Improved process representation in the simulation of the hydrology of a meso-scale semi-arid catchment

Okello, Aline M.L.Saraiva; Masih, Ilyas; Uhlenbrook, Stefan; Jewitt, Graham P.W.; Van der Zaag, Pieter V.

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

water

MDPI

Article

Improved Process Representation in the Simulation of the Hydrology of a Meso-Scale Semi-Arid Catchment

Aline M. L. Saraiva Okello [1,2,*], **Ilyas Masih** [3], **Stefan Uhlenbrook** [1,4,5], **Graham P. W. Jewitt** [2] **and Pieter Van der Zaag** [3,5]

1 Department of Water Science and Engineering, IHE Delft, 2601 DA Delft, The Netherlands; s.uhlenbrook@unesco.org
2 Centre for Water Resources Research, University of KwaZulu-Natal, 3209 Pietermaritzburg, South Africa; JewittG@ukzn.ac.za
3 Department of Integrated Water Systems and Governance, IHE Delft, 2601 DA Delft, The Netherlands; i.masih@un-ihe.org (I.M.); p.vanderzaag@un-ihe.org (P.V.d.Z.)
4 World Water Assessment Programme (WWAP), UNESCO, 06134 Colombella, Perugia, Italy
5 Department of Water Resources, Delft University of Technology, 2600 GA Delft, The Netherlands
* Correspondence: a.saraiva@un-ihe.org; Tel.: +27-727390616

check for updates

Abstract: The variability of rainfall and climate, combined with land use and land cover changes, and variation in geology and soils makes it a difficult task to accurately describe the key hydrological processes in a catchment. With the aim to better understand the key hydrological processes and runoff generation mechanisms in the semi-arid meso-scale Kaap catchment in South Africa, a hydrological model was developed using the open source STREAM model. Dominant runoff processes were mapped using a simplified Height Above the Nearest Drainage approach combined with geology. The Prediction in Ungauged Basins (PUB) framework of runoff signatures was used to analyse the model results. Results show that in the headwater sub-catchments of Noordkaap and Suidkaap, plateaus dominate, associated with slow flow processes. Therefore, these catchments have high baseflow components and are likely the main recharge zone for regional groundwater in the Kaap. In the Queens sub-catchment, hillslopes associated with intermediate and fast flow processes dominate. However, this catchment still has a strong baseflow component, but it seems to be more impacted by evaporation depletion, due to different soils and geology, especially in drier years. At the Kaap outlet, the model indicates that hillslopes are important, with intermediate and fast flow processes dominating and most runoff being generated through direct runoff and shallow groundwater components, particularly in wetter months and years. There is a high impact of water abstractions and evaporation during the dry season, affecting low flows in the catchment. Results also indicate that the root zone storage and the parameters of effective rainfall separation (between unsaturated and saturated zone), quickflow coefficient and capillary rise, were very sensitive in the model. The inclusion of capillary rise (feedback from the saturated to unsaturated zone) greatly improved the simulation results.

Keywords: hydrological modelling; runoff processes; STREAM model; Kaap River; semi-arid catchment; Southern Africa

1. Introduction

In many regions of the world, including in most parts of Southern Africa, data availability and resources for detailed field investigations are limited [1]. Large catchments need to be modelled with limited input data, yet the results are needed to manage water resources that are crucial to

livelihoods and the environment [1,2]. The importance of process understanding is critical [3,4] because the results of modelling are used for every day water management—water resource availability assessments, water resource development, water releases from dams, environmental flow assessments, etc. [1]. Modellers therefore need to strike a balance between model complexity and data availability—even though more complex models strive to represent the various hydrological processes, they are more data intensive, and often such data are not readily available [1]. Therefore, approaches that rely on readily available data could provide better avenues to improve hydrological models and understanding of hydrological processes.

The Predictions in Ungauged Basins (PUB) decade of the International Association of Hydrological Sciences (IAHS) provided numerous tools and approaches to better understand hydrological processes in data-limited environments. Blöschl, et al. [3] present a useful summary of some of these approaches and suggest a framework to analyse hydrological processes through the use of runoff signatures [3,5]. Runoff signatures are the temporal patterns of runoff response of catchments, derived from observed or modelled series of flow data [5]. They are intended to extract relevant information about hydrological behaviour, such as to identify dominant processes, and to determine the strength, speed, and spatiotemporal variability of the rainfall–runoff response [6].

The literature suggests several approaches to define dominating runoff processes (DRP) zones. Some researchers make use of very detailed field investigations (e.g., [7]), with intensive drilling, soil mapping, and interpretation of aerial images [8,9]. Others make use of geological surveys, and hydrogeological maps. Some attempts have been made to use more simplified approaches, only requiring readily available DEM, geological maps and land use maps [10,11]. More recently, further simplification was suggested by Savenije [12], by using only topography-derived information based on the co-evolution concept. This approach was tested in catchments in Luxembourg and Thailand [13,14].

Savenije [12] argues that catchments have "organized complexity", and therefore relatively simple models can perform well. Thus, it is necessary to model only dominant hydrological processes at the relevant scale using, for example, the concept of landscape zones e.g., plateaus, hillslopes and wetlands. These zones have different dominant hydrological processes, and can be defined by topographic indicators, such as Height Above the Nearest Drainage (HAND). According to Savenije [12], plateaus are hydrological landscapes with modest slope and deep groundwater and where the dominant runoff mechanism is evaporation excess deep percolation. Plateaus are often used for agriculture and the vertical processes such as evaporation and recharge dominate. Wetlands and riparian zones are mostly dominated by saturation excess overland flow runoff mechanism. Hillslopes are hydrological landscapes where storage excess subsurface flow is the dominant mechanism. They are often covered by forest, and generate significant portions of runoff.

In central and northern European landscapes, the hillslopes perform the functions of drainage and moisture retention [12], in order to sustain the predominantly forest ecosystem. Subsurface drainage occurs through preferential pathways. Hillslopes also establish the subsurface connection to the groundwater storage of plateaus.

Soil can also be a first-order control in partitioning hydrological flow paths, residence times and distributions as well as water storage, particularly in smaller catchments [15,16]. In South Africa, the study of dominant runoff generation processes has been spearheaded by the hydropedology community [17]. Schulze [18] was a pioneer in using pedo-transfer functions and decision support systems [19] to derive relevant information for hydrological models from soil maps. van Tol, et al. [9] provide a useful framework for classification of hillslopes in South Africa, based on their soils/geology/hydropedology. It provides detailed explanations of how different hillslopes respond to rainfall, and how the water flows through them.

Several studies (e.g., [20,21]) have used digital soil mapping to derive parameters for hydrological models. However, detailed soils map information at the level of transboundary river basins, such as the Incomati basin, is not available. The Soil Grids 250 m database [22,23] was produced using digital

soil mapping techniques, and could potentially fill this gap in making more detailed soil information readily available for hydrological modelling.

The majority of studies on hydrological processes and runoff generation processes are conducted in small-scale research catchments. However, the understanding of runoff processes is critical for informing models at meso/large catchment scales, where additional processes come into play (e.g., space-time variability of rainfall, runoff routing, larger heterogeneity of geology, soils, land use and climate [7]). This research therefore focuses on understanding runoff generation processes and applying a hydrological model in a meso-scale catchment. We followed the best practice recommendations for predicting runoff in ungauged basins, suggested by Blöschl, et al. [3]. The input data were carefully selected to best represent catchment conditions. Experimental studies were previously conducted in the catchment using tracers [24] and baseflow separation techniques [25] to understand the runoff generation processes. Furthermore, the approach to derive DRP by Gharari, et al. [13], combined with a process-based hydrological model, was used to better understand the key hydrological processes and runoff generation mechanisms in the Kaap catchment.

The specific objectives of this research, in line with the PUB framework, were:

- Interpret key landscape elements with respect to their hydrological functioning, and gather available data for hydrological modelling in the study area;
- Analyse runoff signatures and processes from available gauged catchments;
- Setup a process-based hydrological model that can utilize spatial (gridded) data and that is easy to adapt to different hydrological processes;
- Gradually increase model complexity and assess model sensitivity to different inputs and parameters; and
- Understand the key hydrological processes and runoff generation mechanisms in the catchment, and how this could improve hydrological modelling.

2. Materials and Methods

2.1. Study Area

The Kaap catchment (1640 km^2) is located in the South African Lowveld in the northeast of the Mpumalanga province, forming a sub-catchment of the Incomati river basin. The Kaap catchment contains three main tributaries: the Queens, the Suidkaap, and the Noordkaap (Figure 1). The elevation ranges from 300 m to 1800 m above sea level. The climate is semi-arid with cool dry winters (April to September) and hot wet summers (October to March). Precipitation ranges from 583 mm year^{-1} to 1243 mm year^{-1} in the highest parts of the catchment [26]. The mean potential evaporation is 1435 mm year^{-1}. Streamflow is highly seasonal with the highest average flow occurring in February and the lowest flow at the end of the dry season in September. The mean observed annual streamflow ranges from 66 mm year^{-1} in the outlet to 144 mm year^{-1} in the tributaries (Table A1, Appendix A). The catchment is fairly well monitored with five streamflow gauges available within the catchment (Figure 1). The dominant land covers in the Kaap Valley are Bushveld and grasslands (68% of the catchment area) as observed in Figure 1B. The upstream areas are mostly covered by exotic pine and *Eucalyptus* plantations, which may tap groundwater (25% of catchment area). Sugar cane, cash crops and citrus trees are found downstream and are irrigated (6% of catchment area). There are some mines and urban settlements, Barberton being the main town. Biotite granite is the predominant formation in the Kaap valley. Sandstones and shales are found in closer proximity to the Kaap River and in the south of the catchment [27]. Other formations present include lava, gneiss, ultramafic rocks, quartzite and dolomite (see Figure 1C).

