in South Africa: Background, in R. E. Schulze (ed.), South African Atlas of Climatology and Agrohydrology, Water Research Commission, Pretoria, South Africa, WRC Report 1489/1/06, Section 21.1.

Schulze, R. E. and Lynch, S. D. 2006. Annual Precipitation, in R. E. Schulze (ed.), South African Atlas of Climatology and Agrohydrology, Water Research Commission, Pretoria, South Africa, WRC Report 1489/1/06, Section 6.2.

Schulze, R. E. and Maharaj, M. 2006. A-Pan Equivalent Reference Potential Evaporation, in R. E. Schulze (ed.), South African Atlas of Climatology and Agrohydrology, Water Research Commission, Pretoria, South Africa, WRC Report 1489/1/06, Section 13.2.

Scott, W. 1817. Rob Roy. Penguin Books, London.

Scottish Natural Heritage. 2002. Loch Lomond, The Trossachs and Breadalbane. Scottish Natural Heritage, Perth, Scotland. [online] <http://www.snh.org.uk/futures/Data/pdfdocs/Loch_Lomond.pdf> viewed 2009-06-07.

Sidley, P. 2005. Typhoid outbreak prompts protests over inadequate water system. BMJ 331:655+. [online] <http://www.bmj.com/cgi/content/extract/331/7518/655> viewed 2009-05-08.

Silberbauer, M. J. 1979. The effect of a storm over the catchment area on the inflow of nutrients to Roodeplaat Dam-internal document Hydrological Research Institute. Technical report, Pretoria.

Silberbauer, M.J. 1995. Monitoring temporal changes in the ionic ratios of surface waters using Maucha diagrams. Poster presented at the Zoological and Aquatic Sciences Conference, Grahamstown, 27-30 June 1995.

Silberbauer, M. J. 1996. Biomonitoring mapping. [online] <http://www.dwaf.gov.za/iwqs/water_quality/NCMP/av/biomon_2007.zip>. Accessed 2009-03-28.

Silberbauer, M. J. 2006. The construction of a hydrologically-correct, annotated 1:500 000 spatial dataset of the rivers of South Africa and contiguous basins - (N/000/00/REH/0701). Technical report, Resource Quality Services, Department of Water Affairs and Forestry, Pretoria, South Africa. [online] <http://www.dwaf.gov.za/iwqs/gis_data/river/River_Report_01.pdf> viewed 2008-08-09.

Silberbauer, M. J. 1997. The application of geographic information systems to water quality monitoring. In Proceedings of Rabat Symposium S3 Remote Sensing and Geographic Information Systems for Design and Operation of Water Resources Systems, volume 242, pp. 189-195. IAHS.

Silberbauer, M. J. and Geldenhuys, W. G. 2008. Using Keyhole Markup Language to create a spatial interface to South African water resource data through Google Earth. In Proceedings of the FOSS4G 2008 conference. [online] < http://www.foss4g.org/index.php/foss4g/2008/paper/view/74/28> viewed 2009-04-10>.

Silberbauer, M. and Geldenhuys, W. 2009. Google Earth - a spatial interface for SA water resource data. PositionIT pp. 42-47. [online] <http://www.eepublishers.co.za/images/upload/PosIT_Apr-May_p_42-4734-37.pdf> viewed 2009-06-16.

Silberbauer, M. J. and Hohls, B. C. 2002. Mapping water quality on a desktop computer - poster. In SAIE&ES & SASAqS joint congress (SASAqS's 37th annual congress) on research, conservation and management of ecosystems in Southern Africa. Southern African Society for Aquatic Sciences. [online] < http://www.dwaf.gov.za/iwqs/water_quality/NCMP/Silberbauer_2002_Mapping_Water_Quality_Desktop_Computer.pdf> viewed 2009-03-24.

Silberbauer, M. J. and Howman, A. M. 2001. National land cover atlas for the resource directed measures programme-N/0000/00/REQ/0700. Technical report, Institute for Water Quality Studies, Department of Water Affairs and Forestry, Pretoria , South Africa.

Silberbauer, M. J. and King, J. M. 1991. Geographical trends in the water chemistry of wetlands in the south-western Cape Province, South Africa. South African Journal of Aquatic Science 17:82-88.

Slocum, T. A., Yoder, S. C., Kessler, F. C., and Sluter, R. S. 2000. Maptime: Software for exploring spatiotemporal data associated with point locations. Cartographica 37:15-32.

Slocum, T. A., Mcmaster, R. B., Kessler, F. C., and Howard, H. H. 2005. Thematic Cartography and Geovisualization. Prentice Hall, second edition.

Slocum, T. A., McMaster, R. B., Kessler, F. C., and Howard, H. H. 2008. Thematic Cartography and Geovisualization. Prentice Hall, third edition.

Small, H. 1998. Florence nightingale's statistical diagrams. In Stats and Lamps Research Conference. Florence Nightingale Museum.

Spence, I. 2005. No humble pie: The origins and usage of a statistical chart. Journal of Educational and Behavioral Statistics 30:353-368.

Spence, I. and Lewandowsky, S. 1991. Displaying proportions and percentages. Applied Cognitive Psychology 5:61-77.

Stiff, H. A. 1951. The interpretation of chemical water analysis by means of patterns. Journal of Petroleum Technology 3:sec.1, 15-16-sec.2, 3.

Strahler, A. N. 1952. Hypsometric (area-altitude) analysis of erosional topography. Geological Society of America Bulletin. 63(11):1117-1142.

Strydom, W. F., van Wyk, E., Maree, G., and Maluleke, T. P. 2002. The evolution of state-of-rivers reporting in South Africa. In 7th International Conference on Public Communication. University of Cape Town.

Strydom, W. F. 2003. Compiling State of Rivers reports and posters: a manual. NAEBP Report Series No 17. Resource Quality Services, Department of Water Affairs and Forestry, Pretoria, South Africa. [online] <http://www.csir.co.za/rhp/reports/reportseries17.pdf> visited 2009-06-06.

Strydom, W. F. 2009a. The impact of state-of-rivers reporting on people's attitudes towards river conservation a case study of the Buffalo and Hartenbos + Klein Brak catchments in South Africa. Master's thesis, Stellenbosch University.

Strydom, W. F. 2009b. Spreading the conservation message – are we doing it right? The Water Wheel 2009:23-25.

Strydom, W. F., Hill, L., and Eloff, E. 2006. Achievements of the River Health Programme 1994-2004: A national perspective on the ecological health of selected South African rivers. Technical report, Department of Water Affairs and Forestry, the Deparment of Environmental Affairs and Tourism, the Water Research Commission and the CSIR, Pretoria.

Strydom, W., Hill, L., and Eloff, E. 2007. The Role of Communication in Governance: The River Health Programme as a Case Study. Springer-Verlag, Berlin, Heidelberg.

Swart, S. J., van Veelen, M., and Nell, U. 1991. Water quality data inventory, volume 1: Drainage regions A, B, C, D, E, F, TR146. Technical report, Hydrological Research Institute, Department of Water Affairs, Pretoria.

Tamás, F. D. and Abonyi, J. 2002. Trace elements in clinker i. a graphical representation. Cement and Concrete Research 32:1319-1323.

Thirion, C. 2007. Module E: Macroinvertebrate response assessment index in river ecoclassification: Manual for ecostatus determination (version 2) WRC Report No. TT 332/08. Technical report, Water Research Commission and Department of Water Affairs and Forestry. [online] <http://www.wrc.org.za/downloads/report%20lists/web%20rpts/ecosyst/TT%20332-CONSERVATION.pdf> viewed 2009-06-17.

Tonjes, D. J., Heil, J. H., and Black, J. A. 1995. A sophisticated ground water analytical tool. Ground Water Monitoring & Remediation 15:134-139.

Trafton, J. G., Trickett, S. B., and Mintz, F. E. 2005. Connecting internal and external representations: Spatial transformations of scientific visualizations. Foundations of Science 10:89-106.

Trickett, S. B. and Trafton, J. G. 2006. Toward a comprehensive model of graph comprehension: making the case for spatial cognition. In: Barker-Plummer D, Cox R, Swoboda N, editors. Diagrammatic Representation and Inference - 4th International Conference on Theory and Application of Diagrams. p. 286-300.

Tufte, E. R. 1977. Review: Political statistics for the United States: Observations on some major data sources. The American Political Science Review 71:305–314.

Tufte, E. R. 1997. Visual Explanations: Images and Quantities, Evidence and Narrative. Graphics Press.

Tufte, E. R. 2001. The visual display of quantitative information. Graphics Press, Cheshire, Connecticut, second edition.

Tufte, E. 2003. Essay: The cognitive style of Powerpoint: Pitching out corrupts within. Cheshire, CT: Graphics Press.

Tukey, J. 1965. The technical tools of statistics. American Statistician 19:23-28.

Underwood, L. M. 1899. The Royal Botanic Gardens at Kew. Science 10:65-75.

Uys, M. and O'Keeffe, J. 1997. Simple words and fuzzy zones- early directions for temporary river research in South Africa. Environmental Management 21:517-531.

Valero, T. 2009. Quelques outils pour manipuler les fichiers KML pour les SIH des projets HYCOS - IRD/OBHI. Technical report, Institut de Recherche pour le Développement.

van Niekerk, H., Silberbauer, M., and 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. Environmental Monitoring and Assessment. [online] <http://dx.doi.org/10.1007/s10661-008-0407-2> viewed 2009-05-08.

van Veelen, M., Nell, U., and Geldenhuys, W. F. 1990. Surface water quality of South Africa 1979-1988. Volume 2: Drainage region A and B. TR145. Technical report, Hydrological Research Institute, Department of Water Affairs and Forestry, Pretoria.

van Vliet, H. R. and Nell, U. 1986. Surface water quality of South Africa, the Vaal River catchment 1979-1983, TR131. Technical report, Hydrological Research Institute, Department of Water Affairs and Forestry, Pretoria.

van Vuuren, L. 2007. Joining hands to manage SA's water resources. The Water Wheel July/August:14-17.

van Wyk, E., Roux, D., Drackner, M., and McCool, S. 2008. The impact of scientific information on ecosystem management: Making sense of the contextual gap between information providers and decision makers. Environmental Management 41:779-791.

Vegter, J. R. 1995. An explanation of a set of national groundwater maps, WRC Report TT 74/95. Technical report, Water Research Commission, Pretoria, South Africa.

Wainer, H. 2007. Visual revelations. Improving data displays: Ours and the media's. Chance 20:8-15.

Wainer, H. 2008. Visual revelations - improving graphic displays by controlling creativity. Chance 21:46-52.

Wainer, H. and Thissen, D. 1981. Graphical data analysis. Annual Review of Psychology 32:191-241.

Walker, W. W. 1999. Simplified procedures for eutrophication assessment and prediction: User manual instruction report W-96-2. Technical report, U.S. Army Engineer Waterways Experiment Station, Vicksburg, Mississippi.

Wallis, W. A. 1949. Statistics of the Kinsey report. Journal of the American Statistical Association 44:463-484.

