Impact of calibration data variability on rainfall-runoff modelling performance in data-limited basins

C.Z. Li [a,b,c], J. Liu [a], F.L. Yu [a], H. Wang [a], N.N. Zhao [a], W.J. Liu [d]

[a] *State Key Laboratory of Simulation and Regulation of Water Cycle in River Basin, China Institute of Water Resources and Hydropower Research, Beijing 100038, P.R. China*
[b] *State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Hohai University, Nanjing 210098, P.R. China*
[c] *State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan 430072, P.R. China*
[d] *State College of Water and Electricity, Hebei University of Engineering, Handan 056021, P.R. China*
Email: lichuanzhe@gmail.com

Abstract: There is a question which comes to the users of hydrological models: which segment of observation data should be used for calibration? Especially when model users face the case of applying models in ungauged or data-limited basins. It is important to make best use of these limited data. Calibration data variability is seldom considered in lumped conceptual rainfall-runoff models although it has significant impacts on modeling performance. In order to study the performance of rainfall-runoff models in data-limited basins (actually, data are non-continuous and fragmented in some basins), we choose to use non-continuous calibration periods to have more independent runoff data for calibration of the SIMHYD model. The Particle Swarm Optimization (PSO) optimization method is used to calibrate the rainfall-runoff models. The Nash-Sutcliffe efficiency (NSE) and percentage water balance error (WBE) are used as performance measures. Average, dry and wet calibration periods are used for study on impact of calibration data variability. Fifty five (55) relatively unimpaired basins all over Australia are tested to obtain general conclusions. The results show that, the rainfall-runoff models have more steady performance when calibrated by average or wet subsets than calibrated by dry subsets, and wet subsets are more suitable for model calibration both in all basins. Calibration data have much impact on arid and semi-arid basins and have little impact on humid and semi-humid basins. The models perform better in humid and semi-humid basins than arid basins. Our results may have useful and interesting implications when hydrological model users at the case of ungauged or data-limited basins.

Keywords: *rainfall-runoff modeling, data-limited, calibration data, variability*

1. INTRODUCTION

Lumped conceptual rainfall-runoff (CRR) models are used widely for different purpose all over the world. Uncertainty associated with CRR models stem from model structure, calibration and validation data, objective functions and assessment criteria, and optimization algorithms. There are many factors affecting optimal parameters. There is a question which comes to the model user's mind: which segment of data should be used for calibration? Especially when model users face the case of applying models in ungauged or data-limited catchments (Boughton, 2007). It's important to make best use of these limited data.

Model performance and optimal parameters estimation are different because of using different calibration data periods (Pilgrim et al., 1988; Post and Jakeman, 1999; Vaze et al., 2010). Wet and dry years would generate different results. Yapo et al. (1996) suggest that the identifiability of model parameters can be strongly related to hydroclimatic conditions. It appears that parameter identifiability can be significantly improved by selecting the wettest period of the data for calibration (Yapo et al., 1996). Gan et al. (1997) summarized that wet years provide better calibration data than dry years because the former contains more information (especially in terms of peak flows) than the later. So, it is necessary to know the possible performance and its tendency by using limited and fragmented data for calibrating models in ungauged and poorly gauged catchments.

The main objective of this study is to demonstrate the importance of calibration data (e.g., different calibration variables) to estimates of optimal parameters and uncertainty performance of the conceptual rainfall-runoff models. It appears that this work can help improve application of CRR in poorly gauged and ungauged catchments. We choose to use non-continuous calibration periods (actually, data are non-continuous and fragmented in some catchments) to have more independent runoff data for model calibration. Average, dry and wet calibration periods are used for study on impact of calibration data variability.

2. DESCRIPTION OF HYDROLOGICAL MODELS AND DATA

2.1. The SIMHYD model

SIMHYD is a lumped conceptual daily rainfall-runoff model. It is driven by daily rainfall and potential evapotranspiration (PET), and simulates daily streamflow. It has been tested and used extensively across Australia (Chiew et al., 2002; Zhang et al., 2008). The version of the SIMHYD model used here has nine parameters. In SIMHYD, daily rainfall first fills the interception store, which is emptied each day by evaporation. The excess rainfall is then subjected to an infiltration function that determines the infiltration capacity. The excess rainfall that exceeds the infiltration capacity becomes infiltration excess runoff.