The predominant soils in the Kaap are rhodic ferrasols, chromic cambisols, and haplic acrisols in the headwater catchments. In the Kaap valley, lithic leptosols and rhodic nitisols dominate [22,23]. In terms of soil texture (Figure 1D), 53% of the catchment is covered by sandy clay loams, 39% by clay loams, and the remainder of the catchment has clays, sandy clays and sandy loams [22,23,25].
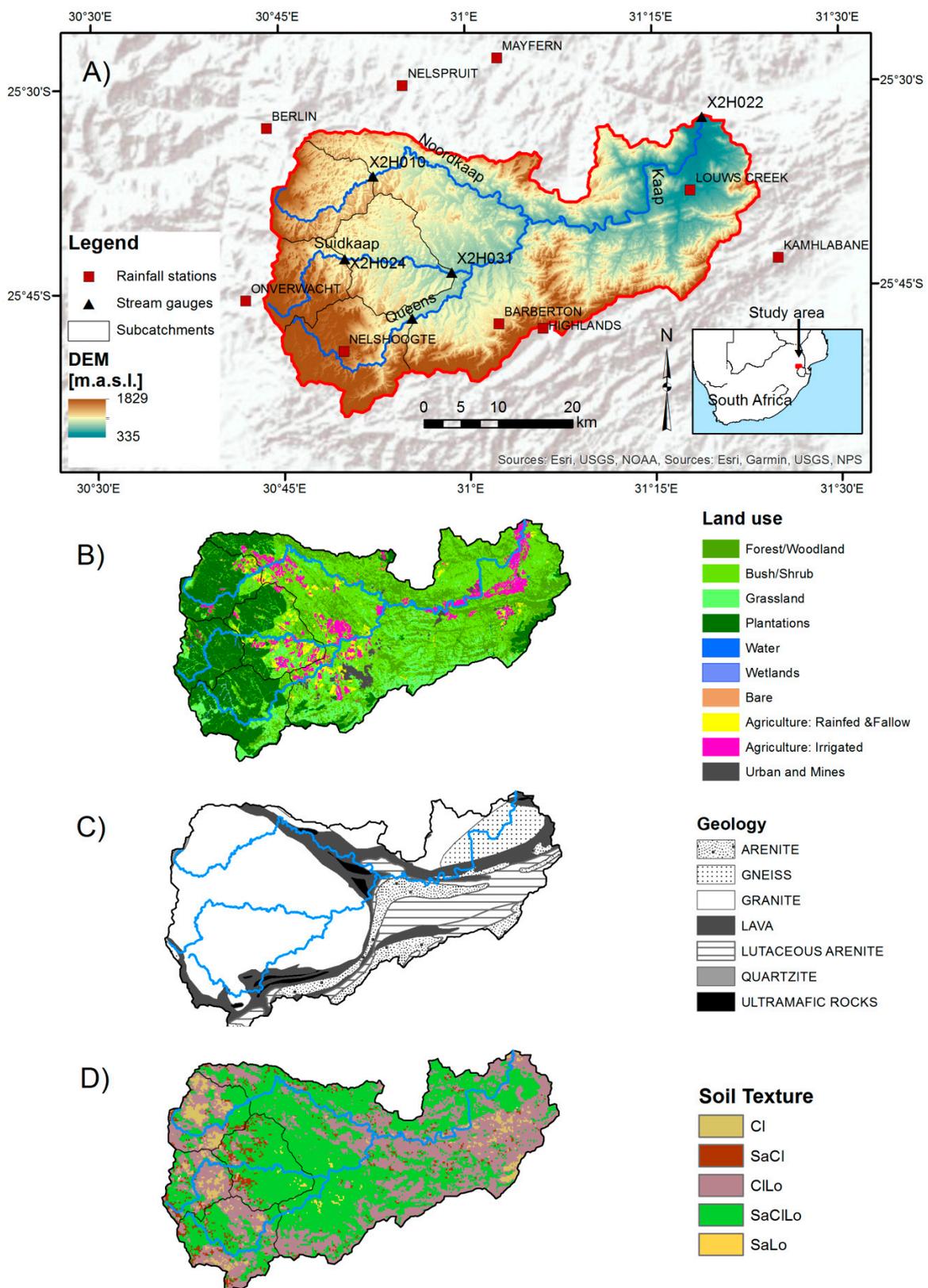
Figure 1. (**A**) Location of Kaap catchment in South Africa (inset) and DEM of Kaap catchment with stream gauges, rainfall stations and sub catchment delineation; (**B**) Land-use and land-cover map of Kaap catchment; (**C**) Geological map [26]; and (**D**) Soil texture based on Soil Grids 250 m dataset [22,23]; the predominantly occurring textures are: Clay (Cl), Sandy clay (SaCl), Clay loam (CL), Sandy clay loam (SaClLo) and Sandy loam (SaLo).

Saraiva Okello, et al. [28] conducted an extensive analysis of rainfall and streamflow in the Incomati basin, and all the gauges within the Kaap were analysed. The analysis revealed that over the past 60 years (1950–2010), no significant upward or downward trend in the catchment rainfall was found, but rather seasonal variability dominated. The streamflow was analysed using the Indicators of Hydrological Alteration tool, and several significant trends were found in the streamflow records. The Noordkaap gauge (X2H010), for example, showed significant decreasing trends in mean monthly flows, low flows, 7-day minimum flow, among others. Further investigation of this shift in the flow regime identified the change of land use, that is, the increase in forestry plantation, as the main driver of the decreasing trends.

Camacho Suarez, et al. [24] conducted an intense tracer study during the rainy season of 2013–2014 in the Kaap catchment. They installed rainfall samplers in two locations in the catchment, and an automatic water sampler at the outlet of the Kaap. Furthermore, grab samples were collected in several locations before rainfall events to provide a snapshot of water quality of the catchment in baseflow conditions. Four major events were sampled and analysed, using isotope and hydro-chemical hydrograph separation, as well as end member mixing analysis. The study revealed great dominance of pre-event water in the streamflow. A three-component hydrograph separation highlighted a major contribution of shallow groundwater, which was enriched with potassium and isotopes. Two main sources of groundwater were identified, the upstream area with fractured granite, characterized by lower ionic content, and the downstream area, with more diverse geology and higher ionic content. Furthermore, a strong correlation was found between antecedent precipitation index and direct runoff. This means that when the catchment is wet from previous rainfall events, and the storages filled, the connectivity of the catchment increases and more direct runoff is generated.

Saraiva Okello, et al. [25] further explored hydrograph separation in the Kaap, using long-term records of water quality, particularly EC. They computed baseflow and quickflow components, using a calibrated recursive digital filter, at monthly and annual scales. The digital filter was calibrated using long-term EC and observed flow data. Hydrograph separation showed that all catchments contribute highly to baseflow.

2.2. Data Used

Hydrological data in the catchment including precipitation, evaporation and streamflow records were collected from the South African Department of Water & Sanitation (DWS, former DWA), the South African Weather Service (SAWS) and the South African Sugarcane Research Institute (SASRI). Figure 1A shows the locations of rainfall and streamflow stations, as well as sub-catchment delineation and topography. To analyze the flow behaviour at the outlet and tributaries, average daily discharges at X2H022 (Outlet), X2H008 (Queens), X2H031 and X2H024 (Suidkaap) and X2H010 (Noordkaap) stream gauges were obtained from the DWS. Land use from the Watplan project was used for this analysis [29]. Topographic information was derived from STRM images with 90 m pixel resolution.

In addition to water use by natural vegetation, the main water users in the catchment are:

- Irrigated sugarcane (98 km^2 area) with a crop water requirement of 92×10^6 m^3 year^{-1} [30]. However, Mallory and Beater [30] report that only 62×10^6 m^3 year^{-1} are supplied from the river.
- Domestic water supply to the Umjindi Local Municipality (over 71,200 population), with a demand of 3.9×10^6 m^3 year^{-1}—this is supplied from an interbasin transfer from the neighbouring Lomati dam (part of the Komati catchment) [30].
- Commercial afforestation (considered a streamflow reduction activity) of 443 km^2, with an estimated streamflow reduction of 40×10^6 m^3 year^{-1} [30].

There are no major reservoirs in the catchment, and the industrial water requirements are considered insignificant [30]. The WR2012 study [31] simulated naturalized flows for the catchment, and these are reported in Table A1 (Appendix A) for comparison. The observed flow constitutes only 57 to 69% of naturalized flows (flows that would occur without human interventions). This means that

31 to 43% of the runoff generated is abstracted due to human activities (irrigation, forestry plantation and other abstractions). The relative difference between observed and naturalized flow is highest for the Kaap and Suidkaap catchments.

2.3. Landscape Classification

SRTM images with 90 m pixel resolution were used to define topography. Furthermore, a landscape analysis was conducted to define zones with similar landscape features, which are presumed to have similar runoff generation processes.

The HAND value was computed, as per the procedure of Rennó, et al. [32] and Gharari, et al. [13]. The HAND value was then combined with the slope map, and thresholds were defined to differentiate Wetlands, Hillslopes and Plateaus (or valley bottom) [12,13] (Figure 2A and Table A1). The thresholds were defined using expert knowledge and some site verifications. Gharari, et al. [13] present an extended calibration procedure to assess sensitivity of HAND model. The thresholds used to define the zones were:

- Stream initiation at 1000 m.
- The HAND threshold to separate wetlands from Plateau and Hillslope was 10 m.
- The slope threshold to separate Hillslope from Plateau was 12%.

Several runs of the model were conducted and compared with verification locations to adjust the parameters.
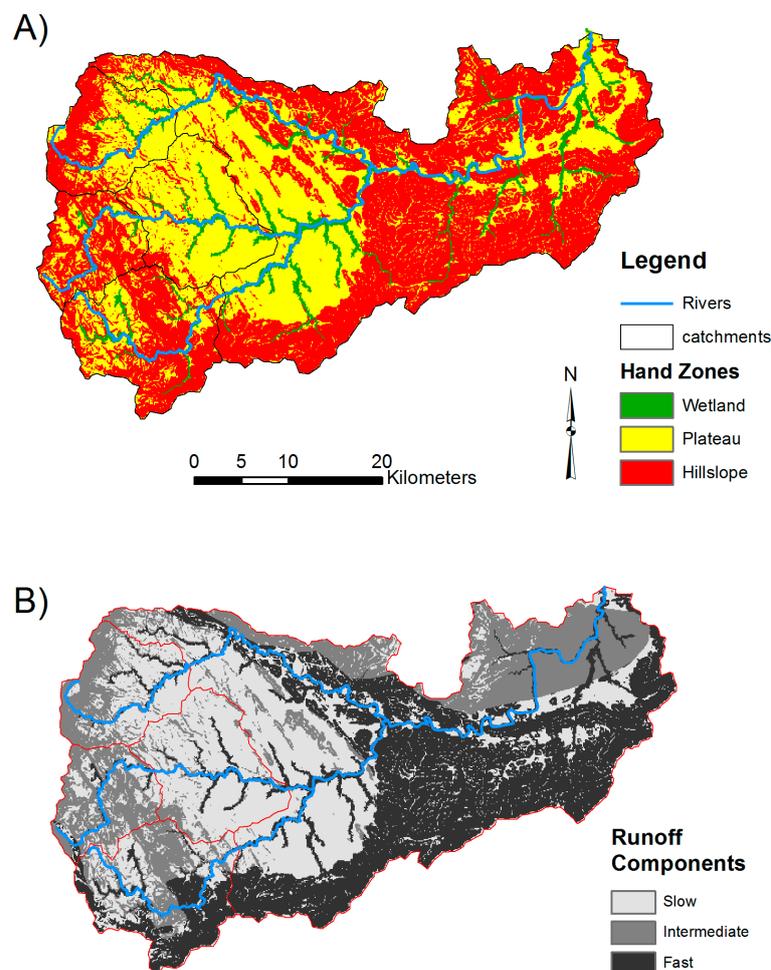

Figure 2. (**A**) HAND zones delineated based on combinations of the HAND value and slope maps; (**B**) Dominant runoff zones, defined based on combination of HAND zones and underlying geology.

2.4. Dominant Runoff Generation Zones

After landscape analysis, and in combination with other physiographic information and previous fieldwork [24], the dominant runoff generation processes were identified in the catchment. The combination of the HAND zones and the geology map helped define zones of slow flow, intermediate (or delayed flow) and fast flow (Figure 2B). All wetlands and sealed areas (urban areas and mines) were considered fast flow generation areas. The plateaus had two dominant mechanisms: Plateaus with underlying geology consisting of quartzite and gneiss were classified as intermediate (or delayed) zones, because both vertical and horizontal flows occur. Plateaus with weathered granite and sedimentary rocks were considered slow flow zones because the vertical percolation and recharge to deep groundwater through fissures of the bedrock is the predominant process [24].

Hillslopes, due to the steep topography, have mostly quickflow occurring through overland flow. When the hillslope has granite, quartzite or gneiss geology, some delayed runoff occurs, as subsurface lateral flow dominates. However, antecedent precipitation can change the dominant processes, in which case, quickflow is sourced from the intermediate runoff zone as well [24].

2.5. The STREAM Model

The Spatial Tools for River basin Environmental Analysis and Management (STREAM) model [33] has been used in several locations internationally and at different spatial/temporal resolutions. It is a spatially distributed and conceptual model, where the non-linear behaviour of the river basins is explained by a combination of thresholds and linear reservoirs. The model is based on a raster GIS which calculates the water balance of each grid cell and routes this through a stream channel network which is based on the digital elevation model (DEM). There is no routing of the surface runoff—it is removed from the model within the same time step as it is generated. A detailed description of model genesis and configuration can be found in several publications [34–36]. The model was selected because of its ability to use distributed (raster) data, and ease of configuration in open source PCRaster dynamic programming language. The main model parameters and variables are presented in Table 1.

