

Adoption of suitable objective functions for auto-calibration in eWater Source: An application to Upper Hunter catchments in Australia

Md Mahmudul Haque^a, Golam Kibria^a

^a *WaterNSW, Parramatta, NSW, Australia*
Email: mahmudul.haque@waternsw.com.au

Abstract: The paper is aimed at exploring the applicability of different objective functions during calibration of a catchment adopting Sacramento rainfall-runoff model in eWater Source platform. A total of 10 objective functions are in-built in eWater Source platform, from where a hydrologist/modeller can choose a particular function from the list. The question is which one would give better calibration results considering the same calibration period and fixed optimisation algorithm for a given catchment. In this study, the testing has been done on two catchments in Upper Hunter region in Australia. In order to obtain a good calibrated parameter set, a modeller has to look for a number of goodness of fit tests (e.g. Nash Sutcliffe Efficiency (NSE), Flow duration curve and volume bias) depending on the applicability of the calibration. It has been found in the testing that no single objective function has been able to generate good results for a combination of goodness of fit tests. For example, if “NSE daily and log flow duration” is used in the calibration, it gives a reasonable NSE values but produces a bigger volume bias. On the other hand, if “Minimise absolute bias” is used, it generates very low volume bias and produces a good match between cumulative volume of simulated flow and observed flow, but it produces very poor flow duration curve. Therefore, a modeller need to carefully check and explore a number of objective functions to produce a good calibration results, sometimes it is often needed to do manual calibration after conducting auto-calibration.

Keywords: *Rainfall-runoff modelling, Sacramento, eWater Source, calibration*

1. INTRODUCTION

Calibration of a rainfall runoff model is a mandatory step in runoff estimation process of a catchment when observed data are available. eWater Source, evolved recently, which is Australia's national hydrological modelling platform, has several rainfall-runoff models (e.g. Sacramento, SIMHYD and AWBM) embedded into it (eWater 2017). One of the great challenges in runoff estimation process is to calibrate a rainfall-runoff model into eWater Source platform. To facilitate the calibration process, eWater Source platform has both the auto and manual calibration options. Moreover, it has a total of 10 objective functions (e.g. NSE daily & Bias Penalty, Minimise absolute bias and NSE daily & log flow duration) and 4 optimisation algorithms (i.e. Shuffled complex evolution (SCE), Uniform random sampling, Rosenbrock, SCE and then Rosenbrock) in built to do calibration. Now the question is, whether auto calibration option is capable of producing a good match between observed and simulated runoff or not?

In this paper, both the auto and manual calibration options of eWater Source model have been explored in two gauged catchments in Upper Hunter region, Australia to check the ability of producing good calibration results. In order to check the calibration results, a number of goodness of fit tests have been adopted (e.g. Nash-Sutcliffe Efficiency (NSE), Volume Bias and Flow Duration Curve). The reason for that, one particular objective function in the calibration process may good at producing better NSE values but may not be good at delivering better volume balance.

2. STUDY AREA

The Hunter Valley (Figure 1) is located around 130 kilometers north of Sydney. The catchment is bordered by the Barrington, Mount Royal, Great Dividing, Liverpool and Watagan Ranges (NSW Environment Protection Authority 2003). It is the largest coastal catchment in NSW, and has an area of around 21,500 km². The Hunter River begins in the Mount Royal Range on the western side of the Barrington Tops and flows around 460 km to enter the sea at Newcastle.

The Hunter Valley is often divided into two sub-regions: (i) Upper Hunter and (ii) Lower Hunter. Upper Hunter contains the area of Muswellbrook, Singleton and Upper Hunter, whereas Lower Hunter contains the area of Newcastle, Lake Macquarie, Matiland, Cessnock, Port Stephens and Dungog. The upper Hunter region comprises three Local Government Areas (LGAs): Upper Hunter, Muswellbrook and Singleton Shires. The Upper Hunter region includes the headwaters and upper reaches of the Hunter River, one of Australia's 'iconic' rivers and the major source of water for the region. The region is characterized topographically by low undulating hills that range from 80-180 m. Slopes vary between 2-10% and are typically 6%. The Upper Hunter valley covers rugged mountain ranges in the north, undulating farmland in the central and western regions, and widespread fluvial/estuarine flatland coastal areas. The valley's central and southern regions consist of gently undulating topography.

