



Christopher D Mahlathi, who is a senior engineer at the Council for Scientific and Industrial Research (CSIR), is currently pursuing a PhD in Civil Engineering (Water) at Stellenbosch University. With a keen focus on water and wastewater management, he holds qualifications in Chemical Engineering and Water Utilisation

Engineering, complemented by proficiency in numerical modelling of both natural and man-made systems. His recent research work included the development and piloting of a decision-support tool aimed at optimising wastewater treatment costs while steadfastly upholding water quality standards at the receiving rivers. He is interested in innovation and efficiency to engineer solutions that promise significant impact through technological advancement and ingenuity.

Contact details:

Council for Scientific and Industrial Research
Next Generation Institutions and Enterprises
PO Box 320
Stellenbosch 7599
South Africa
E: cdmahlathi@csir.co.za / cdmahlathi@gmail.com



Dr Isobel C Brink, who is a Senior Lecturer in the Department of Civil Engineering at Stellenbosch University, specialises in Water Treatment and Environmental Engineering. She is interested in new innovations in water quality improvement methods and undertakes projects with wide application in this regard. Her recent research includes

simple point-of-use systems for potable water treatment, river water quality modelling and the use of green infrastructure for surface runoff water quality improvement.

Contact details:

Department of Civil Engineering
Stellenbosch University
Private Bag X1
Matieland 7602
South Africa
E: icbrink@sun.ac.za



Dr Josefina Wilms holds a PhD in computational fluid dynamics from Stellenbosch University. She has an extensive professional background, having worked as a data scientist at both the CSIR (Council for Scientific and Industrial Research) in South Africa and Deltares in the Netherlands. Currently, she is employed at

the GeoForschungs Zentrum (GFZ) in Oberpfaffenhofen, Germany, where she is involved in performing simulations for Next Generation Gravity Missions (NGGM). Her expertise and experience in computational fluid dynamics and data science are valuable assets in her current role, where she contributes to advancements in gravity mission technology and research.

Contact details:

Deutsches GeoForschungs Zentrum
Claude-Dornier Street 1
Building 401
Wessling 82234
Germany
E: jowilms@gfz-potsdam.de

River water quality modelling in South Africa: Considerations, sourcing and accessing of input data

C D Mahlathi, I C Brink, J M Wilms

River water quality modelling relies on a good dataset for model input, calibration and validation to develop a reliable model. Determination of data requirements, and data sourcing towards this purpose in South Africa, are not elementary. Data availability is key throughout the modelling processes, making the sourcing of good data equally imperative. This paper provides starting considerations, and compiles a list of available data sources applicable to river water quality modelling in South Africa. It also provides a background on the included data sources and a general overview of the extent of the databases.

Keywords: calibration, validation, water quality data, hydrodynamic data, database

INTRODUCTION

Water quality modelling in South African rivers requires compliance with certain standards for both hydrodynamic and water quality data, as determined by the precision and accuracy needed for the model's output and the modelling goals (Daggupati *et al* 2015; Guillaume *et al* 2019). The availability and quality of data should be assessed early in the modelling process to inform the approach and to limit complexities. It is crucial to ensure that sufficient and appropriate data is present to generate realistic estimates of the desired model parameters (Guillaume *et al* 2019).

Modellers must investigate the quality and quantity of available data in the early stages of the modelling process to ensure suitability for modelling goals. However, sourcing relevant and accurate data for river water modelling in South Africa can be challenging. Modellers often encounter uncertainties in model choice and modelling goals, as well as uncertainties in where and how to begin the modelling process.

South Africa has a comprehensive river monitoring network and system, which is publicly available and organised by the South African Departments of Human Settlements, and Water and Sanitation. This covers major river systems in the

country and includes flow gauges, rain gauges and water quality sampling. Despite the extent of the network, data gaps can occur due to equipment breakdowns and outdated data (Slaughter *et al* 2017). Additional data collected by third parties and private companies to fill gaps in certain regions can be requested at a cost.

The assessment of water quality using established field-based techniques is commonly acknowledged for its precision, given the up-close examination of water quality factors. Nonetheless, this method is prohibitively expensive and unworkable when it comes to overseeing large geographical areas (Modiegi *et al* 2020). In recent years, the monitoring of water quality data in South Africa has steadily decreased, particularly since 2010, due to rising costs and a budget that has not kept up with monitoring needs, combined with a lessened capacity to adhere to sampling frequencies, leading to the closure of several monitoring stations (Silberbauer 2020).

The decline in water quality data monitoring has far-reaching implications for the management and protection of South Africa's water resources. This reduction in monitoring capacity reduces the ability to detect changes in water quality in a timely manner and to then respond accordingly,

Mahlathi CD, Brink IC, Wilms JM. River water quality modelling in South Africa: Considerations, sourcing and accessing of input data. *J.S. Afr. Inst. Civ. Eng.* 2024;66(1), Art. #1278, 10 pages. <http://dx.doi.org/10.17159/2309-8775/2024/v66n1a1>

Table 1 Spatial resolution recommendation table relevant to river models (relevant table section extracted from Baffaut *et al* 2015 and used here with permission)

Process category	Point (< 1 m ²)	Small catchment (100 m ² to 50 ha)	Watersheds (50 ha to 50 km ²)	River basin (> 50 km ²)
Water movement	Runoff generation, infiltration, evapotranspiration, perched water table, and preferential flow	Concentrated flow, subsurface flow, drainage, buffers, ponds, wetlands, variable source areas, exfiltration, and interflow	+ Streamflow, bank storage, riparian areas, groundwater flow, aquifer recharge, flood plain, point discharges, and water withdrawals	+ Reservoirs and major hydraulic structures
Nutrients and other agrochemicals	Soil/plant interactions; leaching to perched, shallow, and deep aquifers; sorption, transformation, and degradation in vadose zone and shallow aquifers	+ Transport with drainage and subsurface flow, transformations and degradations in ponds and wetlands	+ Stream transport, water/air and water/streambed exchanges, in-stream transformations and degradation, riparian area, algae and aquatic plants, and point discharges	+ Cycling in major hydraulic structures

making it more challenging to ensure the sustainability and preservation of these critical resources.

The availability and suitability of data for river water quality modelling in South Africa can pose challenges for modellers. It is important to assess the purpose of the modelling project and determine the necessary quantity and quality of data to ensure its usability. This requires careful planning and consideration of the limitations imposed by limited data.

The aim of this paper is therefore to provide starting considerations for modelling endeavours, present the results of extensive research on available databases in South Africa in the form of a comprehensive list (as of October 2023), and provide information on how to access data required for river water quality modelling in South Africa.

STARTING CONSIDERATIONS: RESOLUTION, CALIBRATION AND VALIDATION

The modelling process should balance efficiency and effectiveness by matching the complexity of the model with the complexity of the problem it aims to simulate and solve (requisite simplicity). This section

of the paper covers considerations such as model approach, resolution, calibration and validation strategies, approaches for types of water quality variables, model environment, and handling of loading data for different types of rivers.