The model was used as a tool to test our process understanding in the studied catchments and to highlight shortcomings in process representation in the model. The model structure included some of the main processes expected in a semi-arid catchment such as precipitation, interception, evaporation, and runoff generation (Figure 3). After interception, the effective rainfall is partitioned using the *cr* coefficient between the unsaturated and saturated zones. The portion in the unsaturated zone is available for the transpiration process, which is computed using the soil water balance and is regulated by the maximum unsaturated zone storage, *Sumax*. The portion in the saturated zone can generate runoff, if certain groundwater storage thresholds are exceeded. Initially, the capillary rise process, whereby water from the saturated zone returns to the unsaturated zone, was not simulated, but in later model runs this process was also. Previous research showed that plantation forest (*Eucalyptus*) can access water from great depths, depleting the groundwater storage [37–39]. To mimic this process, the *Sumax* parameter, which was defined based on rooting depth of dominant land use and available water content (based on soil hydraulic properties), was made much larger under forest and plantation land-uses. Where shallow vegetation predominates, the *Sumax* parameter was set at lower values. Note that the hydrological model did not consider direct abstractions of water for irrigation and other purposes.
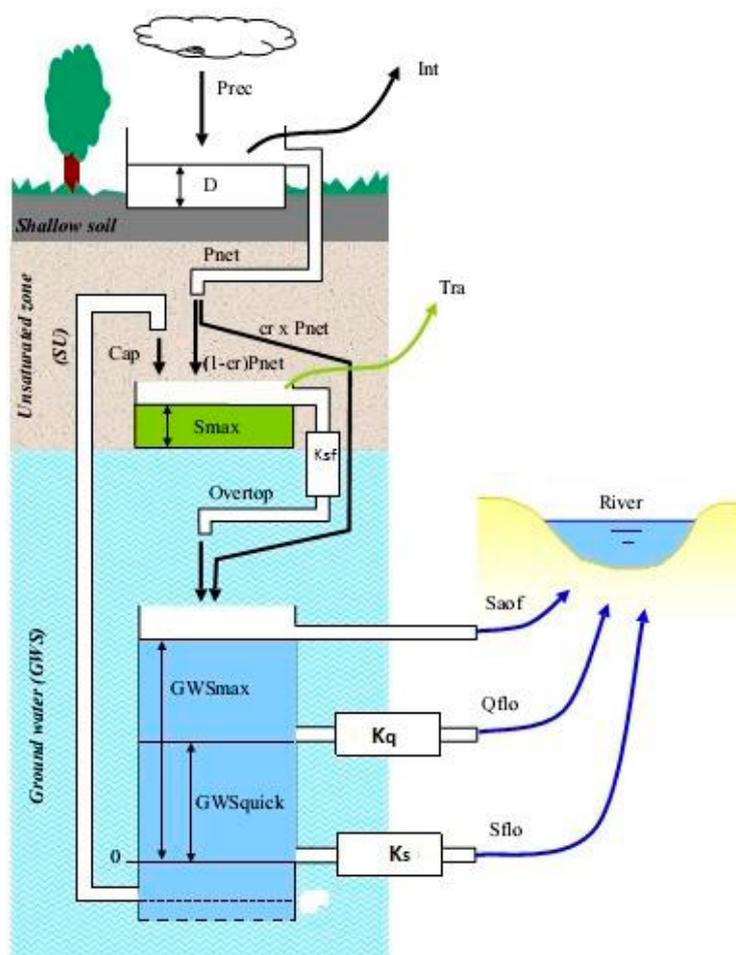
Figure 3. Model configuration [34,36].

2.6. Model Inputs, Parameters and Setup

The hydrological model was configured to simulate stream flow for the period 2003–2013 for the Kaap catchment, with a daily time step and 90 m cell grid size.

Daily rainfall station data (Figure 1) were used for precipitation. The data were interpolated with the Inverse Distance Weighing method (IDW) and were corrected with an elevation factor derived from the Mean Annual Precipitation map [40], according to the methodology described in Sieber and Uhlenbrook [41]. Interception was defined using a fixed daily threshold coefficient *D* based on the land use and land cover map, listed in Table 2. Potential Evaporation was also derived from the station data, and interpolated using IDW method. Actual Transpiration was computed from the soil moisture water balance in the unsaturated zone. The *Sumax* parameter was derived from a combination of available water content (field capacity minus wilting point of each soil type) and rooting depth of each respective land cover. All model parameters were derived from careful analysis of the literature, local expert knowledge and by calibration, as explained in Table 1.

Table 1. Model parameters and variables.

Parameter	Unit	Description	Value	Estimation Method
Ku	day	Overtop timescale	0.5	Recession curve Analysis
Ksf	day	Saturation overland flow timescale	1	Recession curve Analysis
Kq	day	Quickflow timescale	5	Recession curve Analysis
Ks	day	Slow flow (baseflow) timescale	100	Recession curve Analysis

Table 1. *Cont.*

Parameter	Unit	Description	Value	Estimation Method
Suini	mm	Initial storage in unsaturated zone	20	Calibration
GWSini	mm	Initial ground water storage in saturated zone	20	Calibration
LP	-	Reduction of potential evapotranspiration	0.5	Literature
Cr	-	Unsaturated/saturated zone separation coefficient	0–1	Derived from slope, soil texture and land-use land-cover map [42]
Zr	m	Rooting depth	0.5–2.5	Literature/Land-use land-cover map
D	mm/d	Interception threshold	0–4	Literature/Land-use land-cover map
Qc	-	Quickflow coefficient	0–1	Calibration/Soil Texture
Sumax	mm	Maximum storage in unsaturated zone	0–500	Field capacity and rooting depth
GWSmax	mm	Maximum ground water storage in saturated zone	25lnGWSdem	[36]
Cflux	mm/d	Maximum capillary rise threshold	0–2	[35]
GWSmin	mm	Minimum ground water storage threshold to initiate capillary rise	$(0$–$0.5) \times GWSmax$	Modified after [35]

Table 2. Final model parameters dependent on land use and land cover map.

Land Use and Land Cover	%Total Area	*D* (mm/d) *	*Zr* (m) *
Forest/Woodland	19.9%	4	2.5
Bush/Shrub	31.6%	2	1
Grassland	13.9%	2	0.8
Plantations	23.3%	4	2.5
Water	0.2%	0	0.5
Wetlands	0.7%	1	0.5
Bare	0.3%	1	0.5
Agriculture: Rain-fed, Planted pasture, Fallow	2.6%	2	1.5
Agriculture: Irrigated	5.8%	2	2
Urban and Mines	1.7%	1	0.5

* *D* is the interception threshold and *Zr* is the rooting depth, used to compute the *Sumax* parameter.

2.7. Model Simulations

Several model configurations with stepwise variation of model inputs, parameters and processes of differing complexity were tested. The following were the main simulation comparisons conducted:

- Rainfall input (Station data with Thiessen regionalization, with Inverse Distance Weighing and elevation correction, and Remote sensing precipitation from Chirps database).
- Unsaturated/saturated zone separation coefficient *Cr* (-).
- Maximum ground water storage in saturated zone parameter *GWSmax* (mm), derived from DEM or from HAND maps.
- Implementation of capillary rise process, with different thresholds of *Cflux* (mm/d).
- Combinations of capillary rise and different *cr* parameters.
- Implementation of capillary rise with initiation threshold *GWSmin* (*GWSmin* = [0.5, 0.2, 0.1, 0.01] × *GWSmax*).
- Maximum storage in unsaturated zone parameter *Sumax* (mm).

In addition, the HBV model [43,44] was set up for the catchments for comparison. The model was configured using similar input data (precipitation, temperature and potential evaporation), but only vegetation and elevation band zones were used to discretize the model. Automatic calibration was applied to obtain the best performing parameter sets.

2.8. Runoff Signatures and Assessment of Model Performance

The Prediction in Ungauged Basins (PUB) book [3] suggests a framework for hydrological understanding of catchments, by focusing on their runoff signatures. There are a myriad of possible signatures, but we choose to focus on the key signatures suggested by Blöschl, et al. [3], which are commonly used in the region as well (e.g., [45]): annual runoff, seasonal runoff, flow duration curve (FDC), low flows, floods, and runoff hydrographs.

The model performance was also assessed visually and statistically using different indicators of goodness of fit of the hydrographs: the Nash-Sutcliffe efficiency (NSE), the Logarithmic Nash-Sutcliffe efficiency (LogNSE), Bias and percentage Bias (PBias), Mean absolute error (MAE), the Pearson R^2, Root mean square error (RMSE), and Kling-Gupta efficiency (KGE) coefficient [46,47]. The NSE varies between $-\infty$ and 1.0, with 1.0 being the optimal value. Values between 0 and 1 are considered acceptable, whereas less than 0 is unacceptable performance. LogNSE has a similar range, but the flow values are transformed into logarithmic to better analyse low flows. Bias, MAE and RMSE have the same unit as observed flow, whereas PBias is the percentage of bias in relation to mean flow; the closer to 0, the better the model performance, with low-magnitude values indicating accurate model simulation. Positive values indicate model underestimation bias, and negative values indicate model overestimation bias. R^2 varies between 0 and 1, whereas KGE varies between and $-\infty$ and 1. In both cases, values between 0.7 and 1 are considered good; between 0.5 and 0.7, acceptable; and below 0.5, poor. The KGE also offers diagnostic insights into the model performance because of the decomposition into correlation, bias term and variability term. From a hydrologic perspective usage of KGE assists in reproducing temporal dynamics, as well as preserving the distribution of flows; therefore, this was adopted as the main indicator of goodness of fit.

3. Results

3.1. Model Parameterization

The final STREAM model parameters used for the comparison simulations are listed in Tables 2 and 3, and illustrated in Figure 4. Several manual calibration runs were conducted, where each parameter was varied while others kept constant. The best performing parameter sets were retained for subsequent simulations.

For the *Sumax* parameter, the areas with forest and plantations had higher *Sumax* values, because these occur on locations with deeper soils and stronger baseflow, indicative of larger water storage; research has shown that these vegetation types can tap deep water stores.

The *cr* parameter was derived from a combination of land use, soil texture and slope [42]. The *qc* parameter, however, was mainly driven by soil type. Coarser soils, such as sandy loams or sandy clays have a lower *qc* threshold, because these soils allow for quicker response, the threshold to initiate quickflow being lower. The finer clayey soils, in contrast, can hold water for longer periods of time, increasing the *qc* threshold value. The *GWSmax* parameter followed closely the elevation pattern, and the relationship derived by Gerrits [36] was used.

Table 3. Quickflow separation coefficient, based on soil texture.