2.2. Study of catchments and data

Daily streamflow data from 55 unimpaired catchments all over Australia are used in this study. Unimpaired streamflow is defined as streamflow that is not subject to regulation or diversion. The data is a subset of the Australian dataset collated for an Australian Land and Water Resources Audit project (Peel et al., 2000).

Monthly rainfall was estimated from gridded daily rainfall (Peel et al., 2000). The spatial resolution of the gridded daily rainfall is 5 km by 5 km based on interpolation of over 6000 rainfall stations in Australia. The interpolation uses monthly rainfall data, ordinary Kriging with zero nugget and a variable range. Monthly rainfall for each point is converted to daily rainfall using the daily rainfall distribution from the station closest to that point. Catchment-averaged rainfall was estimated from the daily rainfall (Peel et al., 2000). Mean daily potential evapotranspiration was calculated based on the Priestley-Taylor equation (Priestley and Taylor, 1972). Details of the calculation can be found in Raupach et al. (2001).

The catchments parameters were calculated using the above data. The catchments range in area from 51 to 1891 km^2, mean annual rainfall from 587 to 2886 mm, and mean annual runoff from 44 to 2095 mm. The vegetation in the catchments includes native woodlands, open forests, native and managed grass, and agricultural crops. The catchments cover soil types from sand through loams to clays with large differences in soil properties such as saturated hydraulic conductivity and water holding capacity. The runoff coefficient of the catchments ranges from 0.08 to 0.89, and the index of dryness ranges from 0.37 to 2.19, representing diverse hydrological and climatic conditions.
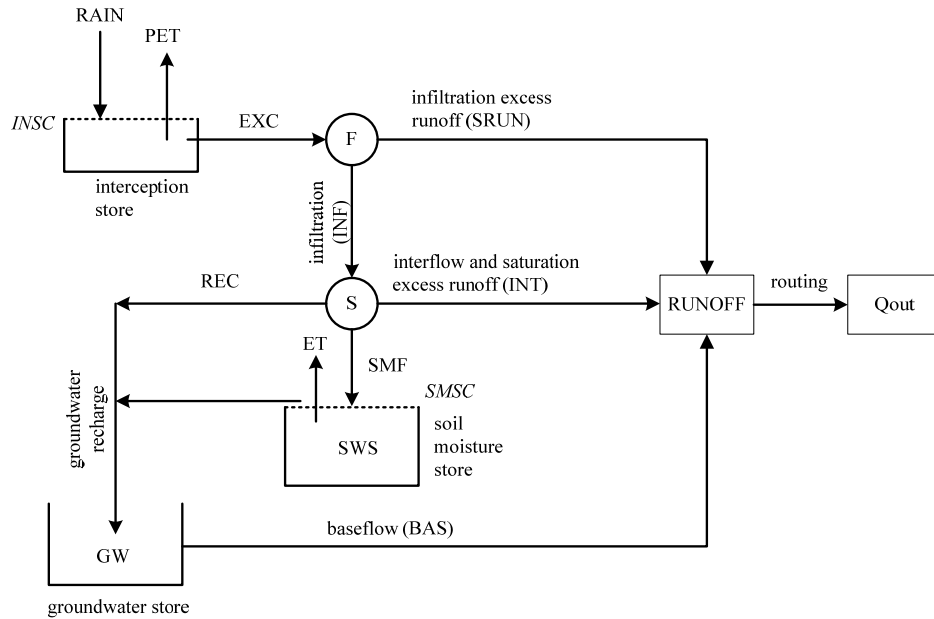
PET = areal potential evapotranspiration (input data)
EXC = RAIN - *INSC*, EXC > 0
INF = lesser of { COEFF exp (-*SQ* × SMS/S*MSC*) , EXC }
SRUN = EXC - INF
INT = *SUB* × SMS/*SMSC* × INF
REC = CRAK × SMS/SMSC × (INF - INT)
SMF = INF - INT - REC
ET = lesser of { 10 × SMS/*SMSC* , PET }
BAS = *K* × GW