Ware, C. 2004. Information Visualization, Second Edition: Perception for Design (Interactive Technologies). Morgan Kaufmann.

Ware, C. 2008. Visual Thinking: for Design (Morgan Kaufmann Series in Interactive Technologies). Morgan Kaufmann, illustrated edition.

Weber, J. R. and Word, C. S. 2001. The communication process as evaluative context: What do nonscientists hear when scientists speak? BioScience 51:487-495.

Wells, N. 1999. ASTRA.BAS: a program in QuickBasic 4.5 for exploring rose diagrams, circular histograms and some alternatives. Computers & Geosciences 25:641-654.

Whalley, J. I. 1980. The art of calligraphy - Western Europe and America. Bloomsbury Books, London, UK.

Whitfield, A. K. 2003. Available scientific information on individual South African estuarine systems (updated). Technical report, Grahamstown, South Africa.

Wilson, J., Dai, M., Jakupovic, E., Watson, S. and Meng F. 2007. Supercomputing with toys: harnessing the power of NVIDIA 8800GTX and Playstation 3 for bioinformatics problem. Computational systems bioinformatics / Life Sciences Society Computational Systems Bioinformatics Conference. 6:387-390.

Wolfenden, L. 1999. Impact of state of the environment reports on environment agency staff. Master's thesis, Imperial College of Science, Technology and Medicine, London.

Appendices

1. Karst diagram

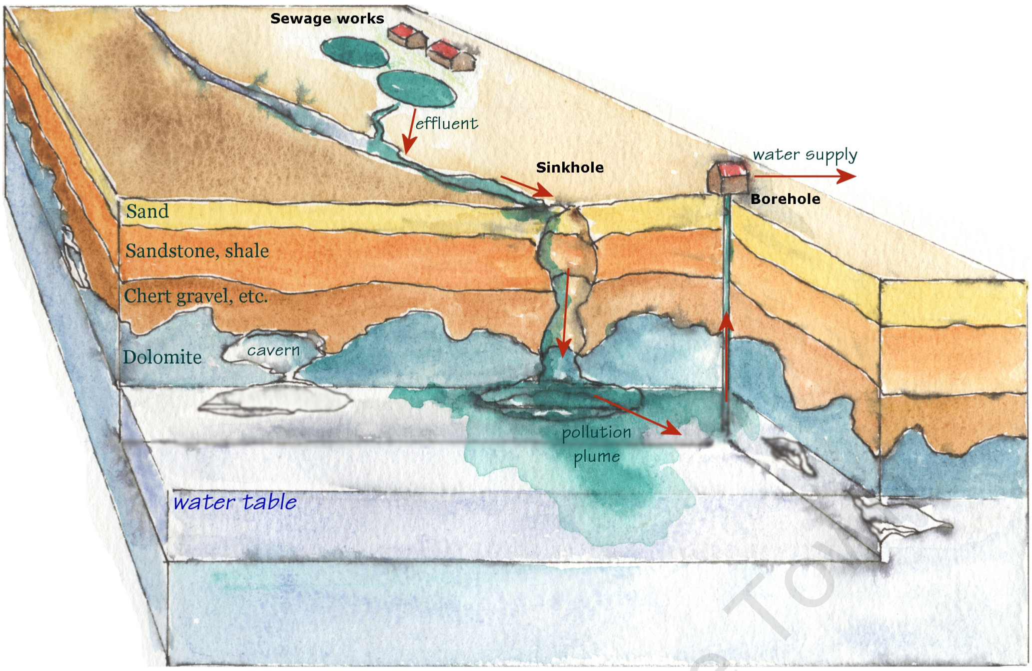


Figure A1-1. (Draft not used for publication, showing that ink and wash is still a viable illustration technique in the computer age: Computer software was nevertheless handy for lettering.) Possible route for contamination of groundwater in a dolomitic or karst area, especially in cases where excessive abstraction has lowered the water table. The diagram shows how sinkhole formation can open "windows" into the water table, allowing much more rapid transfer of pollutants than might be expected by diffusion through geological layers. After diagrams in Currens (1995), Alpha et al. (ca. 1997), Jones and Wagener (2004), and Zhang (2006).

Alpha, T. R., Galloway, J. P., and Tinsley III, J. C. Karst topography-black and white and color paper models - Open-file Report 97-536-A. Technical report, U.S. Department of the Interior, U.S. Geological Survey. [online] < http://www.usgs.gov/education/animations/karst97-536/karst.pdf> viewed 2009-07-04.

Currens, J. C. (1995) Generalized block diagram of the western Pennyroyal Karst. Kentucky Geological Survey and University of Kentucky, Lexington.

Jones & Wagener, 2004, Ekurhuleni metropolitan municipality: first progress report on sinkholes at Bapsfontein, Report no. JW21/04/9401, Jones & Wagener Consulting Civil Engineers, 59 Bevan Road, PO Box 1434, Rivonia, 2128, South Africa (as cited in Zhang, 2006).

Zhang, J (2006) Groundwater management plan for Delmas municipal council for VGC Consulting Engineers (Pty) Ltd. Report number: 2005.08.399 May 2006.

2. Political cartoons

Political cartoons are a powerful visual commentary on current events, using deceptively simple cues. This appendix provides local background information necessary to understand the rather bleak humour in the cartoons used in the text.

Figure A2-1. Illustration by Tony Grogan of the reaction of the Northern Cape administration to the participation of a government ornithologist, Dr M. Anderson, in public comment against the alleged mismanagement of a Kimberley lake containing a flamingo reserve (from the Sunday Independent newspaper). The allegations included poor management of sewage treatment works discharging into the lake and collusion between property developers and local government.

(Anderson later resigned.)

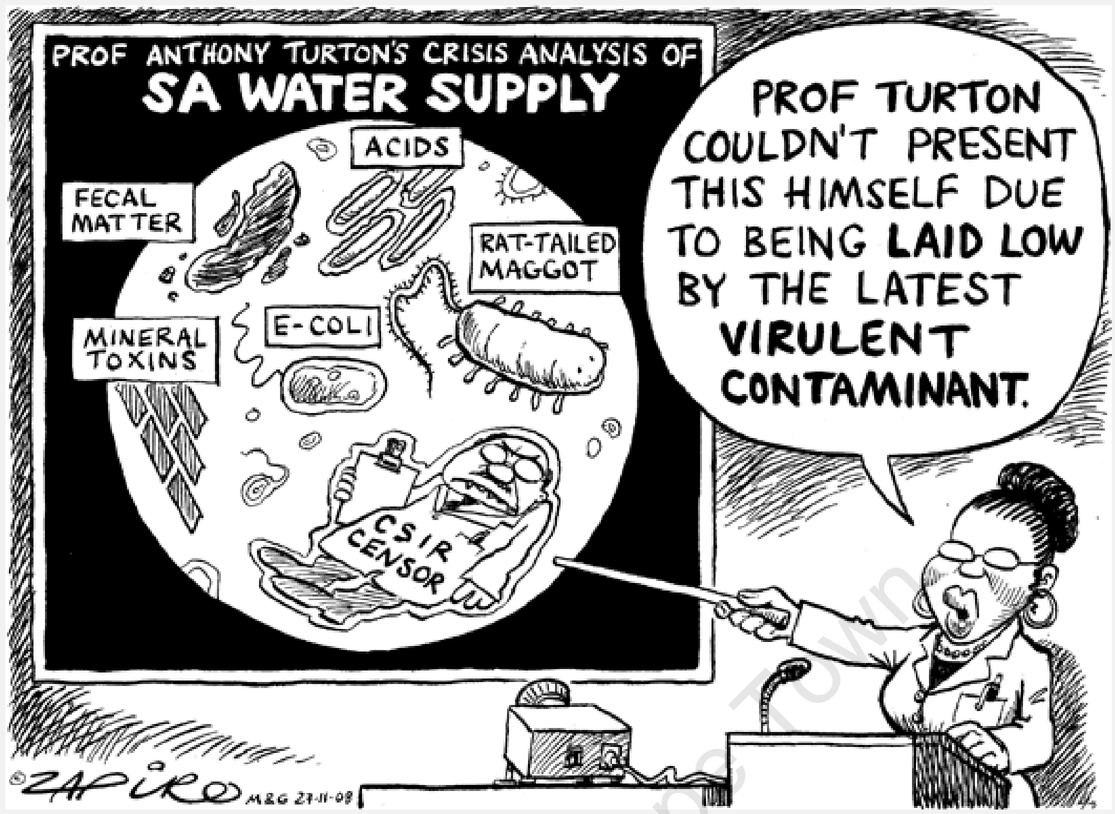
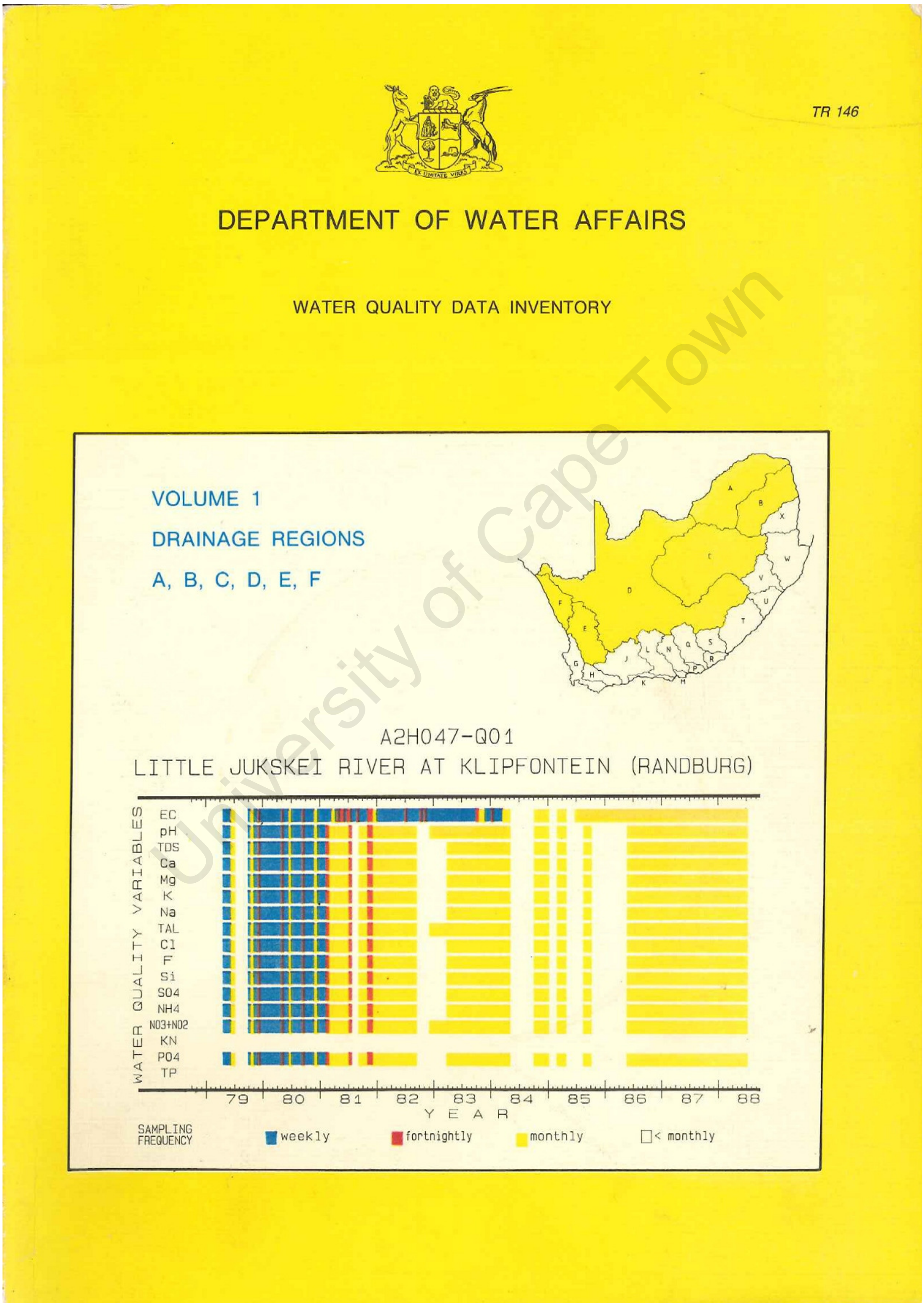