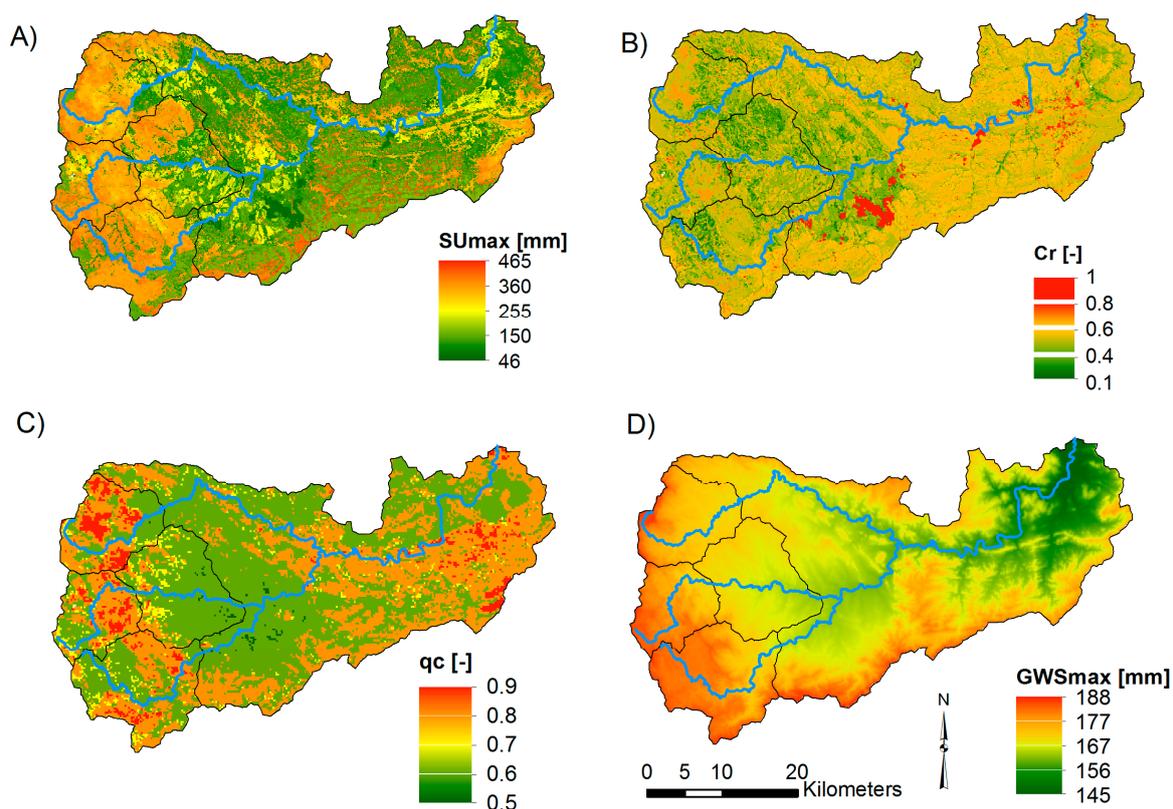| Soil Texture	*Qc* [-]
Clay (Cl)	0.9
Sandy clay (SaCl)	0.7
Clay loam (CL)	0.8
Sandy clay loam (SaClLo)	0.6
Sandy loam (SaLo)	0.5

Figure 4. Maps of final model parameters used. (**A**) *Sumax* map (mm); (**B**) *Cr* coefficient (-); (**C**) *Qc* coefficient (-); (**D**) *GWSmax* (mm).

3.2. Model Simulations

Over 70 model runs were conducted, but only a sample of four representative runs will be presented and discussed (Table 4). Overall, the model simulations were able to capture the flow dynamics well. However, in several runs, the model overestimated peak flows and baseflows, especially at the Kaap outlet and in Queens.

Table 4. Characteristics of model runs presented in the results section.

Run	*Cflux* (mm/d)	*GWSmin*	Description
53	0	0	Model without capillary rise implemented
60	1	0	Model with capillary rise implemented but no *GWSmin*
64	2	0.1*GWSmax*	Model with capillary rise implemented and *GWSmin*
67	2	0.01*GWSmax*	Model with capillary rise implemented and *GWSmin*

A comparison of the goodness of fit indicators was done to see which model better represents the actual catchment conditions (Table 5). Overall, the Pearson correlation was good, ranging from 0.75 to 0.84, meaning that the simulated flows generally followed well observed flow pattern. The NSE was poor to acceptable, mostly due to the overestimation of flow in some runs (e.g., run 53 and run 60), and the seasonality of the flow. In the best performing runs, the NSE was 0.5 to 0.66. Run 60 simulated zero flows during the low flow season, and thus it was not possible to calculate LogNSE. In terms of KGE, which is the most integrated indicator, run 64 was the best for Noordkaap and Suidkaap catchments, with KGE of 0.67 and 0.75, respectively, whereas run 67 was the best for the Queens and Kaap catchments, with KGE of 0.79 and 0.83, respectively. The Bias was very high, especially in the Kaap outlet. These results show how different model setups are needed for the

different catchments. However, there is a similarity between Noordkaap and Suidkaap, and also between the Queens and the Kaap.

Furthermore, a visual analysis of the different hydrological signatures was conducted to further understand which processes were better represented by each model setup. From the other model simulations (not reported here), the parameters *cr* and *qc* proved to be very sensitive.

Table 5. Goodness of fit indicators for the selected model runs, at monthly time scale.

	Bias	KGE	LogNSE	MAE	NSE	PBias	RMSE	Pearson R^2
	(m^3/month)	(-)	(-)	(m^3/month)	(-)	(%)	(m^3/month)	(-)
Run53								
X2H010	22.8	−0.29	−0.77	22.8	−2.46	170.9	28.1	0.83
X2H008	48.4	−1.23	−1.26	48.4	−3.97	479.1	59.2	0.83
X2H031	52.1	−0.91	−1.66	52.6	−5.72	269.3	64.3	0.76
X2H022	371.7	−2.76	−1.38	371.7	−11.44	7299.9	424.5	0.84
Run60 [a]								
X2H010	−0.4	0.68	na	8.1	0.53	−16.6	10.4	0.84
X2H008	14.4	0.32	na	18.8	−0.13	88.9	28.2	0.82
X2H031	5.3	0.62	na	17.8	0.05	8.6	24.1	0.76
X2H022	85.6	0.12	na	102.2	−0.46	186.9	145.3	0.84
Run64								
X2H010	−4.2	0.67	0.38	6.7	0.52	−15.4	10.5	0.77
X2H008	7.2	0.61	0.53	12.1	0.50	87.4	18.7	0.83
X2H031	−0.8	0.75	0.58	10.6	0.52	10.6	17.1	0.75
X2H022	57.7	0.32	0.10	72.4	0.42	1829.5	91.6	0.84
Run67								
X2H010	−8.4	0.22	−6.10	9.6	0.30	−58.9	12.6	0.79
X2H008	1.4	0.79	−0.44	11.3	0.55	−2.3	17.7	0.82
X2H031	−9.8	0.34	−4.14	13.7	0.37	−49.1	19.7	0.76
X2H022	0.7	0.83	0.35	47.5	0.66	197.7	70.6	0.84

[a] LogNSE could not be computed for run 60 because there were months with zero flow simulated. Note: KGE is the Modified Kling Gupta Efficiency, LogNSE is the Logarithmic Nash Sutcliffe efficiency coefficient, MAE is the Mean Absolute Error, NSE is the Nash Sutcliffe efficiency coefficient, RMSE is the Root Mean Squared Error and PBias is the Percent Bias.

3.3. Comparison of Runoff Signatures

3.3.1. Annual Runoff

The annual runoff, which is a key component of the water balance, was computed for all hydrological years. There was difficulty in closing the water balance in the initial runs, when a simple setup without capillary rise (or feedback from the groundwater storage to the unsaturated zone) was implemented. Figure 5 illustrates the results of run 64, compared to observed flow for the four catchments.

Regarding the annual dynamics, one can see that the model tends to better capture the flows generated in wetter years than in drier years. This may be due to more uncertainty in the storage conditions during drier years, and the impact of water abstractions for irrigation. The naturalized flow is also 17 to 51% higher than observed flow, which implies that water abstractions and reductions in streamflow could be up to 50%, particularly in the dry season.

One important aspect to consider is that in semi-arid and sub-humid areas, the evaporation component of the water balance is very large. Actually, the potential evaporation is much larger than rainfall, which reflects in more than 90% of the water balance being attributed to evaporation, and only 10% or less to runoff generation (Table 6). Therefore, uncertainties related to the computation of evaporation, such as the parameters used, interception, and the interpolation of input data, can greatly affect the results of model simulations. This is illustrated by the great difference between evaporation estimates from different model runs (Table 6). A comparison was

also made between evaporation generated by the water balance model, and evaporation from remote sensing products (Table 6). A more detailed description of the remote sensing evaporation products is available in the Supplementary Material. Comparing monthly and annual scales revealed that both ALEXI [48,49] and CMERST [50] products generally overestimate actual evaporation, whereas SSEBop [51,52] results in an underestimate.
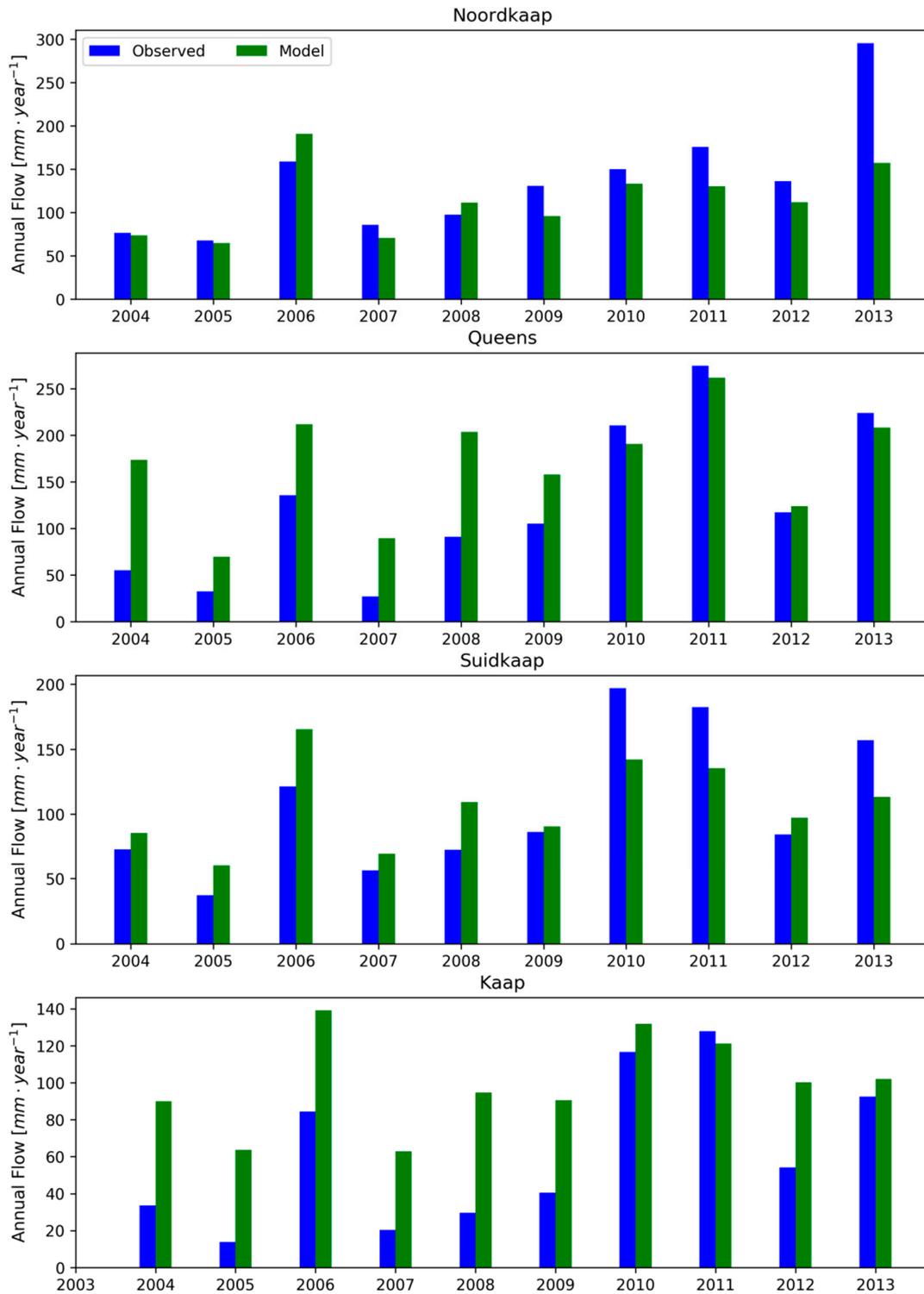

Figure 5. Annual Flow for the four catchments, using results of run 64.

Table 6. Average annual water balance, including evaporation from remote sensing products for comparison, for the period 2003-2013.

