Journal of Hydrology: Regional Studies 59 (2025) 102387

Contents lists available at ScienceDirect

Journal of Hydrology: Regional Studies

journal homepage: www.elsevier.com/locate/ejrh

South African soil, land cover and weather generator file databases for SWAT applications

Jay le Roux [a,*], Ndifelani Mararakanye [b], Michael van der Laan [c,1], Leushantha Mudaly [c], Harold Louw Weepener [d], Johan van Tol [e]

[a] *Department of Geography, Associated Academic: Afromontane Research Unit, University of the Free State, Bloemfontein 9300, South Africa*
[b] *Department of Geography, University of the Free State, Bloemfontein 9300, South Africa*
[c] *Department of Plant and Soil Science, University of Pretoria, Pretoria 0028, South Africa*
[d] *Agricultural Research Council – Soil, Climate and Water, Pretoria 0001, South Africa*
[e] *Department of Soil, Crop and Climate Sciences, Associated Academic: Afromontane Research Unit, University of the Free State, Bloemfontein 9300, South Africa*

ARTICLE INFO

Dataset link: South African soil, land cover and weather generator file databases for SWAT

Keywords:
QSWAT
ArcSWAT
Input datasets
Big data
Data evaluation
South Africa

ABSTRACT

Study region: South Africa.
Study focus: The focus of the study is to develop soil, land cover and weather generator file datasets for Soil and Water Assessment Tool (SWAT) applications in South Africa. The first objective was to format national datasets for use as baseline to run the SWAT model in South Africa. The second objective was to evaluate the performance of the baseline input data by applying the national datasets in four (previously simulated) research catchments.
New hydrological insights for the region: The input datasets comprise of geo-spatial datasets at a national scale to run ArcSWAT or QSWAT (graphical user interface for SWAT in ArcGIS and SWAT+ in QGIS, respectively) in South Africa including: SWAT catchment outline data (tertiary and quaternary); Land cover maps at 20–30 m resolution including South African National Land Cover (2014, 2018, 2020) linked to SWAT land cover codes; A soil map with SWAT attribute data derived from pedotransfer functions of the Land Type Database of South Africa useable at a scale of 1:250,000; Weather statistics (WGN) files for 12 weather stations obtained from the Agricultural Research Council in South Africa. The national baseline data is an important step forward in hydrological modelling by assisting modellers to set-up and run the SWAT model in South Africa.

1. Introduction

The Soil and Water Assessment Tool (SWAT) is widely utilized to assess hydrological processes such as streamflow, water erosion, sediment yield dynamics and nutrient inputs/outputs (Zhao et al., 2024). SWAT is routinely coupled within GIS platforms, which offer unprecedented flexibility in the representation and organization of spatial data (Chen and Mackay, 2004; Gassman et al., 2014). SWAT is a catchment-scale and continuous time model operating on a daily time-step to simulate water, sediment and chemical fluxes in large

* Corresponding author.
E-mail addresses: Lerouxjj@ufs.ac.za (J. le Roux), nmararak@gmail.com (N. Mararakanye), VanDerLaanM@arc.agric.za (M. van der Laan), leushantha.mudaly@up.ac.za (L. Mudaly), weepenerh@arc.agric.za (H.L. Weepener), vanTolJJ@ufs.ac.za (J. van Tol).
[1] Current address: Agricultural Research Council – Soil, Climate and Water, Pretoria 0001, South Africa

https://doi.org/10.1016/j.ejrh.2025.102387
Received 10 February 2025; Received in revised form 9 April 2025; Accepted 10 April 2025
Available online 16 April 2025
2214-5818/

catchments with varying climatic conditions, soil properties, stream channel characteristics, land use, and management practices (Arnold et al., 2012; Bieger et al., 2017; Srinivasan et al., 1998). The model considers most hydrological and sedimentological aspects into one simulation package, including factors controlling runoff on hillslopes and streamflow in river channels, as well as sediment generation, channel transport, and deposition in sinks (Gassman et al., 2007). Although SWAT and its baseline input datasets were developed for use in the USA, the model has gained international acceptance and has been widely applied to support various large catchment (10–10 000 km^2) modelling studies across the world (e.g. Aloui et al., 2023; Francesconi et al., 2016; Tan et al., 2019; van Griensven et al., 2012; Zhao et al., 2024). Abbaspour et al. (2019) states that the SWAT model averages more than 550 peer-reviewed publications per annum.

SWAT has also been applied in Africa with a total of 206 publications listed on the SWAT website database over a 15-year period (2005–2019) (Akoko et al., 2021). According to the review by Akoko et al. (2021), the SWAT model is widely used in African countries being suitable for large-scale applications in data-scarce areas and ungauged catchments. Applications vary between model parameterization and dataset inputs (Chawanda et al., 2020; Nkwasa et al., 2020), the evaluation of water resources and streamflow (e.g. Guzha et al., 2018), erosion and sedimentation (e.g. Opiyo, 2024), nonpoint source pollution modelling (e.g. Zhang et al., 2024), and more notably, land-use management and climate change contexts (e.g. Ebodé et al., 2024). SWAT has also been applied in South Africa (SA) to support various large catchment modelling studies.

The first published application of SWAT was conducted in the Mkabela Catchment (41 km^2) near the town Wartburg in the KwaZulu-Natal Province (le Roux et al., 2013). SWAT effectively identified sediment source areas, as well as storages where connectivity is reduced at the catchment scale. The first large catchment application of SWAT was conducted in the Upper Olifants River Catchment (11,464 km^2) to model flow, sediment and nutrient outputs (Dabrowski et al., 2013). Results illustrate that SWAT effectively identifies drivers (point source over non-point source pollution) and specific source areas (sub-catchments) responsible for high ortho-phosphate loading in large catchments. The largest catchment application of SWAT was conducted in the Mzimvubu River Catchment (19 826 km^2), the only large river network in SA without a dam (le Roux, 2018). Modelling the flow and sediment yield made it possible to estimate life expectancies for 2 proposed dams on the Tsitsa River (between 43 and 55 years). Another catchment application of SWAT worth mentioning was conducted in the upper uMgeni Catchment (498 km^2), in which the Mkabela Catchment is nested (Scott-Shaw et al., 2020). By means of scenario analysis, results illustrated a 16 % increase in annual streamflow if wetlands with a 20-m buffer were restored to a natural state. More recent studies applied SWAT to assess the long-term impacts of land use/land cover (LULC) and climate change on streamflow (Mabuda et al., 2024; Woyessa, 2024) and sediment and/or nutrient loads (Mararakanye et al., 2022). Their results indicate that climate and LULC changes correspond to significant changes in streamflow and sediment/nutrient outputs. In addition, several recent studies applied a hydropedological approach to reflect flowpaths through detailed routing in SWAT in different research catchments in SA (Harrison et al., 2022; Julich et al., 2022; Smit et al., 2024; van Tol et al., 2020, 2021; van Zijl et al., 2020). Results confirm that hydropedological insights can be used in data-scarce areas as soft data to improve the representation of hydrological processes and model accuracy. One of the biggest challenges for the studies mentioned above, however, was to obtain appropriate input data for use in SWAT, especially soil data.

The problem is hydrological models interfaced in a GIS with their own geo-spatial input datasets are absent in most resource constrained countries such as SA (Akoko et al., 2021; Glenday et al., 2024). Input data preparation and model set-up is a laborious task, especially due to the lack of appropriate and representative data. A large part of modelling effort goes into the construction of input datasets (Akoko et al., 2021; Escamilla-Rivera et al., 2022). At a global scale, Abbaspour et al. (2019) prepared datasets of soil, land use, actual evapotranspiration, as well as weather datasets that could serve as standard inputs in SWAT models. However, the application of global datasets at a catchment scale could lead to large hydrological variances and uncertainty in catchment simulations (Escamilla-Rivera et al., 2022; Koo et al., 2020). Akoko et al. (2021) states that a major drive supporting improved data collection and sharing is required in Africa. Datasets are needed to expand SWAT databases which could aid model set-up and parameterization, as well as standardize SWAT modelling efforts in Africa. This study is an important step forward to address this challenge by providing appropriate data for use in SWAT in SA. In this context, the aim of the study is to develop soil, land cover and weather generator file datasets for SWAT applications in SA, including ArcSWAT-2012 and QSWAT. The aim was achieved by means of the following objectives. The first objective was to format national data for use as baseline to run the SWAT model (ArcSWAT-2012 or QSWAT) in SA. ArcSWAT-2012 is the graphical user interface for SWAT and ArcGIS software extension (Srinivasan et al., 1998), whereas QSWAT is the graphical user interface for SWAT+ and QGIS software extension (Bieger et al., 2017). The second objective was to apply the national baseline datasets in four (previously simulated) research catchments to compare the results (flow and sediment yield) of the two different input data models (catchment and national), as well as hydrological accuracy against measured streamflow data. Comparing the results and accuracies of the two input data models (catchment and national) allowed appraisal of the performance of the baseline input data.

The input datasets are stored in the Water Research Observatory (WRO) data portal: https://www.waterresearchobservatory.org/data-and-resources/hydrological-data-and-modelling (van der Laan et al., 2024), as well as on Mendeley Data: https://data.mendeley.com/datasets/5nzr53w6jp/2. The national input datasets assist SWAT-users to set up and run the model in SA, including digital elevation data, catchment outlines, land cover maps and codes, soil map and attribute table, and weather statistics for specific coordinates (required as input by the SWAT model). These datasets consist of more detailed and higher resolution data than the global datasets of Abbaspour et al. (2019) and provide better representation of soil characteristics in SA than the recent global DSOLMap of López-Ballesteros et al. (2023). Database comparisons are discussed in the Results and discussion section. The datasets not only save time with model set-up, it also assists in the standardization of SWAT modelling efforts in SA.

