Using hydropedological characteristics to improve modelling accuracy in Afromontane catchments

Rowena Louise Harrison\textsuperscript{a,},*, Johan van Tol\textsuperscript{b}, Michele L Toucher\textsuperscript{c}

\textsuperscript{a} Department of Soil, Crop and Climate Sciences, University of the Free State, Bloemfontein, PO Box 339, Bloemfontein 9300, Republic of South Africa

\textsuperscript{b} Department of Soil, Crop and Climate Sciences, University of the Free State, Bloemfontein, South Africa, and Afromontane Research Unit, University of the Free State, QwaQwa, Republic of South Africa

\textsuperscript{c} Grasslands-Wetlands-Forests Node, South African Environmental Observation Network, Pietermaritzburg, Republic of South Africa

\textbf{ARTICLE INFO}

\textbf{Keywords:}
Hydropedology
Afromontane catchments
Hydrological modelling
Soil science
Catchment hydrology

\textbf{ABSTRACT}

\textbf{Study region:} Three Afromontane catchments in the Cathedral Peak experimental research site, within the uKhahlamba-Drakensberg escarpment, KwaZulu-Natal, South Africa.

\textbf{Study focus:} Gaining insight into the hydropedological behaviour of catchments enables a deeper understanding of the unique lateral flow dynamics of a landscape and how these affect the hydrological cycle. This study aimed to highlight the importance of understanding the hydropedological behaviour of soils to improve modelling accuracy.

\textbf{New hydrological insights:} Two sets of SWAT+ models were set up for each catchment. The default lateral time, which is the measure of the time required for water to flow through the catchment before being discharged into the stream, was used in the first set up. Specific lateral time inputs, derived from hydropedological soil maps, were utilised in the second model set up and the results compared against observed streamflow. The specific lateral time inputs were based on measured hydraulic properties of the soils coupled with the location of hydrological response units within hydropedological soil maps created for each catchment. The specific lateral time inputs improved modelling accuracy in all statistical parameters used, \(R^2\) (i.e., 0.550–0.903), PBIAS (i.e., 19.742–18.239), ST DEV (i.e., 63.42–51.81), NSE (i.e., 0.316–0.864) and KGE (i.e., 0.630–0.807). This study has highlighted that relevant soil information, based on reliable site-specific data, is essential in hydrological modelling.

\section{Introduction}

One of the important components of addressing challenges related to water resource management is understanding the interaction between water resources and the soil profile (Kahmen et al., 2005, Smith, 2014, Wei et al., 2016, Zhang et al., 2015). This is because soil and water are the fundamental elements in understanding the hydrological response of catchments and therefore the way in which a catchment responds to different management regimes (Bouma, 2016).

In a catchment, each soil type is expected to have a unique influence on its hydrology (van Tol et al., 2013) which also directly regulates soil water flows (Mamera and van Tol, 2018). Soils therefore play a crucial role in rainfall-runoff processes and constituent loading. Soil properties that relate to the rate of infiltration, or ability to store water, significantly affect the water balance in...
watersheds (Geroy et al., 2011, Krpec et al., 2020). These soil properties include texture, organic matter content, bulk density/-porosity, hydraulic conductivity, and water retention characteristics. Some of these properties e.g., texture and organic matter, are easily measured and typically recorded during routine soil surveys. These properties are then used to predict those properties which are laborious and expensive to measure (e.g., water retention and hydraulic conductivity) through pedotransfer functions (Bouma, 1989; Vereecken, 1992, Vereecken et al., 2016). Poor results in hydrological modelling are however produced when simulation models assume homogeneity in soil properties and particularly pedotransfer functions as these functions are only applicable in areas where they were developed (Bouma et al., 2011, Bouma, 2016). To reduce the uncertainty in model outputs, realistic site-specific input data are needed (Robinson et al., 2012).

Hydopedology is an interdisciplinary science, incorporating the concepts of pedology, soil physics, and hydrology to understand soil–water interactions at various scales (Lin, 2003, Lin et al., 2005, 2006). It can therefore provide a significant contribution to the better understanding of soil-water interactions in a specific landscape. Considering hydopedology in watershed modelling has the potential to improve accuracies as its approach is to partition precipitation into infiltration and runoff, thereby redistributing water in soils and landscapes (Bryant et al., 2006) and highlighting the effects of lateral flow on the streamflow dynamics of a landscape (Me et al., 2015).

Understanding the unique lateral flow dynamics of a landscape and how these strongly alter the flow patterns of water, is an important component in hydrological modelling (Bouma, 2006). This is due to runoff potential being influenced by soil properties such as depth to seasonal water table, saturated hydraulic conductivity, and depth to impermeable barriers. These properties determine the rate of infiltration during both dry periods and after prolonged wetting (Bryant et al., 2006, Neitsch et al., 2002), having an influence on the streamflow dynamics of catchment areas.

The KwaZulu-Natal Drakensberg is a mountain escarpment that forms the watershed between the interior catchments of Lesotho and the rivers in KwaZulu-Natal, thus enabling this province to contribute to a quarter of South Africa’s streamflow (Nel, 2009, Whitmore, 1970). This escarpment is therefore crucial for runoff generation, and it is consequently classified as a strategic water source area (Le Maitre et al., 2018). The runoff generation from the montane catchments of the Drakensberg not only supplies KwaZulu-Natal’s water needs but is important in maintaining the water requirements of the economic hub, Gauteng (Nel, 2009).

The main aim of this study was therefore to highlight the importance of understanding the hydopedological behaviour of soils within three Afromontane catchments of this escarpment to improve modelling accuracy. The hypothesis for this study tests that hydopedologic information is needed to understand how water moves through a watershed and model the length of time required for...
water to flow through the catchment as it dictates the partitioning of rainfall between runoff and infiltration. This is achieved through the following objectives: digital soil mapping of the hydropedological behaviour of the three catchments, utilising detailed soil data, particularly of the hydraulic properties of the soils, and using modelling to simulate a more accurate representation of the flow dynamics of the three catchment areas. Our aim was not to calibrate our SWAT+ models. According to Oreskes et al. (1994) the term calibration refers to the process of the manipulation of independent variables of the model to obtain a match between the simulated results and the observed results. It is therefore the process of adjusting model parameters so that the model is forced into particular margins. We did not aim to do this but rather we used the best available data for the catchments to set the models up and then improved the models with the use of detailed soil information in the form of specific lat_time information. We did not perform calibrations on any parameters of the model.

