Journal of Hydrology xxx (xxxx) xxxx

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Research papers

Automatic calibration of a whole-of-basin water accounting model using a comprehensive learning particle swarm optimiser

Lei Gao[a,*], Mac Kirby[b], Mobin-ud-Din Ahmad[b], Mohammed Mainuddin[b], Brett A. Bryan[a,c]

[a] CSIRO Land and Water, Private Mail Bag 2, Glen Osmond, SA 5064, Australia
[b] CSIRO Land and Water, GPO Box 1666, Canberra, ACT 2601, Australia
[c] Centre for Integrative Ecology, Deakin University, Burwood, VIC 3125, Australia

ARTICLE INFO

This manuscript was handled by G. Syme, Editor-in-Chief, with the assistance of Ashok Mishra, Associate Editor

Keywords:
Model calibration
Parameterisation
Particle swarm optimisation
Hydrological models
Rivers
Irrigation

ABSTRACT

We present a two-step framework for calibrating complex, many-parameter hydrological models at basin-scale. The framework first calibrates parameters for each catchment/sub-basin sequentially and then fine-tunes parameters as needed. We implemented a comprehensive learning particle swarm optimiser (CLPSO) as the calibrator and applied the two-step CLPSO tool in calibrating parameters of a water accounting model for the Murray-Darling Basin, Australia. The visual and quantitative results indicated that our tool produced satisfactory calibration and prediction outcomes for the model's intended purpose. The comparison experiments demonstrated that the calibration framework and the CLPSO were competent in calibrating large-scale hydrological models. This framework can guarantee spatial coherence, balance objective trade-offs among all catchments, and calibrate many parameters at a low computational cost. By providing better parameter estimates in complex whole-of-basin hydrological models, our calibration tool has the potential to increase the development and application of these models, and thereby improve the management of large river basins.

1. Introduction

Hydro-economic models have been increasingly used in capturing the complexity of interactions between water and economy, and helping water resource managers address water scarcity and reduce water demand conflicts (Graveline, 2016; Harou et al., 2009; Mainuddin et al., 2007; Qureshi et al., 2007). These hydro-economic models often integrate management options and economic values into a detailed hydrological component, which represents appropriate complexity and heterogeneity of a water resource system. The hydrological component, in some cases, needs to reflect large, basin-scale hydrological dynamics that is affected by policy changes. However, most hydrological models, especially conceptual rainfall-runoff models are calibrated at a catchment scale, capitalising on the reduced hydrologic, physiographic, and socio-economic complexity at this scale. In contrast, large-scale whole-of-basin hydrological models are characterised by expanded model structures covering heterogeneous environmental and socio-economic conditions, complex reach-connection relationships, and require a substantial number of model parameters to be estimated.

In the relatively few studies that have addressed hydrological model calibration across multiple sites/catchments/sub-basins, either a single parameter-set is employed for all catchments (e.g., Engeland et al., 2006; Khakbaz et al., 2012; Wang et al., 2012), or multiple catchments/sites are treated individually from upstream to downstream (e.g., Ajami et al., 2004; Hughes et al., 2014). However, since hydrologic systems often exhibit a large degree of spatial heterogeneity in their characteristics (Grayson and Bloeschl, 2000), approaches which regard model parameters as identical for all catchments are inadequate. Similarly, multi-site calibration experiments (e.g., Andersen et al., 2001) have demonstrated that the calibration performance in downstream catchments may be affected by the performance upstream. This suggests that uncertainties and/or compensation of errors will propagate downstream and affect the overall calibration of the model. Thus, the calibration of individual sites ignoring hydrological connectivity, while an improvement on single-site approaches, is also insufficient. Meanwhile, this calibration fashion showed up low efficiency.

To address this, some researchers have proposed a system-wide approach, enabling a simultaneous calibration for the systematic evaluation of the trade-offs among objective functions from all catchments (e.g., Shrestha and Rode, 2008). But this approach raises two challenges. One is that when applied at a whole-of-basin scale, a large number of model parameters for related catchments may need calibrating simultaneously. For example, a hydrological model for the Murray-Darling Basin (MDB) requires calibrating model parameters for

* Corresponding author at: CSIRO Land and Water, Private Mail Bag 2, Waite Road, Glen Osmond, SA 5064, Australia.
E-mail address: lei.gao@csiro.au (L. Gao).

https://doi.org/10.1016/j.jhydrol.2019.124281
Received 23 March 2019; Received in revised form 15 October 2019; Accepted 23 October 2019
Available online 24 October 2019
0022-1694/

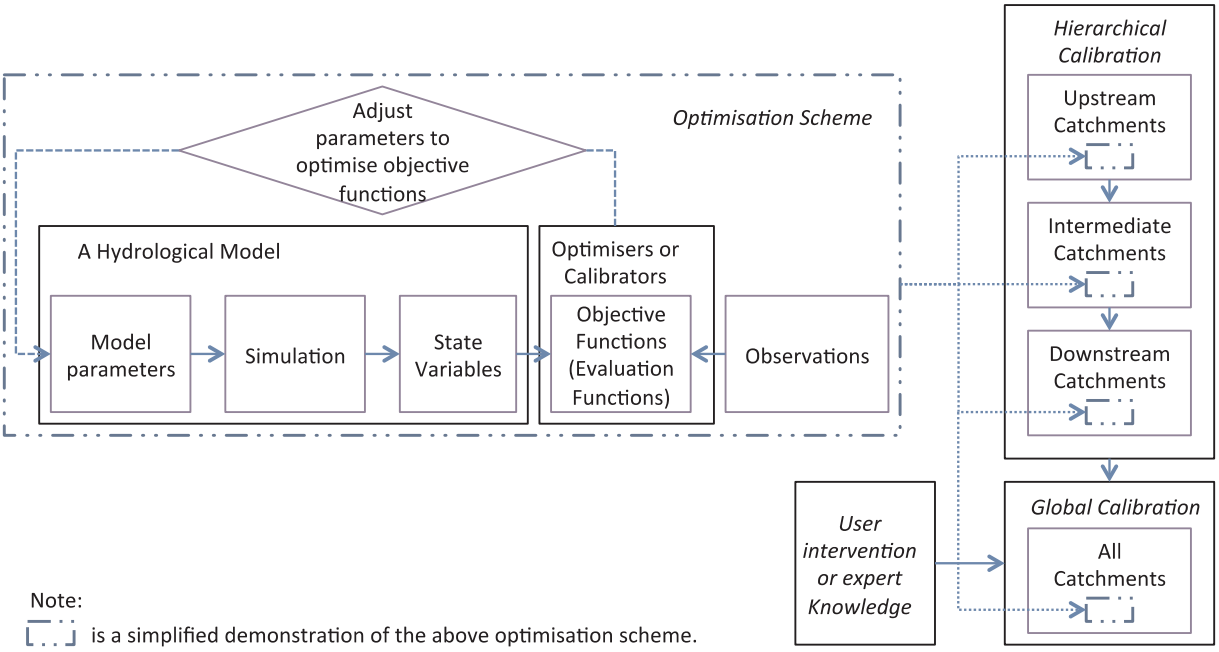

Fig. 1. The proposed two-step calibration framework.

58 connected catchments (CSIRO, 2008a). The second challenge is the trade-offs among objective functions for all catchments. An objective function provides a measurement of similarity between observed and estimated outputs (Molina-Navarro et al., 2017). The similarity degree (calibration space), data availability, and uncertainty levels for catchments differ; therefore, a normalisation procedure is required before balancing the objective functions from all catchments. However, finding a suitable weighting function used in the system normalisation procedure is difficult (Hughes et al., 2015).

To overcome these challenges, this paper adds new dimension to large-scale hydrologic model calibration by enhancing individual catchment-level/sub-basin-level step-wise parameter calibration (introduced by Rajib et al. (2018)) via a novel two-step hierarchical-global approach. The hierarchical calibration is responsible for guaranteeing catchment-level spatial coherence, allowing model calibration at a whole-of-basin scale with low computational cost. The global calibration is used for fine tuning of parameters based on model predictive needs or expert knowledge.

There has been a historical trend away from manual calibration methods, towards more automatic procedures. Two calibration strategies are commonly used to automatically explore the parameter space: local and global methods (Kavetski et al., 2018). Robustness (capability of consistently locating the optimal parameter set) and computational cost (often represented by the number of objective function evaluations or CPU time required to find the optimal parameter set) are two key performance indicators for evaluating these automatic calibrators (Kavetski et al., 2018). Local calibration methods, such as Gauss-Newton-type methods (Qin et al., 2018b), usually converge fast but tend to get trapped in a local optimum when solving multimodal problems. Popular global methods to automatically calibrate hydrological models are evolutionary algorithms (EAs) (Maier et al., 2014), such as genetic algorithms (GA) (Seibert, 2000), shuffled complex evolution (SCE) (Duan et al., 1992), and simulated annealing (SA) (Sumner et al., 1997). A large number of studies have concluded that EAs perform better than local methods in finding the global optimum (e.g., Gao et al., 2006; Madsen et al., 2002), but tend to be computationally expensive (Qin et al., 2018a).

A relatively new EA, particle swarm optimisation (PSO), was first proposed by Kennedy and Eberhart (1995). It is an algorithm inspired by the social behaviour of animals, such as bird flocking and fish schooling. PSO has gained popularity lately and has been applied widely, including to the automatic calibration of hydrological models (Jiang et al., 2013; Thiemig et al., 2013; Zambrano-Bigiarini and Rojas, 2013). Compared to conventional EAs generally, which have strong global-search ability, but low convergence speed and computational efficiency (Zhang et al., 2009), PSO is attractive due to its simplicity of implementation and its ability to quickly converge to a reasonably good solution. However, similar to GA, the standard PSO may get trapped in local optima, especially when solving complex multimodal problems (Liang et al., 2006). Here, a comprehensive learning particle swarm optimiser (CLPSO) (Gao and Hailu, 2010; Liang et al., 2006) is used to improve the standard PSO's performance on complex multimodal problems, aiming to outperform conventional EAs in both robustness and computational cost.

