

Potential improvements to the Australian Water Resources Assessment system landscape (AWRA-L) model

A. Ramchurn and A. J. Frost

Bureau of Meteorology, Australia
 Email: a.ramchurn@bom.gov.au

Abstract: The Australian Water Resources Assessment system landscape model (AWRA-L), is being developed by CSIRO and the Bureau of Meteorology with the intention of providing robust estimates of water yield (runoff and baseflow), evapotranspiration, soil moisture, and aquifer recharge across Australia, specifically for retrospective Water Resource Assessment and National Water Account purposes. This 0.05° (~5km) gridded soil and groundwater balance model is undergoing continual conceptual and parameter estimation development, to reduce the uncertainty and error in the water balance estimates. Significant improvements in the model performance have been achieved to date from the initially parameterised AWRA-L v0.5, yet key areas of weaker performance remain.

This paper identifies areas where AWRA-L v3.0 can be improved when compared to standard and/or simpler models, it provides explanation for the current performance and suggests model improvements. AWRA-L results are compared to other standard rainfall runoff models (GR4J, Sacramento) and the newly developed SpringSIM model. The models are calibrated to streamflow for a set of 305 calibration catchments across Australia, and then validated in 304 additional catchments using nearest-neighbour calibration catchment as parameter set donor (split-sample catchment sampling to test performance in ungauged catchments) as well as using the donor catchment with a minimum separation distance of 50 km, or 200 km (testing the influence of distance on validation performance).

AWRA-L performs adequately using global calibration in the calibration catchments (Figure 1 shows model calibration F metric and bias B), but the results are inferior to locally calibrated models. However, the AWRA-L local calibration results, while improving on the globally calibrated results in calibration catchments, showed a markedly lower performance than the other models. Encouragingly the AWRA-L global calibration approach resulted in the least bias in validation catchments which confirmed its applicability continent wide when ungauged catchments are likely to be frequently more than 50km from the nearest gauged catchment. Furthermore, AWRA-L performed relatively well when distance was taken into consideration in validation. SpringSIM performed best in calibration, but poorest in validation.

Several areas of potential improvement in AWRA-L relate to spatialisation of properties controlling overland flow generation, streamflow routing, baseflow generation, and soil drainage. It was also found that AWRA-L could be improved structurally if mechanisms to "lose" water by means of an outlet threshold or to increase outflow by means of an interflow generation process were available. It may also be possible to improve results by ensuring that starting soil moisture states were realistic in validation.

Keywords: *AWRA, SpringSIM, Sacramento, GR4J, Rainfall-runoff modeling*

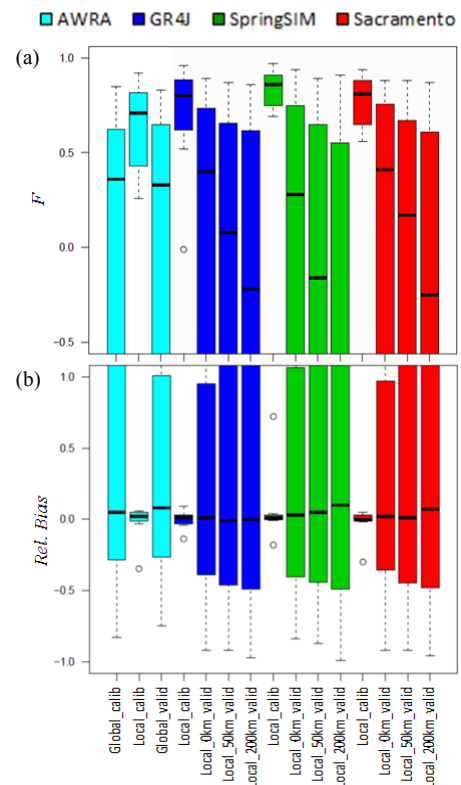


Figure 1. Calibration statistics for landscape water balance models AWRA-L, GR4J, SpringSIM and Sacramento (a) F calibration metric for streamflow and (b) bias in calibration of streamflow. For AWRA-L, results shown are for global calibration, local calibration and application of global parameters to validation catchments, For GR4J, SpringSIM and Sacramento: results are shown for calibration to 305 catchments (Local-calib) and validation in 304 catchments using 0 km, 50 km or 200 km minimum separation distance to donor catchment (Local_valid)

1. INTRODUCTION

The Australian Water Resources Assessment (AWRA) system (Van Dijk et al., 2011) is being developed by CSIRO and the Bureau of Meteorology with the intention of providing credible retrospective water balance for use in production of the Bureau's retrospective annual National Water Account (www.bom.gov.au/water/nwa) and Australian Water Resource Assessment Report (www.bom.gov.au/water/awra). AWRA-L (Van Dijk, 2010), the landscape soil water balance component of the system, estimates of water yield (runoff and baseflow), evapotranspiration, soil moisture, and aquifer recharge across Australia. This gridded model is designed to facilitate the incorporation of national/global datasets, such as remotely sensed evapotranspiration and soil moisture and streamflow, through calibration (e.g. Zhang et al., 2011) and assimilation (Renzullo et al., 2013; van Dijk and Renzullo, 2011) to improve the reliability of the water balance estimates made.