The choice of the modelling approach is influenced by the intended purpose or goal of the model, shaping the direction in which the modelling process is undertaken (Tredennick *et al* 2021). The data requirements of each approach may vary, depending on the complexity and processes of the modelling environment (empirical versus mechanistic). Insufficient data may impact the reliability and validity of results.

Model resolution

Model resolution covers the spatial and temporal aspects of the environment that is modelled. The first step in model creation is to assess data needs in terms of quantity, quality and resolution to ensure accuracy and reliability. Planning for final validation and calibration is critical. However, limitations in available data may impact initial modelling decisions, such as generating data, using mathematical methods, or adjusting goals and strategy based on the limited data. Outcomes from two key reviews by Baffaut *et al* (2015)

and Daggupati *et al* (2015) summarise calibration and validation strategies for variable data availability scenarios and serve as examples of considerations, but do not constitute a comprehensive summary of all information sources on the topic. Baffaut *et al* (2015) highlight spatial and temporal considerations for hydrologic and water quality models. Tables 1 and 2 were extracted from these sources to summarise the temporal and spatial resolutions for modelling various water environments and water quality parameters.

In Table 1 processes active over a range of spatial scales are presented (the + symbol indicates that the processes simulated at smaller scales remain active at the current scale). The spatial resolution of the model can be determined by the size of the river reaches segmentation, subjective to the modelling objectives. This should be the smallest spatial element which the model output should represent.

The study recommends having representative input and calibration data available at a spatial and temporal extent and resolution that align with the model and study objectives. The model should be calibrated in successive steps, matching the critical processes and modelling the objectives' spatial and temporal resolutions.

Table 2 Recommended temporal resolutions for modelling natural river processes (extracted from Baffaut *et al* 2015 and used here with permission)

Process	Relevant time scales									
	Seconds	Minutes	Hours	Days	Weeks	Months	Years	Decades	Centuries	
Hydrodynamic processes										
Infiltrations	✓	✓	✓	✓	–	–	–	–	–	
Evapotranspiration	–	✓	✓	✓	–	–	–	–	–	
Preferential flow	✓	✓	✓	✓	–	–	–	–	–	
Soil moisture redistribution	–	✓	✓	✓	✓	–	–	–	–	
Runoff/overland	–	✓	✓	✓	–	–	–	–	–	

Continued on page 4

Table 2 continued

Process	Relevant time scales								
	Seconds	Minutes	Hours	Days	Weeks	Months	Years	Decades	Centuries
Quick return flow	–	✓	✓	✓	–	–	–	–	–
Field drainage	–	–	✓	✓	✓	–	–	–	–
Groundwater recharge / deep percolation	–	–	✓	✓	✓	✓	✓	✓	✓
Groundwater discharge/depletion	–	–	–	✓	✓	✓	✓	✓	–
Channel flow (small watershed < 0.5 km ²)	–	✓	✓	✓	–	–	–	–	–
Channel flow (large watershed 0.5 to 50 km ²)	–	–	✓	✓	✓	–	–	–	–
River flow (basin > 50 km ²)	–	–	✓	✓	✓	✓	–	–	–
Biological processes									
Plant (row crops, grasses, trees) growths	–	–	✓	✓	✓	✓	✓	✓	–
Carbon cycle: decomposition of fresh organic materials (crop residues, manure)	–	–	–	✓	✓	✓	✓	–	–
Accumulation and decomposition of soil organic matter	–	–	–	–	–	✓	✓	✓	✓
Bacteria growth and die-off	–	✓	✓	✓	–	–	–	–	–
Nutrients and pesticides: nitrification, denitrification, urea hydrolysis	–	✓	✓	✓	✓	✓	–	–	–
N and P mineralisation	–	–	✓	✓	✓	✓	✓	✓	✓
Degradation	–	–	✓	✓	✓	✓	✓	✓	–
Mixing by earthworms	–	–	–	–	✓	✓	✓	–	–
Algae growth/eutrophication	–	–	✓	✓	✓	✓	–	–	–
Erosion and sedimentation processes									
Detachment (rill/interrill)	✓	✓	✓	–	–	–	–	–	–
Gully erosion	✓	✓	✓	✓	–	–	–	–	–
River channel bed and bank erosion	–	–	✓	✓	–	–	–	–	–
Sediment transport	✓	✓	✓	✓	✓	✓	✓	✓	–
Landscape/stream deposition	–	✓	✓	✓	✓	✓	✓	✓	✓
Lake sedimentation	–	–	–	✓	✓	✓	✓	✓	✓
Geomorphologic adjustment	–	–	–	–	–	✓	✓	✓	✓
Physical and chemical processes									
Adsorption/desorption	✓	✓	✓	✓	✓	–	–	–	–
Solute transport, including leaching	✓	✓	✓	✓	–	–	–	–	–
Oxygen depletion	–	✓	✓	✓	–	–	–	–	–
Groundwater chemistry (mineral dissolution and chemical precipitation)	–	–	–	✓	✓	✓	✓	✓	–

Tables 1 and 2 can be used as a guide to regulate the input data required for models. Data quality and quantity required for model calibration and validation can be discerned from the recommended calibration and validation processes available for application at the preference of the modeller.

Model calibration and validation

Daggupati *et al* (2015) recommends a guideline for calibrating and validating hydrodynamic and water quality models. The guideline includes careful consideration, execution, and documentation of the three major calibration and validation elements, as listed in Table 3.

Elements 2 and 3 of the calibration and validation guidelines are linked to data quality and quantity requirements, as highlighted in this paper for South African data sources. Further details on each element can be found in the source article by Daggupati *et al* (2015). The allocation of available data for model validation and

Table 3 Important model calibration and validation elements to consider for hydrodynamic and water quality modelling (extracted from Daggupati *et al* (2015))

Calibration validation	Characteristics	Description
Element 1	Goals of model use	Include the purpose and end use of the model, spatial and temporal scales, as well as model accuracy and precision
Element 2	Data and parameters used in calibration and validation	The use of measured data to compare with model output and selection of model parameters are considered
Element 3	Model calibration and validation processes	Include warmup period, strategy complexity, spatio-temporal allocation of calibration and validation data, manual versus automatic calibration, as well as additional diagnosis

calibration requires splitting the data into two parts, and various data allocation methods exist in literature, such as temporal split-sample, spatial proxy basin, differential split-sample, and proxy basin differential split-sample. The reader is referred to Daggupati *et al* (2015) for more information.

The aim of this paper is to guide the selection of data sources for water quality modelling in South African river systems. However, before sourcing data, it is crucial for the modeller to comprehend the data requirements of the model as prescribed by relevant studies and guidelines, such as those discussed by Daggupati *et al* (2015) and Baffaut *et al* (2015).

