Transmission loss estimation for ephemeral sand rivers in Southern Africa

Simon A. Mathias[a,*], Sim M. Reaney[b,c], Piet K. Kenabatho[d]

[a]*Department of Engineering, Durham University, Durham, UK*
[b]*Department of Geography, Durham University, Durham, UK*
[c]*Institute of Hazard, Risk and Resilience, Durham University, Durham, UK*
[d]*Department of Environmental Science, University of Botswana, Gaborone, Botswana*

Abstract

Ephemeral sand rivers represent an important water resource in Southern Africa. These rivers only flow for a few days in a year. However, much of this water infiltrates the underlying river bed sediments where it is protected from evaporation and utilized by farmers throughout the dry season. Despite their importance, little is known about how much recoverable water is annually stored within the sand. A particular difficulty concerns obtaining reliable estimates of transmission losses (the amount of water that infiltrates the river bed). The objective of this article was to develop an improved methodology for quantifying transmission loss from ephemeral sand rivers by calibrating a lumped rainfall-runoff model to observed river flow data. Fifteen years of daily river flow data were obtained from four sand rivers in Botswana, namely, Shahshe, Ntshe, Tati and Metsimotlhabe. These data were supplemented with meteorological data from AgMERRA (Ruane et al., 2015) and precipitation data from CHIRPS (Funk et al., 2015). Our simplified rainfall runoff model had four unknown parameters including a river bed infiltration factor, a surface storage capacity, a river bed storage capacity and an average river channel width. Posteriori parameter distributions were derived using a GLUE (Beven and Binley, 1992) methodology. Our study confirms that upper and lower bounds for transmission loss can be obtained by calibrating a lumped

rainfall runoff model to a single set of river flow gauging data. Transmission loss was found to represent between 55% and 85% of the total surface runoff at these locations.
Keywords: Sand river, Transmission loss, Probability distributed model, Catchment water balance, Arid zone hydrology

1. Introduction

Many ephemeral rivers in Southern Africa flow over granitic basins forming sand filled eroded channels referred to as sand rivers (Walker et al., 2018). Surface water flows occur for only a few days during the annual wet season (Shaw et al., 1994). However, much of the flowing water infiltrates into the underlying sand where it is protected from evaporation and utilized by farmers during the dry season. Unfortunately, sand rivers are increasingly under threat due to illegal sand mining and there is an urgent need to protect these important water resources (Makaba, 2017). Despite their importance, little is known about how much recoverable water is annually stored within the sand.

There have been several attempts to develop groundwater flow models to describe the water balance within such alluvial deposits (e.g. Mansell and Hussey, 2005; Love et al., 2011; Mpala et al., 2020), but a particular difficulty concerns obtaining reliable estimates of transmission loss (Hughes, 2019). Transmission loss is a commonly used term to collectively quantify reductions in streamflow associated with river bed infiltration, evaporation from the river channel, and loss to stream banks or floodplains as water travels downstream (Shanafield and Cook, 2014).

*Corresponding author. Tel.: +44 (0)1913343491, Fax: +44 (0)1913342301, E-mail address: s.a.mathias@durham.ac.uk

In some cases, transmission loss studies have focused on observing flood induced vertical distributions of moisture content within river bed sediments (Parissopoulos and Wheater, 1992; Dahan et al., 2008). Alternatively, infiltrometers have been used to measure the infiltration capacity of river bed sediments whilst the river channel is dry (Dunkerley, 2008). Such methods are useful for observing fine-scale data both spatially and temporally. However, it remains unclear how to upscale these measurements for catchment water balance.

Lange (2005) and Morin et al. (2009) determined flood induced infiltration rates by calibrating a numerical model, describing transient open channel flow coupled with river bed infiltration, to river flow rate observations. The advantage here is that estimates of infiltration rate represent a larger scale observation. However, both of the aforementioned studies focused on the Kuiseb River in Namibia, which is a special case because it has been continuously monitored by 14 river flow gauge stations. These studies were only made possible because good quality time-series flow data were available as both inputs and outputs of river reaches, such that sufficient information content was available for a meaningful model calibration. For most arid ephemeral rivers of concern, such data are not available and alternative methods are required (Love et al., 2011; Jarihani et al., 2015).

Hughes (2019) presents a simplified method of estimating transmission losses using monthly flow duration curves. The approach can be described as follows. River channel flow excluding transmission losses are simulated using the monthly time-stepping, Pitman rainfall-runoff model (Hughes, 2013). A theoretical flow duration curve, excluding transmission losses, is developed. A simple conceptual model for transmission losses is then calibrated to a monthly flow duration curve based on observed flow rates from the catchment of concern, using the Pitman model as an input. Such a method is naturally able to provide close correspondence to observed river flow rates due to the calibration process. However, models with a large number of parameters, such as the Pitman model, are subject to equifinality (Beven and Binley, 1992).

The objective of this article is to develop an improved methodology for quantifying transmission loss by calibrating a lumped rainfall-runoff model to observed river flow data. Our innovation in this context concerns minimising hydrological model complexity and maximising information content in the river flow data so as to reduce the impact of equifinality.

Most lumped rainfall-runoff models comprise a soil water accounting procedure (SWAP) and a routing function (Beven, 2011). The SWAP transforms daily precipitation and potential evapotranspiration data to surface runoff data. The routing function describes how the surface-runoff is collected across the catchment and delivered to the outlet of concern. The routing function leads to an attenuation of the surface runoff time-series.

Most long-term flow records in Southern Africa are limited to daily data. When dealing with flash floods in ephemeral sand rivers, individual events tend to have short recession periods often no longer than a couple of hours. This means there is little point in trying to apply a routing function because the flow attenuation is on the scale of the observation time-steps.

Excluding a routing function from a rainfall-runoff model has the advantage of reducing the number of model parameters requiring calibration. However, there will often be a delay between the surface runoff time-series and the observed flow rate, which means calibrating SWAP by directly comparing surface runoff data with daily river flow data is unlikely to be successful.

One option is to calibrate a SWAP to a flow duration curve derived from daily flow observations. However, there will be some rare events that exceed one day in duration and will exhibit attenuation, not incorporated in the model (due to the exclusion of a routing function). Another option is to calibrate a SWAP to monthly river flow data. However, this temporal aggregation unnecessarily gives up valuable information in the daily time-scale variations.

An alternative idea is to calibrate the SWAP to cumulative daily river flow data. This has the advantage of minimising the impact of unknown delay times whilst maintaining the value of daily variations. But the problem here is that the quality of data collection varies considerably from one year to the next, and a bad year, early on in the time-series, will massively bias the model calibration process as compared to a bad year at the end of the time-series. Our innovative approach is to calibrate a SWAP to annual cumulative daily flows (i.e., a cumulative daily flow time-series, which is reset to zero at the end of each dry season). This way, we are able to: (1) maintain daily variations, (2) minimise the effect of delay between surface runoff and river flow, and (3) eliminate the bias caused by the timing of poor data quality years.

The SWAP we utilize includes a surface runoff component and a river channel transmission loss component. Surface runoff is calculated using a one parameter form of the probability distributed model (PDM) (Moore, 2007), previously used by Mathias et al. (2016). River channel transmission loss is determined as a fixed proportion of the surface runoff rate minus the rate of river water evaporation. The underlying assumption is that river bed infiltration is a function of the water depth in the river channel, which in turn is a function of the river flow rate. The storage capacity of the underlying river bed sediments, associated with the sand river, is treated as a fixed volume storage tank, which is emptied every dry season, due to a combination of anthropogenic abstraction, seepage into underlying aquifers and evaporation. River bed infiltration is assumed only to occur when there is available storage in the river bed sediments.

The resulting model has just four unknown parameters: (1) the river bed infiltration factor, (2) the surface storage capacity for the PDM, (3) the river bed storage capacity, and (4) the average river channel width. Note that the catchment area and length of river channel network are estimated using a digital elevation model.

In this article we apply our model to 15 years of daily river flow observations from four sand rivers in Botswana. Catchment averaged precipitation and evapotranspiration are obtained using the remote sensing data packages, CHIRPS (Funk et al., 2015) and AgMERRA (Ruane et al., 2015), respectively. The rainfall runoff model is conditioned to annual cumulatively daily river flows. Posteriori probability distributions for the four model parameters are obtained using the generalised likelihood uncertainty estimation (GLUE) method (Beven, 2011). Cross correlation analysis between the four model parameters are discussed. Probability distributions are then derived for runoff and transmission loss coefficients.