Model parameters and description
INSC   interception store capacity (mm)
COEFF maximum infiltration loss (mm)
SQ   infiltration loss exponent
SMSC soil moisture store capacity (mm)
SUB   constant of proportionality in interflow equation
CRAK constant of proportionality in groundwater recharge equation
K   baseflow linear recession parameter
XE   parameter of the Muskingum routing method
KE   parameter of the Muskingum routing method

Figure 1. Structure of the lumped rainfall-runoff model SIMHYD.

3.   METHODOLOGY

3.1.   Experimental design

In order to investigate the impacts of data variability on model calibration, three different 15%-year subsets with different climatic conditions are compiled as follows. Total annual precipitation is calculated for the observation period. Then, three subsets are chosen from this period, representing relatively dry, average, and wet years. Each subset of years contains 15% of the whole observation period (Figure 2). Dry years are dark downward diagonal bars, average years are trellis bars, and wet years are solid bars. And then, the three sub-periods are used to calibrate hydrological model. After model calibration, verification tests are made on the whole records of each catchment.
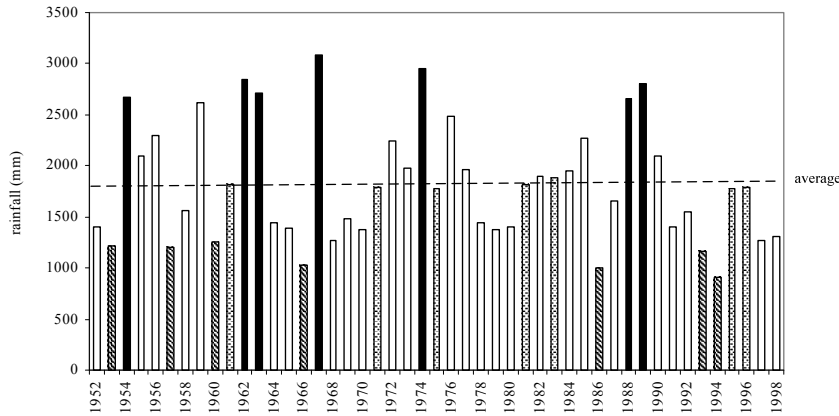
Figure 2. Three subsuts are split from the observation period according to the total annual precipitation. Dry (dark downward diagonal bars), average (trellis bars), and wet (solid bars) years.

3.2. Assessment of model performance criteria

The popular Nash and Sutcliffe efficiency (hereinafter NSE, Nash and Sutcliffe, 1970) and Water Balance Error percentage (hereinafter WBE, Hogue et al., 2006) have been used as objective functions.

The Nash Sutcliffe efficiency (NSE) is defined as:

$$\text{NSE} = 1 - \frac{\sum_{i=1}^{N}(Q_{obs,i} - Q_{sim,i})^2}{\sum_{i=1}^{N}(Q_{obs,i} - \overline{Q_{obs,i}})^2}$$

Where $Q_{sim,i}$ and $Q_{obs,i}$ are the simulated daily runoff and observed daily runoff, respectively, $\overline{Q_{obs,i}}$ is the arithmetic mean of the observed runoff, i is the ith day and N is the total days sampled.

The absolute WBE is a measure of the bias in the calibration results from the observed flow, which is expressed as:

$$\text{absolute } WBE = \frac{\sum_{i=1}^{N}\left|Q_{sim,i} - Q_{obs,i}\right|}{\sum_{i=1}^{N}Q_{obs,i}} \times 100\%$$

NSE and absolute WBE are also used to measure the model performance for each validation test. A good model performance should have both a high NSE value and a low absolute WBE value.

