



Scientific and practical tools for dealing with water resource estimations for the future

D. A. Hughes

Institute for Water Research, Rhodes University, Grahamstown, South Africa

Correspondence to: D. A. Hughes (d.hughes@ru.ac.za)

Received: 11 March 2015 – Accepted: 11 March 2015 – Published: 12 June 2015

Abstract. Future flow regimes will be different to today and imperfect knowledge of present and future climate variations, rainfall–runoff processes and anthropogenic impacts make them highly uncertain. Future water resources decisions will rely on practical and appropriate simulation tools that are sensitive to changes, can assimilate different types of change information and flexible enough to accommodate improvements in understanding of change. They need to include representations of uncertainty and generate information appropriate for uncertain decision-making. This paper presents some examples of the tools that have been developed to address these issues in the southern Africa region. The examples include uncertainty in present day simulations due to lack of understanding and data, using climate change projection data from multiple climate models and future catchment responses due to both climate and development effects. The conclusions are that the tools and models are largely available and what we need is more reliable forcing and model evaluation information as well as methods of making decisions with such inevitably uncertain information.

1 Introduction

The previous (PUB) and present (Panta Rhei) science decades of the International Association of Hydrological Sciences (IAHS) recognised the need for improved understanding of catchment hydrology, incorporating different types of uncertainty into hydrological estimations and for translating science advances into practice (Hrachovitz et al., 2013; Montanari et al., 2013; Pomeroy et al., 2014). From the perspective of practising water resources engineers, the first two objectives may seem incompatible with the third (Pappenberger and Beven, 2006). However, the future flow regimes of rivers will be different to today and these differences are highly uncertain due to imperfect knowledge of present and future climate variations, rainfall–runoff processes and anthropogenic impacts. If future water resources development decisions are to be made, practical and appropriate simulation tools are needed, which represent natural hydrological functions and anthropogenic impacts. To be appropriate they need to be sensitive to changes, able to assimilate different types of change information and flexible enough to accommodate improvements in our understanding of existing and future processes. They need to include methods of represent-

ing uncertainty and generate outputs that provide information useful for uncertain decision-making methods (Matrosov et al., 2013).

This paper presents some examples of the tools that have been developed to address these issues within the southern Africa region. The examples are designed to demonstrate methods of dealing with uncertainty in present day situations due to lack of hydrological understanding and data associated with both natural and developed water resources. The examples also include the use of climate change projection data from multiple climate models and possible future catchment response changes due to both climate and development effects. All of the tools are developed to support the application of one of the most widely used hydrological models in southern Africa; the Pitman model (Hughes, 2013). The model is a semi-distributed, monthly time step, conceptual, rainfall–runoff model that also includes functions to represent anthropogenic impacts on water resources availability, including managed forestry plantations, small farm dams and large reservoirs (and associated abstractions and operating rules), as well as direct river abstractions.

2 Uncertainties in natural hydrological responses

Kapangaziwiri et al. (2012) presented an overview of an approach to uncertainty analysis that was based on simple Monte Carlo sampling of the parameter space that could be defined in different ways. The output ensembles were then examined relative to regional constraints on assumed hydrological behaviour following the approach suggested by Yadav et al. (2007). One problem that emerged in large basins with multiple sub-basins was that downstream ensembles that fall within the constraint bounds could be made up of combinations of upstream ensembles that are both within and outside the constraint bounds. Tumbo and Hughes (2015) present a revised approach for the 86 000 km² Great Ruaha River basin in Tanzania. The first step uses regional or local (i.e. based on gauge data for a specific sub-basin) constraints to determine up to 5000 uncertain parameter sets that produce sub-basin incremental flow patterns that fall within pre-defined constraint bounds and are therefore considered behavioural. The second step samples from the saved parameter sets and runs the model 10 000 times for the basin as a whole with all of the downstream routing links. From both science and practical perspectives the advantage is that all of the incremental flow inputs are known to be behavioural, and the downstream outputs are simply different combinations of upstream behavioural inflows. A further advantage is that different levels of uncertainty can be applied at different points within the basin, depending on the amount of, or confidence in, the available data used to determine the constraints. The focus of attention for reducing uncertainty should therefore be on those sub-basins with wide constraint uncertainty bounds and the type of information required to reduce the bounds.

