Australian Government
Bureau of Meteorology

# Evaluation of the Australian Landscape Water Balance model (AWRA-L v7)

A comparison of AWRA-L v7 against Observed Hydrological Data and Peer Models

Citation: Frost, A. J., Shokri, A., Keir, G., Bahramian, K., Azarnivand, A. (2021). Evaluation of the Australian Landscape Water Balance model: AWRA-L v7. Bureau of Meteorology Technical Report.

| Version number/type | Date of issue |
|---|---|
| 1.0.10 | 30/08/2021 |



Published by the Bureau of Meteorology

Contact details

Dr. Andrew Frost
a/Team Lead, Hydro-informatics & Surface observations
Research Program, Science and Innovation Group
Bureau of Meteorology
PO Box 413, Darlinghurst NSW 1300
Tel: +61 2 9296 1517
Email: Andrew.Frost@bom.gov.au

Dr. Ashkan Shokri
Senior Hydrologist, Hydro-informatics & Surface observations
Research Program, Science and Innovation Group
Bureau of Meteorology
PO Box 1289, Melbourne Vic 3001
Tel: +61 3 9669 4303
Email: Ashkan.Shokri@bom.gov.au

Water Sector email
Water Sector Program, Business Solutions Group

Email: water@bom.gov.au

## Summary

This technical report details the scientific evaluation of the Bureau of Meteorology (hereafter called the Bureau) operational Australian Water Resources Assessment Landscape (AWRA-L version 7) modelling system. The evaluation used a range of the best available measurements/estimates of hydrological variables available nationally, including streamflow, soil moisture, actual evapotranspiration (ET), vegetation cover, terrestrial water storage, and groundwater recharge. In addition, the performance of the operational AWRA-L version 7 model (hereafter called AWRA-L v7) is compared to the previous AWRA-L versions (v5 and v6), and two other national, gridded, land-surface models i.e. CABLE-SLI and WaterDyn. Runoff simulated by AWRA-L v7 is also compared with simulated streamflow from individual conceptual rainfall-runoff models which were calibrated using streamflow observations for gauged catchments and nearest neighbour regionalisation for predictions in ungauged basins.

AWRA-L and the peer models are assessed and compared according to various performance statistics for each set of evaluation data. Select key indicators of AWRA-L model performance are provided. These benchmark statistics provide a baseline against which model improvements can be compared using the same comparison data. Aspirational targets for overall performance are also provided. In addition to the scientific evaluation against observed hydrological datasets and other peer models, annual national maps, and monthly catchment time-series of the outputs from AWRA-L are presented and compared to available observations to demonstrate and explain local model performance.

The results show that AWRA-L v7 performs relatively well according to streamflow nationally (295 unimpaired catchments used in calibration, 291 separate catchments used in validation) in comparison to WaterDyn and CABLE, reflecting that AWRA-L is calibrated to streamflow. Comparison to locally calibrated, nearest neighbour regionalised rainfall-runoff models show that while AWRA-L calibrated nationally does not perform as well in calibration (as there is one set of parameters applying nationally), performance for ungauged basins approaches that of the locally calibrated models; giving confidence in the use of AWRA-L nationally for runoff prediction; along with other components of the water balance.

AWRA-L improves significantly in this version compared to v5 and v6, predominantly due to a conceptual change introduced relating to baseflow, causing reduced bias even though streamflow is not weighted as high in calibration compared to previous versions. Performance according to streamflow also improves over catchments with significant impervious areas (according to testing over 13 additional catchments containing greater than 5% impervious area), due to introduction of a new Impervious hydrological response unit in addition to the existing shallow (grass) and deep (trees) rooted hydrological response units.

AWRA-L v7 also performs well according to probe-based point measurements of root zone (0-90cm) soil moisture from 51 sites in South Eastern Australia (SASMAS and OzNet networks). 30 further soil moisture sites are added to this analysis from the OzFlux and

CosmOz networks, demonstrating the good performance according to profile soil moisture nationally. AWRA-L v7 has improved top layer soil moisture (0-10 cm) according to median daily correlation compared to v6, while not at the expense of root zone performance.

AWRA-L v7 performs better for evapotranspiration on a monthly scale than v5 and v6 according to correlation with 25 flux tower measurements nationally and considerably better according to median correlation on a daily scale. The main reason for the improvement was the fixed alignment of the input data (e.g. maximum temperature and solar radiation) and better top layer soil moisture estimates.

When comparing AWRA-L v7 against peer models, CABLE is similar in terms of soil moisture performance. Even though AWRA-L v7 ET improves over v5 and v6 for daily correlation, CABLE performs marginally better for monthly correlation. This is not unexpected as CABLE was developed primarily as a land/atmosphere exchange model, which was calibrated to flux tower and derived catchment ET. WaterDyn also performs slightly better than AWRA-L for ET, but both these peer models perform worse for streamflow and root zone soil moisture.