2. Data and methods

2.1. Study area

This study uses observed daily river flow data, provided by the Department of Water and Sanitation of Botswana, from four sand rivers, namely, Shashe, Ntshe, Tati and Metsimotlhabe. The locations of the four gauging stations are shown in Fig. 1 and given in Table 1. Note that Shashe, Ntshe and Tati are next to each other whereas Metsimotlhabe is 500 km south west. These catchments have similar, geology, land cover, catchment areas (between 525 km^2 and 2530 km^2) and channel network densities (between 0.176 per km to 0.202 per km). All four catchments are situated on Precambrian Basement outcrops and are overlain by a combination of grassland, shrubland and savannah (Upton et al., 2018). However, the dominant soil type around Shashe, Ntshe and Tati is mostly classified as a clay Luvisol whereas around Metsimotlhabe it is classified as a sandy loam Lexisol (Nachtergaele et al., 2009).

Details concerning the catchment areas, A [L^2], and the upstream river channel lengths, L_r [L], are shown in Table 1. Catchment areas and catchment boundaries were derived from 15 arc-second HydroSHEDS drainage direction data (Lehner et al., 2008) using the D8 algorithm (Jenson & Domingue, 1988). Upstream river channel lengths were derived from 15 arc-second HydroSHEDS river network shape-files. The associated catchment boundaries, river network and digital elevation data (also from HydroSHEDS) are shown in Figs. 1a and b.

We also need estimates of the area of the river channel network, A_r [L^2], for each river catchment. The river channel network area, $A_r = L_r W_r$, where W_r [L] represents an average river channel width within the catchment. Estimates of river channel width at discrete points along the river channel networks were obtained using satellite images from Google Earth.

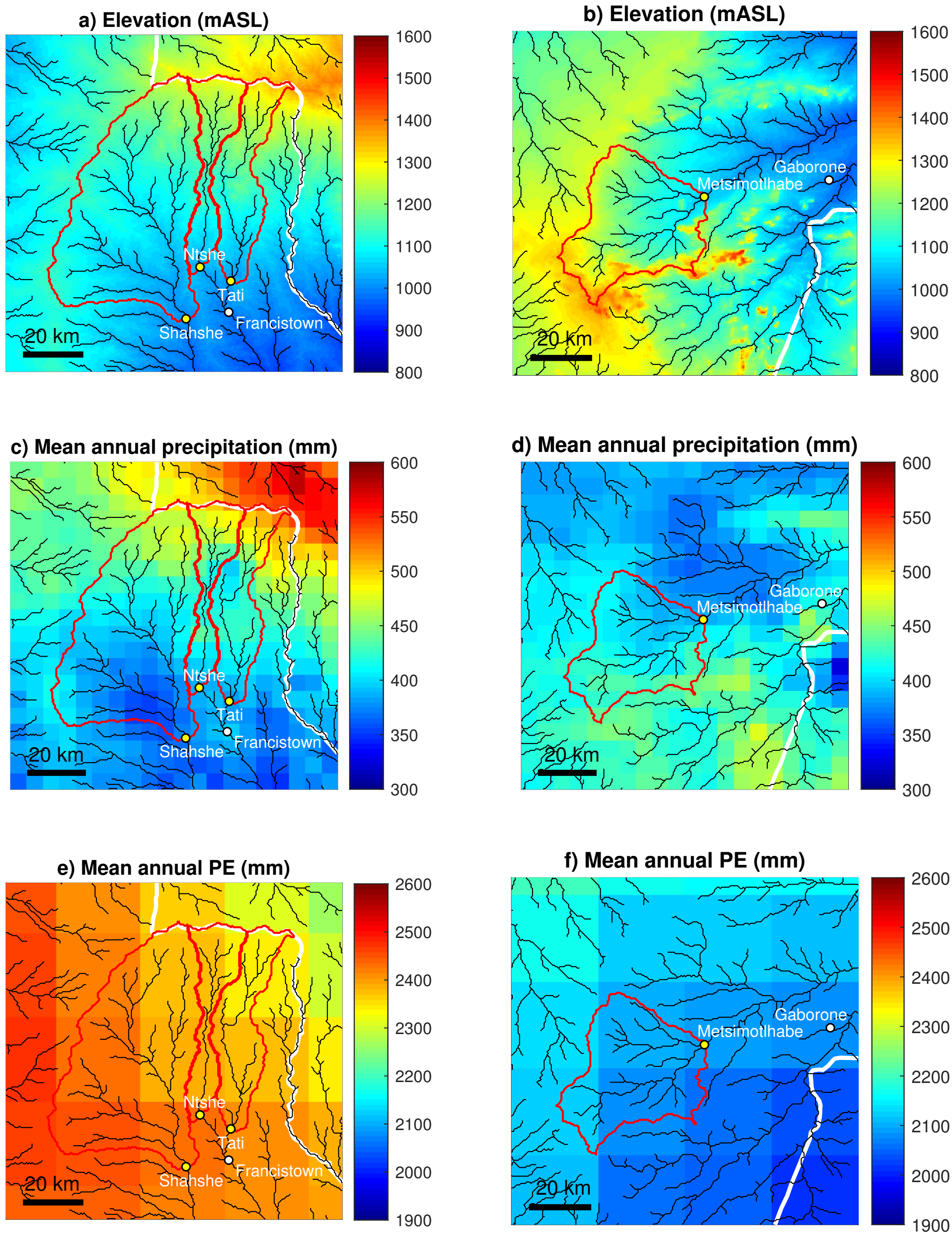

Figure 1: Maps showing the study areas. The yellow circular markers show the locations of the gauging stations used for the rivers Shanshe, Ntshe, Tati and Metsimotlhabe. The red lines show the corresponding catchment areas. The black lines show the HydroSHEDS river channel networks (Lehner et al., 2008). The white lines show the international borders between Botswana and Zimbabwe (a, c, e) and between Botswana and South Africa (b, d, f). The white circular markers show the locations of the nearest big cities. a) and b) show digital elevation data from HydroSHEDS (Lehner et al., 2008). c) and d) show mean annual precipitation, for the period of 1980 to 2010, according to CHIRPS (Funk et al., 2015). e) and f) show mean annual reference crop potential evapotranspiration (PE), for the period of 1980 to 2010, according to FAO56 (Allen et al., 1998) using meteorological data from AgMERRA (Ruane et al., 2015).

Table 1: Background details for the four river flow gauge stations. Mean hydrometeorological data are averages over the period from 1980 to 2010.

River	Shashe	Ntshe	Tati	Metsimotlhabe
Gauge location	Shahshe Mooke	Ntshe Weir	Tati Wear	Thamaga
Latitude and Longitude	(-21.20, 27.38)	(-21.04, 27.43)	(-21.08, 27.52)	(-24.68, 25.56)
Gauge number	4361	4532	4511	2421
Area (km^2)	2530	525	765	1290
River channel length (km)	447	103	154	260
Mean annual precipitation (mm)	425	461	454	429
Mean annual PE (mm)	2410	2380	2370	2100

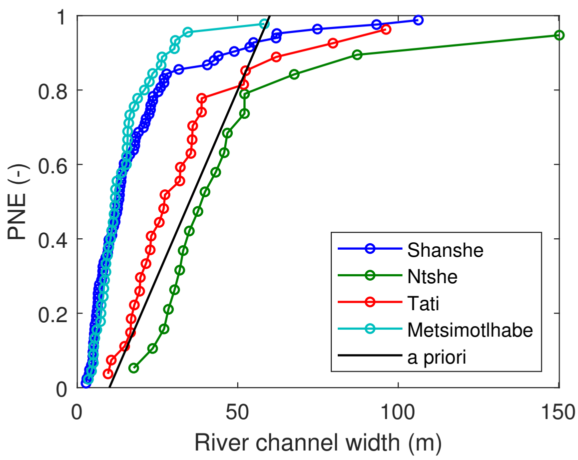

Figure 2: Cumulative distribution functions for river channel width observations within each of the river catchments studied. PNE stands for probability of non-exceedance.

125 The cumulative distribution functions (CDF) for these measurements are shown for each of the

four catchments in Fig. 2. Values range from 5 to 150 m. The median values for river channel width were 12 m for Shanshe and Metsimotlhabe, 27 m for Tati and 38 m for Ntshe. Median values for river channel width are inversely correlated with catchment area. For our subsequent analysis, the a priori distribution for W_r will be treated as a uniform random distribution ranging between 10 m and 60 m (see the line labelled "a priori" in Fig. 2).