In this study, we present a transparent and automated, hierarchical-global CLPSO calibration tool for basin-scale hydrological model calibration. We applied our calibration tool to a water accounting model for the MDB, Australia's largest river system (Kirby et al., 2013). The model was designed for and has been used to supply scenario information for a range of economic analyses (e.g., Kirby et al., 2014; Kirby et al., 2015). The hydrological model discussed in this paper provides monthly estimates of runoff, river flow, diversions and losses for the MDB. Compared with daily water accounting models, monthly models have the advantage of needing only monthly data and have lower computational cost (Wang et al., 2011), so that iterative solutions (EA-based calibration, for example) could be obtained. We evaluate the performance of the calibration tool and discuss its utility for broad application to whole-of-basin models of large, complex river systems.

2. Methods

2.1. Two-step automatic calibration framework

To enable automatic calibration of model parameters at a whole-of-basin scale, we proposed a calibration framework (Fig. 1). The framework can be applied in both optimisation-based and probabilistic/Bayesian calibration (to determine a probability distribution over parameters) (e.g., Wöhling and Vrugt, 2011). In this paper, the calibration procedure means parameter "tuning", i.e., finding the best parameter values by optimising model outputs to observed data.

Table 1
Pseudo-code for hierarchical calibration with n threads.

Algorithm 1: Executing hierarchical calibration with n threads

```
Initialise all nodes#1 in a hierarchical tree structure and their father nodes#2
For each node o /* identify the virtual node of top upstream nodes */
  If (o =O0) /* o is the virtual node */
    co=True /* co represents whether o has been calibrated */
  Else
    co=False
  End If
End For
Repeat
  For each node o with co= False
    If each father so of node o is calibrated (i.e., cso= True) And the
calibration queue q is not full (the size of q is n)
      Add o into q
    End If
  End For
  For each node oi in q /* i ≤ n */
    Assign oi to threadi
    Calibrateoi
    coi=True
  End For
Until all nodes are calibrated
```

#1 This algorithm regards the structure of the whole basin as a complex tree structure in computer sciences. The nodes (i.e., catchments in the basin) in the hierarchical tree store the information (a symbol represents whether a node has been calibrated) of their father nodes (a father node is a direct upstream catchment of a catchment/node and a catchment may have multiple father nodes).

#2 The father nodes of top upstream nodes are set as a virtual one O_0.

The framework includes a hierarchical-global calibration strategy. The hierarchical calibration aims to maintain basin spatial coherence by calibrating catchments/sub-basins from upstream to downstream in the order. Many catchments in the hierarchical calibration can be calibrated in parallel (the parallel hierarchical calibration algorithm is presented in Section 2.2). In this procedure, once upstream catchments are calibrated, their parameters are fixed and discharges at their outlets are used to calibrate downstream catchments. In other words, the calibration of a catchment requires inflows from all upstream catchments and only adjusts parameters at the current catchment by using the fixed and calibrated parameters for all the upstream catchments. This procedure implicitly assigns higher priorities to upstream catchments. The purpose of this calibration procedure is twofold: to avoid optimising a large number of model parameters simultaneously, and to provide a benchmark of all catchments for the global calibration. Global calibration involves a fine-tuning of model parameters that have been optimised in hierarchical calibration procedure. It is typically required when upstream catchments are allocated with higher priorities than their downstream catchments, the optimised parameter-set is not physically meaningful, or additional calibration criteria are not satisfied (for example, model parameters need to be optimised with a certain level of maximum flow). User intervention or expert knowledge is needed during the global calibration process (for example, to assign weights/priorities to different catchments, and/or adjust the searching ranges for parameters).

The optimisation schemes in hierarchical (for upstream, intermediate, and downstream catchments) and global calibration are similar and depicted on the left-hand side of Fig. 1. The implementation of such an optimisation scheme includes an objective function (or an evaluation function), a simulation model, an optimiser/calibrator to search the parameter space, and a period of historical data against which to calibrate the model iteratively.

Depending upon different calibration purposes, a single objective measure or multi-objective formulation can be used in calibration. Multi-objective calibration methods of hydrological models have been developed (e.g., Engeland et al., 2006; Ercan and Goodall, 2016; Shrestha and Rode, 2008). The intended applications of the water accounting model are related to water allocation in the MDB and high flows are prioritised over low flows. Therefore, this work focuses on calibrating flow duration curves (Booker and Snelder, 2012; Li et al., 2010; Vogel and Fennessey, 1994; Zhang et al., 2015) rather than reproducing the exact hydrographs. The flow duration curve depicts the relationship between magnitude and frequency of flow, showing the percentage of time for which specified flow is equalled or exceeded during a given period. It is a useful visual and analytical tool for evaluating flow variability (Kundu et al., 2016) and widely used for water resource assessment and planning, particularly for water allocation problems (Sadegh et al., 2016; and the references therein; Vogel and Fennessey, 1995). It is also appropriate for the intended purpose of the MDB model, for which exact timings of flows (i.e., exact hydrographs) are of lesser concern. Thus, a single objective function on the calibration of flow duration curves is selected. The objective function represents a numerical measure of the difference between simulated outcomes and observed values. The calibration aims to minimise the objective function.

The objective function for the hierarchical calibration (or likelihood function for Bayesian calibration) is described as the absolute distance between modelled estimates and observations (see Eq. (1)).

$$F_i = \sum_{j=1}^{M_i} |c_{i,j} - o_{i,j}| \tag{1}$$

where F_i is the objective function or likelihood function for catchment i, M_i is the number of monthly observations, $c_{i,j}$ is the jth ranked calculated monthly estimate for catchment i in the hierarchical calibration, and $o_{i,j}$ is the jth ranked monthly observation for catchment i.

The objective function for the global calibration used in this study is given as Eq. (2).

$$F = \sum_{i=1}^{N} [\lambda_i \bullet \frac{\sum_{j=1}^{M_i} |c_{i,j}^{'} - o_{i,j}|}{F_i}] \tag{2}$$

where F is the global objective function or likelihood function, N is the number of catchments, M_i is the number of monthly observations for catchment i, λ_i represents a weight assigned to catchment i, $c_{i,j}^{'}$ is the jth ranked calculated monthly estimate for catchment i in the global calibration, and F_i is the optimised objective function or likelihood function for catchment i in the hierarchical calibration.

2.2. A fast hierarchical calibration algorithm

We also developed an algorithm to perform hierarchical calibration in parallel processing mode (with n computer threads or n processors) with less computational time (Table 1).

2.3. Calibrators

Our framework can accommodate different calibrators or optimisers. Known to be highly robust, SCE is the most widely used calibrator in the practical and research hydrology communities; however, it tends to be computationally costly (Kavetski et al., 2018; Tolson and Shoemaker, 2007). Here, we have implemented the CLPSO and an improved SA (He and Wang, 2007) as calibrators in this framework. Both calibrators are claimed to achieve high robustness at low cost. A semi-automatic calibrator was also implemented only for hierarchical calibration for comparison with the CLPSO and the SA. The semi-automatic calibrator is based on Microsoft Office Excel's Solver and is therefore difficult to extend to global calibration. Next, we briefly overview the CLPSO and the SA method.

Inspired by the social behaviour of animals, PSO has been designed as a population-based optimisation algorithm (Kennedy and Eberhart, 1995). Each solution candidate is regarded as a *particle*, and a population of candidates form a *swarm*. These particles move around in the search space and the position of a particle is a potential solution. The

movements of a particle rely on its previous best position *pbest*, and the swarm's best known position *gbest*. A particle can adjust its velocity to alter the search trajectory based on its own knowledge and other particles' knowledge. In this way, PSO integrates the advantages of local and global search techniques to locate the global optimum by balancing the exploitation and exploration aspects in the search space.

For the sake of PSO's performance on complex multimodal problems, Liang et al. (2006) employed a comprehensive learning strategy to extend the conventional PSO. In contrast with the conventional PSO, the comprehensive learning strategy (1) uses all particles' best known positions *pbest*s in the population as exemplars to determine a particle's direction; (2) allows a particle's dimension to learn from the corresponding dimension of another particle's *pbest*; and (3) forbids a particle to learn from exemplars across all generations, in order to make the particle learn from good exemplars and reduce inefficient search on poor directions. The learning strategy allows the optimiser to more effectively utilise the information in the swarm to generate high quality solutions to most multimodal problems. The CLPSO has been widely extended to solve other types of optimisation problems, such as constrained mixed-variable optimisation problems (Gao and Hailu, 2010) and multi-objective optimisation problems (Huang et al., 2006). The pseudo code for the CLPSO is demonstrated in Supplementary Material Table A1. We refer the readers to (Gao and Hailu, 2010; Liang et al., 2006) for details concerning algorithm design, performance measurement, and applications.

SA is an automatic optimiser that is motivated from an analogy between physical annealing in metallurgy and the strategy of solving optimisation problems. The conventional SA is a single-solution metaheuristic optimiser (Kirkpatrick et al., 1983), whose central idea is to perturb the solution continuously, evaluate the quality of the solution, and use a certain probability to accept the worse solution for a more extensive search. Starting with an initial solution, SA produces a new solution in a predefined neighbourhood and evaluates the new solution in each iteration. The new solution is compared with the best solution so far in terms of the fitness/evaluation function. A better new solution is always accepted, while a worse new solution is accepted with a Boltzmann probability (this is a mechanism to probabilistically avoid getting trapped in local optima). We implemented a SA calibrator described in He and Wang (2007). The pseudo code for the algorithm is presented in Supplementary Material Table A2.

2.4. Evaluating calibration performance

It is necessary to conduct model validation since it can reflect whether the model needs further calibration refinement. To quantify the calibration performance, we adopted three quantitative evaluation statistics as evaluation indicators: Nash-Sutcliffe efficiency (NSE), percent bias (PBIAS), and ratio of the root mean square error to the standard deviation of measured data (RSR), which were recommended by Moriasi et al. (2007) as the most effective model evaluation techniques. Moriasi et al. (2007) also provided model performance ratings for the three statistics based on evaluation experiments on hydrographs. Here, we used their model performance ratings (Supplementary Material Table B1) as rough criteria to evaluate the performance of the two-step CLPSO calibration. To validate the model, we chose monthly time series flows (hydrographs) for graphical evaluation (a short-term observed time series flow data, for instance, 15 years, is used for validation and the remaining time series data for calibration), and flow duration curves based on simulated monthly time series flows for quantitative evaluation.