AWRA-L v7 shows considerably better performance for deep drainage over v6 and v5 in terms of correlation with the annual time-series recharge dataset located in South Australia, while there is consistent performance against a national collated average annual recharge dataset. It is noted that modelled drainage does not follow rainfall variation patterns enough yet in v7, with recharge biases influenced strongly by saturated hydraulic conductivity.

Two new observations are added to the calibration objective function in AWRA-L v7, including vegetation fraction (derived from MODIS) and Terrestrial Water Storage (from GRACE). AWRA-L v7 outperforms v5 and v6 for both Vegetation Fraction (Fveg) and Terrestrial Water Storage (TWS) when compared over the validation catchments. Adding these two new observations to the calibration process enabled parameters related to previously unobserved parts of the model be tuned (reflecting the entire water balance and vegetation dynamics). Notably, although the calibration objective weights streamflow lower, the overall performance is better across the water balance and particularly for streamflow. This reflects the valuable contribution of the new observations in constraining the water balance coupled with appropriate changes to the model structure.

AWRA-L v7 outperforms previous versions for almost all metrics. This improvement is a result of the recent updates in inputs, model structure and calibration process as detailed in a companion model description report (Frost and Shokri, 2021). Overall, the water balance verification statistics give confidence in the use of AWRA-L for water resources monitoring and assessment nationally.

## Table of Contents

## List of Tables

## List of Figures

## List of Acronyms

ACCESS: Australian Community Climate and Earth System Simulator climate model

AMSR-E: Advanced Microwave Scanning Radiometer for the Earth Observing System

ASCAT: Advanced Scatterometer aboard the MetOp-A satellite

ASRIS: Australian Soil Resource Information System

AWAP: Australian Water Availability Project

AWRA-L: Australian Water Resources Assessment Landscape Model

AWRA-R: Australian Water Resources Assessment River Model

AWRAMS: Australian Water Resource Assessment modelling system

BoM: Bureau of Meteorology

CABLE: Community Atmosphere Biosphere Land Exchange model

CMRSET: CSIRO MODIS reflectance-based Scaling ET

CSIRO: Commonwealth Scientific and Industrial Research Organisation

DINGO: Dynamic INtegrated Gap filling and partitioning for OzFlux

ET: Evapotranspiration

Fveg: Fraction of vegetation cover

GRACE: Gravity Recovery and Climate Experiment

LAI: Leaf Area Index

NWA: National Water Account

MODIS: Moderate Resolution Imaging Spectroradiometer

RWI: Regional Water Information

SASMAS: Scaling and Assimilation of Soil Moisture and Streamflow

SLI: Soil-Litter-Iso

TWS: Terrestrial Water Storage

WIA: Water in Australia

WIRADA: Water Information Research and Development Alliance

# 1 Introduction

Prolonged extreme drought and resulting water shortages within Australia during the 'Millennium drought', over the period 1997 to 2009, led to the implementation of the federally mandated Water Act (2007) towards better monitoring of water availability and water use nationwide. As a result, the Australian Bureau of Meteorology (the Bureau) was given additional responsibilities including collating water data from jurisdictional agencies and analysing and reporting on water status, in addition to its existing weather and flood forecasting responsibilities.

