Two Stage History Matching for Hydrology Models via Machine Learning

Dewi Tjia, Ritu Gupta and Muhammad Alam

May 27, 2019

Two Stage History Matching for Hydrology Models via Machine Learning

Dewi Tija, Ritu Gupta and Muhammad Alam

Abstract. The reliability of a hydrological model (HydM) highly depends on how well the model parameters are estimated through the history matching (HisM) process. Direct HisM (DHisM) that calibrate input parameters by iteratively executing a model is widely applied in water resource estimation. The computational time of DHisM is prohibitive, as a single run on the model may take several hours. In practice calibration accuracy is compromised to arrive at a solution. Therefore, it is desirable to develop a proxy model that can replace hydrological model in the HisM process. In this study, we propose a two stage HisM, wherein we first develop a proxy model for HydM using artificial neural network techniques. Next we apply, ant colony optimisation (ACO_R) and robust parameter estimation (ROPE) methods for calibrating the parameter of HydM. This methodology is illustrated for the Dandalups catchment of Western Australia to calibrate five global parameters of Land Use Change Incorporated Catchment (LUCICAT) by matching 33 annual daily streamflow peaks. The results reveal that the replacing the LUCICAT by proxy model reduces the computational time by more than 90% with similar accuracy to DHisM and show higher consistency (via standard deviation of RMSE) and reduction of parameter uncertainty compared to DHisM.

Keywords: history matching, proxy models, history matching methods.

1. Introduction

Hydrological model (HydroM) is an important tool for water resource estimation and management. This model comprises of mathematical equations that describes complex natural physical process of a catchment. Inputs of hydrological model are climatic data and catchment characteristic. Catchment characteristics need to be calibrated to match historical response, such as streamflow, salinity, groundwater level. The process of adjusting input parameters of a catchment to reduce the discrepancy between model prediction and historical data is known as history matching (HisM).

HisM is a challenging task as hydrological model is often nonlinear and computational complex. For the past four decades, many researchers have conducted studies to automate HisM for challenging problems[6][23]. HisM methods typically work on an initial set of input parameters, and this set is recursively updated to minimise an objective function describing differences between the observed data and model response. In this process, a hydrological model is run several hundred and thousand times to obtain an optimal set of input parameters. According to Tjia's study [20], for a HisM approach, such as ant colony algorithm and robust parameter estimation, the computational time for HisM of a hydrological model take a few hours to a few days. For more complex hydrological model, for instance when physical processes are represented by grid cells, a hydrological model becomes computationally expensive [13]. Hence, the implementation of HisM using a HydroM directly, which is named as direct HisM approach, could be computationally challenging. In this study, we propose two stage HisM by first developing a proxy model for a HydM, then using this proxy model to find a set of optimal input parameter of HydroM.

A very well-known modelling method, namely artificial neural network (ANN) based on resilient backpropagation with one, two and three layers are used to develop proxy models in this paper. ANN has been widely applied in many fields, such as science, engineering [2][21] and medical science [24].

We apply two HisM methods of ant colony optimisation for continuous domain problem (ACO_R) and robust parameter estimation (ROPE). Both of these methods have been successful in solving the HisM problems in various fields of computer modelling. ACO_R which is inspired by foraging behavior of real ants has been used for HisM in water resource management [1][19] and engineering application [10][16]. ROPE based on geometrical position of a data in data set has been used in rain runoff modelling [12][17].

This paper is organised as follows. Section 2 describes material and methods used in this paper including the study area, the hydrology model (LUCICAT), description of ANN model and HisM methods (ACO_R and ROPE) and simulation settings. Section 3 provides results and discussions. Finally, conclusions and recommendations for further research are presented in Section 4.

2. Material and Methods

2.1. Study Area

Land Use Change Incorporated Catchment (LUCICAT) hydrological model is developed for the Dandalups catchment, Western Australia. The Dandalups is an important area as the two reservoirs within the catchment, namely North and South Dandalup supply Perth's annual water consumption. The location of the Dandalups, surface water bodies and gauging stations are depicted in Figure 1.The area of the catchment is about 700 km^2. The catchment has elevation ranging from 10 AHD (Australian Height Datum) and 500 AHD, with an average slope between 1.2% and 16.2%. The catchment has a Mediterranean climate characterised with a hot and dry summer season. The average annual precipitation is about 1,100 mm. The Dandalups has historical rainfall and streamflow data from 1 January 1960 to 14 November 1992 (33 years). HisM in this study is based on Node 14 (Figure 1) since this node is the most influential node for the catchment.