2.5. Comparison experiments

To demonstrate the benefits of the two-step CLPSO calibration tool, we compared its calibration results with those obtained from a semi-automatic calibration framework (to show the relative benefits of both the two-step framework and the CLPSO) and from a two-step SA calibration approach (to show the relative benefit of the CLPSO).

The semi-automatic calibration was implemented in the early version of the water accounting model (spreadsheet version without VBA calculations). Microsoft Office Excel's Solver was used to conduct the calibration of model parameters, and can be regarded as a semi-automatic optimiser. Similar to Hughes et al. (2014), model parameters within each catchment from upstream to downstream were calibrated by the semi-automatic calibrator in a cascading fashion (simulated outflows from upstream catchments were then used to calibrate the downstream catchments). This functionality has not at the time of writing been developed in the VBA code version, because the Solver cannot call a macro: that is, the Solver cannot call the VBA water accounting model.

To compare the calibration accuracy between the CLPSO and the Solver, as well as between the CLPSO and the SA, a measure was developed as a ratio of objective functions estimated by the CLPSO and another approach/calibrator for a catchment, as shown in Eqs. (3) and (4).

$$ratio\_so_i = \frac{F_i^{CLPSO}}{F_i^{Solver}} \tag{3}$$

$$ratio\_sa_i = \frac{F_i^{CLPSO}}{F_i^{SA}} \tag{4}$$

where $ratio\_so_i$ (or $ratio\_sa_i$) is calibration performance ratios of the CLPSO to the Solver (or the SA) for catchment i, and F_i^{CLPSO}, F_i^{Solver}, F_i^{SA} are optimal objective functions obtained by the CLPSO, the Solver, and the SA calibrator for catchment i, respectively.

Convergence speed is another important performance indicator of an optimiser, which reflects the computational efforts and the searching efficiency of an optimiser. To evaluate convergence rates of calibration experiments using the CLPSO and the SA, we selected a normalised measurement called "difference ratio", as shown in Eq. (5).

$$d_i = \frac{bf_{i,k} - bf_{i,\max\_k}}{bf_{i,\max\_k}} \tag{5}$$

where d_i represents the difference ratio for catchment i, $bf_{i,k}$ is the best objective function value at the kth fitness/objective function evaluation (FFE) for catchment i, and $bf_{i,\max\_k}$ is the best objective function value at the maximum FFE for catchment i. We set the number of FFE as 80,000 in the comparison experiments.

3. Case study — MDB water accounting model

3.1. Study area and data sources

The MDB covers more than one million square kilometres (one-seventh) of mainland Australia. It is the largest river system in Australia and has long played an important role in the nation's agricultural sector. It accounts for approximately 40% of the gross value of the nation's agricultural production and approximately 60% of this country's irrigation water use (CSIRO, 2008b). In this study, the MDB is divided into 58 major catchments (Fig. 2), which are aggregates of the rainfall-runoff sub-catchments identified in CSIRO (2008a). The connection relationship of the 58 catchments is shown in Fig. 3, where some rivers in the basin are disconnected, or are distributaries that end in wetlands. For these rivers, the outflow does not become the inflow to another river reach, but ends as evapotranspiration. A discharge is split into downstream catchments based on actual historical pattern (and channel capacity constrain).

The climate dataset was extended from CSIRO (2008a) to 2009 by Vaze et al. (2011). The land use data were sourced from Bryan et al. (2009a,b, 2004) which also included water-use calculations for major agricultural crops. A flow database for the case study area was

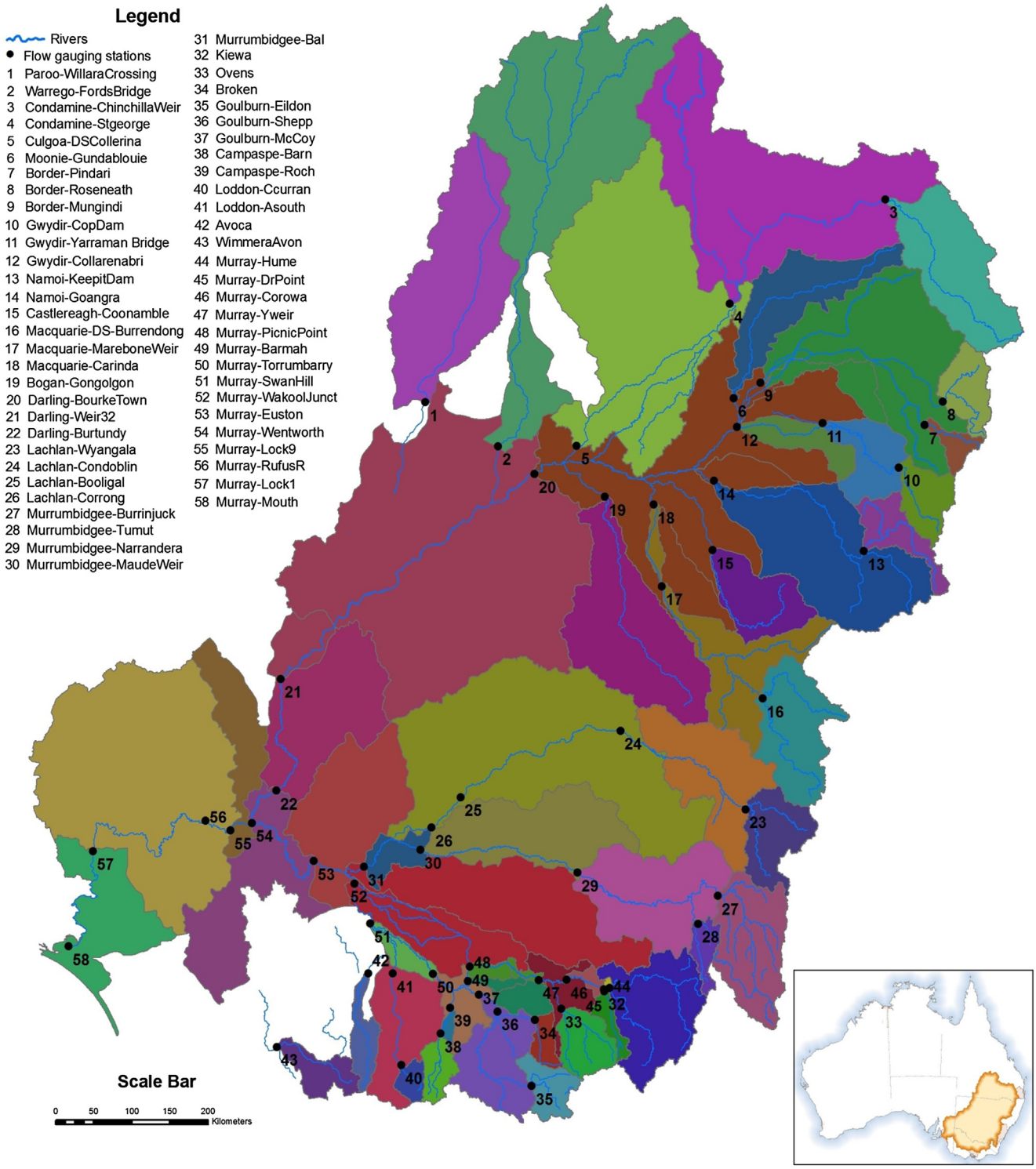

Fig. 2. The MDB and the identified 58 major catchments.

established by integrating data sourced from websites of NSW Government WaterInfo (http://waterinfo.nsw.gov.au/), Queensland Department of Natural Resources and Mines (https://www.dnrm.qld.gov.au/), Victoria's Department of Environment, Land, Water and Planning (http://www2.delwp.vic.gov.au), and South Australia's Department of Environment, Water, and Natural Resources (http://www.environment.sa.gov.au/). Annual data on irrigation were sourced from the MDBA report (MDBA, 2011).

3.2. Model description

The MDB water accounting model includes three principal sub-models. A rainfall-runoff model partitions monthly rainfall into runoff and evapotranspiration using a Budyko model with monthly storage (2008). A river flow and storage model implements a weak form of routing by allowing for temporary hold up of water from one month to the next. Irrigation demand and supply is calculated using a crop coefficient irrigation model. We applied a mass balance model to each

