



Article

Impacts of Soil Information on Process-Based Hydrological Modelling in the Upper Goukou Catchment, South Africa

Edward Smit *  and Johan van Tol 

Department of Soil, Crop and Climate Sciences, University of the Free State, Bloemfontein 9300, South Africa; vantoljj@ufs.ac.za

* Correspondence: 2015037738@ufs4life.ac.za

Abstract: Although soils form an integral part of landscape hydrological processes, the importance of soil information in hydrological modelling is often neglected. This study investigated the impact of soil information on streamflow modelling accuracy and hydrological process representation. Two different levels of soil information were compared to long-term streamflow in the upper Goukou catchment (230 km²), South Africa, over a period of 23 years using the Soil Water Assessment Tool (SWAT+). The land-type soil map (LTSM) dataset was less detailed and derived from the best, readily available soil dataset for South Africa currently. The hydrological soil map (HSM) dataset was more detailed and was created using infield hydropedological soil observations combined with digital soil-mapping techniques. Monthly streamflow simulation was similar for both soil datasets, with Nash–Sutcliffe efficiency and Kling–Gupta efficiency values of 0.57 and 0.59 (HSM) and 0.56 and 0.60 (LTSM), respectively. It is, however, important to assess through which hydrological processes were these streamflow values generated as well as their spatial distribution within the catchment. Upon further assessment, the representation of hydrological processes within the catchment differed greatly between the two datasets, with the HSM more accurately representing the internal hydrological processes, as it was based on infield observations. It was concluded that hydropedological information could be of great value in effective catchment management strategies since it improves representation of internal catchment processes.

Keywords: hydrological processes; SWAT+ model; hydropedology; land type soil map; hydrological soil map



Citation: Smit, E.; van Tol, J. Impacts of Soil Information on Process-Based Hydrological Modelling in the Upper Goukou Catchment, South Africa. *Water* **2022**, *14*, 407. <https://doi.org/10.3390/w14030407>

Academic Editor: Renato Morbidelli

Received: 8 December 2021

Accepted: 24 January 2022

Published: 29 January 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Soils are a primary control mechanism in determining hydrological processes within a landscape by partitioning precipitation into the different components of the water balance. This is the result of the ability of soils to absorb, store, and transmit water at different spatial and temporal scales [1]. Hydrological processes in the soil determine the volume, variability, and residence times of water resources within a landscape, which in turn affects the functionality and diversity of ecosystems [2]. However, due to the logistical impracticality of measuring hydrological processes at landscape-scale, these processes remain most practically defined using hydrological models to simplify and represent real-world hydrological systems [3,4].

As the need for more sustainable water resources has grown, the need for more accurate hydrological models has followed [5,6], particularly in ungauged basins [7]. Soil information is an important input parameter in physically-based hydrological models [8,9], yet the required soil information is often not readily available [10]. Reasons include that existing soil maps were not primarily designed for hydrological modelling purposes [11], and detailed hydrological soil surveys are costly and time-consuming to conduct. Hence, soil information employed in hydrological modelling has remained relatively crude compared

to other spatial inputs, such as topography and land-use datasets, subsequently decreasing the overall accuracy of modelling outcomes [12].

Modern hydrological models also rely heavily on a calibration period, resulting in models that can accurately simulate a specific point observation. This has become known as equifinality [13,14], where different model configurations all lead to similar results. Although these models and corresponding modelling approaches might provide statistically accurate model simulations in terms of point observations, such as streamflow gauges, their value in simulating scenario changes beyond the current environmental conditions for which they were calibrated is questionable [15].

The combined advances in digital soil mapping [16] and hydopedology [17–19], a multidisciplinary field combining hydrology and soil science, has enabled the production of large-scale hydrological soil datasets for modelling purposes [20–25]. Several studies have indicated that more detailed soil information improves modelling accuracy [17,21,23,26–28]. For example, the Soil Land Inference Model (SoLIM) was used to create a more detailed soil map in a data-scarce catchment in Portugal. The more detailed soil map led to a 7% increase in prediction accuracy compared to legacy soil data, combined with a reduction in parameter uncertainty [24]. In South Africa, a study of urbanized catchments (640 km², 550 km², 54 km²) also determined that hydopedological soil information improved modelling accuracy for all catchment sizes compared to readily available soil information [29]. Other studies, however, found that more detailed soil information did not statistically improve hydrological modelling accuracy [30–33]. However, none of these studies used hydrological soil information specifically created for modelling purposes.

The value of hydrological soil information may also extend beyond the ability to accurately model long-term streamflow predictions. The argument is that hydrological soil information may serve as an effective “soft data” tool, especially in ungauged basins where streamflow data is absent, to better represent internal hydrological processes within a catchment [34]. The use of “soft data” in hydrological model configuration and calibration has gained traction in recent years [34–37], where soft data is defined as information that may not necessarily be measured directly but can be related to a hydrological process or processes.

Hence, the primary objective of this study was to evaluate the impact of detailed hydrological soil information on process-based hydrological modelling. This was achieved by: (i) statistically comparing two levels of soil information in predicting streamflow accuracy in a catchment using long-term streamflow data and (ii) comparing spatial distribution of hydrological processes between the two soil datasets to highlight the importance of accurate internal catchment process representation.

2. Materials and Methods

2.1. Study Area

The 230 km² upper Goukou catchment is located within the Western Cape province of South Africa. Elevation of the catchment area is between 92 and 1512 m.a.s.l. (Figure 1), and it forms part of the mountains of the Cape Fold Belt [38]. The predominant geological formations include quartzitic sandstone from the Table Mountain Group, shale and siltstone from the Bokkeveld Group, as well as conglomerate and sandstone from the Uitenhage Group [39]. The region experiences a bimodal rainfall pattern with peaks in spring and autumn. Annual rainfall is strongly related to topography and ranges from 1200 to 1400 mm on the highest mountain peaks, while southern slopes receive between 500 and 800 mm on average [40]. Daily maximum temperatures average at 22 °C in January (summer) and 16 °C in July (winter) [41]. Average daily minimum temperatures are approximately 15 °C in January and 7 °C in July [41]. Vegetation in the catchment primarily consists of Fynbos and Renosterveld as well as an alluvial wetland [42].

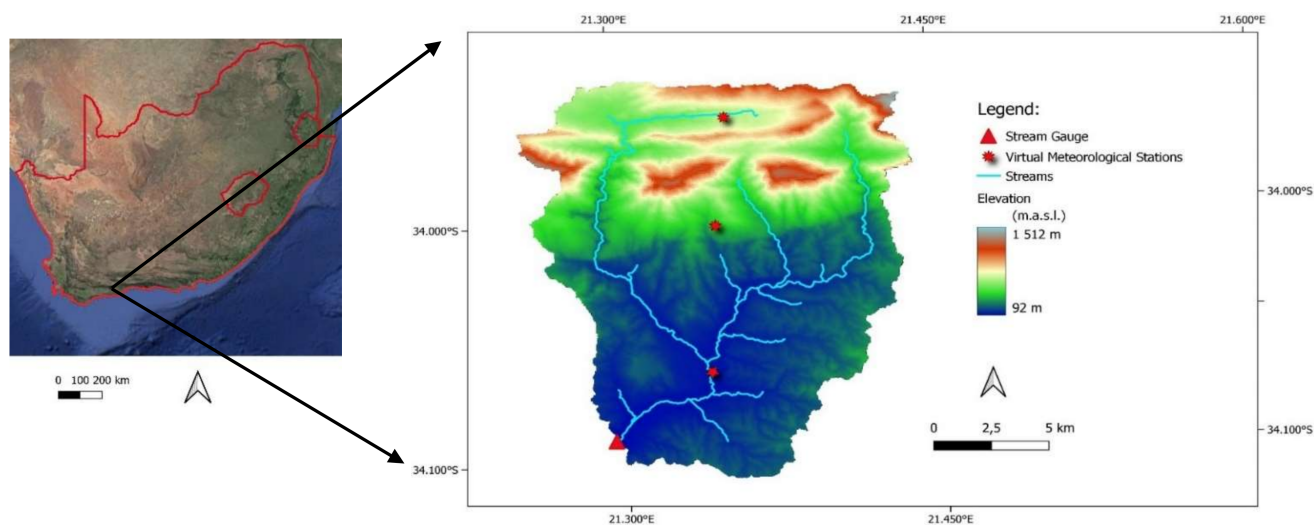


Figure 1. Map of the study area indicating streams, the stream gauge, and positions of virtual meteorological stations.

