Technical note

An inorganic water chemistry dataset (1972–2011) of rivers, dams and lakes in South Africa

Jan Marten Huizenga[1]*, Michael Silberbauer[2], Rainier Dennis[3] and Ingrid Dennis[3]

[1] Department of Geology, Unit for Environmental Sciences and Management, North-West University, Private Bag X6001, Potchefstroom, 2520, South Africa
[2] Resource Water Quality Services, Department of Water Affairs, South Africa
[3] Centre for Water Science and Management, Unit for Environmental Sciences and Management, North-West University, Private Bag X6001, Potchefstroom, 2520, South Africa

ABSTRACT

A national dataset of inorganic chemical data of surface waters (rivers, lakes, and dams) in South Africa is presented and made freely available. The dataset comprises more than 500 000 complete water analyses from 1972 up to 2011, collected from more than 2 000 sample monitoring stations in South Africa. The dataset includes the major ion chemical composition and numerous calculated variables that can, amongst others, be used to determine accuracy of the analysis. The methods described here have potential for improving quality control measures in water chemistry laboratories by detecting anomalous samples. The processed data are available in Excel spreadsheets and can be downloaded from the website of the Centre for Water Science and Management based at the North-West University (www.waterscience.co.za/waterchemistry/data.html).

Keywords: national data set, inorganic water chemistry, South Africa, surface waters

INTRODUCTION

The Department of Water Affairs in South Africa has had an extensive water monitoring programme in place since the early 1970s, which includes more than 2 000 monitoring sites in lakes, dams and rivers covering the entire country (Fig. 1). This monitoring programme has resulted in the availability of hundreds of thousands of chemical analyses for the major ions.

Up to now, selected data can only be obtained from the Department of Water Affairs upon request. Further, the dataset includes numerous incomplete analyses (see section 'Modifications of the dataset' for details) and also does not include any variables that can be used to test the accuracy of the analysis. In this paper, we describe how we have modified and reorganised these data addressing the abovementioned issues. The data are freely available in the form of user-friendly Excel spread sheets. We anticipate that the easy access of this dataset will be beneficial towards hydrogeochemical and environmental research in South Africa and will further enhance the South African National Chemical Monitoring Programme (e.g., Van Niekerk et al., 2009).
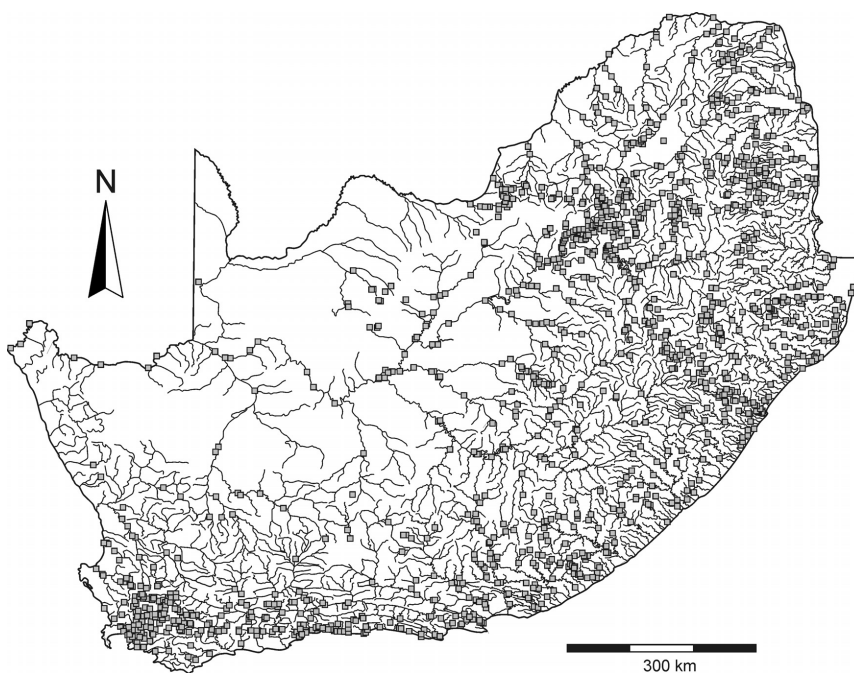

Figure 1
Map of South Africa showing the rivers and Department of Water Affairs sample monitoring stations (grey squares)