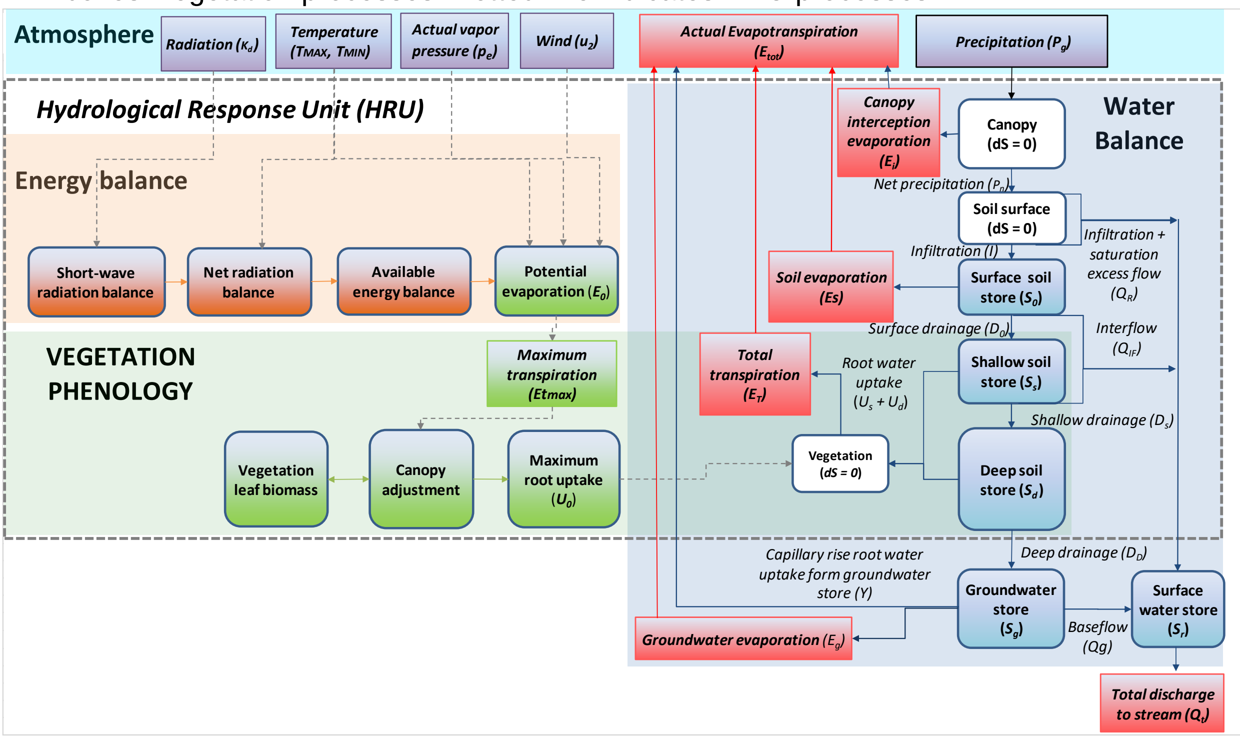
The Australian Water Resources Assessment (AWRA) Modelling System underpins the Bureau's water information services for national water resource assessment reporting, water use accounting and situation monitoring. The modelling system has been developed by the Bureau and CSIRO over the last decade and is run operationally within the Bureau to provide both situational awareness and national retrospective water resource assessment.

The AWRA-L (landscape) model (schematically shown by Figure 1) runs on a daily timestep and 0.05° grid (approximately 5 km) simulating the landscape water balance for Australia from 1911 to yesterday. Key outputs from the AWRA-L model include surface runoff, soil moisture, evapotranspiration, and deep drainage. Outputs from the model are available through the website interface (http://www.bom.gov.au/water/landscape) or by request as a registered user.

AWRA-L is optimised to the whole water balance using a national streamflow dataset along with satellite derived soil moisture (ASCAT), evapotranspiration estimates (CMRSET), Terrestrial Water Storage (TWS; GRACE) and vegetation coverage (MODIS). The model is validated against a wide range of observational datasets including point scale soil moisture probe data, flux tower estimates (of evapotranspiration and soil moisture) and groundwater recharge estimates. The modelling system has recently been released as a community modelling system (https://github.com/awracms/awra_cms), enabling application and development by the wider research community.

AWRA-L model development was initiated through CSIRO in 2010, with the initial model AWRA-L v0.5 developed by van Dijk (2010). Operational modelled outputs have been made publicly available by the Bureau since November 2015 (using AWRA-L version 5; see Frost et al., 2016; Viney et al., 2015), and the modelled fluxes have been used internally and externally for various climatological, flood, water and agriculture applications across Australia. In November 2018 AWRA-L version 6 was released (Frost et al., 2018), showing improvement across the water balance. This report documents the evaluation of AWRA-L v7 released in 2021, with a companion report describing the model in detail (Frost and Shokri, 2021).

Figure 1. AWRA-L conceptual structure. Purple: climate inputs; Blue rounded boxes: water stores; Red boxes: water flux outputs; Brown: energy balance; Green rounded boxes: vegetation processes. Dotted line indicates HRU processes.




AWRA-L v7 contains changes compared to AWRA-L v6 in relation to a) dynamic climate inputs and static spatial inputs b) structural changes including incorporation of impervious hydrological response unit, baseflow mechanism and top layer soil drainage equations, and c) calibration method, data and objective function. These changes were incorporated on the basis of evaluation across the water balance described here.

This report evaluates and compares the hydrologic performance of the AWRA-L v5, v6 and v7 models, a national water balance model (WaterDyn) and a global biogeochemical land surface scheme (CABLE), applied regionally. AWRA-L, WaterDyn and CABLE were also compared to conceptual rainfall runoff models to see how they perform relatively for streamflow (given a range of outputs are provided by the national models).