2.2. SWAT+ Setup, Topography, Land-Use, Streamflow, and Climate Inputs

SWAT+ (rev 60.5.2) download from <https://swat.tamu.edu/software/plus/> (accessed on the 15 April 2021) was used as hydrological model within QSWAT+ (v. 2.0.6) to set up the watershed. SWAT+ is a process-based, semi-distributed catchment-scale model and is a revised version of the well-known Soil and Water Assessment Tool (SWAT) [43]. The SWAT model is widely used to simulate water quality and quantity as well as assessing the impacts of physical changes to a catchment, such as land use and climate changes. SWAT+ divides the catchment into landscape units (LSUs), which comprise of a number of similar hydrological response units (HRUs). An HRU is a homogenous area in terms of soils, land use, and slope. The model then calculates various components of the water balance, such as infiltration, overland flow, lateral flow, percolation, evapotranspiration, as well as discharge to the stream for each HRU. A complete description of the SWAT model is given by Neitsch et al. [44], and for changes in the SWAT+ version, see Bieger et al. [45].

2.2.1. Topography and Land Use

Elevation of the catchment area (Figure 1) was obtained from a 30-m Shuttle Radar Topography Mission (SRTM) Digital Elevation Model (DEM) [46]. Current land use was obtained from the 2013–2014 SA National Land-Cover Map dataset [47]. Land cover was re-grouped according to SWAT land uses with pre-defined parameters for each use similar to real-world conditions (Figure 2). Table 1 provides the corresponding generic SWAT land use.

Table 1. 2013–2014 SA National Land Cover and the corresponding SWAT land-use name and codes.

2013–2014 SA National Land Cover	SWAT Land Use	SWAT Land-Use Code
Water	Water	WATR
Wetlands	Wetlands	WETL
Shrubland (Sparse)	Range-brush	RNGB
Shrubland (Dense)	Range-grasses	RNGE
Cultivated agriculture	Agricultural land generic	AGRL
Cultivated irrigation	Pasture	PAST
Commercial forestry	Forest-evergreen	FRSE

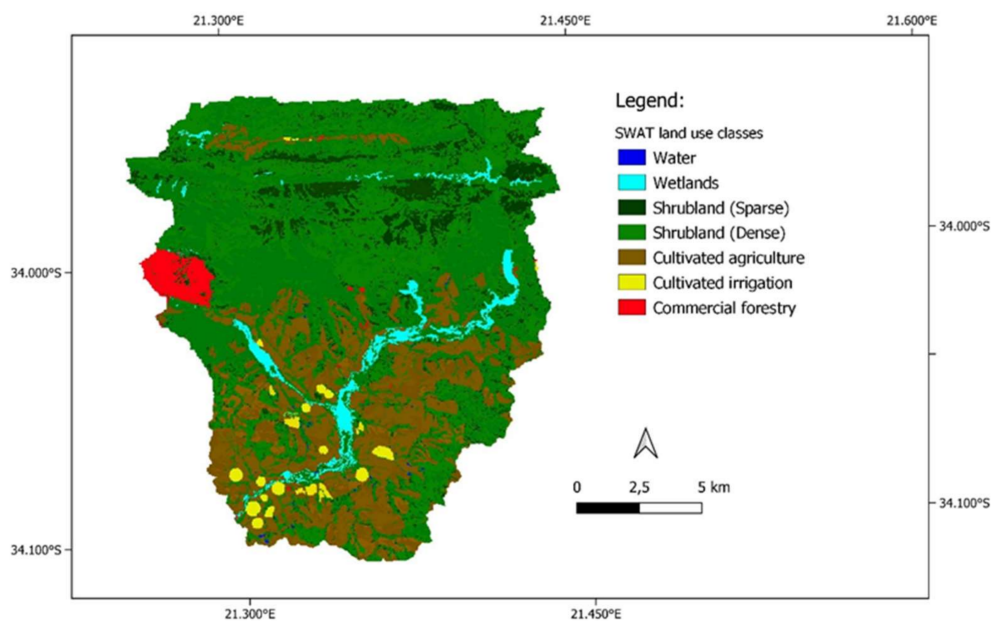


Figure 2. Land uses of the study area.

2.2.2. Streamflow and Climate Data

Daily rainfall as well as maximum and minimum temperatures were obtained from a nearby weather station of the South African Weather Service. Daily solar radiation, relative humidity, and wind speed were obtained from the Climate Forecast System Reanalysis project [48], run by the National Center for Environmental Prediction. In order to account for the orographic rainfall effect of the mountains, three virtual weather stations were created to imitate differential rainfall distribution over the catchment. To generate daily rainfall values for each of the virtual weather stations, rainfall coefficients were calculated by dividing the mean annual precipitation of each virtual station [40] by the mean annual precipitation of the actual weather station (1959–2013). These coefficients were then used to calculate daily rainfall values for each virtual station by multiplying daily measured rainfall from the actual weather station by the coefficient for each virtual station.

Long-term daily streamflow data was obtained for the Goukou River at the outflow gauging station (H9H005) of the catchment, managed by the Department of Water and Sanitation. The daily streamflow was aggregated to monthly discharge volumes, which were used to compare model simulations [49].

2.3. Soil Datasets

SWAT+ requires a soil dataset as an input spatial layer. Information on soil horizons, such as depth, particle size distribution, saturated hydraulic conductivity (K_{sat}), bulk density (BD), carbon content, and available water capacity (AWC), are all prerequisites for each soil layer. The soil information provided as input parameter affects the distribution of hydrological processes throughout the catchment. Two levels of soil information were used in this study.

2.3.1. Land Type Soil Map (LTSM)

The first soil dataset was derived from the South African Land Type Survey maps [50,51], developed between 1972 and 2002, representing the best available soil dataset covering the whole of South Africa at a scale of 1:250,000. A land type polygon is defined as “a homogeneous, unique combination of terrain type, soil pattern and macroclimate zone.” The Land Type Survey identified 7070 unique land type polygons based on some 400,000 soil observations (approximately 1 observation per 300 ha).

A total of 11 different land types are present within the catchment area. Except for K_{sat} , all the SWAT+ soil parameters for the LTSM dataset are available from Schulze [50] and are summarized in Table 2. The ROSETTA model [52] was used to derive K_{sat} for different horizons from the texture classes according to each land type.

Table 2. Land types and their respective hydropedological groups and main SWAT+ hydraulic properties of the LTSM dataset.

Land Type	Hydro-Group	Horizon	Depth (mm)	BD ¹ (g/cm ³)	AWC ² (mm/mm)	K_{sat} ³ (mm/h)	Clay (%)	Silt (%)	Sand (%)
Db12	A	A	200	1.46	0.09	200	10	5	85
		B	500	1.47	0.08	100	15	5	80
Db133	A	A	200	1.47	0.08	100	15	5	80
		B	350	1.49	0.08	50	20	5	75
Db14	B	A	290	1.46	0.09	200	11	5	84
Db21	B	B	360	1.4	0.09	100	10	5	85
Dc32	A	A	200	1.46	0.09	40	25	5	70
		B	330	1.47	0.09	20	30	5	65
Fa43	B	A	310	1.45	0.09	200	6	5	89
Gb21	B	A	200	1.45	0.09	200	11	5	84
		B	650	1.46	0.1	100	16	5	79
Gb22	B	A	200	1.45	0.09	200	9	5	86
		B	220	1.46	0.1	100	14	5	81
Ib168	A	A	180	1.48	0.08	200	5	5	90
Ic38	A	A	230	1.49	0.09	200	3	5	92
Ic49	A	A	150	1.47	0.08	200	5	5	90

Note: ¹ BD, bulk density; ² AWC, available water capacity; ³ K_{sat} , saturated hydraulic conductivity.