2.2. *Hydrometeorological data*

Gridded daily mean temperature, wind speed, relative humidity and incoming shortwave radiation were acquired at 0.25° resolution from the AgMERRA data package (Ruane et al., 2015). These data were used to calculate reference crop potential evapotranspiration (PE) according to FAO56 (Allen et al., 1998). Gridded daily precipitation data were acquired at 0.05° resolution from the CHIRPS data package (Funk et al., 2015).

The AgMERRA data package was chosen because it currently provides the most comprehensive gridded meteorological dataset (in terms of providing temperature, wind speed, humidity and shortwave radiation) for the sub-Saharan Africa. AgMERRA combines reanalysis data, gauged data and satellite date (Ruane et al., 2015). AgMERRA also provides precipitation data. However, the CHIRPS precipitation data package was chosen instead due to its higher spatial resolution. CHIRPS combines gauged data and satellite data (Funk et al., 2015) and has a significant track record of use in sub-Saharan Africa (Dinku et al., 2018; Sacre Regis et al., 2020; Ngoma et al., 2021).

Four sets of catchment averaged daily precipitation and PE data were obtained by spatially averaging the gridded data over the four catchment areas. Mean annual precipitation and mean annual PE for the four river catchments are shown in Table 1. All four catchments have similar mean annual precipitation (between 425 mm and 461 mm). However, Metsimotlhabe is slightly cooler than the others and consequently has a slightly lower mean annual PE (2100 mm as compared to between 2370 mm and 2410 mm).

The spatial distributions of mean annual precipitation across the study areas are shown in Figs. 1c and d. All four catchments show an orographic effect on precipitation although this is less acute for Metsimotlhabe, where the mean annual precipitation is more uniform. The spatial distributions of mean annual PE across the study areas are shown in Figs. 1e and f. These are pretty much uniform across individual catchments although it is noted that PE is substantially less around Metsimotlhabe as compared to the other catchments because Metsimotlhabe is significantly cooler.

Plots of catchment averaged monthly PE and precipitation are shown, for each of the four catchment areas, in Figs. 3a, c, e and g. It can be seen that monthly PE is almost always considerable greater than monthly precipitation and that precipitation mostly only occurs between the months of October and April. Furthermore, the highest values of PE coincide with the highest values of precipitation.

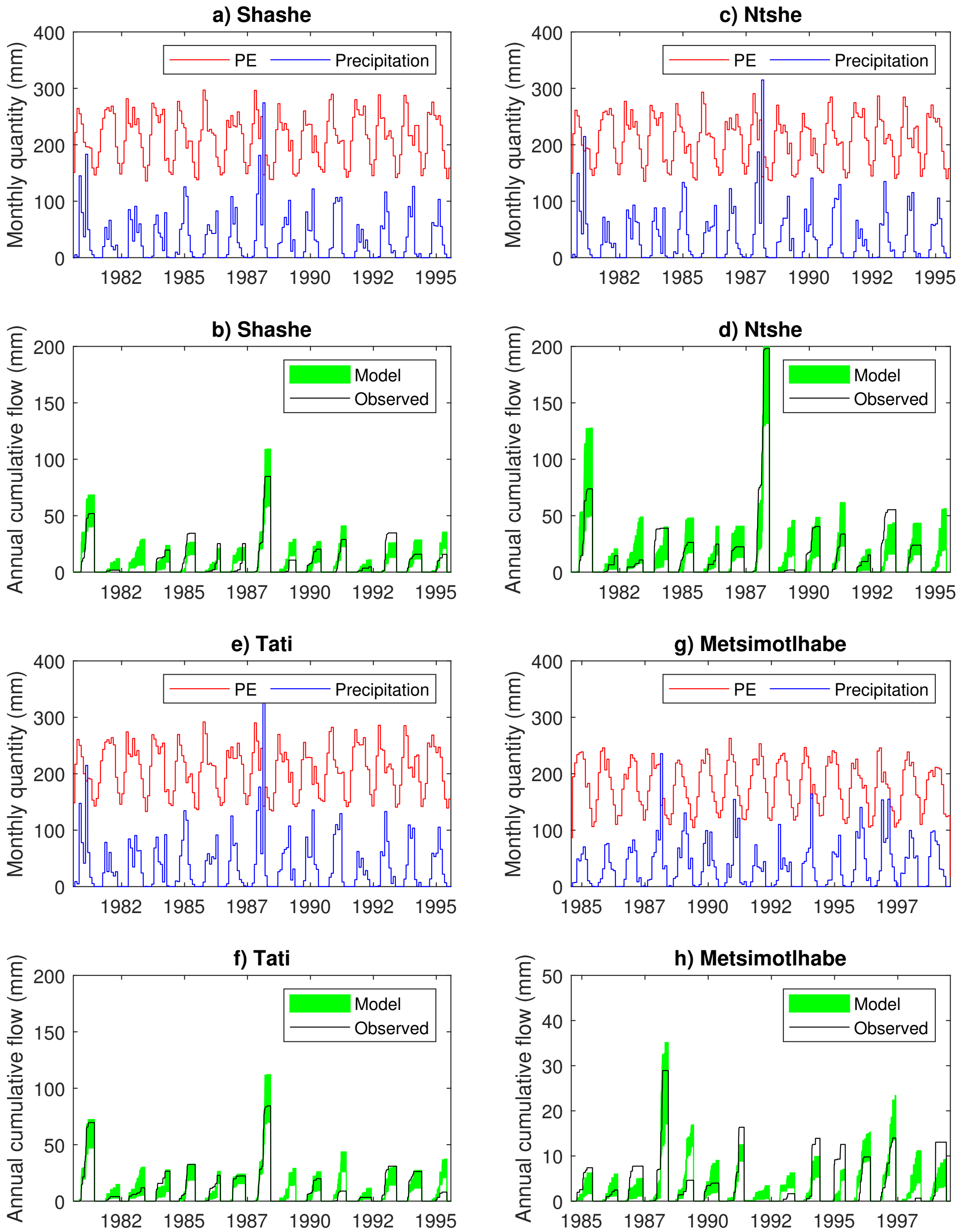

Figure 3: Catchment averaged monthly reference crop potential evapotranspiration (PE) and precipitation (a, c, e, g) along with corresponding annual cumulative daily river flows, normalised by dividing by catchment area, (b, d, f, h) for the four catchment areas upstream of the gauging station locations listed in Table 1. The black lines are the observed river flow data. The green shaded areas are envelopes obtained using the posteriori parameter distributions associated with Model 3 (see Fig. 5).

Corresponding annual cumulative daily river flows, normalised by dividing by catchment area, are shown in Figs. 3b, d, f and h. Annual cumulative daily river flows are obtained by taking the cumulative river flows starting from 1st of June in each year, such as to capture the entire wet season in a single year. As with precipitation, non-zero river flows only occur between the months of October and April. The magnitude of normalised river flows in Shashe and Tati are quite similar. The normalised river flows in Ntshe are generally larger, potentially suggesting less transmission loss. It is also noted that Ntshe is the smallest of the three catchments. Normalised river flows in Metsimotlhabe are significantly lower than in the other three catchments. This could be indicative of higher transmission losses. Note that Metsimotlhabe has a slightly lower mean annual PE as compared to the other three catchments, so this reduced river flow effect is not due to an increase in dryness.

2.3. Hydrological model

2.3.1. Catchment water balance model

The catchment is assumed to be comprised of four compartments: (1) the vegetative canopy, (2) the soil outside of the river channel, (3) the river channel, and (4) the underlying river bed sediments (Fig. 4).

The water balance for the vegetative canopy takes the form

$$\frac{dS_c}{dt} = q_p - E_c - q_c \tag{1}$$

where S_c [L] is the volume of water stored within the canopy per unit area of canopy covered land, t [T] is time, q_p [LT^{-1}] is the precipitation rate, E_c [LT^{-1}] is the rate at which water is evaporated from the canopy and q_c [LT^{-1}] is the combined rate of throughflow and stemflow of water that reaches the soil surface.

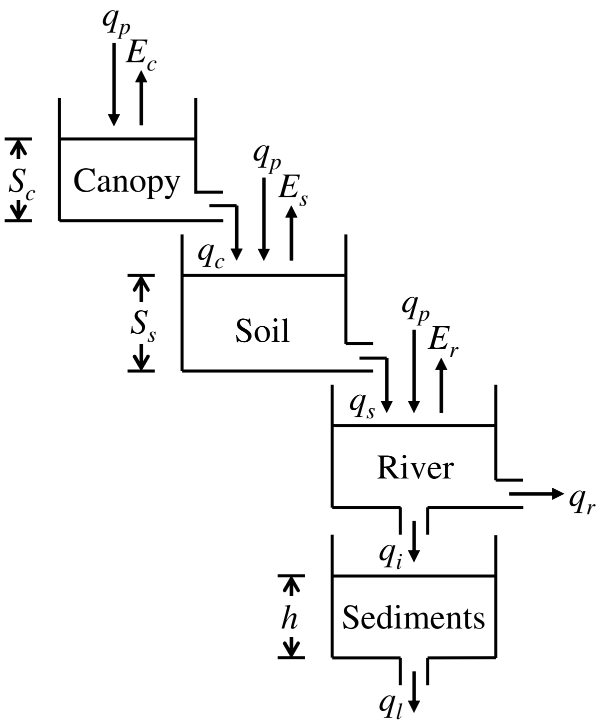

Figure 4: Schematic of hydrological model.