Additionally, validation of water quality models can also be done using long-term simulations and frequency distributions. This involves comparing the simulated results with the available observed data. One approach to calibrate water quality models using long-term simulations and frequency distributions is through a method known as “split-sample validation”. The prevalent method for split-sample validation in the literature is the two-period approach, which involves dividing the calibration and validation periods into roughly equal sections. This technique offers the advantage of ease of implementation and minimising model runtime, particularly beneficial for computationally intensive hydrological and water quality models (Arsenault *et al* 2018). Another approach is to use a Monte Carlo simulation, which uses random sampling from the frequency distribution of the input data to generate multiple realisations of the model output (Atiem & Harmanciölu 2006). The model is then calibrated by comparing the frequency distribution of the observed data to the

distribution of the model output generated by the Monte Carlo simulation. A decision on which method is suitable for a specific application can be determined by the purpose of the study and further consultation with relevant literature.

Model environment and water quality parameters

To accurately model water quality in rivers and reservoirs, different data sets and understanding of processes are required. Rivers require continuous monitoring of water quality parameters, and an understanding of the transport of water and pollutants within the river. On the other hand, modelling water quality in reservoirs requires an understanding of storage, mixing, and release of water, as well as interactions with the surrounding environment, and often requires more diverse and detailed data sets such as bathymetric data, meteorological data, and information on water inputs and outputs. These differences in data requirements reflect the unique challenges and complexities of modelling water quality in each system.

The differentiation between the data requirements for non-conservative and conservative water quality parameters in river systems is based on the type of water quality parameters being modelled. Water quality parameters, whether conservative or non-conservative, are subjected to distinct modelling approaches. Conservative water quality parameters, like total dissolved solids, maintain a consistent concentration in the water column and do not engage in chemical interactions with the environment. Hence, proxy parameters, like electrical conductivity levels, are observed and linked to variations in the concentration of conservative parameters within the river system. This approach

is widely noted in the majority of water quality parameter studies found in the literature (Ranjith *et al* 2019). In contrast, non-conservative water quality parameters, like nutrients, can interact chemically with the environment, causing concentration changes over time. Modelling non-conservative water quality parameters, as demonstrated by reviewed assessments of fate and transport models (Addis *et al* 2023), necessitates the acquisition of high-quality measurement data. This data should encompass information on the sources and sinks of these substances, providing a comprehensive understanding of the intricate processes governing their behaviour in river systems.

In conclusion, the data requirements for modelling non-conservative water quality parameters in rivers are more complex and extensive compared to those for modelling conservative water quality parameters.

The data requirements for representing point and non-point sources of pollution loads in water systems also differ. Point sources refer to identifiable sources of pollution, such as wastewater treatment works, whereas non-point sources refer to diffuse sources of pollution, such as agricultural runoff. To accurately model point sources, detailed data on their discharge characteristics and the pollutants they release into the water system is required. This data may be obtained from monitoring and reporting systems, as well as regulations and guidelines for wastewater treatment works. In South Africa it is likely that data on wastewater treatment works is available through various sources, including governmental organisations and public utilities. However, further research and investigation may be needed to determine the availability and accuracy of such data.

SOUTH AFRICAN RIVER MODELLING DATA SOURCES

The Water Research Commission (WRC) in South Africa can provide valuable data resources for water quality modelling. The WRC has a daily rainfall database up until 2000 (Lynch 2004), as well as a database of daily air temperature (Schulze & Maharaj 2004), which can be useful for hydrologic and meteorologic applications in water quality modelling. These databases provide valuable information for researchers and modellers looking to understand the impacts of weather and climate on water quality in South Africa.

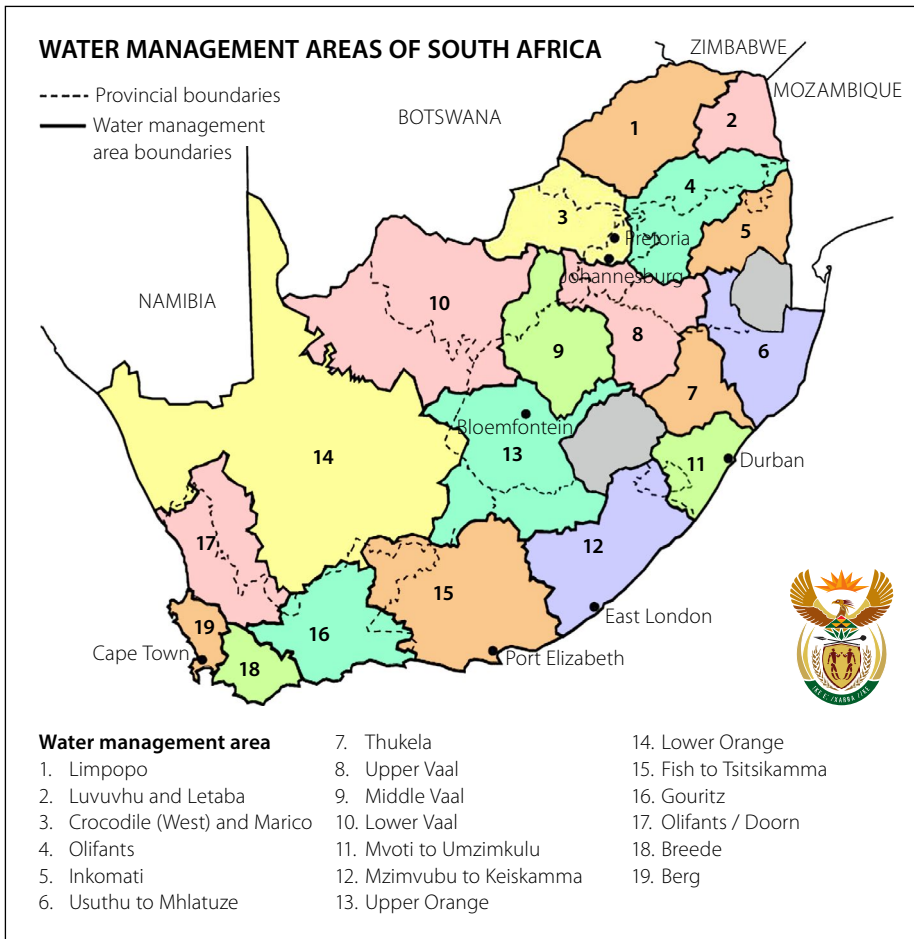


Figure 1 Map showing water management areas in South Africa (DWS 2016)

Data sources for river modelling in South Africa include both free and paid options from public and independent sources. The South African Department of Water Affairs (DWA) water quantity and quality monitoring network is a primary data source for water quality models, as reported and utilised in studies by Dabrowski (2014), Mahlathi *et al* (2016) and Slaughter *et al* (2017). It contains measuring station data for 19 water management areas, as shown in Figure 1.

This database can be accessed through the DWA website or dashboard online. Its network of monitoring stations covers most of the river systems in the country, but data augmentation may be required in certain cases, depending on the modelling purpose (Slaughter 2017). Independent sources are available and can be explored to augment data where needed. These data sources are summarised in Table 4 and discussed in more detail.

