

# Regionalising low-flow responses in large tropical catchments: a comparison of rainfall-runoff modelling and regression approaches

C. Petheram<sup>a</sup>, P. Rustomji<sup>a</sup>, B. Neal<sup>b</sup> and A. Woodman<sup>b</sup>

<sup>a</sup> CSIRO Land and Water

<sup>b</sup> Sinclair Knight Merz

Email: [cuan.petheram@csiro.au](mailto:cuan.petheram@csiro.au)

**Abstract:** A prolonged drought across southern Australia has led to renewed interest in water resource development of northern Australia (NA), with an increase in demand for runoff predictions from ungauged catchments in Australia's tropics. However, wet-dry tropical environments have an extended dry season with periods of low or no flow, and from an ecological perspective, dry-season flows are of vital importance (e.g. Erskine et al., 2003). The dry-season is also the period when potential consumptive water use is likely to be high. Thus it is important to ensure dry-season flows are well predicted.

One of the key challenges in undertaking robust water resource assessments across NA is the relatively low density of hydrological data. Using data from 105 catchments in tropical Australia, five daily rainfall-runoff models (RRM) and three methods of regionalising model parameters were compared for the simulation of dry-season flows. To ensure low-flows were well modelled the approach of Petheram et al. (2012) was adopted, where the simulated and observed terms in the objective function were raised to the power of  $\lambda$ . Here six calibration runs were undertaken for each model, using values of 1.25, 1, 0.75, 0.5, 0.25 and 0.05 for  $\lambda$ . To select the 'best all-round' parameter set we used the method of Petheram et al. (2012), where for each model and for each catchment the 'best all-round' calibrated parameter set was selected based on the weighted combination of different Nash-Sutcliffe Efficiency metrics. We then compared the best performing RRM and best method of regionalisation against regression-based predictive methods for a wet-dry tropical environment.

We found that the adoption of multiple criteria to select an optimal parameter set resulted in an improved ability to simulate low flows with no loss in predictive capacity for high flows. An educated transposition of parameter sets from gauged to ungauged catchments was found to be better than random assignment of model parameters, whilst assigning model parameters on the basis of spatial proximity outperformed physical similarity methods. For simulating the lower-half of the flow duration curve no clear method was best for regionalising model parameters and no method was better than randomly assigning intact model parameter sets. The best performing multi-model ensemble (Sacramento and IharesClassic) and the best method of parameter regionalisation (spatial proximity) performed similarly to statistical regression approaches in predicting mean annual flows. However, the regression approaches demonstrated more skill predicting low-flow metrics. There may be opportunities to improve low-flow NSE metrics under calibration mode through 'smarter' calibration procedures. However, it is thought unlikely that this will result in improved performance under prediction mode.

Erskine, W.E., G.W. Begg, P. Jolly, A. Georges A., A. O'Grady, D. Emaus, N. Rea, P. Dostine, S. Townsend, A. and Padovan (2003). Recommended environmental water requirements for the Daly River, Northern Territory, based on ecological, hydrological and biological principles., Darwin, NT.

Petheram, C., P. Rustomji, F.H.S. Chiew, and J. Vleehsouwer (2012). Rainfall-runoff modelling in northern Australia: a guide to modelling strategies in the tropics. Submitted to special issue on Tropical Hydrology, Journal of Hydrology.