The water balance for the soil outside the river channel takes the form

$$\frac{dS_s}{dt} = \left(1 - \frac{A_c}{A_s}\right) q_p + \frac{A_c}{A_s} q_c - E_s - q_s \tag{2}$$

where S_s [L] is the volume of water stored within the soil per unit area of soil outside the river channel, A_c [L^2] is the area of land covered by vegetative canopy within the catchment, A_s [L^2] is the area of the catchment excluding the river channel, E_s [LT^{-1}] is the rate of evapotranspiration from the soil and q_s [LT^{-1}] is the rate of surface runoff. The bedrock underlying the soil is assumed to be impermeable.

The water balance for the river channel is assumed to be instantaneous and takes the form

$$q_r = \frac{A_s q_s + A_r(q_p - E_r)}{A} - q_i \tag{3}$$

where q_r [LT^{-1}] is the river flow rate per unit area of catchment, A [L^2] is the total area of catchment, A_r [L^2] is the area of the river channel network, E_r [LT^{-1}] is the rate of evaporation from the river channel when open water is present (assumed to be when $q_r > 0$) and q_i [LT^{-1}] is the rate of infiltration, per unit area of catchment, into the underlying river bed sediments, that form the river bed.

Solving Eq. (1) for q_c and then substituting this into Eq. (2) leads to

$$\frac{dS}{dt} = q_p - E_a - q_s \tag{4}$$

where

$$S = S_s + \left(\frac{A_c}{A_s}\right) S_c \tag{5}$$

and

$$E_a = E_s + \left(\frac{A_c}{A_s}\right) E_c \tag{6}$$

2.3.2. Actual evapotranspiration

The evaporation from the canopy and the river channel both represent examples of open-water evaporation. In contrast, the evapotranspiration from the soil represents the combined process of evaporation from the soil pores and transpiration from vegetation utilizing the soil water. This latter process is assumed to be represented by the aforementioned FAO56 reference crop PE, E_0 [LT^{-1}]. Allen et al. (1998, Table 12) suggest that for shallow open water systems (< 2 m depth), the ratio of open-water evaporation to reference crop evapotranspiration is 1.05. Therefore, for simplicity we will assume that open-water evaporation is the same as the reference crop evapotranspiration. Similar to Mathias et al. (2016), it is therefore assumed that:

$$E_c = \begin{cases} E_0, & S_c > 0 \\ 0, & S_c = 0 \end{cases} \tag{7}$$

$$E_s = \begin{cases} E_0 - \left(\frac{A_c}{A_s}\right) E_c, & S_s > 0 \\ 0, & S_s = 0 \end{cases} \tag{8}$$

$$E_r = \begin{cases} E_0, & A_r(E_0 - q_p) \leq A_s q_s \\ q_p + \left(\frac{A_s}{A_r}\right) q_s, & A_r(E_0 - q_p) > A_s q_s \end{cases} \tag{9}$$

If we assume that the soil always has some moisture when the canopy has some moisture, Eq. (6) then leads to

$$E_a = \begin{cases} E_0, & S > 0 \\ 0, & S = 0 \end{cases} \tag{10}$$

2.3.3. *Spatially uniform storage capacities*

Let c_c [L] and c_s [L] be spatially uniform storage capacities for the canopy and soil, respectively, such that $S_c \in [0, c_c]$ and $S_s \in [0, c_s]$ and

$$q_c = \begin{cases} q_p, & S_c = c_c \\ 0, & S_c < c_c \end{cases} \tag{11}$$

and

$$q_s = \begin{cases} (1 - \frac{A_c}{A_s})q_p + \frac{A_c}{A_s}q_c, & S_s = c_s \\ 0, & S_s < c_s \end{cases} \tag{12}$$

If we further assume that the soil never becomes waterlogged until the canopy capacity is exceeded, it can be further stated that

$$q_s = \begin{cases} q_p, & S = c \\ 0, & S < c \end{cases} \tag{13}$$

where

$$c = c_s + \left(\frac{A_c}{A_s}\right)c_c \tag{14}$$

In this way, the canopy and soil water conservation statements have been combined into a single water storage statement described by Eqs. (4), (6) and (13), defined by just a single descriptive parameter, c [L], which represents the combined storage capacity of the canopy and the soil.

Love et al. (2010) recently highlighted the importance of canopy interception for water balance in a catchment study from the Zimbabwe portion of the Limpopo basin. However, the mathematical analysis above shows that it is not possible to distinguish between canopy evaporation and soil evapotranspiration at a catchment scale when the capacity parameters, c_s and c_c, are unknown.

2.3.4. Probability distributed model

At any given time, an area within the catchment, A_w [L^2], contains waterlogged soil and canopy such that additional precipitation leads to the generation of surface runoff. Moore (2007) considers the storage capacity at any point within the catchment, c [L], to be a random variable defined by a probability density function, $f(c)$ [L^{-1}]. Let C [L] be the maximum value of c observed within the area, A_w. It can then be stated that $A_w = F(C)A_s$ where $F(C)$ [-] is the probability of c not exceeding, C, defined as

$$F(C) = \int_0^C f(c)dc \tag{15}$$

It can be further asserted that

$$q_s = F(C)q_p \tag{16}$$

Importantly, the relationship between the depth of water stored in the catchment, S, and the cumulative distribution function, $F(c)$, is (Moore, 2007; Mathias et al., 2016)

$$S = \int_0^C [1 - F(c)]dc \tag{17}$$

the significance of which being that, once the mathematical form of $f(c)$ is defined, the rate, q_s, can be determined explicitly from S.

Following Mathias et al. (2016), we adopt the one parameter exponential distribution function

$$f(c) = \frac{1}{S_{\max}} \exp\left(-\frac{c}{S_{\max}}\right) \tag{18}$$

where $S_{\max}$ [L] represents an empirical scaling parameter for c. Substituting Eq. (18) into Eq. (15) and then Eq. (17) leads to the result (Mathias et al., 2016)

$$F(C) = \frac{S}{S_{\max}} \tag{19}$$

and therefore

$$q_s = \begin{cases} q_p, & S = S_{\max} \\ \left(\frac{S}{S_{\max}}\right) q_p, & S < S_{\max} \end{cases} \tag{20}$$

It can also be seen that $S_{\max}$ represents the maximum possible value of S. $S_{\max}$ is hereafter referred to as the surface storage capacity.

2.3.5. Proportional loss model for river bed infiltration

It is further assumed that the rate of river bed infiltration, q_i, can be determined using a constant proportional loss coefficient, w [-], hereafter referred to as the river bed infiltration factor, such that

$$q_i = \begin{cases} 0, & h = h_{\max} \\ w q_m, & h < h_{\max} \end{cases} \tag{21}$$

where h [L] is the volume of water, per unit area of catchment, stored in the underlying river bed sediments, $h_{\max}$ [L] represents the river bed storage capacity and q_m is the rate of net input of water into the river channel found from (consider again Eq. (3))

$$q_m = \frac{A_s q_s + A_r(q_p - E_r)}{A} \tag{22}$$

The river bed storage equation takes the form

$$\frac{dh}{dt} = q_i - q_l \tag{23}$$

where q_l [LT^{-1}] represents the rate at which water is lost from the river bed sediments (per unit area of catchment) due to a combination of anthropogenic abstraction, seepage into underlying aquifers and additional evaporation.

The river flow rate per unit area of catchment is found from

$$q_r = q_m - q_i \tag{24}$$

Unfortunately we do not have information about how much water is abstracted from the sediments and how much is likely to seep into underlying aquifers. There is also uncertainty about the so-called extinction depth, beyond which evaporation from bare soils becomes significantly reduced (Gong et al., 2020). Therefore, for simplicity, q_l is fixed to zero and h is set to zero at the beginning of each hydrological year (i.e., 1st June). The assumption here is that river bed sediments are completely dry at the end of each dry season, which is pretty much true in the four Botswanan river catchments of concern.