2.3.2. Hydrological Soil Map (HSM)

The second soil dataset was a hydrological soil map (HSM) dataset that was developed by combining an infield hydropedological survey and a digital soil-mapping tool (SoLIM) [53–55], based on hydrological mapping techniques as described by Van Zijl et al. [56].

Environmental covariates, including elevation, slope, profile curvature, and normalized vegetation index (NDVI), were obtained for the catchment area [46,51–53]. NDVI values were obtained for the catchment from the National Oceanic and Atmospheric Administration Climate Data Record NDVI database, where the 2013 average NDVI values for each pixel were calculated. These covariates were used to create a conditioned hypercube sampling method for the catchment area. For the hydropedological survey, a total of seven hillslope transects were selected to represent the soils of the catchment. Soil observations were made at crest, midslope, footslope, and valley bottom terrain positions or a combination of these terrain positions where applicable. A total of 50 soil observations were made by hand auger within the catchment, both on the hillslope transects and at easily accessible critical landscape areas. Legacy wetland soil data [57] were also used due to the inaccessibility of the wetland area. Soil morphology was classified in accordance with “Soil Classification: A Natural and Anthropogenic System for South Africa” [58]. The classified soil observations were then grouped by hydropedological type [59] for digital mapping purposes (Table 3).

Table 3. Hydropedological soil types of South Africa, their dominant properties, and corresponding WRB, IUSS reference groups.

South African Hydropedological Soil Type [59]	South African Soil Forms [58]	WRB/IUSS Reference Groups [60]	Defining Characteristics
Recharge (deep)	Carolina, Dundee, Valsrivier	Acrisols, Nitisols	Soils showing no signs of wetness; vertical drainage through and out of the soil is dominant.
Recharge (shallow)	Mispah, Glenrosa	Leptosols	Shallow soils with chromic colours in the topsoil; underlying bedrock is permeable and drainage out of profile dominant.
Interflow (A/B)	Estcourt, Katspruit, Idutywa	Stagnosols, Planosols, Plinthosols	Hydromorphic properties between top and subsoil signify periodic saturation. Typically duplex soils with textural discontinuity between top and subsoils, resulting in a perched water table at A/B horizon interface and interflow.
Interflow (soil/bedrock)	Sepane, Pinedene, Avalon	Acrisols, Stagnosols, Plinthosols	Hydromorphic properties at the soil/bedrock interface indicate saturation due to relatively impermeable bedrock. Perched water table at the bedrock interface will result in interflow.
Responsive (shallow)	Graskop, Glenrosa, Mispah	Leptosols	Shallow soils with bleached colours in the topsoil indicate that underlying bedrock is relatively impermeable. Small storage capacity will be exceeded following rain events and promote overland flow generation.
Responsive (saturated)	Champagne	Gleysols	Gleyed subsoils indicate long periods of saturation, typical of wetland soils. Soils will respond quickly to rain events and promote overland flow due to saturation excess.

The hillslope transects and soil observations were conceptualized and using SoLIM, the HSM was created by means of a rule-based design (Table 4). SoLIM was used to spatially map four identified hydropedological soil types, where SoLIM selects the soil mapping unit with the highest probability of occurrence for each pixel of the DEM grid. Based on infield investigations of soil distribution patterns, a 30-m buffer zone was created for streams above 350 m.a.s.l to represent responsive saturated soils, while the wetland was delineated using the national wetland map of South Africa [42]. The SoLIM rules used to create the HSM were designed to reflect field observations made in the catchment during the hydropedological soil survey. The HSM was validated using 50 soil observations throughout the catchment. The final HSM had an evaluation point accuracy of 74% and a Kappa statistic value of 0.68, indicating a substantial agreement with reality.

Table 4. The SoLIM rules used to create the HSM.

Soil Map Unit	Elevation (m)	Profile Curvature	NDVI *
Recharge (shallow)	>350 m	–	>0.22
Rock dominated	>350 m	–	<0.22
Interflow (A/B)	<350 m	<0	-
Interflow (soil/bedrock)	<350 m	>0	-

Note: * NDVI, Normalized Vegetation Index.

Undisturbed core samples were collected from diagnostic horizons. The undisturbed core samples were used to determine bulk density (BD) and particle size distribution. K_{sat} was determined using the ROSETTA model [52], while available water capacity (AWC) was obtained through pedotransfer functions [61]. Lastly, the hydraulic properties used as SWAT+ inputs for the different horizons of each hydrological soil type were obtained by averaging the infield measured values for each soil form in the specific hydrological soil type (Table 5).

Table 5. Soil types and their respective hydropedological groups and main SWAT+ hydraulic properties of the HSM dataset.

Soil Type	Hydro-Group	Horizon	Depth (mm)	BD ¹ (g/cm ³)	AWC ² (mm/mm)	K _{sat} ³ (mm/h)	Clay (%)	Silt (%)	Sand (%)
Interflow (A/B)	C	A	150	1.03	0.163	30	25	40	35
		B	800	1.2	0.199	10	33	49	18
		C	900	1.2	0.199	93	33	49	18
Interflow (soil/bedrock)	A	A	150	1.06	0.171	70	22	43	35
		B	1000	1.21	0.194	11	43	45	12
		C	1400	1.21	0.193	8	41	45	14
Recharge (shallow)	D	A	150	1.5	0.097	65	10	5	85
Responsive (saturated)	B	A	250	1.5	0.117	34	16	12	71
Rock dominated	D	A	150	1.9	0.01	65	10	5	85
Wetland	B	A	400	0.84	0.135	70	30	32	38
		B	4500	1.26	0.074	10	50	25	25

Note: ¹ BD, bulk density; ² AWC, available water capacity; ³ K_{sat}, saturated hydraulic conductivity.

2.3.3. Validation and Statistical Comparison of Simulations

Uncalibrated model simulations using the two soil datasets were conducted for the years 1987 to 2013. This included an initial settling period of four years and a 23-year simulation period. For statistical comparison, four widely used statistical indicators were employed, namely the coefficient of determination, percent bias (PBIAS), Nash–Sutcliffe efficiency (NSE), and Kling–Gupta efficiency (KGE).

Percent bias (PBIAS) measures the average tendency of the simulated data to be larger or smaller than their observed counterparts. The optimal value of PBIAS is 0.0, with low-magnitude values indicating accurate model simulation. Positive values indicate model underestimation bias, while negative values indicate model overestimation bias [62]. PBIAS is generally expressed in percentage and is calculated using Equation (1):

$$PBIAS = \frac{\sum_{i=1}^n V_{oi} - \sum_{i=1}^n V_{ei}}{\sum_{i=1}^n V_{oi}} \% \quad (1)$$

where V_{oi} and V_{ei} are, respectively, the observed and simulated volumes of water for day i .

Nash–Sutcliffe efficiency (NSE) is a normalized statistic that determines the relative magnitude of the residual variance (“noise”) compared to the measured data variance (“information”). NSE indicates how well the plot of observed versus simulated data fits the 1:1 line [63]. NSE is computed using Equation (2):

$$NSE = \left[\frac{\sum_{i=1}^n (V_{oi} - V_{ei})^2}{\sum_{i=1}^n (V_{oi} - \bar{V}_o)^2} \right] \quad (2)$$

where V_{oi} and V_{ei} are, respectively, observed and estimated; discharge of day i , \bar{V}_o , is the mean of the observed discharges. The optimum value is 1.0, with higher values indicating better model performance.

The Kling–Gupta efficiency (KGE) that incorporates correlation, variability bias, and mean bias [64] is increasingly used for model calibration and evaluation. It is expressed using Equation (3):

$$KGE = 1 - \sqrt{(r - 1)^2 + \left(\frac{\sigma_e}{\sigma_o} - 1\right)^2 + \left(\frac{\bar{Q}_e}{\bar{Q}_o} - 1\right)^2} \quad (3)$$

where r is the correlation coefficient between observed and simulated flows, and σ_o is the standard deviation of the observed streamflow, whereas σ_e represents the standard deviation of the simulated streamflow. \bar{Q}_e and \bar{Q}_o represent the mean of the simulated and observed discharges, respectively.