These models were compared against catchment streamflow, point estimates of flux tower derived evapotranspiration across Australia, point estimates of profile soil moisture (0-90 cm) derived from OzNet and SASMAS over the Murrumbidgee and Upper Hunter Catchments, and OzFlux and CosmOz networks across Australia. Satellite derived estimates of evapotranspiration (CMRSET, SLST) and soil moisture (AMSR-E, ASCAT) are also compared to the point-based observations. The models are also compared against satellite derived estimates of TWS and vegetation fraction.

The three models were further compared to a collated national long-term average recharge dataset and a set of annual recharge time-series data within South Australia.

This report is structured as follows:

- Chapter 2: Forcing and evaluation data
- Chapter 3: Model descriptions
- Chapter 4: Evaluation approach
- Chapter 5: Evaluation according to data
- Chapter 6: Evaluation for reporting purposes
- Chapter 7: Conclusions

# 2 Data

Various organisations provide monitoring of rainfall and streamflow across the nation; albeit coverage is sparse in some areas (e.g. arid interior of Western Australia). Monitoring of other hydrological fluxes and stores (e.g. soil moisture, ET, deep drainage) is less prevalent. However, since 2000 many universities and research groups have established an excellent ground-based network for physical measurement of hydrological fluxes in select catchments and locations around Australia. The Bureau has utilised these hydrological catchment-based and point-based datasets for scientific evaluation of AWRA-L model performance in this report.

## 2.1 Climate forcing data

All models were forced using the daily gridded Australian Water Availability Project (AWAP) climate data set that consists of air temperature (daily minimum and maximum) and daily precipitation from January 1$^{st}$, 1911 to yesterday (Jones et al., 2009). The climate data is interpolated from station records and provided on a 0.05° (approximately 5 km) grid across Australia. AWAP Actual vapour pressure has been added to the temporally dynamic inputs in AWRA-L v7, replacing an internal calculation based on minimum daily temperature. Additionally, daily solar exposure (downward shortwave radiation) is produced from geostationary satellites (Grant et al., 2008) and aggregated to the same 0.05° AWAP grid. The solar radiation record is from 1990 to yesterday, with the Himawari-8 satellite used since 23rd March 2016. Prior to that date the GMS-4, GMS-5, GOES-9 and MTSAT-1R satellites were used. All model simulations cover at least the period of 1950 until 2017. It is noted that as the aforementioned daily climate data are available in different 24 hour time frames (i.e. 9am to 9am for rainfall data, 12am to 12am for solar radiation measurements), the data is assigned to the most relevant day. For more details please see section 1.4.1 of description report (Frost and Shokri, 2021)

## 2.2 Evaluation data

### 2.2.1 Streamflow

A set of 782 unimpaired catchments with gauged flow records of reasonable length across Australia were collated by Zhang et al. (2013) according to the following criteria: (a) catchment area is greater than 50 km$^2$ (due to the ~5 x 5 km grid scale of input climate data, as 'it is considered difficult to adequately characterise the catchment rainfall in catchments smaller than about 50 km$^2$ using such coarse gridded data'), (b) the stream is unregulated (no dams or reservoirs), (c) no major impacts of irrigation or land use change, (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: www.bom.gov.au/water/geofabric were split for use in calibration and validation of

AWRA-L (see Zhang et al., 2013 section 5.3). The following catchments were excluded from the initial dataset: (a) if greater than 5000 km$^2$ as there is currently no streamflow routing processes in AWRA-L and (b) if nested to ensure independence of records. The spatial distribution of remaining 586 catchments reserved for calibration and validation of AWRA-L is shown in Figure 2; with regional divisions showing areas of similar climate. It is noted that AWRA-L runoff is summed according to weighting the cells that intersect the catchment area for comparison to streamflow, as there is no routing in AWRA currently.