AWRA-L (currently v3.0) has gone through conceptual alterations and improvements since its initial uncalibrated version AWRA v0.5 described in van Dijk (2010). AWRA-L v3.0 differs from the description in van Dijk (2010) in that the following alterations have been made: (a) it uses a windspeed climatology rather than 3.5 m/s across Australia (McVicar et al., 2008), (b) it calculates the daylight fraction of a day, rather than using 0.5 across Australia, (c) vegetation height is estimated through Landsat based forest/non-forest maps (Opie et al., 2011), and (d) the model is calibrated to a national streamflow dataset, rather than using default parameter values. The application of a global calibration procedure, where a single set of parameters across all catchments is modulated to reproduce streamflow over a national streamflow dataset, has produced the greatest improvement to date (Viney et al., 2011).

Viney (2010) and Viney et al. (2013) describe comparisons of AWRA-L v0.5 and v3.0 respectively, with peer rainfall-runoff models, and show the marked improvement between the two versions. While the results of AWRA-L v3.0 are marginally superior compared to the peer models according to the defined calibration/validation metric (that focus on monthly and annual variability), key areas of unsatisfactory performance remain in terms of algorithms and spatial parameterization:

- large biases in stream flow estimation are present in some areas of the country (>100% in the Perth region) – making the current globally parameterized AWRA-L unsuitable for Bureau of Meteorology reporting purposes in those areas.
- relatively poor performance according to daily Nash-Sutcliffe efficiency - an important indicator of how well the model reproduces daily flow variability.
- Unrealistic trends in soil moisture have also been identified in two contrasting areas: an upward trend in soil moisture over 100 years in the arid centre of Australia; and no filling of the soil store in high rainfall Tasmanian environments.

This paper investigates the current level of model performance and reasons for it, towards suggesting alternate conceptualizations/parameterizations for improving AWRA-L, through comparison to a range of other rainfall-runoff models in calibration and validation mode. Spatial patterns in statistics are identified to investigate where differing conceptualizations perform better. Relationships of model performance and catchment properties are also investigated.

2. MODELS

2.1. AWRA-L

AWRA-L v3.0, as used operationally by the Bureau of Meteorology, is a grid distributed biophysical model that simulates water stores and fluxes in the vegetation. A description of AWRA-L v3.0 is provided in Peeters et al. (2013). Spatial resolution is 0.05° (ca. 5 km) as dictated by the resolution of the available daily climate forcing data. The water balance for the three soil reservoirs (top soil, shallow soil and deep soil) is calculated for each hydrological response unit within the grid cell while the water balance for the surface water and groundwater reservoir is computed for the entire grid cell. In this study, AWRA-L is tested in lumped catchment model mode, which is achieved by providing lumped catchment scale inputs to the model rather than gridded data, and simulating for each catchment individually rather than cells within the catchment. **Figure 2(a)** shows the AWRA-L conceptual catchment representation for each of the grid catchment/cells.

2.2. Comparative models

The models used in this study for comparison to AWRA-L v3.0 are GR4J (Perrin et al., 2003), Sacramento (Burnash, 1995) and SpringSIM (Ramchurn, 2012). GR4J (4 parameters) and Sacramento (13 parameters) models are widely accepted and used, and present two differing levels of parsimony for benchmarking purposes, with Sacramento having a comparable number of optimized parameters to AWRA-L (19 parameters). The newly developed SpringSIM model (12-parameters - conceptual representation shown in **Figure 2 (b)**) is also being used

for comparison as it explicitly attempts to address effects of extended droughts while ensuring the successful representation of low flows, high flows and total flow volumes (vital in terms of estimating water availability in wet and dry years within Australia, and thus useful for identification of potential structural changes for AWRA-L). SpringSIM achieves this through: 1) dynamically determining the unsaturated zone depth based on saturated zone depth; and 2) linking the flow observed at the catchment outlet to the volume of moisture available above a threshold within the average soil profile of the catchment.