2.4. Data Used

The hydro–meteorological data that have been used for the SWAT+ model in this study are precipitation, minimum, and maximum air temperature. Table 6 contains all the input data required by the model.

Table 6. Input data used in the SWAT+ model and their sources.

Data	Data Source
Streamflow	Department of Water and Sanitation [49]
Precipitation and temperature	South African Atlas of Climatology and Agrohydrology [40,41,48]
Digital Elevation Model	Landsat Imagery [46]
Soil (Land-Type Soil Map)	Land-Type Soil Dataset [50–52]
Soil (Hydrological Soil Map)	Infield observations
Land use	2013–2014 South African National Land-Cover Dataset [47]

3. Results

3.1. Soil Datasets

The two soil datasets (HSM and LTSM) differed both spatially (Figure 3) as well as in their hydrological properties (Tables 1 and 4). The LTSM dataset contained a greater variety of soils (11 soil types vs. 6 in the HSM dataset) and hence also a wider array of hydrological soil properties based on pedotransfer functions. Although the HSM dataset had less variety in soil types, it was spatially more representative of the actual catchment soils. The HSM also consisted of accurate infield measured hydrological soil properties, more accurately representing the hydrological soil properties of the catchment.

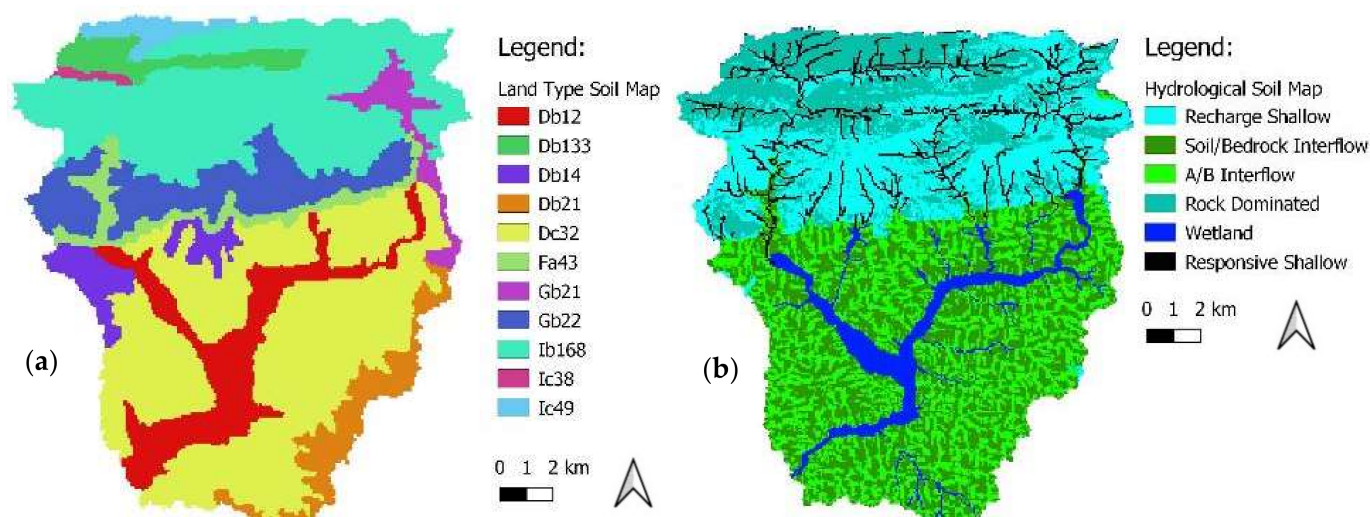


Figure 3. A spatial comparison of the two soil datasets, (a) LTSM and (b) HSM.

3.2. Model Streamflow Simulations

The two model set-ups each contained an identical number of sub–basins (26) and landscape units (2642) since both models used the same DEM dataset. The number of HRUs, however, differed greatly with the HSM dataset containing 4216 HRUs, while the LTSM dataset contained 2836 HRUs. This is the result of the increased level of spatial detail

in the HSM compared to the LTSM dataset as well as both models being run with all HRUs created, meaning that no HRU threshold was applied to either model.

Statistical comparison between the two models over the 23-year simulation period indicated that the two models simulated streamflow very similarly (Figure 4). Both model simulations yielded acceptable correlation values ($R^2 \geq 0.5$) and also produced highly similar NSE and KGE values of 0.57 and 0.59 (HSM) and 0.56 and 0.60 (LTSM), respectively. PBIAS indicated that both models substantially overestimated baseflow conditions within the catchment while underestimating during peak flow periods; however, the HSM performed better than the LTSM, with considerably lower estimation bias.

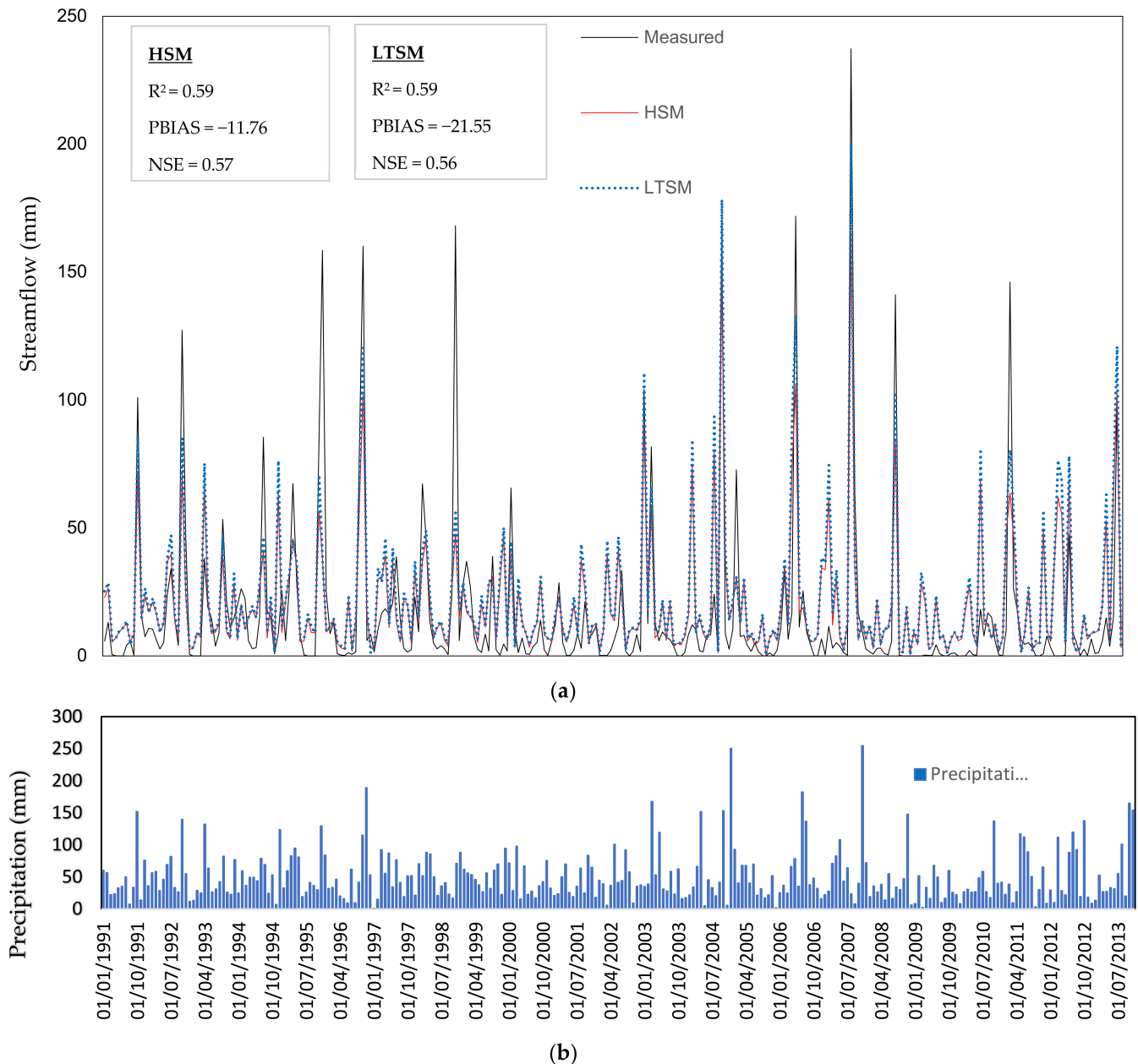


Figure 4. Comparison between the two soil datasets, (a) HSM (hydrological soil map) and LTSM (land-type soil map), in terms of monthly discharge simulations with SWAT for the period 1991 to 2013, together with (b) the monthly measured precipitation values.

