22nd International Congress on Modelling and Simulation, Hobart, Tasmania, Australia, 3 to 8 December 2017
mssanz.org.au/modsim2017

Assimilating stream flow, evapotranspiration and soil moisture data in AWRA-L model with particle filter

H.T. Pham [a, d], **C.N. Do** [b] **and L.J. Renzullo** [c]

[a] *University of New South Wales, Sydney, New South Wales 2052, Australia*
[b] *The University of Adelaide, Adelaide, South Australia 5005, Australia*
[c] *CSIRO Land and Water, Black Mountain, Australian Capital Territory, Australia*
[d] *The University of Danang, University of Science and Technology, Danang, Vietnam*
Email: thanhhung.pham@unsw.edu.au

Abstract: The Australian Water Resource Assessment-Landscape model (AWRA-L) is calibrated against a selection of data from ~500 gauged catchments around Australia to identify a single optimised set of model parameter with an emphasis on improving streamflow prediction. However, this regional approach to AWRA-L calibration can lead to high uncertainty in estimation, especially in ungauged catchments. An approach to help improve prediction in ungauged area is the assimilation of remotely sensed data into hydrological models. Streamflow discharges (Q), satellite soil moisture (SM) and satellite evapotranspiration (ET) observations have been individually assimilated into hydrological models to improve predicted outputs. This paper aims to evaluate performance of both individual and joint assimilation of these three hydrological observations into the AWRA-L model using particle filter technique.

The investigation used collected from six catchments across Australia with areas varying from 70 - 130 km^2. *In-situ* streamflow data from the Hydrological Reference Stations (HRS) are divided by their respective catchment area to generate observations with units that are consistent with AWRA-L modelled streamflow. The European Space Agency Climate Change Initiative (ESA-CCI) soil moisture products were normalised to ensure that observed and simulated soil moisture are comparable. We disaggregated the 8-day CSIRO MODIS ReScaled potential ET (CMRS-ET) product to daily ET estimates using daily potential evapotranspiration (PET) and a linear interpolation method. Forcing data, initial conditions and spatial parameters of six catchments are collected from the datasets used for calibration and validation of the AWRA-L model for 2010.

To address the limitation of high computational time required by the particle filter method in the grid-based AWRA-L model, we adopt a lump catchment approach where forcing inputs, initial conditions, and spatial parameters in each catchment were aggregated into one lumped value. Afterward the aggregated forcing is perturbed cell-wise using a normal distribution with 1000 samples to create a sample-based. These sample-based eventually are informed to the AWRA-L for the data assimilation process.

Four assimilation scenarios were investigated, namely: (1) sole assimilation of Q, (2) sole assimilation of satellite SM, (3) sole assimilation of satellite ET, and (4) joint assimilation of Q, SM and ET. In addition, an open-loop simulation, i.e. without assimilation, was run as a reference. Statistical metrics such as correlation coefficient (R^2), Nash-Sutcliffe model efficiency (NSE), root mean squared error (RMSE), and bias are used to assess the predicted outputs of the data assimilation model and open loop simulation.

Initial results indicated that only assimilating Q was successful in improving ET predictions in all study catchments. The assimilation of ET, however, did not improve streamflow predictions. Although assimilation of soil moisture produces a slight improvement in ET prediction, it degrades the predicted streamflow in this study. Difference with the single assimilations, the joint assimilation of all three observations can improve all predictions compared with the open loop simulation in some catchments. In addition, we found that the model performs poorly in a catchment with very small streamflow. These findings suggest further research efforts in order to fully understand the data assimilation problem. Nevertheless, the results in this paper have demonstrated the potential uses of data assimilation to reduce prediction uncertainty of the AWRA-L model across Australia.