Climate and weather

Climate and weather data includes precipitation, air temperature, solar radiation, relative humidity, and wind speed data. This information is required as input for most mainstream water quality

Table 4 South African water data portals

Data type	Sources	Access	Number of stations/gauges	Website links
Hydrodynamic (river flow, runoff, discharges, groundwater)	GRDC	open, limited to research	317	https://portal.grdc.bafg.de/applications/public.html?publicuser=Publicuser#dataDownload/Stations
	DWA-NWIS	open	285	http://www.dwa.gov.za/niwis2/
	Water Resources of South Africa 2012 Study (WRC 2012)	open		https://waterresourceswr2012.co.za/
Water quality (chemical, biological, physical parameters)	DWA-NWIS	open	4 457	http://www.dwa.gov.za/niwis2/
	WRC 2012	open	unknown	https://waterresourceswr2012.co.za/
Climate and weather (temperature, humidity)	SAWS	open/paid	unknown	https://weathersa.co.za
	ARC-Soil, climate and water	paid	unknown	http://www.agis.agric.za/agisweb/nr_atlas
	Environmental GIS	open	unknown	https://egis.environment.gov.za/
	MiniSass	open	unknown	https://www.minisass.org/en/map/
River stream hydrogeometry (river location, length, depth, width, slope, cross-sectional area)	Google Earth	open	unknown	
	WRC 2012	open	unknown	https://waterresourceswr2012.co.za/
	DWA-NWIS	open	unknown	http://www.dwa.gov.za/niwis2/
Land use (land cover, vegetations)	SANBI	open	unknown	http://bgis.sanbi.org/SpatialDataset
	DWA-NWIS	open	unknown	http://www.dwa.gov.za/niwis2/
	WRC 2012	open	unknown	https://waterresourceswr2012.co.za/
	ESA's land cover portal			https://cds.climate.copernicus.eu/cdsapp#!/dataset/satellite-land-cover?tab=overview

Table 5 Modified SAWS countrywide observational network – number of stations with data relevant to river water quality modelling (Source: South African Weather Service (SAWS 2023))

Station type	Number of stations
Automatic weather stations (AWS)	231
Climate stations	12
Rainfall stations	1180
Automatic rainfall stations	153
Sea surface temperature stations	23
Meteorological radar systems	14
Global atmospheric watch station	1
Dobson ozone spectrophotometer station	2
Baseline surface radiation network station	1

Table 6 Temporal extent and resolution of SAWS climatic data for water quality modelling purpose (Source: South African Weather Service (SAWS 2023))

Data type	Temporal resolution	Temporal extent (starting year until present)	Stations
Rainfall values	daily	1836	all
Surface observations	daily	1884	some
Wind speed and direction	hourly	1950	all
Temperature	hourly	1950	all
Humidity	hourly	1950	all
Pressure	hourly	1950	all
Satellite data	unknown	1992	-
Radar	unknown	1994	-

models such as Soil Water Assessment Tool (SWAT), Water Quality Analysis Simulation Program (WASP) and Qual2kw. Numerous sources with datasets ranging from early 1920 up to the present exist in digital libraries. Examples of these libraries are the United States of America's National Oceanic and Atmospheric Administration (NOAA), and the local databases: the WRC daily rainfall and gridded daily temperature database and the South African Weather Service (SAWS). The SAWS database contains the most South African weather and climate data and is discussed further below.

SAWS

Tables 5 and 6 were created using the information on the SAWS climate data page to respectively display the countrywide observational network (Table 5), and the temporal extent and resolution of the available climatic data (Table 6).

Data can be accessed through a data enquiry. The SAWS website can be used to determine the correct information and the station of interest. The processing

time for data requests varies depending on the type, time frame and extent of the data. The SAWS climate database, which stores the data, can be accessed by clients for a fee, with different pricing for commercial and non-commercial use. It is possible that students may have access to the data at no cost. Rainfall station data, useful for predicting river flow, is also available. Resale of purchased data is not allowed and information on the SAWS Data Policy can be found on the website. SAWS provides weather data through its web-based tool, hydroNET, which uses data and models to make predictions. Documentation on data collection, processing and reporting is available on the data portal to support data reliability.

Stream hydrogeometry

Stream hydrogeometric data includes hydrologic and geometric info (Chapra 1997). No comprehensive database has been found, so methods of estimation are provided. Site-specific literature may

contain limited data for specific study areas. Digital tools (ArcGIS, QGIS) for measuring and estimating river characteristics are available for use by modellers.

The Department of Water and Sanitation – National Integrated Water Information Systems (DWS-NIWIS) website offers Google Earth files of the entire South African river network. The files can be viewed on Google Earth with measuring tools. For missing data, estimation techniques like point estimate, river reach estimate, low-flow analysis, discharge coefficients, and Manning's equation are recommended by Chapra (1997).

Land use and land cover

Land cover data can be determined by analysing satellite and aerial imagery, which captures coverage in the region (forests, wetlands, impervious surfaces, agriculture, and other land and water types) (NOAA 2020). The South African National Land-Cover Dataset (SANLCDS) serves as one of the main suppliers of land use and land cover data in South Africa. The dataset has been generated from 20-metre resolution multi-seasonal Sentinel 2 satellite imagery which represents the current full temporal range of available imagery (SANLCDS 2018).

Data can only be accessed once an account is created on the download link. Land use and land cover models are applied to study water quality in catchments and rivers with examples such as the work of Tahiru *et al* (2020). The work of Petersen *et al* (2017) is an example of a South African study. Additionally, Slaughter and Mantel (2017) used data from the South African National Land-Cover Dataset (SANLCDS) on land cover models to predict non-point nutrient inputs into rivers.

Hydrodynamic data

Table 4 (see page 6) lists sources used to estimate river flows for hydrodynamic models. The data includes geo-hydrodynamic properties, river flows, rainfall, point source discharges and water extraction. Two main data portals for hydrodynamic data are discussed.

The Global Runoff Data Centre (GRDC) (2020) is an international data repository functioning under the patronage of the World Meteorological Organisation (WMO). The database holds global hydrodynamic data, including from South Africa's public and private monitoring networks like the Department of Water