**Keywords:** *Low-flow, rainfall-runoff modelling, northern Australia, regionalisation, regression analysis*

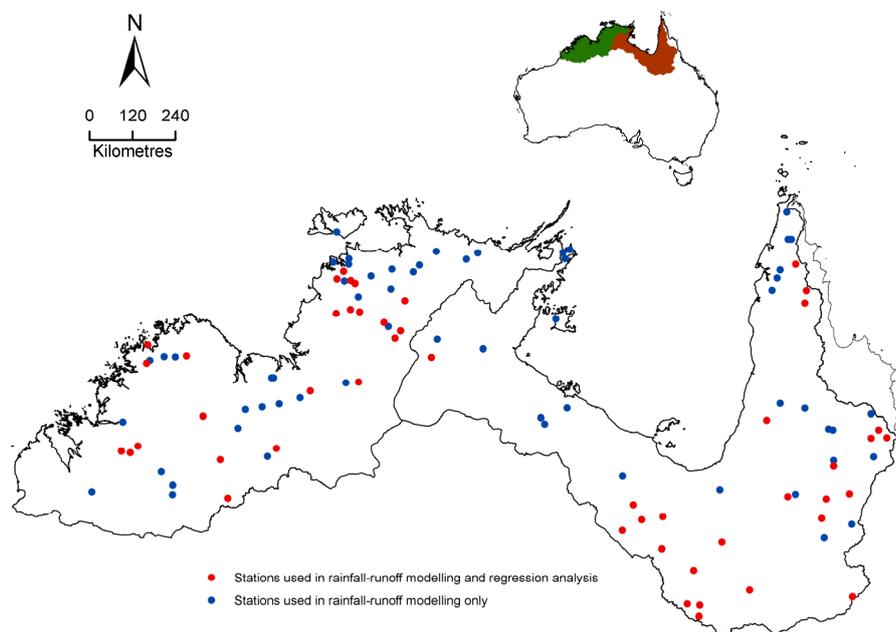
## 1. INTRODUCTION

A prolonged drought across southern Australia has led to renewed interest in water resource development of northern Australia (NA), with an increase in demand for runoff predictions from ungauged catchments in Australia's tropics. However, wet-dry tropical environments have an extended dry season with periods of low or no flow, and from an ecological perspective, dry-season flows are of vital importance (e.g. Erskine *et al.*, 2003). The dry-season is also the period when potential consumptive water use is likely to be high. Thus it is important to ensure dry-season flows are well predicted.

Using data from 105 catchments in tropical Australia, five daily rainfall-runoff models (RRM) and three methods of regionalising model parameters were compared for the simulation of dry-season flows. However, it is well recognised that conceptual RRM do not simulate low-flows as well as mid-to-high flows (e.g. Chiew 2010). In a study of 21 catchments in south-western Norway, Engeland and Hisdal (2009) found regression methods gave better low flow estimates than the HBV semi-distributed conceptual RRM. Here we compare the best performing RRM and method of regionalisation against regression based predictive methods for a wet-dry tropical environment.

### 1.1. Study area

The study region includes Australia's Timor Sea, Gulf of Carpentaria and the most northern section of the North-East Coast Drainage Divisions (Figure 1) – covering an area of over 1.25 million km<sup>2</sup>. Mean annual rainfall varies across the study area by more than an order of magnitude, from about 400 mm yr<sup>-1</sup> south of the Gulf of Carpentaria, to over 4000 mm yr<sup>-1</sup> on the steep coastal escarpments in the North-East Coast Drainage Division. More than 90 percent of rain falls between November and March as the inter-tropical convergence zone extends south and passes over the northern extent of the continent. Inter-annual variability of mean annual rainfall and streamflow are higher than in other parts of the world of the same climate type (Petheram *et al.*, 2008). Baseflow indices and a lag-one correlation coefficient of annual streamflow are very low (Petheram *et al.*, 2008), and together with the generally steep shape of flow exceedance curves, suggest that most of NA has limited hydrologic storage capacity. Exceptions are the rivers set in the Early to Middle Palaeozoic carbonate rocks in the Northern Territory and western Queensland, Cretaceous sandstones in the Northern Territory and Quaternary sands of northern Cape York which exhibit perennial flow (Petheram and Bristow 2008).



**Figure 1 Northern Australia and location of gauging stations used in rainfall-runoff and regression analysis. Inset: Green and brown shading indicate Timor Sea and Gulf of Carpentaria Drainage Divisions respectively.**

## 2. METHOD

### 2.1. Rainfall-runoff model calibration

A database of 105 streamflow gauging stations was compiled (Figure 1). Stations were required to have at least 10 years data and 4 years of wet season data. Sixty percent of catchments had an area greater than 1000 km<sup>2</sup> and 15 percent of catchments had an area greater than 10,000 km<sup>2</sup>.