Figure A2-2. Illustration by Jonathan Shapiro ("Zapiro", the Mail and Guardian newspaper) of the reaction of the CSIR management to a paper by Dr A. Turton (Turton 2008) which made allegedly unsubstantiated claims about political management of science in South Africa during the past twenty-five years. The CSIR withdrew permission for Turton to present the paper at an internal colloquium, but allowed distribution of the electronic version of the paper. The cartoonist has produced a whimsical "microscope slide" containing his impression of various water quality concerns that had appeared in the press during the previous months. Note that he has portrayed the "CSIR Censor" as a white man and the water quality specialist as a black woman to downplay suggestions of victimisation of a white male scientist by a black administration, thus keeping the spotlight on the problem of water quality (compare this with Grogan's cartoon, Appendix 2).

(Turton later resigned.)

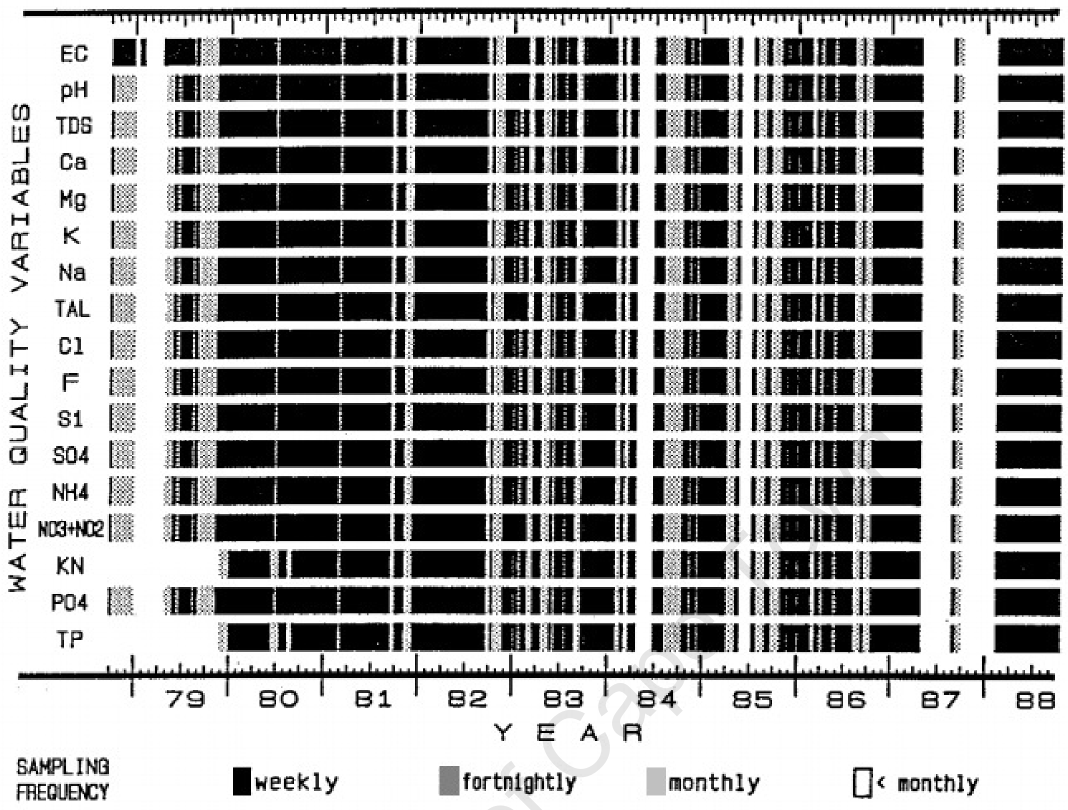
3. Technical Report TR146

Enlargements of the images in Figure 2-1: : Cover and example page of the first volume in the "Yellow Book" water quality monitoring site series, Technical Report TR 146 (Swart et al. 1991).



A2H028-Q01

HARTEBEES SPRUIT AT KAMEELDRIFT



OLD STATION NUMBER : A2M28

LOCATION

LATITUDE AND LONGITUDE : 25 39'03" 28 19'10"
TERTIARY DRAINAGE REGION No. : 123

DESCRIPTION OF CATCHMENT

CATCHMENT AREA (km^2) : 161
CALCULATED MEAN ANNUAL RUNOFF (10^6 m^3) : 9
MAIN TRIBUTARIES UPSTREAM : --
MAIN RESERVOIRS UPSTREAM : --

DESCRIPTION OF MONITORING STATION

TYPE : GAUGING WEIR
AUTOMATIC FLOW RECORDER (Y/N) : Y

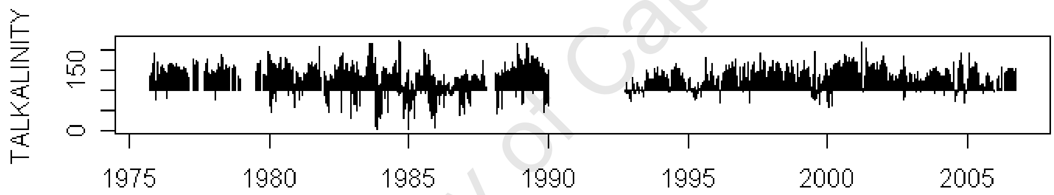
SAMPLING

DATE OF FIRST SAMPLE ANALYSED : 67/05/10
DATE OF LAST SAMPLE ANALYSED : 89/08/29
TOTAL NUMBER OF SAMPLES ANALYSED : 2315
CURRENT SAMPLING FREQUENCY
(Daily, Weekly, Monthly, Quarterly) : W

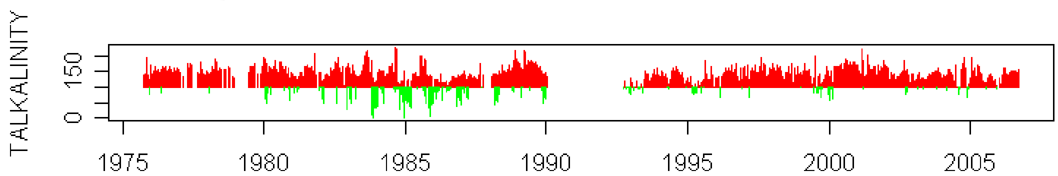
4. R-statistics

Additional examples of the use of the R-statistics package to generate one-page summaries of the sampling intervals, data values and guideline exceedance at a monitoring site. See Chapter 2 for more information.

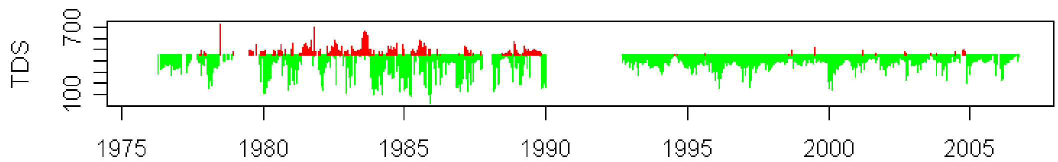
```
xx <- read.csv("C:/data_large/wms/a23/a23_90174.txt")
EC <- xx[[10]]
TDS <- xx[[21]]
plot(EC,TDS)
(z <- line (EC,TDS))
abline(coef(z))
Dat <- as.Date(xx[[3]],"%d-%b-%Y")
plot(Dat,EC)
history()

gln <- rep(100,3990)
gld <- data.frame(Dat,TALKALINITY,gln) [not sure if needed]
plot(Dat,TALKALINITY,type="n")
segments(Dat,TALKALINITY,Dat,gln)
history(20)
```



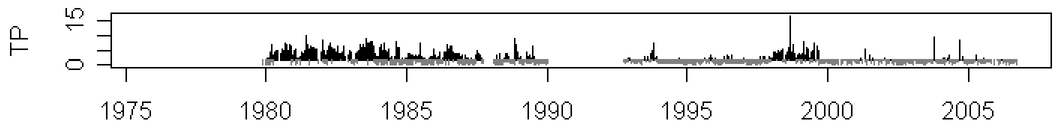

```
plot(Dat,TALKALINITY,type="n")

segments(Dat[TALKALINITY>100],TALKALINITY[TALKALINITY>100],Dat[TALKAL
INITY>100],gln[TALKALINITY>100],col="red")

segments(Dat[TALKALINITY<=100],TALKALINITY[TALKALINITY<=100],Dat[TALK
ALINITY<=100],gln[TALKALINITY<=100],col="green")
```



