



# Parameter and input data uncertainty estimation for the assessment of water resources in two sub-basins of the Limpopo River Basin

Nadia Oosthuizen<sup>1,2</sup>, Denis A. Hughes<sup>2</sup>, Evison Kapangaziwiri<sup>1</sup>, Jean-Marc Mwenge Kahinda<sup>1</sup>, and Vuyelwa Mvandaba<sup>1,2</sup>

<sup>1</sup>CSIR, Natural Resources and Environment, P.O. Box 395, Pretoria 0001, South Africa

<sup>2</sup>Institute for Water Research, Rhodes University, P Bag 94, Grahamstown, 6140, South Africa

**Correspondence:** Nadia Oosthuizen (nadia.oosthuizen.geo@gmail.com)

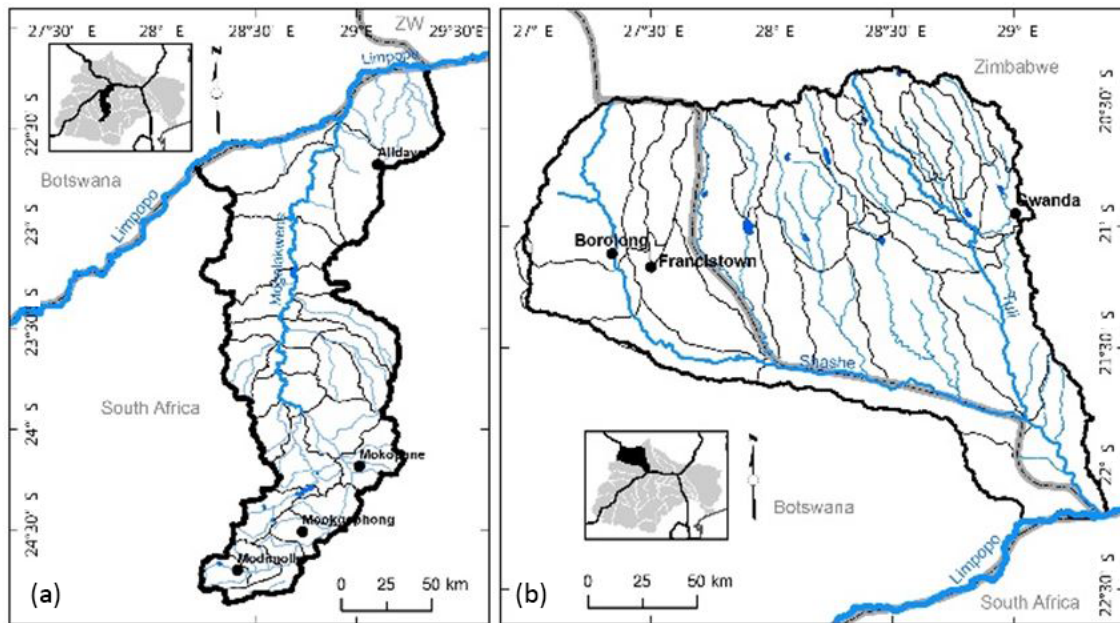
Received: 10 November 2017 – Revised: 23 February 2018 – Accepted: 5 March 2018 – Published: 29 May 2018

**Abstract.** The demand for water resources is rapidly growing, placing more strain on access to water and its management. In order to appropriately manage water resources, there is a need to accurately quantify available water resources. Unfortunately, the data required for such assessment are frequently far from sufficient in terms of availability and quality, especially in southern Africa. In this study, the uncertainty related to the estimation of water resources of two sub-basins of the Limpopo River Basin – the Mogalakwena in South Africa and the Shashe shared between Botswana and Zimbabwe – is assessed. Input data (and model parameters) are significant sources of uncertainty that should be quantified. In southern Africa water use data are among the most unreliable sources of model input data because available databases generally consist of only licensed information and actual use is generally unknown. The study assesses how these uncertainties impact the estimation of surface water resources of the sub-basins. Data on farm reservoirs and irrigated areas from various sources were collected and used to run the model. Many farm dams and large irrigation areas are located in the upper parts of the Mogalakwena sub-basin. Results indicate that water use uncertainty is small. Nevertheless, the medium to low flows are clearly impacted. The simulated mean monthly flows at the outlet of the Mogalakwena sub-basin were between 22.62 and 24.68 Mm<sup>3</sup> per month when incorporating only the uncertainty related to the main physical runoff generating parameters. The range of total predictive uncertainty of the model increased to between 22.15 and 24.99 Mm<sup>3</sup> when water use data such as small farm and large reservoirs and irrigation were included. For the Shashe sub-basin incorporating only uncertainty related to the main runoff parameters resulted in mean monthly flows between 11.66 and 14.54 Mm<sup>3</sup>. The range of predictive uncertainty changed to between 11.66 and 17.72 Mm<sup>3</sup> after the uncertainty in water use information was added.