2. Materials and methods

2.1. Study sites

The Cathedral Peak experimental research catchment site forms part of the Ezemvelo KZN Wildlife Maloti-Drakensberg Park which is a World Heritage Site and is situated in the northern part of the uKhahlamba-Drakensberg escarpment, KwaZulu-Natal, South Africa. The South African National Environment Observatory Network (SAEON) undertakes the monitoring of the catchment site. There are fifteen research catchments within this site, all of which are well defined, and are situated at the head of three isolated Little Berg spurs and are underlain by basaltic lavas, which overlie Claren’s Sandstone (Nann, 1956; Toucher et al., 2016). This area is mainly covered by mesic grasslands of the uKhahlamba Basalt Grassland vegetation type with Northern Afrotemperate Forest in narrow bands along the streams and wetlands in some catchments. The main grassland species are Bromus speciosus, Pentaschistis tysoniana, Cymbopogon nardus and Themeda triandra. While the forest species include Leucosidea sericea and Buddleia salvifolia (Mucina et al., 2006, Toucher et al., 2016). The vegetation of the catchment areas is largely controlled by fire with fire regimes included as management treatments of the catchments since 1945 (Gordijn et al., 2018, Toucher et al., 2016).

The Cathedral Peak research catchments fall within the summer rainfall region of South Africa. The mean annual precipitation (MAP) for the area is approximately 1400 mm, with a gradient. From 1300 mm in the southeast (CP-IX) to 1700 mm in the west (CP-III) (Schulze, 1976). Half of the rainfall events in the catchments are thunderstorms (Bosch, 1979, Everson et al., 1998, Toucher et al., 2016).

The fifteen catchments range in altitude from 1820 m.a.s.l to 2463 m.a.s.l. Topography varies from relatively flat to very steep. (1–39°) with the aspect ranging from north to south facing (Granger and Schulze, 1977; Gordijn et al., 2018). Three catchments were selected for this study and are named CP-III, CP-VI, and CP-IX (Fig. 1). These catchments are similar in shape and size but present different histories in terms of land management practices. Details of each catchment is provided in Table 1.

2.2. Collection of monitoring data

The collection of rainfall data within the catchments has been undertaken since 1950, with streamflow monitoring initiated in the three catchment areas during the late 1940’s and 1950’s. At the outlet of each catchment a concrete weir and stilling hut, with 90-degree V Notches were installed (Fig. 2). These V Notches are 45.72 cm (18 in.) deep and are surmounted by 6 feet wide rectangular notches of varying depth. The stilling ponds were dug to bedrock, and rock walls for the pond were constructed. Details of how early measurements were taken, error checked and processed are given in Toucher et al., 2016. The water stage-height at each weir is currently monitored using an Orpheus Mini (Ott Hydromet GmbH, Germany) at CP-VI weir and a CS451 Stainless steel SDI-12 Pressure Transducers with CR200 loggers at weirs CP-III, CP-VI, and CP-IX. There are two pressure transducers installed at weir CP-VI as this is the core catchment and thus the quality of streamflow records is ensured (Toucher et al., 2016).

Between 1987 and 2015, little to no data were collected in any of the three catchment areas due to funding constraints. Thus, this time period was excluded from the model runs. Furthermore, accidental fires, weir silting, and equipment problems has led to periods

<table>
<thead>
<tr>
<th>Catchment Name</th>
<th>Size (ha)</th>
<th>Altitude Range (m. a.s.l.)</th>
<th>Description of Catchment</th>
</tr>
</thead>
<tbody>
<tr>
<td>CP-III</td>
<td>138.9</td>
<td>1847–2323</td>
<td>Mean annual rainfall of 1564 mm. The catchment is degraded as a result of a forestry experiment in which Pinus patula was planted throughout the catchment in the 1950s and 1960s as well as accidental fires which led to the removal of these trees in 1981. The catchment was rehabilitated with Eragrostis curvula, following the removal of the trees (Toucher et al., 2016). There is however erosion throughout the catchment area, with large portions of the catchment covered by Pietardium sp. (Bracken).</td>
</tr>
<tr>
<td>CP-VI</td>
<td>67.7</td>
<td>1845–2073</td>
<td>Mean annual rainfall of 1340 mm. This catchment covered by mesic grassland of the uKhahlamba Basalt Grassland type which is burned biennially during spring. CP-VI is considered the core catchment with focused, detailed monitoring ongoing in this catchment. A full array of evaporation, soil moisture and groundwater monitoring is undertaken.</td>
</tr>
<tr>
<td>CP-IX</td>
<td>64.5</td>
<td>1823–1966</td>
<td>Mean annual rainfall of 1257 mm. This catchment has been completely protected from fire since 1952 but has experienced accidental burns and wildfires in some years. As a result of fire exclusion, this catchment is dominated by woody scrub (Leucasidea sericea).</td>
</tr>
</tbody>
</table>
of missing data from 2016 to 2021 in CP-III and CP-IX.

2.3. Soil and Water Assessment Tool

For this study, the Soil and Water Assessment Tool (SWAT) + version 1.2.2 with QSWAT version 3.10.9-A Coruña was utilised to set up the hydrological models. Model inputs include spatial information, terrain, climate, soil, and land-use. SWAT+ is an adjusted and more flexible version of SWAT (Arnold et al., 1998, Bieger et al., 2016), which is one of the most widely used hydrologic models in the world, being applied in many watersheds across the globe (Bieger et al. 2016). SWAT is a physically based semi-distributed hydrologic model operating on a daily time step to calculate runoff. SWAT+ combines land processes and land management with channel

Fig. 2. Concrete weir and stilling hut, with 90-degree V Notch installed at CP-VI.