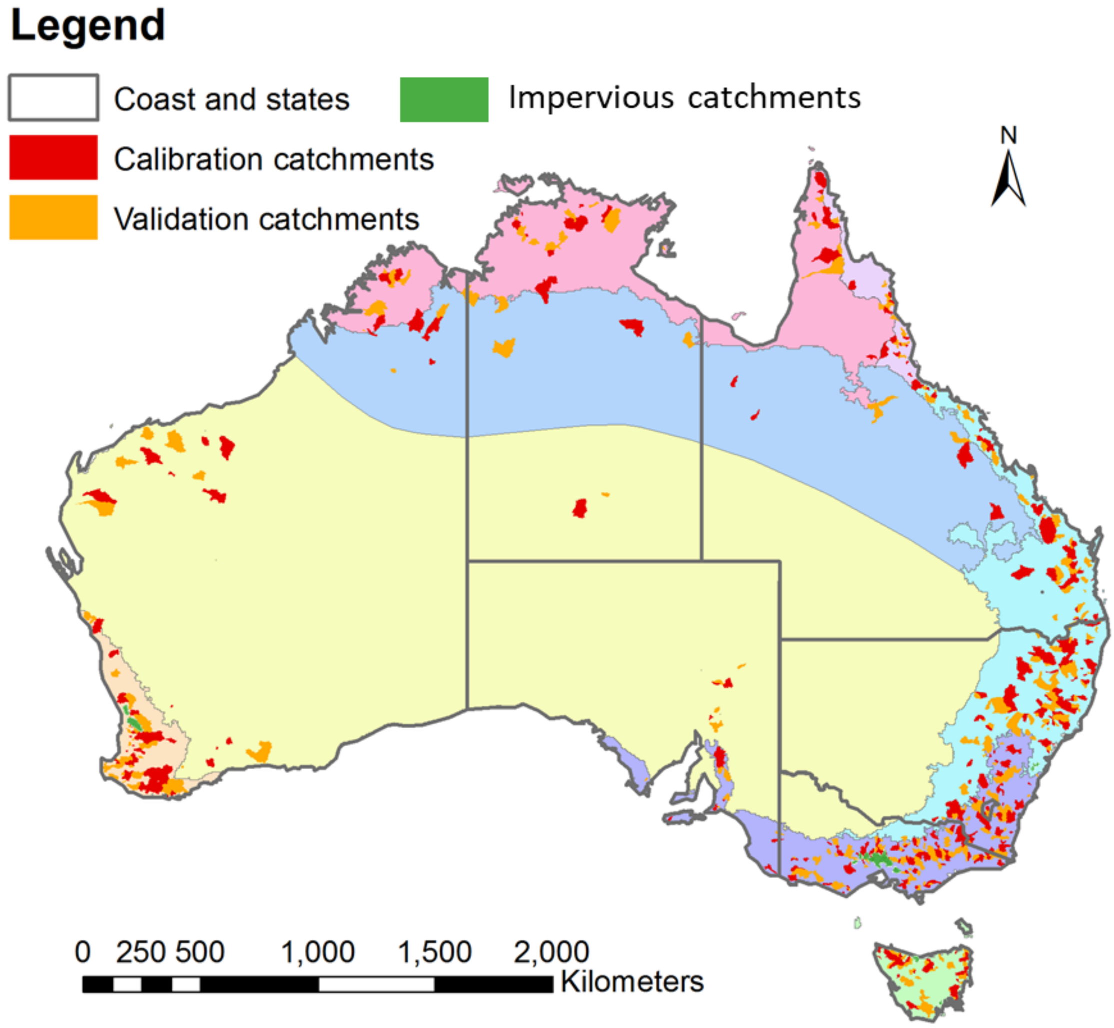



**Figure 2. Location of unimpaired catchments used for model evaluation**

Data from 295 catchments covering the period 1/1/1981-30/12/2011 were used in calibration of AWRA-L while 291 catchments not used in calibration were used for validation. The calibration and validation catchments and period used for evaluation were kept consistent with previously AWRA-L versions for comparability. The spatial delineation of catchments was updated according to Geofabric v3.2 beta within AWRA-L v7, based on more fine scale digital elevation model than used previously. As a result, the catchment areas are more accurate. Further, the 9am-9am flow data was also updated from the Bureau's Water Data online service (see http://www.bom.gov.au/waterdata/). This results in a more accurate and up-to-date record for calibration and evaluation.

A new Hydrologic Response Unit has been added in AWRA-L v7 to represent impervious areas. The calibration and validation set described above excludes sites with more than 10% of their area identified as intensive land use (eg. mining, intensive horticulture, farming) including urban land use. Therefore, to test the performance in urban areas of the model, a set of 13 additional catchments with significant impervious area (minimum 5% and on average 20% of the catchment area according to HRU fractions) is added for evaluation of AWRA, with these sites shown in green in Figure 2 and listed in Table 23 within Appendix E. Select impervious catchments are shown in Figure 3 with satellite imagery underlain to show urban coverage. Hereafter in this report these catchments including impervious area are referred as the impervious catchments.