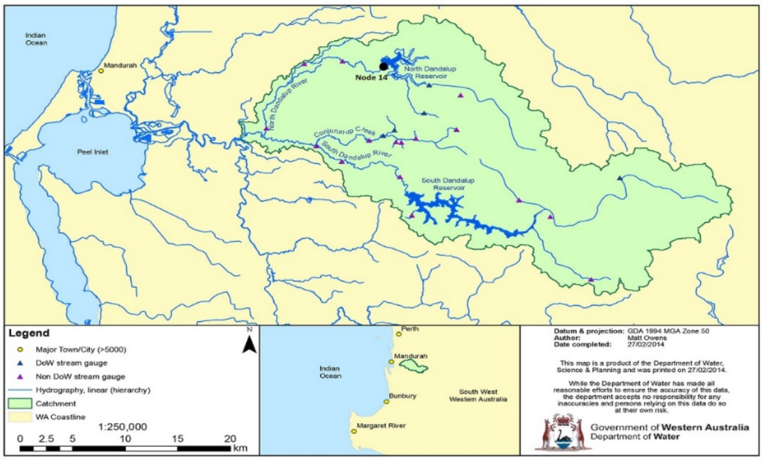


Fig. 1. The Dandalups (Owen, 2014)

2.2. LUCICAT Hydrological Model

LUCICAT hydrological model [3] is a distributed conceptual hydrological model used to predict streamflow and salinity due to the land and climate changes. The model has been used for many applications of catchment modelling in Western Australia [11]. In the LUCICAT model, a catchment is divided into several sub-catchments, known as Response Units (RUs) (Figure 2). Each RU has five interconnected stores, namely the upper store (dry and wet stores), subsurface store and groundwater store. These stores are responsible for generating the streamflow and transporting the salt. In the upper store, surface runoff (Q_{r1}) and interflow (Q_i) are generated. The remaining water in upper store percolate to the subsurface store and groundwater store. When the groundwater level reaches the stream bed, groundwater store creates the streamzone store and generates baseflow to the stream (Q_b). The streamzone store produces additional surface runoff (Q_{r2}).

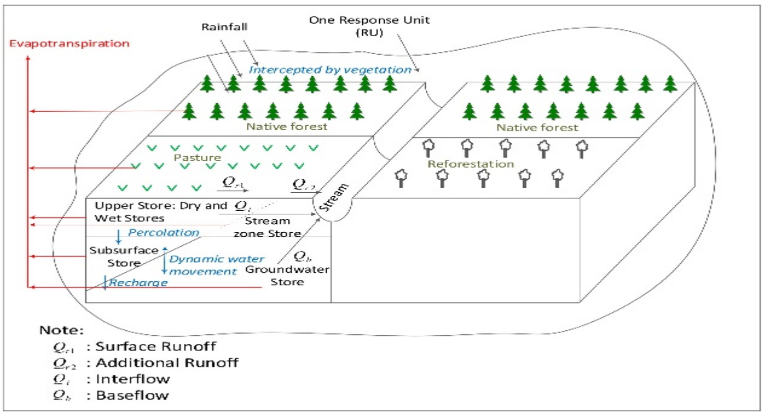


Fig. 2. LUCICAT building model

According to the sensitivity analysis test performed by the Department of Water, Western Australia, LUCICAT has five global sensitive parameters that significantly influence the model response [5]. Table 1 lists the parameters, their role and preset uncertainty ranges.

Table 1. Five global input parameters of LUCICAT model

Parameter	Role	Units	Range
ia	Interflow generation	-	1.8-3.1
K_{uv}	Vertical conductivity of top soil	Mm/d	15.3-27.2
b	Dry water soil moisture exponent	-	0.156-0.56
K_{ll}	Baseflow generation	Mm/d	400-1000
c	Wet water soil moisture exponent	-	0.156-0.56