The Upper Hunter region has diversified vegetation and landscapes, which include the alpine rainforests of the Barrington Tops, the dissected sandstone of the Great Dividing Range, open grasslands and woodlands, and rich alluvial floodplains. Land use in the valley is diverse, with major uses being mining, power generation, urban development and agriculture. Upper Hunter valley has a total catchment area of around 21,500 km². It has been divided into 21 sub-catchments, mainly based on the following criteria: (i) Availability of the gauging stations, (ii) Length of the data period, (iii) Quality of the streamflow data, and (iv) Availability of water extraction data.

In this paper, performance of auto and manual calibration options of eWater Source model has been investigated in two gauged catchments, 210052 and 210084, which is marked in Figure 1. Area of this two sub catchments are 1073km² and 226 km² for 210052 and 210084, respectively.

3. CALIBRATION PROCESS

Each of the sub-catchments has been calibrated against the observed streamflow data from gauge station. Gridded rainfall and evaporation data has been used in the calibration, which have been extracted from Bureau of Meteorology. The SILO Australia-wide gridded dataset contain spatially interpolated rainfall grids at a resolution of 0.05° × 0.05° (i.e. approximately 25 km²) in daily time steps. Streamflow data have been extracted from Hydstra and quality codes of the data have been checked to ensure that data quality is good. Calibration has been done on the available data period where it represents a period of dry and wet conditions. In eWater Source, a total of 10 objective functions are in-built to use in the auto-calibration. Few objective functions can be tailored by giving different weighting in the combined objective functions, for example the objective function "NSE daily & Flow duration" can tailored by giving either more or lesser weighting to the

NSE values in calibration. In the paper, auto calibration has been done with six objective functions (i.e (i) Square-root daily, exceedance and bias, (ii) NSE Log Daily & Bias Penalty, (iii) NSE Log Daily, (iv) NSE Daily and Bias Penalty, (v) NSE Daily and log Flow duration and (vi) Minimise Absolute Bias). Thereafter, looking at the different goodness of fits tests, in particular flow duration curve, cumulative balance, residual plot and time series plot, the best parameter has been selected.

Then, manual adjustment of the parameters has been done to improve the calibration results if needed. Manual adjustment of the parameters has been done in an interactive way by looking at the goodness of fits test results simultaneously. Optimization algorithm has been kept consistent for the auto calibration process which is SCE then Rosenbrock. In the optimization functions for all the auto-calibration, number of shuffles has been kept at “3” and number of iterations has been kept at “100”. The main objective of the calibration for these gauges is to simulate historical runoff by using the calibrated parameter for the last 100 years. This historical runoff will be then used to identify the reliability of the Upper Hunter Water Supply system for different users. Based on these objectives, it is important to have lesser volume bias between simulated and observed flows, and better flow duration curve during the calibration process. In hydrology, it is good practices to cross validate the calibration by subdividing the data set into two subsets: (i) calibration set and (ii) validation set (Gupta et al. 2008). However, in this study the available dataset has been used only for calibration to allow longer period of data in the test. Eventually, for the reliability analysis it will be cross validated by doing both calibration and validation.