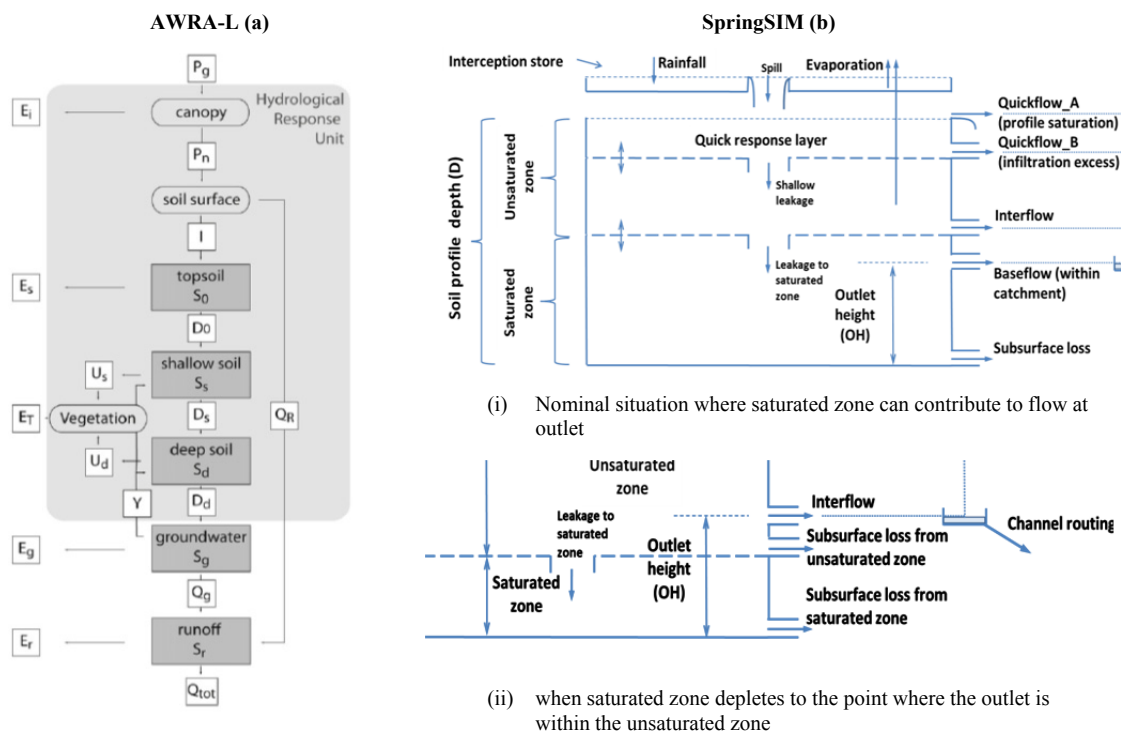


Figure 2 Model conceptual structure of (a) AWRA-L and (b) SpringSIM

3. DATA AND METHODOLOGY

3.1. Data – streamflow and climate data

Daily streamflow and various other catchment properties have been collated for 782 unregulated and unimpaired catchments, reflecting the variety of climatic, geological and hydrological types occurring across Australia, for the main purpose of calibrating and assessing the AWRA-L model (Zhang et al., 2013). The period of record covered by the dataset is 1975-2011, with there being at least 10 years of non-missing record available. These catchments have been randomly partitioned into a set for calibration (305 sites) and validation (304 sites) purposes, with remaining catchments excluded as they either exceed 5000 km² or have other catchments nested within, as AWRA-L does not include a distance based routing capability currently. The data was limited to 1980-2011 within this comparison for calibration, with 1980 used as a warm-up, and 1981 to 2011 used in the comparison; this is in-line with the approach currently employed for the AWRA-L calibration.

Daily climate data (precipitation, solar radiation, vapour pressure, maximum/minimum temperature) was extracted from the Bureau’s 0.05° gridded Australian Water Availability Project data archive (Jones et al., 2009; see www.bom.gov.au/jsp/awap), and averaged for each catchment. Potential evapotranspiration (for input into all models apart from AWRA-L) was calculated using the aforementioned climate inputs according to Morton’s wet environment areal potential evapotranspiration over land (Morton, 1983).

3.2. Calibration and validation

AWRA-L’s performance will be evaluated via a two-step comparison:

1. compare the performances of the models in local calibration with the previously achieved performance of the globally calibrated AWRA-L (produced in Viney et al., 2013) in the calibration catchments.
2. compare the validation performance of the globally calibrated AWRA-L in the validation catchments against the other models using nearest-neighbour, 50km and 200km minimum separation distance regionalisation.