## 1 Introduction

The continued socio-economic development of riparian countries of the Limpopo River leads to pressure on the water resources of the basin. The management of the water resources is, therefore, critical to avoid conflict and ensure equity and accessibility to the resource for both urban and rural populations. There are also various other competing water users such as the environment (environmental water requirements), livestock farming, irrigation, and mining operations. An identification of runoff generating processes, water

use and a clear understanding of their linkages within the basin are needed for improved quantification of the water resources of the basin, a requisite for better resource management. There are several challenges within the Limpopo, including shortages of water (e.g. droughts in the southern part of Zimbabwe and western Botswana), flooding (especially in the Mozambique part of the basin) and water quality issues (e.g. in the Oliphant's sub-basin in South Africa). One way of contributing to finding solutions for some of these problems is the understanding of the process linkages for improved resource quantification. However, in the absence of



**Figure 1.** Map of the Mogalakwena (a) and Shashe (b) sub-basins and their locations in the Limpopo River Basin.

historical observed data (large parts of the basin are virtually ungauged) of the different aspects of the hydrology of the basin (such as rainfall, streamflow, water use, etc.) hydrological models are generally used to generate data and information that inform management and policy decision making.

Ideally, any hydrological model should be based on a sound conceptual understanding of the processes operating in the basin and should be backed by quantitative information that can be used for the parameterization of the model (Hughes et al., 2006, 2010). However, these data are often inadequate in many sub-basins, necessitating the incorporation of the uncertainty related to the estimation process. Given the diversity of the four riparian countries on data collection and management, the quantity and quality of the data in the basin varies and therefore a framework that incorporates estimates of uncertainty should be applied to deal with this challenge. Earlier studies (e.g. Matji and Görgens, 2001) have generally concentrated on the main stem of the basin and this study explains the possible uncertainty related to the modelling process.

## 2 Study Area

Two physically and socio-economically contrasting sub-basins were chosen for the study (Fig. 1). The Mogalakwena River rises as the Nyl River south of Mokopane flows northwards into the Limpopo (Busari, 2008). The Mogalakwena sub-basin, with a drainage area of 19 400 km<sup>2</sup>, is located in South Africa and is very densely populated and industrialized incorporating the towns of Modimolle, Mookgopong, and Mokopane (Fig. 1a). The Shashe sub-basin (Fig. 1b),

shared between Botswana and Zimbabwe, rises on the border of the two countries. It flows south, past Francistown (in Botswana) where after the flow changes to a south-easterly direction along the border for about 362 km until it flows into the Limpopo River.

While the sub-basins are near one another, they however have different physical characteristics such as rainfall, evaporation, soils, geology and land cover which, in turn, influence the runoff rates, water use, and abstractions. Also, since the catchments are located in different countries, the data collection methods vary affecting the data quality and availability. In South Africa, hydrological data can be accessed freely from sources such as the Water Resources Studies (1990, 2005 and 2012), national departments such as the Department of Water and Sanitation (DWS) and the Department of Agriculture, Forestry and Fisheries (DAFF), whereas hydrological data for Botswana and Zimbabwe are more difficult to access, the records of poorer quality often with missing values, and large areas are ungauged. The two study areas are therefore chosen to show the impacts of different physical and hydro-climatic data conditions on estimated water resources.