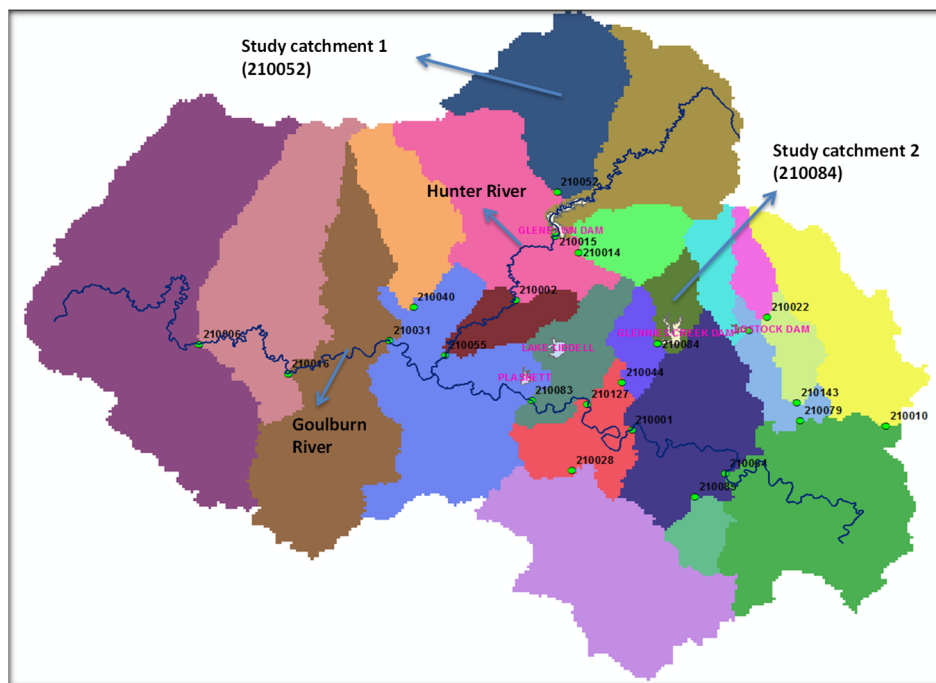


Figure 1. Delineation of sub-catchments in the Hunter region, gauging stations (small green circles) and two major rivers (Goulburn River and Hunter River), and two study catchments.

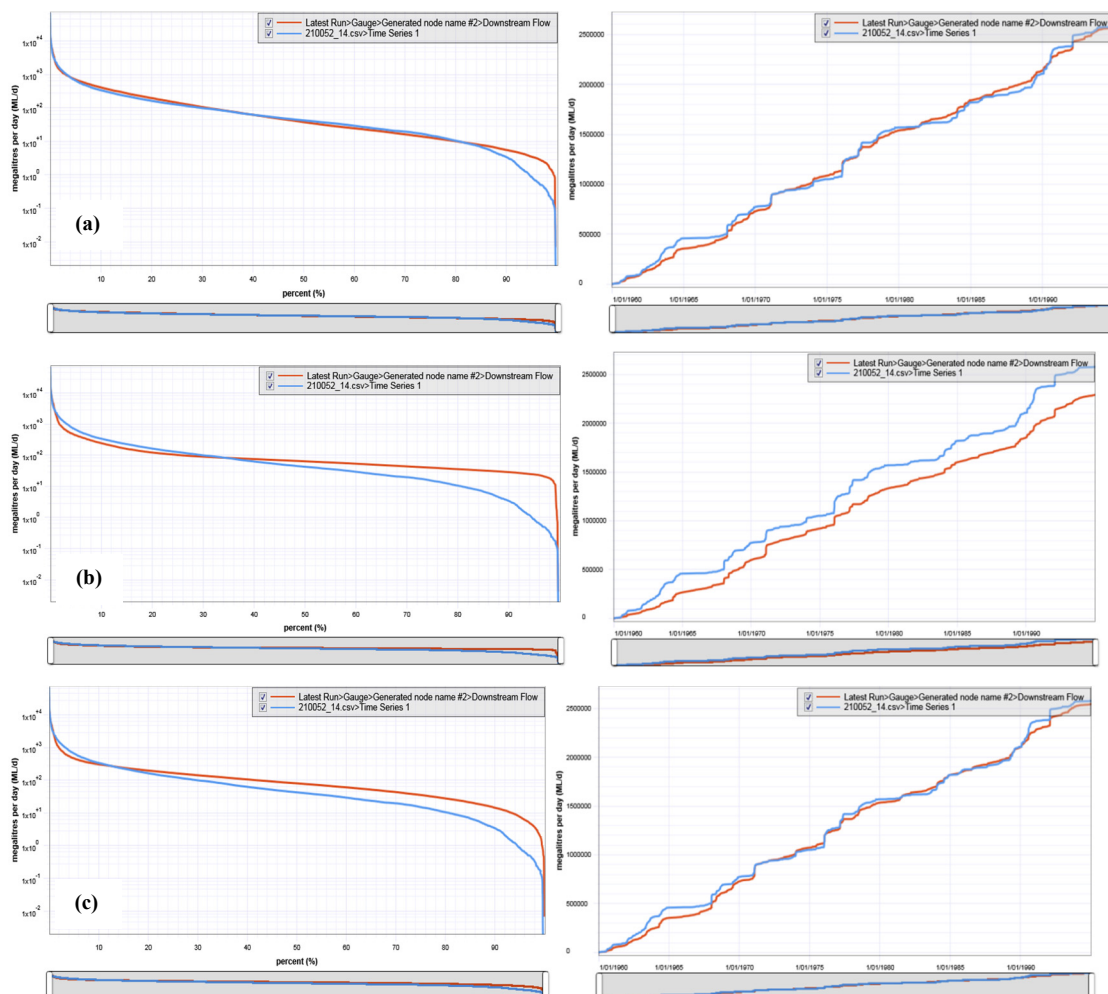
In eWater Source Platform, there are several rainfall-runoff models are embedded into it as mentioned in introduction section. In this paper, calibration of the study catchments has been done through Sacramento Rainfall Runoff model.