	Noordkaap X2H010	Queens X2H008	Suidkaap X2H031	Kaap X2H022
	Mean ± Stdev	Mean ± Stdev	Mean ± Stdev	Mean ± Stdev
Rainfall (mm/year)	1008.0 ± 154.1	1126.6 ± 181.6	946.3 ± 135.9	774.0 ± 121.6
FlowObs (mm/year)	137.5 ± 66.7	127.4 ± 84.5	106.8 ± 55.0	61.4 ± 41.1
RC (%)	13% ± 5%	11% ± 0.1	11% ± 0.0	7% ± 4%
Qnat (mm/year) [a]	222.0 ± 64.0	153.0 ± 70.0	217.0 ± 49.0	106.0 ± 36.0
Fm53 (mm/year)	335.7 ± 67.5	406.4 ± 74.4	316.0 ± 53.9	298.1 ± 48.2
Fm60 (mm/year)	144.7 ± 61.9	210.4 ± 75.8	131.0 ± 49.4	117.2 ± 40.5
Fm64 (mm/year)	113.9 ± 40.5	169.1 ± 59.7	106.9 ± 33.1	99.6 ± 25.6
Fm67 (mm/year)	78.9 ± 45.3	135.7 ± 64.2	71.3 ± 36.4	63.6 ± 28.1
PminFlowOb (mm/year)	870.5 ± 112.8	999.2 ± 121.5	839.5 ± 94.7	712.7 ± 85.8
ETfao (mm/year)	1250.7 ± 69.2	1221.3 ± 93.6	1225.0 ± 89.5	1235.3 ± 83.9
ETm53 (mm/year)	672.0 ± 93.7	717.9 ± 109.3	660.1 ± 88.5	503.7 ± 72.5
ETm60 (mm/year)	864.2 ± 98.0	915.3 ± 109.8	838.2 ± 93.3	681.7 ± 83.8
ETm64 (mm/year)	894.7 ± 120.0	956.4 ± 126.1	853.7 ± 108.3	691.9 ± 99.3
ETm67 (mm/year)	930.0 ± 114.8	990.1 ± 121.4	890.1 ± 106.1	728.9 ± 97.7
ETal (mm/year)	1100.9 ± 54.3	1079.6 ± 55.3	884.7 ± 41.7	861.4 ± 35.8
ETcm (mm/year)	1127.4 ± 269.8	1142.6 ± 56.3	1005.0 ± 50.9	831.2 ± 181.3
ETss (mm/year)	788.5 ± 35.9	733.8 ± 59.8	608.2 ± 68.4	608.4 ± 53.9

[a] Qnat is the average for 2003–2010, given that naturalized flows are only available up to 2010 hydrological year. Note: FlowObsv is observed flow, RC is runoff coefficient, Qnat is the naturalized flow obtained from WR2012 database, Fm53 to 67 are the simulated flows for runs 53 to 67, PminFlowOb is the difference between precipitation and observed flow, ETfao is the potential evaporation using FAO method, ETm53 to 67 are the simulated actual evaporation for model runs 53 to 67, ETal is evaporation from ALEXI product, ETcm is evaporation from the CMERST product, and ETss is evaporation from the SSEBop product.

3.3.2. Seasonal Runoff

The average monthly streamflow graphs (Figure 6) show the strong seasonality of streamflow in the catchments. The flow components analysis demonstrates that quickflow and saturated overland flow are only active for few months of the year (November to March, with peak contributions in January/February). For the Noordkaap catchment, model run 64, which included capillary rise and *GWSmin* threshold, was able to capture the monthly flow pattern relatively well. This pattern was also well captured in the Suidkaap catchment. For the Queens and Kaap catchments, however, this model run overestimates flow, especially during the wet months. Another model configuration (model 67, available in the Supplementary Material), which included a higher coefficient for capillary rise, generated better results for these latter two catchments. Overall, from the results of run 64, the baseflow component accounted for 85% of the flow in the Queens catchment, 95% in the Nordkaap, 94% in the Suidkaap, and 93% in the Kaap. The quickflow contribution ranged between 4 and 13%, whereas saturated overland flow was about 1% or less for all catchments.

3.3.3. Flow Duration Curves

The flow duration curves (Figure 7) give a comprehensive overview of the flow regime. Once more it is evident that the flow variation for the Noordkaap and Suidkaap catchments is fairly well represented. However, for the Queens and Kaap catchments the model overestimates the middle and low flows. This overestimation could be because of the representation of subsurface flow processes in the catchment model. Both Queens and Kaap have higher percentages of area with hillslopes, and also more diverse geology and soils, including a variety of sedimentary rocks. Apparently a more complex representation of flow processes and groundwater is required to capture such variability. Furthermore, the Kaap catchment has higher water abstractions for irrigation—and most of this water is abstracted during the dry season and in drier years. The naturalized flow for the Kaap catchment is 42% higher than the observed flow, which can largely be attributed to water abstractions for irrigation. Most of the

irrigated sugarcane in the catchment is located in the Lower Kaap valley, with an estimated irrigation demand of 92×10^6 m³/year.
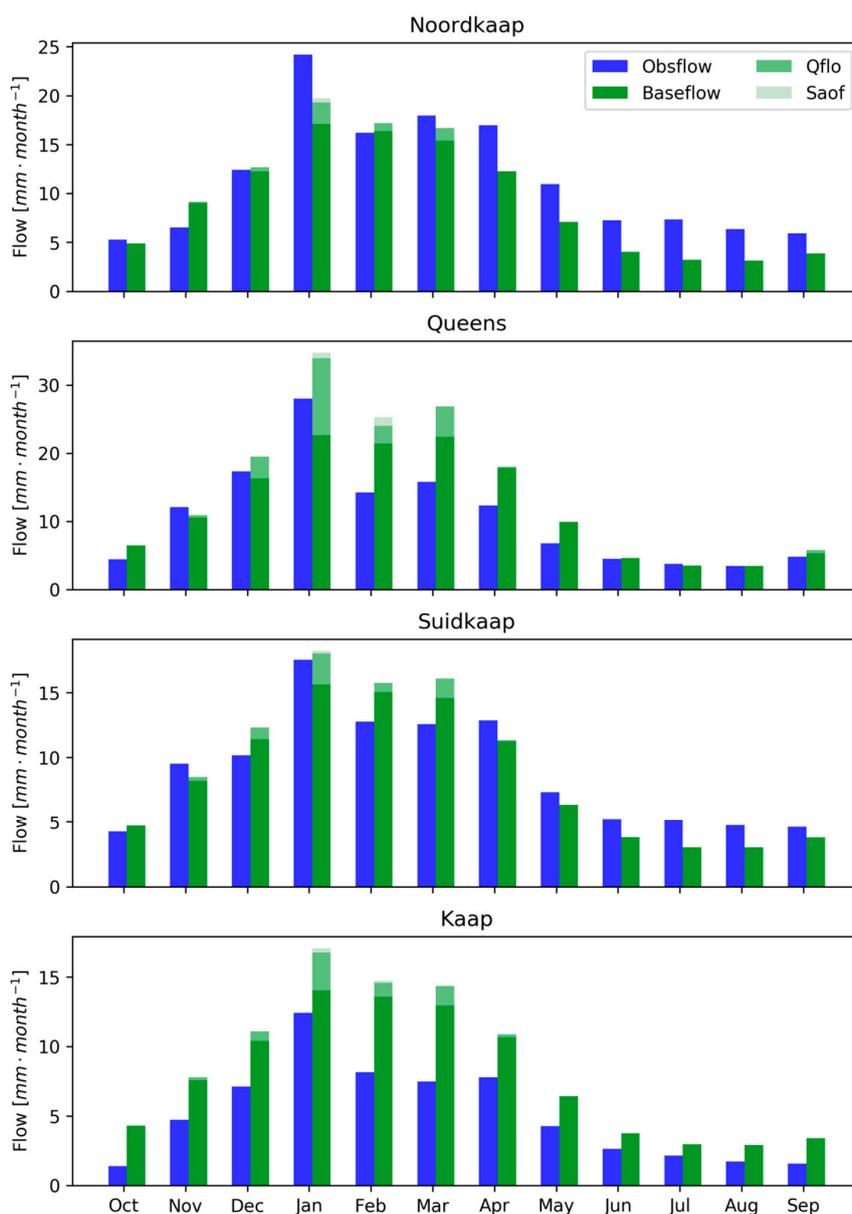

Figure 6. Mean monthly observed and modelled flow for the four catchments, using results of run 64. Modelled flow is partitioned in the components: baseflow, saturated overland flow (Saof) and quickflow component (Qflo).

3.3.4. Low Flows and Floods

The low flows and floods can be characterized by their magnitude, frequency and duration. The flow duration curve provides commonly used indicators of low flows and high flow regime. Q95 (flow exceeded 95% of the time) is the most frequently used indicator of low flow. Q75 is frequently used in South Africa for yield estimation, and the Q50 and Qmean are also reported. On the high flow side, Q1 and Q5 are used, as well as floods with 100-year recurrence intervals. However, given that the model dataset is only 10 years long, it was not sufficiently long to perform a flood frequency analysis. For the sake of comparing the model runs, Q1 and Q5 for modelled and observed time series were used instead compared as indicators of high flows.

Table 7 compares flow percentiles for observed and modelled streamflow under different model setups. The slope of the flow duration curve, computed as the slope between Q30 and Q70, is also reported for comparison. This set of signatures reveals once more the difficulty of having one single model setup performing equally well for both high and low flows, while getting the same slope of the FDC. This is likely due to the fact that in these catchments, different processes control the low flow and high flow generation; apparently the simple model structure did not fully capture these differences.
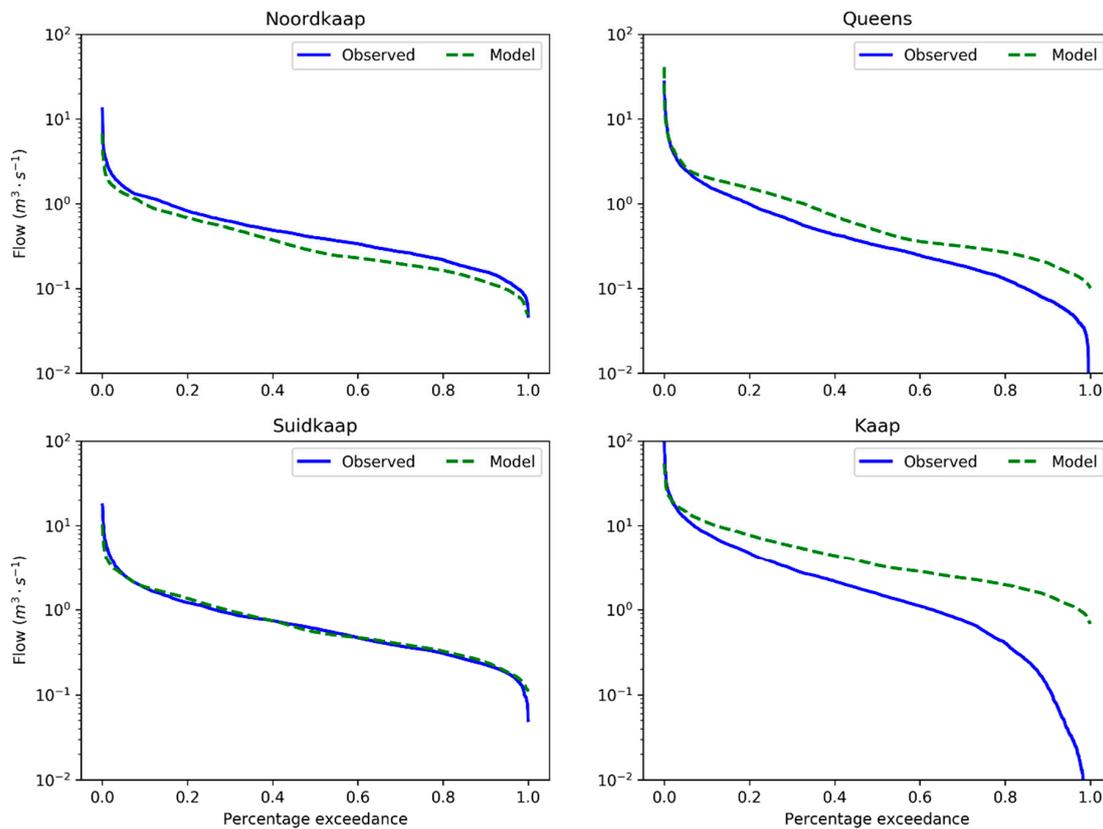

Figure 7. Flow duration curve for the four catchments, using results of run 64.