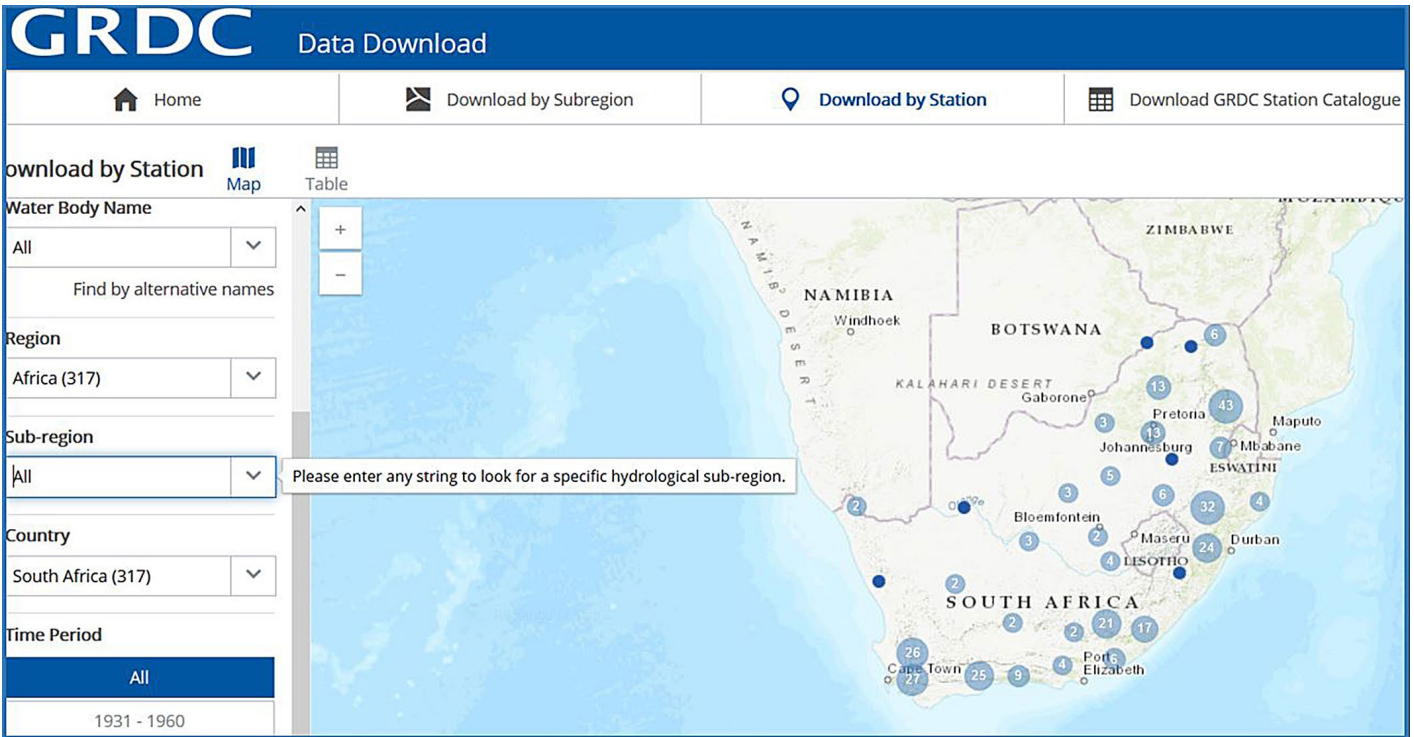


Figure 2 GRDC web portal for hydrodynamic data stations displaying station locations on the map

and Sanitation. Data can be viewed for free on the GRDC portal, but downloading requires a written motivation. Requests for specific river station data can be filtered by region.

For South Africa, applying the country filter produces a map (Figure 1 on page 6) with all 317 stations (Lesotho and Swaziland included). Available data including statistics (percentage missing monthly and daily data, data period range and catchment area) can be seen by clicking on a station point on the map (Figure 2). In addition to the detailed display of station data, a graph that plots the time series of flow data over the

range of the station data is displayed, which allows for quick observations of the time periods with gaps (Figure 3). The spatial and temporal resolution of the data vary for different sections, for example the Olifants River, which has more high-resolution data because of the number of studies (Ashton & Dabrowski 2011; Baker *et al* 2011; Dabrowski 2014; Mahlathi *et al* 2016; Slaughter *et al* 2017; Udall 2018) and projects conducted in the region.

The GRDC site provides runoff data for river hydrodynamic model calibrations and verifications, making it a valuable resource for water quality models. Its

validity and reliability are ensured through collaboration with national water agencies like South Africa's Department of Water and Sanitation. Data usage guidelines, including restrictions on distribution and commercial use, are stated on the GRDC website. Additional guidelines on the use and determination of data are detailed in the document titled Global Runoff Data Centre (GRDC 1995) that is received when access is requested

DWS-NIWIS

The South African National Integrated Water Information System portal

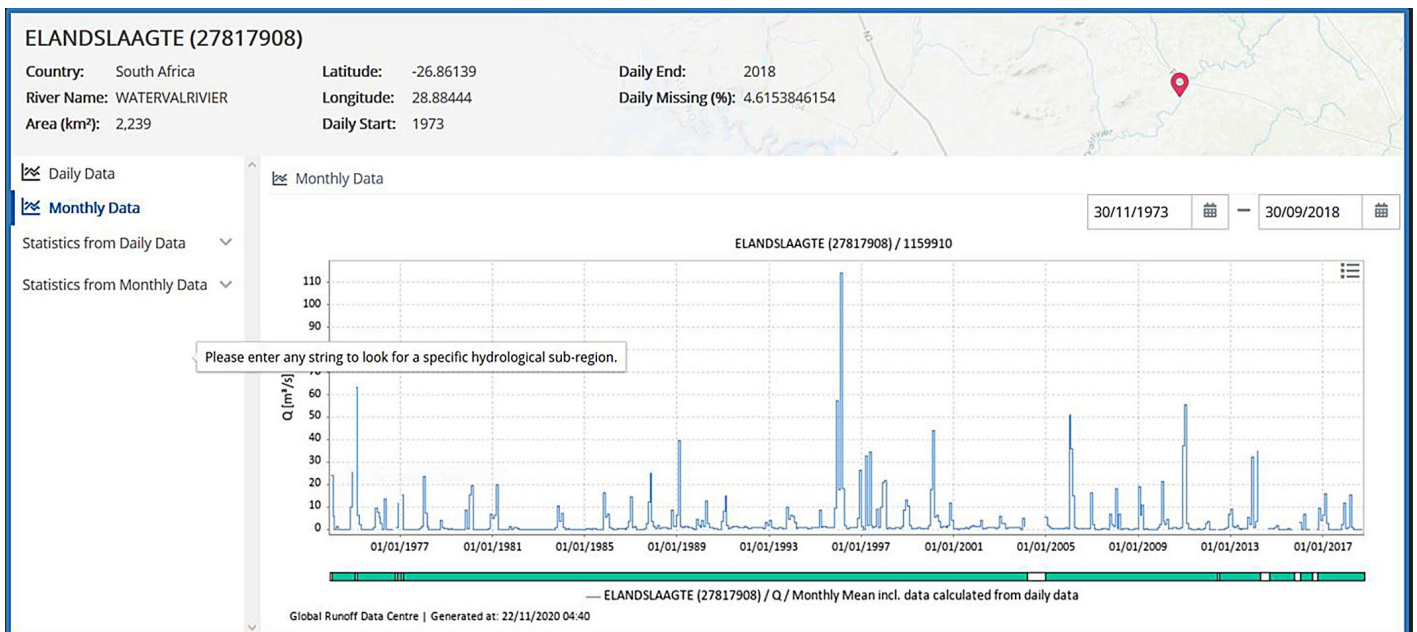


Figure 3 Data portal showing station details and river flow data visualisation in a form of time series plots