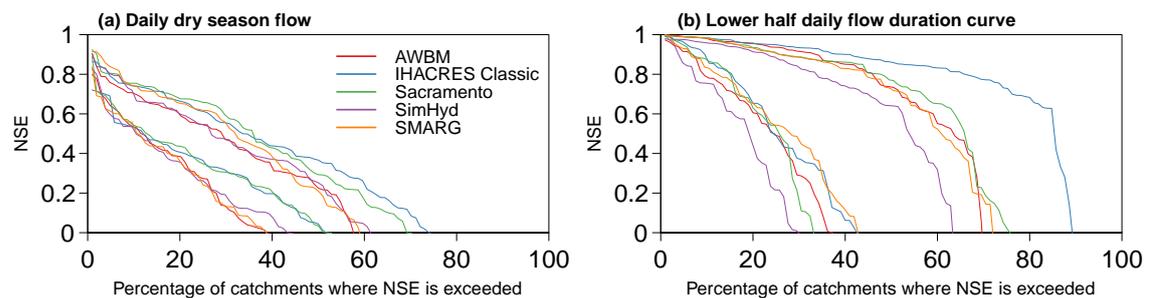
Five lumped conceptual daily RRM were trialled to determine the model or model-combination that best modelled low-flow runoff across NA. In order of decreasing complexity, these models were Sacramento (Burnash *et al.*, 1973), SMARG (Goswami *et al.*, 2002), IHACRES Classic (Croke 2006), SIMHYD (Chiew *et al.*, 2002) and AWBM (Boughton 2004). The rainfall-runoff modelling utilised gridded daily rainfall and Areal Potential Evapotranspiration (APET) data at 0.05 degree resolution (approximately 5 x 5 km), generated from the gridded Silo climate data ([www.nrm.qld.gov.au/silo](http://www.nrm.qld.gov.au/silo)).

Each RRM was calibrated against observed streamflow data from unimpaired catchments with the same parameter values used for all grid cells within a catchment, with a one year warm-up period. Calibration was carried out by minimising the sum of square errors between the simulated and observed values of daily streamflow. To ensure low-flows were well modelled the approach of Petheram *et al.* (2012) was adopted, where the simulated and observed terms in the objective function were raised to the power of  $\lambda$ . Here six calibration runs were undertaken for each model, using values of 1.25, 1, 0.75, 0.5, 0.25 and 0.05 for  $\lambda$ . Values of  $\lambda$  less than 1 generally reduce the relative influence of high-flows in the objective function. The large number of catchments meant that an automated calibration procedure was required. Calibration was automated using two methods applied sequentially: 1) shuffled complex evolution (Duan *et al.*, 1993); and 2) the Rosenbrock method (Rosenbrock, 1960). A log-bias constraint (Viney *et al.*, 2009) was employed to ensure that the total flow volumes were well modelled. Model performance was assessed in terms of Nash-Sutcliffe Efficiency (NSE) and model bias.

To select the ‘best all-round’ parameter set we used the method of Petheram *et al.* (2012), where for each model and for each catchment the ‘best all-round’ calibrated parameter set (denoted by  $\Omega$ ) was selected from the six calibrated sets based upon the highest  $NSE_w$  value as computed by Equation 1.

$$NSE_w = \frac{(10 \times NSE_D + 2 \times NSE_{Ddry} + NSE_{Dfdc} + NSE_M + NSE_{Mdry} + NSE_{Mfdc} + NSE_{LHfdc})}{17} \quad (1)$$

In Equation 1 the subscripts D, Ddry, Dfdc, M, Mdry, Mfdc and LHfdc correspond to daily flow, daily dry season flow, daily flow- duration- curve, monthly flow, monthly dry season flow, monthly flow- duration- curve and lower-half of the daily flow- duration- curve (i.e. compares observed and predicted between the 50 and 100 percentiles of daily flow). In this manuscript the term ‘calibration mode’ refers to the simulation of streamflow in a catchment using a locally derived parameter set. The term ‘prediction mode’ refers to the simulation of streamflow in a catchment using a parameter set donated from another catchment.