Keywords: *AWRA-L model, particle filter, streamflow, soil moisture, evapotranspiration*

1. INTRODUCTION

The Australian Water Resources Assessment (AWRA) - Landscape model (AWRA-L) model is a daily grid-based water balance model which considers interactions between the atmosphere, the soil, groundwater and surface water stores in terms of the model water balance. The current model cell size is 0.05^0 primarily to match the spatial resolution of forcing meteorology datasets. Each model grid is classified into fractions of two hydrological response units (HRU), namely shallow-rooted and deep-rooted vegetation (Figure 1). At each grid cell, the model can estimate groundwater, surface water and soil moisture contents in different layers as well as leaf-biomass for the two HRUs using radiation, energy and water balance equations. The daily 0.05 degree precipitation, solar radiation and temperatures, and parameters at each grid cell are the model input datasets (Van Dijk 2010).

To improve the model predictions accuracy, it is possible to incorporate satellite observations in the AWRA-L model such as using remotely sensed data in model calibration, model evaluation and data assimilation (Van Dijk and Renzullo 2011).

For model calibration, Zhang et al. (2011) used satellite soil moisture and leaf area index to calibrate the AWRA-L model to attempt predictions of streamflow, soil moisture, and vegetation dynamics. Frost et al. (2015) calibrated and compared the AWRA-L model with different combinations of stream flow, evapotranspiration, and soil moisture. Currently, Kunnath-Poovakka et al. (2016) used remotely sensed evapotranspiration and soil moisture to calibrate the AWRA-L model to improve streamflow predictions in catchments with low average runoff. In terms of the model evaluation, Van Dijk et al. (2011) used GRACE data to evaluate the terrestrial water storage derived from the AWRA-L model. Regarding assimilation of satellite data in the AWRA-L model, Renzullo et al. (2014) assimilated remotely sensed soil moisture into the AWRA-L model using an Ensemble Kalman filter (EnKF) technique to predict root-zone soil moisture. Although there have been significant efforts on the model with data assimilation, still more research is required to understand estimates of the model prediction uncertainty and the model performance in terms of water balance.

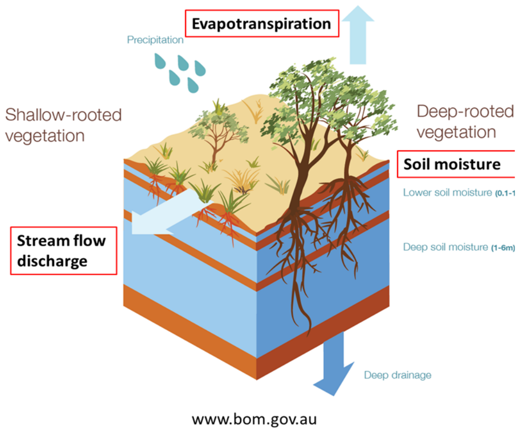

Figure 1. The AWRA-L model

This paper aims to develop both single and joint assimilation of streamflow, soil moisture, and ET observations in the AWRA-L model to improve predictions and quantify uncertainty in model terms. By implementing a joint assimilation to constrain the model with all observations, we attempt to find consistency between all water balance terms, modelled and observed.

2. METHODOLOGY

2.1. Study areas

The meteorological data and parameters in the AWRA-L model vary in space and time across the country. To evaluate robustness and consistency of the model with data assimilation, six catchments located across Australia are selected (Figure 2). Information of the six catchments is shown in Table 1. These selected catchments are small enough to ensure that a time-lag between rainfall and streamflow response at the downstream gauge is less than a day. Here we adopt a lumped catchment approach and aggregate a catchment into a grid cell.