## 3 Data and Methodology

The Pitman model has been widely used for water resources assessment in most of the southern Africa region since its development in the early 1970s (Wilk and Hughes, 2002). The Spatial and Time Series Information Modelling (SPATSIM, Hughes and Forsyth, 2006) version of the Pitman model (Pitman, 1973) was used to quantify water resources of the se-

lected sub-basins. It is a conceptual type model with parameters that are associated with components that represent the main hydrological processes (and human impacts) that operate at a sub-basin scale (Hughes et al., 2010). The SPATSIM version of the model is designed to handle multiple model runs that are required to consider uncertainty issues arising from the quantity and quality of the input data (Kapangaziwiri et al., 2012). This version also has explicit surface-ground water interaction routines (Hughes, 2004) and a wetland function. Table 1 shows the main parameters of the model (Hughes et al., 2006).

Runoff is mainly generated by two model functions. The first is an asymmetrical triangular distribution of catchment absorption rates defined by parameters ZMIN, ZAVE and ZMAX. The second function determines the drainage rate from the main moisture storage ( $S$ , with a capacity of ST, mm). This storage is depleted by evapotranspiration, interflow and groundwater recharge. The maximum interflow (FT, mm month<sup>-1</sup>) and recharge (GW, mm month<sup>-1</sup>) rates occur at ST, while two power functions (parameters POW and GPOW) determine these rates at lower values of moisture storage ( $S$ , mm). Recharge is routed through a groundwater storage function that accounts for evapotranspiration losses, drainage to other catchments and contributions to base flow.

The model also has functions that simulate the impact of human activities like small farm dams, large reservoirs and irrigated agriculture in managed basins. Detailed descriptions of the model can be found in Hughes et al. (2006) and Kapangaziwiri et al. (2012).

### 3.1 Data collection

For this study, the impact of uncertainty related to farm dam and irrigation data was evaluated. The data sources that were used include: the 2012 national water resources assessment study of South Africa (called the WR2012) database, data from the South African Department of Agriculture, Forestry and Fisheries (DAFF) and remote sensing products for the Mogalakwena sub-basin. For Shashe, besides remotely-sensed information, the data from the Limpopo River Basin Monograph study (LIMCOM, 2013) were used.

The farm dam data (and all the other data used in this study) were collected at the scale at which water resources decisions are made (i.e. at the so-called quaternary catchment for the Mogalakwena and at the sub-zone scale for Shashe). In this study, both quaternary and sub-zones are referred to just as catchment. For purposes of simplicity and also based on the model used, the individual small farm dams in each catchment were added up to form one dam at the outlet, whose parameters (e.g. full supply capacity and area) were then subsequently specified and the level of uncertainty where necessary. Only the surface areas of the identified farm dams were estimated from the remote sensing and through a manual digitizing process. The dam capacities had to be calculated using a generalised relationship between capacity

and area at full supply level provided by LIMCOM (2013). While many approaches to the estimation of dam capacities could be used (e.g. Sayl et al., 2016; Hughes and Mantel, 2010 and Sawunyama et al., 2006) this equation was chosen since it was applied successfully in the whole of the Limpopo basin, and is given as:

$$\text{Surface Area} = 0.4 \cdot (\text{Capacity})^{0.7} \quad (1)$$

The calculated volumes were similar to those contain in existing national databases and were therefore assumed adequate to use. For running the uncertainty in the SPATSIM version of the Pitman model, minimum (Min) and maximum (Max) values are required in the model setup. The percentage differences between the average of minimum and maximum of dam volumes or irrigation areas was thus determined to represent the uncertainty. This is given by:  $[\text{maximum average} - \text{minimum average}] / \text{maximum average value} \times 100$  (Sawunyama et al., 2011).

WR2012 irrigation data for Mogalakwena sub-basin were compared to the data received from the South African Department of Agriculture, Forestry and Fisheries (DAFF) and remotely-sensed data. For the Shashe sub-basin, the LIMCOM (2013) data were also compared to areas digitised in Google Earth.

### 3.2 Uncertainty analysis

The analysis focused on the impact of uncertainties related to the main physical parameters of the runoff generation process and water use data (i.e. irrigation and farm dams) on the estimation of sub-basin water resources. Parameter uncertainties were considered first, before the incorporation of water use uncertainties.

