Evaluation of combined contribution of uncertainty sources to total output uncertainty in water resource estimation in South Africa

TENDAI SAWUNYAMA, DENIS A. HUGHES & STEPHEN J. L. MALLORY
Institute for Water Research, Rhodes University, PO Box 94 Grahamstown, 6140, South Africa
tintosawuz@yahoo.com

Abstract While the importance of quantifying different sources of uncertainty is well recognized, there have only been a few attempts within southern Africa to incorporate uncertainty estimates in water resources assessments, and their overall impacts are not well understood. Uncertainties are model, basin, region and climate zone specific, and while the basic principles referred to in the hydrological literature are relevant, they do not provide the specific answers for the region. The focus of this study was on the use of datasets and modelling tools that are frequently used for practical water resources estimation in South Africa. The analyses are based on scenarios of different sources of uncertainty which are then combined and propagated through the model to generate simulation ensembles that include the expected ranges of model output uncertainty. The results indicate that the major source of uncertainty is either rainfall or parameter value estimation depending on the sub-basin. The study suggests that, while input climate data always contributes substantially to total output uncertainty, there may be many situations where parameter uncertainty dominates and there is very little impact from evaporation and water-use uncertainties.

Key words model uncertainty; rainfall–runoff model; South Africa

INTRODUCTION

Previous studies in South Africa on water resources modelling uncertainty have demonstrated that the outputs of a hydrological model are affected by many sources of uncertainty related to hydro-climatic data (Sawunyama & Hughes, 2007), parameter values and the model structure (Hughes *et al.*, 2010). The results suggest that individual sources of uncertainty contribute in different ways under different climate and physiographic conditions (Hughes & Mantel, 2010) and that clear statements about which source of uncertainty is likely to dominate are not generally possible. In this paper an attempt is made to integrate the combined effects of different sources of uncertainty, including uncertainty in water use data on total output uncertainty for a limited number of sub-basins selected in South Africa. Some of the previous studies focused on the simulation of natural flows and ignored uncertainty in estimating water use data that would be required to accurately estimate present day water resource availability. Water use data are among the most unreliable sources of information in many developing countries because observations of actual (rather than licensed) water use are relatively uncommon. The water uses considered in this study are streamflow reductions due to managed forest plantations and direct runoff-of-river irrigation abstractions. The main objective of this paper is to assess the combined effects of different sources of uncertainty on the overall model output uncertainty.

DATA AND METHODS

Historical (1920–1990) spatially-averaged monthly rainfall and mean monthly potential evaporation data for selected South African sub-basins were obtained from the national water resources database (WR90; Midgley *et al.*, 1994). These data were used to calibrate the modified version (including surface–groundwater interactions) of the monthly Pitman model (Hughes, 2004) against all observed streamflow data available during this period. The same rainfall stations used in the WR90 study were used to generate new sets of rainfall data using the Inverse Distance Weighting (IDW) approach. Three sub-basins, covering a wide range of hydro-climatic conditions were selected (Fig. 1), two gauged (K40A; 72km^2 and U20B; 353 km^2) and one ungauged (G10B; 126 km^2). Information on forest plantations and irrigated areas as well as on storage volumes and surface areas of small farm dams is provided in Midgley *et al*. (1994) and the updated WR2005 reports (Middleton & Bailey, 2008). Additional information is available from the Water Use Authorisations and Registration Management System (WARMS) database which is being developed by DWAF as part of the process of registering existing lawful water uses within South Africa. The afforestation effects on streamflow reductions are based on fixed changes to interception (PI) and evaporation demand (FF) parameters of the Pitman model.

Fig. 1 Map of South Africa showing Water Management Areas and selected sub-basins.

There are several reported methods that have attempted to account for all sources of uncertainty, such as the Bayesian Total Error Analysis (Kavetski *et al.*, 2006) and the Integrated Bayesian Uncertainty Estimator (Ajami *et al.*, 2007). However, they are statistical or data-driven approaches and are not very appropriate for data scare regions such as southern Africa. In this study combined contributions of climate input data, model parameter (assumed to include model structure uncertainty) and water use uncertainties are quantified. The analysis was performed by creating scenarios of different sources of uncertainty, combining their contribution and propagating them through the Pitman hydrological model to generate simulation ensembles that include the expected range of model output uncertainty. The scenarios include different formulations of rainfall input, potential evapotranspiration estimates, water use data and parameter sets (Table 1). In this study, uncertainty propagation was evaluated based on both the simulated mean annual runoff (MAR) and the simulated yield of hypothetical reservoirs at each sub-basin outlet. The simulated yields are based on a mean annual achieved abstraction at a 90% level of assurance.