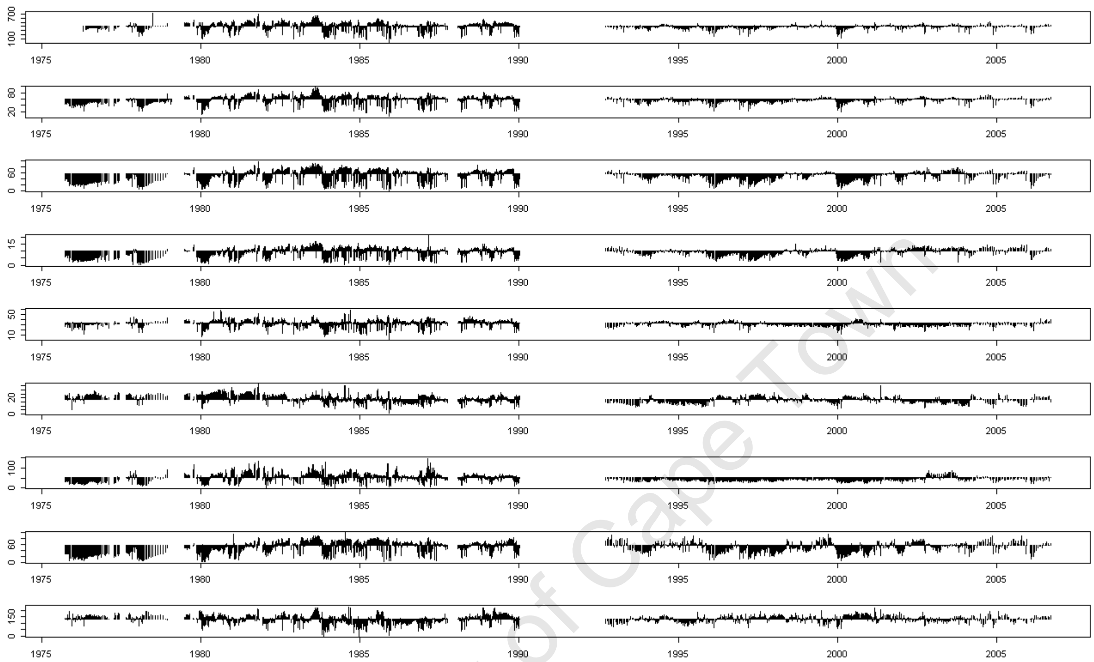

```
plot(Dat,TDS,type="n")
segments(Dat[TDS<=450],TDS[TDS<=450],Dat[TDS<=450],gln[TDS<=450],col=
"green")
segments(Dat[TDS>450],TDS[TDS>450],Dat[TDS>450],gln[TDS>450],col="red
")
```



```
gln <- rep(summary(TP)[3],length(TP))
plot (Dat,TP,type="n")
segments(Dat,TP,Dat,gln)
segments(Dat[TP<gln],TP[TP<gln],Dat[TP<gln],gln[TP<gln],col="gray50")
```




```
format(Sys.time(), "%Y-%m-%d %H:%M:%S")
```



```
 op <- par(mfrow = c(10,1), mar=0.1+c(2,2,2,2))
gln <- rep(summary(TDS)[3],length(TDS))
plot (Dat,TDS,type="n")
segments(Dat,TDS,Dat,gln)
gln <- rep(summary(CONDUCTIVITY)[3],length(CONDUCTIVITY))
plot (Dat,CONDUCTIVITY,type="n")
segments(Dat,CONDUCTIVITY,Dat,gln)
gln <- rep(summary(SODIUM)[3],length(SODIUM))
plot (Dat,SODIUM,type="n")
segments(Dat,SODIUM,Dat,gln)
gln <- rep(summary(POTASSIUM)[3],length(POTASSIUM))
plot (Dat,POTASSIUM,type="n")
segments(Dat,POTASSIUM,Dat,gln)
gln <- rep(summary(CALCIUM)[3],length(CALCIUM))
plot (Dat,CALCIUM,type="n")
segments(Dat,CALCIUM,Dat,gln)
gln <- rep(summary(MAGNESIUM)[3],length(MAGNESIUM))
plot (Dat,MAGNESIUM,type="n")
segments(Dat,MAGNESIUM,Dat,gln)
gln <- rep(summary(SULPHATE)[3],length(SULPHATE))
plot (Dat,SULPHATE,type="n")
segments(Dat,SULPHATE,Dat,gln)
gln <- rep(summary(CHLORIDE)[3],length(CHLORIDE))
plot (Dat,CHLORIDE,type="n")
segments(Dat,CHLORIDE,Dat,gln)
gln <- rep(summary(TALKALINITY)[3],length(TALKALINITY))
plot (Dat,TALKALINITY,type="n")
```

```
segments(Dat,TALKALINITY,Dat,gln)
```

5. Maucha generation

This is an outline of the functioning of the program to generate Maucha ionic diagrams. The actual code is on the enclosed DVD.

```
Obtain concentrations in milligrams per litre from database.

Check that values are available for:

      K+

      Na+

      Ca++

      Mg++

      SO4=

      Cl-

      TAL as HCO3-

      Eighth item (CO3=) left blank or used for high NO3

Convert to milliequivalents per litre and perform ion balance
([cations]=[anions])

Calculate geometry:

      Eight-pointed star, with cations to the right and anions to the
      left.

      Make area of each point proportional to ionic concentration of
      each ion.

Generate symbol as geometric shapes in ArcInfo or a grouped object in
ArcView.
```

6. DWAF site codes

At first, a site's primary and secondary drainage region and the site type (reservoir or hydrological) were embedded in the site code according to the departmental standard (McDonald 1989). When the water quality monitoring site code changed, a more complex geographical site selection method was necessary. This prompted the development of a process to get to the required site using Google Earth, as described in Chapter 4, rather than expending a great deal of effort on developing a spatial interface.

The hydrological site code (TR) consists of

A2 secondary drainage region

H type of station

027 site sequence within secondary drainage

Q type of monitoring

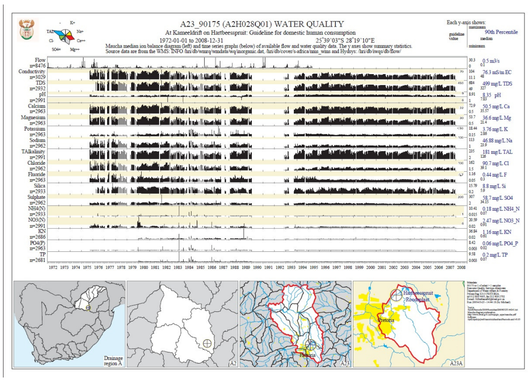
01 sub-site, e.g. in a dam

DWAF's WMS database designers, when migrating the water quality data from the HIS to the WMS changed the meaningful A2H027Q01 format to a supposedly sequential index separate from the description and locality format.

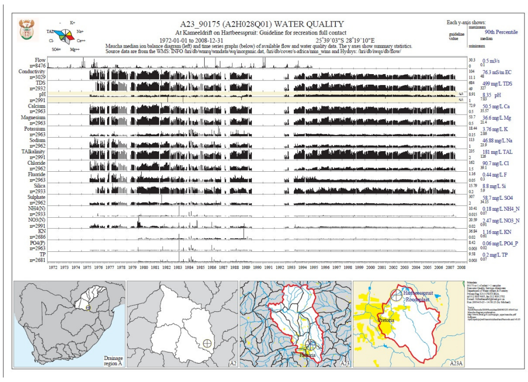
Data capturers soon misused the new sequential WMS index by inserting an initial digit to indicate the region, creating long numbers that exceeded the column width of the rigid WaterMarque card-type data format.

Convoluted coding was required to recognise and extract the HIS pattern embedded in the WMS description in order to determine secondary drainage region and type. For backward compatibility with the rigid WaterMarque data standard, the two types of site code were retained. This is instructive as it illustrates the hazards of embedding information in a database code, and the knock-on effects of arbitrary changes to database conventions.

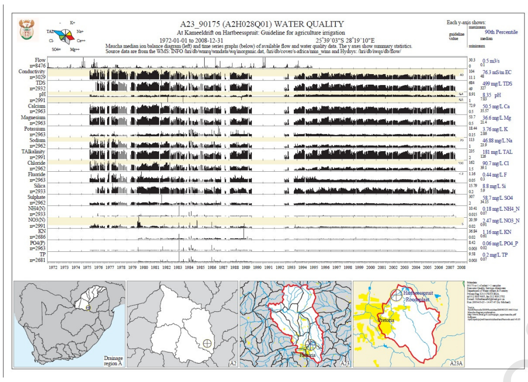
7. Barcode graphs with guidelines (thumbnails)