In this study, the parameter values for each run of the model are independently randomly sampled from the inputs by making use of a Normal (defined by the mean and standard deviation) frequency distribution. As soon as a parameter set generates a simulation that satisfies all of the constraints, it is saved to the SPATSIM database (Ndzabandzaba and Hughes, 2017). However, when 1000 output behavioural parameter sets have been found the model terminates. Since the constraints define the uncertainty in the hydrological response behaviour of each of the selected sub-basins (Yadav, et al., 2007; Westerberg et al., 2011, 2014) all of the saved parameter sets represents behavioural responses (Beven, 2012).

## 4 Results and Discussion

### 4.1 Uncertainty results

A significant difference between the irrigation coverage is observed for the Mogalakwena catchments giving relatively huge uncertainty. However, the difference between the irrigated areas sourced from the LIMCOM (2013) compared to the areas digitised from Google Earth is marginal, resulting

**Table 1.** A list of the parameters of the Pitman model including those of the reservoir water balance model (Hughes et al., 2006).

Parameter	Units	Parameter Description
RDF	–	Controls the distribution of total monthly rainfall over four iterations
AI	Fraction	Impervious fraction of sub-basin
PI1 & PI2	mm	Interception storage for two vegetation types
AFOR	%	% area of sub-basin under vegetation type 2
FF	–	Ratio of potential evaporation rate for Veg2 relative to Veg1
PEVAP	mm	Annual sub-basin evaporation
ZMIN	mm month <sup>-1</sup>	Minimum sub-basin absorption rate
ZAVE	mm month <sup>-1</sup>	Mean sub-basin absorption rate
ZMAX	mm month <sup>-1</sup>	Maximum sub-basin absorption rate
ST	mm	Maximum moisture storage capacity
SL	mm	Minimum moisture storage below which no GW recharge occurs
POW	–	Power of the moisture storage – runoff equation
FT	mm month <sup>-1</sup>	Runoff from moisture storage at full capacity (ST)
GPOW	–	Power of the moisture storage – GW recharge equation
GW	mm month <sup>-1</sup>	Maximum ground water recharge at full capacity, ST
R	–	Evaporation-moisture storage relationship parameter
TL	months	Lag of surface and soil moisture
CL	months	Channel routing coefficient
DDENS	–	Drainage density
T	m <sup>2</sup> d <sup>-1</sup>	Ground water transmissivity
S	–	Ground water storativity
GWSlope	–	Initial ground water gradient
AIRR	km <sup>2</sup>	Irrigation area
IWR	Fraction	Irrigation water return flow fraction
EffRf	Fraction	Effective rainfall fraction
NirrDmd	Ml yr <sup>-1</sup>	Non-irrigation demand from the river
MAXDAM	Ml	Small dam storage capacity
DAREA	%	Percentage of sub-basin above dams
A, B	–	Parameters in non-linear dam area-volume relationship
IrrAreaDmd	km <sup>2</sup>	Irrigation area from small dams
CAP	Mm <sup>3</sup>	Reservoir capacity
DEAD	%	Dead storage
INIT	%	Initial storage
RES 1–5	%	Reserve supply levels (percentage of full capacity)
ABS	Mm <sup>3</sup>	Annual abstraction volume
COMP	Mm <sup>3</sup>	Annual compensation flow volume

in small uncertainty. Nevertheless, the uncertainty still had an impact on the low flows of the sub-basins. Shashe catchments BS4 and BS5 have little irrigated areas and the difference between the irrigated areas provided by LIMCOM (2013) and manual digitizing was too small to impact the results.

The overall calculated uncertainty of the Mogalakwena sub-basin varied between 30 and 100 % and between 5 and 100 % for the farm dams and irrigation data respectively. For the Shashe sub-basin, where farm dams and irrigation existed, the calculated uncertainty varied between 5 and 90 % and between 10 and 100 % farm dams and irrigation respectively.