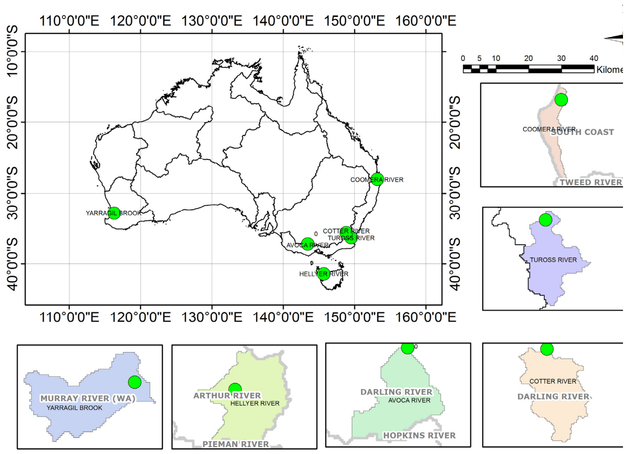

Figure 2. Locations of the six study catchments

Table 1. Information of the six study catchments

ID	Station name	River name	State	Area (km^2)	Lat	Lon
146010	Army Camp	Coomera	QLD	90	-28.03	153.19
218001	Tuross Vale	Tuross	NSW	97	-36.27	149.51
312061	Guilford Junction	Hellyer	TAS	100	-41.42	145.68
408202	Amphitheatre	Avoca	VIC	83	-37.18	143.41
410730	Gingera	Cotter	ACT	130	-35.59	148.82
614044	Yarragil Formation	Yarragil Brook	WA	71	-32.81	116.16

2.2. Data and data pre-processing

The forcing input datasets, initial conditions and spatial parameters of each catchment are selected from the benchmarking data of the AWRA-L model as a grid-based format. The model uses these input data to drive energy and water balance equations at all grid cells of a catchment. However particle filtering (Moradkhani et al. 2005) requires many particles (~1000's) per grid cell, which makes the application to spatial data assimilation intractable. So to reduce the dimensionality of the problem, and save on computing time, we adopt a lumped catchment approach to aggregate forcing input, initial conditions and model parameters of a catchment to a lumped value. We then perturb the lumped value into a sample-based. The AWRA-L model with particle filter will read input data from the sample-based rather than reading a large number of the grid-based catchment samples. The overview of the data pre-processing and data assimilation is shown in Figure 3.

Daily streamflow (Q) at catchment outlets are collected from the Hydrological Reference Stations (HRS). The observed streamflow in GL is divided by a corresponding catchment area in km^2 to generate streamflow observations in mm.

Evapotranspiration (ET) is derived from the CMRS-ET that is based on the MODIS reflectance and short wave infrared data, and gridded meteorological surfaces (Guerschman et al. 2009). A linear relationship between the 8-day CMRS-ET and the daily Australian 5km potential evaporation developed by CSIRO is used to disaggregate the composite data into daily ET observations.

Soil moisture (SM) are extracted from combination products of passive and active microwave ESA-CCI (Liu et al. 2012) at grid cells corresponding to catchment locations. So that simulated and observed data are comparable, soil moisture observations and simulations are normalised before implementing further calculations.

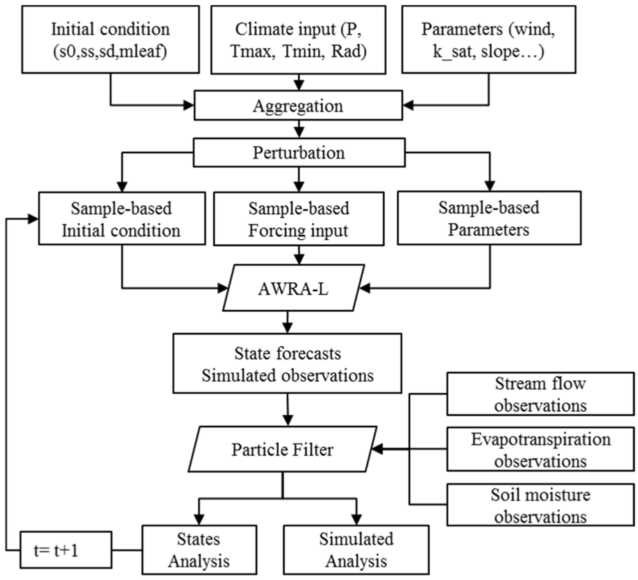

Figure 3. Overview of the research processes

2.3. Particle filter

The overview of the AWRA-L model coupling with particle filter approach is shown in Figure 3. This study uses streamflow, evapotranspiration and soil moisture observations to constrain the model and to estimate uncertainty of the model outputs. The forecast and analysis processes are presented as follows:

- Aggregate initial conditions, forcing inputs and spatial parameters of a catchment into a grid cell,
- Perturb the aggregated precipitation, temperatures and solar radiation to generate a sample-based with 1000 samples using Equation (1), where $\epsilon_t \sim U[a, b]$ for the precipitation error and $\epsilon_t \sim N(\mu, \sigma_\epsilon)$ for the other forcing input errors. Here, a and b are equal to 0.4 and 1.6 respectively. The statistical errors (μ and σ_ϵ) of the other forcing input data presented in Jones et al. (2009),

$$u_t^i = u_t^0 + \epsilon_t, \quad (1)$$

- Calculate weights using simulated data $\hat{y}_t^i$ and observed data y_t^o with assumption of a Gaussian distribution. The variance of streamflow, evapotranspiration or soil moisture (σ^2) are set to a corresponding observation multiplied by an arbitrary alpha (0.2).

$$\omega_t^i = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(\frac{-(\hat{y}_t^i - y_t^o)^2}{2\sigma^2}\right), \quad (2)$$

- Normalise the weights to ensure the sum of normalised weights is equal to 1,
- Use the resampling systematic method to resample particles to generate a new sample,
- Update the states corresponding with the new samples, and
- Use the updated states at time step *t* as initial conditions of the model at the next time step *t+1*.

2.4. Study scenarios and assessment

To evaluate the model performance in terms of the model water balance, we explored four assimilation scenarios including (1) assimilation of Q alone, (2) assimilation of ET alone, (3) assimilation of SM alone, and (4) the joint assimilation of all three observations. These scenarios are implemented over six catchments which have different climate conditions and different land properties to update ten model states including top soil, shallow root zone, deep root zone, leaf biomass in two HRUs as well as surface water and groundwater.

First, single assimilations of streamflow, evapotranspiration and soil moisture are implemented at all six catchments. Second, a joint assimilation of all three observations is applied to six catchments. An open-loop simulation, i.e. without data assimilation, is also performed to produce a reference set of simulated observations. Finally, the data assimilations and the open-loop simulation results were compared to assess the performance.

Different statistical metrics are used as the assessment criteria for different types of hydrological observation assimilation. Correlation coefficient (R^2), Nash-Sutcliffe model efficiency (NSE) and bias are used to assess the predicted streamflow, whereas sole R^2 is used to evaluate soil moisture prediction because of different units and measured layers in soil moisture. Lastly, R^2 and root mean squared error (RMSE) metrics are used to measure accuracy of ET predictions. To evaluate the model prediction uncertainty, 95% confidence intervals are also evaluated along with the predicted results in each scenario.

3. RESULTS AND DISCUSSION

We applied single and joint assimilation of three hydrological observations at all six catchments. Here, we present the results of four scenarios at catchment ID408202 (Table 2). The results of four scenarios at the other catchments (ID of catchment corresponding to the gauge ID in Table 1) are summarised in Table 3.

3.1. Single assimilation

Table 2. Statistical metrics of four scenarios at the gauge ID408202

Observations	Metrics	ID408202				
		Scenario 1	Scenario 2	Scenario 3	Scenario 4	Open-loop
Stream flow	R^2	0.87	0.28	0.42	0.75	0.73
	NSE	0.63	0.07	0.15	0.54	0.46
	Bias	-0.12	-0.39	-0.33	-0.23	-0.29
ET	R^2	0.81	0.92	0.81	0.83	0.81
	RMSE	0.65	0.43	0.63	0.60	0.72
Soil moisture	R^2	0.63	0.65	0.66	0.64	0.62