2.3. Artificial Neural Network (ANN)

ANN method is inspired by the successful parallel working neurons in the brain. In ANN, artificial neurons are structured into several layers, which are input, hidden and output layers. Neurons in each layer were connected with sypnases. Figure 3a depicts ANN for the simplest ANN and ANN with one hidden layer. An ANN model is represented by the following relationship:

$$o(\boldsymbol{x}) = g(w_0 + \boldsymbol{w}^T\boldsymbol{x}) \tag{1}$$

and $g(.)$ is the activation function with characteristic as a bounded non-decreasing and differentiable function. The synaptic weights are iteratively updated to minimise the discrepancy between the predicted output and the actual training response using learning algorithm, such as backpropagation and resilient backpropagation

(Rprop). This process is known as training the model. Since in backpropagation, the weights lie in the layer far from the output layer, they receive slower learning (Miller, 1994), we use Rprop as learning algorithm [7][4].

2.4. History matching methods

Ant Colony Optimization based on Continuous Domain (ACO_R)

ACO_R was firstly introduced by Socha and Dorigo [18] to handle continuous optimisation problems. The idea of ACO_R lies on the construction of a solution archive, **X** with dimension of k rows (number of solutions) by p columns (number of input parameters). This solution archive is depicted in Figure 3b. Every cell of the archive has i^{th} unknown input parameter of j^{th} model solution denoted as x_{ij}. For each solution x_i, the objective function $h(x_i)$ weight w_i are computed. The weights are calculated according to the following formula:

$$w_i = \frac{1}{qk\sqrt{2\pi}} \exp(\frac{-(i-1)^2}{2q^2k^2}) \tag{2}$$

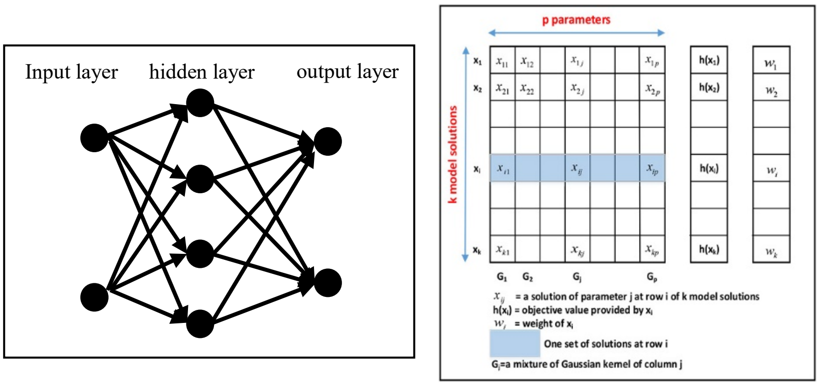
where q is the parameter control for selecting model in the solution archive. A mixture of Gaussian kernels is used to construct new members based on the information of the solution archive. For input parameter j, $G_j(\mathbf{x})$, comprises weighted sum of k one-dimensional Gaussian functions ($g_{ij}(x)$) expressed as follows:

$$G_j(\boldsymbol{x}) = \sum_{i=1}^{k} w_i \frac{1}{\sigma_{ij}\sqrt{2\pi}} \exp(\frac{(x-\mu_{ij})^2}{2(\sigma_{ij})^2}) \tag{3}$$

where $j \in \{1, \dots, p\}$. $G_j(\mathbf{x})$ has parameters a mean vector $\boldsymbol{\mu}_j$ and a vector of standard deviations $\boldsymbol{\sigma}_j$. Each element of $\boldsymbol{\mu}_j$ is defined by the corresponding value in the solution archive, $\mu_{ij} = x_{ij}$ for $i = 1, \dots, k$. Each component of $\boldsymbol{\sigma}_j$ is calculated by:

$$\sigma_{ij} = \xi \sum_{e=1}^{k} \frac{|x_{ej} - x_{ij}|}{k-1} \tag{4}$$

where ξ is pheromone evaporation that controls how fast the algorithm forgets the worse solutions. $\mathbf{X}_{new}$ with matrix size of $m \times p$ are constructed by taking samples from mixture of Gaussian kernels. $\mathbf{X}_{new}$ is combined with **X** and ranked. The m worst solutions are removed from the table to keep the size of archive the same and assign these solutions as **X**.