#### 4.2 Water resources simulation of the Mogalakwena sub-basin

Based solely on the uncertainty related to model parameters, the simulated monthly flows ranged between a minimum of 21.53 Mm<sup>3</sup> and a maximum of 24.68 Mm<sup>3</sup>; resulting in a mean monthly flow of 22.62 Mm<sup>3</sup>. Results indicate that the whole range of flows (i.e. high, medium and low flows) is impacted when uncertainty related to natural parameters is considered. Adding the uncertainty related to water use (i.e. irrigation and farm dams) data resulted in minimum simulated flows of 22.15 Mm<sup>3</sup> and a maximum of 24.99 Mm<sup>3</sup>. High flows are more impacted than low flows.



### 4.3 Water resources simulation for the Shashe sub-basin

The difference between the model parameters and water use uncertainties is barely noticeable for some catchments of the Shashe. This is because the cumulative volumes of the farm dams was too small to have a noticeable impact on the total predictive uncertainty. The same applies to the irrigation water use uncertainty. Considering uncertainty related to the natural parameters of runoff generation, the simulated mean monthly flow was  $14.54 \text{ Mm}^3$  with a maximum of  $17.15 \text{ Mm}^3$  and a minimum of  $11.66 \text{ Mm}^3$ . When the uncertainty related to the water use data was added the minimum value of the simulated flow remained the same, whereas the maximum increased to  $17.72 \text{ Mm}^3$ .

## 5 Conclusions

The main objective of the study was to demonstrate that estimated water resources are capable of spanning a wide range of plausible or probable values when considering the uncertainties in model parameter estimation, our understanding of the hydrological processes prevalent in a basin and how those processes are represented in the model as well as the use of limited and often poor quality historical observed data to calibrate the model. This study gives insight into how the simulation of the natural water resources of the Mogalakwena and Shashe sub-basins would be expected to vary when expected uncertainty is incorporated into the estimation process. It can be concluded that uncertainty (in this case related to parameter and water use data) plays an important role in the estimation of water resources as demonstrated in the sub-basins. The uncertainty related to the estimation of water use data tends to affect the medium to low flows more than just considering parameter uncertainty. Farm dams absorb the flows during the high flow season, whereas irrigation would be important during the low flow season. This study provides a limited illustration of how the identification and quantification of parameter and water use uncertainties can provide insight into the possible impacts of using a database such as the national water resources assessment study (WR2012) without examining the quality of the data. This study is a step forward in providing insight into water resources estimations based on the data that are routinely used in the assessment processes without examining the possible impact of the uncertainties related to those data.

**Data availability.** Data not publicly available as they were obtained from a consultant with the express caveat that they would only be used for the project and not shared publicly. However, these data will soon be available on the Limpopo Management Information System (LIMIS) which is managed by the Limpopo River Basin Commission who contracted the consultant to collect the data.

**Competing interests.** The authors declare that they have no conflict of interest.

**Special issue statement.** This article is part of the special issue “Understanding spatio-temporal variability of water resources and the implications for IWRM in semi-arid eastern and southern Africa”. It is a result of the IAHS Scientific Assembly 2017, Port Elizabeth, South Africa, 10–14 July 2017.

**Acknowledgements.** This paper forms part of a project titled: “Upstream-Downstream Hydrological Linkages in the Limpopo River Basin” which is funded by the Water Research Commission (WRC) of South Africa (Project No. K5/2439/1) and the CSIR parliamentary grant (ECHS052).