**Figure 2 Summary of calibration (top cluster of lines) and prediction (bottom cluster of lines; spatial proximity method) for 2 low-flow NSE measures. From Petheram *et al.* (2012)**

### 2.2 Comparing models and methods

To compare model performance in both calibration and prediction modes across each of the different NSE measures, a single metric was devised to summarise the overall performance of a model or method for all 105 catchments. This metric, denoted by  $NSE-A_{75}$ , was computed as the average NSE of the best 75 percent of catchment models, where catchments with NSE values of less than zero were assigned a NSE of zero.

Three methods of regionalising ‘intact’ model parameter sets were examined; spatial proximity, catchment similarity and random assignment. Catchment similarity method assigned donor-target catchment pairs on the basis of 1) the summed ranking of 6 catchment and climate attributes (drainage density, elevation, slope, mean annual rainfall (MAR), mean dry season rainfall (dry), 20th percentile (P20) daily rainfall); or 2) the ranking of individual catchment and climate attributes.

The spatial proximity and physical similarity regionalisation approaches of assigning donor–target pairs were compared to the 2.5th and 97.5th percentile values from 500 randomly selected donor –target pairs. All three methods of assigning donor-target pairs transposed whole parameter sets.

### 2.3 Regression analysis

A multiple linear regression analysis for prediction of hydrological metrics in NA was undertaken by SKM (2009) using a similar selection of gauging stations to those used in this analysis. SKM (2009) imposed the criteria that gauging stations had more than 90 percent of their record complete between 1972 and 1987. Hydrologic metrics were derived from these observed data between 1972 and 1987 and these were regressed against 34 physiological and seven climate attributes. The hydrological metrics investigated were:

- mean annual, wet season and dry season flow
- coefficient of variation of annual flow
- 10th percentile daily flow
- 20th percentile daily flow
- 50th percentile daily flow
- 80th percentile daily flow
- percentage of time a gauge was at cease-to-flow level
- baseflow index (BFI).

A comparison of the rainfall-runoff modelling and regression approaches was undertaken using 49 gauging stations that were common to both approaches. For the rainfall runoff modelling, the optimal model combination using both locally calibrated parameters and a parameter set transposed from a donor catchment (by the best regionalisation method) was used to generate a modelled daily streamflow timeseries from which the above metrics were computed. These RRM based metrics were then compared to those computed directly from the historic streamflow data (between 1972 and 1987) and also to those computed using the regression model.

Two additional pairs of calibration and cross-verification RRM simulations were undertaken. Each pair varied slightly from the original run (see Table 1). These additional simulations were undertaken to further investigate and illustrate the effect of: 1) modifying the objective function and selecting the best ‘all-round’ calibrated parameter set ( $\Omega$ ); and 2) simulating runoff using parameters only from those catchments that were common to both the rainfall-runoff modelling and multiple linear regression dataset; on the prediction of specific hydrologic metrics.

**Table 1 Rainfall-runoff simulations for comparison with multiple linear regression results**

Model configuration	Power transform	No. catchments used in cross-verification
A <sup>1</sup>	$\lambda = \Omega$	105
B	$\lambda = 1$	105
C <sup>2</sup>	$\lambda = \Omega$	49

1. Original run and anticipated best case.

2. Rainfall-runoff simulations were only undertaken using parameters from those catchments that were common to both the rainfall-runoff modelling dataset and the SKM (2009) multiple linear regression dataset.

## 3. RESULTS

### 3.1. Best method of RRM regionalisation

A comparison of the methods of regionalisation for the five RRMs and the IhacresClassic - Sacramento ensemble is presented in Figure 3. Calibration values are included for context (cross-hair symbol). The range of NSE-A<sub>75</sub> values which captures 95 percent of randomly selected target-donor catchment

combinations is illustrated by the grey bar. What we would expect is that an educated selection of donor – target catchment pairs would result in higher NSE values than the randomised approach. The method of regionalising RRM parameters was insensitive to the exact RRM chosen. For daily dry-season flow the spatial proximity method consistently outperformed the physical similarity approaches. The second best method involved using physical similarity based on all landscape and climate attributes. For simulating the lower-half of the flow duration curve (FDC) there was no consistently best method for transposing model parameters, nor was there a method that was significantly different to randomly assigning donor-target catchments for all model types ( $p < 0.05$ ). More detailed results are provided in Petheram *et al.* (2012).