(a) (b)

Fig. 3. Example of ANN (a) and an ACO$_R$ solution archive (b)

Robust Parameter Estimation (ROPE)

ROPE was firstly proposed by Bardossy and Singh [2]. ROPE is geometrical approach based on data depth. Data depth is a measure of the degree of centrality of a point with respect to a multivariate data set. Many variants of data depth, in this study we focus on halfspace depth. Halfspace depth of points of a point $z \in R^p$ with respect to a p-dimensional data set is defined as the minimum number of data points lying on a closed halfspace plan with boundary of the plane through z.

$$D(z, \boldsymbol{X}) = minimum \ \# \ (i: u^T x_i \geq u^T z) \tag{5}$$

where $\boldsymbol{X} = \{\boldsymbol{x_i} = (x_{i1}, \ldots., x_{ip})$. $\boldsymbol{u}$ ranges over all vector in R^p, $||\boldsymbol{u}|| = 1$.

ROPE initially generates a set of n input parameters and calculates the corresponding objective function. p percent of this data set is identified as a set of good performing model parameters. Then, the algorithm searches a set of robust parameters within a good performing model parameter set by using the concept of halfspace depth, to generate m new input parameters.

2.5. Simulation Settings

Historical Response: This is 33 annual daily streamflow peaks for the period between 1960 and 1992. These peaks were considered the most influential in the choice of model parameters.

Proxy Model: ANN Rprop [8] with one, two and three hidden layers were trained using data set of size of 50, 100 and 200 generated using Latin Hypercube Design (LHD). The average performance ANN Rprop for each size was evaluated by using testing data sets, a size of 100 and the best average performance ANN model was selected for HisM.

History matching: LHD was used to design initial parameter settings for HisM algorithms. LHD with S-optimality criterion of size of 50 was used for ACO_R, while we used maximin criterion of size of 1000 for ROPE as LHD with S-optimality criterion is computationally extensive when the size of design is large as for ROPE. LHDs are generated via DoW R package [9]. ACO_R and ROPE were executed to minimise the objective function defined as:

$$RMSE = \sqrt{\frac{\sum_{j=1}^{33}\left(Y_{flow_j}^{h} - Y_{flow_j}^{s}\right)^2}{33}} \tag{6}$$

where $Y_{flow_j}^{h}$ and $Y_{flow_j}^{s}$ are the historical streamflow and the proxy model response for j^{th}, $j \in \{1, \ldots, 33\}$. HisM process was repeated until a stopping criterion was achieved. For ACO_R stopping criteria was a difference of less than 0.005 between the best RMSE of four consecutive iterations. For ROPE, HisM process was terminated if the mean RMSE of two consecutive iterations was less than 0.005. The tuning parameters of HisM methods were adopted from Tjia's study [20]. For ACO_R, k=m=50, ξ=0.5 and q=0.05 and ROPE, the initial settings (n) and new generation of solutions (m) were set each 1000 and the new solutions were selected from the best 10% (p) solutions of previous iterations. The performances of HisM using both algorithms were evaluated in terms of accuracy (mean RMSE), consistency (standard deviation of RMSE), computational time and reduction of initial parameter uncertainty for six experiments.

2.6 Simulation System

All coding routines of ACO_R and ROPE were written in R statistical software. For DHisM, LUCICAT model was called and executed in R environment when algorithms updated the input parameter and calculated RMSE. The simulation was performed using a PC with an Intel ® Core ™ i5-2500 CPU at 3.3 GHz, 8 Gb of RAM.

3. Result and Discussion

3.1. Performance of Proxy Models

The performances of ANN models for one, two and three hidden layers from various training data size are illustrated in Table 2.

Table 2. The best performances of ANN model from three training data size

Training data	Number of hidden layers			p-value
	one	two	three	
Size of 50	h1=70	h1=10, h2=80	h1=10, 2=50, h3=70	
Mean RMSE	0.207	0.199	0.197	0.311
Stdev	0.019	0.015	0.015	0.531
Comp.time (s)	24.0	50.7	89.0	0.000
Size of 100	h1=50	h1=10, h2=50	h1=10, 2=50, h3=70	
Mean RMSE	0.144	0.117_a	0.116_a	0.001
Stdev	0.000	0.007	0.006	0.064
Comp.time (s)	59.5	115.8	148.7	0.000
Size of 200	h1=60	h1=10, h2=50	h1=10, 2=10, h3=100	
Mean RMSE	0.101	0.079_a	0.078_a	0.042
Stdev	0.004_a	0.005_a	0.001	0.031
Comp.time (s)	388.9	435.7	1,224.7	0.000

Note: h1, h2, h3 are number of neurons in the first, second and third hidden layers respectively. Models with the same subscript represent pairs with equal means RMSE at 5% level of significance based on Tukey's multiple comparison, while variance is based on F-test for pairwise comparison with Bonferroni correction.