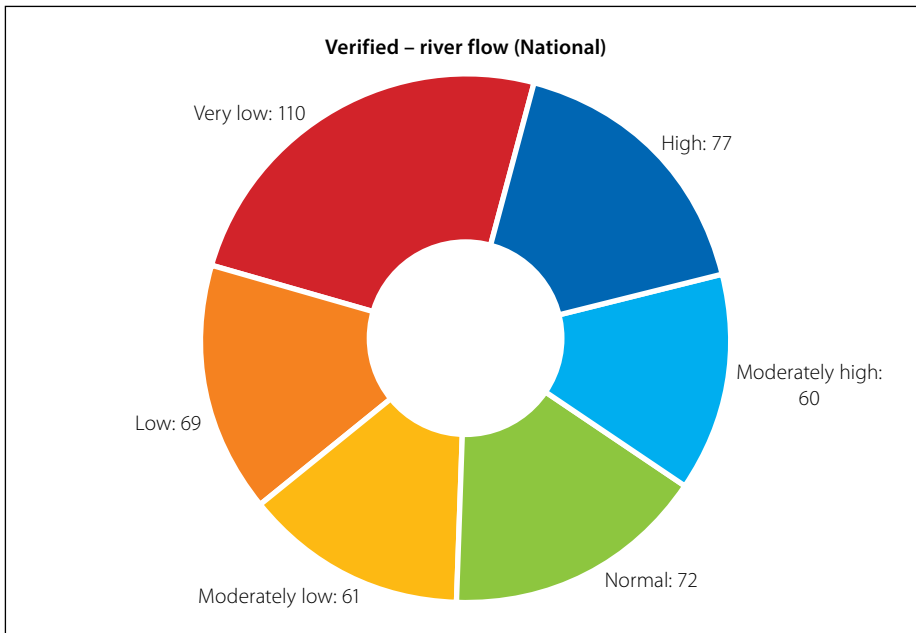


Figure 4 NIWIS chart showing the number of stations with verified river flow data with colour coding for the number of stations and level of monitoring ranging from low to high (redrawn from DWS-NIWIS portal 2015)

(DWS-NIWIS 2015) is an online platform managed by the Department of Water and Sanitation. It provides information products through dashboards and a web GIS interface that displays maps and locations of data stations for Surface Water, Ground Water, and Water Quality monitoring networks. The platform facilitates efficient analysis and reporting of water data across South Africa. The Water Quantity dashboard provides access to hydrodynamic data, including river flows, surface water storage, groundwater availability status, groundwater level status, and water transfers. Figure 4 refers.

River flow data

The DWS-NIWIS dashboard displays river flow data for 19 water management areas, with a varying number of stations. Verified river flow data is graphically depicted, with colour codes indicating low- to high-flow systems, as shown in Figure 5, and unverified data is available on the main database. The graphical representation allows users to easily identify the flow conditions of a particular river.

River flow data can be downloaded directly from the dashboard in the form of a *.csv*, *.xls* or *.pdf* file. The file contains date columns corresponding to river flow rates.

In addition, there are columns showing high-, moderate- and low-flow rates.

In the absence of a good set of measured hydrodynamic data, the WRC 2012 website provides comprehensive data, information, tools and models (Pitman and ACRU hydrological models, and the WRYM, WRPM, WReMP water management models) for water resource practitioners to study and plan their water resources. The website features GIS maps, the WRSM2000 (Pitman) rainfall-runoff model, and various databases, reports and spreadsheets. The WRSM model, managed by Allan Bailey, determines catchment-based rainfall, calibrates simulated flows against observed data, and produces naturalised flows and present-day flows. To fully understand the model and information system, it is recommended to attend the WRSM/Pitman and WRC 2012 information system courses; course schedule information can be obtained from Allan Bailey of Bailey and Pitman Water Resources (Pty) Ltd at allankb@netactive.co.za. Access to the website requires registration.

Temporal and spatial data augmentation methods may be required to correct the temporal resolutions of data from monthly to daily. Sources such as Slaughter *et al* (2015), and Hughes and Slaughter 2015 provide methods for disaggregation of flow data from monthly to daily flows. Recently, the work of Mahlathi *et al* 2022 details the impact of model input data augmentation methods on model output that can be used to evaluate some data augmentation methods.

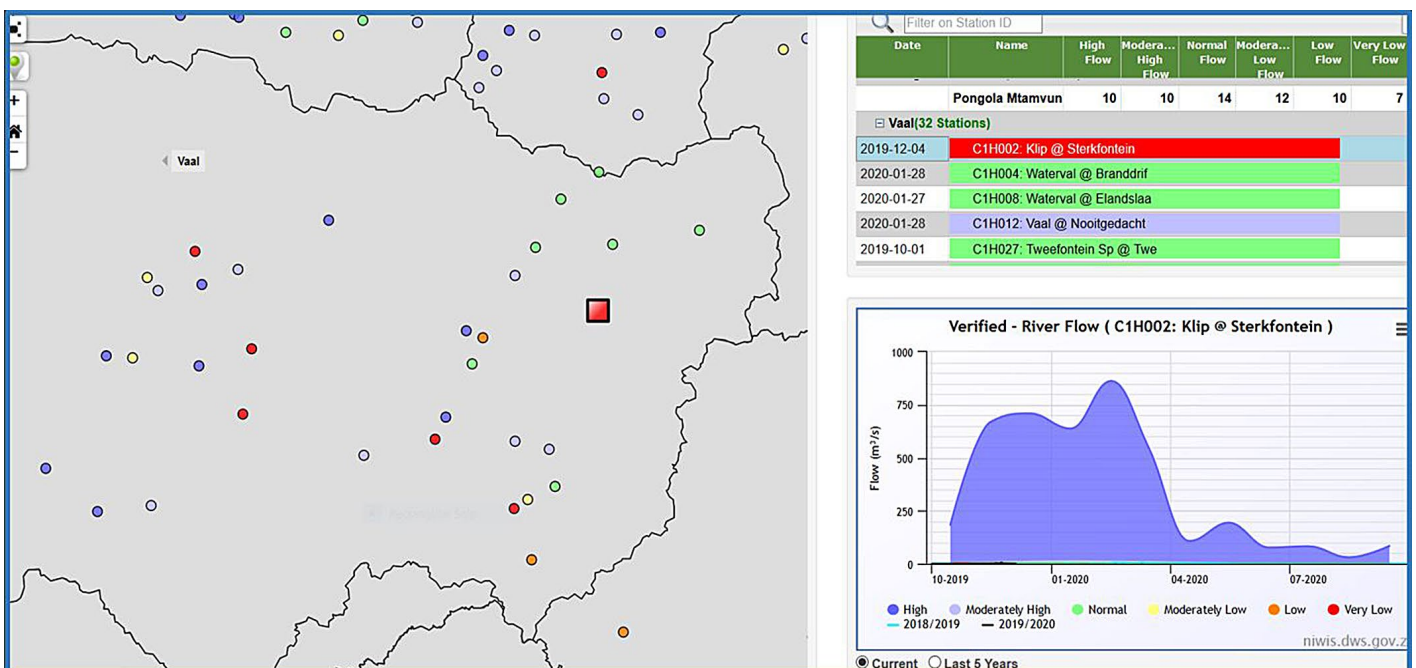


Figure 5 Map visualisation for Vaal catchment showing stations with colour codes (DWS-NIWIS portal 2015)

Water quality data

Water quality data is necessary to set up, calibrate and validate river water quality models. The data covers biological parameters (e.g. biological oxygen demand), chemicals (e.g. dissolved oxygen), and physical characteristics (e.g. water temperature). The main source of water quality data in South Africa is DWS-NIWIS and additional data can be obtained from third-party sources at a fee.