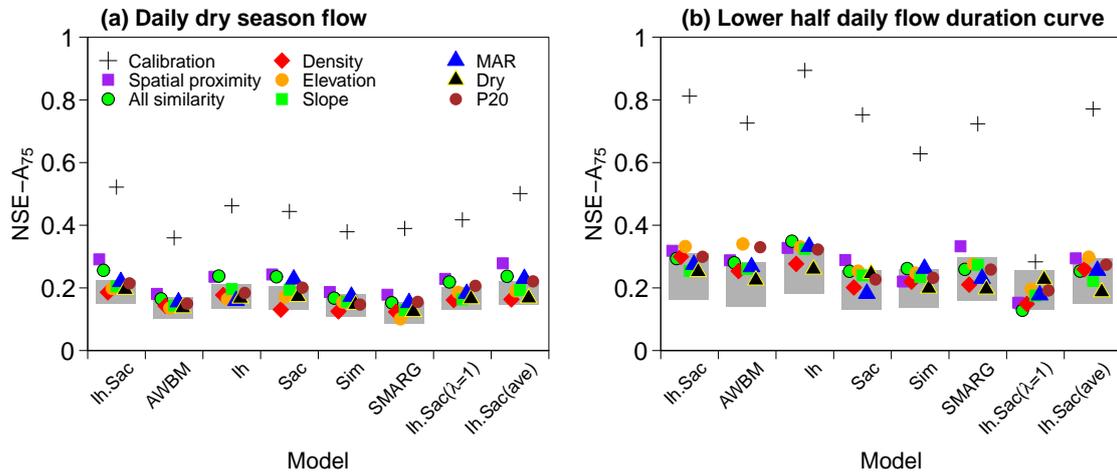


Figure 3 Comparison of regionalisation methods for five models and the IhacresClassic-Sacramento model ensemble. Calibration results shown by cross-hair. Grey box illustrates the 95 percent range (i.e. 2.5 to 97.5 percentile values) of NSE–A<sub>75</sub> values of 500 randomly selected intact parameter sets. From Petheram *et al.* (2012)

### 3.2. Comparison with regression analysis

The best performing multi-model ensemble (Sacramento and IhacresClassic) and the best method of parameter regionalisation (spatial proximity) were used in comparison with the statistical regression approaches. The results are presented in Table 2. It can be seen from this table that the best performing multi-model ensemble (Sacramento and IhacresClassic) and the best method of parameter regionalisation (spatial proximity) performed similarly to statistical regression approaches in predicting mean annual flows. However, the regression approaches demonstrated more skill predicting low-flow metrics.

**Table 2 Comparison of NSE for rainfall-runoff modeling and regression analysis**

Hydrological indicator	Calibration modelled vs observed (NSE)			Prediction modelled vs observed (NSE)			Regression modelled vs observed (NSE)
	A	B	C	A	B	C	
Model configuration							
Mean annual flow	0.98	0.98	0.98	0.79	0.77	0.75	0.77
Mean dry season flow	0.96	0.73	0.96	0.41	0.36	0.59	0.74
50 <sup>th</sup> percentile daily flow	0.96	0.62	0.96	0.16	0.09	0.48	0.31
80 <sup>th</sup> percentile daily flow	0.84	0.36	0.84	-0.06	-0.24	0.09	0.41
Cease-to-flow	0.40	-0.68	0.40	-0.21	-1.03	-0.21	0.58
Base flow index	0.73	0.51	0.73	0.34	0.11	0.30	0.66