Tumbo and Hughes (2015) discuss some of the practical issues with setting up an uncertainty model of this type. These include ensuring that the constraints are compatible with each other, an example being to ensure that the groundwater recharge and low flow (Q_{10}) constraints are compatible. They also include ensuring that the a priori input parameter ranges are compatible with the constraints such that 5000 behavioural results will be generated when the parameter are independently randomly sampled within a total of 100 000 model runs. A further practical issue is the length of time needed for a model run. Given a basin setup that includes 30 sub-basins and a time series length of 85 years, the normal run time for the second step is approximately 4 h on a standard PC. Tristram et al. (2014) demonstrated that this time could be drastically reduced by using graphics processors and re-coded software, while additional re-coding is underway to make use of the multiple processing cores on standard modern PCs. Initial trials suggest improvements in speed of about 3.5 times.

3 Adding water use and associated uncertainties

During the second step of the uncertainty model run, additional uncertainty in parameters that are not included during the first step (designed to focus on natural hydrology simulations) can be added. This is where uncertainties in the impacts of land use change, reservoir storage, abstractions and return flows, etc. can be added to simulate various scenarios of development. Hughes and Mantel (2010) investigated the uncertainties in simulating changes related to irrigation from small farm dams in three catchments of South Africa. The impacts were found to be variable, as was the confidence in the information that was available to quantify the volumes of storage and abstractions. Nevertheless, the Pitman model was able to generate realistic uncertainty bounds that bracketed the observed flows. Hughes (2014) used an unconstrained uncertainty version of the Pitman model to investigate how well the model can replicate temporal variability in observed data using two examples from Australia, one from Mali and one from South Africa. The approach adopted was to generate 10 000 ensembles and compare the optimum parameter sets based on objective functions using different periods of the observed data. The results indicated that the model was robust enough to simulate temporal variability in the Australian and Mali catchments that were largely caused by climatic variations, as well as variability in the South African catchment that was attributed to commercial afforestation under more-or-less stationary climate conditions. The South African catchment (NoordKaap River in Mpumalanga Province) is re-visited in this study.

The catchment is 126 km² in area, with approximately 70 % forest cover at the present time, and has stream flow data records extending back to 1948, when it is assumed that afforestation (mixed eucalypts and pines) activities began. Unfortunately, the historical details of the afforestation activities are not available but there are clear signals in the observed stream flow data that reductions in flow were already occurring in the early 1960s. The results of the Hughes (2014) uncertainty simulations were used to extract representative parameter sets for the first (no forest) and last (70 % forested) parts of the stream flow record and the simulations compared to appropriate periods of the observed flow data (Table 1). If both simulations are compared with the full observed data set the results are very poor as evidenced by the flow duration curves (FDCs) shown in Fig. 1. In terms of both FDCs and time series comparisons, parameter set 1 results match the observed well until about 1962 and then begin to progressively over-estimate. Parameter set 2 results strongly under-estimate up to about 1962, and then start matching the observed data far better after the mid-1970s. The two simulated time series were then merged using a simple linear approach:

Table 1. Model parameters and results for the NoordKapp River for the two simulation periods.