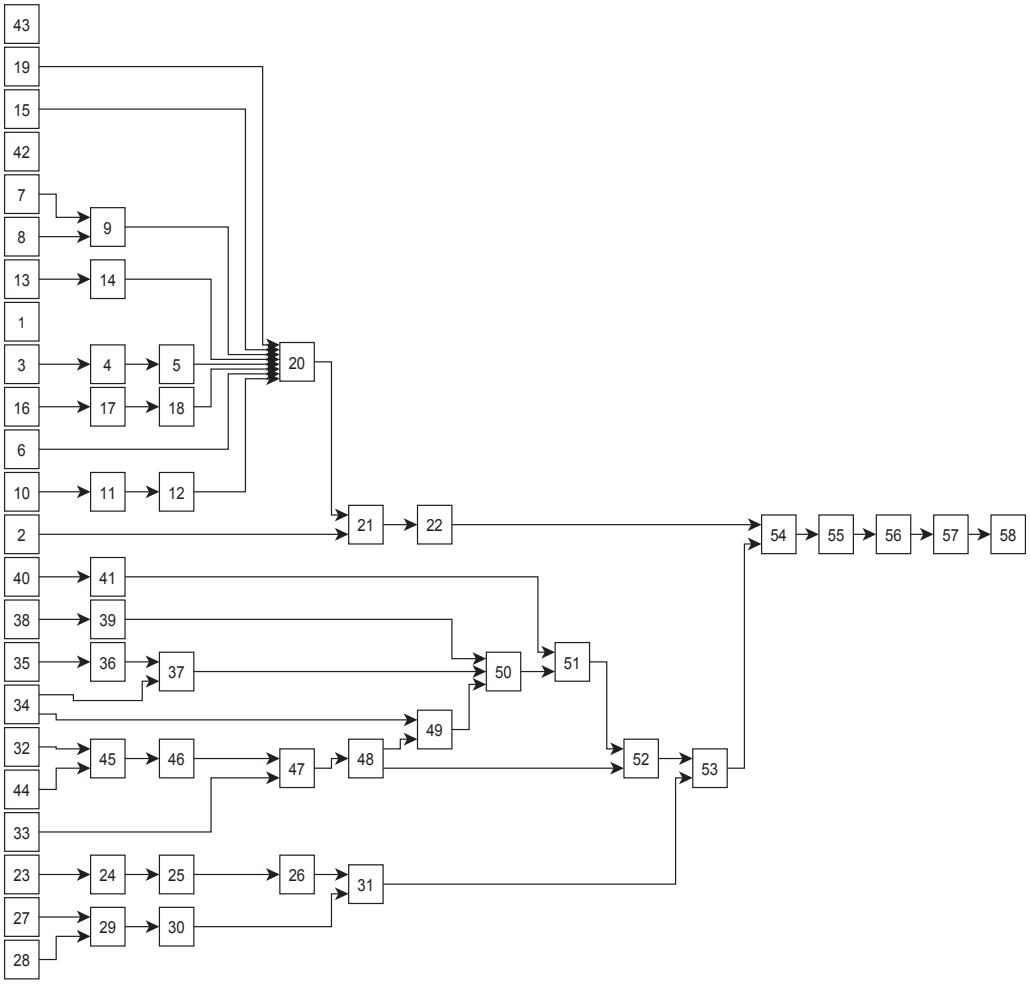

Fig. 3. The connection relationship of the 58 MDB catchments.

catchment (Supplementary Material Fig. C1), although some elements are missing in some catchments (some lack irrigation, for instance). We did not consider the contribution of groundwater to water balance.

A basic mass balance relationship applies to each key entity (such as a catchment, a river, a dam, or a whole basin) in the MDB water accounting model:

$$\sum inflows - \sum outflows - \sum \Delta storages = 0 \tag{6}$$

In this way, the rainfall-runoff model can be written as,

$$P - I - R_o = 0 \tag{7}$$

where P is rainfall at the land surface and partitioned into runoff R_o and infiltration I.

In Eq. (7), P is the supply limit, the relationship between P and the unfilled space of a generalized surface storage can be built using a Budyko-like equation (Budyko, 1974):

$$\frac{I}{\Delta S_{smax}} = \left(\frac{(P/\Delta S_{smax})^{a_1}}{1 + (P/\Delta S_{smax})^{a_1}} \right)^{1/a_1} \tag{8}$$

where ΔS_{smax} is the capacity limit (see Kirby et al. (2013) for more details), a_1 determines the sharpness of the curve in the relationship between infiltration and the incident rainfall, and thus, how much runoff is generated by rainfall. Larger values of a_1 mean more infiltration and lower rainfall.

The evapotranspiration can be modelled using a similar form as Eq. (8):

$$\frac{ET}{ET_{pot}} = \left(\frac{(S_s^{t-\Delta t}/ET_{pot})^{a_2}}{1 + (S_s^{t-\Delta t}/ET_{pot})^{a_2}} \right)^{1/a_2} \tag{9}$$

where ET_{pot} is the potential evapotranspiration, S_s is the surface storage, t is time, Δt is the time step (one month), and a_2 is a parameter to smooth the transition from S_s to ET_{pot}.

Water in the generalized surface store can be increased by infiltration, but reduced by evapotranspiration, as shown in Eq. (10):

$$S_s^t = S_s^{t-\Delta t} + I - ET \tag{10}$$

River flows can be modelled as a reach water balance:

$$Q_o = Q_i + Q_t + R_o - D - L - \Delta S_r \tag{11}$$

where Q_o, Q_i, and Q_t represent the outflow, the inflow, and tributary flows of a reach, respectively; R_o is the runoff from the adjacent catchment; D is diversions, L represents flow losses, and ΔS_r is the difference between reach storage at two time steps ($\Delta S_r = S_r^t - S_r^{t-\Delta t}$).

L is calculated as Eq. (12):

$$L = Q_i(1 - c_{loss}) \tag{12}$$

where c_{loss} ($0 \le c_{loss} \le 1$) is a parameter.

The reach storage can also be modelled as a function of Q_i, which is specified in Eq. (13).

$$S_r = c_1 \bullet Q_i \tag{13}$$

where c_1 is a parameter.

We refer readers to Kirby et al. (2013) for the sub-model of the irrigation demand and supply, as well as further details on modelling

Table 2
Model parameters to be optimised.

Parameter	Range	Description
S_{max}	0.25–1.5	These three parameters are influential only in the upstream catchments. They have very little influence in the lower parts of the Murray, the Darling, and some other larger rivers.
a_1	0.25–5	
a_2	0.25–5	
c_1	0–1	This parameter is influential lower down in the system, and will generally be 0 in upstream catchments.
c_{loss}	0–1	This parameter is influential lower down in the system, and will generally be 1 in upstream catchments.

methods and application of the water accounting model in scenario simulation. The model was developed in a Microsoft Excel spreadsheet, where the bulk of the calculations were done in a Visual Basic for Applications (VBA) macro. The spreadsheet contains all the input data needed to define the model; the VBA macro calculates runoff, flows and diversions; and the results are written to output worksheets. Graphing and subsidiary calculations may be performed as required in the output worksheets. The model parameters of an earlier spreadsheet only version (without the VBA macro) were calibrated using Solver in Excel, which was used to minimise the absolute distance between modelled estimates and observations in Eq. (1) by varying model parameter values. This cannot be done in the VBA version of the model, because Solver cannot call a macro.

3.3. Model parameters

The model parameters to be optimised (Table 2) were identified by the water accounting model developers. For complex models in which developers lack confidence in identifying influential parameters, a global sensitivity analysis (Gao and Bryan, 2016; Gao et al., 2016; Saltelli et al., 2008) can provide insights about the mapping of model inputs to model outputs. The runoff into any reach must equal the sum of the outflow, losses, diversions, and changes to storage minus the sum of the inflows. This holds true for any period, from a single month to the full length of the record under consideration. By varying S_{max}, and parameters a_1 and a_2 in Eqs. (8) and (9), the sum of the runoff over the full period is set to approximately equal the sum over the full period of the outflows and changes to storage less the sum of the inflows. Larger values of S_{max}, and smaller values of a_1 and a_2 lead to more non-linear behaviour in which runoff ratios decline more sharply in extended dry periods, and with greater peak flows in wet periods.

The storage in a reach affects the peak flows and the rate of recession from peak flows. A proportion of the flow is stored, and then becomes available for the flow in the next time period. Larger values of c_1 in Eq. (13) lead to greater reductions in peak flows and longer recession curves. c_1 can be adjusted to minimize the sums of squares of deviations between measured and observed flows. The major rivers lose water particularly in the lower sections, and the value of c_{loss} is similarly adjusted to improve the fit of calculated flow duration curves to observed ones.

4. Results

Based on the calibration methodology described above, a number of experiments were carried out to examine the effects of the proposed calibration solution and the performance of CLPSO calibrator. We began by investigating the calibration effects of the water accounting model.

4.1. Calibration and validation performance of the MDB model under the CLPSO-based two-step framework

Since the intended purpose of the MDB model is to evaluate flow variability for water reallocation under different climate conditions (see Kirby et al., 2014; Kirby et al., 2015), this work focuses on the calibration of flow magnitude and frequency during the period when climatic and hydrologic records are available. The graphical comparisons between observed flow duration curves and simulated ones using the optimal parameter set for 57 catchments (there is no observation data for the River Murray at the mouth, which is the outlet of the 58th catchment in the model) are shown in Fig. 4. The figure indicates that the proposed calibration solution leads to satisfactory simulation of flow duration curves, especially for high flows that are important in integrated hydro-economic scenario simulation.

Performance indicators of calibrated models for 57 catchments are demonstrated in Table 3. We found that, aside from the site Warrego-FordsBridge, the NSE values for calibrating flow duration curves for 56 catchment sites ranged from 0.76 to 1.00. These values indicated that simulated flow duration curves perform very well ($0.75 < NSE < 1.00$) based on model performance ratings (Supplementary Material Table B1). The RSR values for the above 56 catchment sites varied from 0.05 to 0.49, meaning model performance for residual variations of simulated flow duration curves have a very good range ($0.00 \leq RSR \leq 0.50$). The PBIAS values for all flow duration curve calibrations ranged from − 20.34% to 24.92%. These values show the average magnitude of simulated flow duration curves range from *satisfactory* to *very good* based on model performance ratings. Based on performance ratings of the three types of indicators, the performance of 47 calibrated catchments could be evaluated as *very good* with $0.75 < NSE < 1.00$, $0.00 \leq RSR \leq 0.50$, and $PBIAS < \pm 10$.

We used 15 years' observed time series flow data from 1995 to 2009 for validation and the remaining time series data for calibration. Fig. 5 shows the calibrated and validated monthly flows in the Murray River at Hume and the Darling River at Burtundy. As could be seen from the visual inspection of hydrographs in the figure, the calibration is successful, with relatively good prediction. The calibrated time series flows fit the observation well (the simulations capture most of high and low flow events), although the calibration was conducted against flow duration curves, rather than time series flow data. The fitting performance of validated time series flows is slightly worse than that of calibrated ones, but exhibits sufficient accuracy for water resource prediction in our hydro-economic applications.

4.2. Effects of global calibration on catchment calibration

The global calibration is a required and useful procedure to adjust the parameter set optimised by the hierarchical calibration. Practically, according to different needs, there are a few ways to set up global calibration targets and carry out calibration experiments. We used the objective function described in Eq. (2) as the global calibration target and adjusted the searching ranges of some parameters in order for better calculation performance. The weights assigned to catchments were set as identical. Four catchment observation sites (Campaspe-Barn, Campaspe-Roch, Murray-SwanHill, and Murray-WakoolJunct) were selected to demonstrate the effects of hierarchical and global calibration (Fig. 6).