## 4. DISCUSSION

The use of RRMs to simulate low flows in the tropical environment of NA appears problematic. Although the models were in part calibrated and then selected on the basis of their ability to simulate low flows, the form of Equation 1 is such that there was still a slight bias towards medium and high flows. Had the objective function been of the same form as the low-flow NSE equations used in this analysis, then the low-

flow NSE measures would have undoubtedly been higher under calibration mode. However, it cannot be taken for granted that this would result in improved NSE results under prediction mode. For example, the difference in  $NSE_{FDC_{lh}-A_{75}}$  and  $NSE_{D_{dry}-A_{75}}$  under calibration and prediction modes was much greater (0.33 and 0.48 respectively) than the difference in medium and high flow NSE measures under calibration and prediction modes (e.g. 0.17 for  $NSE_D$  and 0.08 for  $NSE_{FDC}$ ). Furthermore Table 2 shows that for the Sacramento and IhacresClassic ensemble, the NSE of mean annual flow and NSE mean dry season flow were similar under calibration mode (0.98 cf 0.96 respectively). However, under prediction mode (using spatial proximity method) the RRM's were able to simulate mean annual flow (0.79) much better than they were able to simulate mean dry season flow (0.41). A large difference in  $NSE-A_{75}$  under calibration and prediction modes was also observed for the other low-flow related hydrologic indices, namely 50th and 80th percentile of daily flow, cease-to-flow and BFI (Table 2).

There is no one reason why simulating low flow processes under prediction mode should be worse than under calibration mode (as judged under  $NSE_{FDC_{lh}}$  and  $NSE_{D_{dry}}$ ). It is likely that the spatial proximity and physical similarity methods used here for regionalising model parameters do not adequately capture the spatial variability inherent in those processes in the landscape that represent storage and flow attenuation. In the case of the physical similarity method, it is also likely that inappropriate 'lumped' landscape attributes are being used to judge similarity in low flow hydrology. It may be that we cannot measure, map or capture in a single term the controlling landscape attributes necessary to represent low flow hydrology. It is also highly likely that lumped, conceptual RRM's may have in-built structural limitations that cannot easily be overcome when it comes to simulating low flows.

Although both the rainfall-runoff modelling and regression methods were assessed over the same time period and only using stations common to both approaches, different sets of stations were used to derive the regression equations and transpose RRM parameters. When only those stations common to both approaches were used to transpose RRM parameters (49 as opposed to the original 105 catchments), there was a slight increase to the low flow NSE measures (model configurations A and C in Table 2), indicating that the low flow results are less sensitive to station density than having a similar calibration and assessment time period. Conversely, the high flow NSE metrics under prediction mode were observed to decline slightly when using only those stations common to both the approaches, indicating that station density may be slightly more important than having a similar calibration and assessment time period. Nevertheless the difference between using the subset of 49 stations was slight, indicating the conclusions are relatively robust.

## 5. CONCLUSION

- The adoption of multiple-criteria to select an optimal parameter set resulted in an improved ability to simulate low flows with no loss in predictive capacity for mid to high flows.
- An educated transposition of parameter sets from gauged to ungauged catchments was better than random assignment of model parameters, whilst assigning model parameters on the basis of spatial proximity outperformed physical similarity methods. For simulating the lower-half of the FDC no clear method was best for regionalising model parameters and no method was better than randomly assigning intact model parameter sets.
- The best performing multi-model ensemble (Sacramento and IhacresClassic) and the best method of parameter regionalisation (spatial proximity) performed similarly to statistical regression approaches in predicting mean annual flows. However, the regression approaches demonstrated more skill predicting low-flow metrics.
- There may be opportunities to improve low-flow NSE metrics under calibration mode through 'smarter' calibration procedures. However, it is unlikely that this will result in improved performance under prediction mode.
- There may be opportunities in the future to incorporate regression-based results into the calibration/regionalisation process to improve low-flow simulations

## ACKNOWLEDGEMENTS

The first author was supported by Water in Northern Australia stream of the Water for Healthy Country Flagship. The authors are grateful to the Department of Water (Western Australia), Natural Resources, Environment, the Arts and Sport (Northern Territory) and the Queensland Department of Environment and Resource Management (DERM) for collecting and supplying streamflow data. We acknowledge comments provided by Dr Jai Vaze and Dr Fazul Karim and two anonymous reviewers of the manuscript.