Fig. 3. Hydropedological soil group maps utilised for the SWAT+ model for (A) CP-III, (B) CP-VI, and (C) CP-IX. Adapted from Harrison and van Tol (in review).
processes to create a more spatially represented outcome of the interactions and process within a watershed (Bieger et al. 2016), with improved runoff routing capabilities (Bieger et al., 2019, Kakarndee and Kositsakulchai, 2020). The SWAT+ model divides subbasins into water areas and landscape units (LSUs). The LSUs are then subdivided into Hydrological Response Units (HRUs), which are designated as separate spatial objects, with their hydrologic interaction, defined by the user. For example, an HRU can have an interaction with an aquifer as either overbank flow, lateral flow, or surface flow and this can be set by the user based on site specific details (Bailey et al., 2020).

2.4. Model inputs

2.4.1. Terrain, land-use, and climate

Data inputs for the SWAT+ model include a Digital Elevation Model (DEM), land cover information, climatic data as well as soil information. A 5 m resolution DEM (Ezemvelo, 2016) as well as the 2013–2014 SA land cover map information with a 30 m resolution (Geoterrainage, 2015) were utilised. Given that the land cover of CP-III was plantations of *Pinus patula* until 1981, historic model runs of this catchment area utilised a separate land cover map, created to represent this more accurately. Climatic data were obtained from SAEON. Precipitation and temperature data have been recorded on a daily basis from 1950, with additional climatic data (solar radiation, relative humidity, and wind speed) available from 2012 (Toucher et al., 2016). Precipitation data were obtained for each individual catchment, while the remaining climatic data including temperature, relative humidity, wind speed and solar radiation were obtained from the Mikes Pass weather station located approximately 2 km (CP-III), 4 km (CP-VI), and 6 km (CP-IX) from the catchments. While more site specific than utilising data from the global SWAT+ database, the use of climatic data from the Mikes Pass weather station is seen as a limitation to the study as a result of the distance between catchments and this station.

2.4.2. Soil maps

The soil maps utilised in this study were created following a digital soil mapping exercise for the three catchment areas. The procedure used for the digital soil maps (DSMs), is detailed in Harrison and van Tol (in review) and is briefly described here. Soil maps of the three catchments were classified using the South African classification system (Soil Working Group, 2018) and then regrouped into hydropedological soil types, namely, shallow recharge soils, deep recharge soils, interflow soils, and saturated responsive soils, based on the classifications from van Tol and Le Roux, (2019). These groups of soils convey water differently and thus have different hydropedological behaviour. The ArcSIE (Soil Inference Engine) version 10.2.105 was used to create the DSMs. A rules-based approach was first utilised based on knowledge of the catchments as well as the outcomes of the creation of Digital Terrain Models (DTMs) with the following environmental control variables applied to the rules: wetness index, slope, elevation, and planform curvature. The rules applied were aimed at producing the optimal relationships between soil type and a particular DTM (de Menezes et al., 2014, Zhu et al., 1997). The initial maps created following the rules-based approach were then validated based on the information gained during soil surveys undertaken within each of the catchment areas. The maps were refined according to the validation points taken during these surveys. These final hydropedological soil group maps (Fig. 3) were then utilised as the input maps for the SWAT+ model.

2.4.3. Soil parameters

Detailed soil information was obtained from a combination of laboratory-based assessments for both particle size analysis utilising the pipette method (Gee and Bauder, 1986) and carbon analysis using the Leco element analyzer, as well as the work by Kuenene et al. (2007). This work provided details on the bulk density (Bd) and available water capacity (AWC). Saturated hydraulic conductivity (Ks) was calculated utilising HYDRUS-1D, version 4.17.0140 (Šimůnek et al., 2008) from the particle size analysis for the different horizons for each soil group.

Kuenene et al. (2011) utilised a combination of measured hydrograph and soil water content data to create drainage curves for the main soils associated with CP-VI. These drainage curves were created from measuring the soil water content at appropriate time.

![Fig. 4. Hydrograph for CP-VI for the period 24/03/1991-07/10/91, together with the average flow rates per day during six phases, marked P1 to P6 (Kuenene et al., 2011).](image-url)
intervals from a saturated profile until the decrease in soil water content becomes negligible (Hensley et al., 1993). Soil water content was measured using neutron water meter access tubes for a period of four years, between 1991 and 1994, along a 200 m transect in CP-VI. Over the four years approximately weekly soil water content measurements were made at 0.25 m depth intervals. Evapotranspiration values were then subtracted from the total change in soil water content to give the actual change in soil water content due to drainage. The volumetric water content versus time data is used to construct a drainage curve for the particular profile. This was undertaken in three separate profiles of the soils associated with CP-VI to represent deep soils (1800 mm), moderately deep soils (1600 mm) and shallow soils (1300 mm). The main components of the streamflow in CP-VI were identified to be attributed to overland flow and subsurface flow from the vadose zone. This information resulted in the formation of a series of time-steps (in days) that shows how water flows through the catchment and contributes to the streamflow. An example of a drainage curve produced by Kuenene et al. (2011) is shown in Fig. 4. Here the different phases, marked P1 to P6 of the hydrograph are displayed in relation to the time taken for water to move through the catchment.

Utilising these results, along with the hydropedological soil group maps, it was hypothesized that the data obtained from the drainage curves formed in CP-VI could be transferred to the lateral time (lat_time) it takes for each hydropedological soil group to contribute to the streamflow. The lat_time is the flow time required for water to flow through the catchment before being discharged into the stream. Within SWAT+ the default value is set to 0 by default, which allows the model to calculate the travel time based on the soil hydraulic properties input into the model (Neitsch et al., 2002).

Based on the information obtained from the drainage curves and subsequent identification of the travel time required for water to move through CP-VI and be discharged within the streamflow (Kunene et al., 2011), the lat_time input data were then manually specified within the model. This was undertaken to test the hypothesis that this information will improve the accuracy of the model simulations.