Water chemistry data of South African surface waters

Inorganic water chemistry data up to 1998 were obtained from the CSIR (Environmentek): the 'Water Quality on Disc, version 1.0'. Data from 1999 up to 2011 were obtained from the

TABLE 1 Summary of dataset		
Variable(s)	Comments	Relevant reference
Concentrations of major ions in mg/ℓ	Data obtained from the Department of Water Affairs	N/A
pH	Data obtained from the Department of Water Affairs	N/A
Concentration of major ions in mmol/ℓ	Calculated from concentrations of major ions in mg/ℓ	N/A
Measured electrical conductivity ($EC_{meas.}$) in mS/m	Data obtained from the Department of Water Affairs Recalculated EC values in mS/cm are also given (1 mS/m = 10 mS/cm).	N/A
Concentration of $(HCO_3)^-$ and $(CO_3)^{2-}$ in mmol/ℓ	Calculated from pH and total alkalinity (TAL) using Eqs. (1) and (2).	e.g., Appelo and Postma (2005)
Stoichiometric charge balance (SCB)	Calculated from major ion concentrations in meq/ℓ using Eq. (3). Accurate analysis: −5% < SCB < +5%.	e.g., Zhu and Andersen (2002); Appelo and Postma (2005)
Ionic strength	Calculated from major ion molar concentrations using Eq. (4).	e.g., Appelo and Postma (2005)
Calculated electrical conductivity (EC_{calc})	Calculated using Eqs. (5) and (6)	McCleskey et al. (2012)
Conductivity imbalance (δ_{EC})	Calculated using Eq. (7) for a temperature of 20°C. Accurate analysis: −15% < δ_{EC} < +15% for South African surface waters	McCleskey et al. (2012)
Calculated total dissolved solids (TDS_{calc})	Sum of all ions in mg/ℓ. Total alkalinity is used for the carbonate species. Si is included as SiO_2.	N/A
Total dissolved solids imbalance (δ_{TDS})	Calculated using Eq. (8). Accurate analysis: −15% < δ_{EC} < 0% for South African surface waters.	N/A
Water quality parameters for South African surface waters	Calculated using Eqs. (9), (10) and (11)	Huizenga (2011)
Sodium adsorption (SAR) and adjusted SAR	SAR is calculated using Eq. (12). The adjusted SAR is calculated following the method described in Lesch and Suarez (2009).	Lesch and Suarez (2009)

Department of Water Affairs as CSV files. Chemical variables that were determined include the following: pH, electrical conductivity (EC, mS/m), total alkalinity (measured as $CaCO_3$ in mg/ℓ), the total dissolved solids (TDS, in mg/ℓ), and the concentrations of the following ions (all in mg/ℓ): sodium (Na^+), potassium (K^+), calcium (Ca^{2+}), magnesium (Mg^{2+}), ammonium (($NH_4)^+$), silica (Si), fluoride (F), orthophosphate (($PO_4)^{3-}$), chloride (Cl^-), sulphate (($SO_4)^{2-}$), nitrate and nitrite combined (($NO_3)^- + (NO_2)^-$). Each individual analysis is characterised by a sample station identification code and/or point identification number, and a sampling date. A separate Excel file is available that comprises a brief description of the sample locality including GPS coordinates, the specific sample taken (e.g., spring, river, treatment works, etc.), total number of samples taken at the locality, and the first and last date of sample collection.

MODIFICATIONS OF THE DATASET

The original dataset has been extensively modified, which includes the removal of incomplete analyses and the addition of extra chemical variables calculated from concentrations of the major ions (Table 1).

Removal of incomplete analyses

A chemical analysis is considered to be incomplete (though not necessarily incorrect) if one of the following ions was not analysed for: Na^+, Ca^{2+}, Mg^{2+}, K^+, Cl^-, $(SO_4)^{2-}$, and $(HCO_3)^- + (CO_3)^{2-}$ (measured as the total alkalinity). These chemical species are critically important for the calculation of variables that can be used to test the analysis for its accuracy (e.g., charge balance, ionic strength, calculated EC and calculated TDS). Other ions including $(NH_4)^+$, Si, F^-, $(PO_4)^{3-}$, and $((NO_3)^- + (NO_2)^-)$ generally occur in such small molar concentrations that they are not essential for the calculation of the abovementioned variables. Further, chemical analyses without a pH were also considered to be incomplete as the pH is essential to calculate the concentrations of $(HCO_3)^-$ and $(CO_3)^{2-}$ from the total alkalinity. Incomplete analyses were removed from the dataset.