Parameters/Obj. functions	No forest	70 % forest	Explanation
Obs. data period	Oct 1947– Sep 1962	Oct 1990– Sep 2005	Observed data period used to calculate the objective functions
PI (mm) FF (–)	1.5 0.0	4.0 1.4	Interception (PI) and forest evaporation scaling factor (FF).
ZMIN (mm month ⁻¹) ZMAX (mm month ⁻¹)	20.0 600.0	140.0 1000.0	Limits of the triangular distribution of catchment absorption that are used to estimate surface runoff.
FT (mm month ⁻¹) GW (mm month ⁻¹)	60.0 40.0	30.0 40.0	Maximum interflow runoff (FT) and groundwater recharge (GW) rates.
R (–)	0.5	0.0	Actual evapotranspiration parameter (lower value gives more effective soil storage drying).
Riparian (%)	0.2	0.6	Riparian strip determining rates of evapotranspiration from groundwater.
NSE NSE {ln}	0.490 0.682	0.647 0.704	Nash–Sutcliffe efficiency statistic. No transformation (NSE) and natural log transformation (NSE {ln}).
% Bias % Bias {ln}	0.55 2.68	3.05 31.40	% bias in the mean monthly flows for untransformed and natural log transformed data.

- up to DATE1: $MS_j = S1_j$;
- between DATE1 and DATE2:
 $MS_j = S1_j \cdot x_j + S2_j \cdot (1 - x_j)$; $x_j = M_j / \Sigma M_j$;
- after DATE2: $MS_j = S2_j$

where $S1_j$ (no forest) and $S2_j$ (70 % forest) are the two simulated time series for months j , MS_j is the merged series and x_j represents the merging weight for the two simulations. M_j is the month count after DATE1, while ΣM_j represents the total number of months between DATE1 and DATE2. DATE1 and DATE2 were changed to find values that gave the best objective functions for MS_j and the total observed data period. There were few differences in the overall statistics if DATE1 was near the end of the 1950s and DATE2 was between the start and end of the 1970s. Starting the merge in October 1959 and ending in September 1975 gave one of the best results and this has been included in Fig. 1. The Nash–Sutcliffe statistics based on no transformation (NSE) and natural log transformations (NSE {ln}) are 0.590 and 0.745, respectively, which compare quite favourably with the shorter period “calibration” results given in Table 1.

From a practical perspective, the NoordKaap example represents a relatively simple way to explore different impacts of water use (in this case through land use change effects) over different historical periods. The simulations reported in

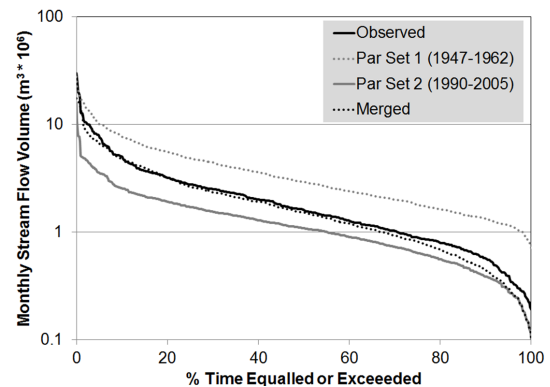


Figure 1. Flow duration curves for the NoordKaap River example based on data for 1948 to 2005, including observed, simulated using parameter set 1 (calibrated for 1947 to 1962; no afforestation) and parameter set 2 (calibrated for 1990 to 2005; 70 % afforestation) and a merged simulated time series (see text for details).

Hughes (2014) were based on uncertainty analysis and the exploration of multiple possible parameter sets, while this study has applied a simple linear merging approach. The parameters in Table 1 suggest that before afforestation there were greater amounts of surface and interflow runoff coupled with lower rates of evapotranspiration loss. These results are

consistent with existing concepts about the impacts of commercial forest plantations in South Africa (Görgens and Van Wilgen, 2004). It is possible that the $S1_j$ simulations do not represent natural conditions, but are more consistent with preparation and early-planting conditions associated with an increase in access roads and drainage, together with some clearance of the natural vegetation (bush and thicket). Unfortunately, all similar gauged catchments in this region are also forested and there are no available data to determine parameter sets for truly natural conditions.