2. Materials and methods

2.1. Formatting national baseline data

National data formatted included catchment outline data, land cover data, soil data, and weather statistics. Metadata of the national datasets were drafted based on ISO 19115 Geospatial metadata standards (see Appendix A: Metadata). Catchment outline data were obtained and prepared from the hydrologically corrected 90 m SRTM DEM and derived products of Weepener et al. (2012), including minor corrections/updates in 2018 (DWS, 2022). The limitation of this DEM is it has a relatively course resolution. The reason for using this DEM instead of DEMs with finer a resolution is because it has been hydrologically corrected and subsequently provided appropriate tertiary and quaternary catchment outlines at a national scale. However, users can decide to use any other DEM or catchment outlines such as the South African Atlas of Climatology and Agrohydrology (Schulze, 2007). Latter mentioned database is currently improved by the Centre for Water Resources Research at the University of KwaZulu-Natal in SA, including a set of sub-quaternary catchment boundaries (called quinary catchments) that are nested within the DWS 2018 quaternary catchment boundaries.

National Land Cover maps (SANLC, 2014; 2018; 2020) with 72–73 land cover classes at 20 m resolution were linked to 27 land cover types in the SWAT database. See Appendix B indicating the list of land cover types in the SWAT database (excluding parameters) that was linked to the National Land Cover (2014, 2018 and 2020) maps of SA. The limitation of linking the SANLC maps to land cover types in the SWAT database is subsequent use of default SWAT parameters outside of conditions for which it was developed. Therefore, phenological plant development is not based on local conditions and plant growth cycles.

Soil texture and hydraulic parameter values were assigned to the Land Types of SA, a national soil, climate and terrain database with polygons demarcatable at a scale of 1:250 000 (ARC Land Type Database, 2012). Pedotransfer functions based on the same approach of van Tol et al. (2013), van Zijl et al. (2016) and van Tol et al. (2020) were used to generate the required soil parameters. Hydraulic soil parameters are provided in Schulze (2007), which are based on the work of Schulze et al. (1985), Schulze et al. (1991) and Schulze and Horan (2005). Table 1 indicates the descriptions and methodology used to assign soil parameter values to Land Types at a national scale. The limitations of using land type data as the basis for soil modelling inputs were discussed in detail by van Tol and van Zijl (2020) and van Tol and van Zijl (2022). Briefly, a land type is not a soil polygon but rather an area characterized by homogeneous soil distribution patterns. Within a land type, vastly different soils can occur. In this study, we used weighted average

Table 1
Descriptions and methodology used to assign soil parameter values to each soil component (ARC Land Type Database, 2012) at a national scale.

Soil parameter	Methodology/reasoning
Number of layers in the soil	Two soil layers/horizons were incorporated into each soil component of the ARC Land Type Database (2012).
Depth from soil surface to bottom of layer (mm)	Depth descriptions/classes in the ARC Land Type Database (2012) and Schulze (2007) were used to assign depth to each Land Type of SA.
Maximum rooting depth of soil profile (mm)	As above.
Soil Hydrologic Group (A,B,C,D) in terms of runoff potential, Soil Group A = low, B = moderately low, C = moderately high, D = high.	Soil hydrological groups were based on the broad soil patterns given in the ARC Land Type Database (2012) as follows: A for deep and freely drained apedal soils with humic topsoils as well as podzols; B for apedal soils with plinthic subsoils or deep alluvial soils; C for shallow soils or planosols comprising sandier topsoil abruptly overlying more clayey subsoil; D for rock outcrops.
Available water capacity of the soil layer (mm H_2O/mm soil)	For each Land Type, Schulze (2007) calculated plant available water content as the difference between water content at field capacity and permanent wilting point.
Saturated hydraulic conductivity (mm/hr)	Values were derived from the Rosetta Model (Schaap et al., 2001) based on the soil texture classes for each Land Type.
Bulk density (Mg/m^3 or g/cm^3)	Bulk density (BD) was estimated using porosity (PO) data in Schulze (2007) for each Land Type: PO = 1-BD/2.65.
Soil albedo (non-dimensional value between 0 and 1)	Albedo values were assigned to broad soil patterns in the ARC Land Type Database (2012) ranging between 0.25 for light-coloured sands to 0.7 for dark clays.
Texture of soil layer [optional]	Texture classes were assigned using clay classes given to each Land Type.
Clay content with diameter of < 0.002 mm (% soil weight)	Clay content in the A-horizon was assigned using the average topsoil clay classes given to each Land Type. Clay content in the B-horizon was assigned to each Land Type by adjusting the clay values of the A-horizon to clay-factors given in le Roux et al. (2023).
Silt content with diameter of 0.05–0.002 mm (% soil weight)	Due to the lack of data, silt content for A and B horizons were assigned values between 10 % and 22.5 %, increasing with increase in clay as follows (le Roux et al., 2023): percentage of Land Type with < = 6 % clay = 10 % silt; 6.1–15 % clay = 15 % silt; 15.1–25 % clay = 17.5 % silt; 25.1–35 % clay = 20 % silt; 35.1–55 % clay = 22.5 % silt.
Sand content with diameter of > 2 mm (% soil weight)	Sand content for A and B horizons were assigned as follows: Sand = 100 % – (%clay + %silt).
Rock fragment content (% soil weight)	Rock content was based on the agricultural restriction/rock (MB) classes in the ARC Land Type Database (2012) as follows: MB0 = 0 % (no rock); MB1 = 20 %; MB2 = 50 %; MB3 = 20 %; MB4 = 100 % (no soil).
Organic carbon content (% soil weight)	A soil organic carbon map of SA of Schulze and Schütte (2020) was used to assign average carbon values for A and B horizons per Land Type.
USLE K factor in SI units t/ha per unit 'erosivity'	Erodibility units established by le Roux et al. (2008) were assigned to each Land Type.

values for each land type to obtain soil parameters. It is important to note that the land type survey generally applied a depth criterion of 1.2 m. Soils, particularly in the eastern part of South Africa, are often considerably deeper than this, which can impact storage capacity, plant-available water, and recharge when simulated using land type inputs. Additionally, the Rosetta pedotransfer function used for estimating saturated hydraulic conductivity relies on particle size distribution but does not distinguish between types of clay. Highly weathered soils may contain significant amounts of kaolinitic clay, which tends to have high conductivity and behaves similarly to sandy soils. In contrast, even small increases in montmorillonitic clay can significantly reduce hydraulic conductivity. These limitations should be considered when interpreting modelling results, and future research should continuously aim to improve soil inputs for modelling purposes.

Weather Generator (WGN) input files consist of weather statistics including precipitation, temperature, solar radiation, relative humidity and wind speed. WGN files are required by SWAT to generate representative daily climate data for simulated catchments in two instances: when the user specifies that simulated weather will be used, or when measured data is missing. WGN files were created by acquiring and interpreting climate data from the Agricultural Research Council (ARC) Agroclimatology weather stations in different climate zones in SA. The completeness of climate data was the most important consideration when selecting weather stations. Using the SWAT Weather Database of Essenfelder (2016), two sets of (12) WGN files were prepared covering the periods 1981–2000 and 2001–2020 respectively. The limitation worth mentioning is the national baseline datasets exclude climate data (besides weather generator files), particularly rainfall data, which is necessary to consider the spatial distribution of rainfall throughout a catchment. Climate data are usually not freely available in developing countries and often incomplete at a regional scale (Schuol and Abbaspour, 2007; Glenday et al., 2024).

2.2. *Application of baseline input data in four research catchments*

The second step was to apply the input data in four (previously simulated) research catchments, including Middle Olifants River Catchment (MORC), Lower Vaal River Catchment (LVRC), Mkabela Catchment (MC) and Tsitsa River Catchment (TRC) (see Fig. 1). These catchments were simulated using the same weather data, over the same timeframes, as before. Simulations were performed using ArcSWAT-2012, which is a graphical user interface for SWAT and ArcGIS® software extension. Comparing the results and accuracies of the two input datasets (original input versus baseline input), allowed appraisal of the performance of the baseline input data. Table 2 provides summary site descriptions of the four research catchments.

2.3. *Model configuration and parameterization of the four research catchments*

Table 3 provides a summary of the input data that were originally utilized to configure and parameterize the four research

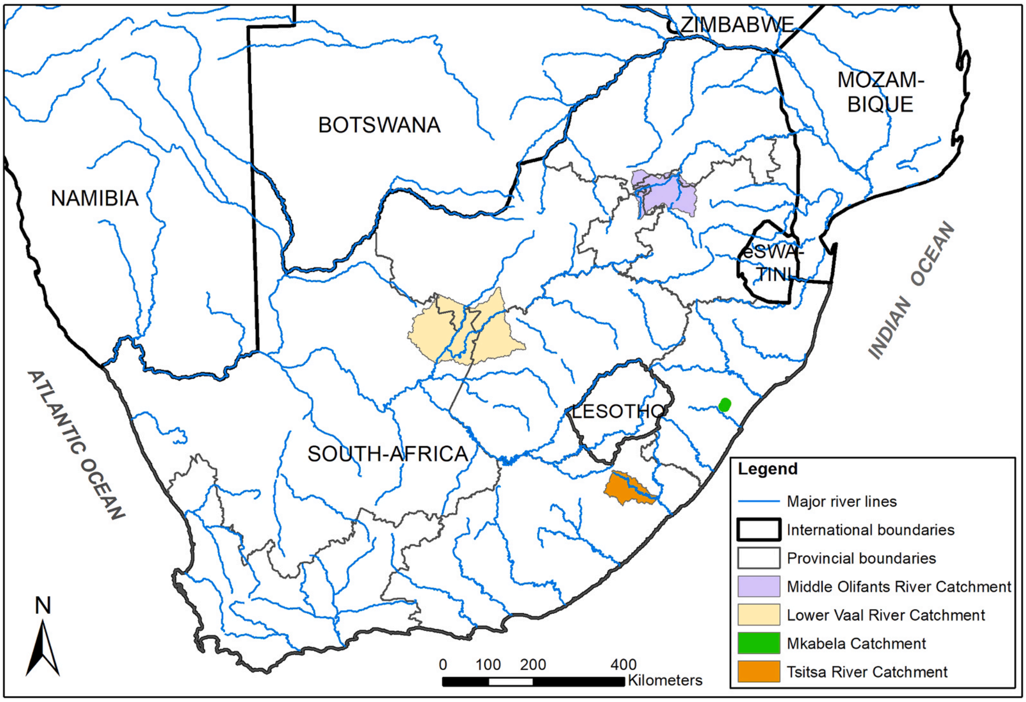

Fig. 1. Location map of South Africa including the locations of the four research catchments in which the SWAT model was applied using two (catchment and national) input datasets.

