Evaluation of AWRA-L: the Australian Water Resource Assessment model

A.J. Frost a, A. Ramchurn a, M. Hafeez a, F. Zhao a, V. Haverd b, J. Beringer c and P. Briggs b

aThe Bureau of Meteorology, Australia   bCSIRO, Australia   cUniversity of Western Australia

Email: a.frost@bom.gov.au

Abstract: The Australian Water Resource Assessment Modelling System (AWRAMS) provides nationally consistent water balance estimates at a national to regional scale for the past and present using observations where available, and modelling otherwise. The AWRA Landscape model, AWRA-L, is a 0.05° (~5km) gridded soil and groundwater balance model which provides credible estimates of landscape water yield (runoff and baseflow), evapotranspiration, soil moisture, and aquifer recharge across Australia, for retrospective Water Resource Assessment, National Water Account and soil moisture monitoring purposes. This paper evaluates and compares the hydrologic performance of various versions of the AWRA-L model with a national water balance model (WaterDyn) and a global biogeochemical land surface scheme (CABLE), applied regionally. The versions of AWRA-L evaluated are: (a) v4.5 nationally calibrated to streamflow, (b) v5.0 including updated soil drainage properties, nationally calibrated to streamflow and satellite evapotranspiration (CMRSET) and soil moisture (AMSR-E); and (c) v5.0 also calibrated according to eight separate regions to fit local conditions better. The models were compared against catchment streamflow, point estimates of flux tower derived evapotranspiration across Australia, and point estimates of 0-90cm profile soil moisture over the Murrumbidgee and Upper Hunter Catchments (see Figure 1). Satellite derived estimates of evapotranspiration (CMRSET, SLST) and soil moisture (AMSR-E, ASCAT) are also presented for comparison purposes.

Figure 1. Comparison of CABLE, WaterDyn, AWRA-L (various versions) against monthly (a) streamflow (Nash-Sutcliffe efficiency over 291 catchments), (b) flux tower derived estimates of evapotranspiration (correlation), and 0-90cm profile soil moisture (correlation) from the (c) Murrumbidgee and (d) Upper Hunter Catchments.

AWRA-L reproduces streamflow relatively well over the 291 catchments reserved for validation. For ET, CABLE/WaterDyn are better than AWRA-L. AWRA-L/CABLE perform similarly for profile soil moisture, with WaterDyn worse. Regional AWRA-L calibration performs best for streamflow, but worse for soil moisture and ET. Performance of each of the models is explainable according to their design and data used in calibration. AWRA-L v5.0 (using streamflow, ET and soil moisture in calibration) is currently chosen for operational use in the Bureau of Meteorology due to its superior performance across the variables of interest.

Keywords: Water resource assessment, soil moisture, runoff, evapotranspiration, validation
1. **INTRODUCTION**

The Australian Water Resource Assessment Modelling System (AWRMS; see Vaze et al., 2013; Hafeez et al., 2015) is being developed by the Australian Commonwealth Scientific and Industrial Research Organisation (CSIRO) and Bureau of Meteorology (BoM) towards supplying water balance estimates published by the BoM for regular Water Resource Assessment and National Water Account. Further, AWRA-L daily updated gridded outputs – of potential and actual evapotranspiration, soil moisture, and runoff – are available through the Bureau’s website, with additional outputs and derived datasets/maps available to registered users (Smith et al., 2015).

This paper evaluates the AWRA-L model for its ability to reproduce streamflow, evapotranspiration derived from flux tower estimates, and soil moisture estimates from in-situ probes, towards demonstrating the utility of the model for water reporting purposes. Monthly summary statistics are presented and reasons for model performance are explained.

2. **DATA**

2.1. **Streamflow**

A set of 782 unimpaired catchments across Australia were collated Zhang et al. (2013), according to the following criteria: (a) catchment areas is greater than 50 km², (b) the stream is unregulated (no dams or reservoirs), (c) no major impacts of irrigation and land use, (d) observed record has at least 10 years of data between 1975 and 2011. The catchments (delineated using the BoM’s national catchment Geofabric product: http://www.bom.gov.au/water/geofabric/) were collated towards being used in evaluation and calibration of AWRA-L and other models, as these models do not have the ability to model effects of human impact, and therefore use of data from impaired catchments does not provide an accurate test for the models' water balance behaviour. The spatial distribution of catchments reserved for calibration and validation of AWRA-L is shown in Figure 2(a). In testing the models used in this paper all data was used where available covering the period 1/1/1981-30/12/2011 (covering the calibration period). 295 catchments were used in calibration of AWRA-L while 291 catchments not used in calibration are used for validation.