We begin with Scenario 1 by assimilating streamflow observations alone into the AWRA-L model. The predicted ET and SM are compared with the observations as a measure of performance. Assimilation of Q was observed to reduce the error between the simulations and the observations, especially during high rainfall events (Figure 4a and b). It is seen that three metrics of R^2, NSE and bias of Q are improved in comparison with these metrics of the simulated streamflow in the open-loop (Table 2). In addition, there are slightly improvements in predicted ET values in terms of RMSE value compared with these metrics of the open-loop simulation (Table 2). The predicted SM is relatively similar to the output of the open-loop simulation (Figure 4d). This means that assimilated Q somewhat improves the predicted ET and SM. The same results are found in the other catchments in this study, excepting the catchment ID614044 (Table 3). This result will be discussed in the last paragraph of this section.

In scenario 2, ET observations are assimilated in the AWRA-L model to improve the estimates of ET and the model predictions. As expected the modelled ET matches well the ET observations (Figure 5c). Although assimilation of ET observations can improve SM contents (Table 2 and Figure 5d), it degrades the streamflow values, especially in the peak events (Figure 5b and c). This is because the model increases the amount of water for ET and SM, and simultaneously decreases the amount of water for streamflow to ensure the model water balance. Values of R^2 and RMSE from the data assimilation in the other catchments are found to be better than these values from the open loop simulation.