Local calibration: Calibration of the models was facilitated through the use of the **hydromad** package (Andrews et al., 2011) within the R environment which contains a range of optimization, standard rainfall-runoff/routing

models and data visualisation/summary routines for lumped hydrological analysis. The package contains bundled versions of the Sacramento and of the GR4J models, while AWRA-L and SpringSIM were coded as Fortran DLLs. The default PORT (Gay, 1990) optimization algorithm was used for SpringSIM/GR4J (which was found to run quickly and robustly when compared to other algorithms), while in the interest of speed (due to the PORT algorithm taking excessively long time due to the number of parameters using the same settings as the other models), an initial solution search using Dynamically Dimensioned Search (Tolson and Shoemaker, 2007) was applied before using the PORT algorithm to obtain a quick/robust estimation of Sacramento and AWRA-L. It is recognized that using differing optimization algorithms for differing models may affect the results, however testing indicated that the results were robust for each model and the broad conclusions of this paper would not be effected by using consistent algorithms.

All the models were calibrated to maximize the following objective function for each catchment:

$$F = (E_d + E_m)/2 - 5 | \ln(1 + B) |^{2.5}$$

where E_d and E_m are respectively the Nash-Sutcliffe efficiencies of daily and monthly streamflow, and B is the bias (total prediction error divided by total observed streamflow). This objective function combines the objective function introduced in Viney et al (2009), with the E_m component included to improve monthly variability reproduction, aligned with the Bureau's annual/monthly reporting requirements.

Global calibration: The operational version of AWRA-L v3.0 was calibrated by CSIRO staff using a global calibration technique, where the AWRA-L model parameters are modulated to maximize the mean of the 25th, 50th, 75th and 100th percentiles of the F values of the 305 calibration catchments. This approach aims to exclude poor quality stream flow and climate observations from the calibration. 3 repeats of the Shuffled Complex Evolution (Duan et al., 1993) are applied to ensure robust optimization. The Microsoft Trident Workflow Environment is used to optimize the model written in C#, with optimisation applied on a 64 core cluster of 3 GHz machines, with the objective function evaluation parallelised to enable pragmatic optimization (~12 hrs). This is the first known successful approach at optimizing a landscape hydrology model in such a way.

Validation: For the locally calibrated models, validation using the nearest neighbour regionalization approach is used, whereby the parameters optimised for the closest calibrated catchment is applied to the "validation" catchment. This is used for comparative purposes for the globally calibrated AWRA-L. A limitation of this approach is that the regionalization method used for the lumped models does not take into account regional catchment characteristics in calibration, where catchment similarity indices could potentially have been used. The local calibration regionalization is also trialed using 50km and 200km separation distances, to investigate the effect of distance on regionalization performance.

Statistics: Performance was evaluated using catchment B , E_d , E_m , F and $\log E_d$ in calibration and validation, and selected statistics are presented in the results.

4. RESULTS

4.1. Performance in calibration

Figure 1(a) and (b) shows the calibration/validation performance according to F and B respectively. SpringSIM performed better than the other models in calibration. GR4J and Sacramento offered similar levels of performance scores, except for a more marked difference observed in the South-West Western Australia (SWWA: not shown).

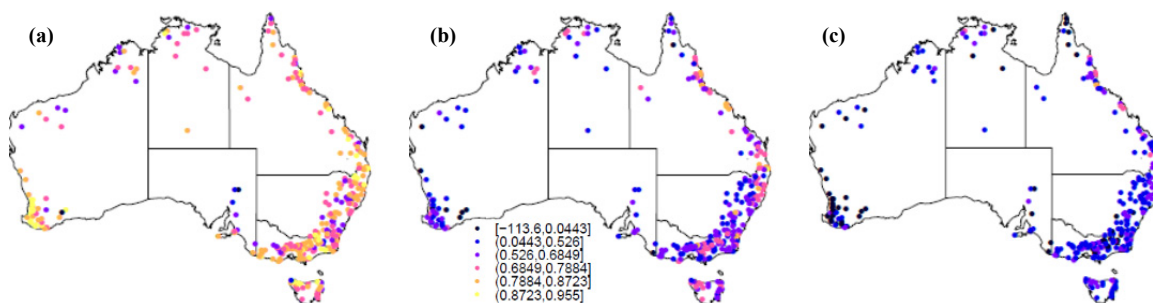


Figure 3 E_d for (a) Local calibration SpringSIM, (b) Local calibration AWRA, (c) Global calibration AWRA

This result cannot be solely attributed to the higher number of parameters (13, including calibration of the starting saturated zone level), as the Sacramento model is being calibrated with a similar number of parameters. SpringSIM is believed to perform best in calibration because of its more effective way of accounting for the soil moisture history and the application of the threshold level to limit what fraction of the soil moisture contributes to baseflow in the catchment. Sacramento contains these features also, however initial soil moisture store was not calibrated for that model (due to that not being possible in the framework used), which may explain the better

performance of SpringSIM. All models performed well in terms of B in local calibration due to the high importance placed on low bias in the objective function used.