After removal of the incomplete analyses, the total number of analyses is 509 919. The relative distribution of number of complete water analyses over the different primary catchment areas is shown in Fig. 2, illustrating that the large majority of the analyses are available from the north-western (Limpopo, Olifants and Vaal catchments) and south-eastern parts of South Africa (Berg and Breede catchments). Most water analyses are available from the late 1970s to the late 1990s; from 2000 onwards there is a significant decline in the amount of water analyses available (Fig. 3).

Addition of chemical variables

We included several chemical and water quality variables (Table 1) that were calculated from the major ions to make the dataset as complete and attractive as possible for any user. The most important additions are listed and described below.

Concentrations of (bi)carbonate

The concentrations of $(HCO_3)^-$ and $(CO_3)^{2-}$ are essential in order to obtain a complete analysis of all the anions present in

Figure 2
The primary catchment areas of South Africa, showing the frequency of water chemistry analyses in each catchment as a percentage of the total number of complete analyses in the database as described in the text.
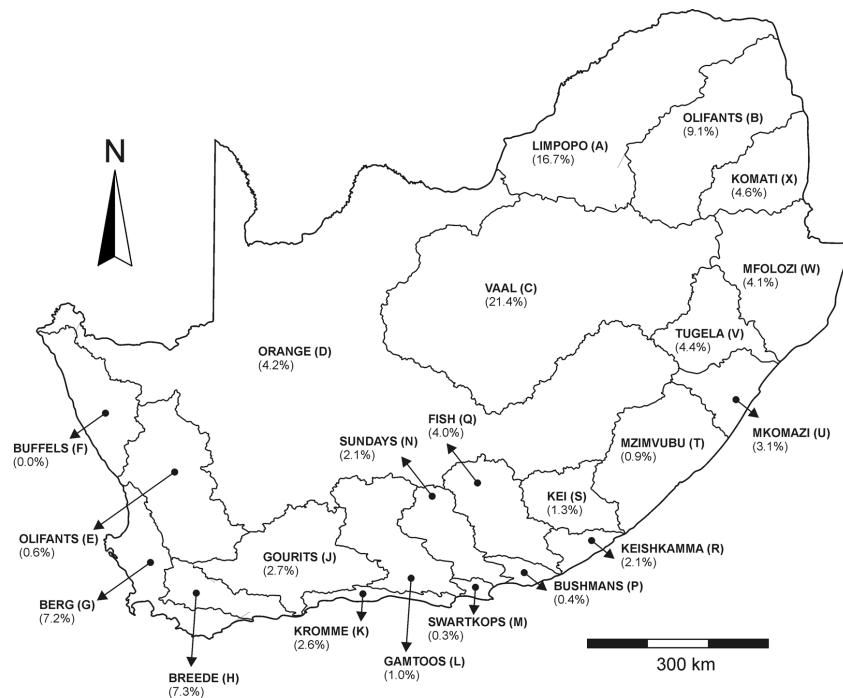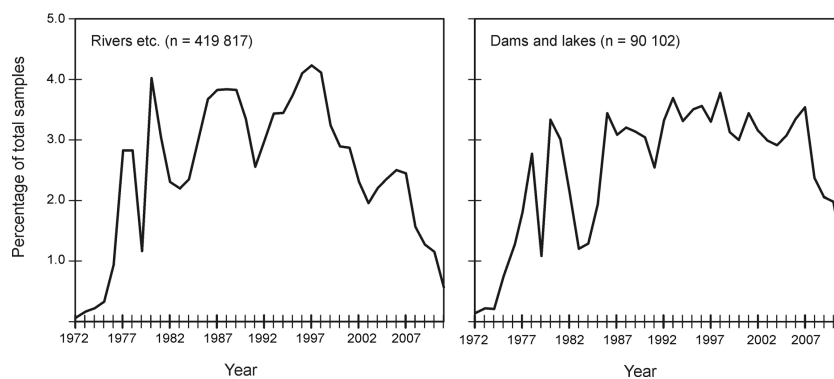