Since Campaspe-Barn and Murray-SwanHill are upstream catchments of Campaspe-Roch and Murray-WakoolJunct, respectively, over-calibration of these two upstream catchments results in poor effects of the calibration on downstream catchments (Fig. 6(a)–(d)). The global calibration is able to mitigate the deficiency. As shown in the bottom row in Fig. 6, after the global calibration was executed, the optimal

Fig. 4. The calibration effects for all catchments under the two-step calibration framework and with the CLPSO calibrator. The two ratios of objective functions estimated by the two-step CLPSO and the semi-automatic calibration (Excel Solver), and by the two-step CLPSO and the two-step SA for each catchment are also indicated.

objective function values of the upstream catchments increased, comparing to those obtained by the hierarchical calibration. The increase of the optimal objective function values means worse fitting effects between observed and simulated flow duration curves, and this is also reflected in the graphical hydrographs. However, the concessions of fitting effects made by upstream catchments bring improvements in the calibration of downstream catchments. The minimum objective function values for Campaspe-Roch and Murray-WakoolJunct are only 24% and 15% of those in the hierarchical calibration, respectively. Visually, the calibration effects of Campaspe-Roch and Murray-WakoolJunct are dramatically improved. The minimum objective function value for the combined upstream and downstream catchments (such as Campaspe-Barn and Campaspe-Roch) was significantly improved in both cases.

Table 3
The performance of the two-step CLPSO calibration based on the three statistics.

ID	Catchment Name	NSE	PBIAS	RSR	ID	Catchment Name	NSE	PBIAS	RSR
1	Paroo-WillaraCrossing	0.93	1.03	0.27	30	Murrumbidgee-MaudeWeir	0.82	0.80	0.42
2	Warrego-FordsBridge	**−0.61**	−15.97	**1.27**	31	Murrumbidgee-Bal	0.99	−2.36	0.12
3	Condamine-ChinchillaWeir	1.00	0.91	0.06	32	Kiewa	0.99	1.05	0.08
4	Condamine-Stgeorge	0.97	−2.66	0.17	33	Ovens	0.99	0.87	0.08
5	Culgoa-DSCollerina	0.99	−2.20	0.09	34	Broken	0.97	13.01	0.17
6	Moonie-Gundablouie	0.99	0.99	0.08	35	Goulburn-Eildon	0.99	−1.06	0.11
7	Border-Pindari	0.97	7.19	0.17	36	Goulburn-Shepp	1.00	1.79	0.05
8	Border-Roseneath	1.00	0.15	0.06	37	Goulburn-McCoy	0.98	13.44	0.13
9	Border-Mungindi	0.97	−0.40	0.17	38	Campaspe-Barn	0.97	20.04	0.18
10	Gwydir-CopDam	0.99	2.29	0.10	39	Campaspe-Roch	1.00	1.99	0.06
11	Gwydir-Yaraman Bridge	0.98	4.05	0.13	40	Loddon-CCurran	0.98	3.36	0.14
12	Gwydir-Collarenabri	0.97	24.92	0.16	41	Loddon-ASouth	0.90	4.20	0.32
13	Namoi-KeepitDam	0.95	2.75	0.22	42	Avoca	0.87	−5.91	0.36
14	Namoi-Goangra	0.99	−0.01	0.07	43	WimmeraAvon	0.99	1.95	0.11
15	Castlereagh-Coonamble	1.00	0.88	0.06	44	Murray-Hume	0.99	0.16	0.09
16	Macquarie-DS-Burrendong	0.99	0.97	0.08	45	Murray-DrPoint	0.98	−0.27	0.14
17	Macquarie-MareboneWeir	0.88	0.74	0.35	46	Murray-Corowa	0.98	4.28	0.15
18	Macquarie-Carinda	0.99	1.44	0.10	47	Murray-YWeir	0.88	−20.34	0.35
19	Bogan-Gongolgon	1.00	1.02	0.07	48	Murray-PicnicPoint	0.96	4.35	0.21
20	Darling-BourkeTown	0.98	2.08	0.13	49	Murray-Barmah	0.82	13.60	0.42
21	Darling-Weir32	0.97	3.16	0.17	50	Murray-Torrumbarry	0.90	12.68	0.31
22	Darling-Burtundy	0.76	0.43	0.49	51	Murray-SwanHill	0.89	−7.23	0.33
23	Lachlan-Wyangala	0.99	3.23	0.10	52	Murray-WakoolJunct	0.99	4.82	0.10
24	Lachlan-Condoblin	0.97	4.33	0.17	53	Murray-Euston	0.99	2.40	0.09
25	Lachlan-Booligal	0.95	6.39	0.22	54	Murray-Wentworth	0.96	−3.99	0.21
26	Lachlan-Corrong	0.97	7.15	0.19	55	Murray-Lock9	0.94	0.10	0.25
27	Murrumbidgee-Burrinjuck	1.00	−0.95	0.07	56	Murray-RufusR	0.92	23.02	0.28
28	Murrumbidgee-Tumut	0.99	−1.32	0.08	57	Murray-Lock1	0.94	13.05	0.24
29	Murrumbidgee-Narrandera	0.99	−0.98	0.10		Murray-Mouth	n/a	n/a	n/a

4.3. The two-step calibration versus the semi-automatic calibration

The two-step calibration framework with the CLPSO as the calibrator performs much better than the semi-automatic calibration approach with Excel Solver (see the ratio in Eq. (3) for each catchment shown in Fig. 4), although the solution accuracy of the Solver on calibrating 4 in 57 catchments is superior to that of the two-step calibration. The average calibration performance ratio of the two-step calibration to the semi-automatic calibration over 57 catchments is 0.59, meaning that the two-step calibration has approximately 40% improvement in flow duration curve calibration across all catchments.

4.4. The CLPSO versus the SA under the two-step calibration framework

The solution accuracy of the SA is better than that of the Solver, but still worse than that of the CLPSO – only 14 in 57 catchments calibrated by the SA perform better than those calibrated by the CLPSO (Fig. 4).

We have illustrated the superiority of the CLPSO calibrator in solution accuracy. Below, we report the convergence characteristics of the CLPSO calibrator. Fig. 7 presents the means and standard deviations of

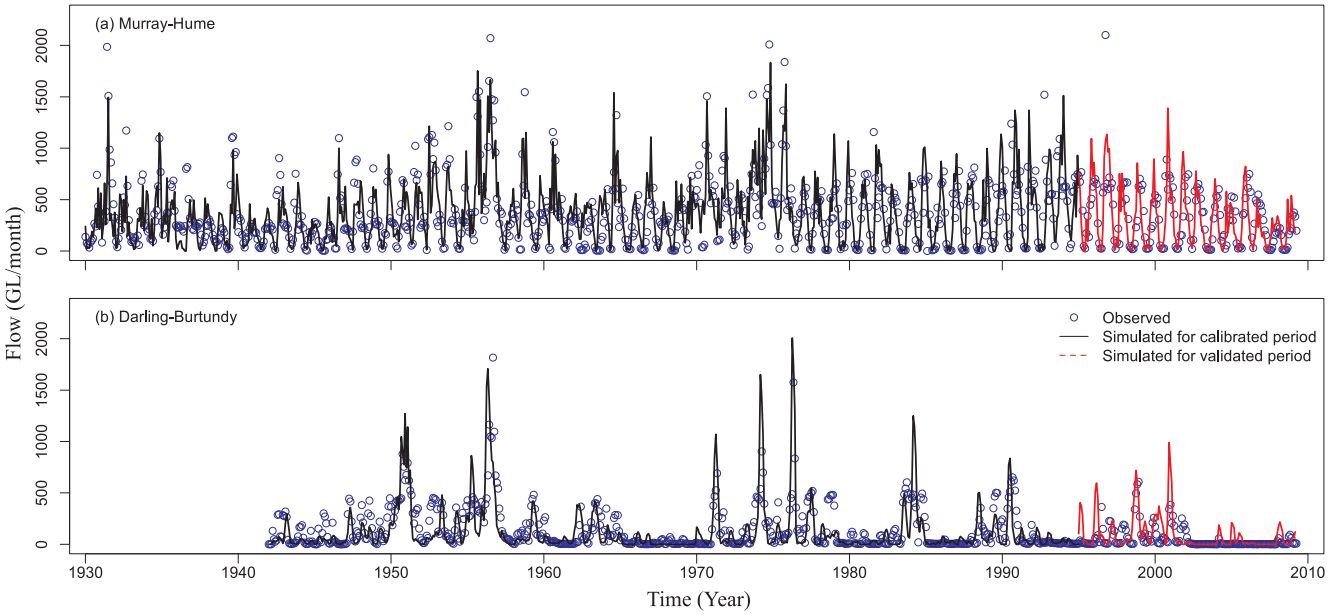

Fig. 5. Hydrograph of observed and simulated flows during both calibrated and validated periods in the Murray River at Hume (top) and the Darling River at Burtundy (bottom). See Fig. 2 for location.