Table 7. Quantiles of high and low flow, and slope of flow duration curve (between Q30 and Q70) for the different model runs compared to observed flow (FlowObs).

		Q01	**Q05**	**Q50**	**Q75**	**Q90**	**Q95**	**Q99**	**Qmean**	**SlopeFDC**
		m^3/s	m^3/s	m^3/s	m^3/s	m^3/s	m^3/s	m^3/s	m^3/s	(-)
Noordkaap	FlowObs	3.12	1.60	0.40	0.24	0.16	0.12	0.09	0.59	1.48
	Fm53	5.68	3.03	1.12	0.66	0.47	0.41	0.30	1.34	1.51
	Fm60	2.96	1.68	0.38	0.02	0.00	0.00	0.00	0.58	3.33
	Fm64	2.04	1.32	0.27	0.18	0.12	0.10	0.07	0.46	1.74
	Fm67	1.92	1.23	0.10	0.01	0.01	0.01	0.01	0.32	2.97
Queens	FlowObs	5.94	2.49	0.32	0.16	0.07	0.05	0.03	0.73	1.55
	Fm53	12.27	6.49	1.76	1.03	0.75	0.64	0.49	2.32	1.32
	Fm60	8.11	3.49	0.76	0.09	0.00	0.00	0.00	1.20	2.98
	Fm64	5.89	2.69	0.48	0.29	0.20	0.16	0.12	0.96	2.01
	Fm67	5.54	2.58	0.26	0.02	0.02	0.01	0.01	0.77	3.14
Suidkaap	FlowObs	5.71	2.57	0.60	0.34	0.23	0.18	0.11	0.91	1.44
	Fm53	12.21	6.22	2.18	1.26	0.91	0.78	0.58	2.62	1.50
	Fm60	5.93	3.19	0.69	0.04	0.00	0.00	0.00	1.09	3.29
	Fm64	4.05	2.56	0.55	0.36	0.24	0.19	0.13	0.89	1.62
	Fm67	3.75	2.34	0.17	0.03	0.02	0.02	0.01	0.59	2.79

Table 7. *Cont.*

		Q01	Q05	Q50	Q75	Q90	Q95	Q99	Qmean	SlopeFDC
		m^3/s	m^3/s	m^3/s	m^3/s	m^3/s	m^3/s	m^3/s	m^3/s	(-)
Kaap	FlowObs	25.06	12.20	1.55	0.57	0.12	0.04	0.00	3.28	1.72
	Fm53	65.57	36.15	13.10	7.52	5.47	4.74	3.52	15.49	1.53
	Fm60	32.04	18.76	3.76	0.30	0.00	0.00	0.00	6.09	3.21
	Fm64	22.61	14.67	3.35	2.18	1.48	1.14	0.87	5.18	1.62
	Fm67	20.46	13.06	1.09	0.15	0.11	0.09	0.06	3.30	2.77

3.3.5. Hydrographs

Figure 8 shows the hydrographs for catchments aggregated at monthly time scale for the entire simulation period. During a sequence of wet years (2010 to 2013), the model was able to represent the flow dynamics fairly well with no systematic under- or over-prediction of flows. This confirms that the model performed better under wetter conditions than under drier ones. In drier years, better characterization of the evaporation processes is required, as well as groundwater discharge, storage and water abstractions, as this greatly influences the water balance. Small, localised rain events are also very difficult to capture with the given monitoring network. Water abstractions for irrigation and other uses are relatively higher in drier years, so explicit representation of the irrigation management and other water uses would be required to improve the model.
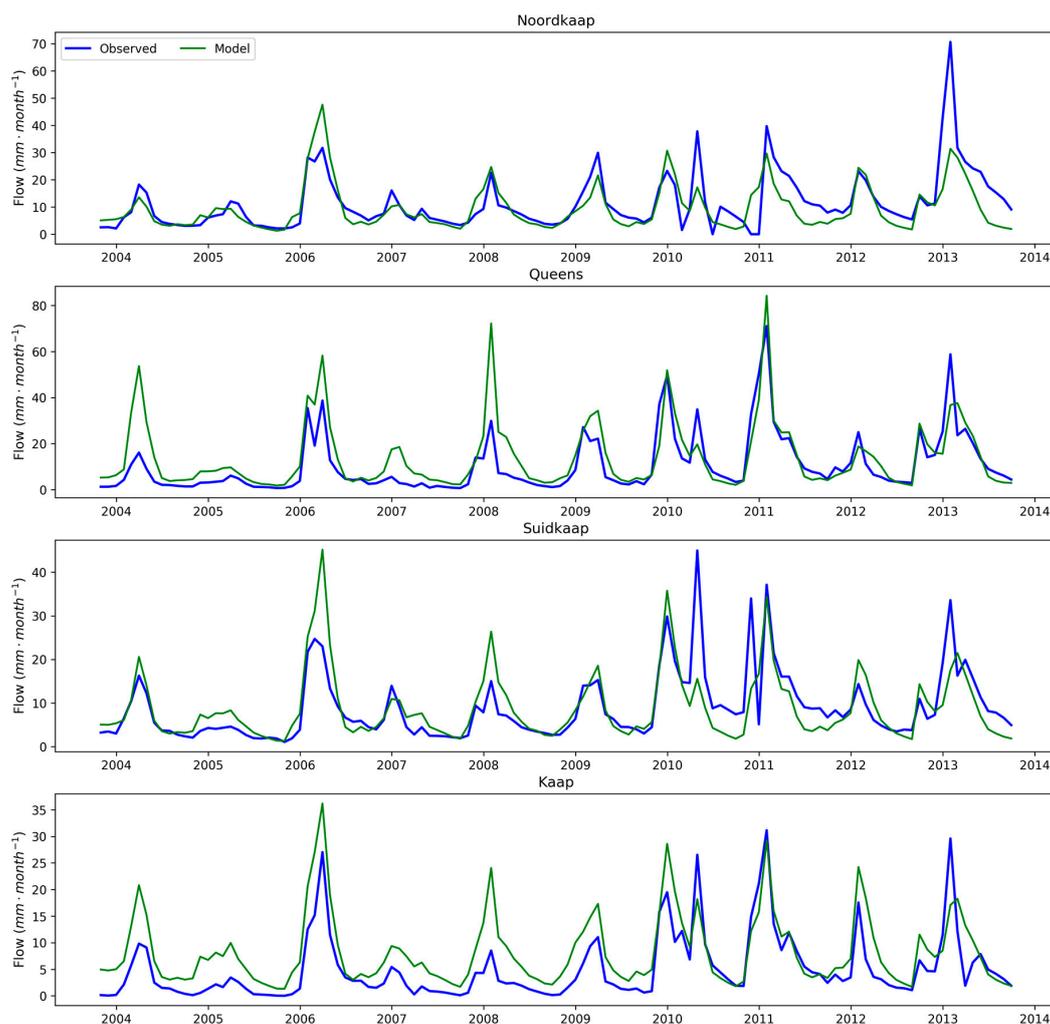

Figure 8. Monthly hydrographs for the four catchments, using results of run 64.

4. Discussion

4.1. Implications for Hydrological Process Understanding

The research provided a framework to evaluate model performance considering a range of information sources. The use of landscape features to map dominant runoff mechanisms also assisted in estimating parameters for the STREAM model. Model parameter estimates derived from this study could guide parameterisation in the case of other semi-arid catchments with similar characteristics. Furthermore, the research highlighted gaps in understanding of key hydrological processes in the study catchments, which seem typical for semi-arid areas. The computation of all evaporation fluxes and the accurate quantification of water uses is very important for the water balance estimates of semi-arid areas, given that the greatest component of the water balance is attributed to these processes. The use of information from previous hydrograph separation studies (using digital filters or water quality data) proved useful to aid in understanding the flow components and dominant flow generation processes in the catchments as well. Therefore, it is relevant for other areas to explore these types of data and information, in addition to traditional hydrological data.

Notably, the soil grids dataset that is freely available at global scale proved useful as model input.

4.2. Implications for Water Resource Management

The results reveal that there are great heterogeneities in the catchments studied. Water resource management decisions are made at the scale of the catchments studied; therefore, it is important that a better process understanding is included in current management models and tools.

We could see that especially the Queens and Kaap catchment seem to have higher levels groundwater/surface water interaction. These catchments are more impacted by water abstractions and evaporation, particularly during dry season and drier years, which is evident from the steeper shape of the flow duration curves in the low flow portion (Figure 7). Therefore, the impacts of pollution and water abstraction could be more critical for these catchments, particularly on their low flows.

The Noordkaap and Suidkaap catchments, in contrast, are dominated by subsurface runoff, as a result of deeper soils and mostly fractured granite lithology. The baseflow is steadier in these catchments, and they are likely the recharge areas of the regional groundwater body. Research using hydro-chemical and stable isotopes in the Kaap outlet over the rainy season of 2013–2014 [24] revealed that 64 to 98% of flow in the Kaap was from shallow and deep groundwater components. During wet conditions, up to 41% of total runoff was attributed to direct runoff, and strong correlations were found between antecedent precipitation conditions and direct runoff.

Saraiva Okello, et al. [25] also found high contributions of baseflow to total flow in the Kaap catchment and tributaries using calibrated recursive digital filters for hydrograph separation. They reported baseflow contributions ranging from 45 to 70%, with very high inter and intra annual variability.

Water abstractions for irrigated agriculture significantly impact the streamflow in the Kaap, particularly in the dry season and in drier years. This impact seems to be more pronounced at the Kaap outlet, and in the Queens catchment, and is attributed to the different processes that are dominant in these catchments. The Suidkaap catchment, in spite of also experiencing high streamflow reductions (mostly due to commercial forestation), still sustains significant baseflow contributions. Some water is imported to augment irrigation in the Kaap valley, but these water imports do not significantly affect the water resources of the Kaap catchment, as return flows are small.

The reduction of water quantity in the Kaap negatively affects the water quality in the catchment. Saraiva Okello, et al. [25] reported that higher loads of EC and other water quality parameters, due to reduced dilution capacity of the system, particularly during the dry season and in dry years. Deksissa, et al. [53] and Slaughter and Hughes [54] supported this finding, attributing it to the combination of abandoned mines and irrigation return flows that occur in the Kaap. Therefore,

in order to improve water quality in the Kaap and in the Crocodile catchment overall, it is important to reduce water abstractions and/or better control and restrict water pollution in the sources.

4.3. Input Uncertainty and Model Structure

The rainfall input greatly impacts the runoff generation. The regionalization of rainfall based on available stations using the IDW method could have induced a greater number of rainy days and reduced the magnitude of rain events. This results in higher interception rates, and underestimation of flow peaks. An attempt was made to compare the rainfall regionalization with the Thiessen method, and with remote sensing rainfall (Chirps), but still the IDW method provided better simulation results, likely due to the altitude correction using the MAP pattern.

Currently, the model only accounts for human activity in terms of modified land use, and not through explicit water abstractions. Improved monitoring of water use in the catchment would greatly assist in better hydrological simulation, as such information can be used to develop an irrigation routine in the model.

Another issue of uncertainty is the configuration of the groundwater reservoirs. From the shape of the observed hydrograph it appears that the recession is not linear, but rather logarithmic or another non-linear function. The research using tracers [24] revealed that there are two distinct groundwater components in the Kaap outlet, which can be indicative of different reservoirs that operate with distinct dynamics. Shallow groundwater responds quickly in rainfall events, and has the highest contribution to flow, particularly when the antecedent moisture in the catchment is already high. The other groundwater component is from deeper sources, which could be the regional groundwater, recharged in the headwaters of the catchment. This component is responsible for sustaining the baseflow during most of the year. In the months of February to April, when the catchments are already quite wet, most of the runoff is generated through direct runoff [25].