Hence, despite substantial differences between the HSM and LTSM datasets, both spatially and regarding physical soil properties, these differences were not carried over into statistical differences for long-term streamflow simulations.

3.3. Hydrological Processes

When comparing internal catchment processes, the impact of soil information on hydrological modelling became evident. Based on its importance in accurate streamflow predictions, average annual lateral flow and overland flow were used as examples to illustrate the differences in hydrological process representation between the two soil datasets. The HSM resulted in a very clear pattern, showing how the bulk of lateral flow volumes originate in the mountainous regions of the catchment and diminishes towards the wetland (Figure 5). This coincides with the transition from sandy, shallow mountain soils to deeper, clayey interflow soils on the lower slopes, to extremely deep alluvial soils within the riparian zone. The LTSM, on the other hand, produced a haphazard spatial distribution of lateral flow generation processes with no regard for the delineated riparian wetland.

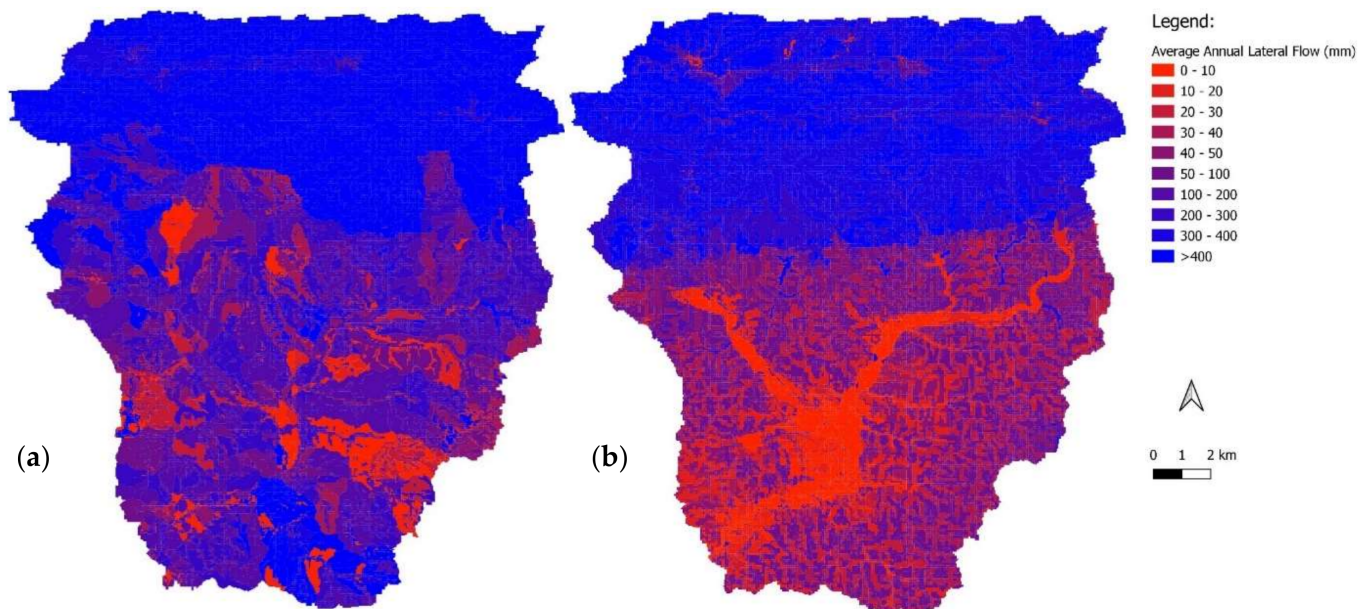


Figure 5. The average annual simulated lateral flow ($\text{mm}\cdot\text{year}^{-1}$) at HRU level for the (a) LTSM and (b) HSM.

When analysing overland flow or surface runoff generation processes, the differences between soil datasets became even more apparent (Figure 6). The HSM simulated the majority of surface runoff within the mountainous regions of the catchment, with comparatively small volumes of surface runoff (<10 mm) occurring on the southern footslopes. The LTSM dataset generated localized pockets with high runoff volumes, with the majority of these falling outside the mountainous area of the catchment.

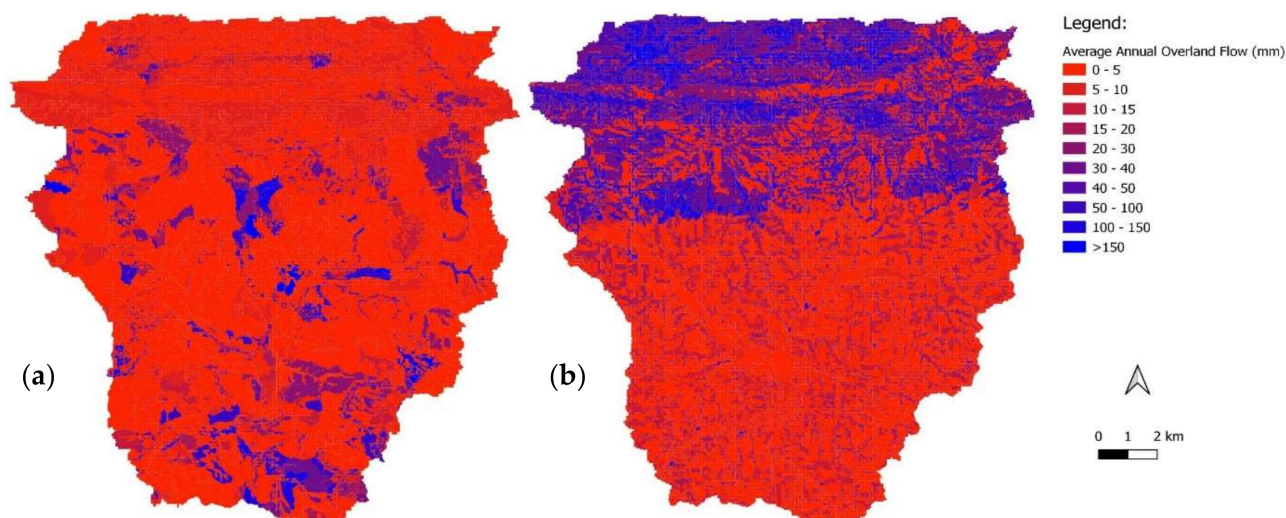


Figure 6. The average annual simulated overland flow ($\text{mm}\cdot\text{year}^{-1}$) at HRU level for the (a) LTSM and (b) HSM.

4. Discussion

4.1. Long-Term Streamflow Simulations

According to Knoben et al. [65], a KGE value greater than -0.41 implies that a model prediction is a better fit than the mean observed values. Therefore, it can be accepted that both uncalibrated model simulations in the current study produced satisfactory KGE values. Moriasi et al. [66] presented improved evaluation criteria for hydrologic and water-quality models. For streamflow simulations $R^2 > 0.6$, $\text{NSE} > 0.5$ and $\text{PBIAS} \leq 15\%$ were regarded as satisfactory. Although both models in the current study just fall short of the $R^2 > 0.6$ norm, both models produced acceptable NSE values above the 0.5 threshold. The HSM model also stayed within the 15% PBIAS limit (11.76%), resulting in the HSM slightly outperforming the LTSM model. Examples of similar hydrological modelling studies obtained R^2 values of anywhere between 0.15 and 0.74 [27–29].