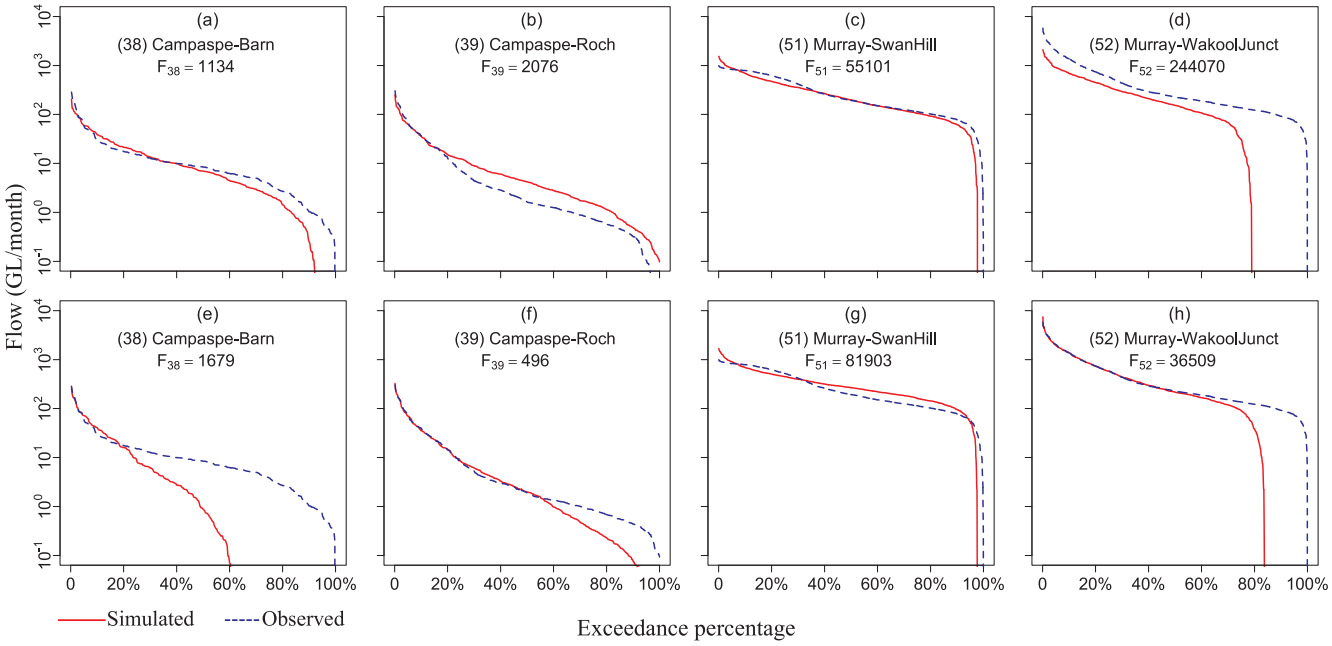

Fig. 6. The calibration effects from the hierarchical ((a)–(d)) and the global calibration ((e)–(h)) for four case study catchments 38, 39, 51, and 52. The graphical comparison between observed (in dash blue lines) and calibrated (in solid red lines) flow duration curves, as well as optimal objective function values obtained are presented in each subfigure. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

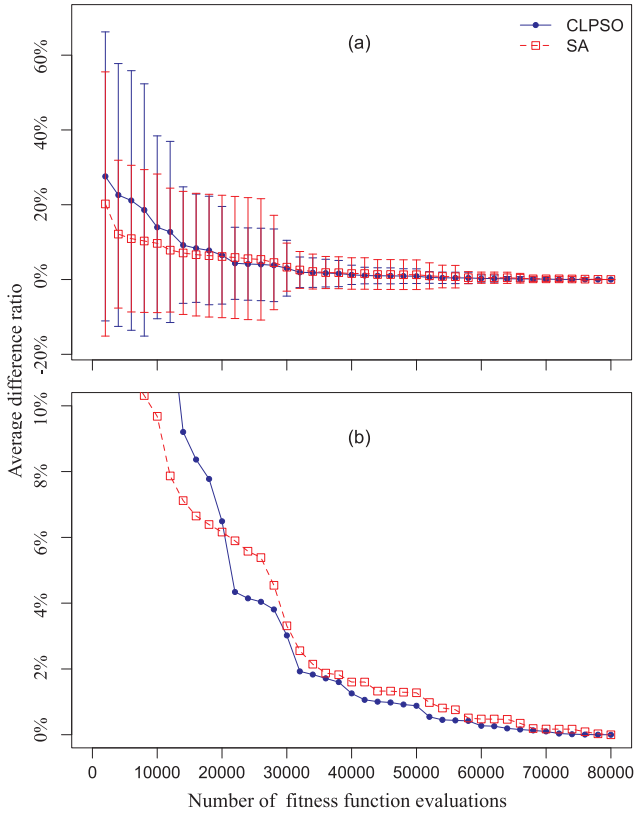

Fig. 7. The convergence performance of the CLPSO calibrator: (a) the means and standard deviations of difference ratios over the 57 catchments and (b) the means of difference ratios ranging from 0% to 10.

difference ratios (see equation (5)) over 57 catchments. In the CLPSO calibrator, the maximum number of FFEs equals the product of population size n of swarm S and the total number of iterations/generations. As FFE equals 40,000, the average difference ratio, which shows the average difference between the current best fitness value and final best fitness value, is 1.26%. When FFE is 60000, the ratio is only 0.27%. This value demonstrates the calibration effect with 60,000 FFEs is good enough, since the remaining 20,000 FFE efforts only improve less than 0.3%. The standard deviation continuously decreases with values 2.59% at 40,000 FFE and 0.96% at 60,000 FFE. The dynamics of the standard deviations indicates the convergence of the CLPSO is stable.

The convergence performance of the CLPSO was also compared to the SA calibrator (Fig. 7). At the initial 15,000 FFEs, the SA calibrator shows good convergence capability and the average difference ratios obtained from it are well below those obtained from the CLPSO. After that, the persistent convergence of the CLPSO calibrator leads to lower difference ratios relative to those in the SA case. The standard deviations of difference ratios obtained by the SA calibrator are greater, indicating that, in terms of the convergence stability/robustness, the CLPSO is superior to the SA.

5. Discussion

5.1. Effectiveness of the two-step calibration solution

It is difficult to effectively calibrate a basin-scale hydrological model with complex model structure, non-linear component interactions, and a substantial number of model parameters ($5 \times 57 = 285$ in our case). We have presented a hierarchical-global calibration approach that performs calibration from upstream catchments to downstream ones and fine-tunes the hierarchical calibration results. We have shown that the hierarchical calibration has very good performance on flow duration curve calibration. This calibration procedure based on flow duration curve also provided good calibration performance and acceptable prediction ability for time series flows (hydrographs). The global calibration procedure offers an opportunity to adjust the results from the hierarchical calibration based on specific needs. The example provided in Section 4.2 demonstrated that the calibration performance for downstream catchments could be significantly improved in the global calibration process by sacrificing some performance for upstream catchments. The hierarchical calibration used a divide-and-conquer approach to reduce computation demands and maintain spatial coherence. The hierarchical calibration results can be further tailored to a

specific application based on user or expert information in the global calibration process, and expert knowledge can be incorporated to reduce uncertainty.

The intended applications of the calibrated model are in exploring water allocation strategies, and thus fitting quality in high flows is emphasised. Towards this end, we focused on calibrating flow duration curves and chosen a measure of distance between ranked simulated and observed monthly flows. The two-step calibration framework (with the CLPSO as the calibrator) led to very good calibration performance in 47 of the total 57 calibrated catchments based on the indicators of NSE, RSR, and PBIAS (Table 3), with additional five catchments being evaluated as "Good" and four being measured as "Satisfactory". The catchment Warrego-FordsBridge was the only exception that was evaluated as "Unsatisfactory" due to a relative high mismatch in high flows. Both the visual calibration (Fig. 4) and validation (Fig. 5) results from the calibration framework indicate that the calibrated model is qualified for our intended purpose of water resources prediction and allocation.

5.2. Comparing the two-step calibration with existing calibration approaches

The two-step calibration is an automatic calibration framework, allowing for a fast and flexible way to produce very skillful model simulations. In practice, many water accounting models are still manually calibrated. However, manual calibration is much slower, and to obtain satisfactory calibration performance greatly depends on the knowledge and experience of the modeller. The automatic calibration framework allows for the accommodation of different calibrators and supports a parallel processing mode. The selection of a better calibrator can further improve the efficiency of the calibration framework.

The calibration of a large-scale (such as basin-scale) water accounting model is often conducted either manually or automatically on each component (such as catchment, gauge, or reach) of the basin system separately. Some work (e.g., Ajami et al., 2004) considered little on fluxes from outside of the component within the system. When these calibrated individual components are linked together and work as a whole system, omitted fluxes from upstream components will influence the downstream behaviours, result in inconsistencies, and propagate simulation errors across all downstream components (Hughes et al., 2016). The other work (e.g., Hughes et al., 2014) considered rive 'network' calibration (as opposed to the above isolated component calibration), where simulated fluxes from the upstream components were then used to calibrate the next downstream components. However, these studies implicitly set higher priorities to upstream components without any system-level adjustment. In this way, the calibrated parameters might be physically meaningless, violate calibration criteria for the whole river system, or suffer from the over-fitting especially where observations were erroneous. Meanwhile, these work lacked of a parallel mechanism to support efficient automatic calibration. The proposed framework can effectively reduce spatial error propagation, and calibrate model parameters in an efficient and systematic fashion.

Compared with the semi-automatic calibration using the Microsoft Office Excel Solver to conduct catchment-by-catchment calibration of model parameters within each catchment, the two-step calibration using the CLPSO has improved 40% in flow duration curve calculation and is more accurate in 93% of 57 catchments.

The framework is able to avoid dealing with highly dimensional and many-objective calibration tasks. With the rise of the number of parameters to be optimised, finding the optimum is increasingly difficult for calibrators and the computational load increases exponentially, especially where the calibration space is rough and multimodal. In this case study, applying a system-wide calibration to simultaneously deal with the calibration problem with 285 (5×57) parameters and goodness-of-fits of 57 catchments is difficult and computationally expensive. While the two-step calibration presents a hierarchical way to reduce the computational cost and then manages the tradeoffs of all components' goodness-of-fits in terms of expert knowledge or model predictive needs.

5.3. Calibration performance of the CLPSO

Effectiveness, efficiency, and robustness are three significant aspects to evaluate calibration performance and compare different calibrators (Gao and Hailu, 2010). The CLPSO outperformed the SA in solution accuracy in 75% of the modelled catchments. The mean and standard deviation trends of difference ratios (Fig. 7) reflect superiority of the CLPSO to the SA in calibration efficiency and robustness. The CLPSO improved standard PSO in enabling all searching particles to learn from good particles and to largely avoid low-efficient exploration of poor directions. This improvement gives the CLPSO an outstanding capability of dealing with complex multimodal problems. The calibration results and comparisons with the SA demonstrate that the CLPSO can improve calibration performance with regard to effectiveness, efficiency, and robustness.

A global sensitivity analysis (Gao and Bryan, 2016; Gao et al., 2016; Saltelli et al., 2008) is recommended to help modellers to screen out non-influential parameters so that the calibration performance can be further improved by reducing the number of model parameters to be calibrated.

5.4. Limitations and future directions

There are a number of limitations and possibilities for further investigation with respect to improve the calibration of the MDB water accounting model.