The Sacramento model is a continuous rainfall-runoff model used to generate daily streamflow from rainfall and evaporation records. The model was developed at US National Weather Service (NWS) in the early 1970s (Moore & Bell 2001, Podger 2004). It is a lumped model that considers the whole catchment as a set of several sub homogeneous units and uses soil moisture accounting to simulate the water balance within the catchment. Soil moisture storages are the main component in the model that determines the amount of runoff generation based on the size and relative size of the storages. The storages are filled by rainfall and reduced by evaporation and by flow of water from the storage. The runoff is transformed into surface flow through unit hydrograph. Lateral flows which also occur from the soil moisture stores are combined with the surface flow to yield the total stream flow.

4. RESULTS AND DISCUSSION

Calibration results for the catchment of 210052 are presented in Figure 2. Due to the limitation of page numbers, only flow duration curve and cumulative volume plots are presented. Calibration period was 1/1/1960 to 31/12/1994. From Figure 2, it is found that “Square-root daily, exceedance and bias” objective function has given relatively good match between simulated and observed flow for both the cumulative volume and flow duration plots among all other objective functions. “NSE daily” has not been able to produce good match for any of the plots. “NSE daily & bias penalty” has given good match for cumulative plots but produced poorer results for flow duration curve. “NSE daily & log flow duration” has given good results for flow duration curve but has not given good match for cumulative volume plot. “Minimise absolute bias” has been found better for cumulative volume match but found poorer for simulating flow duration curve. “NSE log daily and bias penalty” has produced good results for flow duration curve but nor for cumulative volume plots. These results indicate that a modeller should explore a number of objective functions during calibration process to take better results for the mentioned two plots.

Calibration statistics for the gauge 210052 at daily time steps in terms of NSE values, volume bias and Pearson correlation coefficient (r) is given in Table 1. It is found that “NSE daily” has given the highest NSE values but produced biggest volume bias (-11.13%). On the other hand, “minimise absolute bias” has produced best results for volume bias but given poorest results in terms of NSE value (0.28). Among all other objective functions compared, “Square-root daily, exceedance and bias” has simulated relatively better results for both the NSE and volume bias. Pearson correlations coefficients are found to be relatively same for all of them. These results also indicate that a modeller should look for a number of goodness of fit test and several objectives function to choose the best calibration parameter set. Since auto calibration with the function “Square-root daily, exceedance and bias” has produced a good match between simulated and observed flows, no manual calibration has been done at this gauge.