Table 2
Summary site descriptions of the four research catchments including Middle Olifants River Catchment (MORC), Lower Vaal River Catchment (LVRC), Mkabela Catchment (MC) and Tsitsa River Catchment (TRC).

Description	MORC	LVRC	MC	TRC
Location: Province and coordinates	Limpopo and Mpumalanga between 24°48′52″ - 25°48′23″ south and 28°20′41″ - 29°48′21″ east	North West and Free State between 27°7′18″ - 28°17′35″ south and 24°52′44″ - 25°45′41″ east	KwaZulu-Natal between 29º 21' 12'' - 29º 27' 16'' south and 30º 36' 20'' - 30º 41' 46'' east	Eastern Cape between 30º 46' 58'' - 31º 28' 55'' south and 27º 55' 56'' - 29º 13' 47'' east
Drainage area (km^2)	10 256	6 150	41	4 924
Topography, altitude range (m a.s.l.)	Flat to steep mountainous, 800–1 900	Flat with no distinct features, 1 183–1 431	Undulating to steep slopes, 880–1 057	Very steep mountain slopes to nearly level valley floors, 168–2 730
Climate	Semi-arid with mean annual rainfall of 620 mm	Semi-arid with mean annual rainfall of 400 mm	Sub-humid with mean annual rainfall of 825 mm	Sub-humid with mean annual rainfall of 625 mm
Vegetation type (Mucina and Rutherford, 2006); Main land use (SANLC, 2020)	Savannah; Irrigated and dryland crop farming, as well as livestock	Savannah; Irrigated and dryland crop farming, as well as livestock	Savannah; Sugarcane cultivation with minor forestry, pasture and vegetable plot	Grassland; Extensive grazing with areas of pine and gum plantations and maize cultivation
Geological data from the South African Council for Geoscience	Various formations of the Vaalian Group formed 2600–1800 MY dominated by Bushveld Igneous Complex including by granophyre and granite	Various formations of the Ventersdorp Supergroup of the late Archaean to early Proterozoic sequences formed 2400–3000 MY including mudrock, rhyolite, minor sandstone, biotite granite and network of dolerite intrusions	Formations of the Natal Group of the Cambrian Age formed 541–485 MY including sandstones and small pocket of Ecca sedimentary rocks	Succession of sedimentary layers of the Beaufort Group of the Triassic age formed 252–201 MY, including Adelaide mudrock and mudstones, overlain by sandstone and siltstone
Soil (ARC Land Type Database, 2012)	Varying soils including soils with a plinthic catena in mostly low-lying areas and shallow soils with minimal development on hard or weathering rock in high areas with steep slopes	Mostly freely drained sandy soils with low water holding capacity	Mostly shallow sandy soils on steep and convex hillslopes with little water holding capacity and deeper sandy soils with soft or hard plinthic sub-horizons that is permeable to water	Although soils in the catchment vary significantly, those from the mudstone parent material in the central part of the catchment are associated with duplex soils that are highly erodible with widespread erosion

catchments, including topographic, land use, soil, climate and hydrological data. First, topographic and drainage networks of the catchments were partitioned into sub-catchments that are comparative in size and representing all relevant river tributaries, as well as ensuring that flow monitoring points spatially overlay with sub-catchment outlet points for model validation. A total of 60, 27, 19 and 13 sub-catchments were delineated in the MORC, LVRC, MC and TRC, respectively.

Land use-cover data were derived from remote sensed data (SPOT or Landsat) acquired at different periods ranging between 2006 and 2018, creating between five and twelve land cover classes (per catchment). These land use-cover classes were linked to the land cover types in the SWAT database. In order to represent the variable soils in each catchment, textural and soil hydraulic parameter values were assigned to each soil component, namely Land Types (ARC Land Type Database, 2012). Pedotransfer functions similar to van Tol et al. (2010); (2020); (2021) were used to generate the required hydraulic parameters. The overlay of land cover and soil maps created 813, 625, 130 and 202 in the MORC, LVRC, MC and TRC, respectively.

SWAT also requires climate parameters, including precipitation, temperature, solar radiation, relative humidity and wind speed. Daily precipitation and temperature data were acquired from (two to four) meteorological stations of the ARC and/or South African Weather Service Data Portal (SAWS) over 30 years (38 years for LVRC). Since not all the stations have full records of the required parameters, incomplete precipitation and temperature records were patched with the most complete and closest stations within or near (<2 km) the catchment boundaries.

Hydrological parameters included flow contributions from point or inlet sources in catchments and accounted for diversion of flow for irrigation where applicable. In addition, reservoirs were defined in sub-catchments where needed, ranging from small farm dams (approximately 1 ha) to large (>5 000 ha) reservoirs. The MC also incorporated five wetlands (5–22 ha) to receive loadings from the sub-catchments where they are located (le Roux et al., 2013). The Penman-Monteith equations were used to calculate potential (and actual) evapotranspiration for each catchment, considering soil moisture and crop development (Aouissi et al., 2016).

Management practices include tillage, nutrient applications, irrigation schedules and harvest. These practices affect the water balance and sediment and nutrient load generation through the impacts of the plant growth cycle on evapotranspiration. Due to the lack of data on management practices, however, assumptions had to be made in order to provide appropriate input to the catchment models (see le Roux et al., 2023). For example, in the MORC, it was assumed that 150 kg N ha^{-1} and 40 kg P ha^{-1} are applied for an average maize yield of 10 t ha^{-1} under irrigation. Sources of irrigation water were specified where needed (in MORC, LVRC and MC), including reservoir inflow and outflow data supplied by the Department of Water and Sanitation - Integrated Regulatory Information System (http://ws.dwa.gov.za/IRIS).

Finally, model simulations were conducted between 3 and 38 years, preceded by a 1–3 year warm-up period to get the hydrological

Table 3
Summary of topographic, land use-cover, soil, climate, and hydrological input data used to parameterize the research catchments including Middle Olifants River Catchment (MORC), Lower Vaal River Catchment (LVRC), Mkabela Catchment (MC) and Tsitsa River Catchment (TRC).

Input data	MORC	LVRC	MC	TRC
	Topographic and drainage network data			
DEM;	ASTER-GDEM 15	SRTM DEM (NASA LP DAAC)	Unpublished contour-derived DEM (GISCOE)	SRTM DEM (NASA)
GSD[a] (m)		30	20	90
	Land use-cover, soil and slope combinations			
Land use-cover data	SANLC (2014)	Landsat 8 (2018)	SPOT 5 (2006)	SPOT 5 (2011)
GSD[a] (m)	30	30	10	10
Soil data, in addition to ARC Land Type Database (2012).	South African Atlas of Climatology and Agrohydrology (Schulze, 2007).	Harmonized World Soil Database (FAO/IIASA/ISRIC/ISS-CAS/JRC, 2012).	Pedological soil map with textural profile descriptions (Lorentz et al., 2012).	South African Atlas of Climatology and Agrohydrology (Schulze, 2007).
Usable scale	1:250,000	1:250,000	1:100,000	1:250,000
HRUs:	613	825	130	202
Slope class %	< 8, 8–30, > 30	Single slope class	0–5, > 5	0–5, > 5
Thresholds[b]%	LU 10, S 10, Sl 10	LU 10, S 10, Sl 10	LU 5, S 10, Sl 20	LU 5, S 10, Sl 20
	Climate data			
Data source	ARC SAWS	ARC SAWS CFSR	ARC	ARC
Number of stations	4	3	1	2
Timeframe	1984–2015	1980–2018	1977–2008	1978–2007
	Hydrological parameters			
Reservoirs (#)	3	1	9	0
Area (ha)	178 – 1580	2 129	1 – 10.3	0
Volume (million m^3)	28–362	5 068	0.012 – 0.405	0
Simulation period (years)	28	38	3	6
(Warm-up)	(2)	(3)	(1)	(1)
Calibration	1987–2001	1980–2006	2006	2007
Validation	2001–2015	2007–2018	2007–2008	2008–2012

[a] . GSD: ground sample distance; HRU thresholds
[b] for LU: land use, S: soil, Sl: slope.

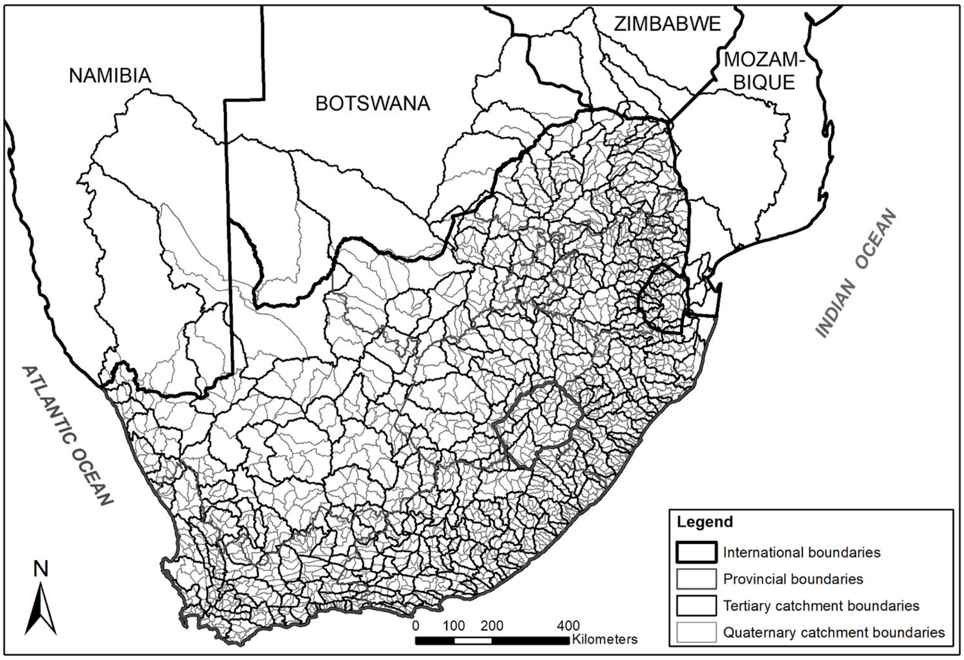

Fig. 2. Tertiary and quaternary catchment boundaries of South Africa.

cycle fully operational. Model performances were determined by means of the coefficient of efficiency (NSE) of Nash and Sutcliffe (1970), as well as the coefficient of determination (r^2). A per cent deviation method (*Dv*) of Martinec and Rango (1989) was used as a measure of goodness-of-fit between simulated and measured streamflow data at the main catchment outlets. A temporal split-sample approach was used to split the observation data into two periods for calibration and validation (see Table 3).