The clear overestimation of baseflow and underestimation of peak flow for both model simulations resulted in an overall overestimation of streamflow values ($\text{PBIAS} = -11.76$ and -21.55). Such an overestimation could be corrected by adjusting groundwater parameters, e.g., by decreasing the threshold depth of the groundwater before re-evaporation occurs (REVAMPM), increasing the threshold depth for return flow to occur (GWQMN), or increasing the coefficient of re-evaporation from the groundwater (GW_REVAP) for selected HRUs. The most likely explanations for the observed overestimation of baseflow are (a) the large degree of uncertainty in terms of rainfall distribution over the catchment area, which could not be accurately quantified by means of virtual rainfall stations, and/or (b) the catchment location in the transition zone between winter and summer rainfall region, adding to the complexity of rainfall variability within the catchment.

Although the two models did perform remarkably similar, the HSM did slightly outperform the LTSM when considering PBIAS values. The more detailed soil information of the HSM provided some degree of improvement in simulation accuracy in the overall water balance. This may be attributed to more accurate representation of soil physical parameters, such as profile depths, hydraulic conductivity values, and available water capacities.

4.2. Importance of Internal Catchment Processes

Figures 5 and 6 represent the dominant hydrological processes (subsurface lateral flow and overland flow) of the upper Goukou catchment. These figures illustrate the large discrepancies between the two soil datasets when simulating different hydrological processes. These differences are a result of how hydrological processes are simulated by the

SWAT+ model. The HSM dataset, based on infield observations and measurements, more accurately represents actual hydrological processes of the catchment. The high validation accuracy of the HSM, as well as the reasonable modelling efficiency without calibration, emphasizes the accuracy of the HSM to accurately represent hydrological processes within the catchment.

There is a noticeable difference in spatial transition between higher and lower subsurface lateral flow areas between the two datasets. For the HSM data, there is a clear transition from high lateral flow volumes in the mountains to lower volumes on the footslopes and finally minimal lateral flow within the alluvial floodplain. Subsurface lateral flow is calculated by SWAT using a kinematic storage model, which simulates the movement of water in a two-dimensional cross-section of a hillslope. The kinematic approximation method assumes that the flow paths are parallel to the bedrock and that the hydraulic gradient equals the slope of the hill. Primarily, subsurface lateral flow is calculated as a product of the excess soil water volume after saturation, the hydraulic gradient of the HRU, as well as the drainable porosity of the soil. It is therefore expected that lateral flow simulations would decrease as the slope decreases from mountainous soils, footslope soils, and floodplain soils.

The LTSM simulated far higher lateral flow volumes for the catchment than the HSM, which could be attributed to the overall shallower nature of the LTSM soils compared to HSM soils, especially at the footslope and valley-bottom terrain positions. Shallower soils allow for less water to be stored within the soil profile, which means that more water percolates down to the soil/bedrock interface and is available for lateral flow. The distribution of lateral flow as simulated by the LTSM was also not primarily determined by soil information, as was the case with the HSM simulation, but rather by a combination of soil, land use, and topography for each HRU.

The mountains of the upper Goukou catchment form part of the areas identified as strategic water-source areas for surface water in South Africa [67]. However, the LTSM dataset simulated low volumes of surface runoff for this critical area of the catchment (Figure 6a). Differences between the soil datasets in terms of overland flow volumes in the mountainous regions of the catchment were mainly attributed to differences in hydraulic soil information, as both datasets contained similarly shallow mountain soils (<300 mm). The HSM dataset comprised of far lower K_{sat} values and higher clay percentages than the LTSM. The HSM data also accounted for rock-dominated soils that cover most of the mountainous areas of the catchment. These differences resulted in less permeable topsoils for the HSM, which led more runoff generation to occur. The SWAT model calculates surface runoff using the Soil Conservation Service curve number or the Green–Ampt infiltration method. Surface runoff only occurs when excess water is available at the soil surface. Topsoil hydraulic conductivity, available water capacity, and bulk density all impact the potential volume of surface runoff generated. Infield measurements of these soil properties are imperative to accurately simulating surface runoff.

The inability of the LTSM to accurately simulate this important water balance component in a strategic water source area illustrates why the accurate representation of internal catchment processes are imperative to water-management strategies. This also highlights the importance of soil information in ungauged basins. For example, because the HSM provides a more accurate representation of hydrological processes over the catchment, it might aid in improved planning of surface drainage infrastructure, improve rehabilitation and maintenance of wetland water regimes, as well as assist in spatially quantifying areas for optimal clearing of invasive plant species.

Internal catchment processes should be adequately reflected through the hydrological model structure [35,68]. It was clear that the more detailed HSM dataset produced for hydrological modelling using infield soil observations did add value to the hydrological modelling process. The true value of more detailed soil information may therefore lie in the more accurate representation of internal catchment processes and its effect on catchment management strategies. For instance, if the upper Goukou catchment had been ungauged,

the improved hydrological process representation of the HSM would enable more appropriate catchment development and management practices to transpire compared to the LTSM. This gives credence to the use of hydrogeological information as a “soft data” tool to better represent internal catchment processes, especially in ungauged basins [29].

5. Conclusions

This study presented the findings of uncalibrated hydrological simulations using two different levels of soil information in the upper Goukou catchment, South Africa. Detailed hydrogeological soil information, using infield soil measurements and digital soil mapping techniques, resulted in a more accurate representation of internal catchment hydrological processes.

Although the statistical improvement in long-term streamflow simulation were not as significant as expected, improved hydrological process representation does give credence to the importance of more detailed soil information for improved catchment management strategies. This is particularly true for ungauged basins, where long-term streamflow gauging stations are absent for calibration purposes.

Author Contributions: Conceptualization, J.v.T.; methodology, E.S. and J.v.T.; formal analysis, E.S. and J.v.T.; investigation, E.S.; writing—original draft preparation, E.S.; writing—review and editing, J.v.T.; visualization, E.S. and J.v.T.; funding acquisition, J.v.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by South African National Biodiversity Institute.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: Data sharing not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Park, S.J.; Mcsweeney, K.; Lowery, B. Identification of the spatial distribution of soils using a process-based terrain characterization. *Geoderma* **2001**, *103*, 249–272. [[CrossRef](#)]
2. Wenninger, J.; Uhkenbrook, S.; Lorentz, S.A.; Leibungut, C. Identification of runoff generation processes using combined hydrometric, tracer and geophysical methods in a headwater catchment in South Africa. *Hydrol. Sci. J.* **2008**, *53*, 65–80. [[CrossRef](#)]
3. McDonnell, J.J.; Woods, R.A. On the need for catchment classification. *J. Hydrol.* **2004**, *299*, 2–3. [[CrossRef](#)]
4. Devia, G.K.; Ganasri, B.P.; Dwarakish, G.S. A review on hydrological models. *Aquat. Procedia* **2015**, *4*, 1001–1007. [[CrossRef](#)]
5. Gassman, P.W.; Reyes, M.R.; Green, C.H.; Arnold, J.G. The soil and water assessment tool: Historical development, applications, and future research directions. *Trans. ASABE* **2007**, *50*, 1211–1250. [[CrossRef](#)]
6. Sawicz, K.; Wagener, T.; Sivapalan, M.; Troch, P.A.; Carrillo, G. Catchment classification: Empirical analysis of hydrologic similarity based on catchment function in the eastern USA. *Hydrol. Earth Syst. Sci.* **2011**, *15*, 2895–2911. [[CrossRef](#)]
7. Sivapalan, M. Prediction in ungauged basins: A grand challenge for theoretical hydrology. *Hydrol. Process.* **2003**, *17*, 3163–3170. [[CrossRef](#)]
8. Beven, K. Surface water hydrology—Runoff generation and basin structure. *Rev. Geophys.* **1983**, *21*, 721–730. [[CrossRef](#)]
9. Worqlul, A.W.; Ayana, E.K.; Yen, H.; Jeong, J.; MacAlister, C.; Taylor, R.; Gerik, T.J.; Steenhuis, T.S. Evaluating hydrologic responses to soil characteristics using SWAT model in a paired-watersheds in the Upper Blue Nile Basin. *Catena* **2018**, *163*, 332–341. [[CrossRef](#)]
10. Lilly, A.; Boorman, D.; Hollis, J.D. The development of a hydrological classification of UK soils and the inherent scale changes. *Nutr. Cycl. Agroecosyst.* **1998**, *50*, 299–302. [[CrossRef](#)]
11. Zhu, A.X.; Mackay, D.S. Effects of spatial detail of soil information on watershed modelling. *J. Hydrol.* **2001**, *248*, 54–77. [[CrossRef](#)]
12. Rinderer, M.; Siebert, J. Soil information in hydrologic models: Hard data, soft data, and the dialog between experimentalists and modelers. In *Hydrogeology: Synergistic Integration of Soil Science and Hydrology*; Lin, H.S., Ed.; Elsevier: Amsterdam, The Netherlands, 2012; pp. 515–536. [[CrossRef](#)]
13. Beven, K.; Freer, J. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *J. Hydrol.* **2001**, *249*, 11–29. [[CrossRef](#)]
14. Beven, K.A. Manifesto for the equifinality thesis. *J. Hydrol.* **2006**, *320*, 18–36. [[CrossRef](#)]
15. Kirchner, J.W. Getting the right answer for the right reason: Linking measurements, analysis, and models to advance the science of hydrology. *Water Resour. Res.* **2006**, *42*, W03S04. [[CrossRef](#)]
16. McBratney, A.B.; Mendonça Santos, M.L.; Minasny, B. On digital soil mapping. *Geoderma* **2003**, *117*, 3–52. [[CrossRef](#)]