4 Adding climate change impacts and uncertainties

Hughes et al. (2014) refer to a previous approach to incorporating climate change uncertainties into hydrological modelling. This was based on the use of rainfall and temperature data for 9 downscaled GCMs obtained from the Climate Systems Analysis Group (CSAG) of the University of Cape Town (Hewitson and Crane, 2006). These consist of daily rainfall, maximum and minimum temperature for baseline (1961 to 2000), near-future (2046 to 2065) and far-future (2081 to 2100) periods. As the statistical characteristics of the baseline rainfall simulations are very different, across the 9 GCMs, to the available historical data, Hughes et al. (2014) proposed a bias correction approach to generate corrected near-future rainfall time series that could be used with a hydrological model established using historical rainfall data. There was less difference in the predicted temperature scenarios across the 9 models. There are three potential practical problems with this approach. The first is that the sequences of near-future simulated flows for the 9 GCMs cannot be compared because they all have different starting conditions. The second is that the near-future simulations are only 20 years compared with the 85 year (1920 to 2005) historical data simulations that are currently used in South Africa for water resources availability assessments. The third is that there is quite a lot of data preparation and the hydrological model has to be run 9 times.

An alternative approach has been adopted in this study that is based on generating 500 rainfall ensembles for the same period as the historical data and that are designed to reflect uncertainties in future rainfall given the changes between the baseline and near-future patterns predicted by the 9 downscaled GCMs. The rainfall ensembles are generated by a single pre-processing program, after which an uncertainty version of the Pitman model is run that combines the 500 rainfall ensembles with 500 random parameter samples to generate a total of 250 000 uncertainty ensembles. Given the similarities in the predictions of future temperatures, the historical evaporation demand seasonal distributions are scaled by the average of the scaling factors (baseline to near-future) suggested by the 9 GCMs (see Hughes et al., 2014 for more details).

The pre-processing program calculates the calendar month means and standard deviations for all the rainfall data using

a square root transformation, which was found by Hughes et al. (2014) to give the lowest skewness values (and therefore the best approximation to Normal distributions) across all calendar months and GCMs. The delta change in means (ΔM_{ji}) and standard deviations (ΔSD_{ji}) are then calculated for each calendar month (j) and GCM (i) and their ranges determined across all 9 GCMs (ΔM_{min_j} to ΔM_{max_j} and ΔSD_{min_j} to ΔSD_{max_j}). The historical means and standard deviations are also calculated (HM_j and HSD_j). The 500 rainfall ensembles are then generated using Monte Carlo samples from the range of delta changes that are applied to the historical transformed data (HP_k). For each ensemble a fixed random value (FR) between 0.1 and 0.9 is generated and an additional random value between -0.1 and 0.1 generated for each month (k) of the time series and for the means and standard deviation changes (MR_k and SDR_k). The rainfall value for each ensemble month is then determined using steps (i) to (v):

- i. Determine the historical standard variate based on the appropriate calendar month mean and standard deviation: $HV_k = (HP_k - HM_j)/HSD_j$.

- ii. Generate the random delta change values for the calendar month mean and standard deviations:

$$\Delta M_k = (FR + MR_k) \cdot (\Delta M_{max_j} - \Delta M_{min_j}) + \Delta M_{min_j}$$

$$\Delta SD_k = (FR + SDR_k) \cdot (\Delta SD_{max_j} - \Delta SD_{min_j}) + \Delta SD_{min_j}$$

- iii. Apply the delta change values to the historical calendar month (j) means and standard deviations to generate the future ensemble values:

$$FM_j = HM_j \cdot \Delta M_k \text{ and } FSD_j = HSD_j \cdot \Delta SD_k$$

- iv. Re-scale the historical standard variate by the future ensemble means and standard deviations:

$$FP_k = FM_j + FSD_j \cdot HV_k$$

- v. Finally, back-transform the FP_k values.