3. Results and discussion

The national baseline datasets are presented as a series of maps that are stored on a data portal system, followed by comparison of the results and accuracies of the two input datasets (catchment and national input) of the four research catchments.

3.1. Online data portal system for South Africa

The national input datasets to run the SWAT model are stored in the WRO data portal: https://www.waterresearchobservatory.org/data-and-resources/hydrological-data-and-modelling, as well as on Mendeley Data: https://data.mendeley.com/datasets/5nzr53w6jp/2. The portal provides:

- SWAT catchment outline data (tertiary and quaternary) including the hydrologically corrected SRTM DEM of SA at 90 m resolution (Weepener et al., 2012). Fig. 2 illustrates the tertiary and quaternary catchment boundaries of SA;
- South African National Land Cover (SANLC, 2014; 2018; 2020) including the user lookup table that links the grids with the SWAT land cover codes. Fig. 3 illustrates the SWAT land cover map of SA derived from SANLC (2020) with 72 land cover classes at 20 m resolution;
- Soil map and user lookup table including SWAT attribute data for each Land Type of SA (ARC Land Type Database, 2012). Fig. 4 illustrates the Land Types of SA usable at a scale of 1:250,000;
- Weather statistics (WGN) files required as input by the model, including two sets of (12) weather statistics (WGN) covering the periods 2001–2020 and 1981–2000, respectively. Fig. 5 illustrates the locations of the 12 ARC weather stations used to create the WGN files, superimposed over the rainfall erosivity factor (R) map of le Roux et al. (2008). SWAT users should select the WGN file closest to a respective study site (catchment).

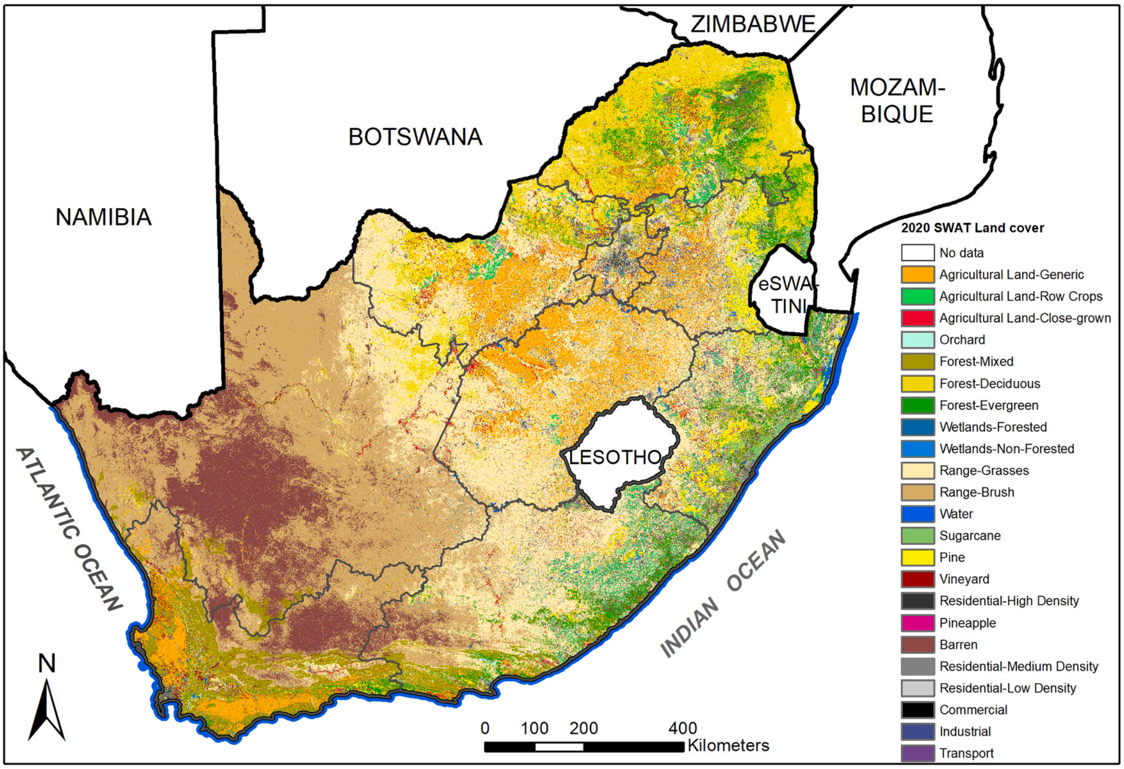

Fig. 3. SWAT land cover map of SA derived from SANLC (2020) with 72 land cover classes at 20 m resolution.

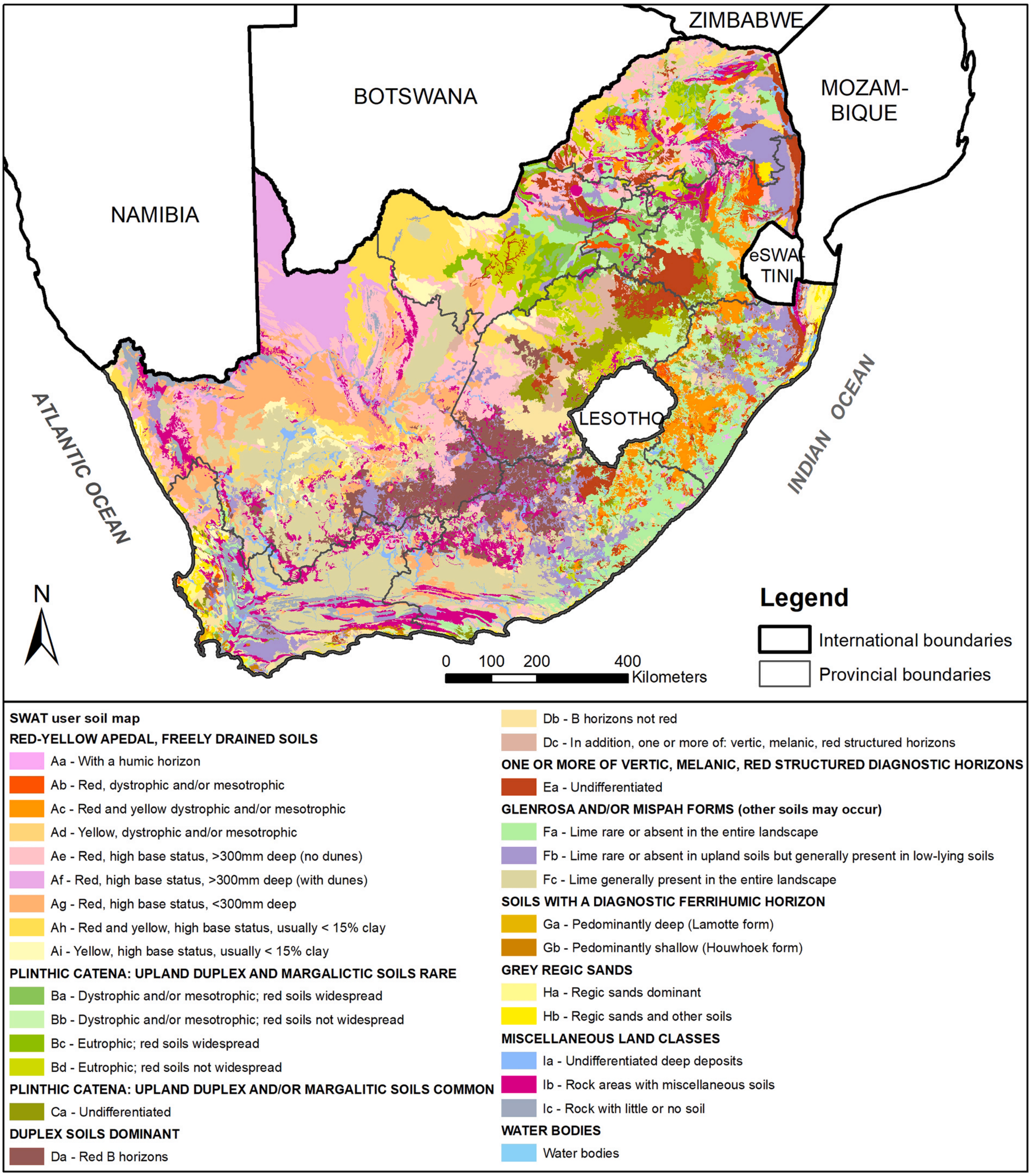

Fig. 4. Land Types of SA usable at a scale of 1:250,000 (ARC Land Type Database, 2012).

3.2. *Comparison between catchment and national data model results*

Simulations with original catchment input data are referred to as 'catchment data model(s)', whereas ArcSWAT simulations with the national baseline data are referred to as 'national data model(s)'.