17. Lin, H.S. Hydropedology: Bridging disciplines, scales, and data. *Vadose Zone J.* **2003**, *2*, 1–11. [[CrossRef](#)]
18. Lin, H.S.; Kogelman, W.; Walker, C.; Bruns, M.A. Soil moisture patterns in a forested catchment: A hydropedological perspective. *Geoderma* **2006**, *131*, 345–368. [[CrossRef](#)]
19. Lin, H.S.; McDonnell, J.J.; Nimmo, J.R.; Pachepsky, Y.A. Hydropedology: Synergistic intergration of soil science and hydrology in the Critical Zone. *Hydrol. Process.* **2015**, *29*, 4551–4561. [[CrossRef](#)]
20. Julich, S.; Breuer, L.; Frede, H.-G. Integrating heterogeneous landscape characteristics into watershed scale modelling. *Adv. Geosci.* **2012**, *31*, 31–38. [[CrossRef](#)]
21. Thompson, J.A.; Roecker, S.; Gunwald, S.; Owens, P.R. Digital soil mapping: Interactions with and applications for hydropedology. In *Hydropedology: Synergistic Integration of Soil Science and Hydrology*; Lin, H.S., Ed.; Elsevier: Amsterdam, The Netherlands, 2012; pp. 665–709. [[CrossRef](#)]
22. Van Tol, J.J.; van Zijl, G.M.; Riddell, E.S.; Fundisi, D. Application of hydropedological insights in hydrological modelling of the Stevenson Hamilton Research Supersite, Kruger National Park, South Africa. *Water SA* **2015**, *41*, 525–533. [[CrossRef](#)]
23. Van Zijl, G.M.; Van Tol, J.J.; Riddell, E.S. Digital mapping for hydrological modelling. In *Digital Soil Mapping Across Paradigms, Scales and Boundaries*; Zhang, G., Brus, D., Liu, F., Song, X., Lagacherie, P., Eds.; Springer: Singapore, 2016; pp. 115–129. [[CrossRef](#)]
24. Wahren, F.T.; Julich, S.; Nunes, J.P.; Gonzalez-Pelayo, O.; Hawtree, D.; Feger, K.H.; Keizer, J.J. Combining digital soil mapping and hydrological modelling data in a data scarce watershed in north-central Portugal. *Geoderma* **2016**, *264*, 350–362. [[CrossRef](#)]
25. Van Zijl, G.M.; Van Tol, J.J.; Bouwer, D.; Lorentz, S.A.; Le Roux, P.A.L. Combining historical remote sensing, digital soil mapping and hydrological modelling to produce solutions for infrastructure damage in Cosmo City, South Africa. *Remote Sens.* **2020**, *12*, 433. [[CrossRef](#)]
26. Romanowicz, A.A.; Vancloostera, M.; Rounsevelb, M.; La Junesseb, I. Sensitivity of the SWAT model to the soil and land use data parametrisation: A case study in the Thyle catchment, Belgium. *Ecol. Modell.* **2005**, *187*, 27–39. [[CrossRef](#)]
27. Bossa, A.Y.; Diekkrüger, B.; Igué, A.M.; Gaiser, T. Analyzing the effects of different soil databases on modelling of hydrological processes and sediment yield in Benin (West Africa). *Geoderma* **2012**, *173*, 61–74. [[CrossRef](#)]
28. Diek, S.; Temme, A.J.A.M.; Teuling, A.J. The effect of spatial soil variation on the hydrology of a semi-arid Rocky Mountain catchment. *Geoderma* **2014**, *235*, 113–126. [[CrossRef](#)]
29. Van Tol, J.J.; van Zijl, G.M.; Julich, S. Importance of detailed soil information for hydrological modelling in an urbanized environment. *Hydrology* **2021**, *7*, 34. [[CrossRef](#)]
30. Geza, M.; McCray, J.E. Effects of soil data resolution on SWAT model stream flow and water quality predictions. *J. Environ. Manag.* **2008**, *88*, 393–406. [[CrossRef](#)]
31. Wang, X.; Melesse, A.M. Effects of STATSGO and SSURGO as inputs on SWAT model's snowmelt simulation. *JAWRA J. Am. Water Resour. Assoc.* **2006**, *42*, 1217–1236. [[CrossRef](#)]
32. Mukundan, R.; Radcliffe, D.; Risse, L. Spatial resolution of soil data and channel erosion effects on SWAT model predictions of flow and sediment. *J. Soil Water Conserv.* **2010**, *65*, 92–104. [[CrossRef](#)]
33. Adem, A.A.; Dile, Y.T.; Worqlul, A.W.; Ayana, E.K.; Tilahun, S.A.; Steenhuis, T.S. Assessing digital soil inventories for predicting streamflow in the headwaters of the Blue Nile. *Hydrology* **2020**, *7*, 8. [[CrossRef](#)]
34. Seibert, J.; McDonnell, J.J. On the dialog between experimentalist and modeler in catchment hydrology: Use of soft data for multicriteria model calibration. *Water Resour. Res.* **2002**, *38*, 23.1–23.14. [[CrossRef](#)]
35. Winsemius, H.C.; Schaefli, B.; Montanari, A.; Savenije, H.H.G. On the calibration of hydrological models in ungauged basins: A framework for integrating hard and soft hydrological information. *Water Resour. Res.* **2009**, *45*, W12422. [[CrossRef](#)]
36. Yen, H.; Bailey, R.T.; Arabi, M.; Ahmadi, M.; White, M.J.; Arnold, J.G. The role of interior watershed processes in improving parameter estimation and performance of watershed models. *J. Environ. Qual.* **2014**, *43*, 1601–1613. [[CrossRef](#)] [[PubMed](#)]
37. Van Tol, J.J.; Bieger, K.; Arnold, J.G. A hydropedological approach to simulate streamflow and soil water contents with SWAT+. *Hydrol. Process.* **2021**, *35*, 6. [[CrossRef](#)]
38. Shone, R.W.; Booth, P.W.K. The Cape Basin, South Africa: A review. *J. Afr. Earth Sci.* **2005**, *43*, 196–210. [[CrossRef](#)]
39. Rogers, J. Stratigraphy and geomorphology of three generations of regressive sequences in the Bredasdorp Group, southern Cape province, South Africa. In *Symposium on the Geomorphology of Southern Africa*; A.A. Balkema: Rotterdam, The Netherlands, 1988; pp. 407–434.
40. Dent, M.; Lynch, S.; Schulze, R. Mapping mean annual and other rainfall statistics over southern Africa. In *South African Atlas of Climatology and Agrohydrology*; Schulze, R.E., Ed.; WRC Report No. 1489/1/06, Section 7.3; Water Research Commission: Pretoria, South Africa, 2007.
41. Schulze, R.E.; Maharaj, M. Daily minimum temperatures. In *South African Atlas of Climatology and Agrohydrology*; Schulze, R.E., Ed.; WRC Report No. 1489/1/06, Section 7.5; Water Research Commission: Pretoria, South Africa, 2007.
42. SANBI (South African National Biodiversity Institute). The Vegetation Map of South Africa, Lesotho and Swaziland. Mucina, L., Rutherford, M.C., Powrie, L.W., Eds.; Version 2012. Available online: <http://bgis.sanbi.org/SpatialDataset/Detail/18> (accessed on 14 September 2019).
43. Arnold, J.G.; Srinivasan, R.; Muttiah, R.S.; Williams, J.R. Large area hydrologic modelling and assessment, Part I: Model development. *J. Am. Water Resour. Assoc.* **1998**, *34*, 73–89. [[CrossRef](#)]
44. Neitsch, S.L.; Williams, J.; Arnold, J.; Kiniry, J. *Soil and Water Assessment Tool Theoretical Documentation Version 2009*; Technical Report No. 406; Texas Water Resources Institute: College Station, TX, USA, 2011.