Hydrologic model predictability is subject to not only the calibration method itself, but also the availability and quality of observation data used for calibration. Our step-wise calibration approach is applicable to hydrologic models built for managed or 'working' rivers which are comprehensively gauged. There are many examples of such rivers throughout the world and therefore many relevant applications for our approach. For other rivers that have insufficient streamflow gauge stations, spatially distributed remotely sensed data, such as the evapotranspiration (ET) data from the Moderate Resolution Imaging Spectroradiometer (MODIS), can be used as an effective data resource for the calibration of hydrologic models. For example, Rajib et al. (2018) improved hydrologic model predictability by calibrating 29 sub-basins 'separately' with 29 representative time-series of MODIS ET (one for each sub-basin) and only one time-series of streamflow data from the outlet of the basin. Additionally, through the calibration process, uncertainty in the calibration data is converted into uncertainty in the estimated model parameters and other inferred quantities. Our future work includes the integration of Bayesian inference schemes (Peeters et al., 2018; Renard et al., 2010) into the proposed step-wise calibration framework, allowing the quantification of this 'derived' parametric uncertainty.

Multi-objective optimisation methods (Gao et al., 2014; Marler and Arora, 2004) are increasingly used in the calibration of hydrological models (Bekele and Nicklow, 2007; Rajib et al., 2016) to produce a Pareto set of optimal parameters in regard to several calibration objectives. What we would like to explore in the future is incorporating Pareto dominance into the CLPSO in order to allow this calibrator to handle calibration with several objective functions. For example, in the case study model, multiple objectives to be calibrated include flow duration curves, time-series flows, and diversions.

One major difficulty regarding the calibration of the water accounting model is the expensive computational demands. Therefore, one aspect that we would like to investigate is translating the model from Microsoft Excel VBA macro to other languages (such as Python) and utilising high performance computing technologies (Bryan, 2013).

6. Conclusions

Whole-of-basin hydrological models are needed to better manage the great river basins of the world which are threatened by a range of processes including climate change, over-extraction, and declines in water quality and flows (Best, 2019; Bryan et al., 2018). However, the lack of an effective means of calibrating these complex, many-parameter hydrological models has potentially restricted their development and application. We developed a new and effective calibration framework to improve the simulation accuracy of basin-scale hydrological models. It involves a hierarchical-global calibration strategy, aiming at effectively estimating a large number of model parameters and balancing trade-offs among objective functions for all catchments. The strategy includes a parallel processing algorithm to enable fast calibration for the hierarchical calibration, and expert knowledge is usually required to facilitate the global adjustment of calibration outcomes. Considering excellent performance of locating global optima, we implemented the CLPSO as a main calibrator in this framework, although other methods could also be used. Several calibration experiments were conducted in order to examine the effects of the two-step CLPSO calibration tool. The visual and quantitative calibration effects indicated that the calibration tool could lead to satisfactory model outcomes with sufficient accuracy for the model's intended purpose. The calibrated tool was also able to accurately simulate and predict time series flows. Global calibration significantly improved the calibration results obtained by the hierarchical calibration. This work offers a useful tool for calibrating complex, whole-of-basin hydrological models and reduces a significant barrier to the more widespread development of large-scale river basin modelling to support the better management of the world's river basins.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the CSIRO, Australia. L. Gao was supported by a CSIRO Julius Career award. The authors thank Rodrigo Rojas, Luk Peeters, and Dirk Mallants for their valuable comments. The authors are also thankful to two anonymous reviewers who have provided comments that have greatly improved the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhydrol.2019.124281.

References

Ajami, N.K., Gupta, H., Wagener, T., Sorooshian, S., 2004. Calibration of a semi-distributed hydrologic model for streamflow estimation along a river system. J. Hydrol. 298 (1–4), 112–135. https://doi.org/10.1016/j.jhydrol.2004.03.033.

Andersen, J., Refsgaard, J.C., Jensen, K.H., 2001. Distributed hydrological modelling of the Senegal River Basin — model construction and validation. J. Hydrol. 247 (3–4), 200–214. https://doi.org/10.1016/s0022-1694(01)00384-5.

Bekele, E.G., Nicklow, J.W., 2007. Multi-objective automatic calibration of SWAT using NSGA-II. J. Hydrol. 341 (3), 165–176. https://doi.org/10.1016/j.jhydrol.2007.05.014.

Best, J., 2019. Anthropogenic stresses on the world's big rivers. Nat. Geosci. 12 (1), 7–21. https://doi.org/10.1038/s41561-018-0262-x.

Booker, D.J., Snelder, T.H., 2012. Comparing methods for estimating flow duration curves at ungauged sites. J. Hydrol. 434–435, 78–94. https://doi.org/10.1016/j.jhydrol.2012.02.031.

Bryan, B.A., 2013. High-performance computing tools for the integrated assessment and modelling of social–ecological systems. Environ. Modell. Softw. 39, 295–303. https://doi.org/10.1016/j.envsoft.2012.02.006.

Bryan, B.A., Barry, S., Marvanek, S., 2009a. Agricultural commodity mapping for land use change assessment and environmental management: an application in the Murray-Darling Basin, Australia. J. Land Use Sci. 4 (3), 131–155. https://doi.org/10.1080/17474230802618722.

Bryan, B.A., et al., 2018. China's response to a national land-system sustainability emergency. Nature 559 (7713), 193–204. https://doi.org/10.1038/s41586-018-0280-2.

Bryan, B.A., Hajkowicz, S., Marvanek, S., Young, M.D., 2009b. Mapping economic returns to agriculture for informing environmental policy in the murray-darling Basin, Australia. Environ. Model. Assess. 14 (3), 375–390. https://doi.org/10.1007/s10666-008-9144-8.

Bryan, B.A., Marvanek, S., 2004. Quantifying and Valuing Land Use Change for ICM Evaluation in the Murray-Darling Basin 1996/97-2000/01. CSIRO Land and Water Stage 2 Report for the Murray-Darling Basin Commission.

Budyko, M.I., 1974. Climate and Life. Academic Press, San Diego.

CSIRO, 2008a. The Murray-Darling Basin Sustainable Yields Project. Overview and reports available at http://www.csiro.au/science/MDBSY.html (accessed July 2012).

CSIRO, 2008b. Water availability in the Murray-Darling Basin – A report to the Australian Government from the CSIRO, CSIRO, Australia.

Duan, Q., Sorooshian, S., Gupta, V., 1992. Effective and efficient global optimization for conceptual rainfall-runoff models. Water Resourc. Res. 28 (4), 1015–1031. https://doi.org/10.1029/91wr02985.

Engeland, K., Braud, I., Gottschalk, L., Leblois, E., 2006. Multi-objective regional modelling. J. Hydrol. 327 (3–4), 339–351. https://doi.org/10.1016/j.jhydrol.2005.11.022.

Ercan, M.B., Goodall, J.L., 2016. Design and implementation of a general software library for using NSGA-II with SWAT for multi-objective model calibration. Environ. Modell. Softw. 84, 112–120. https://doi.org/10.1016/j.envsoft.2016.06.017.

Gao, L., et al., 2014. A systems model combining process-based simulation and multi-objective optimisation for strategic management of mine water. Environ. Modell. Softw. 60, 250–264.

Gao, L., Bryan, B.A., 2016. Incorporating deep uncertainty into the elementary effects method for robust global sensitivity analysis. Ecol. Modell. 321, 1–9. https://doi.org/10.1016/j.ecolmodel.2015.10.016.

Gao, L., et al., 2016. Robust global sensitivity analysis under deep uncertainty via scenario analysis. Environ. Modell. Softw. 76, 154–166.

Gao, L., Ding, Y.-S., Ying, H., 2006. An adaptive social network-inspired approach to resource discovery for the complex grid systems. Int. J. Gen. Syst. 35 (3), 347–360.

Gao, L., Hailu, A., 2010. Comprehensive learning particle swarm optimizer for constrained mixed-variable optimization problems. Int. J. Comput. Int. Sys. 3 (6), 832–842.

Graveline, N., 2016. Economic calibrated models for water allocation in agricultural production: a review. Environ. Modell. Softw. 81, 12–25. https://doi.org/10.1016/j.envsoft.2016.03.004.

Grayson, R., Bloeschl, G., 2000. Spatial Processes, organizations and patterns. In: Grayson, R., Bloeschl, G. (Eds.), Spatial Patterns in Catchment Hydrology. Cambridge University Press, Cambridge, U. K., pp. 3–16.

Harou, J.J., et al., 2009. Hydro-economic models: concepts, design, applications, and future prospects. J. Hydrol. 375 (3–4), 627–643. https://doi.org/10.1016/j.jhydrol.2009.06.037.

He, Q., Wang, L., 2007. A hybrid particle swarm optimization with a feasibility-based rule for constrained optimization. Appl. Math. Comput. 186 (2), 1407–1422.

Huang, V.L., Suganthan, P.N., Liang, J.J., 2006. Comprehensive learning particle swarm optimizer for solving multiobjective optimization problems: research Articles. Int. J. Intell. Syst. 21 (2), 209–226.

Hughes, J.D., Dutta, D., Vaze, J., Kim, S.S.H., Podger, G., 2014. An automated multi-step calibration procedure for a river system model. Environ. Modell. Softw. 51, 173–183. https://doi.org/10.1016/j.envsoft.2013.09.024.

Hughes, J.D., Kim, S.S.H., Dutta, D., Vaze, J., 2016. Optimization of a multiple gauge, regulated river-system model. A system approach. Hydrol. Process. 30 (12), 1955–1967. https://doi.org/10.1002/hyp.10752.

Hughes, J.D., Kim, S.S.H., Yang, A., Dutta, D., Vaze, J., 2015. Whole of system calibration of river models: weighting functions and their effect on individual gauge and system performance. In: Weber, T., McPhee, M.J., Anderssen, R.S. (Eds.), 21st International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand, Gold Coast, Australia, pp. 2054–2060.

Jiang, Y., Li, X., Huang, C., 2013. Automatic calibration a hydrological model using a master–slave swarms shuffling evolution algorithm based on self-adaptive particle swarm optimization. Expert Syst. Appl. 40 (2), 752–757. https://doi.org/10.1016/j.eswa.2012.08.006.