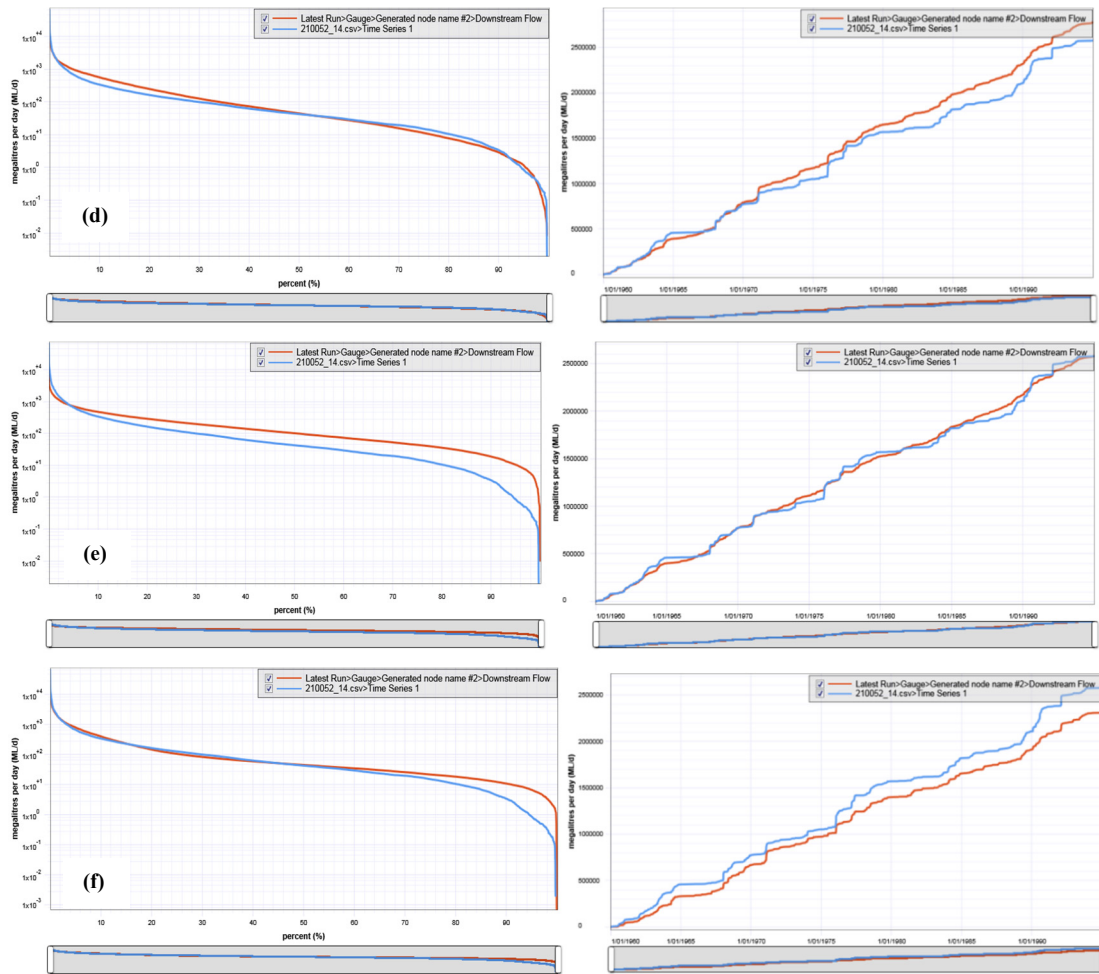


Figure 2. Plot of flow duration curve in log scale and cumulative volume in normal scale for the calibration of the gauge 210052 for the following objective function in auto calibration, (a) Square-root daily, exceedance and bias, (b) NSE daily, (c) NSE daily & bias penalty, (d) NSE daily & log flow duration, (e) Minimise absolute bias and (f) NSE log daily & bias penalty. Red line indicates the simulated flow and the blue line indicates observed flow.

Table 1. Calibration statistics for daily time series at the gauge 210052 for several objective functions through auto calibration options of eWater Source model

	Square-root daily, exceedance and bias	NSE daily	NSE daily & bias penalty	NSE daily & log flow duration	Minimise absolute bias	NSE log daily and bias penalty
NSE	0.6	0.62	0.61	0.49	0.28	0.46
Volume bias (%)	-0.14	-11.13	-1.28	7.72	0.04	10.33
r	0.78	0.79	0.78	0.72	0.64	0.72

Calibration results for the catchment of 210084 are presented in Figure 3. Calibration period was 5/11/1969 to 31/12/1979. Visual inspection of the plots for the above mentioned six objective functions of auto calibration has been done and only the best plot among them has been reported here. Moreover, plots for manual calibration have been presented. It is found that “minimise absolute bias” has produced the best results among all other objective functions in auto-calibration. In comparison to Figure 3(a) and 4(b), manual calibration has produced better match between observed and simulated flow for both the cumulative volume and flow duration curve than auto calibration.