![Figure 2](image)

**Figure 2.** Location of (a) unimpaired catchments used for model evaluation and the eight calibration regions shaded and (b) ET flux towers and Murrumbidgee/Upper Hunter catchments with soil moisture monitoring. Rainfall gradient is also shown to give an indication of the range of conditions sampled within Australia.

2.2. **Evapotranspiration**

Daily evapotranspiration estimates were derived from flux stations from the OzFlux network (http://www.ozflux.org.au; see Figure 2(b) for locations and Table 3 placed at end of paper due to size for details), with latent heat obtained using the DINGO (Dynamic INtegrated Gap filling and partitioning for OzFlux) methodology for processing raw flux tower data (Beringer et al., Biogeosciences Special issue, in prep.). Eddy covariance datasets were quality assured and quality controlled (QA/QC) using the OzFlux standard processing protocol OzFluxQCv2.8.5. Eamus et al., (2013) provide the first summary of the QA/QC processes and corrections involved in the OzFluxQC protocol.
Gaps in missing data were filled using the DINGO processing package. This advanced processing technique ingests OzFlux L3 datasets and applies a linear interpolation to gaps of < 2 hours. For gaps > 2 hours, DINGO searches for the 10 closest Australian Bureau of Meteorology (BoM) monitoring sites from a localised database, selects the BoM site with the best correlation with the flux dataset, and then gap fills using the BoM data. Temperature, humidity, pressure, precipitation and wind speed are all gap filled in this way. Solar radiation is gap filled using gridded AWAP or MODIS data and soil moisture and temperature gaps are filled using CABLE BIOS2 outputs (Haverd et al., 2013). Artificial neural networks (ANN), described in Beringer et al., (2007) are used to gap fill for the fluxes of water and heat along with conduction into the ground. The period 2001-2013 was used for evaluation, being the intersection of years available for all models available, and had a median of 25% months available for the 25 sites tested (after infilling using the DINGO).

**Satellite retrieval based estimates:** of ET covering 2001-2010 were compared to the observed point estimates towards evaluating possible use in AWRA-L calibration. CMRSET satellite ET (CSIRO MODIS reflectance-based Scaling ET; Guerschman et al., 2009) and SLST: CSIRO developed Simplified Land Surface Temperature Algorithm (Van Niel et al., 2012), which specifically deals with scaling up from instantaneous satellite retrievals, to monthly ET quantities.

### 2.3. Soil moisture

**Murrumbidgee:** Time series of volumetric soil moisture at various depths within the soil profile (0-30cm, 30-60cm, 60-90cm) for 38 sites across the Murrumbidgee catchment, NSW (see Figure 2b) were used in evaluation of the models. These time series were derived from reflectometer measurements from the OzNet network, setup and maintained by the University of Melbourne and Monash University (Smith et al., 2012). The reflectometers were calibrated according to independent measurements (Rüdiger et al., 2010), and had a median of 72% of monthly data available over the December 2001 – May 2012 period considered here.

**Upper Hunter:** Time series of water content reflectometer measurements of soil moisture at various depths within the profile (0-5cm, 0-30cm, 30-60cm, 60-90cm) within the Upper Hunter River, NSW (Rüdiger et al., 2007). These time series were collated as part of the Scaling and Assimilation of Soil Moisture and Streamflow (SASMAS) project monitoring sites (managed by the University of Newcastle). There were 13 sites with profile (0-90cm) data available, with a median 88% of monthly data available over the period 2002-2011.

**Satellite retrieval based estimates:** of soil moisture were compared to the point estimates where available. ASCAT is a Technische Universität Wien (TUW) product (Bartalis et al., 2007), active Advanced Scatterometer aboard the MetOp-A satellite covering 1/07/2007-31/12/2011. Also the Vrije Universiteit Amsterdam (VUA)-NASA AMSR-E product (Owe et al., 2008) derived from passive Advanced Microwave Scanning Radiometer for the Earth Observing System aboard the Aqua polar orbiting satellite, covering 20/06/2002-01/10/2011. The methods used to derive satellite data here are further discussed in Renzullo et al (2014).

### 2.4. Models evaluated

**AWRA-L** is a three soil layer (top: 0-10cm, shallow: 10cm-100cm, deep: 100cm-600cm), two hydrological response unit model (shallow rooted versus deep rooted) hydrological model. The two most recent operational versions of AWRA-L are compared: AWRA-L v4.5 (denoted v4.5 in the figures; Viney et al., 2014) and AWRA-L v5.0 (denoted v5QES in the figures; Viney et al., 2015) towards highlighting causes of differences between versions. The models differ in two ways:

a. According to calibration: AWRA-L v4.5 uses calibration to ~300 catchments streamflow records spread across Australia spanning the time period 1981-2001. AWRA-L v5.0 also uses CMRSET satellite ET and AMSR-E satellite soil moisture averaged over the catchments for calibration.

b. According to the input spatial data fields describing the soil saturated conductivity (Ksat): v4.5 had spatial data fields derived from ASRIS (Johnston et al., 2003), v5.0 uses regression functions to estimate spatially smoother Ksat values from soil physical properties.