The AWRA-L local calibration results are enlightening, they show a markedly lower performance to the other locally calibrated models although with a greater number of parameters freed (19). It suggests that there may be problems with the AWRA-L structure and with the effective usage/parameterization of spatially-distributed parameters. In both global and local calibration, AWRA performs worst in Western Australia. Reasons for the spatial distribution of Ed values in global calibration shown in Figure 3 are investigated in Section 5.

4.2. Performance in validation

Locally calibrated models: Validation performance for the locally calibrated models drops with separation distance (Figure 1). This deterioration in performance with distance is expected due to less similarity between catchment and climate characteristics. GR4J and Sacramento have the best performance for each separation distance, although SpringSIM (with the best results in calibration), has the poorest in validation.

AWRA global calibration: AWRA-L v3 in validation does not deteriorate relative to calibration, a benefit of the global calibration approach. While it performs worse than the nearest neighbour GR4J and Sacramento simulations, performance is better than all other models at greater separation distances.

5. DISCUSSION OF AWRA MODEL PERFORMANCE

Some AWRA-L parameters are estimated as a function of local conditions (e.g. soil type, rain, vegetation) and the algorithms currently used may contribute to a lower local calibration performance. The parameter values are calculated using relationships derived by van Dijk (2010), from an initial analysis of AWRA v0.5 outputs against various climate/catchment attributes.

Overland runoff (Q_R) is related to net rainfall (P_n : rainfall after initial infiltration and interception loss) via $Q_R = ((1 - f_{sat})P_n / (P_n + P_{ref}) + f_{sat})P_n$, where $f_{sat} = \max(f_{water}, S_G / S_{Gref})$ gives the saturated area fraction and P_{ref} is a ‘reference storm size’, defining the depth of rainfall at which half of it runs off (currently fixed at 250mm for the whole continent). The saturated area fraction is the largest of either the groundwater storage level (S_G) relative to a spatialised reference groundwater level (S_{Gref}), or the fraction of the area covered by water, $f_{water} = \min(0.005, 0.007 \times S_r^{0.75})$, where S_r is the aggregated overland runoff, baseflow and carryover storage. Mean annual precipitation (MAP), is used to spatialise the parameter S_{Gref} via $S_{Gref} = S_{gref_scale} MAP^{S_{gref_shape}}$, where $S_{gref_scale} = 9$ and S_{gref_shape} is a fitted parameter.

Drainage of the soil stores ($D_z = f_D S_z$) is, in unsaturated conditions ($w_z \leq 1$) controlled by the soil storage (S_z) and the drainage fraction, $f_D = K_{FC} \exp[-\beta(1 - w_z)]$, where $\beta = 4.5$ and $w_z = S_z / S_{zFC}$ (S_{zFC} is a fitted spatially constant available water at field capacity). The drainage fraction at field capacity (K_{FC}) is related spatially to average climate wetness $H = MAP / ME0$; $ME0$ is the mean potential evapotranspiration, by $K_{FC} = \min(0.3, K_{FC_scale} H^{K_{FC_shape}})$.

Table 1 Comparison of AWRA-L parameters that are globally and locally calibrated at 3 sites with poor fit to streamflow data. Grey – Unfitted constant parameter. Pink local parameters markedly differ from global ones. Orange local and global spatialised parameters differ significantly.