RESULTS

Table 2 presents an example of one set of results (using MAR and yield estimates) where the evaporation inputs are varied and all other inputs fixed, while Table 3 summarises all combinations of uncertainty using the estimated yield values. The model output uncertainty assessment is also illustrated using flow duration curves (FDCs) in Figs 2, 3 and 4.

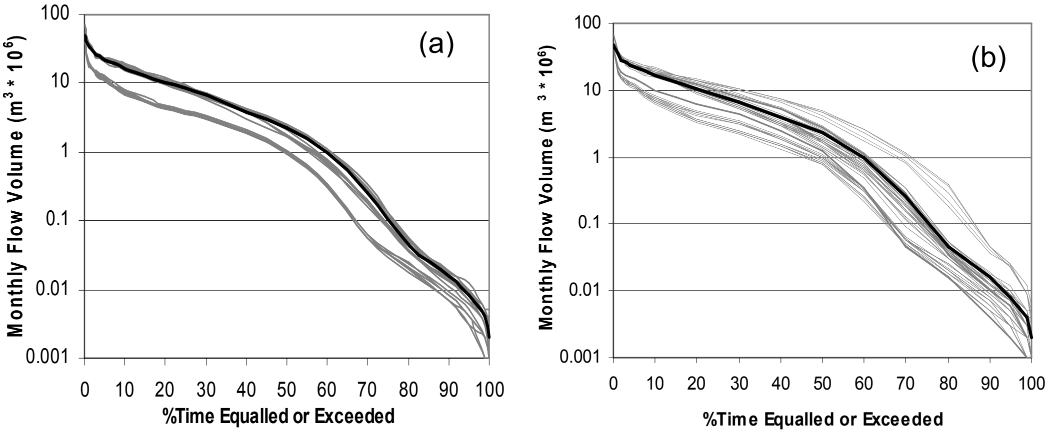

Fig. 2 Ensembles of FDCs for G10B: (a) no parameter uncertainty; (b) all sources of uncertainty included. The bold lines represent best estimate flows.

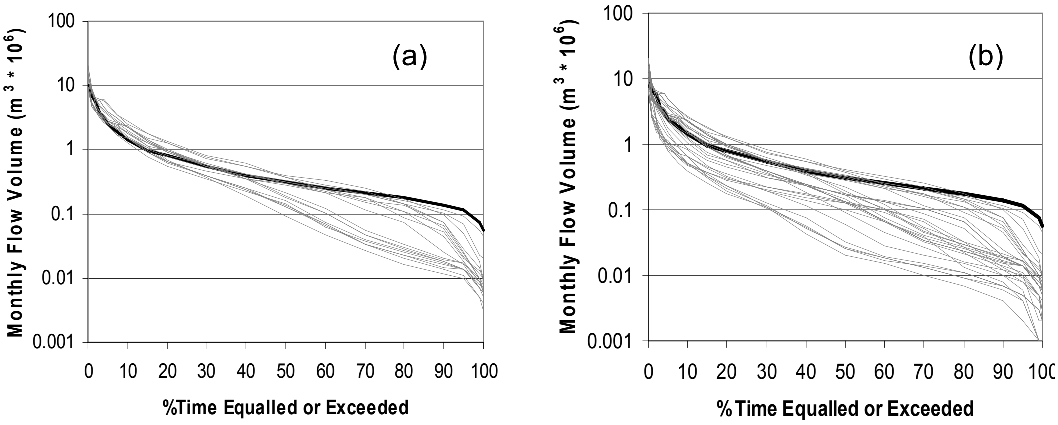

Fig. 3 Ensembles of FDCs for K40A: (a) no parameter uncertainty; (b) all sources of uncertainty included. The bold lines represent observed flows for 1965–2000.