The fixed random value (FR) ensures that each ensemble is not a complete mixture of delta changes from all GCMs, while the use of the same random number for the mean and standard deviation changes assumes that a GCM with higher means will have higher standard deviations, an assumption that is only partly supported by an analysis of the GCM data. The additional random values (MR_k and SDR_k) were added to allow for some uncertainties in the delta change values within a single ensemble and the need for this requires further evaluation. The method assumes that the historical sequences

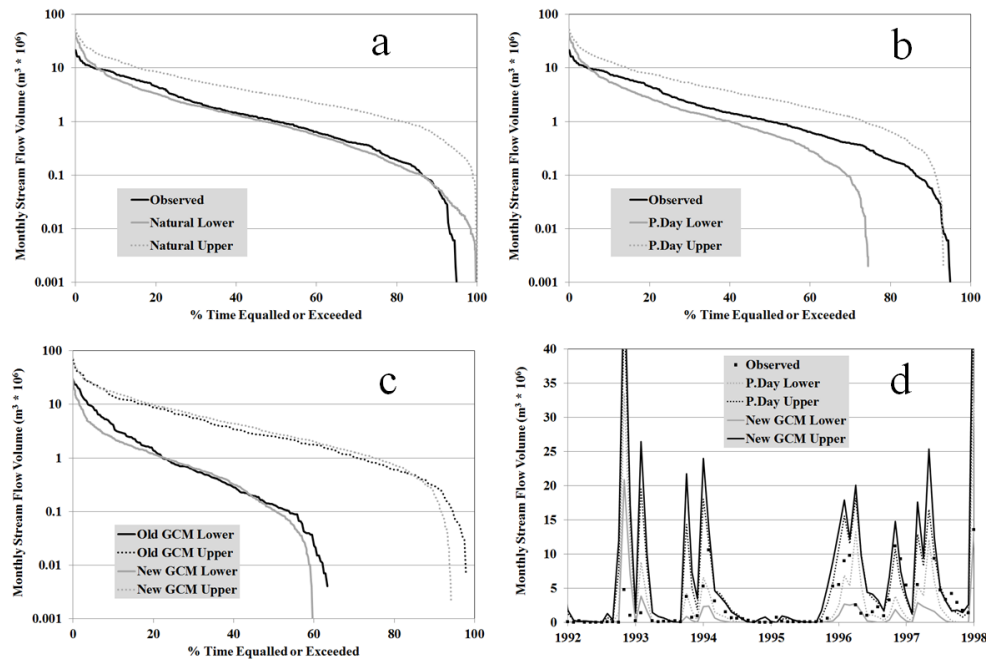


Figure 2. Upper Caledon uncertainty example. (a) Natural uncertainty bounds; (b) uncertainty bounds including present day water use; (c) near-future climate change uncertainty bounds; (d) representative time series with present day and climate change uncertainty bounds.

of wet and dry periods will be more-or-less preserved into the future and this appears to be valid for South Africa (Hughes et al., 2014).

The approach has been applied to a headwater sub-catchment (474 km²) of the Caledon River, a major tributary of the Orange River. The Caledon River was the subject of an earlier study of model uncertainty and climate change impacts on hydrology (PhD thesis in preparation) and the results of using bias corrected rainfall data for the near-future period and all 9 GCMs were available for comparison purposes. The natural hydrology is impacted by abstractions directly from the river and from many small farm dams (for irrigation and domestic water). Figure 2 illustrates the results of applying the uncertainty model to (a) historical natural conditions, (b) present day historical conditions, and (c) both methods of accounting for climate change with present day water use. Figure 2a–c all use the total number of available months to construct flow durations curves (FDCs), while Fig. 2d shows a short time series comparing the observed data (plotted as points to show missing data) with the uncertainty bounds under historical and near-future climate change conditions. Figure 2b indicates that the uncertainty bounds generally bracket the observed data except for high flows, which are expected to be under-measured due to rating curve limitations. The proximity of the observed data to the upper uncertainty bound is likely a result of non-stationarities in the observed data (increased water use in more recent years). Figure 2c shows little difference in the results of running the model 9 times using the bias corrected near-future GCM

data (“Old GCM”) compared to pre-processing to determine the range of delta changes, generating rainfall ensembles and then running the model once with a combination of rainfall and parameter sampling (“New GCM”). The FDCs for the “Old GCM” are the minimum and maximum values of the uncertain FDCs generated for all 9 GCMs.