| Parameter | Process/spatialised parameter | Global | Site/local value | | |
|----------------------|---------------------------------------|--------|------------------|-----------|-----------|
| | | | 312159 | 609017 | 405205 |
| S_{gref_scale} | Overland runoff | 9 | 9 | 9 | 9 |
| S_{gref_shape} | | 1.2 | 45.55 | 45.83 | 29.16 |
| S_{ref} | Local | | 2.70E+31 | 2.90E+16 | 2.30E+17 |
| | Global | | 57.3 | 22.9 | 42.7 |
| $F_{drainFC_scale}$ | Soil drainage capacity | 0.71 | 4.33 | 1.7 | 2.44 |
| $F_{drainFC_shape}$ | | 1.69 | 3.39 | 0.29 | 3.19 |
| K_{FC} | Local | | 42.3 > 0.3 | 1.5 > 0.3 | 6.8 > 0.3 |
| | Global | | 2.2 > 0.3 | 0.3 > 0.3 | 1.2 > 0.3 |
| K_{gw_scale} | Baseflow discharge | 0.01 | 2 | 0.01 | 0.71 |
| K_{gw_shape} | | 0.06 | 0.06 | 0.06 | 0.06 |
| K_{gw} | Local | | 2.08 | 0.01 | 0.73 |
| | Global | | 0.01 | 0.01 | 0.01 |
| K_{rout_scale} | Streamflow routing | 0.1 | 1.1 | 0.11 | 0.1 |
| K_{rout_int} | | 0.05 | 0.5 | 0.05 | 0.3 |
| K_r | Local | | 3.3 | 0.5 | 0.6 |
| | Global | | 0.3 | 0.4 | 0.3 |
| c_{Gsmax1} | Stomatal conductance / photosynthesis | 0.02 | 0.02 | 0.04 | 0.04 |
| c_{Gsmax2} | | 0.02 | 0.02 | 0.04 | 0.04 |
| U_{d01} | Max root water | 6.57 | 6.99 | 5.57 | 5.58 |
| f_{rac_Ud0} | | 0.66 | 0.99 | 0.55 | 0.76 |
| S_{OFC} | Thin soil field cap. | 12 | 9.4 | 13.4 | 10 |
| S_{sFC} | Shallow soil FC | 26 | 58.7 | 13.4 | 41.7 |
| S_{dFC} | Deep soil FC | 1229 | 117 | 1335 | 2083 |
| ER_frac_ref1 | Evap/rain ratio in storms | 0.08 | 0.1 | 0.23 | 0.15 |
| S_{sis1} | Canopy rainfall capacity | 0.04 | 0.04 | 0.8 | 0.03 |
| | | 0.03 | 0.03 | 0.51 | 0.45 |
| $F_{soilEmax1}$ | Relative soil evaporation | 0.34 | 0.2 | 1 | 1 |
| $F_{soilEmax2}$ | | 0.99 | 0.2 | 1 | 0.66 |
| $InitLoss1$ | Initial retention capacity | 23 | 23 | 23 | 23 |
| $InitLoss2$ | | 4.65 | 20.3 | 2.11 | 2.29 |
| Relative bias | Local | | -0.015 | 0.042 | 0.003 |
| | Global | | -0.16 | 1.89 | 0.04 |
| Daily NSE | Local | | 0.76 | 0.28 | 0.61 |
| | Global | | 0.57 | -2.59 | 0 |
| SpringSIM NSE | Local | | 0.81 | 0.94 | 0.81 |
| Rain (mm/d) | Catchment characteristic | | 4.67 | 2.18 | 3.65 |
| PET (mm/d) | | | 2.43 | 3.67 | 2.66 |
| Rain/PET (H) | | | 1.96 | 0.59 | 1.38 |

Groundwater discharge ($Q_g = [1 - \exp(-K_g)]S_g$) is spatialised through a relationship with H , where K_g is given by $K_g = K_{g_scale} H^{K_{g_shape}}$ and **streamflow routing** ($Q_{out} = [1 - \exp(-K_r)]S_r$) through a relationship with $ME0$ where K_r is given by $K_r = K_{r_scale} ME0 + K_{r_int}$.

In Table 1 we present catchment characteristics and calibration statistics and compare locally calibrated AWRA-L parameters to globally calibrated values to observe differences at 3 example sites with a poor fit to stream flow data. A site (609017) in Western Australia was chosen due to this area showing the poorest performance overall, particularly in terms of overestimating runoff. A site in Tasmania (312159) was chosen as, while AWRA performs reasonably in terms of streamflow statistics, the soil stores never reach saturation. Finally a poorly performing site within the Murray-Darling Basin in South Eastern Australia (405205) was chosen. The results indicate that:

- a) The spatial distribution of the K_{FC} (field drainage capacity) parameter is constrained because a limited range of values (0-0.3) is allowed for this parameter. In both global and local calibration, this results in no spatial differentiation between drainage fractions in areas of differing hydraulic conductivity (Table 1 – K_{FC}). Thus inadequate drainage in fast draining areas (sandy areas of Central/ Western Australia) and excessive drainage in slower draining areas (Tasmania) occur as a result of the average fit obtained in global calibration, while in local calibration other parameters of the model get adjusted to provide a better fit to data;
- b) The effect of a single continent-wide value for the Pref parameter (250mm), is to cause the same proportion of infiltration for any given rainfall depth everywhere, irrespective of catchment topsoil layer or slope characteristics; this acts in combination with the groundwater store reference size (S_{gref}), related to MAP, to define overland runoff. The local calibration results in very large groundwater store sizes that indirectly cause reduced overland runoff because of automatically low saturated area proportion. The connection between groundwater store and surface runoff is likely to be inappropriate as it relates a quick surface response to what should be a much slower evolving rate of moisture in the deeper layers of the soil profile. As f_{sat} now tends to be very small, insufficient overland runoff gets generated in locations such as Tasmania because of the fixed value of Pref;
- c) This leads to flow from the groundwater store (baseflow recession) being accelerated in local calibration, indicating that the model is trying to maximise the event flow generation in order to improve the fit to high to intermediate flows (e.g. at Site 312159 where local calibration bias is -16%). Aquifer/soil properties such as clay content are likely to help to capture the spatial variation in K_g , possibly instead of H (wetness) whose influence is largely reduced in the global calibration ($K_{gw_scale} = 0.01$). Local calibration seems to provide a mechanism for faster flow release than baseflow that is missing in the AWRA-L structure (e.g. akin to interflow represented in other models);
- d) Similarly, the approach to spatialising the parameters related to streamflow routing also looks ineffective. The parameters obtained in local calibration (K_{rout_int} at sites 312159 and 405205) cause much faster flow routing than global values, which helps to explain the lower daily NSE results of the global AWRA-L. Further work is needed to link physical properties of the catchment that may have a bearing on the hydraulics of the river system to this parameter;
- e) In Western Australia, the global calibration of AWRA-L tends to have large positive bias (e.g. site 609017, Bias = 189%). In local calibration, the model tries to withhold runoff by increasing store capacities (S_{OFC1} , S_{sls1} , S_{sls2}) and evaporation (ER_frac_refl , $F_{soilEmax1}$, $F_{soilEmax2}$). However, even these adjustments do not result in greatly improving the daily NSE (NSE = 0.28). Potential reasons for this are that the AWRA-L model currently does not have losses/gains from groundwater nor the ability for baseflow thresholding (whereas the other models do). This relatively poor result locally then translates into the global calibration effectively ignoring the area in calibration, as all F values below the 25% do not affect the calibration (hence allowing very large biases in the area).