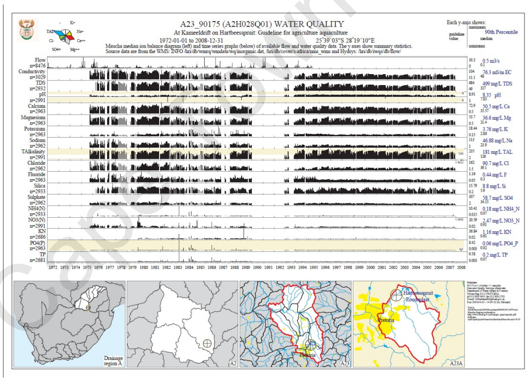
a) Domestic: human consumption



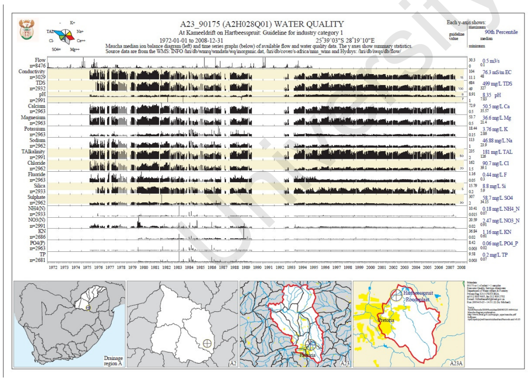
b) Recreation: full contact



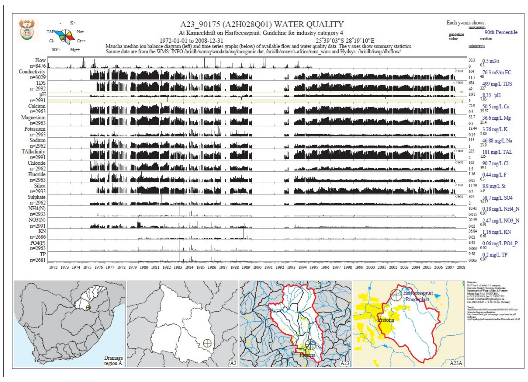
c) Agriculture: irrigation



d) Agriculture: aquaculture



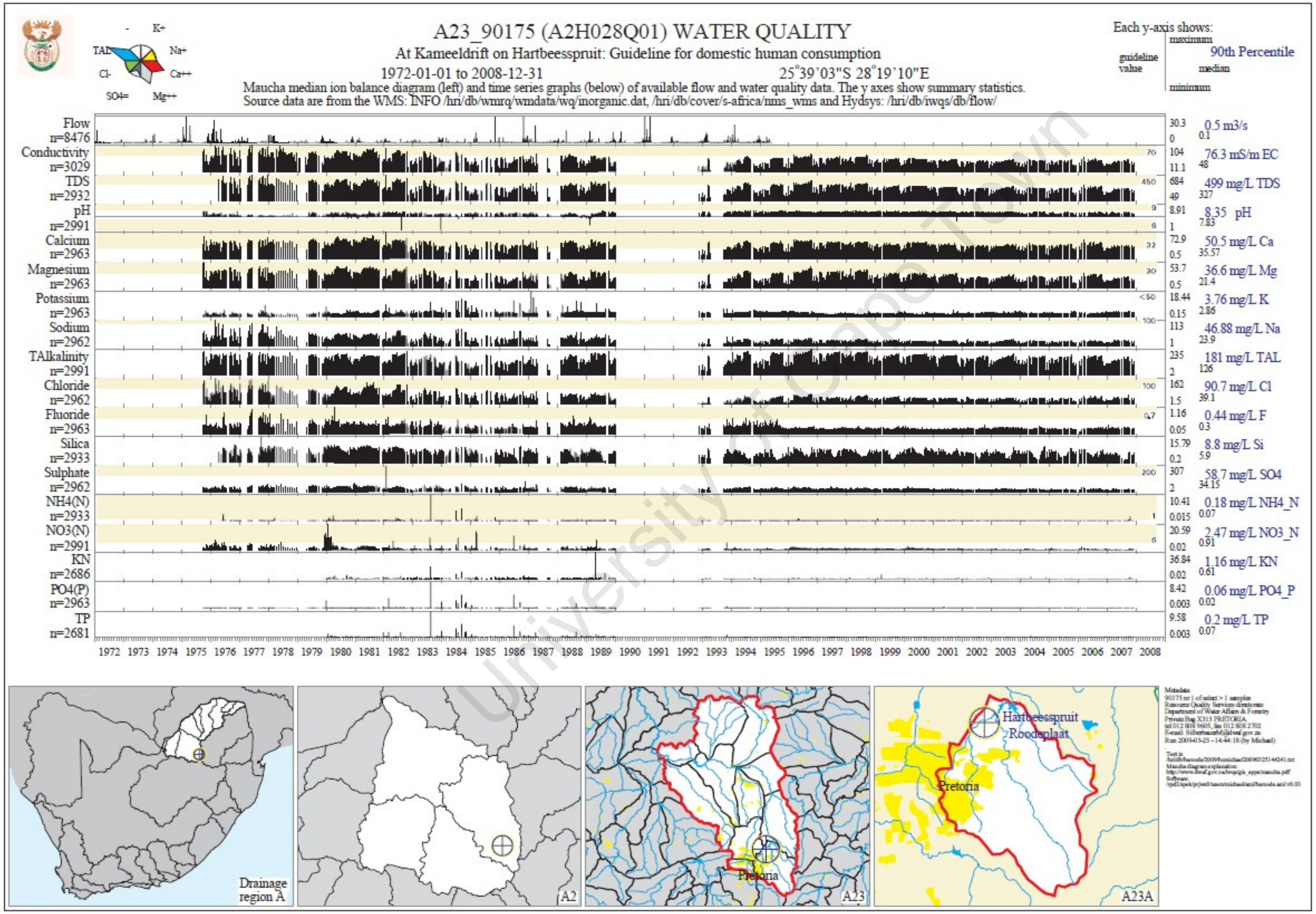
e) Industry: category 1



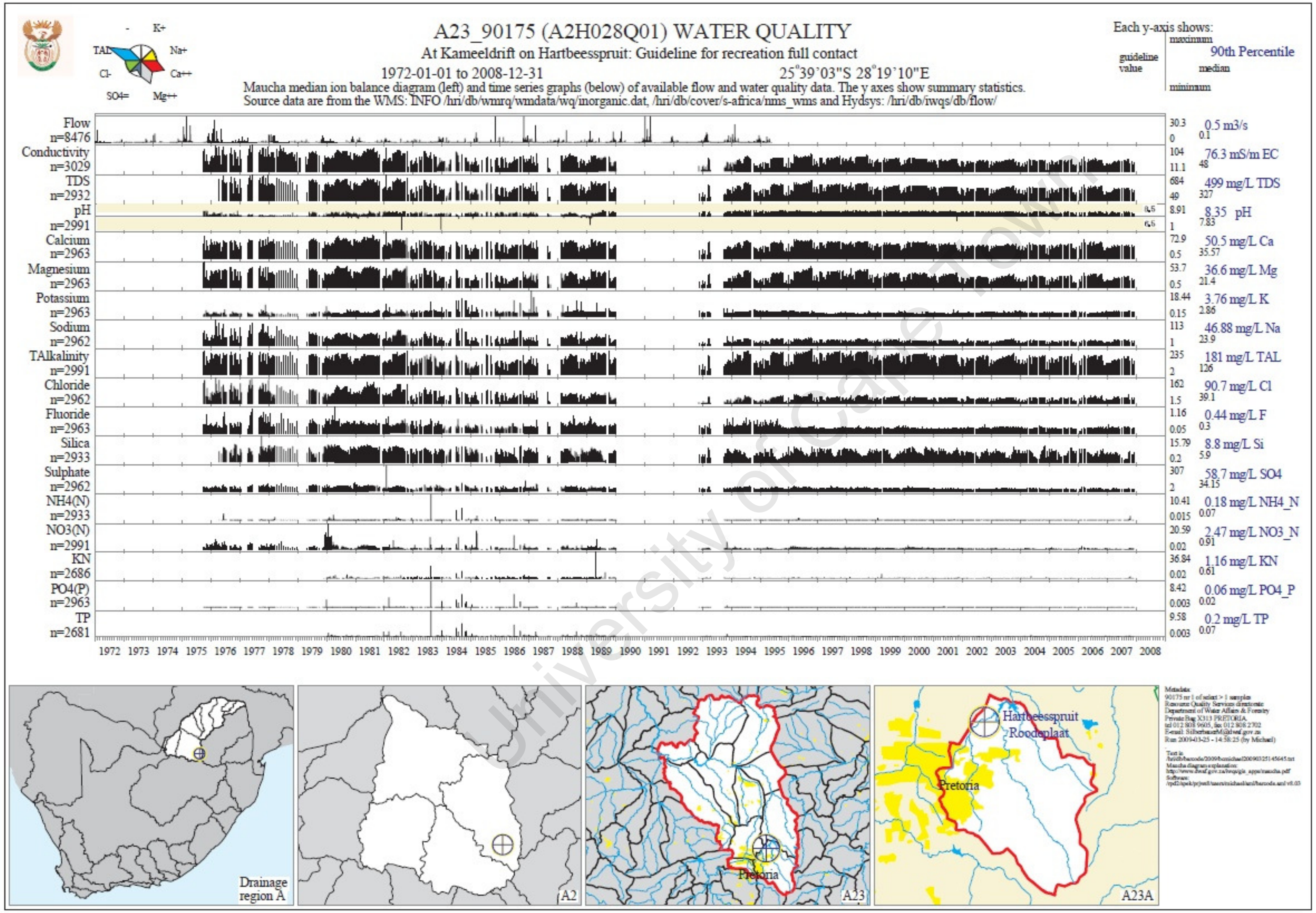
f) Industry: category 4

Figure A7-1. Examples of Barcode output with DWAF (1996) guidelines for six user types as shaded rectangles behind the time series plots. The actual guidelines for these uses include solutes not plotted by Barcode, so this type of display is only an indicator of problems, not of suitability for use. Larger copies of the graphs appear on the following pages.