3.3. Optimization algorithms

The Particle Swarm Optimisation (PSO) was used to optimize the parameters of the rainfall-runoff model. PSO is a population-based stochastic optimization technique developed by Eberhart and Kennedy in 1995, inspired by the social behavior of bird flocking or fish schooling (Eberhart and Kennedy, 1995). PSO shares many similarities with evolutionary computation techniques such as Genetic Algorithms (GA). It originates from the swarm paradigm, called particle swarm, and is expected to provide the so-called global or near-global optimum. The PSO method has been successfully used in several rainfall-runoff model parameter optimizations (Chau, 2006; Gill, et al., 2006).

4.   RESULTS AND DISCUSSION

4.1.   Impacts on model calibration

Calibration values of the SIMHYD model performances are summarized as the distributions of NSE and absolute WBE values obtained for the 55 catchments in Figure 3. The NSE values for average and wet calibration periods are better than those for dry calibration period. The 25th percentile, median, and 75th percentile NSE values calibrated for the average periods are generally greater than those of the dry periods. However, the average and wet calibration periods yield similar results. The standard deviation of the NSE for the average, dry and wet calibration periods are 0.14, 0.21 and 0.12 respectively. Average and wet calibration periods have lower standard deviations indicating that the NSE values of 55 catchments tend to be close to the mean values, while dry calibration periods has higher standard deviation indicating that the NSE values of 55 catchments are more spread out. The results suggest that SIMHYD has more steady performance when calibrated by average or wet periods than calibrated by dry periods.
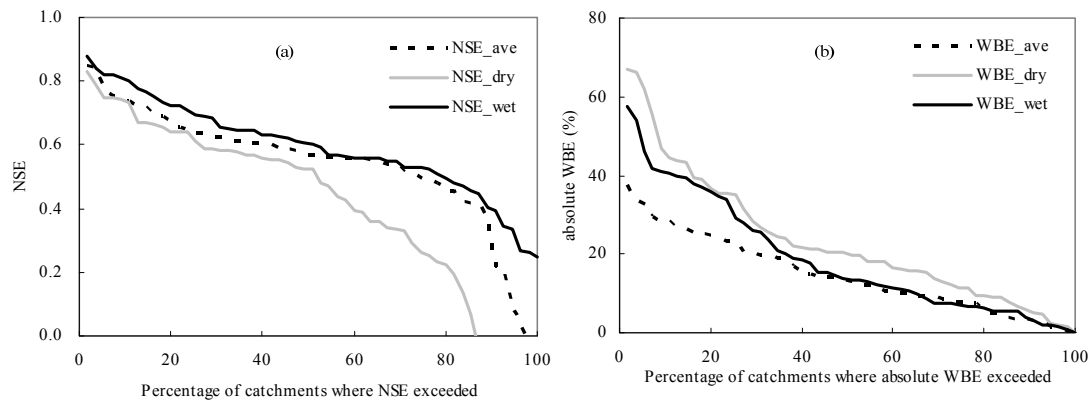
For absolute WBE results (see Figure 3), unlike NSE results, the 75th and averaged absolute WBE values of dry calibration period are about 1.60% and 0.73% lower than those of average calibration periods and about 1.77% and 2.13% lower than those of wet calibration periods. The median absolute WBE values of the dry calibration period are about 1.74% and 0.62% higher than those of average and wet calibration period respectively. The 25th percentile absolute WBE values of the dry calibration period are about 1.10% higher than those of average calibration period, and about 2.52% lower than those of wet calibration period. The 25th percentile, median, 75th percentile and averaged absolute WBE values of average calibration period are about 0.17% to 3.62% lower than those of wet calibration period. The absolute WBE standard deviation values of average, dry and wet calibration periods are 9.82%, 9.26% and 12.45% respectively. Average and dry calibration periods have closer and lower standard deviations indicating that SIMHYD model has more steady performance when calibrated by average or dry periods. It is obviously that larger errors are associated with higher flows (Sorooshian et al., 1983).
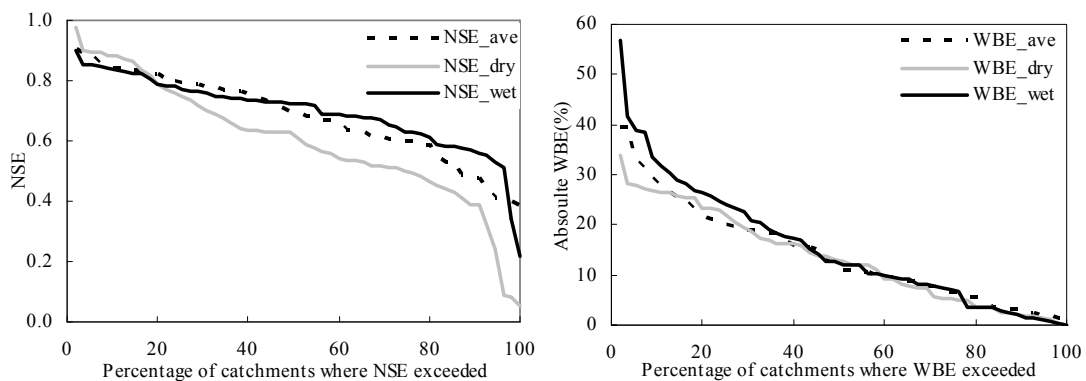
Figure 3. Summary of calibrated Nash-Sutcliffe Efficiency (NSE) and absolute Water Balance Error percentage (WBE) values under average, dry and wet calibration periods over 55 catchments.