5 Discussion and conclusions

It was not the main purpose of this paper to demonstrate the validity of the Pitman model for simulating water resources assessments in southern Africa as this has been done a number of times in the past (Hughes, 2013). The purpose was to illustrate the practical application of the model where the scientific issues of uncertainty analysis and potential change (Montanari et al., 2013) are not neglected. The 2-step approach for simulating uncertain natural hydrological regimes (Tumbo and Hughes, 2015) offers a practical and flexible approach to uncertainty analysis of natural hydrology as well as anthropogenic impacts. The NoordKaap afforestation example illustrates that the model is sensitive to at least some land use impacts and that a simple approach can be used to merge natural and changed conditions over an extended time period. If necessary it would be straightforward to replace the simple linear merging used here, with a more complex non-linear merging process involving more than two simulations.

The approach used to incorporate climate change effects based on multiple climate model outputs is also simple and can be applied to as many climate models or emission sce-

narios as required. With the previous approach (Hughes et al., 2014), increasing the number of climate models would increase the number of model runs, while the approach presented is based on a single pre-processing program and one run of the model. It assumes that all of the GCM predictions are equally likely and that the effects of temperature on evaporation demand are similar for all GCMs. It is possible that future GCM outputs will change as new methods of simulating climate and of downscaling are generated by the climate science community (Kalognomou et al., 2013). The new approach makes it far easier to incorporate these updates into existing model setups. The approach could possibly be further enhanced by using a stochastic rainfall generator (Srikanthan and Pegram, 2009), where more statistical properties of the climate change projections are accounted for than just the calendar month means and standard deviations. The example provided uses the same uncertainty parameter sets that were established for historical conditions. However, if there is evidence to suggest that atmosphere-land feedback mechanisms exist (such as changes to root zone storage; Gao et al., 2014), then these can be incorporated as changes in the parameter ranges.

Ultimately, the practical value of the simulations using the methods described here are highly dependent upon the data used to set up the model (climate forcing data, local and regional output constraint information and water use or land use change information). The value is also dependent upon appropriate decision-making processes that can incorporate uncertainty (Matrosov et al., 2013).

Acknowledgements. The author is grateful to T. Mohobane for the use of the results from the earlier climate change assessments on the Caledon River catchment and to the Climate Systems Analysis Group of the University of Cape Town for the downscaled GCM data. Some of the research for this paper was supported by the Water Research Commission of South Africa, while T. Mohobane was supported by a PhD bursary by the Carnegie Corporation of New York under the Regional Initiative for Science Education (RISE) programme.