A shortcoming of the above approach is that river bed sediments may reach full water capacity too early on within a wet season, due to q_l not being accounted for until the end of the dry season. However, for some sand rivers, the real-time losses associated with q_l are likely to be sufficiently high such that river bed infiltration is unlikely to be river bed storage capacity limited. Where this is the case, modelled river flow rates will be insensitive to the value of h_{max} providing h_{max} is sufficiently large. The above simplified modelling approach can therefore be used to help determine where river bed infiltration is not river bed storage capacity limited.

2.3.6. Three different model structures with varying complexity

The above set of equations can be used to derive three different model structures of varying complexity.

Model 1 assumes that river channel transmission loss represent an indistinguishable part of the losses accounted for within the PDM such that

$$q_r = q_s \tag{25}$$

Model 2 assumes that river channel transmission losses are distinguishable from surface infiltration outside of the sand river network but that water stored within the river bed sediments never gets close to maximum capacity (i.e., it can be assumed that $h_{\max} \to \infty$). In this way, $q_i = wq_m$ and

$$q_r = (1 - w)q_m \tag{26}$$

Model 3 is the same as Model 2 except the river bed sediments have a finite capacity such that q_i and q_r must be found from Eqs. (21) and Eq. (24), respectively. Model 3 represents the most complete model adopted in this study.

2.4. Model implementation and parameter estimation

The hydrological model used for our analysis is completely defined by Eqs. (4), (9), (10) and Eqs. (20) to (24). The differential equations are solved using an Euler explicit time-stepping scheme with a daily time-step, as described in Appendix B of Mathias et al. (2016). Daily reference crop PE, E_0, and daily precipitation, q_p, are treated as catchment averaged quantities and derived as described in Section 2.2. The area of the catchment, A, and the length of the river channel, L_r, have been determined using the spatial data described in Section 2.2 and are provided in Table 1.

Remaining unknown parameters include the river bed infiltration factor, w, the surface storage capacity, $S_{\max}$, the river bed storage capacity, $h_{\max}$, and the river channel width, W_r. This latter parameter is used to determine the area of the river channel network, i.e., $A_r = L_r W_r$.

A priori probability distributions for these four unknown parameters are assumed to be bounded uniform random distributions. The w parameter is assumed to be between 0 and 1, which represents the entire physical range. The $S_{\max}$ parameter is assumed to be between 0 and 300 mm on the basis that the maximum $S_{\max}$ value previously observed in 120 UK catchments by Mathias et al. (2016) was 230 mm. The $h_{\max}$ parameter is assumed to be between 0 and 500 mm on the basis that we are less certain about $h_{\max}$ as compared to $S_{\max}$. The W_r parameter is assumed to be between 10 m and 60 m on the basis of the observed median values for river channel width observed in Fig.

2.

Posteriori probability distributions for these four unknown parameters, for each of the four rivers are acquired using a GLUE methodology (Beven and Binley, 1992) as follows:

1. Values of w, $S_{\max}$, $h_{\max}$ and W_r are randomly sampled from the specified a priori probability distributions.
2. Theoretical river flow data, q_r, are determined using each of the three model structures (see Section 2.3.6).
3. The Nash Sutcliffe Efficiency (NSE) (Nash and Sutcliffe, 1970) comparing the annual cumulative daily river flows from the modelled and observed data is calculated, using each of three models. Data where the observed annual cumulative daily flow are zero are excluded.
4. Steps 1 to 3 are repeated 20,000 times. (Some preliminary analysis was performed with 100,000 realisations and the results were found not to be significantly different to when only 20,000 realisations were performed.)
5. Posteriori parameter distributions for each of the unknown parameters are obtained by only retaining the top 1% of the realisations, in terms of highest NSE, for each of the three model structures.

Note that Model 1 only has one unknown parameter, $S_{\max}$. Model 2 has three unknown parameters, w, $S_{\max}$ and W_r. Model 3 has four unknown paramters, w, $S_{\max}$, $h_{\max}$ and W_r.

3. Results

Table 2 shows the range of NSE values obtained from the top 1% realisations for each of the three model structures and for each of the four rivers. For all four rivers, it can be seen that Model

2 is significantly better at matching the observed data as compared to Model 1. Note that Model 1 leads to negative NSE values for Metsimotlhabe because of the limited range of $S_{\max}$ considered for the a priori distributions. The improvement provided by Model 3 as compared to Model 2 is much less significant with the exception of Ntshe.

A comparison between the envelope of annual cumulative flows from the top 1% realisations for Model 3 with the observed annual cumulative flows is shown for each river in Figs. 3c, d, f and h. As indicated by the NSE values, the models are much better at capturing the data for Shashe, Ntshe and Tati as compared to Metsimotlhabe. This difference is likely due to the quality of flow measurement being better at these three gauging stations. Note that the models predict non-zero flows during years where no flows are recorded. These zero flow periods are in fact due to an absence of data records for those years. These years are not included in the NSE calculation.

Table 2: Nash Sutcliffe Efficiency (NSE) range for the top 1% realisations for each of the three model structures and for each of the four rivers.

	Shashe		Ntshe		Tati		Metsimotlhabe	
	Min.	Max.	Min.	Max.	Min.	Max.	Min.	Max.
Model 1	0.630	0.630	0.729	0.729	0.606	0.606	-0.538	-0.507
Model 2	0.770	0.830	0.734	0.760	0.783	0.837	0.439	0.465
Model 3	0.750	0.834	0.741	0.850	0.763	0.837	0.437	0.486

3.1. Posteriori parameter distributions

Fig. 5 shows the a priori and posteriori cumulative distribution functions (CDF) for the four unknown parameters, based on the top 1% best performing realisations, for each of the four rivers.

Fig. 5a shows the CDF for the river bed infiltration factor, w, using Model 2 (dashed lines) and Model 3 (solid lines). Recall that Model 2 assumed an infinite $h_{\max}$ (river bed storage capacity) whereas Model 3 assumed a finite $h_{\max}$.

The posteriori w distributions for Shashe, Tati and Metsimotlhabe are quite similar, suggesting a range of w values from 0.55 to 0.85. Furthermore, the CDFs from Model 2 and Model 3 closely follow each other for these rivers, suggesting that transmission loss is not $h_{\max}$ limited for these cases.

In contrast, a much wider range of w values are possible for Ntshe, with some Model 3 values lower than 0.2. Furthermore, there is a wider discrepancy between the CDFs for Model 2 and Model 3 for this river, suggesting that transmission loss is likely to be $h_{\max}$ limited in this case. It is noted that the Ntshe catchment is the smaller of the four catchments (see Table 1) and that it produces more water per unit area of catchment as compared to the other three rivers (compare Figs. 3b, d, f and h), supporting the idea that there is less transmission loss here. Perhaps smaller catchments have less of an opportunity to erode out deeper sand river channels in their granitic basins. Interestingly, Ntshe also has a larger median river channel width as compared to the other rivers (see Fig. 2).

Fig. 5b shows the CDF for the surface storage capacity parameter, $S_{\max}$. The dash-dot lines are for Model 1 where $q_r = q_s$. It can be seen that conditioning the PDM directly to the observed river flow data leads to very high values for $S_{\max}$, all of which are greater than 125 mm. In fact, the CDF for Metsimotlhabe suggests that $S_{\max}$ is greater than 300 mm for that river, leading to the negative values of NSE in Table 2.

Mathias et al. (2016) applied the same single parameter PDM to 120 different river catchments in the UK. In their study they found that S_{max} was always less than 230 mm and mostly less than 100 mm. The reason why S_{max} is coming out so high for Model 1 in this current context is that the calibration process is compensating for transmission loss occurring within the river channel. When transmission loss is explicitly accounted for (i.e., as in Model 2 and Model 3), the CDFs for S_{max} give a 95% probability that $S_{\text{max}} < 115$ mm for Metsimotlhabe and a 95% probability that $S_{\text{max}} < 51$ mm for Tati (see solid lines for Model 3 in Fig. 5b). The S_{max} values are even less for Shashe and Ntshe.

It is also noted that the S_{max} CDFs for Model 2 and Model 3 closely correspond with each other for Shashe, Tati and Metsimotlhabe, suggesting that the surface runoff component of the river flows is also not h_{max} limited for these rivers. However, for Ntshe, there is a notable difference between Model 2 and Model 3 CDFs when $S_{\text{max}} > 36$ mm.