3.2.1. *Streamflow results*

Graphical comparisons of observed versus simulated mean monthly streamflow for simulations with catchment and national data are presented in Fig. 6. Graphically, streamflow simulations with catchment and national data appear similar in each catchment, with a

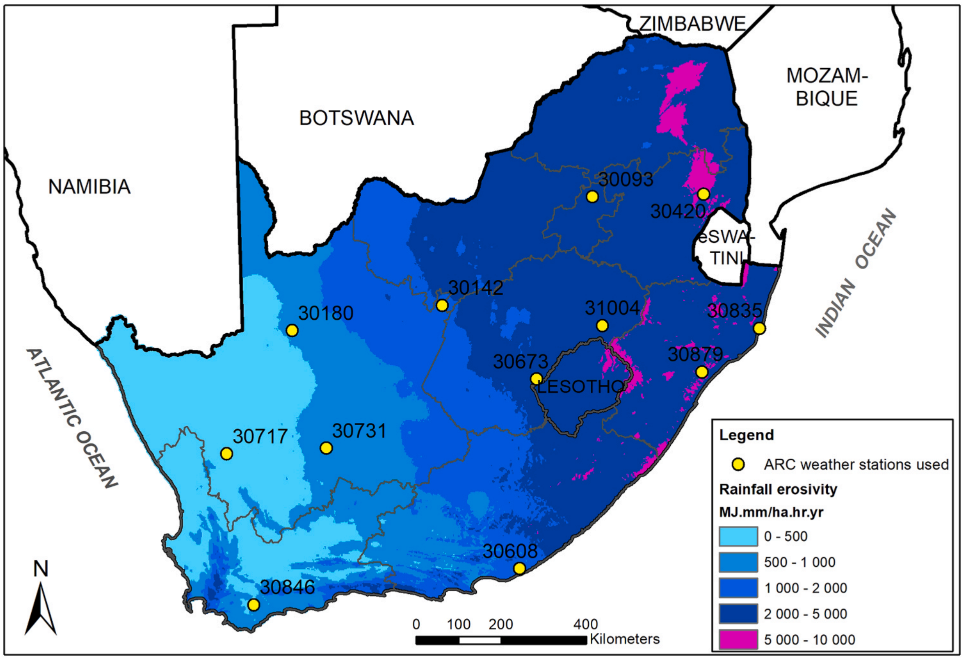

Fig. 5. Locations of the 12 ARC weather stations used to create the WGN files, superimposed over the rainfall erosivity factor map of le Roux et al. (2008).

good level of agreement between observed and simulated mean monthly streamflow. The catchment data models were slightly superior compared to the national data model during validation, as shown by the higher NSE, r^2 and *Dv* values (see Table 4). On average, the catchment and national data models over-predicted streamflow by 20 % and 30 % as determined by *Dv*, the goodness of fit expressed by NSE was 60 % and 55 %, and r^2 was 77 % and 73 % (respectively).

Between the four catchments, the LVRC performed the best, followed by the MC and then TRC, whereas the MORC performed the poorest (both MORC data models over-predicted streamflow by 47 %). Model performance depends largely on the detail or quality of input data (Kiros et al., 2015; Tan et al., 2020). Similar to other studies (Mararakanye et al., 2020; Schuol and Abbaspour, 2007; Tan and Yang, 2020), differences in the performance between catchments was largely influenced by the quality of rainfall data. For example, the MORC utilized four weather stations located in the east of the catchment. The use of four weather stations in such a large catchment (10 256 km^2) possibly caused errors in daily rainfall in other parts of the catchment and subsequent over-predicted streamflow output. Although the MC used only one weather station, the MC is the smallest catchment (41 km^2) of the four, with less probability of uneven rainfall distribution within the catchment. Another reason for the slightly superior performance of the catchment data models is due to differences between land use-cover datasets (further discussed in the Sediment yield section below).

3.2.2. Sediment yield results

To spatially illustrate sediment source areas, the average annual sediment yield for each catchment, for both the catchment and national data models, is shown in Fig. 7. Although the average sediment yield of the catchment and national data models are similar, the catchment and national data models identified different sediment source areas (except for the MORC). In the TRC for example, although the average sediment yield of the catchment and national data models are similar (0.85 and 0.72 t/ha/yr respectively), the national data model identifies the lower half of the TRC as important sediment source areas (>2 t/ha/yr), whereas the catchment data have moderate sediment yield values (1.0–2.0 t/ha/yr) throughout the catchment. The spatial differences in sediment yield between the catchment and national data models are possibly attributed to land use-cover variances since the topography, soils and climate parameters in both data models are similar.

Spatial differences in sediment yield between catchment and national data models are mainly attributed to land use-cover variances since the other input datasets (DEMs and soil input data) are in essence similar between data models. Each case study utilized DEMs with similar spatial resolutions (15–30 m). Furthermore, in each case study the soil input data for catchment and national data models were mostly derived from the ARC Land Type Database (2012) of SA. Similar methodology was followed in assigning of the required parameter values to Land Types (see Tables 2 and 3). The methodology and reasoning of the national baseline data were based on similar pedotransfer functions used to assign soil parameter values to the Land Types in each of the four catchments. Land use-cover classes, however, were obtained from different sources with different acquisition dates. In the TRC for example, land use-cover for the catchment data model was created by means of unsupervised classification on SPOT 5 imagery acquired in 2011, whereas for the national data model, the land use-cover was obtained from the 2018 land cover dataset (which was created from Sentinel 2 satellite imagery). Besides the difference in dates acquired, the spatial and temporal scales of the imagery used to create land use-cover maps

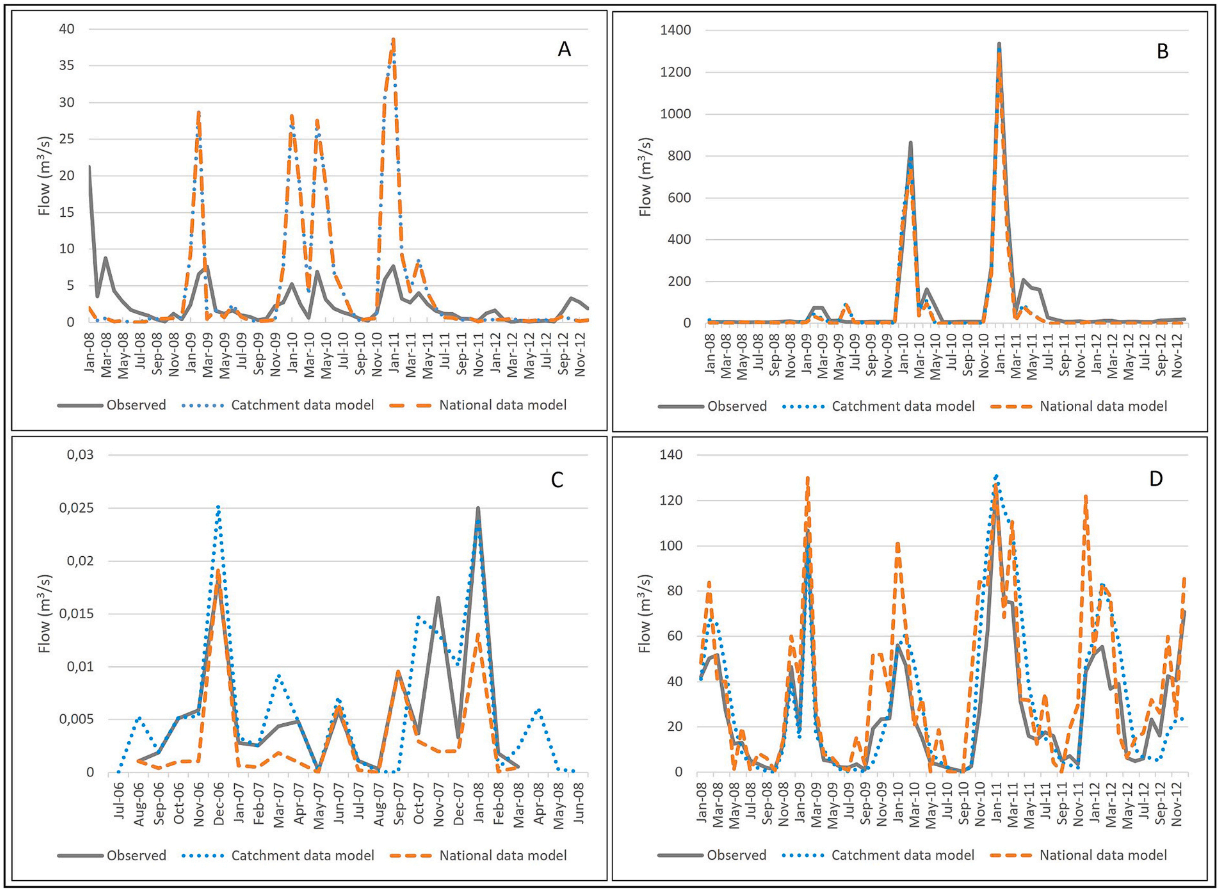

Fig. 6. Graphical comparison of observed monthly streamflow (in m^3/s) with the catchment and national data models in the (A) MORC (2008–2012), (B) LVRC (2008–2012), (C) MC (2006–2008) and (D) TRC (2008–2012).

Table 4
Performance metrics (r2, NSE and Dv in %) obtained from monthly streamflow validation for catchment and national data models including the MORC, LVRC, MC and TRC, as well as the mean of each data model.

	r^2 accuracy (%)		NSE accuracy (%)		*Dv* over-prediction (%)	
	Catchment	National	Catchment	National	Catchment	National
MORC	40	40	10	10	47	47
LVRC	97	96	97	97	17	25
MC	82	71	60	47	1.6	19
TRC	88	87	75	65	14	29
Mean	**77**	**73**	**60**	**55**	**20**	**30**

are also different. SPOT 5 multispectral imagery acquired in 2011 has a spatial resolution of 10 m and an annual (single season) temporal resolution, whereas the 2018 land cover dataset was derived from Sentinel 2 satellite imagery with 20 m spatial resolution and ten days (multi-seasonal) temporal resolution. Subsequently, in the national data model of the TRC, more barren land occurs in the lower half of the catchment than in the catchment data model. Barren land is associated with no vegetation cover, which accounts for relatively high sediment yield in these sub-catchments (le Roux, 2018). Despite these variances, the national baseline data appears to be an efficient input dataset capable of modelling streamflow and sediment dynamics at a catchment scale.