References

- Gao, H., Hrachowitz, M., Schymanski, S. J., Fenicia, F., Sriwongstanon, N., and Savenije, H. H. G.: Climate controls how ecosystems size the root zone storage capacity at catchment scale, *Geophys. Res. Lett.*, 41, 7916–7923, 2014.
- Görgens, A. H. M. and Van Wilgen, B. W.: Invasive alien plants and water resources in South Africa: Current understanding, predictive ability and research challenges, *S. Afr. J. Sci.*, 100, 27–33, 2004.
- Hewitson, B. and Crane, R.: Consensus between GCM climate change projections with empirical downscaling: precipitation downscaling over South Africa, *Int. J. Climatol.*, 26, 1315–1337, 2006.
- Hrachowitz, M., Savenije, H. H. G., Blöschl, G., McDonnell, J. J., Sivapalan, M., Pomeroy, J. W., Arheimer, B., Blume, T., Clark, M. P., Ehret, U., Fenicia, F., Freer, J. E., Gelfan, A., Gupta, H. V., Hughes, D. A., Hut, R. W., Montanari, A., Pande, S., Tetzlaff, D., Uhlenbrook, S., Wagener, T., Winsemius, H. C., and Woods, R. A.: A decade of Predictions in Ungauged Basins (PUB) – a review, *Hydrolog. Sci. J.*, 58, 1198–1255, 2013.
- Hughes, D. A.: A review of 40 years of hydrological science and practice in southern Africa using the Pitman rainfall–runoff model, *J. Hydrol.*, 501, 111–124, 2013.
- Hughes, D. A.: Simulating temporal variability in catchment response using a monthly rainfall–runoff model, *Hydrolog. Sci. J.*, doi:10.1080/02626667.2014.909598, in press, 2014.
- Hughes, D. A. and Mantel, S. K.: Estimating the uncertainty in the impacts of small farm dams on stream flow regimes in South Africa, *Hydrolog. Sci. J.*, 55, 578–592, 2010.
- Hughes, D. A., Mantel, S., and Mohobane, T.: An assessment of the skill of downscaled GCM outputs in simulating historical patterns of rainfall variability, *Hydrol. Res.*, 45, 134–147, 2014.
- Kalognomou, E.-A., Lennard, C., Shongwe, M., Pinto, I., Favre, A., Kent, M., Hewitson, B., Dosio, A., Nikulin, G., Panitz, H.-J., and Büchner, M.: A diagnostic evaluation of precipitation in CORDEX models over Southern Africa, *J. Climate*, 26, 9477–9506, 2013.
- Kapangaziwiri, E., Hughes, D. A., and Wagener, T.: Constraining uncertainty in hydrological predictions for ungauged basins in southern Africa, *Hydrolog. Sci. J.*, 57, 1000–1019, 2012.
- Matrosov, E. S., Woods, A. M., and Harou, J. J.: Robust Decision Making and Info-Gap Decision Theory for water resource system planning, *J. Hydrol.*, 494, 43–58, 2013.
- Montanari, A., Young, G., Savenije, H., Hughes, D. A., Wagener, T., Ren, L., Koutsoyiannis, D., Cudennec, C., Grimaldi, S., Bloeschl, G., Sivapalan, M., Beven, K., Gupta, H., Arheimer, B., Huang, Y., Schumann, A., Post, D., Srinivasan, V., Boegh, E., Hubert, P., Harman, C., Thompson, S., Rogger, M., Hipse, M., Toth, E., Viglione, A., Di Baldassarre, G., Schaeffli, B., McMillan, H., Schymanski, S. J., Characklis, G., Yu, B., Pang, Z., and Belyaev, V.: “Panta Rhei – Everything Flows”: Change in hydrology and society – The IAHS Scientific Decade 2013–2022, *Hydrolog. Sci. J.*, 58, 1256–1275, 2013.
- Pappenberger, F. and Beven, K.: Ignorance is bliss: Or seven reasons not to use uncertainty analysis, *Water Resour. Res.*, 42, W05302, doi:10.1029/2005WR004820, 2006.
- Pomeroy, J. W., Spence, C., and Whitfield, P. H. (Eds.): Putting Prediction in Ungauged Basins into Practice, Canadian Water Resources Association, Canada, 2014.
- Srikanthan, R. and Pegram, G. G. S.: A nested multisite daily rainfall stochastic generation model, *J. Hydrol.*, 371, 142–153, 2009.
- Tristram, D., Hughes, D. A., and Bradshaw, K.: Accelerating a hydrological uncertainty ensemble model using Graphics Processing Units (GPUs), *Comput. Geosci.*, 62, 178–186, 2014.
- Tumbo, M. and Hughes, D. A.: Uncertain hydrological modelling: Application of the Pitman model in the Great Ruaha River Basin, Tanzania, *Hydrolog. Sci. J.*, doi:10.1080/02626667.2015.1016948, in press, 2015.
- Yadav, M., Wagener, T., and Gupta, H. V.: Regionalisation of constraints on expected watershed response behaviour, *Adv. Water Resour.*, 30, 1756–1774, 2007.