Calibration statistics for the gauge 210084 at daily time steps in terms of NSE values, volume bias and Pearson correlation coefficient is given in Table 2. It is found that results of “minimise absolute bias” and

“manual calibration” are in the acceptable range, all other objective functions have given poorer results either in terms high volume bias or low NSE values. The objective function “Square-root daily, exceedance and bias” has found better during calibration of 210052 but has produced poorer results for 210084. The calibration results for these two sites indicate that a modeller should explore several objective functions in auto calibration to find the best calibrated parameter set. Also, a modeller should consider doing manual calibration when auto calibration cannot produce good match between observed and simulated flows. In this study, out of two catchments, manual calibration was needed for one catchment (210084) to get better calibration results than auto-calibration. In future, this study will be extended by doing auto and manual calibration at several catchments in Upper Hunter.

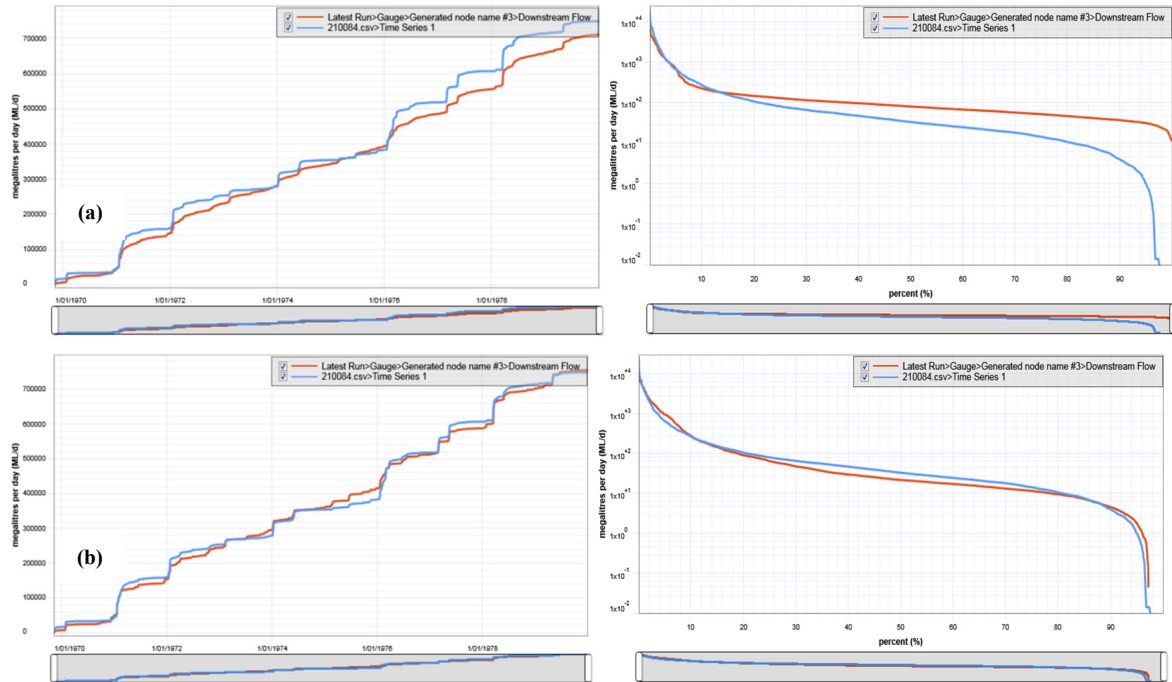


Figure 3. Plot of flow duration curve in log scale and cumulative volume in normal scale for the calibration of the gauge 210084 for auto calibration, (a) Minimise absolute bias and, (b) Manual calibration. Red line indicates the simulated flow and the blue line indicates observed flow.