The HBV model [43,44] was set up for the catchments using similar input data, but only vegetation and elevation band zones were used to discretize the model. Precipitation, temperature and potential evaporation were used to drive the model and automatic calibration was applied to get the best performing parameter sets. Results of this model (available in the Supplementary Material) were comparable to the STREAM model outputs. The HBV was able to better simulate the water balance in the Kaap outlet, but the shape of the hydrographs and FDCs were better captured by the Stream model. This confirms that the distribution of input data, as well as the understanding of dominant runoff generation zones indeed assisted in informing model parameters and model simulation. However, both models still lack the complexity to fully capture the runoff processes occurring in the Kaap and tributaries.

The assessment of model performance should not only rely on statistical measures, but also on other aspects such as shape of the hydrograph, flow duration curves, among other hydrological signatures.

4.4. Limitations and Gaps in Process Understanding

Even though great effort was made in selecting the best available input data, and making use of all available information and process understanding, there are some limitations and gaps in process understanding.

The representation of evaporation processes in the model is simplified, but is largely consistent with the water balance. The interception process is represented by a single daily threshold, and the interception storage is not simulated dynamically.

There is also the limitation of available data for potential evaporation. It is likely that the spatial variability of evaporation is higher than that simulated, given that stations used to interpolate climatic data were located in low altitudes. Furthermore, temporal variability of potential and actual evaporation, based on the vegetation cover growing stage and physiology occurs. These aspects were not captured in the current model configurations.

An attempt was made to use remote sensing actual evaporation products, such as ALEXI [48,49], SSEBop [51,52] and CMERST [50]. However, these products have different time scales (weekly and monthly, respectively) and there were challenges with temporal interpolation of this data. We were able to downscale the ALEXI product and use it to drive the model, but the results were disappointing. Comparison at monthly and annual scale revealed that both ALEXI and CMERST products generally overestimate actual evaporation, whereas SSEBop results in underestimation. A potential way forward would be to use an ensemble product, and explore bias correction of the evaporation.

The runoff generation module of STREAM requires further development for model applications at daily time steps. The routing of runoff, the lateral flow process and the percolation were some of the gaps in the current model setup. In literature, most publications applying the STREAM model were done at monthly [34,36,55,56] and weekly [35] time steps. The reported daily model applications were also at much coarser spatial scale [57–59].

5. Conclusions

This study combined hydrological modelling with mapping of dominant runoff generation processes, and a runoff signatures approach to improve the understanding of hydrological processes and runoff generation in a semi-arid African catchment.

Several data sources, parameter input values, and model structures were explored, in order to better understand the dominant processes in the catchment. Runoff response was sensitive to parameters related to the partitioning of rainfall between unsaturated and saturated zone *cr*, as well as the thresholds for initiation of quickflow *qc*. However, the inclusion of the feedback process from the saturated zone to the unsaturated zone, termed capillary rise, proved critical to improve model simulations. This was particularly the case for the Kaap and Queens catchments, which have a more diverse geology, coarser soils and hillslope zone. The Noordkaap and Suidkaap catchments have mostly fractured granite for bedrock, have deeper soils, and more plateaus, which results in more subsurface flow occurring.

The results of model simulations were analysed using the hydrological signatures framework as well as standard goodness of fit parameters. Annual runoff, seasonal runoff, flow duration curves and hydrographs of the different model runs were compared. The annual runoff showed that these catchments have high inter-annual variability, driven mostly by the variability of rainfall. The models were able to better simulate flows in wetter years (2010–2013) than in drier years (2004–2006). The seasonal flow analysis also revealed that there is strong seasonality in the flow generation. The capillary rise process in the model required a minimum threshold of initiation of the process (*GWSmin*) to avoid that the groundwater storage would run completely dry, which is not the case in the observed series of streamflow.

FDCs were the signature that best revealed the performance of different model simulations. In most cases the model was able to capture the slope of the FDC up to Q50/60, but missed the slope during the low flows. This was the case especially in the Kaap and Queens catchments. This finding reflects the importance of improving the representation of the evaporation and groundwater-surface water interaction processes, as well as water abstractions in the model setups, to better simulate the low flows.

Finally, the daily and monthly hydrographs were compared, and goodness of fit parameters computed between observed and modelled streamflow. Even though the goodness of fit results were average (0.75 to 0.84 Pearson R^2, 0.5 to 0.66 NSE in the best simulations), visual comparison shows that the models were able to capture the flow variability well, but missed the simulation of peak flows and overestimated baseflows.

A comparison was also made between the Stream and HBV models. This yielded very similar results in terms of goodness of fit statistics for headwater catchments, but HBV performed better for the Kaap outlet. However, when the results were visually compared in terms of the various

signatures used, the Stream model better captured the hydrograph shape, and the flow duration curve, particularly for baseflows.

This study clearly shows that there is no single model setup that can represent all the processes equally well for all the catchments. Due to the differences in landscape, geology, soils and land-use and land cover, different model configurations are better suited for each catchment. However, the distribution of input data, as well as the understanding of dominant runoff generation zones assisted in informing the Stream model. There is a benefit in combining process studies and modelling. The models highlight the shortcomings in process understanding, illustrating gaps in our knowledge. Process studies in this catchment assisted in filling some of this knowledge gaps, but other shortcomings were identified. Future improvements in the model should include the explicit accounting for irrigation and water transfers.

In terms of water management, the research findings reveal that the Queens and Kaap catchments are more sensitive to pollution, particularly during low flows, due to higher level of groundwater/surface water interactions. It is important to improve monitoring of water use, given the high impact of water abstractions in the catchment. The use of remote sensing products could assist in this, but more research is required for bias correction and calibration of products. Improvement in the calculation of actual evaporation is also required, as this constitutes the major component of the water balance, and there is high uncertainty in parameters used, and different evaporation products.

Appendix A

Table A1. Physiographic and hydroclimatic characteristics of Kaap catchment and tributaries.

Streamgauge		Noordkaap X2H010	Queens X2H008	Suidkaap X2H031	Kaap Total X2H022
Sub-basin area (km^2)		126	180	262	1640
HAND zones	Wetland	8%	7%	7%	8%
	Plateau	51%	32%	61%	39%
	Hillslope	42%	61%	32%	53%
Soil texture	Clay	19%	5%	5%	4%
	Sandy clay	4%	7%	9%	4%
	Clay loam	42%	42%	25%	39%
	Sandy clay loam	34%	46%	60%	53%
	Sandy loam	0%	0%	1%	0%

Table A1. *Cont.*

Streamgauge		Noordkaap X2H010	Queens X2H008	Suidkaap X2H031	Kaap Total X2H022
Geology	Granite	97%	58%	98%	52%
	Lava	0%	28%	1%	16%
	Arenite	0%	2%	0%	9%
	Ultramafic rocks	0%	4%	0%	2%
	Quartzite	3%	0%	0%	0%
	Gneiss	0%	0%	1%	6%
	Lutaceous arenite	0%	7%	0%	14%
LULC	Forest/Woodland	14%	12%	9%	20%
	Bush/Shrub	11%	9%	17%	32%
	Grassland	7%	18%	10%	14%
	Plantations	62%	60%	52%	23%
	Water	0%	0%	0%	0%
	Wetlands	0%	1%	1%	1%
	Bare	0%	0%	0%	0%
	Agriculture: Rainfed, Planted pasture, Fallow	2%	0%	5%	3%
	Agriculture: Irrigated	3%	0%	5%	6%
	Urban and Mines	0%	0%	0%	2%
Mean Annual Runoff observed (mm/y)		149	99	120	66
Mean Annual Runoff naturalized (mm/y) [a]		216	146	210	116
Mean annual Precipitation (mm/y)		1101	1016	905	900
Mean annual potential evaporation (mm/y)		1425	1369	1451	1435

[a] Source: the WR2012 study [31].

References

1. Hughes, D.A. Hydrological modelling, process understanding and uncertainty in a southern African context: Lessons from the northern hemisphere. *Hydrol. Process.* **2016**, *30*, 2419–2431. [CrossRef]
2. Hughes, D.A.; Jewitt, G.; Mahé, G.; Mazvimavi, D.; Stisen, S. A review of aspects of hydrological sciences research in Africa over the past decade. *Hydrol. Sci. J.* **2015**, *60*, 1865–1879. [CrossRef]
3. Blöschl, G.; Sivapalan, M.; Wagener, T.; Viglione, A.; Savenije, H. *Runoff Prediction in Ungauged Basins: Synthesis Across Processes, Places and Scales*; Cambridge University Press: Cambridge, UK, 2013.
4. Hrachowitz, M.; Savenije, H.H.G.; Blöschl, G.; McDonnell, J.J.; Sivapalan, M.; Pomeroy, J.W.; Arheimer, B.; Blume, T.; Clark, M.P.; Ehret, U.; et al. A decade of Predictions in Ungauged Basins (PUB)—A review. *Hydrol. Sci. J.* **2013**, *58*, 1198–1255. [CrossRef]
5. Viglione, A.; Parajka, J.; Rogger, M.; Salinas, J.L.; Laaha, G.; Sivapalan, M.; Blöschl, G. Comparative assessment of predictions in ungauged basins—Part 3: Runoff signatures in Austria. *Hydrol. Earth Syst. Sci.* **2013**, *17*, 2263–2279. [CrossRef]
6. McMillan, H.; Westerberg, I.; Branger, F. Five guidelines for selecting hydrological signatures. *Hydrol. Process.* **2017**, *31*, 4757–4761. [CrossRef]
7. Uhlenbrook, S. An empirical approach for delineating spatial units with the same dominating runoff generation processes. *Phys. Chem. Earth Parts A/B/C* **2003**, *28*, 297–303. [CrossRef]
8. Scherrer, S.; Naef, F. A decision scheme to indicate dominant hydrological flow processes on temperate grassland. *Hydrol. Process.* **2003**, *17*, 391–401. [CrossRef]
9. Van Tol, J.J.; Le Roux, P.A.L.; Lorentz, S.A.; Hensley, M. Hydropedological Classification of South African Hillslopes. *Vadose Zone J.* **2013**, *12*. [CrossRef]
10. Müller, C.; Hellebrand, H.; Seeger, M.; Schobel, S. Identification and regionalization of dominant runoff processes a GIS-based and a statistical approach. *Hydrol. Earth Syst. Sci.* **2009**, *13*, 779–792. [CrossRef]
11. Hümann, M.; Müller, C. Improving the GIS-DRP approach by means of delineating runoff characteristics with new discharge relevant parameters. *ISPRS Int. J. Geo-Inf.* **2013**, *2*, 27–49. [CrossRef]
12. Savenije, H.H.G. HESS Opinions "Topography driven conceptual modelling (FLEX-Topo)". *Hydrol. Earth Syst. Sci.* **2010**, *14*, 2681–2692. [CrossRef]

13. Gharari, S.; Hrachowitz, M.; Fenicia, F.; Savenije, H.H.G. Hydrological landscape classification: Investigating the performance of HAND based landscape classifications in a central European meso-scale catchment. *Hydrol. Earth Syst. Sci.* **2011**, *15*, 3275–3291. [CrossRef]

14. Gao, H.; Hrachowitz, M.; Fenicia, F.; Gharari, S.; Savenije, H.H.G. Testing the realism of a topography-driven model (FLEX-Topo) in the nested catchments of the Upper Heihe, China. *Hydrol. Earth Syst. Sci.* **2014**, *18*, 1895–1915. [CrossRef]

15. van Zijl, G.; Le Roux, P. Creating a conceptual hydrological soil response map for the Stevenson Hamilton Research Supersite, Kruger National Park, South Africa. *Water SA* **2014**, *40*, 331–336. [CrossRef]

16. Soulsby, C.; Tetzlaff, D.; Rodgers, P.; Dunn, S.; Waldron, S. Runoff processes, stream water residence times and controlling landscape characteristics in a mesoscale catchment: An initial evaluation. *J. Hydrol.* **2006**, *325*, 197–221. [CrossRef]

17. Van Huyssteen, C.W. A review of advances in hydropedology for application in South Africa: Review article. *S. Afr. J. Plant Soil* **2008**, *25*, 245–253. [CrossRef]

18. Schulze, R. Hydrological characteristics and properties of soils in Southern Africa 1: Runoff response. *Water SA* **1985**, *11*, 121–128.