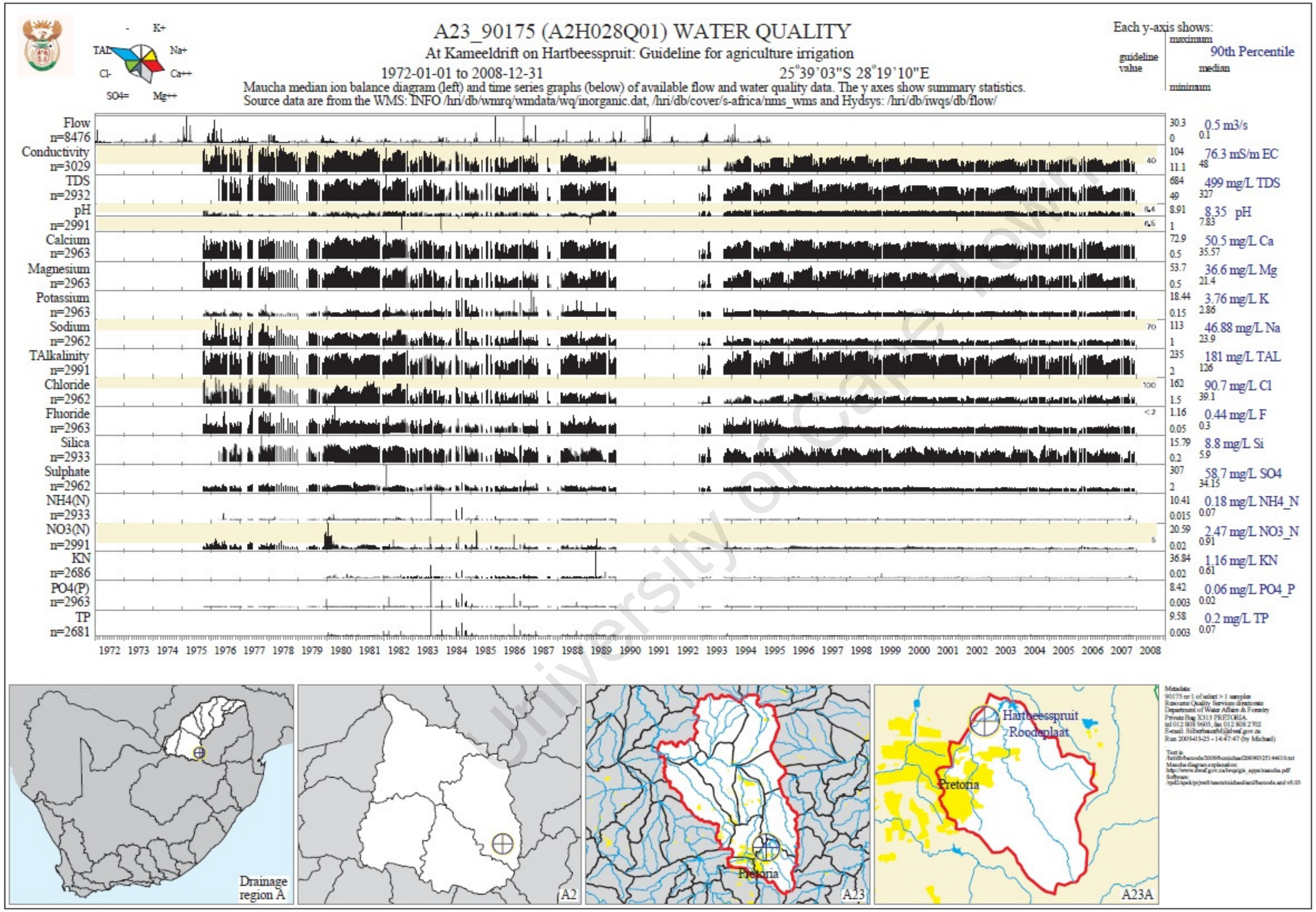
8. Barcode graphs with guidelines



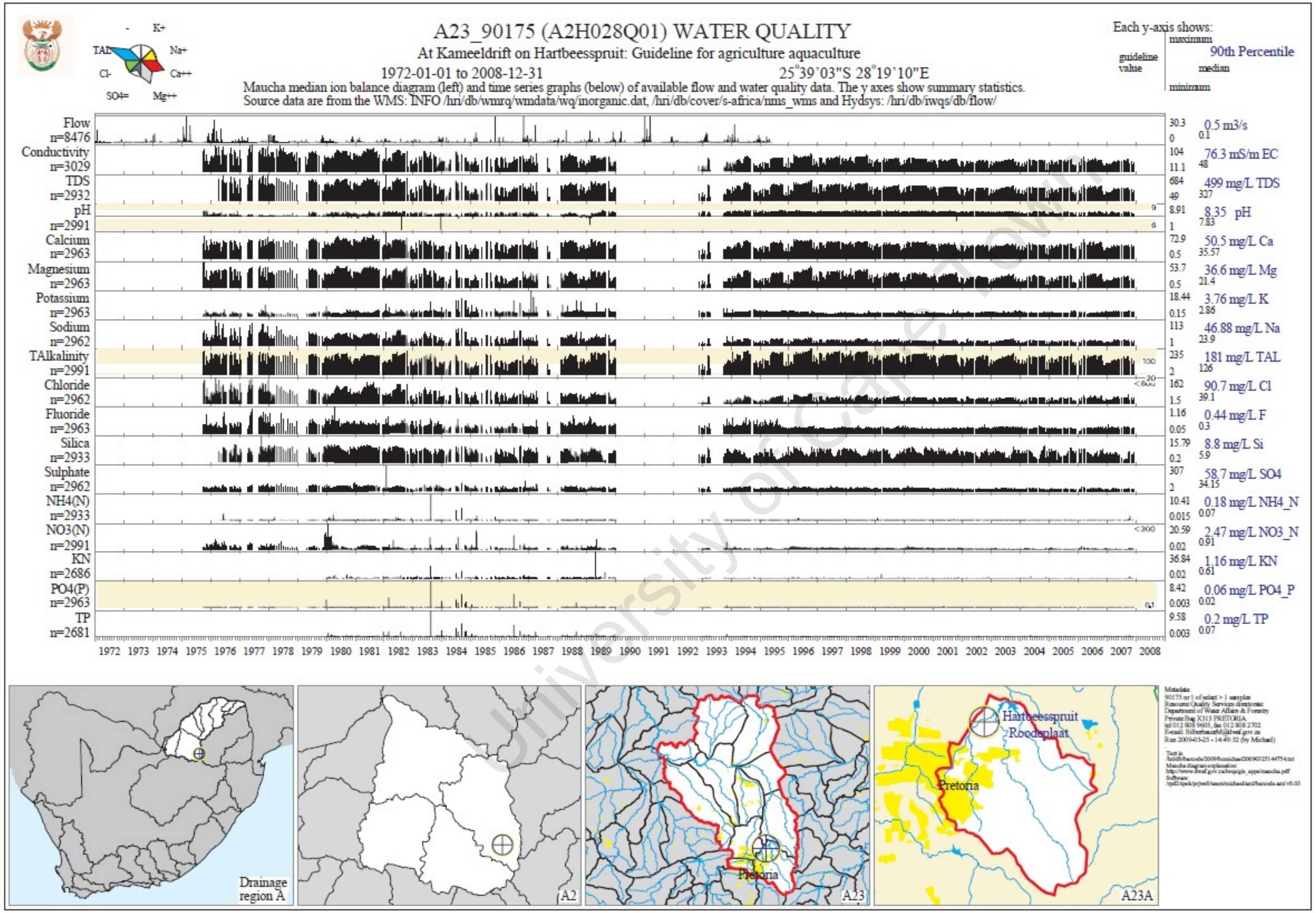
a) Domestic: human consumption



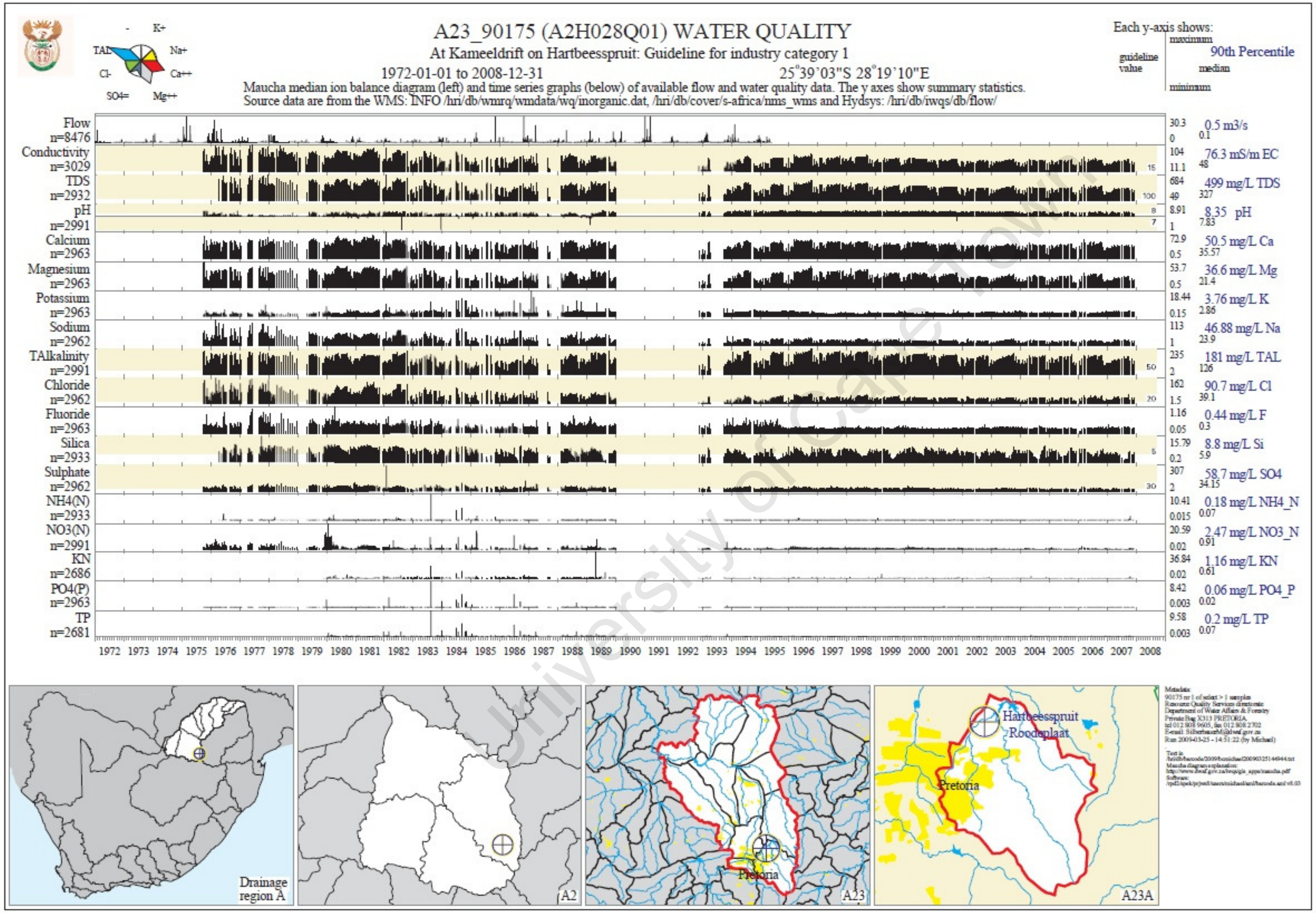
b) Recreation: full contact



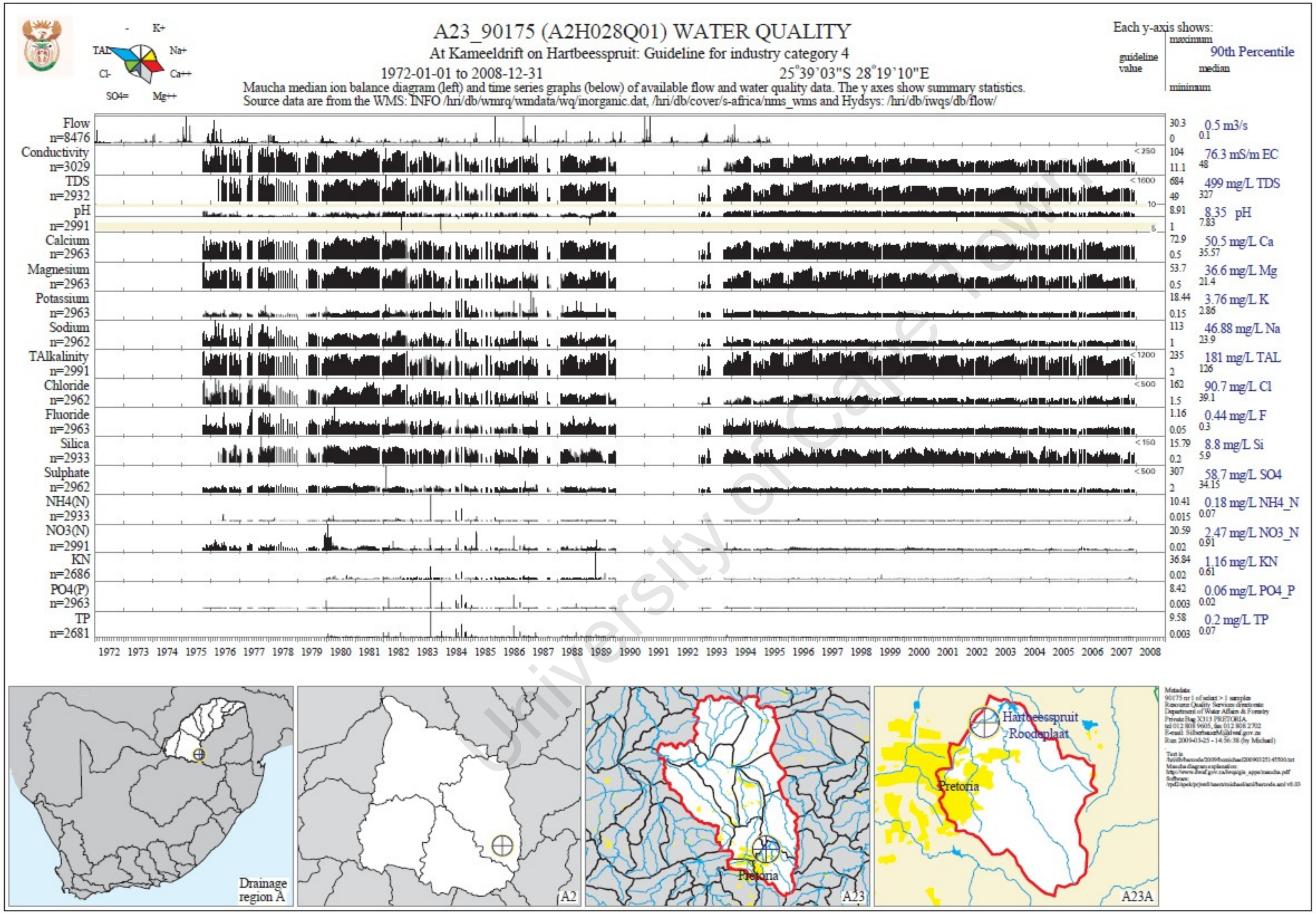
c) Agriculture: irrigation



d) Agriculture: aquaculture



e) Industry: category 1



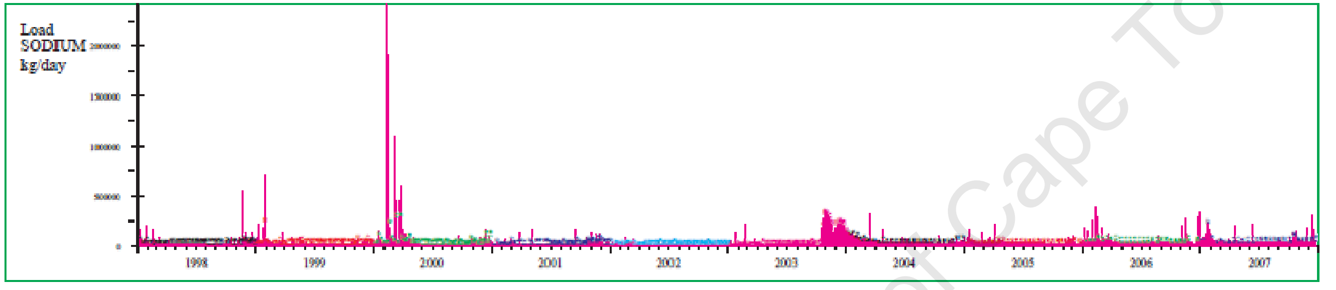
f) Industry: category 4

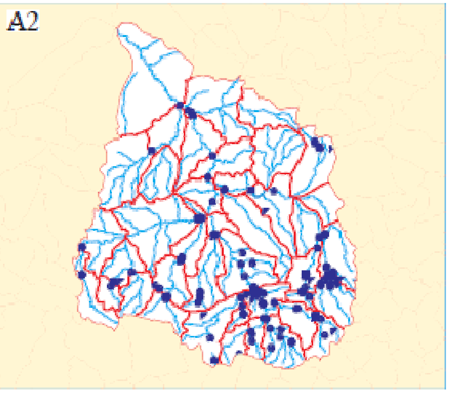
10. Load summary page

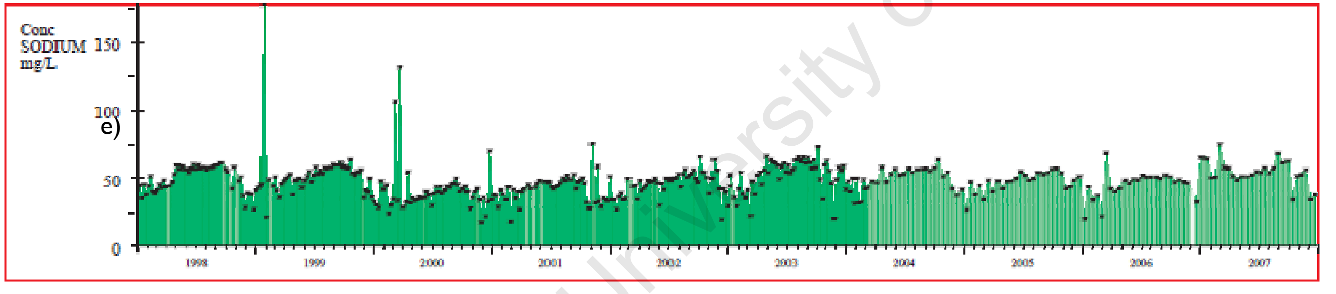
A2H012Q01: SODIUM ... 1998-01-01 to 2007-12-31

Flow data: /lel/db/tmp/db/flow/a2h012a01
Chem data: /lel/db/wms/wmdata/wq/MACRO090815.dat
3636 loads calculated: Log flow range = 5.17 to 7.78 .
Method 1: Sum(q*c)/n = 43131 kg SODIUM/ day
Method 2: (q_avg*c_sum)/n = 388017 kg SODIUM/ day
Method 3: (q_avg*Sum(q*c))/n = 36357 kg SODIUM/ day
Regression: logc=(-0.167*logq)+(2.695) r^2=0.16
Method 4: (Sum(q*c)/c)*(n*q_avg/Sum(q)/(^slope-1)) = 37694 kg SODIUM/ day

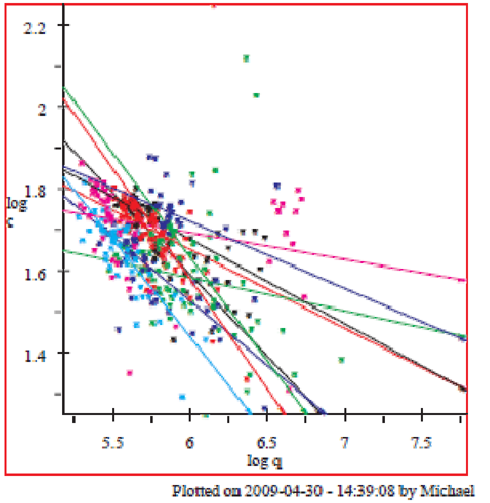
Year	n Obs	Mean m3/d	Method1 kg	Method4 kg
1998	48	744644	10442604	12921352
1999	51	598458	11154320	10762464