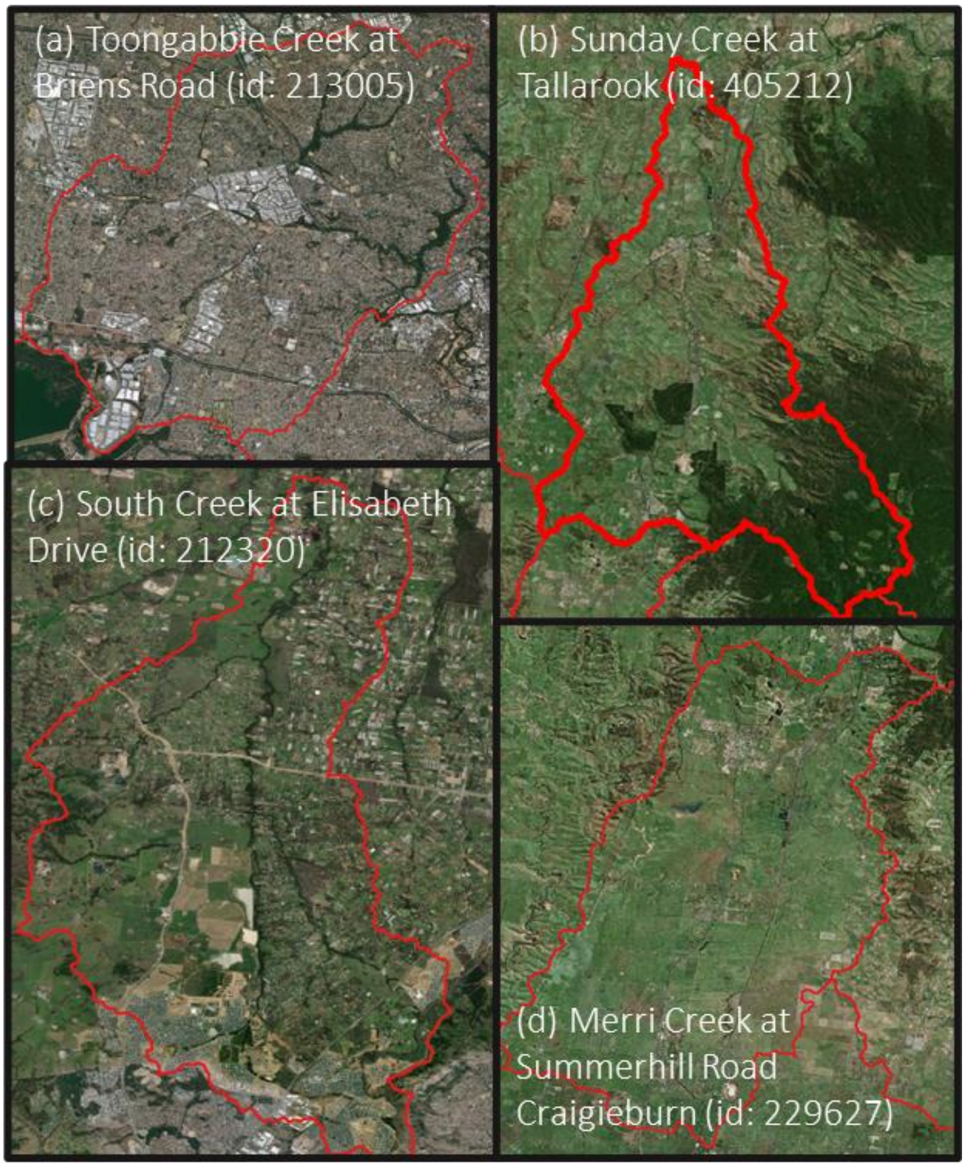



**Figure 3. Location of selected impervious catchments adjacent to major cities including (a) Toongabbie Creek at Briens Road (86.35% impervious coverage) and (c) South Creek at Elisabeth Drive (10.30% impervious coverage) close to Sydney as well as (b) Sunday Creek at Tallarook (8.87% impervious coverage) and (d) Merri Creek at Summerhill Road Craigieburn (16.52% impervious coverage) close to Melbourne.**

### 2.2.2 Soil moisture

Soil moisture data sets from a range of networks distributed nationally (see Figure 4) have been used for evaluation of the modelled outputs, as detailed below.