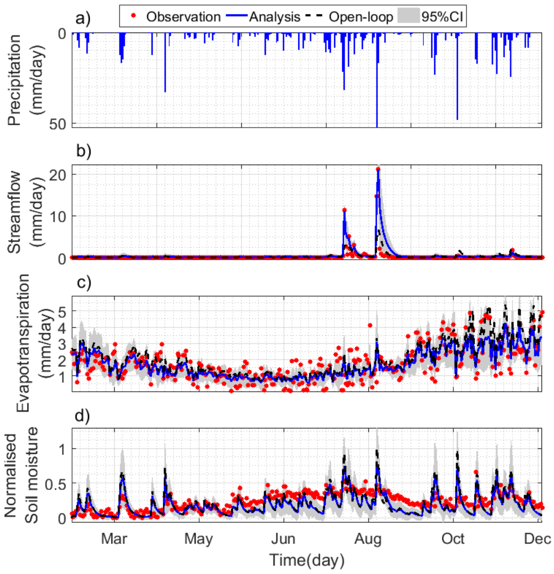

Figure 4. Single assimilation of streamflow

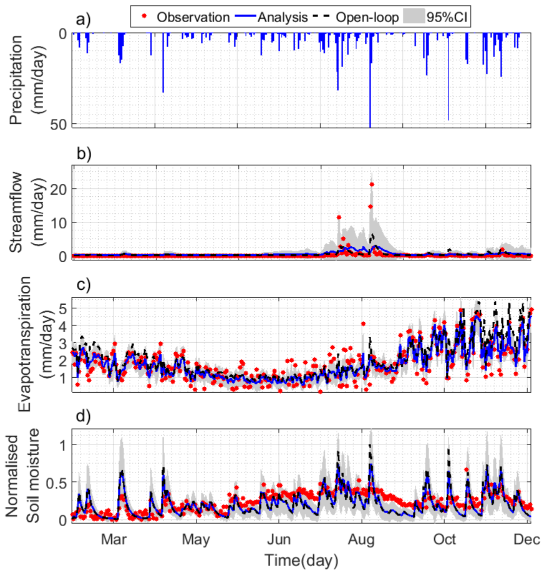

Figure 5. Assimilation of evapotranspiration

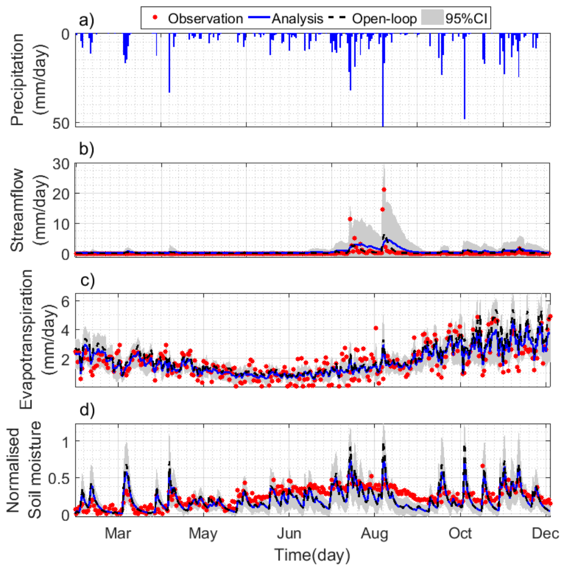

Figure 6. Single assimilation of soil moisture

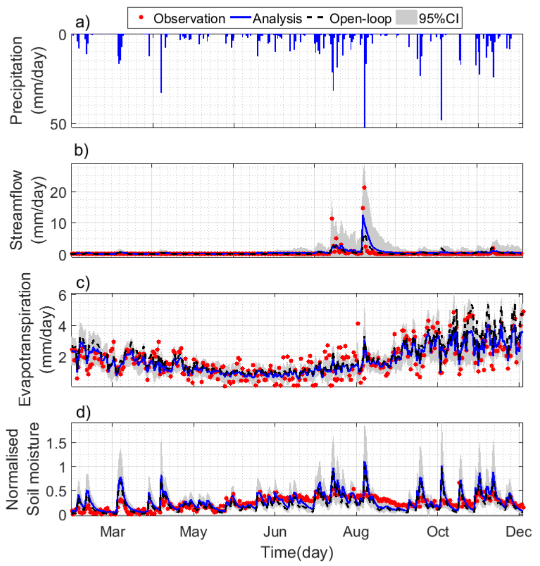

Figure 7. Joint assimilation of Q, ET and SM

In the third scenario, SM observation is assimilated in the AWRA-L model to improve estimates soil moisture content in the top soil layer and the predictions of Q as well as ET. The results show that the predicted streamflow values are underestimated in comparison with observations in the high intense rainfall events (Figure 6b). However the predicted streamflow discharges in the low flows are overestimations compared with observations, as negative bias value is found in this case (Table 2). Similar to the scenario 2 where only ET is assimilated, single assimilation of SM observations slightly improves the predicted ET (Table 2) in comparison with the results of the open-loop simulation. However the interactions between three hydrological observations detrimentally affects the outflow in this case, resulting in in this case. Again, the same results of the scenario 3 are found in the other study catchments (Table 3), excluding catchment ID146010 results. The assimilation of SM in this catchment improves both predictions of streamflow and ET. This inconsistent result was also found in the previous studies, for example Sun et al. (2016) indicated that assimilation of SM can improve the estimates of Q, whereas Yan and Moradkhani (2016) showed that the assimilated SM degrades the estimate of streamflow. Further research is needed to resolve this inconsistency.

It should be noted that the results for catchment ID614044 are different to the others, which may be due to conflict between input data and observations. The data in this area indicated that there are high solar radiation values, very arid, almost zero of Q and SM contents in a top soil layer; thereby the estimate of ET is expected to be small. However ET observations are still high from January to April. This mismatch can be a result of uncertainty in observations due to the retrieval algorithm or external factors such as wind, soil type, and vegetation. As a result, the statistical metrics are very poor (Table 3). It can be seen that in a very dry condition, the model cannot capture the observations in both cases with or without the data assimilation. This result suggests a further research of the model in short-term and long-term dry condition catchments as well as the effects of uncertainty in observations.