3.3. Comparison with other studies/datasets

For land cover, the current study used three National Land Cover maps (SANLC, 2014; 2018; 2020) with 72–73 land cover classes at 20 m resolution, subsequently corresponding to 27 land cover classes in the SWAT database. Abbaspour et al. (2019) used two global land cover datasets including the Global Land Cover Characterization (GLCC) acquired in 1993 with 24 land cover classes at 1 km resolution, and Global Landuse (GlobCover acquired in 2006 with 23 land cover classes at 300 m resolution. For soil, the current study

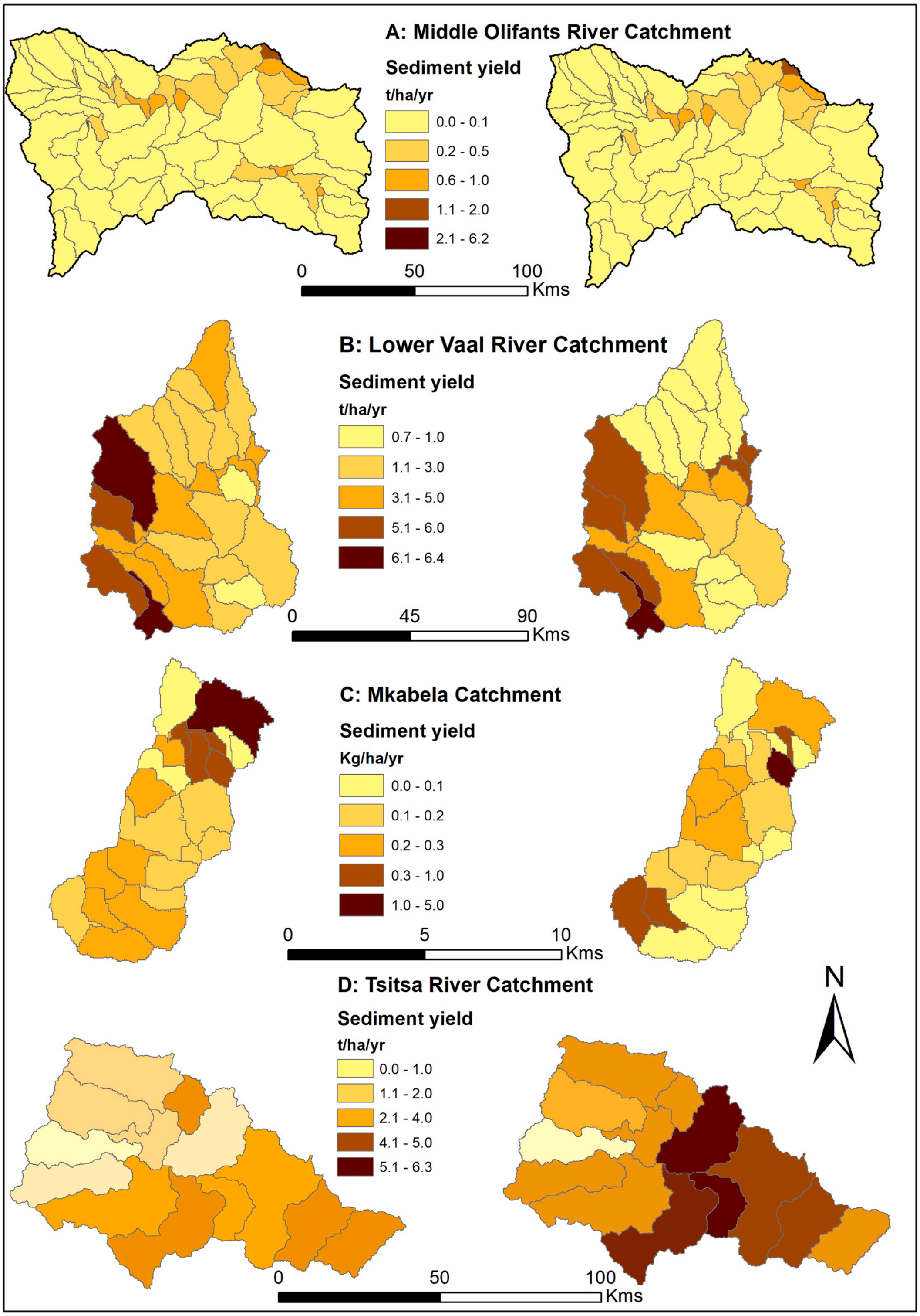

Fig. 7. Spatial comparison of average annual sediment yield (in t/ha/yr) simulated by the SWAT model with catchment data (on the left) and national data (on the right) in the: (A) MORC (1989–2015), (B) LVRC (1980–2018), (C) MC (2006–2008) and (D) TRC (2018–2012).

used the ARC Land Type Database (2012) of SA, a national database with 16,557 soil records usable at a scale of 1:250,000. Abbaspour et al. (2019) formatted two global soil maps including the FAO-UNESCO (2003) Soil Map of the World with only 4931 soil records at 1:5000,000 scale, and the Harmonized World Soil Database version 1.21 with 16,328 soil records at approximately 1 km (30 arc seconds) resolution (FAO/IIASA/ISRIC/ISS-CAS/JRC, 2012). Each soil database mentioned above provides a limited description of hydraulic parameters and, therefore, requires pedotransfer functions to generate essential parameters such as hydraulic conductivity, available water capacity and bulk density (see Abbaspour et al., 2019; le Roux et al., 2023). However, the pedotransfer functions of Abbaspour et al. (2019) are based on soils from around the globe, providing parameters that are universally applicable. It is therefore postulated that the national soil dataset provides a better representation of soil characteristics in SA than the global soil dataset of Abbaspour et al. (2019). Similarly, it is postulated that the national soil dataset also provides a better representation of soil characteristics in SA than the recently produced global DSOLMap of López-Ballesteros et al. (2023). The DSOLMap is a global soil map developed for SWAT+ at 250 m resolution. Soil properties for SWAT+ were extracted from soil data hosted by the OpenGeoHub foundation including soil texture and coarse fractions, organic carbon content, bulk density and soil water content at field capacity and at wilting point at six depths (0, 100, 300, 600, 1000, and 2000 mm) (Hengl et al., 2019). Other soil properties required by SWAT were derived by an ensemble of pedotransfer functions established by Abbaspour et al. (2019) and Ross et al. (2018), including USDA hydrologic soil group, hydraulic conductivity, soil erodibility factor and moist soil albedo. Soil mapping units were generated based on USDA textural classes for each of the six soil horizons (López-Ballesteros et al., 2023). Although the DSOLMap data includes six horizons with a spatial resolution of 250 m, the information used in the analysis is derived from the interpretation of small-scale maps based on machine learning predictions from limited soil profile observations. Subsequently, it is postulated that DSOLMap units provide only a broad and aggregate picture of soil conditions, which disguises more local variation in SA. Furthermore, hydrological evaluation of the DSOLMap was conducted in a small, forested catchment (47 km^2) in the north of Spain, consisting of mainly one dominant soil (Mollisols) and land use (forestry). The DSOLMap still needs to be evaluated in SA and other parts of the world.

4. Conclusion and recommendations

One of the biggest challenges to set-up and run the SWAT model in different parts of the world is to obtain appropriate input data, especially soil data (Akoko et al., 2021). This study addressed this challenge by providing appropriate soil, land cover and weather generator file datasets for use in SWAT applications in SA. The limitations inherent in the datasets include the following. The limitation of linking the SANLC maps to land cover types in the SWAT database is subsequent use of default SWAT parameters outside of conditions for which it was developed. The main limitation of using geospatial data as the basis for soil modelling inputs is that these soil units is not a soil polygon but rather an area characterized by homogeneous soil distribution patterns. Within a Land Type, for example, vastly different soils can occur. In this study, we used weighted average values for each soil unit to obtain soil parameters. The last important limitation worth mentioning is the national baseline datasets exclude climate data (besides weather generator files), particularly rainfall data, which is necessary to consider the spatial distribution of rainfall throughout a catchment. Despite these limitations, the input datasets presented here consist of more detailed and higher resolution data than the global datasets of Abbaspour et al. (2019) and provide a better representation of soil characteristics in SA than the global DSOLMap of López-Ballesteros et al. (2023). The national baseline data is an important step forward in hydrological modelling in SA by assisting modellers to set-up and run the SWAT model. Not only will the datasets save time with model set-up, but they will also assist in the standardization of SWAT modelling efforts in SA.

Although the input data could be used 'as is', it is recommended to supplement, improve and/or replace the input data with recent data. Since climate data are not freely available in developing countries, the best alternative is to use satellite-derived gridded weather data in empirical relationships (Mararakanye et al., 2020). Another recommendation is to improve the soil input dataset by means of digital soil modelling (DSM) techniques, including the use of slope gradient and plan curvature of different terrain units (see e.g. van Tol et al., 2020; van Zijl et al., 2020). Further improvements can be achieved by means of a hydropedological approach in the simulation of soil water contents to obtain a more accurate representation of the dominant hydrological processes in catchments (e.g. Harrison et al., 2022; Smit et al., 2024; van Tol et al., 2021). It is also recommended to compare the performance of the national soil data with different global soil datasets (FAO-UNESCO Soil Map, the Harmonized World Soil Database, and the DSOLMap). Further performance evaluation is needed in other catchment areas under different conditions. Evaluation of other water balance components such as evapotranspiration is also recommended. The input datasets should be updated continually with new data, especially land use-cover data, and can be expanded to cross-bordering catchments and to other African countries.

CRediT authorship contribution statement

le Roux Jay: Writing – review & editing, Writing – original draft, Validation, Supervision, Project administration, Methodology, Funding acquisition, Formal analysis, Conceptualization. **Mararakanye Ndifelani:** Writing – original draft, Validation, Methodology, Formal analysis. **van der Laan Michael:** Supervision, Funding acquisition, Data curation, Conceptualization. **Mudaly Leushantha:** Writing – original draft, Validation, Methodology, Formal analysis. **Weepener Harold Louw:** Resources, Methodology, Formal analysis. **van Tol Johan:** Supervision, Methodology.