The idea that transmission loss is not h_{max} limited, for Shashe, Tati and Metsimotlhabe, is further corroborated by the posteriori distributions for h_{max}, given in Fig. 5c. With the exception of Ntshe, these CDFs are all approximately uniformly distributed above a lower bound value of 90 mm, suggesting that as long as h_{max} is set greater than 90 mm, the river flow response is insensitive to this parameter. In contrast, a more complicated distribution is noted for Ntshe, where transmission loss is likely to be h_{max} limited.

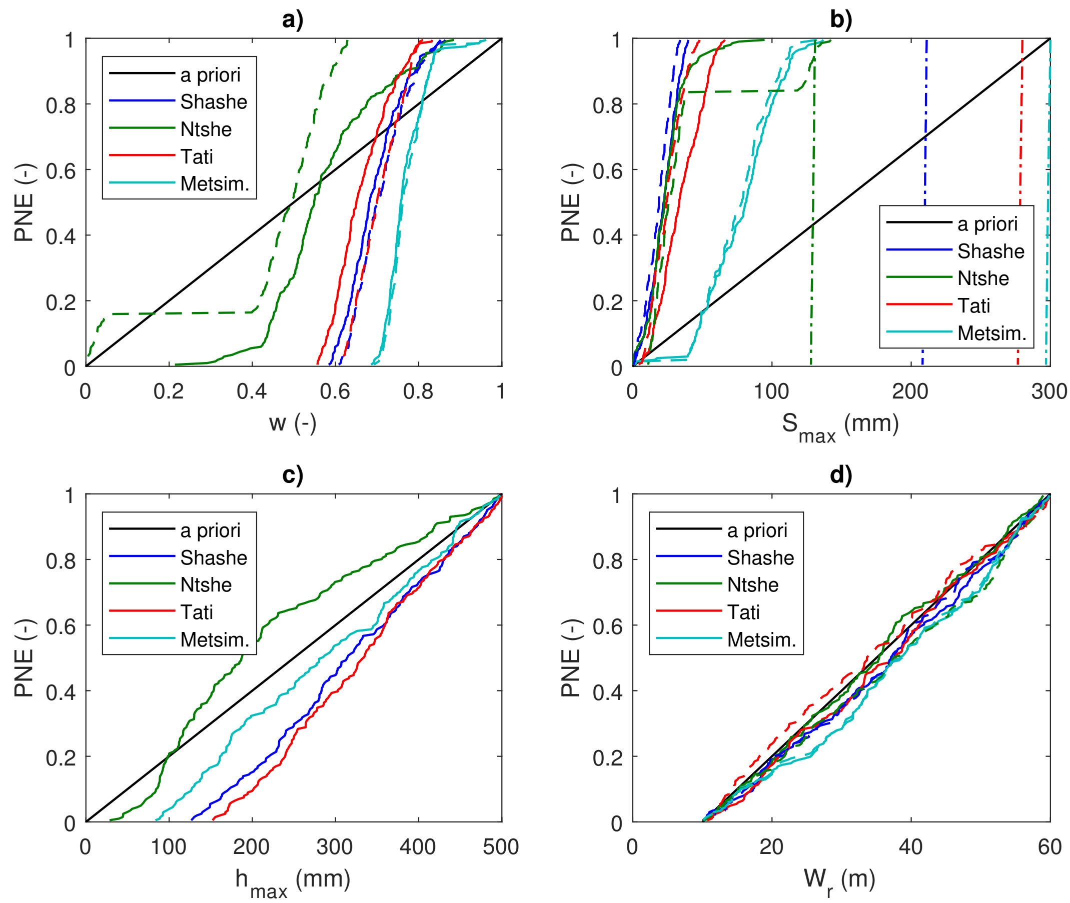

Figure 5: A priori cumulative distribution functions (CDF) and posteriori CDF for each of the four unknown parameters and each of the four rivers, Shashe, Ntshe, Tati and Metsimotlhabe. PNE stands for probability of non-exceedance. The dash-dot lines, dashed line and solid lines are from Models 1, 2 and 3, respectively (see Section 2.3.6). a) Shows river bed infiltration factor, w. b) shows surface storage capacity, S_{max}. c) shows river bed storage capacity, h_{max}. d) shows river channel width, W_r.

Fig. 5d shows the CDFs for the river channel width, W_r. It can be seen that there is very little deviation between the a priori and posteriori distributions suggesting that river flow values are insensitive this property. The only thing that W_r controls is the open water evaporation from the river channel in Eq. (22), through $A_r = L_r W_r$. Because $A_r \ll A_s$, this open water evaporation term is not that significant, and hence the river flow is largely insensitive to W_r.

3.2. Parameter identifiability and cross-correlation

The Kolmogorov–Smirnov (KS) statistic (Ang and Tang, 1975, p. 277-280) measures the maximum distance between the a priori and posteriori CDFs and provides a simple method of comparing the identifiability of each of the four unknown parameters in Model 3. A parameter with a higher KS statistic is more identifiable than a parameter with a lower KS statistic. The KS statistics for each of four parameters are presented in Table 3. For each of the four rivers, it is found that $S_{\max}$ is the most identifiable parameter, followed by w and then $h_{\max}$. The W_r parameter has a very low KS statistic, quantifying the fact that river flows are insensitive to this parameter.

Table 3: Kolmogorov–Smirnov statistics from using Model 3 for each model parameter and for each of the four rivers.

	Shashe	Ntshe	Tati	Metsimotlhabe
w (-)	0.576	0.362	0.552	0.677
$S_{\max}$ (mm)	0.862	0.785	0.774	0.568
$h_{\max}$ (mm)	0.260	0.171	0.310	0.167
W_r (m)	0.062	0.053	0.058	0.123

Table 4 shows the correlation coefficients between the four model parameters within their posteriori distributions. Mostly the parameters are not that correlated with the exception of w, which is strongly negatively correlated with $S_{\max}$. The reason for this is as follows. The higher the value of S_{max}, the more the precipitation is stored within the soil and evaporated, the less the water runs off into the river channel network, the less the water needs to be infiltrated into the river bed sediments to match the modelled river flow with observed flow rates, and the lower the necessary value for w. It is also noted that there is a moderate negative correlation between S_{max} and h_{max} for Ntshe.

Fig. 6a shows plots of w against S_{max} within the Model 3 posteriori distributions for each of the four rivers. Despite the strong negative correlation between w and S_{max}, the model calibration process is able to identify upper and lower bounds for both parameters. The results show that the value of w is between 0.55 and 0.85 for Shashe, Ntshe and Metsimotlhabe. Values of S_{max} range between 40 mm and 135 mm for Metsimotlhabe and between 0 and 65 mm for Shashe and Tati. The results are more complicated for Ntshe due to additional correlation between S_{max} and h_{max} .

Table 4: Correlation coefficients for each model parameter pair when using Model 3 for each of the four rivers. * indicates that a correlation is statistically significant.

	Shashe	Ntshe	Tati	Metsimotlhabe
w and S_{max}	-0.851*	-0.762*	-0.896*	-0.929*
w and h_{max}	0.105	0.021	0.119	0.110
w and W_r	0.098	0.029	-0.080	-0.164*
S_{max} and h_{max}	-0.159	-0.401*	-0.074	-0.113
S_{max} and W_r	-0.125	0.072	0.075	0.228*
h_{max} and W_r	0.060	-0.098	-0.030	0.106

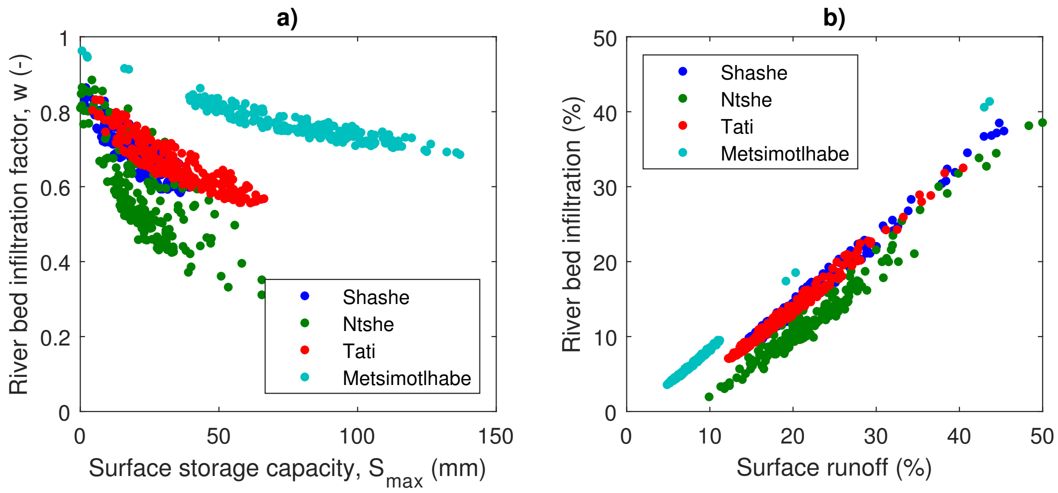

Figure 6: a) Plots of river bed infiltration factor, w, against surface storage capacity, $S_{\max}$, for each of the Model 3 posteriori distributions. b) Plots of river bed infiltration coefficient against surface runoff coefficient for each of the Model 3 posteriori distributions.