Three further AWRA-L versions are compared to a) investigate the effect of using different combinations of data within the AWRA-L v5 (v5Q: using streamflow and ET, v5QES: streamflow and soil moisture, and v5Q: streamflow only). V5Q is used to compare to v4.5 to isolate the effect of the change in spatial data fields, using the updated spatial Ksat fields but calibration to streamflow only. Finally a version of AWRA-L v5QES (denoted v5 Reg) is calibrated independently to eight regions across Australia (see Figure 2a) to investigate the possible improvements to parameterization regionally. Models compared, calibration used and model acronyms used are summarized in Table 1.
Table 1. AWRA-L versions compared and calibration data used

<table>
<thead>
<tr>
<th>Model Acronym</th>
<th>Calibrated to</th>
</tr>
</thead>
<tbody>
<tr>
<td>AWRA-L v4.5</td>
<td>Catchment streamflow nationally</td>
</tr>
<tr>
<td>AWRA-L v5Q</td>
<td>Catchment streamflow nationally</td>
</tr>
<tr>
<td>AWRA-L v5QE</td>
<td>Catchment streamflow and satellite evapotranspiration nationally</td>
</tr>
<tr>
<td>AWRA-L v5QS</td>
<td>Catchment streamflow and satellite soil moisture nationally</td>
</tr>
<tr>
<td>AWRA-L v5QES</td>
<td>Catchment streamflow and satellite evapotranspiration and soil moisture nationally</td>
</tr>
<tr>
<td>AWRA-L v5Reg</td>
<td>Catchment streamflow, satellite evapotranspiration and soil moisture over eight regions</td>
</tr>
</tbody>
</table>

Two other national gridded models are also evaluated for comparison purposes:

a. The WaterDyn model, developed by CSIRO Marine and Atmospheric Research (Raupach et al., 2009), as part of the Australian Water Availability Project (AWAP). This is another daily national 0.05° grid-based biophysical model of the water balance between the atmosphere and soil. WaterDyn has two soil layers, and is run using various spatial datasets including thickness of soil and saturated volumetric water content of upper/lower soil layers, while constant saturated hydraulic conductivity values were used nationally.

b. The CABLE-SLI model, The CSIRO Atmosphere Biosphere Land Exchange (CABLE) model, is a community global land-surface model developed by CSIRO, BoM and other universities (Kowalczyk et al., 2006; Wang et al., 2011). This implementation of CABLE includes 10 soil layers (0.022, 0.058, 0.07, 0.15, 0.30, 0.30, 0.30, 1.20, 3.0, and 4.5 m in thickness from the top/near-surface layer to the bottom respectively). Secondly, the default CABLE v1.4 soil and carbon modules were replaced respectively by the SLI soil model and the CASA-CNP biogeochemical model (Wang et al., 2010). Spatially varying soil properties used by BIOS2 are bulk density, clay and silt fractions, saturated hydraulic conductivity, suction at saturation, field capacity, wilting point, and saturated volumetric water content – see Haverd et al (2013) for further details.

Table 2. Summary of AWRA-L, WaterDyn, and CABLE-SLI model characteristics

<table>
<thead>
<tr>
<th>WaterDyn</th>
<th>CABLE-SLI</th>
<th>AWRA-L (v4.5, v5.0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Developer</td>
<td>CSIRO/BoM/ABARES</td>
<td>CSIRO/BoM + universities</td>
</tr>
<tr>
<td>Purpose</td>
<td>Monitoring terrestrial water balance</td>
<td>Land surface scheme for the Australian Community Climate and Earth-System Simulator (ACCESS)</td>
</tr>
<tr>
<td>Soil layers (spatially varying properties)</td>
<td>2 (depth, saturated volumetric water content)</td>
<td>10 (saturated hydraulic conductivity, field capacity, etc)</td>
</tr>
<tr>
<td>Calibration</td>
<td>Parameter calibration and sensitivity analysis to 6 catchments in Murrumbidgee</td>
<td>Calibration to derived ET (50 catchments across 10 climate zones within Australia) and flux tower data</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Streamflow over ~300 catchments (v4.5) and satellite soil moisture and ET (v5.0)</td>
</tr>
</tbody>
</table>

All models were forced using daily gridded AWAP climate data, with all model simulations covering at least 1950 until 2013. Model soil layers were weighted according to the fraction of overlap they have with the observations they are being compared with (0-90cm for profile).
3. DISCUSSION AND CONCLUSIONS