Table 2 shows that all ANN models of training data size of 50 have similar average performances (p-value=0.311). By contrast, when the training data size increases (100 and 200), there are significant difference in mean RMSE across number of layers (p-value=0.001 and p-value=0.042). Tukey's multiple comparison demonstrates ANN from training data size of 100 and 200 have comparable performances for two and three hidden layers and outperform ANN model with one hidden layer. It is also observed that there is similar consistency in standard deviation of RMSE for all number of hidden layers from training data size of 50 and 100. Unlike these results, ANN (based on training data size of 200) with three hidden layers is more consistent in RMSE than the models with one or two hidden layers. Comparing ANN performances from different size of data training within the same number of hidden layer, it is clearly seen the performances of ANN improves when the size of data used for model training increases. However, the computational time increases (up to 1000 fold) when the size of training data increases, and ANN structures are more complex (one to three hidden layers). The gain is accuracy is not consistent with the increase in computational time.

3.2. History Matching Results

In order to keep the results pragmatic, we only present HisM results using ANN model developed by training data size of 50. Our results show that HisM performances of using ANN model trained by data size of 50 (mean RMSE and standard deviation RMSE) are similar to those provided by ANN models of training data size of 100 and 200. The comparison of HisM results between DHisM using LUCICAT model and two stage HisM, with proxy model developed using ANN as explained above are listed in Table 3. These results are obtained by running six repeated experiments for which each experiment uses 400 model runs. According to t-test analysis for comparing mean RMSE, the performances of both HisM methods with DHisM, and both HisM methods with two stage HisM approach (ANN with one, two and three hidden layers) are comparable (all p-values t-test>0.05). In term of consistency (standard deviation of RMSE), F-test analysis shows that ACO_R with ANN three hidden layers are significantly more consistent than ACO_R in DHisM approach (p-value F-test = 0.001). In ROPE, higher consistency in RMSE is achieved when the algorithm coupled with ANN two (p-value F-test = 0.002) and three hidden layers (p-value F-test = 0.002).

Table 3. Performances of ACO_R and ROPE using direct and Two stage approaches

ACO_R with	LUCICAT	ANN with n hidden layers		
		n=1	n=2	n=3
Mean RMSE	3.665	3.671	3.670	3.658
Stdev	0.012	0.017	0.008	0.002
p-value t-test		0.471	0.426	0.273
p-value F-test		0.534	0.371	0.001
ROPE with	LUCICAT	ANN with n hidden layers		
		n=1	n=2	n=3
Mean RMSE	3.675	3.698	3.672	3.671
Stdev	0.012	0.028	0.002	0.003
p-value t-test		0.139	0.529	0.389
p-value F-test		0.058	0.002	0.002

Further results in Table 4 show that indirect HisM reduces significantly average computational time for DHisM by more than 90% for both ACO_R and ROPE. This concludes that indirect HisM provides very high efficiency in time for HisM. This procedure may become more advantageous when the hydrological model is more complex (using each grid to represent physical hydrological process for instance).

Table 4. Average computational time (s) using LUCICAT and ANN models

	LUCICAT	ANN with n hidden layers		
		n=1	n=2	n=3
ACO_R				
Data acquisition		2,484	2,484	2,484
Model development		24.0	115.8	148.7
HisM process	33,001	0.5	0.6	1.2
Total	33,001	2,508.5	2,600.4	2,633.9
% reduction		92.4	92.1	92.0

ROPE				
Data acquisition		2,484	2,484	2,484
Model development		24.0	115.8	148.7
HisM process	295,074	94.2	146.6	162.2
Total	295,074	2,602.2	2,746.4	2,794.9
% reduction		99.1	99.1	99.1

Another important criterion in evaluating the performances of models and methods are the reduction of initial input parameter uncertainty. Figure 4 illustrates 95% CI of obtained input parameter uncertainty in both direct and two stage HisM (using ANN model developed from 50 training data size). Most of input parameters result from indirect HisM show significant reduction from the initial range of input parameters (Table 1) and within the 95% CI of results of direct HisM. However, both algorithms with current settings are observed to have difficulties in reducing input parameter uncertainties produced by the error surface of ANN model with one hidden layer.