2000	52	1508448	17643686	23335876
2001	50	611980	8420092	10962473
2002	51	410630	5998575	7869732
2003	52	376702	21310818	7302867
2004	51	725772	10587639	12617388
2005	26	741423	11569301	12876540
2006	25	1177951	15103717	18970998
2007	26	890412	10011245	15134885

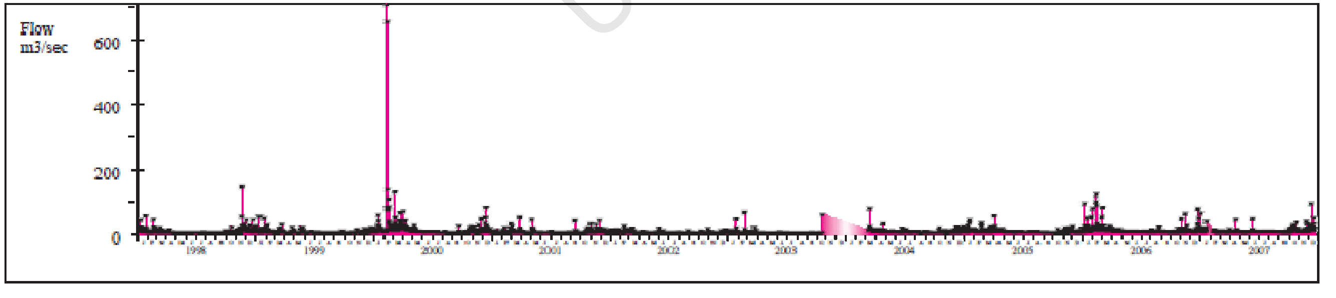






e)





Plotted on 2009-04-30 - 14:39:08 by Michael

11. Data disk

DVD containing digital data.

Please open file index.html to start.

Figure A11-1. Home page of files on enclosed DVD.

Figure A11-2. Page with links to images on enclosed disk, including the Chemclass water quality maps and examples of Barcode maps with guideline values.

Figure A11-3. Page on enclosed disk with links to Google Earth information including demo videos and KML files.

Figure A11-4. Page on enclosed disk with links to animations of water quality.

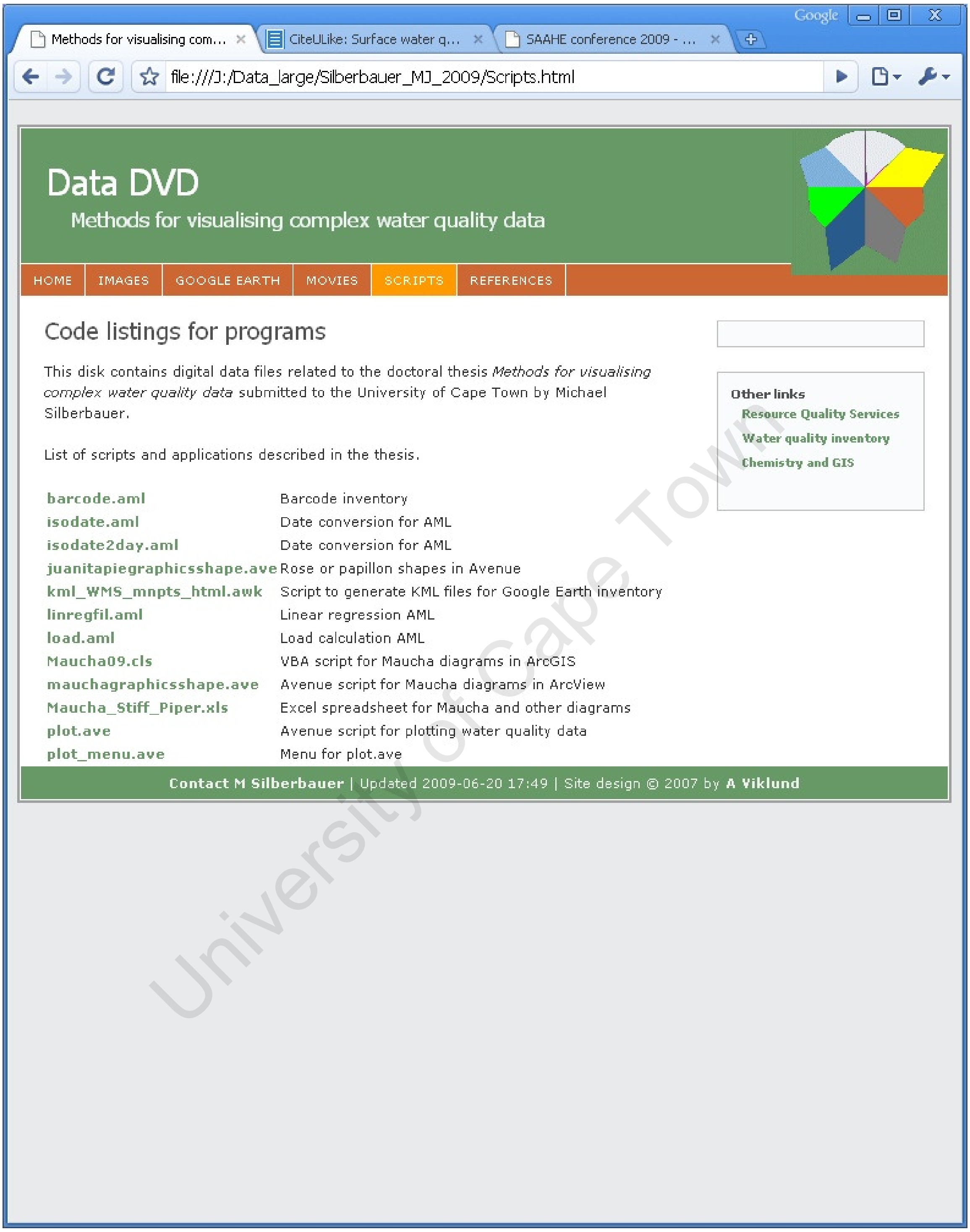


Figure A11-5. Page on enclosed disk with links to source code of scripts described in the text.

Figure A11-6. Page on enclosed disk with links to copies of some references cited, depending on copyright restrictions.