3.2. Joint assimilation

We consider joint assimilation (scenario 4) as a way of mitigating those instances when single source data assimilation does not improve the estimates The results demonstrate that there are improvements in the estimates of Q, ET as well as SM in comparison with the open-loop simulation (Table 2 and Figure 7). Note that in scenario 4 'improvement' cannot be independently verified as all the data are used in the assimilation. Here by improvement we mean agreement with the observation. The same result is found in the catchment ID312061 (Table 3). However in the other catchments, there are improvements in the estimates of ET and soil moistures compared with the open-loop simulation, but the accuracy of predicted streamflow is lower than the simulated streamflow from the model without data assimilation. These findings suggest further investigation required to identify the possible cause of the inconsistencies, as well as revisiting the formulation of the weights in the particle filter.

Table 3. Summary of three single assimilation scenarios at the other five catchments

Scenarios	ID	Data assimilation						Open-loop					
		Stream flow (Q)			Evapotranspiration (ET)		SM	Stream flow			Evapotranspiration		SM
		R^2	NSE	Bias	R^2	RMSE	R^2	R^2	NSE	Bias	R^2	RMSE	R^2
1 (Q)	146010	0.95	0.86	-0.29	0.78	0.99	0.64	0.71	0.49	-1.05	0.72	1.04	0.63
	218001	0.85	0.51	0.10	0.79	0.71	0.65	0.73	0.44	-0.23	0.73	0.75	0.66
	312061	0.99	0.99	-0.13	0.70	0.83	0.40	0.90	0.70	-0.50	0.68	0.95	0.47
	410730	0.99	0.98	0.01	0.85	0.84	0.54	0.9	0.78	0.32	0.76	0.77	0.58
	614044	0.22	-1.91	-0.07	0.04	1.53	0.65	0.23	-2.78	-0.18	0.8	0.9	0.63
2 (ET)	146010	0.71	0.32	0.82	0.97	0.40	0.65	0.71	0.49	-1.05	0.72	1.04	0.63
	218001	0.41	0.16	-0.60	0.95	0.36	0.65	0.73	0.44	-0.23	0.73	0.75	0.66
	312061	0.83	0.66	0.49	0.81	0.63	0.42	0.90	0.70	-0.50	0.68	0.95	0.47
	410730	0.48	0.02	-0.27	0.96	0.39	0.55	0.9	0.78	0.32	0.76	0.77	0.58
	614044	0.43	-3.3	-0.07	0.4	1.27	0.64	0.23	-2.78	-0.18	0.8	0.9	0.63
3 (SM)	146010	0.78	0.60	-0.37	0.84	0.87	0.69	0.71	0.49	-1.05	0.72	1.04	0.63
	218001	0.69	0.46	-0.43	0.79	0.74	0.67	0.73	0.44	-0.23	0.73	0.75	0.66
	312061	0.55	-0.33	-0.21	0.66	0.91	0.49	0.90	0.70	-0.50	0.68	0.95	0.47
	410730	0.69	-0.84	-2.03	0.86	0.82	0.60	0.9	0.78	0.32	0.76	0.77	0.58
	614044	0.21	-2.45	-0.18	0.05	1.56	0.67	0.23	-2.78	-0.18	0.8	0.9	0.63
4 (Q-ET-SM)	146010	0.64	-0.36	-1.31	0.86	0.78	0.67	0.71	0.49	-1.05	0.72	1.04	0.63
	218001	0.67	0.43	0.17	0.83	0.67	0.67	0.73	0.44	-0.23	0.73	0.75	0.66
	312061	0.94	0.73	-0.41	0.69	0.83	0.47	0.90	0.70	-0.50	0.68	0.95	0.47
	410730	0.83	0.63	-0.32	0.88	0.75	0.59	0.9	0.78	0.32	0.76	0.77	0.58
	614044	0.12	-0.42	-0.73	0.08	1.46	0.64	0.23	-2.78	-0.18	0.8	0.9	0.63

4. CONCLUSIONS

As with previous research, this paper explores the assimilation of individual observation of streamflow, ET or soil moisture into AWRA-L. However we also implement the joint assimilation of all three observations to understand the estimates of model uncertainty as well as the interactions between observations in terms of the model water balances. The results to-date demonstrated that:

- assimilation of only streamflow can improve the predictions of the ET in all catchments and soil moisture in some catchments,
- assimilation of solely ET does not always improve the predicted soil moisture and is shown to degrade the prediction of streamflow,
- individual assimilation of soil moisture can improve the ET in most catchments but can worsen the predicted streamflow, and
- joint assimilation could improve all predictions in comparison with the open-loop simulation in some catchments but this result needs more investigation.