Kavetski, D., Qin, Y., Kuczera, G., 2018. The fast and the robust: trade-offs between optimization robustness and cost in the calibration of environmental models. Water Resourc. Res. 54 (11), 9432–9455. https://doi.org/10.1029/2017wr022051.

Kennedy, J., Eberhart, R.C., 1995. Particle swarm optimization. In: IEEE International Conference on Neural Networks, Perth, Australia, pp. 1942–1948.

Khakbaz, B., Imam, B., Hsu, K., Sorooshian, S., 2012. From lumped to distributed via semi-distributed: calibration strategies for semi-distributed hydrologic models. J. Hydrol. 418, 61–77. https://doi.org/10.1016/j.jhydrol.2009.02.021.

Kirby, J.M., Connor, J., Ahmada, M.D., Gao, L., Mainuddin, M., 2014. Climate change and environmental water reallocation in the Murray-Darling Basin: impacts on flows, diversions and economic returns to irrigation. J. Hydrol. 518, 120–129.

Kirby, J.M., Mainuddin, M., Ahmad, M.D., Gao, L., 2013. Simplified monthly hydrology and irrigation water use model to explore sustainable water management options in the murray-darling basin. Water Resourc. Manage. 27 (11), 4083–4097. https://doi.org/10.1007/s11269-013-0397-x.

Kirby, M., Connor, J., Ahmad, M.-U.D., Gao, L., Mainuddin, M., 2015. Irrigator and environmental water management adaptation to climate change and water reallocation

in the murray–darling basin. Water Econ. Policy 01 (03). https://doi.org/10.1142/S2382624X15500095. 1550009.
Kirkpatrick, S., Gelatt, C.D., Vecchi, M.P., 1983. Optimization by simulated annealing. Science 220 (4598), 671–680. https://doi.org/10.1126/science.220.4598.671.
Kundu, D., van Ogtrop, F.F., Vervoort, R.W., 2016. Identifying model consistency through stepwise calibration to capture streamflow variability. Environ. Modell. Softw. 84, 1–17. https://doi.org/10.1016/j.envsoft.2016.06.013.
Li, M., Shao, Q., Zhang, L., Chiew, F.H.S., 2010. A new regionalization approach and its application to predict flow duration curve in ungauged basins. J. Hydrol. 389 (1–2), 137–145. https://doi.org/10.1016/j.jhydrol.2010.05.039.
Liang, J.J., Qin, A.K., Suganthan, P.N., Baskar, S., 2006. Comprehensive learning particle swarm optimizer for global optimization of multimodal functions. IEEE Trans. Evolut. Comput. 10 (3), 281–295.
Madsen, H., Wilson, G., Ammentorp, H.C., 2002. Comparison of different automated strategies for calibration of rainfall-runoff models. J. Hydrol. 261 (1–4), 48–59. https://doi.org/10.1016/s0022-1694(01)00619-9.
Maier, H.R., et al., 2014. Evolutionary algorithms and other metaheuristics in water resources: current status, research challenges and future directions. Environ. Modell. Softw. 62, 271–299. https://doi.org/10.1016/j.envsoft.2014.09.013.
Mainuddin, M., Kirby, M., Qureshi, M.E., 2007. Integrated hydrologic–economic modelling for analyzing water acquisition strategies in the Murray River Basin. Agric. Water Manage. 93 (3), 123–135. https://doi.org/10.1016/j.agwat.2007.06.011.
Marler, R.T., Arora, J.S., 2004. Survey of multi-objective optimization methods for engineering. Struct. Multidiscip. Optim. 26 (6), 369–395. https://doi.org/10.1007/s00158-003-0368-6.
MDBA, 2011. Water Audit Monitoring Reports. Available at http://www2.mdbc.gov.au/nrm/nrm_archives/cap_archives.html#WAM_reports (pre 2007-8) and http://www.mdba.gov.au/services/publications (2007-8 and 2008-9).
Molina-Navarro, E., Andersen, H.E., Nielsen, A., Thodsen, H., Trolle, D., 2017. The impact of the objective function in multi-site and multi-variable calibration of the SWAT model. Environ. Modell. Softw. 93, 255–267. https://doi.org/10.1016/j.envsoft.2017.03.018.
Moriasi, D.N., et al., 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Trans. ASABE 50 (3), 885–900.
Peeters, L.J.M., et al., 2018. Determining the initial spatial extent of an environmental impact assessment with a probabilistic screening methodology. Environ. Modell. Softw. 109, 353–367. https://doi.org/10.1016/j.envsoft.2018.08.020.
Qin, Y., Kavetski, D., Kuczera, G., 2018a. A Robust gauss-Newton algorithm for the optimization of hydrological models: benchmarking against industry-standard algorithms. Water Resour. Res. 54 (11), 9637–9654. https://doi.org/10.1029/2017wr022489.
Qin, Y., Kavetski, D., Kuczera, G., 2018b. A robust gauss-newton algorithm for the optimization of hydrological models: from standard Gauss-newton to robust Gauss-newton. Water Resour. Res. 54 (11), 9655–9683. https://doi.org/10.1029/2017wr022488.
Qureshi, M.E., Connor, J., Kirby, M., Mainuddin, M., 2007. Economic assessment of acquiring water for environmental flows in the Murray Basin*. Aust. J. Agric. Resour. Econ. 51 (3), 283–303. https://doi.org/10.1111/j.1467-8489.2007.00383.x.
Rajib, A., Evenson, G.R., Golden, H.E., Lane, C.R., 2018. Hydrologic model predictability improves with spatially explicit calibration using remotely sensed evapotranspiration and biophysical parameters. J. Hydrol. 567, 668–683. https://doi.org/10.1016/j.jhydrol.2018.10.024.
Rajib, M.A., Merwade, V., Yu, Z., 2016. Multi-objective calibration of a hydrologic model using spatially distributed remotely sensed/in-situ soil moisture. J. Hydrol. 536, 192–207. https://doi.org/10.1016/j.jhydrol.2016.02.037.
Renard, B., Kavetski, D., Kuczera, G., Thyer, M., Franks, S.W., 2010. Understanding predictive uncertainty in hydrologic modeling: the challenge of identifying input and structural errors. Water Resour. Res. 46 (5). https://doi.org/10.1029/2009wr008328.
Sadegh, M., Vrugt, J.A., Gupta, H.V., Xu, C., 2016. The soil water characteristic as new class of closed-form parametric expressions for the flow duration curve. J. Hydrol. 535, 438–456. https://doi.org/10.1016/j.jhydrol.2016.01.027.
Saltelli, A., et al., 2008. Global Sensitivity Analysis: The Primer. John Wiley & Sons, West Susses, England.
Seibert, J., 2000. Multi-criteria calibration of a conceptual runoff model using a genetic algorithm. Hydrol. Earth Syst. Sci. 4 (2), 215–224.
Shrestha, R.R., Rode, M., 2008. Multi-objective calibration and fuzzy preference selection of a distributed hydrological model. Environ. Modell. Softw. 23 (12), 1384–1395. https://doi.org/10.1016/j.envsoft.2008.04.001.
Sumner, N.R., Fleming, P.M., Bates, B.C., 1997. Calibration of a modified SFB model for twenty-five Australian catchments using simulated annealing. J. Hydrol. 197 (1–4), 166–188. https://doi.org/10.1016/s0022-1694(96)03277-5.
Thiemig, V., Rojas, R., Zambrano-Bigiarini, M., De Roo, A., 2013. Hydrological evaluation of satellite-based rainfall estimates over the Volta and Baro-Akobo Basin. J. Hydrol. 499, 324–338. https://doi.org/10.1016/j.jhydrol.2013.07.012.
Tolson, B.A., Shoemaker, C.A., 2007. Dynamically dimensioned search algorithm for computationally efficient watershed model calibration. Water. Resour. Res. 43 (1). https://doi.org/10.1029/2005wr004723.
Vaze, J., et al., 2011. Rainfall-runoff modelling across southeast Australia: datasets, models and results. Aust. J. Water Resour. 14 (2), 101–116. https://doi.org/10.1080/13241583.2011.11465379.
Vogel, R.M., Fennessey, N.M., 1994. Flow-duration curves. I: New interpretation and confidence intervals. J. Water Resour. Plann. Manage. 120 (4), 485–504. https://doi.org/10.1061/(ASCE)0733-9496(1994)120:4(485).
Vogel, R.M., Fennessey, N.M., 1995. Flow duration curves II: a review of applications in water resources planning. JAWRA J. Am. Water Resour. Assoc. 31 (6), 1029–1039. https://doi.org/10.1111/j.1752-1688.1995.tb03419.x.
Wang, Q.J., et al., 2011. Monthly versus daily water balance models in simulating monthly runoff. J. Hydrol. 404 (3–4), 166–175. https://doi.org/10.1016/j.jhydrol.2011.04.027.
Wang, S., et al., 2012. Multi-site calibration, validation, and sensitivity analysis of the MIKE SHE Model for a large watershed in northern China. Hydrol. Earth Syst. Sci. 16 (12), 4621–4632. https://doi.org/10.5194/hess-16-4621-2012.
Wöhling, T., Vrugt, J.A., 2011. Multiresponse multilayer vadose zone model calibration using Markov chain Monte Carlo simulation and field water retention data. Water Resour. Res. 47 (4), W04510. https://doi.org/10.1029/2010wr009265.
Zambrano-Bigiarini, M., Rojas, R., 2013. A model-independent Particle Swarm Optimisation software for model calibration. Environ. Modell. Softw. 43, 5–25. https://doi.org/10.1016/j.envsoft.2013.01.004.
Zhang, X., Srinivasan, R., Zhao, K., Liew, M.V., 2009. Evaluation of global optimization algorithms for parameter calibration of a computationally intensive hydrologic model. Hydrol. Process. 23 (3), 430–441.
Zhang, Y., Vaze, J., Chiew, F.H.S., Li, M., 2015. Comparing flow duration curve and rainfall–runoff modelling for predicting daily runoff in ungauged catchments. J. Hydrol. 525, 72–86. https://doi.org/10.1016/j.jhydrol.2015.03.043.