3.1. Performance of AWRA-L compared to WaterDyn and CABLE

Fig. 1(a) shows that the various versions of AWRA-L produce higher monthly Nash-Sutcliffe Efficiency (NSE) over the 291 validation catchment than WaterDyn with CABLE performing poorest according to streamflow. This is due at least in part to a) AWRA-L being calibrated to streamflow nationally (while WaterDyn and CABLE were not); and b) WaterDyn and CABLE not being designed towards reproduction of streamflow (e.g. simple thresholding mechanism for runoff generation in WaterDyn, and application of Richards equation for drainage without subsequent runoff dynamics for CABLE), whereas AWRA-L is specifically designed to capture runoff dynamics (e.g. mechanisms for interflow and baseflow). Clear comparison of model components is not possible due to differing datasets being used in calibration however. In terms of reproduction of ET at the flux tower locations according to monthly correlation (Fig. 1(b)), WaterDyn and CABLE outperform AWRA-L. This is partly due to CABLE and WaterDyn being designed for reproduction of land surface energy fluxes, and partly due to calibration to flux tower data in the case of CABLE along with smaller time step and greater number of soil layers. The satellite derived CMRSET produces higher correlation than SLST and the other model estimates, noting that part of the reason for its relatively high performance is shorter coverage over the evaluation period compared to the other models (i.e. not covering 2012/13), and also due to calibration of CMRSET to a subset of the OzFlux data. For correlation of modelled soil moisture compared to profile 0-90 cm soil moisture (Fig. 1(c) and (d)), CABLE and AWRA-L perform best over the Murrumbidgee and Hunter sampling regions respectively, although results are similar. WaterDyn performs poorer than CABLE/AWRA-L, particularly for the Hunter sampling area. For the satellite derived estimates, ASCAT performs better than AMSR-E in terms of reproduction of profile soil moisture (noting ASCAT has 5 years less data). The satellite estimates are worse than the models in terms of correlation with the deeper soil moisture probes as satellite based estimates only correspond to the surface soil moisture (~5 cm depth). The top layer (0-5 cm) soil moisture reproduction was reviewed (Fig. 3), and CABLE and WaterDyn performed best. Overall, the results show the different models have different strengths and weaknesses depending on model structure and calibration, with AWRA being strong in terms of streamflow, soil moisture and ET, CABLE strong in terms of ET and soil moisture, and WaterDyn strong in terms of ET and better than CABLE for streamflow. What is striking about this assessment is the complementarity of the three models for the prediction of streamflow, soil moisture and evapotranspiration. The above suggests that an improved model for landscape water balance may be achieved by merging functionality from the various models.

Figure 3. Top layer (0-5 cm) soil moisture monthly correlation for the (a) Murrumbidgee and (b) Hunter sites.

3.2. Performance of various versions of AWRA-L

Streamflow reproduction in AWRA-L has improved from v4.5 to v5.0 (Fig. 1(a)) for the continental calibrations (due to the altered soil layers), although the cases where satellite ET or soil moisture are used in calibration reduces that performance marginally. In the case where regions are used for calibration, a similar increase in performance is observed. ET performance (Fig. 1(b)) has also improved significantly moving from v4.5 to v5.0, with the best performance being when streamflow and AMSR-E soil moisture are used (and not CMRSET ET). The regional calibration is significantly worse according to ET, due to a higher focus being placed on streamflow for the regional calibrations. Profile soil moisture performance (Fig. 1(c) and (d)) decreases for v5 in the Murrumbidgee, but increases in the Hunter, compared to v4.5 (reflecting the influence of the differing soil drainage properties being used). For v5.0, profile soil moisture reproduction is marginally best where ET was used in calibration over the Murrumbidgee. The regional based calibration (using streamflow, ET and soil moisture) performs similarly to the other AWRA-L v5 calibrations for profile soil moisture. v5QS had the highest performance for top layer soil moisture (Fig. 3) of the AWRA continentally calibrated models, while those using ET dropped the performance marginally. Overall, across all of the variables tested, as v5QES had comparable performance to that calibrated using only streamflow, and
marginally better performance according to soil moisture and ET overall, it was chosen for use operationally in v5.0.

ACKNOWLEDGMENTS

The AWRA modelling system was developed and funded under the Water Information Research and Development Alliance (WIRADA) agreement between CSIRO and the BoM. The authors acknowledge the great amount of effort put into the AWRA system research and implementation by many CSIRO and BoM staff over the last eight years; Professor Jeff Walker and Dr. Dongryeol Ryu for discussions regarding/supply of the OzNet dataset; and Dr Tony Wells for supply of the SASMAS dataset.

Table 3. OzFlux ET data used (with attribution to data source noted)
REFERENCES