SpringSIM's performance was better in calibration but poorer than other models in nearest-neighbour validation which indicates that there are features of the model that cannot be readily applied regionally. Two differentiating parameters at the core of the SpringSIM model: the **outlet height (OH)** threshold, and the **starting saturated zone content, (InitGWD)** notionally physical properties of catchments were tested for such behaviour. The effect of using appropriate values for these parameters, is shown in Figure 4 by means of percentile exceedance curves for F scores. The significant improvement in results highlights the importance of appropriately regionalizing the baseflow generation threshold and estimates of the model's starting soil moisture.

6. CONCLUSION

In this paper we show that locally calibrated models of runoff have a strong deterioration in validation performance as the distance to the donor catchment increases - this supports the use of the global calibration approach adopted for AWRA-L. However, AWRA-L performs relatively poorly in local calibration, and this is likely to be explained by deficiencies related to spatialisation of variables controlling overland runoff (Pref, S_{gref}), routing (K_{gw} and K_g) and drainage properties (K_{FC}), and to structural weaknesses in the model. Such weaknesses have been identified and are currently being investigated by the CSIRO development team: the lack of a mechanism to allow sub-surface flux and storage (i.e. when the groundwater store is below river level), linkage of Hortonian runoff to

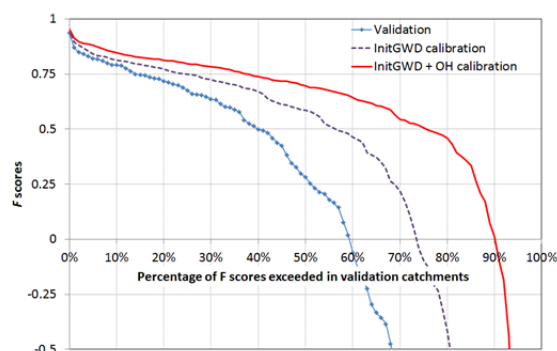


Figure 4 Effect of calibrating outlet height and starting saturated zone moisture on SpringSIM validation results

groundwater contents, lack of interflow generation mechanism. It may also be possible to improve results by ensuring that starting soil moisture states were realistic in validation.

ACKNOWLEDGMENTS

We would like to express our thanks to our Bureau colleagues Carl Daamen, Fangfang Zhao and Sri Srikanthan and the conference reviewers, who valuably reviewed this paper and to the CSIRO AWRA-L team for provision of AWRA-L global calibration results as well as their ongoing efforts to improve the model.