2.5. Model setup

The model was configured individually for each of the three catchments. Once the individual models were created the phenology trigger for the plant communities was set as a moisture trigger for growth. Moisture was chosen as the growth trigger over temperature as a number of studies have highlighted that plant growth highlights are mainly controlled by rainfall rather than temperature in sub-Saharan Africa (Alemayehu et al., 2017, Lotsch et al., 2003, Nkwasa et al., 2021). Furthermore, studies in CP-VI have identified that soil moisture and particularly subsurface flow are major contributors to the hydrology and overall ecological drivers of the catchment (Everson et al., 1998).

The model was initially configured and run using the default lat_time. Following this, a second run was conducted, where a specific lat_time was used, and all other parameters held constant. The specific lat_time inputs were based on the location of the HRU within the hydropedological soil group map which was used as one of the base maps for delineating the HRUs. Depending on the location of the HRU and the associated hydropedological soil group, a specific lat_time number was manually input into the model. These model runs were undertaken for two time periods as a result of the gap in data between 1987 and 2015. One run was therefore based on historic data and one run on current data. This was to establish whether the inclusion of specific lat_time values within the model improves the accuracy of the model both in the long term (historic period) and well as more short term (current period) data sets. For CP-III the time periods were 1957–1983 (historic) and then 2015–2021 (current), for CP-VI the time periods were 1962–1987 and then 2015–2021 and for CP-IX the time periods were 1957–1987 and then 2016–2021.

2.6. Model validation

Observed streamflow data were obtained from SAEON. Both daily streamflow and the sum of the daily streamflow as a monthly timestep was utilised in comparisons with the simulated results. In a review of SWAT papers Gassman et al. (2014) showed that the majority of studies cite Moriasi et al. (2007) with regards to judging the success of their SWAT testing results. The strongest results of these papers are reported for the aggregated annual and monthly timesteps. However, Gassman et al. (2014) also noted that over half of these studies further report relatively strong daily timestep statistics. There is thus an increasing number of SWAT studies reporting successful testing at a daily timestep. The use of monthly and daily timesteps in this study was therefore to highlight the strong statistical results for an improvement in the modelling accuracy.

The performance of the simulated results compared with the observed streamflow was analysed using the Nash-Sutcliffe Efficiency (Nash and Sutcliffe, 1970), as well as the alternative metric, the Kling–Gupta Efficiency (KGE). The Nash-Sutcliffe Efficiency (NSE) indicates how well the plot of observed versus simulated data fits the 1:1 line. NSE ranges between -1 and 1, with 1.0 being the optimal value. Values between 0.0 and 1.0 are generally viewed as being acceptable levels of performance (Knoben et al. 2019). The NSE is however sensitive to peak flows at the expense of a better performance during low flow conditions (Krause et al. 2005). The KGE was developed by Gupta et al. (2009) to address shortcomings in NSE and is increasingly utilised in model calibration and evaluation. However unlike in the NSE, there is no specific meaning attached to KGE = 0. The mean flow is therefore used as a KGE benchmark, and model simulations between -0.41 < KGE < 1 are considered as reasonable performance (Knoben et al., 2019).

The coefficient of determination ($R^2$), percent bias (PBIAS), and the percent difference in standard deviation (ST DEV) were additional criteria used for the performance evaluation. The PBIAS measures the average tendency of the simulated data to be larger or smaller than the observed data. The optimal value of PBIAS is 0.0, with positive numbers indicating an overestimation, and negative values indicating an underestimation of the model (Gupta et al., 1999).

Eqs. (1)–(3) were used to calculate the performance indices:
\[ \text{NSE} = 1 - \left[ \frac{\sum_{i=1}^{n} (Y_{i}^{o} - Y_{i}^{m})^2}{\sum_{i=1}^{n} (Y_{i}^{o} - \bar{Y}_{i}^{o})^2} \right] \]

\[ \text{KGE} = 1 - \sqrt{(r - 1)^2 + \left( \frac{\delta_{\text{obs}}}{\delta_{\text{sim}}} - 1 \right)^2 + \left( \frac{\mu_{\text{sim}}}{\mu_{\text{obs}}} - 1 \right)^2} \]

\[ \text{PBIAS} = \left[ \frac{\sum_{i=1}^{n} (Y_{i}^{o} - Y_{i}^{m}) \times 100}{\sum_{i=1}^{n} (Y_{i}^{m})} \right] \]

where \( Y_{i}^{o} \) is the ith observation for the evaluated model, \( Y_{i}^{m} \) is the ith simulation for the evaluated model, \( \bar{Y}_{i}^{m} \) is the mean of the observed data for the evaluated model and \( n \) is the total number of observations.

\( r \) is the linear correlation between observations and simulations, \( \delta_{\text{obs}} \) is the standard deviation in observations, \( \delta_{\text{sim}} \) the standard deviation in simulations, \( \mu_{\text{sim}} \) the simulation mean, and \( \mu_{\text{obs}} \) the observation mean.

Graphical representations showing the comparison between the observed flow, default lat_time and specific lat_time simulations were created. A flow duration curve was created for each model run as an additional performance diagram for the models. The flow-duration curve is a cumulative frequency curve that shows the percent of time during which specified discharges (the observed streamflow of the catchments) were equalled or exceeded in a given period. Flow duration curves set up at a specific site have a key role to play in the knowledge of the streamflow characteristics at that site. This is due to the flow duration curves providing information on the flow variability of the water regime of a specific site during a specific period of interest (Ridolfi et al., 2018). The flow duration curves show the improvement in the specific lat_time model runs in this study.

### 2.7. Sensitivity analysis

Sensitivity analysis was conducted on the lat_time parameter in order to determine if this parameter contributes greatly to the model outputs (sensitive parameter) or if it has only minor relevance (non-sensitive parameter) to the overall performance of the model simulations (Moreira et al., 2018). The model run for CP-VI current time period was chosen for the sensitivity analysis as the specific lat_time parameters were derived from the drainage curves created for this catchment. The value range for lat_time within the SWAT+ model is 0–180 days. An initial mean value of 65 days was utilised within the sensitivity analysis as CP-VI is largely dominated by HRUs within the hydropedological soil group with a lat_time of 65 days. Lenhart et al. 2002, calculates a sensitivity index, I, as a ratio between the relative change of the model output and the relative change of the parameter (lat_time input). A change of 25% of the entire lat_time range (0–180 days) was utilised. This method does not account for interactions among different parameters.