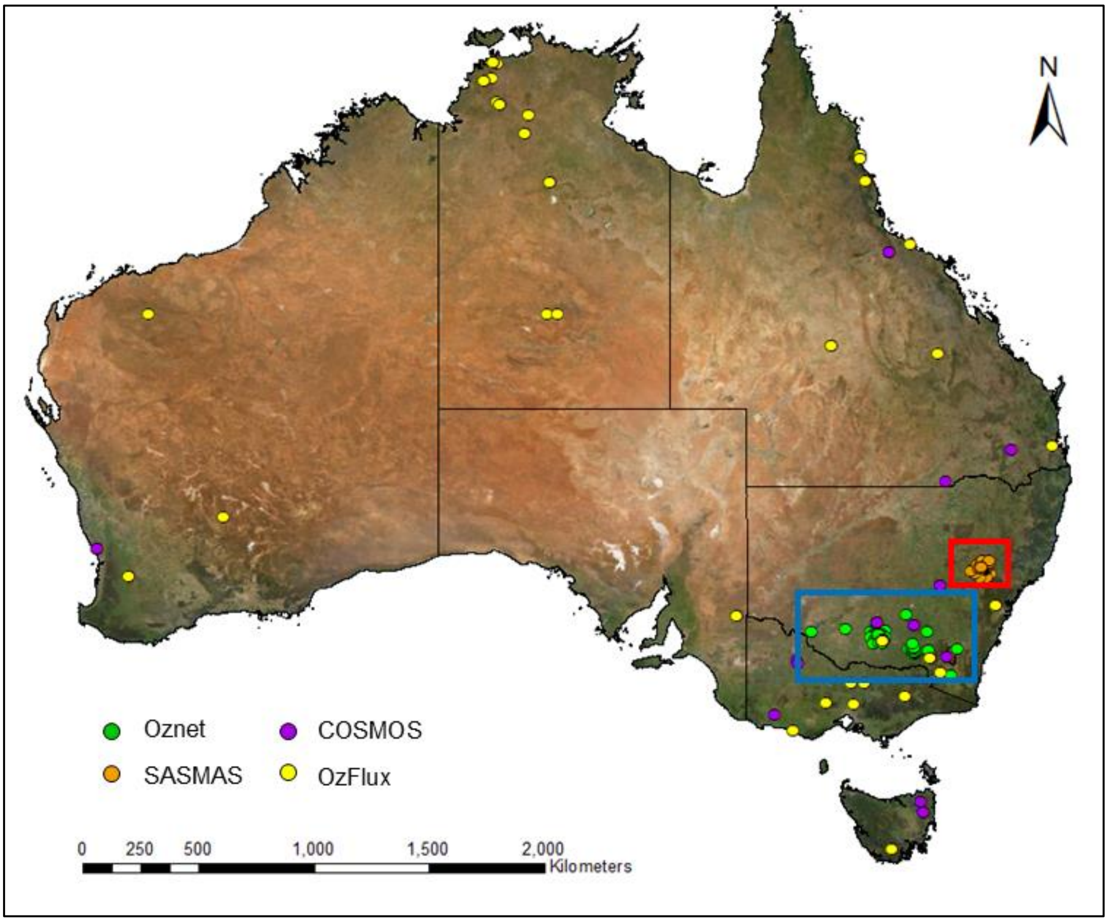
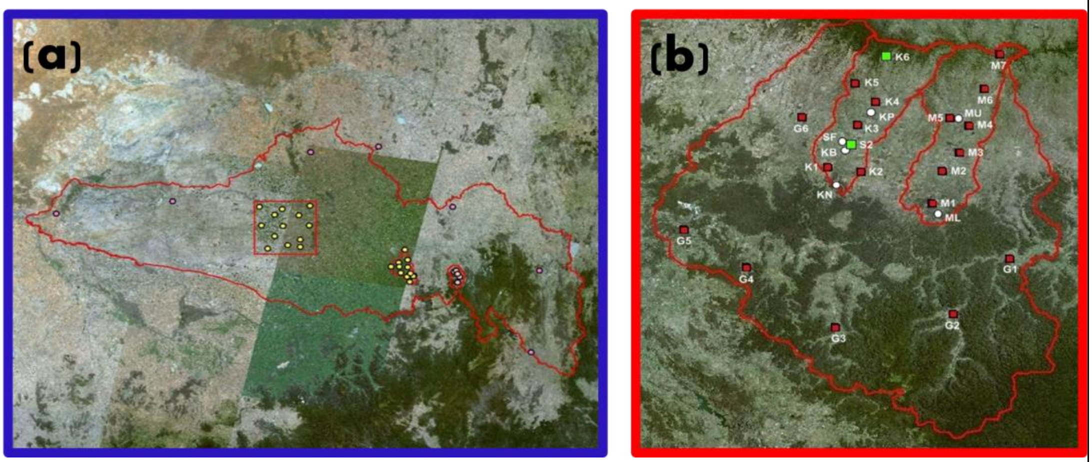



**Figure 4. Spatial distribution of location of available data sets used for evaluation of soil moisture across Australia. The blue and red rectangles are the location of OzNet (Murrumbidgee) and SASMAS (Upper Hunter) networks.**

**OzNet network:** Timeseries of volumetric soil moisture at various depths within the soil profile (0-5cm/8cm, 0-30cm, 30-60cm, 60-90cm) for 38 sites across the Murrumbidgee catchment in NSW (see Figure 4 and for more details Figure 5(a) and Appendix C) were used only in evaluation (and not calibration) of the models. These timeseries 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 67% of monthly data available over the 2001 – 2013 period considered. For soil moisture evaluation, 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).