In this study, we assumed the model parameters are non-stationary and a Gaussian distribution for white noise observations. Future research could consider a joint update of states and parameters and/or non-Gaussian distribution for the data error.

REFERENCES

Frost, A., Ramchurn, A., Hafeez, M., Zhao, F., Haverd, V., Beringer, J., & Briggs, P. (2015). Evaluation of AWRA-l: the Australian water resource assessment model. In, *21st International Congress on Modelling and Simulation, Gold Coast, Australia*

Guerschman, J.P., Van Dijk, A.I., Mattersdorf, G., Beringer, J., Hutley, L.B., Leuning, R., Pipunic, R.C., & Sherman, B.S. (2009). Scaling of potential evapotranspiration with MODIS data reproduces flux observations and catchment water balance observations across Australia. *Journal of Hydrology, 369*, 107-119

Jones, D.A., Wang, W., & Fawcett, R. (2009). High-quality spatial climate data-sets for Australia. *Australian Meteorological and Oceanographic Journal, 58*, 233

Kunnath-Poovakka, A., Ryu, D., Renzullo, L., & George, B. (2016). The efficacy of calibrating hydrologic model using remotely sensed evapotranspiration and soil moisture for streamflow prediction. *Journal of Hydrology, 535*, 509-524

Liu, Y.Y., Dorigo, W.A., Parinussa, R., de Jeu, R.A., Wagner, W., McCabe, M.F., Evans, J., & Van Dijk, A. (2012). Trend-preserving blending of passive and active microwave soil moisture retrievals. *Remote Sensing of Environment, 123*, 280-297

Moradkhani, H., Hsu, K.L., Gupta, H., & Sorooshian, S. (2005). Uncertainty assessment of hydrologic model states and parameters: Sequential data assimilation using the particle filter. *Water Resources Research, 41*

Renzullo, L.J., Van Dijk, A., Perraud, J.-M., Collins, D., Henderson, B., Jin, H., Smith, A., & McJannet, D. (2014). Continental satellite soil moisture data assimilation improves root-zone moisture analysis for water resources assessment. *Journal of Hydrology, 519*, 2747-2762

Sun, L., Seidou, O., Nistor, I., Goïta, K., & Magagi, R. (2016). Simultaneous assimilation of in situ soil moisture and streamflow in the SWAT model using the Extended Kalman Filter. *Journal of Hydrology, 543*, 671-685

Van Dijk, A. (2010). The Australian water resources assessment system. In, *Technical Report 4.Landscape Model (version 0.5) Evaluation Against Observations*

Van Dijk, A., & Renzullo, L.J. (2011). Water resource monitoring systems and the role of satellite observations. *Hydrology and Earth System Sciences, 15*, 39-55

Van Dijk, A., Renzullo, L.J., & Rodell, M. (2011). Use of Gravity Recovery and Climate Experiment terrestrial water storage retrievals to evaluate model estimates by the Australian water resources assessment system. *Water Resources Research, 47*

Yan, H., & Moradkhani, H. (2016). Combined assimilation of streamflow and satellite soil moisture with the particle filter and geostatistical modeling. *Advances in Water Resources, 94*, 364-378

Zhang, Y., Viney, N., Chiew, F., Van Dijk, A., & Liu, Y. (2011). Improving hydrological and vegetation modelling using regional model calibration schemes together with remote sensing data. In, *Proceedings of the 19th International Congress on Modelling and Simulation (MODSIM'11)* (pp. 3448-3454)