Sensitivity was calculated based on the equation by Lenhart et al. 2002:

\[ I = \frac{(y_2 - y_1)/y_0}{2\Delta x/x_0} \]

in which \( I \) is the sensitivity index, \( y_0 \) is the model output calculated from the initial value (\( x_0 \)), which in this case is 65 days. This initial parameter is varied by \( \pm \Delta x \) (25%) yielding \( x_1 \) and \( x_2 \) with corresponding \( y_1 \) and \( y_2 \) values. The calculated sensitivity indices were ranked into four classes (Table 2) as per (Lenhart et al., 2002).

### 3. Results

#### 3.1. Lateral time used in the model hypothesis

The lateral times taken for each hydropedological soil group to contribute to the streamflow within CP-VI is displayed in Table 3. The same lateral times were applied to CP-III and CP-IX as the soil forms and associated characteristics are similar in all three catchments studied. These times were input into each of the models set up for each of the three catchments based on the location of the HRU within the hydropedological soil group map.

#### 3.2. HRU and hydropedological soil group outputs

Both model outputs in all three catchments had the same HRUs as the same input data were used in both model simulations with the

<table>
<thead>
<tr>
<th>Class</th>
<th>Value Range</th>
<th>Sensitivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>0.00–0.05</td>
<td>Small to negligible</td>
</tr>
<tr>
<td>II</td>
<td>0.05–0.20</td>
<td>Medium sensitivity</td>
</tr>
<tr>
<td>III</td>
<td>0.20–1.00</td>
<td>High sensitivity</td>
</tr>
<tr>
<td>IV</td>
<td>&gt; 1.00</td>
<td>Very High sensitivity</td>
</tr>
</tbody>
</table>
exception of specific inputs in the lat_time. For CP-III the HRUs were 800, for CP-VI there were 717 HRUs and for CP-IX there were 712 HRUs created. Comparison of the HRUs with the hydropedological soil group maps revealed that each catchment had different number of HRUs in the four different hydropedological soil groups and this is based on the different topographies of the catchments as well as the various soil inputs and land covers. CP-III is dominated by the deep well drained soils of the Recharge Deep group, which take 24 days to convey water through the catchment before they contribute to streamflow. In CP-VI the specific lat_time taken for the soils to contribute to streamflow is dominated again by soils of the Recharge Deep hydropedological soil group followed closely by the Interflow soil group. In CP-IX, the Recharge Deep hydropedological soil group again dominates, but in this catchment Recharge Shallow soils follow a close second. The different flow dynamics of each soil type as well as the dominating hydropedological soil group therefore affects the streamflow dynamics of the individual catchment.

3.3. Model outputs

The SWAT+ model was first run in CP-VI as this catchment was utilised to create the drainage curves used to input the specific lat_time values for each hydropedological soil group. The models were run again with the same hypothesis used in CP-VI as similar soils and similar hydropedological recharge groups were identified in both CP-III and CP-IX. The model was run twice, with the integration of (1) the ‘default lat_time and (2) the specific lat_time (hydropedological data) incorporated into the model parameters. As stated previously data collection from the three catchments is patchy in some years and these have reduced the number of observed and simulated data inputs in the statistical equations utilised. It has also led to gaps in the graphical representations of the simulated and observed flows.

3.3.1. CP-VI

In CP-VI climatic and streamflow data from 01/01/1961–31/12/1987 (termed historic time period) and then from 01/01/2014–31/03/2021 (termed current time period) were utilised in separate runs of the model. Printed data from 1961 as well as 2014 were not utilised as this was regarded as a warm-up period for the models. Daggupati et al. (2015) explains why a comprehensive guideline for warm-up periods cannot be given due to the complexity of watershed-scale processes. They however recommend a warm-up period of one to four years with this being related to the temporal and spatial scale of the governing processes. Shorter warm-up periods are required when input values are measured as compared to estimated, the watershed is smaller in size, and the

<table>
<thead>
<tr>
<th>Hydropedological Soil Group</th>
<th>Lateral Time (days) required for water to move through the hydropedological soil group as per the drainage curves from Kunene et al. (2011)</th>
<th>Dominant drainage processes and sources of water</th>
</tr>
</thead>
</table>
| Recharge Shallow Soils      | 6                                                                                               | These are soils that are freely drained and do not show any indication of saturation. They are typically shallow in nature (<500 mm). The freely drained B horizon merges with fractured rock or a lithic horizon. These soils typically occur on steeper convex slopes in the higher lying or steeper parts of the catchments. The recharge shallow soils drain rapidly into nearby drainage channels and wetlands.
| Recharge Deep Soils         | 24                                                                | These are soils that are freely drained and do not show any indication of saturation. They are typically deeper than the Recharge Shallow Soils (>500 mm). The freely drained B horizon merges into fractured rock or a lithic horizon. These soils where identified throughout the catchments on gentler convex and concave slopes and away from wetlands and watercourses. Drainage from the vadose zone of the deep recharge soils flows into nearby drainage channels and wetlands.
| Interflow soils             | 65                                                                | These soils have a freely drained upper solum which overlies relatively impermeable bedrock. Hydromorphic properties are identified at this interface and signify periodic saturation associated with a water table. They typically occur on gentler concave slopes in areas delineated as wetlands as well as adjacent to watercourses. Drainage from the deep phreatic zone of the interflow soils flows into drainage channels and wetlands. The wetlands in all three catchment areas remain saturated throughout the year and are thus continuously fed through this drainage process.
| Responsive Saturated        | 85                                                                | These soils display morphological indications of long-term saturation. These soils were identified in the valley bottom positions of the catchments, in permanently saturated wetlands. They typically occur on gentle concave slopes. Drainage from the responsive saturated soils flows through the wetlands and drainage channels. The wetlands in all three catchment areas remain saturated throughout the year and are thus continuously fed through this drainage process as well as the interflow process. |
model is set-up to evaluate soil moisture processes as compared to groundwater processes. As this study is set in three small watersheds, it is studying soil moisture processes and detailed information on the catchments has been input into the models, 1 year warm-up period was regarded as being sufficient.