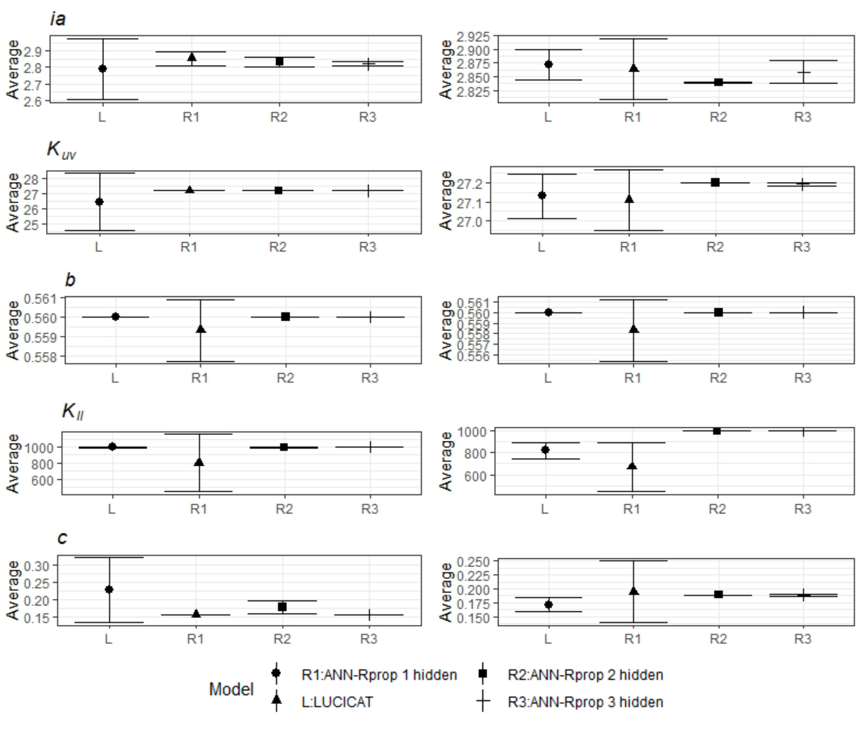


Fig. 4. Reduction of input parameter uncertainty between direct HisM and indirect HisM approaches with ACO_R (left) and ROPE (right).

The input parameters based on two stage HisM methods results in a range of streamflow between 22.5 and 22.9 million m^3/year in comparison to the estimate of 22.6 and 22.9 million m3/year for DHisM, requiring 90% less computational time. The estimates obtained are of similar order.

4. Conclusion

This paper discussed the implementation of two stage HisM using proxy models. The purpose of HisM is to estimate five sensitive input parameters of The Dandalups by matching the historical 33 annual daily peaks of streamflow. ANN based on Rprop algorithm were selected for developing proxy model for LUCICAT model as input in the process of HisM. HisM was then performed using ACO_R and ROPE HisM methods.

The performances of ANN with one, two and three hidden layers improved on average in terms of RMSE and consistency (standard deviation) as training data size increases. ANN from training data size of 50 have comparable performances across number of hidden layers. These findings are different from the results of larger training data size (100 and 200), where ANN with two and three hidden layers perform better than one hidden layer. For ANN, there was tendency of having similar variation in RMSE across number hidden layers when the models come from smaller training data size.. In spite of better performance, ANN models trained by larger training data set required longer computational time.

Two stage HisM results have comparable performances with DHisM methods in terms of mean RMSE. In terms of consistency, lower variation in RMSE values (standard deviation) occur when ACO_R and ROPE, coupled with ANN with two and three hidden layers in comparison to the HisM results with the LUCICAT model. Besides, the ranges provided by indirect HisM are much lower than those provided by direct HisM, with one exception for indirect HisM using ANN with one hidden layer. Furthermore, the application of two stage HisM reduces computational time used in direct HisM by more than 90%. This study shows the successful implementation of HisM using proxy models for the LUCICAT. We had similar experience in implementing two stage HisM for other catchments.