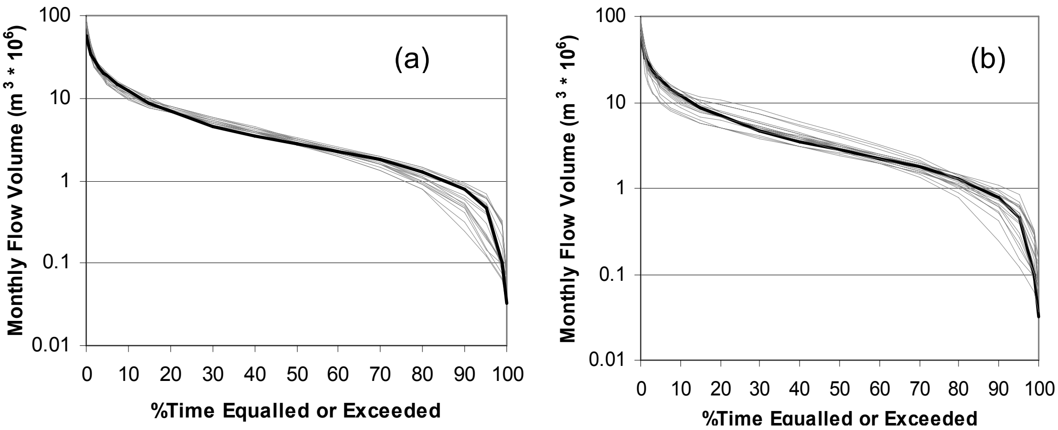

Fig. 4 Ensembles of FDCs for U20B: (a) no parameter uncertainty; (b) all sources of uncertainty included. The bold lines represent observed flows for 1954–2000.

The flow duration curves in Figs 2, 3 and 4 illustrate that the uncertainty bounds are substantially increased when the parameter uncertainty is included. The parameter uncertainty is mainly caused by uncertainty associated with the inability to accurately relate physical property data to the parameter values, and how the Pitman model responds to these effects given that the scale of physical property data does not usually match the model scale. In the gauged sub-basins (K40A and U20B) the available observed flows are included on the graphs, while for the ungauged G10B sub-basin the "best estimate" flows are used for comparisons with the other ensembles. The "best estimate" flows were generated using the combination of WR90 rainfall input (R1), mean monthly evapotranspiration (E1), "best guess" parameter values (PB) together with water use data from the WR90 database (W1) (Table 1).

For G10B (Fig. 2), there are two distinct ensemble groups which reflect uncertainty in rainfall input data due to the high spatial rainfall gradients in this mountainous region. The inclusion of parameter uncertainties (upper and lower bounds) (Fig. 2(b)) contributes a large amount of uncertainty. For K40A (Fig. 3) most ensembles underestimate observed flows, mainly in the low-flow regime, and this may be a reflection of unaccounted for sources of uncertainty or because the "best-guess" parameter set is not very behavioural. For U20B (Fig. 4) the uncertainties due to rainfall, evaporation and water use alone are relatively small and, while parameter uncertainty dominates, the overall uncertainty is much less than in the other sub-basins.

Table 1 Description of scenarios used in the uncertainty analysis.

Scenario	Description
R1	WR90 monthly spatial rainfall (Midgley *et al.*, 1994) (1920–1990).
R2	IDW monthly spatially interpolated rainfall data (1920–2000).
R3	Same as "R2" but with frequency characteristics corrected based on "R1" (1920–2000).
E1	Fixed monthly pan based evaporation demand (Midgley *et al.*, 1994).
E2	Monthly time series evaporation demand based on pan data (start dates are different in different sub-basins).
E3	Monthly time series evaporation demand perturbed from temperature data (1950–2000).
PW	Existing WR90 regional parameters values based on calibration against observed flows (Midgley *et al.*, 1994).
PB	The "best" estimate parameter values estimated by parameter estimation approach of Kapangaziwiri & Hughes (2008).
PL	The lower bound parameter values also estimated by parameter estimation approach of Kapangaziwiri & Hughes (2008).
PU	The upper bound parameter values also estimated by parameter estimation approach of Kapangaziwiri & Hughes (2008).
W1	Irrigation and afforestation water demands based on WR90 database.
W2	Irrigation and afforestation water demands based on either WR2005 or WARMS data.

Table 2 Illustration of ensemble simulations results of MAR and yield (both in $m^3 \times 10^6$ $year^{-1}$).