Figure 3
Relative amount of complete water analyses per year (1972–2011) for rivers (left) and dams/lakes (right)

the solution. These concentrations are required to calculate the electrical conductivity and the ionic strength of the solution, as will be described later. The molar concentrations of $(HCO_3)^-$ and $(CO_3)^{2-}$ are calculated from the molar total alkalinity (TAL_{molar}; TAL_{molar} = TAL (as $CaCO_3$ in mg/ℓ) / molar mass $CaCO_3$) and the pH (e.g., Appelo and Postma, 2005):

$$[(HCO_3)^-] = 2\,(TAL_{molar} - 10^{pH-pKw})/(1 + 2 \cdot 10^{pH+Ka}) \qquad (1)$$

$$[(CO_3)^{2-}] = (TAL_{molar} - [(HCO_3)^-])/2 \qquad (2)$$

where:
 Kw is the dissociation constant for water (10^{-14})
 Ka the equilibrium constant ($10^{-10.3}$) for the reaction
 $(HCO_3)^- + H_2O \Leftrightarrow (CO_3)^{2-} + H_3O^+$.
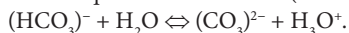
Stoichiometric charge balance

The stoichiometric charge balance (SCB) is widely used to determine whether a chemical analysis of a water sample is accurate. The SCB can be defined as follows (e.g., Appelo and Postma, 2005):

$$SCB\,(\%) = 100 \times (\textstyle\sum [\text{cations}] + |\sum [\text{anions}]\,|)/(\sum [\text{cations}] - |\sum [\text{anions}]\,|) \qquad (3)$$

where:
 $\sum$ [cations] and $\sum$ [anions] denote the sum of the charge-corrected cations and anions (meq/ℓ).

The chemical analysis is considered to be accurate if $-5\% <$ SCB $< +5\%$ (Zhu and Andersen, 2002; Appelo and Postma, 2005). Inaccurate analyses are either due to analytical error or some ions that are presumed to be minor are not included in the chemical analysis (e.g., Appelo and Postma, 2005). For example, it is likely that surface waters affected by acid mine drainage (AMD) will show a negative value for the charge balance because significant amounts of e.g., Fe^{3+}, Al^{3+}, Mn^{2+}, etc., are present in AMD-affected surface waters (e.g., Lee et al., 2002). These elements are not included in the standard analyses by the Department of Water Affairs.

Ionic strength

The ionic strength (I) is a measure of the effect of the charge of the dissolved ions and is calculated as follows:

$$I = \tfrac{1}{2}\textstyle\sum m_i (z_i)^2 \qquad (4)$$

where:
 m_i and z_i denote the concentration (mol/ℓ) and charge of ion i, respectively (e.g., Hem, 1985).

The ionic strength is used in the calculation of, e.g., the electrical conductivity (see next section) and the activity of the ions (e.g., Appelo and Postma, 2005).

Calculated electrical conductivity

The electrical conductivity is calculated (EC_{calc}) using the method described by McCleskey et al. (2012). It is calculated from the chemical composition using the relation:

$$EC_{calc} = \sum \lambda_i \, m_i \qquad (5)$$

where:
   λ_i denotes the ionic molal conductivity for ion i.

The ionic molal conductivity can be calculated from the temperature and ionic strength (I) (McCleskey et al., 2012):

$$\lambda_i = \lambda° - A_T \sqrt{I} / (1 + B \sqrt{I}) \qquad (6)$$

where:
   A_T is a temperature-dependent variable
   B a constant, both of which can be found in Table 1 in McCleskey et al. (2012).

The calculated EC can be compared to the measured EC (EC_{meas}) (McCleskey et al., 2012):

$$\delta_{EC} \,(\%) = 100 \times (EC_{calc} - EC_{meas}) / EC_{meas} \qquad (7)$$

The following should be noted for EC_{calc} in the South African dataset: (i) we used a temperature of 20°C to calculate A and $\lambda°$ (a 1°C temperature difference results in a change of EC_{calc} of ~2% if EC_{calc} is ~500 μS/cm), and (ii) ion pairs were not included in the conductivity calculation as the chemical speciation is not part of the dataset.
   Evaluation of δ_{EC} of close to 1 800 natural water samples by McCleskey et al. (2012) shows that δ_{EC} falls within the range of ±10%. In addition to the SCB, δ_{EC} can be used as an accuracy test for water analyses of surface waters. Keeping in mind the limitations (i.e., exclusion of ion pairs and the fixed temperature of 20°C) and the range of natural samples as reported by McCleskey et al. (2012), we suggest that a δ_{EC} value between ~+15 and ~−15% is acceptable for routine laboratory analysis of natural surface waters in South Africa.