45. Bieger, K.; Arnold, J.G.; Rathjens, H.; White, M.J.; Bosch, D.D.; Allen, P.M.; Srinivasan, R. Introduction to SWAT+, a completely restructured version of the soil and water assessment tool. *J. Am. Water Resour. Assoc.* **2017**, *53*, 115–130. [[CrossRef](#)]
46. USGS (United States Geological Survey). Landsat Images. Available online: <http://landsat.usgs.gov> (accessed on 16 September 2019).
47. Geoterraimage. 2013–2014 South African National Land-Cover Dataset; Report Created for Department of Environmental Sciences; DEA/CARDNO SCPF002: Implementation of Land Use Maps for South Africa; Department of Environmental Affairs: Pretoria, South Africa, 2015.
48. Saha, S.; Moorthi, S.; Pan, H.L.; Wu, X.; Wang, J.; Nadiga, S.; Tripp, P.; Kistler, R.; Woollen, J.; Behringer, D.; et al. The NCEP climate forecast system reanalysis. *Bull. Am. Meteorol. Soc.* **2015**, *91*, 1015–1057. [[CrossRef](#)]
49. DWS (Department of Water and Sanitation). Daily Streamflow. Available online: <https://www.dws.gov.za/Hydrology/> (accessed on 16 September 2019).
50. Schulze, R.E. Soils: Agrohydrological information needs, information sources and decision support. In *South African Atlas of Climatology and Agrohydrology*; Schulze, R.E., Ed.; WRC Report No. 1489/1/06, Section 4.1; Water Research Commission: Pretoria, South Africa, 2007.
51. Land Type Survey Staff. *Land Types of South Africa: Digital Map (1:250,000 Scale) and Soil Inventory Datasets*; ARC-Institute for Soil, Climate and Water: Pretoria, South Africa, 1972–2002.
52. Schaap, M.G.; Leij, F.J.; van Genuchten, M.T. ROSETTA, a computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions. *J. Hydrol.* **2001**, *251*, 163–176. [[CrossRef](#)]
53. Zhu, A.X. A similarity model for representing soil spatial information. *Geoderma* **1997**, *77*, 217–242. [[CrossRef](#)]
54. Zhu, A.X.; Band, L.E. A knowledge-based approach to data integration for soil mapping. *Can. J. Remote Sens.* **1994**, *20*, 408–418. [[CrossRef](#)]
55. Zhu, A.X.; Hudson, B.; Burt, J.; Lubich, K.; Simonson, D. Soil mapping using GIS, expert knowledge, and fuzzy logic. *Soil Sci. Soc. Am. J.* **2001**, *65*, 1463–1472. [[CrossRef](#)]
56. Van Zijl, G.M.; van Tol, J.J.; Tinnefeld, M.; Le Roux, P.A.L. A hillslope based digital soil mapping approach, for hydrogeological assessments. *Geoderma* **2019**, *354*, 113888. [[CrossRef](#)]
57. Job, N.M. Geomorphic Origin and Dynamics of Deep, Peat-Filled, Valley Bottom Wetlands Dominated by Palmiet (*Prionium Serratum*): A Case Study Based on the Goukou Wetland, Western Cape. MSc Dissertation, Rhodes University, Grahamstown, South Africa, 2014.
58. Soil Classification Working Group. *Soil Classification: A Natural and Anthropogenic System for South Africa*; Agricultural Research Council, Institute for Soil, Climate and Water (ARC-ISCW): Pretoria, South Africa, 2018.
59. Van Tol, J.J.; Le Roux, P.A.L. Hydrogeological grouping of South African soil forms. *S. Afr. J. Plant Soil* **2019**, *36*, 233–235. [[CrossRef](#)]
60. Paterson, D.G.; Turner, D.P.; Wiese, L.D.; van Zijl, G.; Clarke, C.E.; Van Tol, J.J. Spatial soil information in South Africa: Situational analysis, limitations and challenges. *S. Afr. J. Sci.* **2015**, *111*, 1–7. [[CrossRef](#)]
61. Hutson, J.L. Estimation of Hydrological Properties of South African Soils. PhD Thesis, University of Natal, Pietermaritzburg, South Africa, 1983.
62. Moriasi, D.N.; Arnold, J.G.; van Liew, M.W.; Bingner, R.L.; Harmel, R.D.; Veith, T.L. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Trans. ASABE* **2007**, *50*, 885–900. [[CrossRef](#)]
63. Nash, J.E.; Sutcliffe, J.V. River flow forecasting through conceptual models, Part I: A discussion of principles. *J. Hydrol.* **1970**, *10*, 282–290. [[CrossRef](#)]
64. Gupta, H.V.; Kling, H.; Yilmaz, K.K.; Martinez, G.F. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *J. Hydrol.* **2009**, *377*, 80–91. [[CrossRef](#)]
65. Knoben, W.J.M.; Freer, J.E.; Woods, R.A. Technical note: Inherent benchmark or not? Comparing Nash-Sutcliffe and Kling-Gupta efficiency scores. *Hydrol. Earth Sys. Sci.* **2019**, *23*, 4323–4331. [[CrossRef](#)]
66. Moriasi, D.N.; Gitau, M.W.; Pai, N.; Daggupati, P. Hydrologic and water quality models: Performance measures and evaluation criteria. *Trans. ASABE* **2015**, *58*, 1763–1785. [[CrossRef](#)]
67. Nel, J.L.; Driver, A.; Strydom, W.; Maherry, A.; Petersen, C.; Hill, L.; Roux, D.J.; Nienaber, S.; van Deventer, H.; Swartz, E.; et al. *Atlas of Freshwater Ecosystem Priority Areas in South Africa: Maps to Support Sustainable Development of Water Resources*; WRC Report No. TT 500/11; Water Research Commission: Pretoria, South Africa, 2011.
68. Arnold, J.G.; Youssef, M.A.; Yen, H.; White, M.J.; Sheshukov, A.Y.; Sadeghi, A.M.; Moriasi, D.N.; Steiner, J.L.; Amatya, D.M.; Skaggs, R.W.; et al. Hydrological processes and model representation: Impact of soft data on calibration. *Trans. ASABE* **2015**, *58*, 1637–1660. [[CrossRef](#)]