Sub-basin	G10B		K40A		U20B	
Realization	MAR	Yield	MAR	Yield	MAR	Yield
R1, E1, PW, W1.	87.4	77.2	10.0	7.3	56.2	47.4
R1, E2, PW, W1.	86.7	78.6	11.4	7.5	63.1	48.7
R1, E3, PW, W1.	91.6	79.1	10.7	7.3	60.7	49.5

In Table 3 the "%Diff" values represent percentage differences between minimum (min) and maximum (max) annual yields and is given by: {100 × (maximum average – minimum average value) / maximum average value}. The minimum and maximum values represent a range of yield values resulting from varying one source of uncertainty while the others are fixed. A "set" represents a number of combinations which are grouped (the data in Table 2 represent set 1) where one source of uncertainty is varied while others are fixed.

For the G10B sub-basin, Table 3 suggests that the ranges of achieved yields (based on averages) due to evaporation and water-use uncertainty are small (3.4 and 3.1%) compared with the ranges associated with rainfall uncertainty (44.6%) and parameter uncertainty (25.7%). Figure 5 summarises the differences between sources of uncertainty in all sub-basins. The pie charts in Fig. 5 are based on the percentage differences of the average minimum and maximum yield estimates given in Table 3. It has been assumed that the sum of these percentage differences (i.e. 76.8% in G10B) represents a measure of the total uncertainty and that the individual differences represent the relative contribution of each source. The results (Fig. 5) indicate that the major source of uncertainty is either rainfall (59% for G10B) or parameter value estimation (64% for K40A and 64% for U20B), depending on the sub-basin. In the three examples, evaporation and water-use uncertainties are relatively small. This information is useful to both hydrologists and water resources managers, and it allows water resources planning decisions to be made about future risks of using uncertain information and where efforts to reduce uncertainty should be focused.

Table 3 Summarised results of combined contribution of different uncertainty sources on simulated mean annual yields ($m^3 \times 10^6$ $year^{-1}$) for G10B sub-basin.

Set	Evaporation		Rainfall		Parameter		Water use	
	Min yield	Max yield	Min yield	Max yield	Min yield	Max yield	Min yield	Max yield
1	77.2	79.1	50.7	77.2	60.5	77.2	60.5	62.5
2	48.3	50.8	48.3	78.6	60.2	78.6	60.2	62.4
3	75.9	76.6	50.8	79.1	63.2	79.1	63.2	65.0
.	.	.	.	.	.	.	.	
.	.	.	.	.	.	.	.	.
24	70.8	72.4	44.2	74.9	–	–	31.5	32.9
Average	58.2	60.3	38.9	70.2	50.9	68.6	54.1	55.9
%Diff.		3.4		44.6		25.7		3.1

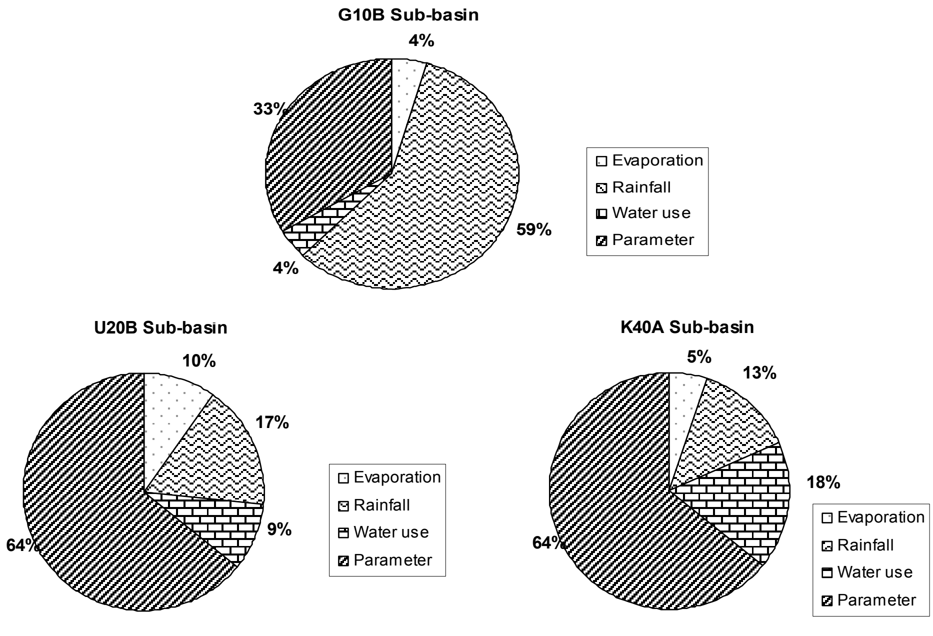

Fig. 5 Pie charts showing the contributions of different uncertainty sources to total output uncertainty based on the % differences in achieved yield from hypothetical reservoirs.