Water quality data can be obtained in a similar manner as hydrodynamic data. The water quality data is available on the water quality monitoring dashboard, which displays 4 457 monitoring stations. The stations are displayed on a table and GIS interface.

The portal provides a GIS interface with water quality data from monitoring points in South Africa. By selecting the river system on the map, users can access data for each station, including water quality information, as well as the locations and compliance data of wastewater treatment plants. Data can be viewed in a graph or downloaded in a .csv file.

Remote sensing data and other sources

In the context of river water quality modelling in South Africa, remote sensing data is a valuable resource. This data, including satellite imagery, aerial photography and hyperspectral data, provides insight into various water quality parameters, such as temperature, turbidity, dissolved oxygen and chlorophyll concentration. In the application of data various analysis and models can be created to estimate water quality parameters, as applied in various studies in literature (Van Deventer *et al* 2018; Li *et al* 2020; Seaton *et al* 2020; Adjovu *et al* 2023). However, it is crucial to combine remote sensing data with ground-based measurements. The drawback is that there is a lack of in situ data due to site cost, as well as a lack of high-level technical expertise, particularly in sub-Saharan Africa, for water quality monitoring (Adjovu *et al* 2023).

The spatial and temporal scales of remote sensing data that can be used for water quality modelling of river systems in South Africa are dependent on the type of data and the sensor utilised. Spatial resolution, a crucial factor in satellite imaging, defines the smallest real-world object captured in a two-dimensional image, determining the applicability of the product. High-resolution satellite imagery

and aerial photography excel at small-scale detail, whereas lower-resolution satellite imagery offers information at a larger scale with reduced detail (Valenzuela *et al* 2022).

The frequency of satellite overpasses, and the duration between acquisitions, significantly influence the temporal scale. While certain satellites furnish daily data, others cover an area only every few days or even weekly. Satellites with high resolution and a brief revisit time provide more frequent observations, enhancing the temporal scale (Trezza *et al* 2018).

It is essential to consider both the spatial and temporal scales when choosing remote sensing data for water quality modelling in South African river systems. The ideal spatial and temporal resolution will depend on the goals and requirements of the modelling project. For instance, high-resolution data may be necessary to accurately capture fine-scale changes in water quality, while lower-resolution data may be adequate for monitoring large-scale trends.

CONCLUSION

The process of finding data that suits the modelling requirement for a South African river system is not obvious. This paper therefore provides a guideline for accessing available river modelling data sources, and discusses relevant starting modelling considerations. Applying this knowledge may save time during the modelling exercise by directing the modeller to initial considerations and the relevant databases towards sourcing the applicable data. It is likely that data preparation will be required before modelling can commence. There are many other sources of data which this paper might not have covered, including independent sources; however, the summary presented here serves as a starting point for sourcing and accessing data for river water quality modelling in South Africa.

REFERENCES

- Addis, T Z, Adu, J T, Kumarasamy, M & Demlie, M 2023. Assessment of existing fate and transport models for predicting antibiotic degradation and transport in the aquatic environment: A review. *Water*, 15(8): 1511–1540. doi: <https://doi.org/10.3390/w15081511>.
- Adjovu, G E, Stephen, H, James, D & Ahmad, S 2023. Overview of the application of remote sensing in effective monitoring of water quality parameters. *Remote Sensing*, 15: 2439–2465. doi: <https://doi.org/10.3390/rs15071938>.

- Arsenault, R, Brissette, F & Martel, J 2018. The hazards of split-sample validation in hydrological model calibration. *Journal of Hydrology*, 566(September): 346–362. doi: <https://doi.org/10.1016/j.jhydrol.2018.09.027>.
- Ashton, P J & Dabrowski, J M 2011. An overview of surface water quality in the Olifants River catchment. WRC Report KV 293/11. Pretoria: Water Research Commission.
- Atiem, I A & Harmanciölu, N B 2006 Assessment of Regional Floods Using L-Moments Approach: The Case of The River Nile. *Water Resources Management*, 20(5): 723–747. doi: <https://doi.org/10.1007/s11269-005-9004-0>.
- Baffaut, C, Dabney, S M, Smolen, M D, Youssef, M A, Bonta, J V, Chu, M *et al* 2015. Hydrologic and water quality modeling: Spatial and temporal considerations. *Transactions of the ASABE*, 58: 1661–1680. doi: <https://doi.org/10.13031/trans.58.10714>.
- Baker, T, September, L & Van Veelen, M 2011. Development of a reconciliation strategy for the Olifants River water supply system – Environmental Screening Report. Aurecon for the Department of Water Affairs. DWA Report P WMA 04/B50/00/8310/13. Pretoria: DWA.
- Chapra, S 1997. *Surface Water-Quality Modeling*. Long Grove, IL: Waveland Press.
- Dabrowski, J M 2014. Applying SWAT to predict ortho-phosphate loads and trophic status in four reservoirs in the upper Olifants catchment, South Africa. *Hydrology and Earth System Sciences*, 18: 2629–2643. doi: <https://doi.org/10.5194/hess-18-2629-2014>.
- Daggupati, P, Pai, N, Ale, S, Zeckoski, R W, Jeong, J, Parajuli, P B *et al* 2015. A recommended calibration and validation strategy for hydrologic and water quality models. *Transactions of the ASABE*, 58: 1705–1719. doi: <https://doi.org/10.13031/trans.58.10712>.
- DWS (Department of Water and Sanitation) 2016. Directorate: Catchment Management Agencies. <https://www.dws.gov.za/IO/cma.aspx>.
- DWS-NIWIS (Department of Water and Sanitation – National Integrated Water Information Systems) 2015. DWS-NIWIS portal. <http://www.dwa.gov.za/niwis2/%0A> (accessed: 12 November 2020).
- GRDC (Global Runoff Data Centre) 1995. Policy guidelines for the dissemination of data. Geneva, Switzerland: World Meteorological Organization.
- GRDC 2020. The Global Runoff Data Centre. Koblenz: Germany: Federal Institute of Hydrology. <https://gdk.gdi-de.org/geonetwork/srv/api/records/cee45bde-a6ad-4b0f-9d63-23d46837b958>.
- Guillaume, J H A, Jakeman, J D, Marsili-Libelli, S, Asher, M, Brunner, P, Croke, B *et al* 2019. Introductory overview of identifiability analysis: A guide to evaluating whether you have the right type of data for your modeling purpose. *Environmental Modeling & Software*,