Calculated total dissolved ions

The calculated total dissolved solids (TDS_{calc}) is the sum of all anions and cations in mg/ℓ. The Si concentration (mg/ℓ) is for this purpose recalculated to the SiO_2 concentration according to SiO_2 (mg/ℓ) = 2.14 Si (mg/ℓ). In the same way as for EC, the TDS_{calc} can be compared to the measured TDS (TDS_{meas}):

$$\delta_{TDS} \,(\%) = 100 \times (TDS_{calc} - TDS_{meas}) / TDS_{meas} \qquad (8)$$

where:
   TDS_{calc} can never exceed TDS_{meas} (i.e., $\delta_{TDS} < 0$). Acceptable values of δ_{TDS} for water samples to be accurate are between 0 and −15%.

SA water characterisation parameters

Based on a statistical evaluation, the surface waters in South Africa can be characterised by 3 factors, namely, chemical weathering (reflected in the total alkalinity), chloride salinisation, and sulphate contamination (Huizenga, 2011). These three factors are calculated from the molar concentrations of the relevant chemical species as follows (modified after Huizenga, 2011):

Chemical weathering (%) = $100 \times ([[HCO_3^-] + 2\,[(CO_3^{2-})]]$
$/([[HCO_3^-] + 2\,[(CO_3^{2-})] + [Cl^-] + 2[(SO_4^{2-})] + 3[(PO_4^{3-})])$ (9)

SO_4 contamination (%) = $100 \times 2[(SO_4^{2-})] / ([[HCO_3^-]$
$+ 2\,[(CO_3^{2-})] + [Cl^-] + 2[(SO_4^{2-})] + 3[(PO_4^{3-})])$ (10)

Cl salinisation (%) = $100 \times [Cl^-] / ([[HCO_3^-] + 2\,[(CO_3^{2-})]$
$+ [Cl^-] + 2[(SO_4^{2-})] + 3[(PO_4^{3-})])$ (11)

Other characterisations are possible based on different combinations of ions (e.g., Day and King, 1995).

Sodium adsorption ratio (SAR)

The sodium adsorption ratio (SAR) is an index of the suitability of water for irrigation (e.g., Lesch and Suarez, 2009). We have included the SAR in the datasets as irrigation for agricultural purposes is one of the prime uses of natural surface waters in South Africa. The SAR is calculated from the molar concentrations of Na, Ca, and Mg:

$$SAR = [Na] / \sqrt{([Ca] + [Mg])} \qquad (12)$$

The SAR may need to be adjusted if the water has relatively high Ca and bicarbonate concentrations (e.g., Lesch and Suarez, 2009), which is the case in certain areas in South Africa, depending on the geology. In that case the Ca concentration in Eq. (10) must be adjusted for the precipitation of $CaCO_3$. We have adopted the method described by Lesch and Suarez (2009) to calculate the adjusted SAR. This process involves numerous calculation steps, which are not repeated here. A detailed description of the calculation method can be found in the paper by Lesch and Suarez (2009).

Dataset information and availability

The files that are available are shown in Table 2. All Excel files are 'values only' files, i.e. the formulae used to calculate variables listed in Table 1 are not included in this file. A calculation template is, however, available that includes all formulae used. The worksheets in all files are protected with the password 'quality' (in lower case) in order to avoid accidental modifications. Data that are not available or variables that could not be calculated due to missing data are indicated with the cell value '-9999'. It must be noted that during the transfer from the original DWA database, results that were flagged 'below detection limit' may appear as zero mg/ℓ.
   All datasets are available free of charge from the Centre for Water Science and Management (North-West University) website: www.waterscience.co.za/waterchemistry/data.html. More information on the analytical methods and detection limits for the different chemical species can be obtained from the Department of Water Affairs website: www.dwa.gov.za/iwqs/report.aspx. This website also shows the e-mail addresses of the people who should be contacted in order to obtain more recent water quality data. Finally, although care was taken in the water analyses and compiling the datasets, it must be noted that neither the Department of Water Affairs nor the authors