DISCUSSION AND CONCLUSIONS

It can be concluded that the uncertainty in water use is associated with the inability to accurately define present-day water uses within sub-basins, representing these in an integrated way using model parameters. The information on "real" water use in many parts of the country is often inaccurate and can constrain the successful implementation of water availability estimation tools. However, within the three example sub-basins, the effects of water use uncertainty are small compared to climate data and parameter uncertainty. G10B and K40A are examples where high spatial rainfall variability occurs and the lack of adequate observations contributes to the uncertainty in climate input data. In G10B, an area with steep topography, orographic effects and inadequate rainfall gauging play a dominant role in overall uncertainty. In some instances (e.g. K40A and U20B), the dominant source of uncertainty is in the quantification of model parameters. The uncertainty in the parameter estimates is made up of uncertainty in the physical property data, as well as in the estimation equations themselves. Part of this problem lies in the scale differences between the physical property data and the model. However, it must be accepted that the methods used in this study to quantify the uncertainty ranges require further development. Some of the differences in the contributions of parameter uncertainty can be attributed to the lack of a consistent approach to interpreting the physical property data.

An important observation is that a collective assessment of uncertainty is more informative than individual assessments. This study provides a limited illustration of how uncertainty sources can be combined in a water resource estimation process and a more consistent strategy must be developed. One approach would be to further develop the methods of estimating the inputs (climate data and parameters) such that some measure of the probability of each possible input is quantified. An alternative approach is currently being developed in South Africa (Hughes *et al.*, 2009) and involves quantifying regional measures of hydrological response (Yadav *et al.*, 2007) that can be used to evaluate the likelihood of simulated ensembles.

REFERENCES

Ajami, N. K., Duan, Q. & Sorooshian, S. (2007) An integrated hydrologic Bayesian multimodel combination framework: Confronting input, parameter, and model structural uncertainty in hydrologic prediction. *Water Resour. Res.* **43**, W01403, doi:10.1029/2005WR004745.

Hughes, D. A. (2004) Incorporating groundwater recharge and discharge functions into an existing monthly rainfall–runoff model. *Hydrol. Sci. J.* **49**(2), 297–311.

Hughes, D. A., Kapangaziwiri, E. & Sawunyama, T. (2010) Hydrological model parameter and results uncertainty estimation for water resource assessments in southern Africa. *J. Hydrol*. **387**, 221–232.

Hughes, D. A. & Mantel, S. (2010) Estimating uncertainties in simulations of natural and modified streamflow regimes in South Africa. In: *Global Change – Facing Risks and Threats to Water Resources* (ed. by E. Servat *et al*.), 358–364. (Proceedings of the Sixth FRIEND World Conference held in Fez, Morocco, November 2010). IAHS Publ. 340, IAHS Press, Wallingford, UK.

Kavetski, D., Kuczera, G. & Franks, S. W. (2006) Calibration of conceptual hydrological models revisited: 1. Overcoming numerical artefacts. *J. Hydrol*. **320**(1–2), 173–186.

Middleton, B. J. & Bailey, A. K. (2008) Water Resources of South Africa, 2005 Study (WR2005), Water Research Commission Report, Contract K5/1491, Pretoria, South Africa.

Midgley, D. C., Pitman, W. V. & Middleton, B. J. (1994) *Surface Water Resources of South Africa 1990*, Volumes I to VI, Water Research Commission Report, Pretoria, South Africa.

Sawunyama, T. & Hughes, D. A. (2007) Assessment of rainfall–runoff model input uncertainties on simulated runoff in southern Africa. In: *Quantification and Reduction of Predictive Uncertainty for Sustainable Water Resource Management* (ed. by E. Boegh *et al.*) (Proc. Symp. HS2004 at IUGG2007, Perugia, July 2007), 98–106. IAHS Publ. 313. IAHS Press, Wallingford, UK.

Yadav, M., Wagener, T. & Gupta, H. (2007) Regionalisation of constraints on expected watershed response behaviour for improved predictions in ungauged basins. *Adv. Water Res.* **30**, 1756–1774.