Declaration of Competing Interest

The authors declare that the submitted work is our own and that copyright has not been breached in seeking its publication. The author declares that the research material in the paper submitted to the Journal has neither been published elsewhere in full nor is being considered elsewhere for publication. Lastly, no generative AI was used.

Acknowledgements

Thanks to the Water Research Commission (WRC) for funding the research project, including project management by Mr W. Nomquphu and Mrs. P. Jaca. The study also benefited greatly from comments by the WRC Reference Group. Special thanks to Dr Stefanie Schutte at the University of KwaZulu-Natal for spatial Carbon data. The paper also benefited greatly from comments by reviewers on earlier versions.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.ejrh.2025.102387.

Data availability

I have shared the link where data is available
South African soil, land cover and weather generator file databases for SWAT (Mendeley Data)

References

Abbaspour, K.C., Vaghefi, S.A., Yang, H., Srinivasan, R., 2019. Global soil, land use, evapotranspiration, historical and future weather databases for SWAT Applications. Sci. Data - Nature 6, 1–11. https://doi.org/10.1038/s41597-019-0282-4.

Akoko, G., Le, T.H., Gomi, T., Kato, T., 2021. A Review of SWAT Model Application in Africa. Water 13 (1313), 1–20. https://doi.org/10.3390/w13091313.

Aloui, S., Mazzoni, A., Elomri, A., Aouissi, J., Boufekane, A., Zghibi, A., 2023. A review of Soil and Water Assessment Tool (SWAT) studies of Mediterranean catchments: applications, feasibility, and future directions. J. Environ. Manag 326 B (116799), 1–20. https://doi.org/10.1016/j.jenvman.2022.116799.

Aouissi, J., Benabdallah, S., Chabaâne, Z.L., Cudennec, C., 2016. Evaluation of potential evapotranspiration assessment methods for hydrological modelling with SWAT—application in data-scarce rural Tunisia. Agr. Water Manag. 174, 39–51. https://doi.org/10.1016/j.agwat.2016.03.004.

Arnold, J.G., Kiniry, J.R., Srinivasan, R., Williams, E.B., Haney, E.B., Neitsch, S.L., 2012. SWAT: input/output file documentation, version 2012. report no. TR-439.. Texas Water Resources Institute, Texas USA, p. 650 ⟨https://swat.tamu.edu/media/69296/swat-io-documentation-2012.pdf⟩. report no. TR-439.

Bieger, K., Arnold, J.G., Rathjens, H., White, M.J., Bosch, D.D., Allen, P.M., Srinivasan, R., 2017. Introduction to SWAT+ , a completely restructured version of the soil and water assessment tool. J. Am. Water Resour. Assoc. 53, 115–130. https://doi.org/10.1111/1752-1688.12482.

Chawanda, C.J., George, C., Thiery, W., van Griensven, A., Tech, J., Arnold, J., Srinivasan, R., 2020. User-friendly workflows for catchment modelling: Towards reproducible SWAT+ model studies. Environ. Model. Softw. 134 (104812), 1–10. https://doi.org/10.1016/j.envsoft.2020.104812.

Chen, E., Mackay, D.S., 2004. Effects of distribution-based parameter aggregation on a spatially distributed agricultural nonpoint source pollution model. J. Hydrol. 295, 211–224. https://doi.org/10.1016/j.jhydrol.2004.03.029.

Dabrowski, J., Oberholster, P., Dabrowski, J., Le Brasseur, J., Gieskes, J., 2013. Chemical characteristics and limnology of Loskop Dam on the Olifants River (South Africa), in light of recent fish and crocodile mortalities. Water SA 39, 675–686. https://doi.org/10.4314/wsa.v39i5.12.

DWS, 2022. Spatial Data and Application Portal. Department of Water and Sanitation, Pretoria, South Africa. Available from: ⟨https://gia.dws.gov.za/portal/home⟩.

Ebodé, V.B., Dzana, J.G., Onguéné, R., Dongué, S.B., Koffi, B., Riotte, J., Mahé, G., Braun, J.J., 2024. Availability of the current and future water resources in Central Africa, case of the large Sanaga catchment in Cameroon. J. Hydrol. Reg. Stud. 53 (101815), 1–23. https://doi.org/10.1016/j.ejrh.2024.101815.

Escamilla-Rivera, V., Cortina-Villar, S., Vaca, R., Golicher, D., Arellano-Monterrosas, J., Honey-Rosés, J., 2022. Effects of finer scale soil survey and land-use classification on SWAT hydrological modelling accuracy in data-poor study areas. J. Water Resour. Prot. 14, 100–125. https://doi.org/10.4236/jwarp.2022.142007.

Essenfelder, A.H., 2016. SWAT Weather Database: A Quick Guide. Version: v.0.16.06. DOI: 10.13140/RG.2.1.4329.1927.

FAO/IIASA/ISRIC/ISS-CAS/JRC, 2012. Harmonized world soil database (version 1.2).

FAO-UNESCO, 2003. The Digital Soil Map of the World, Version 3.6. Land and Water Development Division, Rome, Italy.

Francesconi, W., Srinivasan, R., Pérez-Miñana, E., Willcock, S.P., Quintero, M., 2016. Using the Soil and Water Assessment Tool (SWAT) to model ecosystem services: a systematic review. J. Hydrol. 535, 625–636. https://doi.org/10.1016/j.jhydrol.2016.01.034.

Gassman, P.W., Reyes, M.R., Green, C.H., Arnold, J.G., 2007. The Soil and Water Assessment Tool: historical development, applications, and future research directions. Trans. ASAE/ASABE 50 (4), 1211–1250. https://doi.org/10.13031/2013.23637.

Gassman, P.W., Sadeghi, A.M., Srinivasan, R., 2014. Applications of the SWAT Model Special Section: overview and Insights. J. Environ. Qual. 43, 1–8. https://doi.org/10.2134/jeq2013.11.0466.

Glenday, J., Tanner, J., Rebelo, A., Holden, P., Metho, P., Gwapedza, D., Gokool, S., Jumbi, F., 2024. Modelling uncertainty and reliability for water resource assessment in South Africa. WRC report 3149/1/24. Water Research Commission, Pretoria, South Africa. ISBN 978-0-6392-0627-1.

Guzha, A.C., Rufino, M.C., Okoth, S., Jacobs, S., Nóbrega, R.L.B., 2018. Impacts of land use and land cover change on surface runoff, discharge and low flows: evidence from East Africa. J. Hydrol. Reg. Stud. 15, 49–67. https://doi.org/10.1016/j.ejrh.2017.11.005.

Harrison, R.L., van Tol, J., Toucher, M.L., 2022. Using hydropedological characteristics to improve modelling accuracy in Afromontane catchments. J. Hydrol. Reg. Stud. 39 (100986), 1–15. https://doi.org/10.1016/j.ejrh.2021.100986.

Hengl, T., Collins, T.N., Wheeler, I., MacMillan, R.A., 2019. Everybody has a right to know what's happening with the planet: towards a global commons. Medium (Towards Data Sci.). Zenodo. https://doi.org/10.5281/zenodo.2611127.

Julich, S., Moorcroft, M.A., Feger, K.H., van Tol, J.J., 2022. The impact of overgrazing on water fluxes in a semi-arid watershed – the suitability of watershed scale modeling in a data scarce area. J. Hydrol. Reg. Stud. 43 (101178), 1–14. https://doi.org/10.1016/j.ejrh.2022.101178.

Kiros, G., Shetty, A., Nandagiri, L., 2015. Performance evaluation of SWAT model for land use and land cover changes in semi-arid climatic conditions: a review. Waste Water Treat. Anal. 6 (3), 1–7. https://doi.org/10.4172/2157-7587.1000216.

Koo, H., Chen, M., Jakeman, A.J., Zhang, F., 2020. A global sensitivity analysis approach for identifying critical sources of uncertainty in non-identifiable, spatially distributed environmental models: a holistic analysis applied to SWAT for input datasets and model parameters. Environ. Model. Softw. 127 (104676), 1–11. https://doi.org/10.1016/j.envsoft.2020.104676.

Land Type Database, A.R.C., 2012. Land Types of South Africa: Maps (69 sheets) and Memoirs (39 books). Agricultural Research Council -Soil, Climate and Water, Pretoria, South Africa.

le Roux, J.J., 2018. Sediment yield potential in South Africa's only large river Network without a dam: implications for water resource management. Land Degrad. Dev. 29 (3), 765–775. https://doi.org/10.1002/ldr.2753.