References

1. Ananda Babu, K., Shrivastav, R.K.: Developing objective function by ant colony optimization for continuous domains (ACO_R): Sondur Reservoir. International Journal of Latest Trends in Engineering and Technology (IJLTET) 2, 58-64 (2013).
2. Bardossy, A., Singh, S.K.: Robust estimation of hydrological model parameters. Hydrology and Earth System Sciences 12, 1273-1283 (2008).
3. Bari, M.A.: A distributed conceptual model for stream salinity generation processes: A systematic data based approach. Doctor of Philosophy thesis, The University of Western Australia (2005).
4. Cui, Y., Yi, Z., Duan, J., Shi, D., Wang, Z.: A Rprop-Neural-Network-Based PV Maximum power point algorithm with short-circuit current limitation. Electrical Engineering and System Science. IEEE ISGT NA (2019).
5. Department of Water: LUCICAT user manual (2012).

6. Duan, Q., Sorooshian, S., Gupta, V.: Effective and efficient global optimization for conceptual rainfall-runoff models. Water Resources Research 28(4), 1015-1031 (1992).
7. Fernandez, P.O., Teixeira, J. P., Ferreira, J., Azevedo, S.: Training Neural Network by Resilient Backpropagation Algorithm for Tourism Forcasting. Management Intelligence System, 41-46 (2013).
8. Fritsch, S., Guenther, F., Wright, M.N, Suling, M., Mueller, S.M.: Package ‘neuralnet’ version 1.44.2 (2019).
9. Groemping, U.: DoE.wrapper: Wrapper package for design of experiments functionality, r package version 0.31. https://cran.r-project.org/web/packages/DoE.base/index.html, last accessed 1/11/2017.
10. Hajizadeh, Y., Christie, M., Demyanov, V.: Ant colony optimization for history matching and uncertainty quantification of reservoir models. Journal of Petroleum Science and Engineering 77, 78-92 (2011).
11. Joyce, L.R.: The hydrological impacts of climate change and variability in the Murray Hotham catchment, Western Australia. Honours Dissertation. School of Environment Systems Engineering, The University of Western Australia (2007).
12. Krausse, T., Cullmann, J.: Toward a more representative parameterization of hydrologic models via synthesizing the strength of particle swarm optimization and robust parameter estimation. Hydrology and Earth Sciences 16, 603-629 (2012).
13. Kuntsmann, H., Heckl, A., Rimmer, A.: Physically based distributed hydrological modelling of the upper Jordan catchment and investigation of effective model equation. Advances Geosciences 9, 123-130 (2006).
14. Miller, M.: Rprop-description and implementation details. Technical Report. Institut fur Logic, Komplexitat und Dedutionsysteme, University of Karlsruhe (1994).
15. Owen, M.: The Dandalups Catchment scale 1:250,000. Department of Water, Western Australia.
16. Simsek, F., Mergen, F.: Optimization of induction motors using continuous ant colony systems. International Journal of Recent Trends in Engineering and Technology 4(3), 62-65 (2010).
17. Singh, S.K., Bardossy, A.: Hydrological model calibration by sequential replacement of weak parameter sets using depth function. Hydrology 2(2), 69-92 (2015).
18. Socha, K., Dorigo, M.: Ant colony optimization for continuous domains. European Journal of Operational Research 185, 1155-1173 (2008).
19. Tjia, D., Gupta, R., Alam, M.: Ants do history matching in hydrology. In: Ozwater 2015 12-14 May 2015, Adelaide, South Australia (2015).
20. Tjia, D.: Statistical Methods for history matching of hydrological model. Doctor of Philosophy thesis, Curtin University, Perth, Western Australia (2016).
21. Tumac, D.: Artificial neural network application to predict the sawability performance of large diameter circular saws. Measurement 80, 12-20 (2016).
22. Vaz, A.G.R, Elsinga, B., van Sark, W.G.J.H.M., Brito, M.C.: An artificial neural network to assess the impact of neighbouring photovoltaic systems in power forecasting in utrect, The Netherlands. Renewable Energy 85, 631-641 (2016).
23. Wang, Q. J.: The genetic algorithm and its application to calibrating conceptual rainfall-runoff models. Water Resource Research 27(9), 2467-2471 (1991).
24. Wang, Y. Li, J., Gu, J., Zhou, Z., Wang, Z.: Artificial neural network for infectious diarrhea prediction using meteorological factors sin Shanghai (China). Applied Soft Computing 35, 280-290 (2015).