Name of file	Format	File size	Short description
TABLE 2			
Overview of available files			
Rivers up to 1998 (A–D)	Excel 2007 (xlsx)	~90 MB	File includes water chemistry data for rivers for the primary catchment areas A to D up to 1998.
Rivers up to 1998 (E–X)	Excel 2007 (xlsx)	~90 MB	File includes water chemistry data for rivers for the primary catchment areas E to X up to 1998.
Rivers 1999–2011	Excel 2007 (xlsx)	~70 MB	File includes water chemistry data for rivers for all primary catchment areas from 1999 up to 2011.
Dams and lakes up to 1998	Excel 2007 (xlsx)	~35 MB	File includes water chemistry data for dams and lakes up to 1998.
Dams and lakes 1999–2011	Excel 2007 (xlsx)	~20 MB	File includes water chemistry data for dams and lakes from 1999 up to 2011.
Sample stations	Excel 2007 (xlsx)	~378 KB	File includes a list of all sample stations with details such as the locality, total number of samples collected, first and last date of collection, at the time of data transfer.
Calculation template	Excel 2007 (xlsx)	~30 KB	Files with all formulae used for the calculated variables that are listed in Table 1.
Sample statistics	Excel 2007 (xlsx)	~15 KB	File includes subdivision of total number of samples per year and per primary catchment area. The numbers are indicated for both rivers and lakes/dams.

can be held responsible for any errors in the dataset provided and any interpretations based thereon.

CONCLUSION AND RECOMMENDATIONS

The extensive water chemistry database for South African surface waters is a valuable source of information about the status of water resources and changes during the past 4 decades. However, the dataset contains many inconsistencies in chemical analysis results, which the user needs to take into consideration before using the data for situation or trend analysis. We have included numerous methods to test this consistency in the available dataset.

We recommend that the Department of Water Affairs incorporate the tests proposed in this paper into their laboratory information management system. Flagging of inconsistencies while the original sample is still available for re-analysis would considerably improve the effectiveness of the national monitoring network at minimal extra cost. This approach could result in a more cost-effective workflow and eliminate wasted sampling effort.

ACKNOWLEDGEMENTS

REFERENCES

APPELO CAJ and POSTMA D (2005) *Geochemistry, Groundwater and Pollution* (2nd edn). AA Balkema, Rotterdam. 649 pp.

DAY JA and KING JM (1995) Geographical patterns, and their origins, in the dominance of major ions in South African Rivers. *S. Afr. J. Sci.* **91** 299–306.

HEM JD (1985) *Study and Interpretation of the Chemical Characteristics of Natural Water* (3rd edn.) US Geological Survey Water-Supply Paper 2252. 263 pp.

HUIZENGA JM (2011) Characterisation of the inorganic chemistry of surface waters in South Africa. *Water SA* **37** 401–410.

LESCH SM and SUAREZ DL (2009) A short note on calculating the adjusted SAR index. *T. ASABE* **52** 493–496.

LEE G, BIGHAM JM and FAURE G (2002) Removal of trace metals by coprecipitation with Fe, Al and Mn from natural waters contaminated with acid mine drainage in the Ducktown Mining District, Tennessee. *Appl. Geochem.* **17** 569–581

McCLESKEY RB, NORDSTROM DK, RYAN JN and BALL JW (2012) A new method of calculating electrical conductivity with applications to natural waters. *Geochim. Cosmochim. Acta* **77** 369–382.

VAN NIEKERK H, SILBERBAUER MJ and HOHLS BC (2009) Monitoring programme revision highlights long-term salinity changes in selected South African rivers and the value of comprehensive long-term datasets. *Environ. Monit. Assess.* **154** 401–411.

ZHU C and ANDERSEN G (2002) *Environmental Applications of Geochemical Modelling.* Cambridge University Press, London. 284 pp.

ERRATUM (6 October 2014)

Erratum for *Water SA* Vol 39 (2) 335–340, originally published in April 2013.

Eq. (3), on p. 337, should be:

SCB (%) = 100 × abs(Σ [cations] − Σ [anions]) ⁄ (Σ [cations] + Σ [anions])