Statistical results are presented in Table 4 for both daily and monthly time steps, with graphical representations of a 5-year period for the monthly comparisons as well as the flow duration curves displayed in Fig. 5.

In the historical 1962–1987 data, both the default simulation as well as the specific lat_time simulations show an overestimation of streamflow compared to the observed flow in all runs of the model with the exception of an underestimated flow in the specific lat_time simulated run for the daily time step 2015–2021. The overestimated runs however show an underestimation of the baseflows and overestimation of the peak flows in the original simulations, with these variations being less pronounced in the specific lat_time simulations (Fig. 3). There was an improvement in the R², ST DEV, and NSE values in all specific lat_time simulations compared to the default simulations. The NSE values for the specific lat_time simulated runs are all categorised as ‘very good’ (>0.65) with the exception of the specific lat_time simulated run for the current time period which is categorised as ‘adequate’ (0.54–0.65) as per the classifications by Moriasi et al. (2007). According to Knoben et al. (2019), KGE values between – 0.41 and 1.0 are considered reasonable, and thus all simulated runs (original and lat_time runs) are classified as reasonable.

The flow duration curves created for the monthly time step for both the historical and current time period shows a marked improvement in the simulated lat_time model run versus the observed streamflow as compared to the default model run and the observed streamflow. This is particularly so in the current time period (Fig. 3) and highlights the improvement in the accuracy of the model runs with the input of the specific lat_time.

### 3.3.2. CP-III

In CP-III climatic data as well as streamflow data from 01/01/1957–31/12/1987 (termed historic time period) and then from 01/01/2015–31/03/2021 (termed current time period) were utilised in separate runs. Printed data from 1958 as well as 2015 were not utilised as these were regarded as warm-up periods for the model. Statistical results are presented for both daily and monthly time steps in Table 5, along with graphical representations of a 5-year period for the monthly comparisons as well as the flow duration curves displayed in Fig. 6.

The historical time period for both the daily and monthly time step showed an underestimation of the simulated flows compared to the observed flows in both the default simulated and specific lat_time simulated runs. The current time period showed an overestimation of the default simulated and specific lat_time simulated model runs. In both time periods the graphical representation of the model runs shows an improvement in the simulation of baseflows as well as peak flows in the specific lat_time runs of the model, with the simulated data following the curves of the observed data more closely (Fig. 6).

As with CP-VI there was an improvement in the R², ST DEV, and NSE values in all specific lat_time simulations compared to the default simulations for both time periods (historic and current) as well as for both time steps used. The NSE values for the specific lat_time simulated runs improved in all models from classifications of ‘unsatisfactory’ and ‘satisfactory’ (>0.5) to adequate (0.54–0.65) and ‘very good’ (0.65) in the specific lat_time simulations. The ‘very good’ classifications were obtained in the current time period for both the daily and monthly time step. However, the consideration of the disjointed input data for this current time period must be taken into consideration when comparing the historical and current time periods. All KGE values for all model runs are between – 0.41 and 1.0 and are considered reasonable.

As with CP-VI the flow duration curves created for the monthly time step for both the historical and current time periods in CP-III show the improvement in the accuracy of the model with the input of the specific lat_time. Again, this is especially apparent in the current time period (Fig. 6).

### 3.3.3. CP-IX

In CP-IX climatic data as well as streamflow data from 01/02/1957–31/12/1987 and then from 01/01/2015–31/03/2021 were utilised in separate runs. Printed data from 1958 as well as 2015 was not utilised as these were regarded as warm-up periods for the models. Statistical results are presented for both daily and monthly time steps in Table 6, along with graphical representations of a 5-year period for the monthly comparisons as well as the flow duration curves displayed in Fig. 7.

The historical time period for the monthly time step showed an overestimation of the simulated flows compared to the observed flows in both the default simulated and specific lat_time simulated runs. The remaining model runs (monthly current time period as well as daily historical and current time periods) showed an underestimated of the default simulated and specific lat_time simulated

<table>
<thead>
<tr>
<th>Time step of data</th>
<th>Dates of model run</th>
<th>Model simulation type</th>
<th>R²</th>
<th>PBIAS</th>
<th>NSE</th>
<th>KGE</th>
<th>ST DEV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monthly</td>
<td>01/01/1962 – 31/12/1987</td>
<td>Default Simulation</td>
<td>0.550</td>
<td>19.742</td>
<td>-0.316</td>
<td>0.630</td>
<td>63.42</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Simulation with specific lat_time</td>
<td>0.903</td>
<td>18.239</td>
<td>0.864</td>
<td>0.807</td>
<td>51.81</td>
</tr>
<tr>
<td></td>
<td>01/01/2015–31/03/2021</td>
<td>Default Simulation</td>
<td>0.491</td>
<td>25.812</td>
<td>-0.023</td>
<td>0.463</td>
<td>66.86</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Simulation with specific lat_time</td>
<td>0.931</td>
<td>14.305</td>
<td>0.903</td>
<td>0.849</td>
<td>50.56</td>
</tr>
<tr>
<td>Daily</td>
<td>01/01/1962 – 31/12/1987</td>
<td>Default Simulation</td>
<td>0.418</td>
<td>19.743</td>
<td>-0.541</td>
<td>0.283</td>
<td>3.16</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Simulation with specific lat_time</td>
<td>0.836</td>
<td>15.942</td>
<td>0.802</td>
<td>0.835</td>
<td>1.73</td>
</tr>
<tr>
<td></td>
<td>01/01/2015–31/03/2021</td>
<td>Default Simulation</td>
<td>0.428</td>
<td>25.912</td>
<td>-0.630</td>
<td>0.216</td>
<td>3.04</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Simulation with specific lat_time</td>
<td>0.744</td>
<td>-21.485</td>
<td>0.597</td>
<td>0.681</td>
<td>2.19</td>
</tr>
</tbody>
</table>
Fig. 5. (A) Graphical representations of the monthly comparisons between the default simulated flow as well as the specific lat_time simulated flow and the observed flow for a 5-year period within the two time periods for CP-IV, and (B) Flow duration curves for the monthly data for the two time periods for CP-VI.