Edited by: Hodson Makurira

Reviewed by: Joel Nobert and Donald Rwasoka

## References

- Beven, K.: Casual models as multiple working hypotheses about environmental processes, *Comptes Rendus – Geoscience*, 344, 77–88, 2012.
- Busari, O.: Groundwater in the Limpopo Basin: occurrence, use and impact, *Environmental Development and Sustainability*, 10, 943–957, 2008.
- DAFF (Department of Agriculture, Forestry and Fisheries): RSA Irrigation Data, Department of Agriculture, Forestry and Fisheries, Pretoria, South Africa, 2015.
- Hughes, D. and Mantel, S. K.: Estimating the uncertainty in simulating the impacts of small farm dams on streamflow regimes in South Africa, *Hydrol. Sci. J.*, 55, 578–592, 2010.
- Hughes, D. A.: Southern Africa FRIEND – The Application of Rainfall-Runoff Models in the SADC Region, Water Research Commission Report No. 235/1/97, Pretoria, South Africa, 1997.
- Hughes, D. A.: Incorporating ground water recharge and discharge functions into an existing monthly rainfall-runoff model, *Hydrol. Sci. J.*, 49, 297–311, 2004.
- Hughes, D. A. and Forsyth, D.: A generic database and spatial interface for the application of hydrological and water resource models, *Computers and Geosciences*, 32, 1389–1402, 2006.
- Hughes, D. A., Andersson, L., Wilk, J., and Savenije, H. H. G.: Regional calibration of the Pitman model for the Okavango River, *J. Hydrol.* 331, 30–42, 2006.
- Hughes, D. A., Kapangaziwiri, E., and Sawunyama, T.: Hydrological model uncertainty assessment in southern Africa, *J. Hydrol.*, 387, 221–232, 2010.
- IWR. SPATSIM: Spatial and Time Series Information Modelling Software, Institute for Water Research, Rhodes University, Grahamstown, South Africa, available at: <https://www.ru.ac.za/iwr/research/software/>, last access: April 2017.
- Kapangaziwiri, E., Hughes, D. A., and Wagener, T.: Incorporating uncertainty in hydrological predictions for gauged and ungauged basins in southern Africa, *Hydrol. Sci. J.*, 57, 1000–1019, 2012.

- Limpopo Watercourse Commission (LIMCOM): Limpopo River Basin Monograph, Final Monograph. Report No. LRBMS-81137945, 2013.
- Matji, M. P. and Görgens, A. H. M.: Update of the Hydrological Model of the Limpopo River Main Stem, Peer-Reviewed Report by the Department of Civil Engineering and Ninham Shand (Pty) Ltd to the Water Research Commission, Pretoria, South Africa, on behalf of the four co-border Countries, 2001.
- Ndzabandzaba, C. and Hughes, D. A.: Regional water resources assessments using an uncertain modelling approach: The example of Swaziland, *J. Hydrol.: Regional Studies*, 10, 47–60, 2017.
- Pitman, W. V.: A mathematical model for generating monthly river flows from meteorological data in South Africa, Hydrological Research Unit, Univ. of the Witwatersrand, Report No. 2/73, 1973.
- Sawunyama, T., Senzanje, A., and Mhizha, A.: Estimation of reservoir storage capacities in Limpopo River Basin using geographical information systems (GIS) and remotely sensed areas: Case of Mzingwane catchment, *Physics and Chemistry of the Earth Parts A/B/C*, 31, 935–943, 2006.
- Sawunyama, T., Hughes, D. A., and Mallroy, S. J. L.: Evaluation of combined contribution of uncertainty sources to total output uncertainty in water resource estimation in South Africa, IAHS Publication, 347, 2011.
- Sayl, K. N., Muhammad, N. S., Yaseen, Z. M., and El-shafie, A.: Estimation the Physical Variables of Rainwater Harvesting System Using Integrated GIS-Based Remote Sensing Approach, *Water Resour. Manag.*, 30, 3299, <https://doi.org/10.1007/s11269-016-1350-6>, 2016.
- Westerberg, I. K., Guerrero, J.-L., Younger, P. M., Beven, K. J., Seibert, J., Halldin, S., Freer, J. E., and Xu, C.-Y.: Calibration of hydrological models using flow-duration curves, *Hydrol. Earth Syst. Sci.*, 15, 2205–2227, <https://doi.org/10.5194/hess-15-2205-2011>, 2011.
- Westerberg, I. K., Gong, L., Beven, K. J., Seibert, J., Semedo, A., Xu, C.-Y., and Halldin, S.: Regional water balance modelling using flow-duration curves with observational uncertainties, *Hydrol. Earth Syst. Sci.*, 18, 2993–3013, <https://doi.org/10.5194/hess-18-2993-2014>, 2014.
- Wilk, J. and Hughes, D. A.: Calibrating a rainfall-runoff model for a catchment with limited data, *Hydrol. Sci. J.*, 41, 3–17, 2002.
- Yadav, M., Wagener, T., and Gupta, H. V.: Regionalization of constraints on expected watershed response behaviour, *Adv. Water Res.*, 30, 1756–1774, 2007.