REFERENCES

- Andrews, F.T., Croke, B.F.W. and Jakeman, A.J., 2011. An open software environment for hydrological model assessment and development. *Environmental Modelling & Software*, 26(10): 1171-1185.
- Burnash, R.J., 1995. The NWS river forecast system—catchment modelling. In: V.P. Singh (Editor), *Computer Models of Watershed Hydrology*. Water Resources Publications, Littleton, Colorado, pp. 311-366.
- Duan, Q.Y., Gupta, V.K. and Sorooshian, S., 1993. Shuffled complex evolution approach for effective and efficient global minimization. *Journal of Optimization Theory and Applications*, 76(3): 501-521.
- Gay, M.D., 1990. Usage Summary for Selected Optimization Routines, AT&T Bell Murray Hill, NJ 07974.
- Jones, D.A., Wang, W. and Fawcett, R., 2009. High-quality spatial climate data-sets for Australia. *Australian Meteorological and Oceanographic Journal*, 58: 233-248.
- McVicar, T.R., Van Niel, T.G., Li, L.T., Roderick, M.L., Rayner, D.P., Ricciardulli, L. and Donohue, R.J., 2008. Wind speed climatology and trends for Australia, 1975-2006: Capturing the stilling phenomenon and comparison with near-surface reanalysis output. *Geophysical Research Letters*, 35(L20403).
- Morton, F.I., 1983. Operational estimates of areal evapotranspiration and their significance to the science and practice of hydrology. *Journal of Hydrology*, 66(1-4): 1-76.
- Opie, K., Newnham, G. and Guerschman, J.P., 2011. Comparison of Land Cover Data Products and their Impact on Continental Water Assessments, CSIRO: Water for a Healthy Country National Research Flagship.
- Peeters, L.J.M., Crosbie, R.S., Doble, R.C. and Van Dijk, A.I.J.M., 2013. Conceptual evaluation of continental land-surface model behaviour. *Environmental Modelling & Software*, 43(0): 49-59.
- Perrin, C., Michel, C. and Andréassian, V., 2003. Improvement of a parsimonious model for streamflow simulation. *Journal of Hydrology*, 279(1-4): 275-289.
- Ramchurn, A., 2012. Improved modelling of low flows and drought impacts in Australian catchments using new rainfall-runoff model SpringSIM, 34th Hydrology and Water Resources Symposium. Engineers Australia, Sydney, Australia, pp. 429-440.
- Renzullo, L.J., Collins, D., Perraud, J.-M., Henderson, B., Jin, H. and Smith, A., 2013. Improving soil water representation in the Australian Water Resource Assessment landscape model through the assimilation of remotely-sensed soil moisture products, 20th International Conference on Modelling and Simulation, Adelaide.
- Tolson, B.A. and Shoemaker, C.A., 2007. Dynamically dimensioned search algorithm for computationally efficient watershed model calibration. *Water Resources Research*, 43(1): W01413.
- Van Dijk, A.I.J.M., 2010. The Australian Water Resources Assessment System. Technical Report 3. Landscape Model (version 0.5) Technical Description CSIRO: Water for a Healthy Country National Research Flagship.
- Van Dijk, A.I.J.M., Bacon, D., Barratt, D., Crosbie, R., Daamen, C., Fitch, P., Frost, A.J., Guerschman, J.P., Henderson, B., King, E.A., McVicar, T.R., Renzullo, L.J., Stenson, M.P. and Viney, N., 2011. Design and development of the Australian Water Resources Assessment system, WIRADA Symposium. CSIRO, Melbourne.
- Van Dijk, A.I.J.M. and Renzullo, L.J., 2011. Water resource monitoring systems and the role of satellite observations. *Hydrol. Earth Syst. Sci.*, 15(1): 39-55.
- Viney, N.R., 2010. A comparison of modelling approaches for continental streamflow prediction, CSIRO: Water for a Healthy Country National Research Flagship.
- Viney, N.R., Perraud, J., Vaze, J., Chiew, F.H.S., Post, D.A. and Yang, A., 2009. The usefulness of bias constraints in model calibration for regionalisation to ungauged catchments. In: R.S. Anderssen, R.D. Braddock and L.T.H. Newham (Editors), 18th World IMACS / MODSIM Congress. Modelling and Simulation Society of Australia and New Zealand and International Association for Mathematics and Computers in Simulation, Cairns, Australia, pp. 3421-3427, http://www.mssanz.org.au/modsim09/17/viney_I7a.pdf.
- Viney, N.R., Van Dijk, A.I.J.M. and Vaze, J., 2011. Comparison of models and methods for estimating spatial patterns of streamflow across Australia, WIRADA Science Symposium. CSIRO, Melbourne, Australia, pp. 30-35.
- Viney, N.R., Vaze, J., Wang, B., Zhang, Y., Yang, A., Vleeshouwer, J., Ramchurn, A. and Frost, A., 2013. Comparison of prediction performance of AWRA-L with other models, CSIRO Water for a Healthy Country Flagship, Australia.
- Zhang, Y.Q., Viney, N.R., Chiew, F.H.S., van Dijk, A.I.J.M. and Liu, Y.Y., 2011. Improving hydrological and vegetation modelling using regional model calibration schemes together with remote sensing data, 19th International Congress on Modelling and Simulation, Perth, Australia, pp. 3448-3454.
- Zhang, Y.Q., Viney, N., Frost, A., Oke, A., Brooks, M., Chen, Y. and Campbell, N., 2013. Collation of Australian modeller's streamflow dataset for 780 unregulated Australian catchments. CSIRO: Water for a Healthy Country National Research Flagship, 117pp.