Table 5
Statistical results for the SWAT + model simulations for CP-III.

<table>
<thead>
<tr>
<th>Time step of data</th>
<th>Dates of model run</th>
<th>Model simulation type</th>
<th>R²</th>
<th>PBIAS</th>
<th>NSE</th>
<th>KGE</th>
<th>ST DEV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monthly</td>
<td>01/01/1958–31/12/1983</td>
<td>Default Simulation</td>
<td>0.529</td>
<td>3.300</td>
<td>0.474</td>
<td>0.723</td>
<td>49.99</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Simulation with specific lat_time</td>
<td>0.785</td>
<td>-1.381</td>
<td>0.767</td>
<td>0.730</td>
<td>39.32</td>
</tr>
<tr>
<td></td>
<td>01/09/2016–31/03/2021</td>
<td>Default Simulation</td>
<td>0.681</td>
<td>-4.994</td>
<td>0.569</td>
<td>0.760</td>
<td>64.29</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Simulation with specific lat_time</td>
<td>0.867</td>
<td>1.251</td>
<td>0.866</td>
<td>0.891</td>
<td>50.96</td>
</tr>
<tr>
<td>Daily</td>
<td>01/01/1958–31/12/1983</td>
<td>Default Simulation</td>
<td>0.300</td>
<td>2.370</td>
<td>-0.591</td>
<td>0.332</td>
<td>2.80</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Simulation with specific lat_time</td>
<td>0.686</td>
<td>2.412</td>
<td>0.634</td>
<td>0.567</td>
<td>1.13</td>
</tr>
<tr>
<td></td>
<td>10/09/2016–31/03/2021</td>
<td>Default Simulation</td>
<td>0.270</td>
<td>6.311</td>
<td>-1.033</td>
<td>0.181</td>
<td>3.21</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Simulation with specific lat_time</td>
<td>0.863</td>
<td>7.652</td>
<td>0.858</td>
<td>0.863</td>
<td>1.76</td>
</tr>
</tbody>
</table>

Fig. 6. (A) Graphical representations of the monthly comparisons between the default simulated flow as well as the specific lat time simulated flow and the observed flow for a 5-year period within the two time periods for CP-III, and (B) Flow duration curves for the monthly data for the two time periods for CP-III.
model runs. As is the case in CP-III, the graphical representation of the model runs shows an improvement in both time periods in the specific lat_time simulation of baseflows as well as peak flows, with the simulated data following the curves of the observed data more closely when compared to the default simulated flow (Fig. 7).

As with CP-VI and CP-III there was an improvement in the $R^2$, ST DEV, and NSE values in all specific lat_time simulations compared to the default simulations for both time periods as well as for both time steps used. The NSE values for the specific lat_time simulated runs are categorised as ‘very good’ ($>0.65$) for both monthly time step periods as well as the current time period daily time step model run. The historical time period daily time step run is categorised as ‘satisfactory’ ($>0.5$). All KGE values for all model runs are between $-0.41$ and $1.0$ and are considered reasonable, with the exception of the default simulated run ($-0.205$) for the current time period daily time step. The KGE values improve with the specific lat_time run for the same time period ($0.770$) and are classified as reasonable.

As with CP-VI and CP-III the flow duration curves created for the monthly time step for both the historical and current time periods in CP-IX show the improvement in the accuracy of the model with a comparison of the default lat_time versus the specific lat_time compared to the observed streamflow. This is apparent in both the historic and current time periods (Fig. 7).

The sensitivity index calculated for the CP-VI monthly timestep, and current time period is $0.092$. This is categorised as Class II and medium sensitivity (Lenhart et al., 2002).

### Table 6
Statistical results for the SWAT+ model simulations for CP-IX.

<table>
<thead>
<tr>
<th>Time step of data</th>
<th>Dates of model run</th>
<th>Model simulation type</th>
<th>$R^2$</th>
<th>PBIAS</th>
<th>NSE</th>
<th>KGE</th>
<th>ST DEV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monthly</td>
<td>01/01/1958–31/12/1987</td>
<td>Default Simulation</td>
<td>0.611</td>
<td>4.801</td>
<td>0.123</td>
<td>0.598</td>
<td>57.83</td>
</tr>
<tr>
<td></td>
<td>01/09/2016–31/03/2021</td>
<td>Simulation with specific lat_time</td>
<td>0.888</td>
<td>7.809</td>
<td>0.843</td>
<td>0.844</td>
<td>44.29</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Default Simulation</td>
<td>0.678</td>
<td>-11.344</td>
<td>-0.063</td>
<td>0.291</td>
<td>50.07</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Simulation with specific lat_time</td>
<td>0.904</td>
<td>-7.173</td>
<td>0.803</td>
<td>0.723</td>
<td>37.69</td>
</tr>
<tr>
<td>Daily</td>
<td>01/01/1958–31/12/1987</td>
<td>Default Simulation</td>
<td>0.344</td>
<td>5.216</td>
<td>-1.373</td>
<td>0.012</td>
<td>2.89</td>
</tr>
<tr>
<td></td>
<td>02/09/2016–31/03/2021</td>
<td>Simulation with specific lat_time</td>
<td>0.725</td>
<td>-51.023</td>
<td>0.327</td>
<td>0.416</td>
<td>1.89</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Default Simulation</td>
<td>0.424</td>
<td>-8.421</td>
<td>-1.826</td>
<td>-0.205</td>
<td>2.42</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Simulation with specific lat_time</td>
<td>0.840</td>
<td>-5.564</td>
<td>0.753</td>
<td>0.770</td>
<td>1.36</td>
</tr>
</tbody>
</table>

Fig. 7. (A) Graphical representations of the monthly comparisons between the default simulated flow as well as the specific lat_time simulated flow and the observed flow for a 5-year period within the two time periods for CP-IX, and (B) Flow duration curves for the monthly data for the two time periods for CP-IX Sensitivity analysis.