## REFERENCES

- Boughton, W., 2004. The Australian water balance model. *Environmental Modelling & Software*, 19(10): 943-956.
- Burnash, R.J.C., R.L. Ferral, and R.A. McGuire (1973). A generalised streamflow simulation system - conceptual modelling for digital computers., U.S. Department of Commerce, National Weather Service and State of California, Department of Water Resources.
- Chiew, F.H.S., (2010). Lumped Conceptual Rainfall-Runoff Models and Simple Water Balance Methods: Overview and Applications in Ungauged and Data Limited Regions. *Geography Compass*. 4: 206–225, doi:10.1111/j.1749-8198.2009.00318.x
- Chiew, F.H.S., M.C. Peel, and A.W. Western (2002). Application and testing of the simple rainfall-runoff model SIMHYD. In: V.S.a.D. Frevert (Editor), *Mathematical Models of Small Watershed Hydrology and Applications* Water Resources Publication, Littleton, Colorado, pp. 335-367.
- Croke, B.F.W., F. Andrews, A.J. Jakeman, S.M. Cuddy and A. Luddy (2006). IHACRES classic plus: A redesign of the IHACRES rainfall-runoff model. *Environmental Modelling & Software*, 21(3): 426-427.
- Duan, Q., V.K. Gupta, and S. Sorooshian (1993). A shuffled complex evolution approach for effective and efficient global minimization. *Journal of Optimization theory and Applications*, 76(3): 501-521.
- Engeland, K., and H. Hisdal (2009). A Comparison of Low Flow Estimates in Ungauged Catchments Using Regional Regression and the HBV-Model. *Water Resources Management*, 23(12): 2567-2586.
- Erskine, W.E., G.W. Begg, P. Jolly, A. Georges, A. O'Grady, D. Emaus, N. Rea, P. Dostine, S. Townsend, and A. Padovan (2003). Recommended environmental water requirements for the Daly River, Northern Territory, based on ecological, hydrological and biological principles., Darwin, NT.
- Goswami, M., K.M. O'Connor, and A.Y. Shamseldin (2002). Structures and performances of five rainfall-runoff models for continuous river-flow simulation., *Proceedings of 1st Biennial Meeting of International Environmental Modelling and Software Society*, Lugano, Switzerland, pp. 476-481.
- Nash, J. E., and J.V. Sutcliffe (1970). River flow forecasting through conceptual models part I — A discussion of principles, *Journal of Hydrology*, 10 (3), 282–290
- Petheram, C., and K.L. Bristow (2008). Towards an understanding of the hydrological factors, constraints and opportunities for irrigation in northern Australia: A review. CRC for Irrigation Futures Technical Report No. 06/08. CSIRO Land and Water Science Report No. 13/08. February 2008
- Petheram, C., T.A. McMahon, and M.C. Peel (2008). Flow characteristics of rivers in northern Australia: Implications for development. *Journal of Hydrology*, 357(1-2): 93-111.
- Petheram, C., P. Rustomji, F.H.S.S. Chiew, and J. Vleehsouwer, (2012). Rainfall-runoff modelling in northern Australia: a guide to modelling strategies in the tropics. Submitted to special issue on Tropical Hydrology, *Journal of Hydrology*.
- Rosenbrock, H.H., (1960). An automatic method for finding the greatest or least value of a function. *Computer Journal*, 3(3): 175-184.
- SKM (2009). Regionalisation of hydrologic indices. A report prepared by Sinclair Knight Merz for the Northern Australia Sustainable Yield Project. , SKM, Melbourne.
- Viney, N.R., J. Perraud, J. Vaze, F.H.S. Chiew, D.A. Post, and A. Yang (2009). The usefulness of bias constraints in model calibration for regionalisation to ungauged catchments, 18th World IMACS/MODSIM Congress, Cairns, Australia.