3.3. Hydrological components as percentages of total precipitation

Further insight into the hydrological processes taking place can be obtained by studying quantities of water associated with different hydrological components as a percentage of total precipitation.

Fig. 7a shows the Model 3 posteriori distributions for total surface runoff as a percentage of total rainfall for the entire study period, hereafter referred to as the surface runoff coefficient (SRC). SRC is found to be between 5% and 12% for Metsimotlhabe and between 10% and 50% for the other three rivers. SRC could be lower for Metsimotlhabe, as compared to the other three catchments, on account of its sandier soil classification.

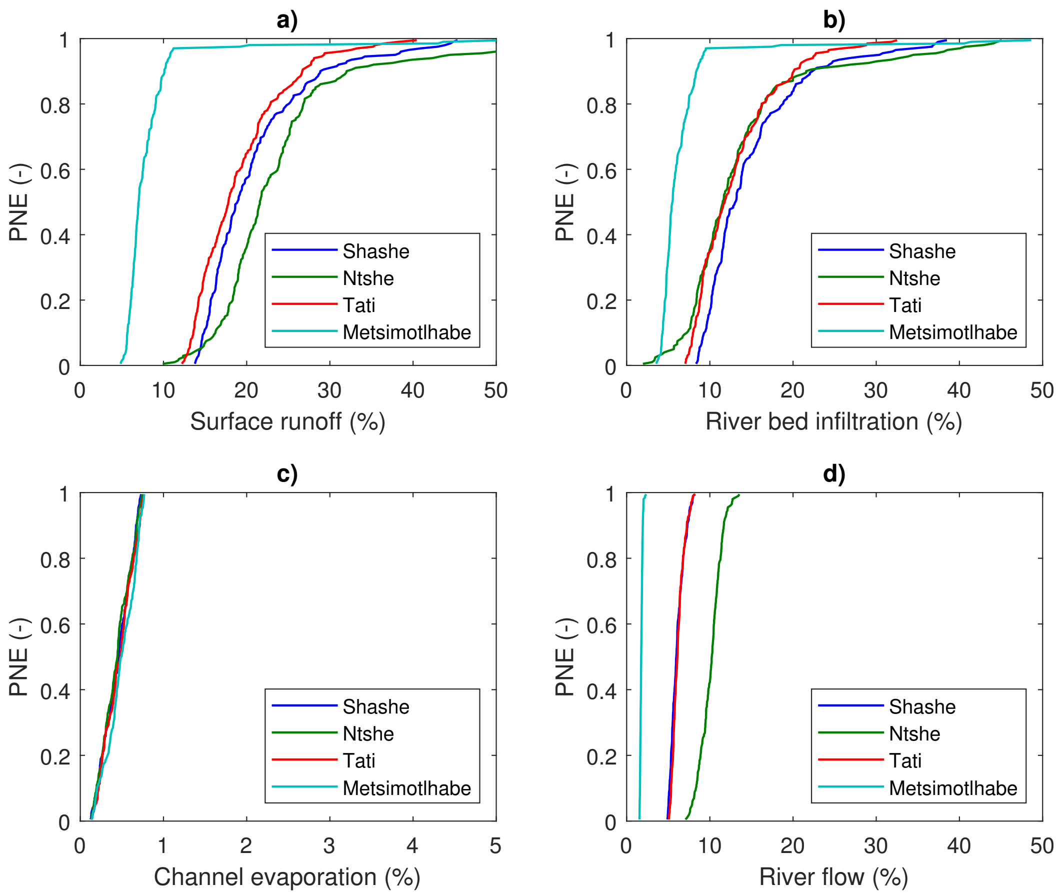

Figure 7: Model 3 posteriori distributions for: a) total surface runoff, q_s, as a percentage of total precipitation; b) total river bed infiltration, q_i, as a percentage of total precipitation; c) total open water evaporation from the river channel, $(A_r/A)E_r$, as a percentage of total precipitation; d) total river flow, q_r, as a percentage of total precipitation.

FAO (1995) present regional scale runoff coefficients for Africa based on the ratio of total "internal renewable water resource" (IRWR) to total precipitation. IRWR is defined as the "average annual flow of rivers and groundwater generated from endogenous precipitation". For Botswana, FAO (1995) cite a total IRWR of 2.9 km^3 per year and a total precipitation of 233.2 km^3 per year, which yields a runoff coefficient of 12.4%. This is much lower than the upper bound SRC of 50%, estimated for Shahshe, Ntshe and Tati. The reason for this is that SRC does not take into account the transmission loss that occurs in the river channel network.

Interestingly, Parida et al. (2006) estimated runoff coefficients for the Notwane river catchment near Gaborone, Botswana. Notwane is not classified as a sand river and is expected to have a relatively low transmission loss. Therefore, their runoff coefficients should be more comparable with our SRC. They observed annual runoff coefficients ranging from 35% to 56%.

Fig. 7b shows posteriori distributions for total river bed infiltration as a percentage of total rainfall, hereafter referred to as the river bed infiltration coefficient (RBIC). RBIC is found to be between 4% and 10% for Metsimotlhabe and between 7% and 45% for the other three rivers. RBIC is lower for Metsimotlhabe on account of its lower SRC. In fact RBIC is strongly correlated with SRC for all of the catchments (see Fig. 6b). Nevertheless, the Monte Carlo simulation has provided a useful set of bounds for this property.

Fig. 7c shows posteriori distributions for total open water evaporation from the river channel as a percentage of total rainfall, hereafter referred to as the open water evaporation coefficient (OWEC). For all rivers this represents a very small component at less than 0.8%. This is because the area of the river channel network is small as compared to the total catchment area and also the river channel network is only flowing for a small proportion of time. However, additional evaporation may take place from water within the river bed sediments when the river is not flowing. This component has not been specifically identified in our analysis but is implicitly accounted for by the drying out of the sediment storage at the end of each dry season (see Section 2.3.5).

Recall that the term transmission loss is a term used to collectively quantify reductions in stream flow associated with river bed infiltration and evaporation from the river channel. Given that OWEC is very small in this context, the RBIC can be thought of as a transmission loss coefficient.

Fig. 7d shows posteriori distributions for total river flow as a percentage of total rainfall, hereafter referred to as the tributary runoff coefficient (TWC). TWC ranges between 1.5% and 2.0% for Metsimotlhabe, 5% and 8% for Shashe and Tati, and 7% and 13% for Ntshe. These figures compare well with Love et al. (2010) who observed event-based TWC in the Zhulube sand river of the northern Limpopo basin in Zimbabwe, mostly ranging between 2% and 10%. Our figures are also in line with the 12.4% runoff coefficient estimated for Botswana by FAO (1995), especially bearing in mind that not all rivers in Botswana are sand rivers.

4. Summary and conclusions

The objective of this article was to develop an improved methodology for quantifying transmission loss from ephemeral sand rivers by calibrating a lumped rainfall-runoff model to observed river flow data. Fifteen years of daily river flow data were obtained from four sand rivers in Botswana, namely, Shahshe, Ntshe, Tati and Metsimotlhabe. Our simplified rainfall runoff model had four unknown parameters including the river bed infiltration factor, the surface storage capacity, the river bed storage capacity and the river channel width. Posteriori parameter distributions were derived using a GLUE methodology.

Water stored within the river bed sediments is controlled by river bed infiltration, anthropogenic abstraction, seepage into underlying aquifers and evaporation. Due to uncertainty about these latter three quantities, a simplified approach was taken whereby losses were ignored during each wet season but assumed to be sufficiently large such that river bed sediments were completely dry by the end of each dry season. Such an approach is not suitable for modelling rivers where channel bed infiltration is storage capacity limited. However, in an inverse modelling sense, it is possible to use such an approach to determine whether river bed storage capacity is affecting river flow rates.

A parallel set of model runs were performed where the river bed sediments were assumed to have an infinite storage capacity. By comparing these results with those from models with finite storage capacities, it was ascertained that river bed infiltration was not river bed storage capacity limited for three of the four rivers studied (Shahshe, Tati and Metsimotlhabe). It was also found that the simulated river flows were insensitive to river channel network area. This insensitivity is because river channel network area only affects open water evaporation from river channels, which turns out to represent less than 0.8% of total precipitation.