### 4. Discussion and conclusions

In this study the distribution of the hydropedological soil groups within each of the catchments played a role in how precipitation flows through the catchment, either through overland flow or infiltration through the soil profile, how it moved through the various areas of the catchment and how it influenced the streamflow. For example, CP-IX had a greater distribution of soils in the Recharge Shallow group as compared to CP-III and CP-VI and thus there is quicker distribution of water flow from the top of this catchment to the streamflow outlet in this catchment. CP-VI had a greater combined percentage of soils in the Interflow and Responsive Saturated groups compared to soils in the Recharge Deep and Recharge Shallow groups and thus there is a longer lateral time taken for water to move through these soils before it contributes to streamflow. The importance of understanding and incorporating specific lateral time periods, taken for water to move through a catchment before it contributes to streamflow was identified in other studies (Jiao et al., 2020, Johnson et al., 2003, Ma et al., 2017, Me et al., 2015). The ability to define the specific lateral time for each HRU in this study...
was seen as a beneficial input of the SWAT+ model which helped to improve the understanding of the catchments’ dynamics.

The statistical parameters used in this study (R², PBIAS, ST DEV, NSE and KGE) show a definitive improvement in modelling accuracy with the input of the specific lat_time measurements per HRU. These improvements are apparent in all three catchments, with CP-VI obtaining the largest increase in R² values, particularly for the monthly time step when comparing the default simulations with the specific lat_time simulations. Improvements in the other statistical parameters are clear in all three catchments with none of the catchments improving the most in a particular statistical validation. This is despite the specific lat_time inputs being derived from drainage curves created for the conditions of CP-VI. It is therefore postulated that the specific lat_time inputs associated with CP-III and CP-IX are due to the similar soil properties identified in these catchments as well as the detailed soil information input into the model which is particular to each catchment.

The sensitivity analysis conducted for this study showed that the sensitivity of the lat_time parameter is classified as medium. Leng et al. (2020) detailed the use of sensitive versus insensitive parameters in improving modelling accuracy and found that the use of an apparent insensitive parameter, such as soil Bd, enhanced their model significantly, with the R², NSE and RSR parameters improving. This suggested complex hydrological processes occurring in the watershed and a wider variety of sensitive and insensitive parameters that need to be applied to the hydrologic model. The same principles can be applied to this study with a number of parameters being sensitive to modelling accuracy including the detailed soil information, the corresponding and detailed digital soil maps which highlight the hydropedological behaviour of the catchments as well as the detailed information obtained from Kuenene et al. (2011) on the soil drainage curves and how these translate into the time taken for lateral water movement to contribute to streamflow.

Numerous research studies have highlighted the importance of detailed soil information on improving SWAT model accuracy (Adem et al., 2020; Chen et al., 2016, Krpec et al., 2020, Peschel et al., 2006). Soil information should be coupled with ecological information of the catchments such as the effects of fire, vegetation as well as rates of evapotranspiration on hydrological modelling accuracy. These aspects have been shown to be all interrelated. In a study by Manson et al. (2007), within the Cathedral Peak area, frequent burning was shown to result in nitrogen limited soils, which in turn affects the vegetation type as well as vegetation condition which affects the soil condition and therefore the flow dynamics of the catchment. A further study in the Cathedral Peak research catchments by Gordijn and O’Connor, (2021) shows that species diversity increased with heterogenous fire regimes over the long term. Improved species diversity has an impact on soil fertility and flow dynamics (Bai et al., 2001; Mason and Zanner, 2005) having an impact on the hydrological flow of the catchment. Evapotranspiration is another important input into hydrological models as it is a vital component of the water cycle (Aouissi et al., 2016; Zhao et al., 2013).

Several studies furthermore emphasise the importance of understanding the hydropedological character of a catchment or study area and translating this knowledge into input variables within hydrologic models (Bouma et al., 2011, Bryant et al., 2006, Sierra et al., 2018, van Tol et al., 2020, and van Tol et al., 2021). In all these studies there is an improvement in the overall accuracy of the models following the input of soil and hydropedological information or an improved accuracy in a certain section of the model. For example, in van Tol et al. (2021), internal catchment processes were reflected more accurately with rerouting of water fluxes between specific HRUs within the SWAT+ model.

In response to mounting environmental challenges, integrative studies are needed, requiring both specialists within a field as well collaborative efforts across disciplines (Hopmans, 2020, Pachepsky, 2010). The use of soil hydraulic properties within a hydrology model can help one to understand the physical processes that control the partitioning and routing of precipitation into evaporation, infiltration, transpiration, recharge, and runoff (Brooks and Vivoni, 2015), thus improving the model runs for specific catchments.

This study contributed to the understanding of the hydropedological character of three Afromontane catchments and how the characteristics of the soils as well as the flow dynamics of the catchments improves modelling accuracy. It has highlighted that relevant information, based on reliable data, is essential to assess not only the current condition of water resources in a given catchment but also past trends and future possibilities (Droogers and Bouma, 2014).

Funding

This research was funded by the Iphakade Earth Stewardship Science Bursary Programme, South Africa

CRediT authorship contribution statement

R Harrison: Conceptualization, Methodology, Formal analysis, Validation, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization. J van Tol: Conceptualization, Methodology, Resources, Data curation, Writing – review & editing, Supervision. M Toucher: Methodology, Data curation, Writing – review & editing.

Acknowledgements

The authors wish to thank the South African Environmental Observation Network (SAEON) for all the logistical support in gathering the data for this study.

Declaration of competing interest

The authors report no declaration of interest.