4.2.   Impacts on model verification performance

Figure 4 illustrates the model efficiency over 55 catchments in the verification stage in terms of NSE and absolute WBE values, respectively. For the wet calibration period, which has the best verification results, the 25th percentile, median and averaged of the NSE values from 55 catchments are generally about 0.04 and 0.09 higher than those of the average and dry periods. The difference between the three calibration periods is more significant in the poorer modelled catchments. The 75th percentile NSE values (lower quartile NSE value showing results from the poorer modelled catchment) are 0.04 and 0.27 higher than those of the average and dry periods. The NSE values of the average and wet calibration periods showed similar results. The NSE values of the average period are also higher than those of the dry period. Absolute WBE values calibrated by average period are lower by 8.38% and 4.04% than those of dry and wet periods for averaged results. Absolute WBE values of wet period are also lower 4.34% than dry periods.

Li *et al.*, Impact of calibration data variability on rainfall-runoff modelling performance in data-limited basins

Figure 4. Percentage of catchments with Nash-Sutcliffe Efficiency (NSE) greater than or equal to a given NSE value (a) and percentage of catchments where absolute Water Balance Error percentage (WBE) values exceeded.

5.   SUMMARY AND CONCLUSIONS

The primary goal of this study is to demonstrate the importance of calibration data (e.g., different calibration variables) to estimates of uncertainty performance of the conceptual rainfall-runoff SIMHYD model. In order to study the performance of rainfall-runoff model in ungauged or data-limited catchments, we choose to use non-continuous calibration periods (actually, data are non-continuous and fragmented in some catchments) to have more independent runoff data for model calibration. Average, dry and wet calibration periods are used for study on impact of calibration data variability. Fifty five (55) catchments are tested to obtain general conclusions.

Our results may have useful and interesting implications when one studies at the case of ungauged or data-limited catchments. These results give some indication of how many years of data should be acquired in order to calibrate the model. This may yield better results rather than just using parameter values estimated from regionalization studies.

As shown in our study, the SIMHYD model generally proved more effective for average and wet years, even model has been extensively applied to some catchment, but it may not work well under dry climatic conditions. If only a short series of observed data is available to the catchment, it would be a gingerly first step to ascertain the wetness level of the data by comparing them with time series of neighbouring catchments (runoff or rainfall data). If the data at hand were found to be during dry years, it is expected that such calibration would not perform well during dry years in dry catchment.

This is a wide area of research, and we should do more. There are many factors impacting the performance and parameter values on rainfall-runoff models. However, we can reduce the uncertainty in runoff prediction in ungauged or data-limited catchments by knowing more tendencies on the performance of rainfall-runoff model under different calibration data conditions.

REFERENCES

Boughton, W.C. (2007). Effect of data length on rainfallerunoff modelling, *Environmental Modelling & Software*, 22, 406-413.
Chau, K. W. (2006). Particle swarm optimization training algorithm for ANNs in stage prediction of Shing Mun River, *Journal of Hydrology*, 329, 363-367.