le Roux, J.J., Mararakanye, N., Mudaly, L., Weepener, H.L., van der Laan, M., 2023. Development of a South African national input database to run the SWAT model in a GIS. WRC report 3053/1/22. Water Research Commission, Pretoria, South Africa. ISBN 978-0-6392-0454-3.
le Roux, J.J., Morgenthal, T.L., Malherbe, J., Sumner, P.D., Pretorius, D.J., 2008. Water erosion prediction at a national scale for South Africa. Water SA 34 (3), 305–314. https://doi.org/10.4314/wsa.v34i3.180623.
le Roux, J.J., Sumner, P.D., Lorentz, S.A., Germishuyse, T., 2013. Connectivity aspects in sediment migration modelling using the soil and water assessment tool. Geosciences 3 (1), 1–12.
López-Ballesteros, A., Nielsen, A., Castellanos-Osorio, G., Trolle, D., Senent-Aparicio, J., 2023. DSOLMap, a novel high-resolution global digital soil property map for the SWAT + model: development and hydrological evaluation. Catena 231 (107339), 1–13. https://doi.org/10.1016/j.catena.2023.107339.
Lorentz, S.A., Kollongei, J., Snyman, N., Berry, S.R., Jackson, W., Ngaleka, K., Pretorius, J.J., Clark, D., Thornton-Dibb, S., le Roux, J.J., Germishuyse, T., Görgens, A.H.M., 2012. Modelling Nutrient and Sediment Dynamics at the Catchment Scale. WRC report 1516/3/12. Water Research Commission, Pretoria, South Africa. ISBN 978-1-4312-0242-0.
Mabuda, M.O., Shokob, C., Dube, T., Mazvimavi, D., 2024. An analysis of the effects of changes in land use and land cover on runoff in the Luvuvhu catchment, South Africa. Remote Sens. Appl. Soc. Environ. 33 (101144), 1–14. https://doi.org/10.1016/j.rsase.2024.101144.
Mararakanye, N., Le Roux, J.J., Franke, A.C., 2020. Using satellite-based weather data as input to SWAT in a data poor catchment. Phys. Chem. Earth 117 (102871), 1–13. https://doi.org/10.1016/j.pce.2020.102871.
Mararakanye, N., Le Roux, J.J., Franke, A.C., 2022. Long-term water quality assessments under changing land use in a large semi-arid catchment in South Africa. Sci. Total Environ. 818 (151670), 1–14. https://doi.org/10.1016/j.scitotenv.2021.151670.
Martinec, J., Rango, A., 1989. Merits of statistical criteria for the performance of hydrological models. Water Resour. Bull. 25, 421–432.
Mucina, L., Rutherford, M.C., 2006. The vegetation of South Africa, Lesotho and Swaziland. Strelitzia 19. South African National Biodiversity Institute, Pretoria, South Africa. ISBN-13: 978-1-919976-21-1.
Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models. J. Hydrol. 10, 282–290.
Nkwasa, A., Chawanda, C.J., Msigwa, A., Komakech, H.C., Verbeiren, B., van Griensven, A., 2020. How can we represent seasonal land use dynamics in SWAT and SWAT+ Models for African cultivated catchments? Water 12 (1541), 1–18. https://doi.org/10.3390/w12061541.
Opiyo, S.B., 2024. Modeling the impacts of land use and land cover changes on streamflow and sediment yield with SWAT model in Migori River watershed, Lake Victoria basin region, Kenya. Afr. Geogr. Rev. 2024, 1–22. https://doi.org/10.1080/19376812.2024.2324279.
Ross, C.W., Prihodko, L., Anchang, J., Kumar, S., Ji, W., Hanan, N.P., 2018. HYSOGs250m, global gridded hydrologic soil groups for curve-number-based runoff modeling, 180091 Sci. Data 5 (1), 1–9. https://doi.org/10.1038/sdata.2018.91.
SANLC, 2014. National Land Cover Data of South Africa. 1990 and 2013-14. Department of Environmental Affairs: Pretoria, South Africa. Available from ⟨https://www.environment.gov.za/projectsprogrammes/egis_landcover_datasets⟩.
SANLC, 2018. National Land Cover Data of South Africa 2018. Department of Environmental Affairs: Pretoria, South Africa. Available from ⟨https://www.environment.gov.za/projectsprogrammes/egis_landcover_datasets⟩.
SANLC, 2020. National Land Cover Data of South Africa 2020. Department of Environmental Affairs: Pretoria, South Africa. Available from ⟨https://www.environment.gov.za/projectsprogrammes/egis_landcover_datasets⟩.
Schaap, M.G., Leij, F.J., van Genuchten, M.T., 2001. ROSETTA: A computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions. J. Hydrol. 251 (3-4), 163–176. https://doi.org/10.1016/S0022-1694(01)00466-8.
Schulze, R.E., 2007. South African Atlas of Climatology and Agrohydrology, WRC Report 1489/1/06. Water Research Commission, Pretoria, South Africa.
Schulze, R.E., Angus, G.R., Guy, R.M., 1991. Making the most of soils information: A hydrological interpretation of southern African soil classifications and data bases. Proceedings, 5th South African National Hydrology Symposium, 3B-2-1 to 3B-2-12, Stellenbosch, NSI, Cape Town, South Africa.
Schulze, R.E., Horan, M.J.C., 2005. AUTOSOILS Revised. University of KwaZulu-Natal, School of Bioresources Engineering and Environmental Hydrology, Pietermaritzburg, South Africa.
Schulze, R.E., Hutson, J.L., Cass, A., 1985. Hydrological characteristics and properties of soils in southern Africa 2: Soil water retention models. Water SA 11, 129–136.
Schuol, J., Abbaspour, K.C., 2007. Using monthly weather statistics to generate daily data in a SWAT model application to West Africa. Ecol. Model. 201 (3–4), 301–311. https://doi.org/10.1016/j.ecolmodel.2006.09.028.
Scott-Shaw, B.C., Hill, T.R., Gillham, J.S., 2020. Calibration of a modelling approach for sediment yield in a wattle plantation, KwaZulu-Natal, South Africa. Water SA 46 (2), 171–181. https://doi.org/10.17159/wsa/2020.v46.i2.8232.
Smit, E., Zijl, G.M., Riddell, E., van Tol, J.J., 2024. Model calibration using hydropedological insights to improve the simulation of internal. Hydrol. Process. Using SWAT+ . Ecol. Model. 38, 1–16. https://doi.org/10.1002/hyp.15158.
Srinivasan, R., Ramanarayanan, T.S., Arnold, J.G., Bednarz, S.T., 1998. Large area hydrologic modeling and assessment part II. Model Appl. J. Am. Water Resour. Assoc. 34, 91–101. https://doi.org/10.1111/j.1752-1688.1998.tb05962.x.
Tan, M.L., Gassman, P.W., Srinivasan, R., Arnold, J.G., Yang, X., 2019. A review of SWAT studies in southeast asia: applications, challenges and future directions. Water 11 (914). https://doi.org/10.3390/w11050914.
Tan, M.L., Gassman, P.W., Yang, X., Haywood, J., 2020. A review of SWAT applications, performance and future needs for simulation of hydro-climatic extremes. Adv. Water Resour. 143 (103662), 1–15. https://doi.org/10.1016/j.advwatres.2020.103662.
Tan, M.L., Yang, X., 2020. Effect of rainfall station density, distribution and missing values on SWAT outputs in tropical region. J. Hydrol. 584 (124660), 1–10. https://doi.org/10.1016/j.jhydrol.2020.124660.
van der Laan, M., Viviers, C., Maseko, S., Schutte, C., Thomson, A., Khoboko, P., Silberbauer, M., le Roux, J.J., Mudaly, L., Weepener, H.L., Hoogenboom, G., Srinivasan, R., Clark, D., Kunz, R., 2024. Development of the Water Research Observatory and case studies on machine learning applications WRC Report 3121/1/23. Water Research Commission, Pretoria, South Africa. ISBN 978-0-6392-0593-9.
van Griensven, A., Ndomba, P., Yalew, S., Kilonzo, F., 2012. Critical review of SWAT applications in the upper Nile basin countries. Hydrol. Earth Syst. Sci. 16, 3371–3381. https://doi.org/10.5194/hess-16-3371-2012, 2012.
van Tol, J.J., Bieger, K., Arnold, J.G., 2021. A hydropedological approach to simulate streamflow and soil water contents with SWAT. Hydrol. Process 35 (6), 1–14. https://doi.org/10.1002/hyp.14242.
van Tol, J.J., le, Roux, P.A.L., Hensley, M., 2010. Soil indicators of hillslope hydrology in the Bedford catchments. S. Afr. J. Plant Soil 27 (3), 242–251. https://doi.org/10.1080/02571862.2010.10639993.
van Tol, J.J., le Roux, P.A.L., Lorentz, S.A., Hensley, M., 2013. Hydropedological classification of South African hillslopes. Vadose Zone J. 12 (4). https://doi.org/10.2136/vzj2013.01.0007.
van Tol, J.J., van Zijl, G.M., Julich, S., 2020. Importance of detailed soil information for hydrological modelling in an urbanized environment. Hydrol 7 (34), 1–15. https://doi.org/10.3390/hydrology7020034.
van Tol, J.J., van Zijl, G.M., 2020. Regional soil information for hydrological modelling. Water Wheel 19 (2), 43–45. ⟨https://hdl.handle.net/10520/EJC-1d51e3655a⟩.
van Tol, J.J., van Zijl, G.M., 2022. South Africa needs a hydrological soil map: a case study from the upper uMgeni catchments. Water SA 48 (4), 335–347. https://doi.org/10.17159/wsa/2022.v48.i4.3977.
van Zijl, G.M., van Tol, J.J., Riddell, E.S., 2016. Digital Soil Mapping for Hydrological Modelling. In: Zhang, G.L., Brus, D., Liu, F., Song, X.D., Lagacherie, P. (Eds.), Digital Soil Mapping Across Paradigms, Scales and Boundaries. Springer Environmental Science and Engineering, Singapore, pp. 115–129. https://doi.org/10.1007/978-981-10-0415-5_10.
van Zijl, G.M., Van Tol, J.J., Bouwer, D., Lorentz, S., le Roux, P.A.L., 2020. Combining historical remote sensing, digital soil mapping and hydrological modelling to produce solutions for infrastructure damage in Cosmo City, South Africa. Remote Sens. 12 (433), 1–18. https://doi.org/10.3390/rs12030433.

Weepener, H.L., van den Berg, H.M., Metz, M., Hamandawana, H., 2012. The development of a hydrologically improved Digital Elevation Model and derived products for South Africa based on the SRTM DEM. WRC report 1908/1/11. Water Research Commission, Pretoria, South Africa. ISBN 978-1-4312-0217-1.

Woyessa, Y.E., 2024. Sustainable Management of Water Resources in a Semi-arid River Basin Under Climate Change: A Case Study in South Africa. In: Li, Y., Chaudhuri, H., Corrêa Rotunno Filho., O., Guseva, N., Bux, F. (Eds.), BRICS Countries: Sustainable Water Resource Management and Pollution Control. Springer, Singapore, pp. 183–209. https://doi.org/10.1007/978-981-99-9581-3_8.

Zhang, P., Chen, S., Dai, Y., Sekadende, B., Kimirei, I.A., 2024. Integrated application of SWAT and L-THIA models for nonpoint source pollution assessment in data scarce regions. Water 16 (800), 1–16. https://doi.org/10.3390/w16060800.

Zhao, J., Zhang, N., Liu, Z., Zhang, Q., Shang, C., 2024. SWAT model applications: from hydrological processes to ecosystem services. Sci. Total Environ. 931 (172605), 1–13. https://doi.org/10.1016/j.scitotenv.2024.172605.