19. Pike, A.; Schulze, R. *AUTOSOILS: A Program to Convert ISCW Soils Attributes to Variables Usable in Hydrological Models*; University of KwaZulu-Natal, School of Bioresources Engineering and Environmental Hydrology: Pietermaritzburg, South Africa, 1995.

20. Van Zijl, G.M.; van Tol, J.J.; Riddell, E.S. Digital soil mapping for hydrological modelling. In *Digital Soil Mapping Across Paradigms, Scales and Boundaries*; Zhang, G.-L., Brus, D., Liu, F., Song, X.-D., Lagacherie, P., Eds.; Springer: Singapore, 2016; pp. 115–129.

21. 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]

22. Hengl, T.; Heuvelink, G.B.; Kempen, B.; Leenaars, J.G.; Walsh, M.G.; Shepherd, K.D.; Sila, A.; MacMillan, R.A.; de Jesus, J.M.; Tamene, L. Mapping soil properties of Africa at 250 m resolution: Random forests significantly improve current predictions. *PLoS ONE* **2015**, *10*, e0125814. [CrossRef] [PubMed]

23. Hengl, T.; Mendes de Jesus, J.; Heuvelink, G.B.M.; Ruiperez Gonzalez, M.; Kilibarda, M.; Blagotić, A.; Shangguan, W.; Wright, M.N.; Geng, X.; Bauer-Marschallinger, B.; et al. SoilGrids250m: Global gridded soil information based on machine learning. *PLoS ONE* **2017**, *12*, e0169748. [CrossRef] [PubMed]

24. Camacho Suarez, V.V.; Saraiva Okello, A.M.L.; Wenninger, J.W.; Uhlenbrook, S. Understanding runoff processes in a semi-arid environment through isotope and hydrochemical hydrograph separations. *Hydrol. Earth Syst. Sci.* **2015**, *19*, 4183–4199. [CrossRef]

25. Saraiva Okello, A.M.L.; Uhlenbrook, S.; Jewitt, G.P.W.; Masih, I.; Riddell, E.S.; Van der Zaag, P. Hydrograph separation using tracers and digital filters to quantify runoff components in a semi-arid meso-scale catchment. *Hydrol. Process.* **2018**, *32*, 1334–1350. [CrossRef]

26. Middleton, B.J.; Bailey, A.K. *Water Resources of South Africa, 2005 Study*; WRC Report No. TT 380/08; Water Research Commission: Pretoria, South Africa, 2009.

27. Sharpe, M.R.; Sohnge, A.P.; Zyl, J.S.V.; Joubert, D.K.; Mulder, M.P.; Clubley-Armstrong, A.R.; Plessis, C.P.D.; Eeden, O.R.V.; Rossouw, P.J.; Visser, D.J.L.; et al. *2530 Barberton*; Council for Geoscience, Ed.; Department of Minerals and Energy Affairs, Government Printer: Cape Town, South Africa, 1986.

28. Saraiva Okello, A.M.L.; Masih, I.; Uhlenbrook, S.; Jewitt, G.P.W.; van der Zaag, P.; Riddell, E. Drivers of spatial and temporal variability of streamflow in the Incomati River basin. *Hydrol. Earth Syst. Sci.* **2015**, *19*, 657–673. [CrossRef]

29. Van Eekelen, M.W.; Bastiaanssen, W.G.M.; Jarmain, C.; Jackson, B.; Ferreira, F.; van der Zaag, P.; Saraiva Okello, A.; Bosch, J.; Dye, P.; Bastidas-Obando, E.; et al. A novel approach to estimate direct and indirect water withdrawals from satellite measurements: A case study from the Incomati basin. *Agric. Ecosyst. Environ.* **2015**, *200*, 126–142. [CrossRef]

30. Mallory, S.; Beater, A. *Hydrology Report for the Crocodile (East) River Catchment*; Department of Water Affairs and Forestry: Pretoria, South Africa, 2009.

31. Bailey, A.; Pitman, W. *Water Resources of South Africa, 2012 Study: User Guide*; WRC Report K5/2143/1; Water Research Comission: Gezina, South Africa, 5 November 2015.

32. Rennó, C.D.; Nobre, A.D.; Cuartas, L.A.; Soares, J.V.; Hodnett, M.G.; Tomasella, J.; Waterloo, M.J. HAND, a new terrain descriptor using SRTM-DEM: Mapping terra-firme rainforest environments in Amazonia. *Remote Sens. Environ.* **2008**, *112*, 3469–3481. [CrossRef]

33. Aerts, J.C.J.H.; Kriek, M.; Schepel, M. STREAM (Spatial tools for river basins and environment and analysis of management options): 'set up and requirements'. *Phys. Chem. Earth Part B Hydrol. Oceans Atmos.* **1999**, *24*, 591–595. [CrossRef]

34. Winsemius, H.C.; Savenije, H.H.G.; Gerrits, A.M.J.; Zapreeva, E.A.; Klees, R. Comparison of two model approaches in the Zambezi river basin with regard to model reliability and identifiability. *Hydrol. Earth Syst. Sci.* **2006**, *10*, 339–352. [CrossRef]

35. Kiptala, J.K.; Mul, M.L.; Mohamed, Y.A.; van der Zaag, P. Modelling stream flow and quantifying blue water using a modified STREAM model for a heterogeneous, highly utilized and data-scarce river basin in Africa. *Hydrol. Earth Syst. Sci.* **2014**, *18*, 2287–2303. [CrossRef]

36. Gerrits, A. *Hydrological Modelling of the Zambezi Catchment for Gravity Measurements*; University of Technology: Delft, The Netherlands, 2005.

37. Dye, P.J. Response of Eucalyptus grandis trees to soil water deficits. *Tree Physiol.* **1996**, *16*, 233–238. [CrossRef] [PubMed]

38. Jewitt, G. Water and Forests. In *Encyclopedia of Hydrological Sciences*; John Wiley & Sons, Ltd.: Hoboken, NJ, USA, 2006. [CrossRef]

39. Scott, D.F.; Lesch, W. Streamflow responses to afforestation with Eucalyptus grandis and Pinus patula and to felling in the Mokobulaan experimental catchments, South Africa. *J. Hydrol.* **1997**, *199*, 360–377. [CrossRef]

40. Schulze, R.; Maharaj, M.; Warburton, M.; Gers, C.; Horan, M.; Kunz, R.; Clark, D. *South African Atlas of Climatology and Agrohydrology*; Pretoria, RSA, WRC Report; Water Research Commission: Pretoria, South Africa, 2007; Volume 1489, p. 6.

41. Sieber, A.; Uhlenbrook, S. Sensitivity analyses of a distributed catchment model to verify the model structure. *J. Hydrol.* **2005**, *310*, 216–235. [CrossRef]

42. Liu, Y.; De Smedt, F. *WetSpa Extension, a GIS-Based Hydrologic Model for Flood Prediction and Watershed Management*; Vrije Universiteit Brussel: Brussels, Belgium, 2004; pp. 1–108.

43. Bergström, S. *The HBV Model: Its Structure and Applications*; Swedish Meteorological and Hydrological Institute: Norrköping, Sweden, 1992.

44. Lindström, G.; Johansson, B.; Persson, M.; Gardelin, M.; Bergström, S. Development and test of the distributed HBV-96 hydrological model. *J. Hydrol.* **1997**, *201*, 272–288. [CrossRef]

45. Mazvimavi, D.; Meijerink, A.M.J.; Savenije, H.H.G.; Stein, A. Prediction of flow characteristics using multiple regression and neural networks: A case study in Zimbabwe. *Phys. Chem. Earth Parts A/B/C* **2005**, *30*, 639–647. [CrossRef]

46. Kling, H.; Fuchs, M.; Paulin, M. Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios. *J. Hydrol.* **2012**, *424–425*, 264–277. [CrossRef]

47. 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]

48. Anderson, M.C.; Norman, J.M.; Diak, G.R.; Kustas, W.P.; Mecikalski, J.R. A two-source time-integrated model for estimating surface fluxes using thermal infrared remote sensing. *Remote Sens. Environ.* **1997**, *60*, 195–216. [CrossRef]

49. Hain, C.R.; Mecikalski, J.R.; Anderson, M.C. Retrieval of an available water-based soil moisture proxy from thermal infrared remote sensing. Part I: Methodology and validation. *J. Hydrometeorol.* **2009**, *10*, 665–683. [CrossRef]

50. Guerschman, J.P.; Van Dijk, A.I.J.M.; Mattersdorf, G.; Beringer, J.; Hutley, L.B.; Leuning, R.; Pipunic, R.C.; Sherman, B.S. Scaling of potential evapotranspiration with MODIS data reproduces flux observations and catchment water balance observations across Australia. *J. Hydrol.* **2009**, *369*, 107–119. [CrossRef]

51. Senay, G.B.; Bohms, S.; Singh, R.K.; Gowda, P.H.; Velpuri, N.M.; Alemu, H.; Verdin, J.P. Operational evapotranspiration mapping using remote sensing and weather datasets: A new parameterization for the SSEB approach. *J. Am. Water Resour. Assoc.* **2013**, *49*, 577–591. [CrossRef]

52. Chen, M.; Senay, G.B.; Singh, R.K.; Verdin, J.P. Uncertainty analysis of the Operational Simplified Surface Energy Balance (SSEBop) model at multiple flux tower sites. *J. Hydrol.* **2016**, *536*, 384–399. [CrossRef]

53. Deksissa, T.; Ashton, P.J.; Vanrolleghem, P.A. Control options for river water quality improvement: A case study of TDS and inorganic nitrogen in the Crocodile River (South Africa). *Water SA* **2003**, *29*, 209–218. [CrossRef]

54. Slaughter, A.R.; Hughes, D.A. A simple model to separately simulate point and diffuse nutrient signatures in stream flows. *Hydrol. Res.* **2013**, *44*, 538–553. [CrossRef]

55. Bouwer, L.M.; Aerts, J.C.J.H.; Droogers, P.; Dolman, A.J. Detecting the long-term impacts from climate variability and increasing water consumption on runoff in the Krishna river basin (India). *Hydrol. Earth Syst. Sci.* **2006**, *10*, 703–713. [CrossRef]

56. Ward, P.J.; Aerts, J.C.J.H.; de Moel, H.; Renssen, H. Verification of a coupled climate-hydrological model against Holocene palaeohydrological records. *Glob. Planet. Chang.* **2007**, *57*, 283–300. [CrossRef]

57. Ward, P.J.; Renssen, H.; Aerts, J.; Van Balen, R.; Vandenberghe, J. Strong increases in flood frequency and discharge of the River Meuse over the late Holocene: Impacts of long-term anthropogenic land use change and climate variability. *Hydrol. Earth Syst. Sci.* **2008**, *12*, 159–175. [CrossRef]

58. Bouwer, L.M.; Biggs, T.W.; Aerts, J.C.J.H. Estimates of spatial variation in evaporation using satellite-derived surface temperature and a water balance model. *Hydrol. Process.* **2008**, *22*, 670–682. [CrossRef]

59. Ward, P.J.; Renssen, H.; Aerts, J.C.J.H.; Verburg, P.H. Sensitivity of discharge and flood frequency to twenty-first century and late Holocene changes in climate and land use (River Meuse, northwest Europe). *Clim. Chang.* **2011**, *106*, 179–202. [CrossRef]