Li *et al.*, Impact of calibration data variability on rainfall-runoff modelling performance in data-limited basins

Chiew, F. H. S., Peel, M. C., and Western A. W. (2002). Application and testing of the simple rainfall-runoff model SIMHYD, in Mathematical Models of Small Watershed Hydrology and Applications, edited by V. P. Singh and D. K. Frevert, pp. 335-367, Water resources Publication, Littleton, Colorado, USA.

Eberhart, R. C., and Kennedy J. (1995). A new optimizer using particle swarm theory, in the Sixth International Symposium on Micro Machine and Human Science. MHS edited, pp.39-43,doi:10.1109/MHS.1995.494215, IEEE Press, Piscataway, N. J.

Gan, T.Y., Dlamini, E. M., Biftu, G. F. (1997). Effects of model complexity and structure, data quality, and objective functions on hydrologic modelling, *Journal of Hydrology*, 192, 81-103.

Gill, M. K., Y. H. Kaheil, A. Khalil, M. McKee, and L. Bastidas (2006). Multiobjective particle swarm optimization for parameter estimation in hydrology, Water Resources Research, 42, W07417, doi:07410.01029/02005WR004528.

Hogue S. T., Gupta H., Sorooshian S. (2006). A 'user-friendly' approach to parameter estimation in hydrologic models, *Journal of Hydrology*, 320, 202-217.

Nash J. E., J. V. Sutcliffe. (1970). River forcasting using conceptual models, 1. A discussion of principles. *Journal of Hydrology*, 10, 280-290.

Peel, M.C., Chiew, F.H.S., Western, A.W., McMahon, T.A. (2000). Extension of unimpaired monthly stream flow data and regionalization of parameter values to estimate stream flow in ungauged catchments. Report to National Land and Water Resources Audit, Cent. For Environ. Appl. Hydrol., Univ. of Melbourne, Parkville, Vict., Australia.

Pilgrim, D.H., Chapman, T.G. and Doran, D.G. (1988). Problems of rainfall-runoff modelling in arid and semiarid regions, *Hydrological Sciences Journal*, 33(4), 379-400.

Post, D.A. and Jakeman, A.J. (1999). Predicting the daily streamflow of ungauged catchments in S.E. Australia by regionalising the parameters of a lumped conceptual rainfall-runoff model, *Ecological Modelling*, 123(2-3), 91-104.

Priestley, C.H.B., Taylor, R.J. (1972). On the assessment of the surface heat flux and evaporation using large-scale parameters. *Monthly Weather Review*, 100, 81-92.

Raupach M. R., J. M. Kirby, D. J. Barrett, P. R. Briggs, H. Lu and L. Zhang (2001). Balances of water, carbon, nitrogen and phosphorus in Australian landscapes: 2. Model formulation and testing, Tech. Rep. 41/01, CSRIO Land and Water, Canberra, ACT, Australia.

Sorooshian, S., Gupta, V.K., Fulton, J.L. (1983). Evaluation of maximum likelihood parameter estimation techniques for conceptual rainfall-runoff models: Influence of calibration data variability and length on model credibility, *Water Resources Research*, 19(1), 251-259

Vaze, J., Chiew, F.H.S., Perraud, J-M., Viney, N., Post, D., Teng, J., Wang, B., Lerat, J. and Goswami, M. (2010). Rainfall-runoff Modelling across Southeast Australia: Datasets, Models and Results. *Australian Journal of Water Resources*, 14(2), 101-116.

Yapo, P. O., Gupta, H. V. & Sorooshian, S. (1996). Automatic calibration of conceptual rainfall-runoff models: sensitivity to calibration data. *Journal of Hydrology*, 181, 23-48.

Zhang, Y. Q., F. H. S. Chiew, L. Zhang, R. Leuning, and H. A. Cleugh (2008). Estimating catchment evaporation and runoff using MODIS leaf area index and the Penman-Monteith Equation, *Water Resources Research*, doi:10.1029/2007WR006563.