12. Correspondence from the Swiss Federal Department of the Environment, Transport, Energy and Communication: Coordinated monitoring of Swiss surface waters

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Département fédéral de l'environnement, des transports, de l'énergie et de la communication (DTEC)
Office fédéral de l'environnement (OFEV)

CH-3003 Berne, OFEV, Stb, EY

Aux services cantonaux de la pêche

Expédié par courrier et par courriel

Référence / n° de dossier: G014-0408
V/réf.:
N/réf.: ZS
Traité par: ZS
Berne, le 30 novembre 2009

Observation coordonnée des eaux de surface en Suisse

Mesdames, Messieurs,

Les eaux superficielles de Suisse faisant l'objet d'analyses régulières, nous disposons d'une foule de données sur l'état chimique de l'eau, même si celles-ci sont parfois peu homogènes. Par contre, les données concernant la biologie sont rares et il est de ce fait impossible d'avoir un bon aperçu de la situation nationale.

Au vu de ce qui précède, la Confédération et les cantons ont décidé de mettre conjointement sur pied un réseau d'observation coordonnée afin de récolter des données comparables à travers tout le pays. C'est à cet effet qu'a été lancé le projet **Observation coordonnée des eaux de surface en Suisse.** Outre la **qualité de l'eau**, l'observation englobe aussi l'**écomorphologie** et la **biologie**. Les poissons constituent des indicateurs particulièrement fiables de ce dernier domaine, car leur état résulte de divers effets environnementaux cumulés.

Telle qu'elle est prévue, l'observation coordonnée des cours d'eau se subdivise en deux champs d'étude:

- L'**observation de base ou permanente (*TREND)*** est un suivi sur le long terme de la même palette de paramètres dans toutes les stations de mesure. Un réseau formé de 112 stations de mesure a été convenu avec les cantons et sera bientôt définitivement arrêté (cf. cartes ci-joint).
- L'**observation spécifique *(SPE)*** aborde des problématiques spécifiques, comme l'évolution de l'état de petits cours d'eau fortement affectés par des activités agricoles (apports de nutriments, de pesticides, etc.)

Outre le relevé des paramètres chimico-physiques classiques, le programme *TREND* comprend un **programme d'analyses biologiques**, qui vise à recenser les macro-invertébrés, les diatomées, les végétaux aquatiques, l'aspect général du cours d'eau et, plus particulièrement, les poissons.

Voici ce qui est prévu pour la **partie poissons**:

- Les relevés du **suivi des poissons *TREND*** et leur interprétation se baseront sur la **méthode Poissons (Niveau R)**[1] du Système modulaire gradué.

[1] http://www.bafu.admin.ch/publikationen/publikation/00592/index.html?lang=fr&show_kat=/publikationen/00011

30.11.2009 12:37h 1/2

- Les stations de mesure pour la biologie doivent être aussi proches que possible des 112 stations dédiées à l'observation des paramètres chimiques (conditions identiques). On y procèdera à des relevés **seulement si le cours d'eau est de faible profondeur.**
- La fréquence et la période des relevés suivent les instructions du Système modulaire gradué. Toutefois, pour les paramètres biologiques (poissons, macro-invertébrés, diatomées et végétaux aquatiques), ils ne se feront qu'une fois par an (pour les poissons, en automne) et à intervalles **quadriennaux**, ceci afin d'optimiser le rapport coût/utilité. Les **premiers relevés sont prévus en 2012.**
- Les **pêches électriques** seront assurées par des **bureaux spécialisés** dans le domaine de la pêche ou par le service cantonal compétent, avec lequel elles doivent impérativement être coordonnées au préalable. De même, les **déterminations** et l'**interprétation** des résultats seront confiés à des experts ou des bureaux spécialisés.
- Outre les prélèvements quadriennaux, il sera procédé à un recensement **annuel** des peuplements de poissons dans une vingtaine de sites de référence (**observation des poissons *SPE***). Ces sites restent à déterminer; aucune proposition n'a encore été discutée avec les cantons.

Financement: La participation financière de la Confédération à l'observation coordonnée des eaux de surface sera plus particulièrement destinée au domaine de la biologie, car c'est là que les écarts entre les investigations prévues et les études que les cantons entreprennent en application de la LEaux sont les plus importants.
Le travail des experts chargés du relevé et du dépouillement des données sera financé par la Confédération. L'attribution centralisée des mandats externes devrait permettre de limiter autant que possible les différences méthodologiques. L'aide des cantons est souhaitée pour les prélèvements (pour les pêches électriques, p. ex.).

Vous recevrez désormais des **emails réguliers** qui vous informeront sur l'avancement de ce projet et des éléments pour lesquels est requise votre participation active, que ce soit pour leur planification ou au niveau opérationnel (pêches électriques dans votre canton, p. ex.). Pour ce qui est du suivi des autres paramètres, nous restons en contact étroit avec les services de la protection des eaux. Nous vous fournissons par ailleurs volontiers des informations complémentaires. N'hésitez donc pas à contacter l'une des personnes ci-dessous, si vous avez des questions, remarques ou suggestions:

- Werner Göggel, section Eaux de surface – morphologie et débits, responsable du projet Observation des eaux de surface; 031 325 41 35; werner.goeggel@bafu.admin.ch
- Daniel Hefti, section Pêche et faune aquatique; 031 322 92 42, daniel.hefti@bafu.admin.ch
- Sabine Zeller, section Eaux de surface – morphologie et débits, Biologie (TREND); 031 325 09 19; sabine.zeller@bafu.admin.ch

Nous nous réjouissons de la collaboration à venir et vous prions d'agréer, Mesdames, Messieurs, nos salutations les meilleures.

Rémy Estoppey
Chef de la section Eaux de surface – morphologie et débits
Division Eaux, OFEV

Erich Staub
Chef de la section Pêche et faune aquatique
Division Gestion des espèces, OFEV

Copie par courriel:
- Michael Eugster (SG), Pius Niederhauser (ZH), Marin Huser (BL), Marc Bernard (VS), Joachim Huerlimann (AquaPlus), Franziska Stoessel (AquaPlus)
- A l'OFEV: Erich Staub, Daniel Hefti, Rémy Estoppey, Werner Göggel, Sabine Zeller, Christian Leu, Paul Liechti, Monika Schaffner, Ulrich Sieber, Adrian Jakob, Markus Wüest

Annexe:
- Stations de mesure retenues pour l'Observation coordonnée des eaux de surface TREND

30.11.2009 12:34h 2/2

ESRI Mapping Center – Ask a Cartographer – Map symbol brewer

http://mappingcenter.esri.com/index.cfm?fa=ask.answers&q=989

Ask A Cartographer

Map symbol brewer

October 22 2009 | 1 comment
Categories: Cartographic Design

Does ESRI have any developments afoot to implement a more comprehensive set of graphical symbols than the currently limited spot, pie and bar? I am thinking on the lines of the work of Dr. Olaf Schnabel, who developed the Map Symbol Brewer (http://www.oschnabel.gmxhome.de/portfolio_en.html). This kind of development would make ArcGIS a useful tool for summarising multivariate point data on a map, and would save users a great deal of time in developing their own shaky scripts to draw things like Maucha, spider, polar and radar diagrams.

The proxy server did not allow an upload to the ESRI site, but an example of what I am asking is at this link: http://www.dwaf.gov.za/iwqs/water_quality/NCMP/natl_assmnt04_maucha.pdf.

The R statistics program has many of these types of symbols, and has a simple mapping procedure for shape files (and it's free). However, the ideal would be to use the symbols in a cartographic environment.

Mapping Center Answer:

There's no effort like this afoot. Once you start digging, there is no such thing as a comprehensive list of graphical symbols--there is no practical limit to what could be symbolized.

The problem of multivariate thematic symbology for maps is challenging, not just to make the map, but to read it as well. Back when all we had was paper, relatively few of these maps got made, but those that did were usually pretty decently designed and executed (meaning expensive). In the digital mapping age, if the will exists to make such a map, it can be done, but because these are all specialty maps, they are on the fringe of most people's expertise, so finding a cartographer is a challenge, and it's, like you indicate, a lot of work.

We are not limited to paper any longer, but because such maps are not mainstream, you're just not likely to see a generic or mainstream tool for making them. But web maps are offering some solutions, particularly by leverating Web 2.0 graphics engines against GIS services. For instance:

http://downloads2.esri.com/mappingcenter2007/maps/worldtopomap/ESRI_demographics_Ex.h

Where instead of using a complex symbol for every polygon, symbology is based on a characteristic that explains something commonly understood, and then clicking on that feature gives better context. I'm not saying this replaces the need for rich multivariate symbology, but it offers an alternative that will work well in many cases.

User Comments | Submit Your Comments

Star and rose map symbols posted by Michael Silberbauer on Nov 3 2009 5:25AM

The ESRI demographics example is very elegant mentioned above, and neatly solves the problem of allowing the interested user to view the detailed information available for a region of interest. However, this differs from the problem that faces people working with water quality data, which is to show multiple variables for many sites scattered across a map. This need not be confusing, so long as one uses common-sense visualisation to avoid confusion, as suggested by authors such as Ware, MacEachren or Slocum. The background map can display other information without distracting attention, for example rivers as lines and land use as polygons.

An example is the use of KML to display monitoring points symbolised with ionic composition Maucha diagrams on a 3D-earth backgound. The user can click on an interesting site to bring up a balloon with metadata, and follow other links in the balloon to see the time series data (see http://www.dwaf.gov.za/iwqs/wms/data/000key.asp). The ability to 'splay' overlapping points when they are selected is essential.

However, to return to the original question about development plans, which was prompted by Olaf Schnabel's contention that, despite the phenomenal developments in spatial analysis software during the past couple of decades, the encoding of multivariate statistics on maps has actually decreased. His thesis suggests a way of overcoming the substantial programming difficulties by standardising the process of generating symbols using a diagram markup language as demonstrated in his Map Symbol Brewer.

My own experience with developing multivariate symbols during the past 20 years supports the view that we have somehow hit a ceiling:

Pascal - generate the symbols, print them and glue them on a paper map
Arc Info - generate the symbols on the actual map with SHADE
ArcView - generate symbols as groups that can be shifted when they overlap
ArcGIS - generate symbols as groups that crash ArcGIS when you bring a mouse anywhere near them

GIS systems surely need to have more advanced symbolisation built in, so that users can concentrate on presenting their data rather than trying to encode concepts in ever more obscure and user-hostile computer languages. Any GIS package that offers more multivariate symbol options than the run-of-the-mill pub lunch (bar and pie) will have an immediate competitive advantage.

Here are a few references on the subject:

Schnabel O. Benutzerdefinierte Diagramm-signaturen in Karten: Konzepte, Formalisierung und Implementation. Eidgenössischen Technischen Hochschule Zürich; 2007. Available from:http://e-collection.ethbib.ethz.ch/eserv/eth:29352/eth-29352-01.pdf. A paper in English on the subject is: Schnabel O, Hurni L. Primitive-based Construction Theory for Diagrams in Thematic Maps. Cartographic Journal. 2009;46(2):136-145. Available from:http://dx.doi.org/http://dx.doi.org/10.1179/000870409X459851.

Tufte E. (assorted)

Ware C. Visual Thinking: for Design (Morgan Kaufmann Series in Interactive Technologies)

Slocum T A, Mcmaster R B, Kessler F C, Howard H H. Thematic Cartography and Geovisualization.