- 120: 104–514. doi: <https://doi.org/10.1016/j.envsoft.2019.07.007>.
- Hughes, D A & Slaughter, A R 2015. Daily disaggregation of simulated monthly flows using different rainfall datasets in southern Africa. *Journal of Hydrology: Regional Studies*, 4: 153–171. doi: <https://doi.org/10.1016/j.ejrh.2015.05.011>.
- Li X, Huang M & Wang, R 2020. Numerical simulation of Donghu Lake hydrodynamics and water quality based on remote sensing and Mike 21. *ISPRS International Journal of Geo-Information*, 9: 94–114. doi: <https://doi.org/10.3390/ijgi9020094>.
- Lynch, S D 2004. Development of a Raster database of annual, monthly, and daily rainfall for southern Africa. WRC Report 1156/1/04. Pretoria: Water Research Commission.
- Mahlathi, C, Siyakatshana, N & Chirwa, E 2016. Water quality modelling and optimisation of wastewater treatment network using mixed integer programming. *Water SA*, 42(4): 563–571. doi: <https://doi.org/10.4314/wsa.v42i4.16>.
- Mahlathi, C D, Wilms, J & Brink, I 2022. Investigation of scarce input data augmentation for modelling nitrogenous compounds in South African rivers. *Water Practice and Technology*, 17(12): 2499–2515. doi: <https://doi.org/10.2166/wpt.2022.146>.
- Modiegi, M, Rampedi, I T & Tesfamichael, S G 2020. Comparison of multi-source satellite data for quantifying water quality parameters in a mining environment. *Journal of Hydrology*, 591(June): 125322. doi: <https://doi.org/10.1016/j.jhydrol.2020.125322>.
- NOAA (National Oceanic and Atmospheric Administration) 2020. What is the difference between land cover and land use? Washington, DC: NOAA. <https://oceanservice.noaa.gov/facts/lclu.html> (accessed: 5 February 2021).
- Petersen, C, Jovanovic, N, Le Maitre, D & Grenfell, M 2017. Effects of land use change on streamflow and stream water quality of a coastal catchment. *Water SA*, 43: 139–152. doi: <https://doi.org/10.4314/wsa.v43i1.16>.
- Ranjith, S, Shivapur, A V, Kumar, P S K, Hiremath, C G & Dhungana, S 2019. Water Quality Model for Streams: A Review. *Journal of Environmental Protection*, 10(12): 1612–1648. doi: <https://doi.org/10.4236/jep.2019.1012097>.
- SANLDCS (South African National Land-Cover Dataset) 2018. Data portal: https://egis.environment.gov.za/data_portal/dataset/sanlcs-2018 (accessed 5 February 2021).
- SAWS (South African Weather Service) 2023. Home page. <https://www.weathersa.co.za/home/overview> (accessed on 5 November 2023).
- Schulze, R E & Maharaj, M 2004. Development of a database of gridded daily temperatures for southern Africa. WRC Report 1156/2/04. Pretoria: Water Research Commission.
- Seaton, D, Dube, T & Mazvimavi, D 2020. Use of multi-temporal satellite data for monitoring pool surface areas occurring in non-perennial rivers in semi-arid environments of the Western Cape, South Africa. *ISPRS Journal of Photogrammetry and Remote Sensing*, 167: 375–384. doi: <https://doi.org/10.1016/j.isprsjprs.2020.07.018>.
- Silberbauer, M 2020. Internet-based applications for interrogating 50 years of data from the South African national water quality monitoring network. *Hydrological Sciences Journal*, 65(5): 726–734. doi: <https://doi.org/10.1080/02626667.2019.1645334>.
- Slaughter, A R, Retief, D C H & Hughes, D A 2015. A method to disaggregate monthly flows to daily using daily rainfall observations: Model design and testing. *Hydrological Sciences Journal*, 4(B): 153–171. doi: <https://doi.org/10.1080/02626667.2014.993987>.
- Slaughter, A R 2017. Simulating microbial water quality in data-scarce catchments: An update of the WQSAM model to simulate the fate of *Escherichia coli*. *Water Resources Management*, 31: 4239–4252. doi: <https://doi.org/10.1007/s11269-017-1743-1>.
- Slaughter, A R & Mantel, S K 2017. Land cover models to predict non-point nutrient inputs for selected biomes in South Africa. *Water SA*, 43(3): 499–508. doi: <https://doi.org/10.4314/wsa.v43i3.15>.
- Slaughter, A R, Cullis, J D S, Addo-Bediako, A & Quinn, N W T 2017. A management-oriented water quality model for data scarce catchments. *Environmental Modelling & Software*, 97: 11–22. doi: <https://doi.org/10.1016/j.envsoft.2017.07.015>.
- Tahiru, A A, Doke, D A & Baatuuwue, B N 2020. Effect of land use and land cover changes on water quality in the Nawuni Catchment of the White Volta Basin, Northern Region, Ghana. *Applied Water Science*, 10(8): 1–14. doi: <https://doi.org/10.1007/s13201-020-01272-6>.
- Tredennick, A T, Hooker, G, Ellner, S P & Adler, P B 2021. A practical guide to selecting models for exploration, inference, and prediction in ecology. *Ecology*, 102(6): 1–14. doi: <https://doi.org/10.1002/ecy.3336>.
- Trezza, R, Allen, R G, Kilic, A, Ratcliffe, I & Tasumi, M 2018. Influence of Landsat revisit frequency on time-integration of evapotranspiration for agricultural water management. *Advanced Evapotranspiration Methods and Applications*, 1: 1–13. doi: <https://doi.org/10.5772/intechopen.80946>.
- Udall, B 2018. 21st Century climate change impacts on Olifants River Flows, South Africa. Centre for Environmental Rights. <https://cer.org.za/wp-content/uploads/2018/05/Udall-Olifants-River-Analysis-FINAL.pdf> (accessed on 5 February 2021).
- Valenzuela, A, Reinke, K & Jones, S 2022. A new metric for the assessment of spatial resolution in satellite imagers. *International Journal of Applied Earth Observations and Geoinformation*, 114: 103051. doi: <https://doi.org/10.1016/j.jag.2022.103051>.
- Van Deventer, H, Smith-Adao, L, Petersen, C, Mbona, N, Skowno, A & Nel, J L 2018. Review of available data for a South African inventory of inland aquatic ecosystems (SAIIAE). *Water SA*, 44: 184–199. doi: <https://doi.org/10.4314/wsa.v44i2.05>.
- WRC (Water Research Commission) 2012. Water Resources of South Africa 2012 Study. <https://waterresourceswr2012.co.za/>.