Only two of the unknown model parameters were of significant importance: the river bed infiltration factor and the surface storage capacity. These latter two parameters were found to be strongly negatively correlated. Nevertheless, it was possible to obtain lower and upper bounds for the river bed infiltration factor for three of the rivers studied (Shahshe, Tati and Metsimotlhabe). Our results have identified that transmission loss represents between 55% and 85% of the total surface runoff at these locations.

Our study confirms that upper and lower bounds for transmission loss can be obtained by calibrating a lumped rainfall runoff model to a single set of river flow gauging data. Results from this study can be used to better inform future water balance studies for similar sand rivers in Southern Africa.

Acknowledgements

The study has benefited from a Global Challenges Research Fund (GCRF) Networking Grant from the Academy of Medical Sciences (GCRFNGR3\1048) and two GCRF research grants from Durham University. The authors are also grateful to the Department of Water and Sanitation, Botswana, for providing the river flow data for Shahshe, Ntshe, Tati and Metsimotlhabe.

References

Allen, R. G., Pereira, L. S., Raes, D., & Smith, M. (1998). Crop Evapotranspiration (guidelines for computing crop water requirements), FAO Irrigation and Drainage paper 56. Food and Agriculture Organization of the United Nations.

Ang, A. H. S., & Tang, W. H. (1975). Probability concepts in engineering planning and design. Basic Principles, vol. 1. Wiley, New York.

Beven, K., & Binley, A. (1992). The future of distributed models: model calibration and uncertainty prediction. Hydrological processes, 6(3), 279-298.

Beven, K. J. (2011). Rainfall-Runoff Modelling: The Primer. John Wiley & Sons.

Dahan, O., Tatarsky, B., Enzel, Y., Kulls, C., Seely, M., & Benito, G. (2008). Dynamics of flood water infiltration and ground water recharge in hyperarid desert. Groundwater, 46(3), 450-461.

Dinku, T., Funk, C., Peterson, P., Maidment, R., Tadesse, T., Gadain, H., & Ceccato, P. (2018). Validation of the CHIRPS satellite rainfall estimates over eastern Africa. Quarterly Journal of the Royal Meteorological Society, 144, 292-312.

Dunkerley, D. L. (2008). Bank permeability in an Australian ephemeral dry-land stream: variation with stage resulting from mud deposition and sediment clogging. Earth Surface Processes and Landforms: The Journal of the British Geomorphological Research Group, 33(2), 226-243.

FAO (1995). Irrigation in Africa in Figures. Water Reports. Food and Agriculture Organization of the United Nations. `http://www.fao.org/3/V8260B/V8260B00.htm#Contents`

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., ... & Michaelsen, J. (2015). The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. Scientific Data, 2(1), 1-21.

Gong, C., Wang, W., Zhang, Z., Wang, H., Luo, J., & Brunner, P. (2020). Comparison of field methods for estimating evaporation from bare soil using lysimeters in a semi-arid area. Journal of Hydrology, 590, 125-334.

Hughes, D. A. (2013). A review of 40 years of hydrological science and practice in southern Africa using the Pitman rainfall-runoff model. Journal of Hydrology, 501, 111-124.

Hughes, D. A. (2019). A simple approach to estimating channel transmission losses in large South African river basins. Journal of Hydrology: Regional Studies, 25, 100619.

Jarihani, A. A., Larsen, J. R., Callow, J. N., McVicar, T. R., & Johansen, K. (2015). Where does all the water go? Partitioning water transmission losses in a data-sparse, multi-channel and low-gradient dryland river system using modelling and remote sensing. Journal of Hydrology, 529, 1511-1529.

Jenson, S. K., & Domingue, J. O. (1988). Extracting topographic structure from digital elevation data for geographic information system analysis. Photogrammetric Engineering and Remote Sensing, 54(11), 1593-1600.

Lange, J. (2005). Dynamics of transmission losses in a large arid stream channel. Journal of Hydrology, 306(1-4), 112-126.

Lehner, B., Verdin, K., & Jarvis, A. (2008). New global hydrography derived from spaceborne elevation data. Eos, Transactions American Geophysical Union, 89(10), 93-94.

Love, D., Uhlenbrook, S., Corzo-Perez, G., Twomlow, S., & van der Zaag, P. (2010). Rainfall–interception–evaporation–runoff relationships in a semi-arid catchment, northern Limpopo basin, Zimbabwe. Hydrological Sciences Journal, 55(5), 687-703.

Love, D., van der Zaag, P., Uhlenbrook, S., & Owen, R. J. S. (2011). A water balance modelling approach to optimising the use of water resources in ephemeral sand rivers. River research and applications, 27(7), 908-925.

Makaba, L. P. (2017). Cultural and Demographic Factors Contributing to Environmental Degradation along the Metsimotlhabe River near Gaborone, Botswana. Journal of Human Ecology, 57(1-2), 70-77.

Mansell, M. G., & Hussey, S. W. (2005). An investigation of flows and losses within the alluvial sands of ephemeral rivers in Zimbabwe. Journal of Hydrology, 314(1-4), 192-203.

Mathias, S. A., McIntyre, N., & Oughton, R. H. (2016). A study of non-linearity in rainfall-runoff response using 120 UK catchments. Journal of Hydrology, 540, 423-436.

Moore, R. J. (2007). The PDM rainfall-runoff model. Hydrol. Earth Syst. Sci, 11(1), 483-499.

Morin, E., Grodek, T., Dahan, O., Benito, G., Kulls, C., Jacoby, Y., ... & Enzel, Y. (2009). Flood routing and alluvial aquifer recharge along the ephemeral arid Kuiseb River, Namibia. Journal of Hydrology, 368(1-4), 262-275.

Mpala, S. C., Gagnon, A. S., Mansell, M. G., & Hussey, S. W. (2020). Modelling the water level of the alluvial aquifer of an ephemeral river in south-western Zimbabwe. Hydrological Sciences Journal, 65(8), 1399-1415.

Nachtergaele, F., van Velthuizen, H., & Verelst, L. (2009). Harmonized World Soil Database (version 1.1). FAO, Rome, Italy and IIASA, Laxenburg, Austria.

Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I—A discussion of principles. Journal of Hydrology, 10(3), 282-290.

Ngoma, H., Wen, W., Ojara, M., & Ayugi, B. (2021). Assessing current and future spatiotemporal precipitation variability and trends over Uganda, East Africa, based on CHIRPS and regional climate model datasets. Meteorology and Atmospheric Physics, 1-21.

Parida, B. P., Moalafhi, D. B., & Kenabatho, P. K. (2006). Forecasting runoff coefficients using ANN for water resources management: The case of Notwane catchment in Eastern Botswana. Physics and Chemistry of the Earth, Parts A/B/C, 31(15-16), 928-934.

Parissopoulos, G. A., & Wheater, H. S. (1992). Experimental and numerical infiltration studies in a wadi stream bed. Hydrological sciences journal, 37(1), 27-37.

Ruane, A. C., Goldberg, R., & Chryssanthacopoulos, J. (2015). Climate forcing datasets for agricultural modeling: Merged products for gap-filling and historical climate series estimation. Agricultural and Forest Meteorology, 200, 233-248.

Sacre Regis M, D., Mouhamed, L., Kouakou, K., Adeline, B., Arona, D., Koffi Claude A, K., ... & Issiaka, S. (2020). Using the CHIRPS Dataset to Investigate Historical Changes in Precipitation Extremes in West Africa. Climate, 8(7), 84.

Shanafield, M., & Cook, P. G. (2014). Transmission losses, infiltration and groundwater recharge through ephemeral and intermittent streambeds: A review of applied methods. Journal of Hydrology, 511, 518-529.

Shaw, P., Shick, A., & Hassan, M. (1994). Bedload sediment transport in the sand rivers of Botswana. Botswana Notes & Records, 26(1), 115-127.

Upton, K., O Dochartaigh, B. E., Key, R., Farr J., & Bellwood-Howard, I. (2018). Africa Groundwater Atlas: Hydrogeology of Botswana. British Geological Survey. http://earthwise.bgs.ac.uk/index.php/Hydrogeology_of_Botswana

Walker, D., Jovanovic, N., Bugan, R., Abiye, T., du Preez, D., Parkin, G., & Gowing, J. (2018). Alluvial aquifer characterisation and resource assessment of the Molototsi sand river, Limpopo, South Africa. Journal of Hydrology: Regional Studies, 19, 177-192.