Table 2. Calibration statistics for daily time series at the gauge 210084 for several objective functions through auto calibration options of eWater Source model and manual calibration

	Manual calibration	Square-root daily, exceedance and bias	NSE daily	NSE daily & bias penalty	NSE daily & log flow duration	Minimise absolute bias	NSE log daily and bias penalty
NSE	0.69	0.59	0.72	0.56	0.53	0.64	0.58
Volume bias (%)	0.81	-31.38	-8.15	-18.33	-49.6	-5.05	-25.91
r	0.83	0.84	0.85	0.8	0.82	0.83	0.83

Table 3. Comparison of simulated and observed high flows for 210052 and 210084

Date	210052			Date	210084				
	Simulated Flow (Square-root daily, exceedance and bias)	Observed Flow (ML/d)	Bias (%)		Simulated (ML/d) (manual calibration)	Simulated(ML/d) (minimise absolute bias)	Observed (ML/d)	Bias (%) (Manual calibration)	Bias (%) (Minimise absolute bias)
12/01/1968	17206	45692	-62	21/01/1971	10944	7448	11950	-8	-38
23/01/1976	17730	64982	-73	24/01/1972	5780	3916	17953	-68	-78

24/01/1976	33734	40390	-16	23/01/1976	4934	3721	11090	-56	-66
30/01/1984	20814	31329	-34	24/01/1976	7907	5174	9946	-20	-48
9/02/1992	33474	53987	-38	3/03/1977	3719	2805	14170	-74	-80
10/02/1992	18420	28708	-36	4/03/1977	8068	5205	9756	-17	-47

It is found during the calibration of 210052 and 210084 that no objective function in auto calibration has produced good match in high flow periods. It has produced under estimation of the high flows. As can be seen in Table 3, for 210052 the under estimation bias varies from 16% to 73%, for 210084, the underestimation bias varies from 17% to 80% during auto calibration. At 210084, manual calibration produces less under estimation bias than auto calibration, but still it varies from 8% to 74%. This under estimation bias for the high flows will be explored in more details in next phase of the study.

5. CONCLUSIONS

In this study, auto and manual calibration has been explored through eWater Source Modelling platform for two catchments in Upper Hunter, Australia. It is found that among the six objective functions used in this study (i.e. (a) Square-root daily, exceedance and bias, (b) NSE daily, (c) NSE daily & bias penalty, (d) NSE daily & log flow duration, (e) Minimise absolute bias and (f) NSE log daily & bias penalty), “Square-root daily, exceedance and bias” gives better results for one catchment and “minimise absolute bias” gives better results for another catchments. Since different objective functions have given better results for the two catchments, a modeller should explore several objective functions in auto calibration to get the better calibrated parameter set that will produce better match between observed and simulated flows. It is also found that auto calibration does not always produce better results. Manual calibration is often needed to produce better calibration results as in one catchment (210084) auto calibration could not produce good calibration results. The results also indicate that a modeller should look for different goodness of fits test to get the best calibrated parameter set, otherwise results will be biased. The high flows could not be well produced by the auto-calibration in eWater Source modelling platform that need to be investigated further.

ACKNOWLEDGMENTS

The authors express their gratitude to Bureau of Meteorology for providing gridded rainfall and evaporation data for the analysis. The authors also express their sincere thanks to Nanda Nadakumar at WaterNSW for assisting in the analysis.

REFERENCES

- eWater (2017). eWater Source — Australia's national hydrological modelling platform - Integrated Water Resources Management (IWRM) combined with water policy and governance capability. (Source: <http://ewater.org.au/products/ewater-source>).
- Gupta, H. V., Wagener, T. & Liu, Y. (2008). Reconciling theory with observations: Elements of a diagnostic approach to model evaluation, *Hydrological Processes*, 22, 3802–3813.
- Moore, R. J. & Bell, V. A. (2001). Comparison of Rainfall-Runoff Models for Flood Forecasting, Environment agency, R&D technical report: W241.
- NSW Environment Protection Authority (2003). Hunter River salinity trading scheme: Working together to protect river quality and sustain economic development. Sydney, NSW Environment Protection Authority.
- Podger, G. (2004). User Guide—Rainfall Runoff Library (RRL), Catchment Modelling Toolkit, CRC for Catchment Hydrology.