www.oznet.org.au/murrumbidgeesm.html
www.eng.newcastle.edu.au/sasmas/SASMAS/sasdata.html

**Figure 5. (a) OzNet Murrumbidgee soil moisture (b) SASMAS Goulburn soil moisture**

**Scaling and Assimilation of Soil Moisture and Streamflow (SASMAS) Network:** 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 in NSW (Rüdiger et al., 2007) were used just for evaluation (and not calibration) of the models; see Figure 4, Figure 5(b) and Appendix C and for more details. These time series were collated as part of the SASMAS project monitoring sites (managed by the University of Newcastle). There were 13 active sites with profile (0-90cm) data available and with a median 75% of monthly data available over the period 2003-2011.

**ASCAT and AMSR-E satellite based gridded estimates:** of soil moisture along with modelled estimates were compared to the point probe based estimates where available (surface soil moisture) to determine their value for evaluation, AWRA-L calibration and as a potential rival modelled product or for eventual assimilation into AWRA-L. ASCAT is a Technische Universitat Wien (TUW) product (Bartalis et al., 2007) with an active Advanced Scatterometer aboard the MetOp-A satellite covering 1/07/2007-31/12/2011. The Vrije Universiteit Amsterdam (VUA)-NASA AMSR-E product (Owe et al., 2008) is derived from passive Advanced Microwave Scanning Radiometer for the Earth Observing System aboard the Aqua polar orbiting satellite and covers a nine-year period between 2002-2011. The methods used to derive satellite data here are further discussed in Renzullo et al. (2014). It is noted that catchment averages of AMSR-E soil moisture have been used for calibration of the AWRA-L model covering the same time period – see section 3.1.

**OzFlux:** Time-series of OzFlux (http://www.ozflux.org.au/monitoringsites/index.html; Beringer, 2016) probe-based measurements of soil moisture at various depths was used in evaluation of the model. Profile and top layer soil moisture data are derived respectively from 14 and 20 active stations, for the period of 2007-2017 (with 2017 end date used here for consistency with previous AWRA v5 and v6 comparison). Since the measurement depth is variable across sites, the soil moisture is estimated for 0-30cm, 30-60cm and 60-90cm layers of soil profile using linear weighting (see Appendix A for further details).

**CosmOz:** The soil moisture measurements derived from the 16 Australian Cosmic Ray Soil Moisture Monitoring Network (CosmOz; http://cosmoz.csiro.au/) sensors (see Figure 4 and Appendix C) were used as a benchmark to evaluate performance of models. This network uses cosmic rays originating from outer space to measure average soil moisture over an area of about 40 hectares to a depth up to 90 cm. Given that the CosmOz dataset is available on an hourly basis, the closest measurement of soil moisture at 10am is assumed to be representative of the 9am daily soil moisture.

CosmOz provides measurements that vary in depth depending on the wetness of the soil and is typically in the range 10 to 30cm. As the measurement depth varies, and does not usually cover depths below 1m, the top (0-10cm) and shallow soil layer (10cm-100cm) in AWRA are weighted linearly according to be equivalent to the CosmOz effective depth (see Appendix A for further details). That weighted AWRA value is then compared to the CosmOz through correlation statistics for period 2010-2017 (with June 2017 end date used here for consistency with previous AWRA v5 and v6 comparison).

### 2.2.3 Terrestrial Water Storage (TWS)

Terrestrial Water Storage (TWS) is defined as the entire water stored on the land surface and in the subsurface. It includes surface and root zone soil moisture, groundwater, snow, ice, water stored in the vegetation, river, and lake water. GRACE mission provides monthly gravity field solutions for characterising TWS anomalies across the world. A mass concentration (mascon) solution of GRACE TWS (RL05 V1.0) delivered by the University of Texas Center for Space Research (CSR) (Save et al., 2016) was used for calibration and evaluation of the AWRA-L model (http://www2.csr.utexas.edu/grace). The product has a monthly temporal and ~300km spatial resolution. TWS is included in the AWRA-L analysis for the first time in with the evaluation of AWRA-L v7. As GRACE represents water balance anomaly values (in terms of mm), AWRA-L is normalised by the mean storage for comparison purposes.

For calibration and validation catchments comparison, a weighting function based on catchment area overlap with GRACE pixel is used here to derive nominal catchment total water storage. Further, the influence of leakage and highly uncertainty data on coastal areas is dealt with using reduced weighting when a GRACE pixel has area covered by the Ocean (see Figure 6).

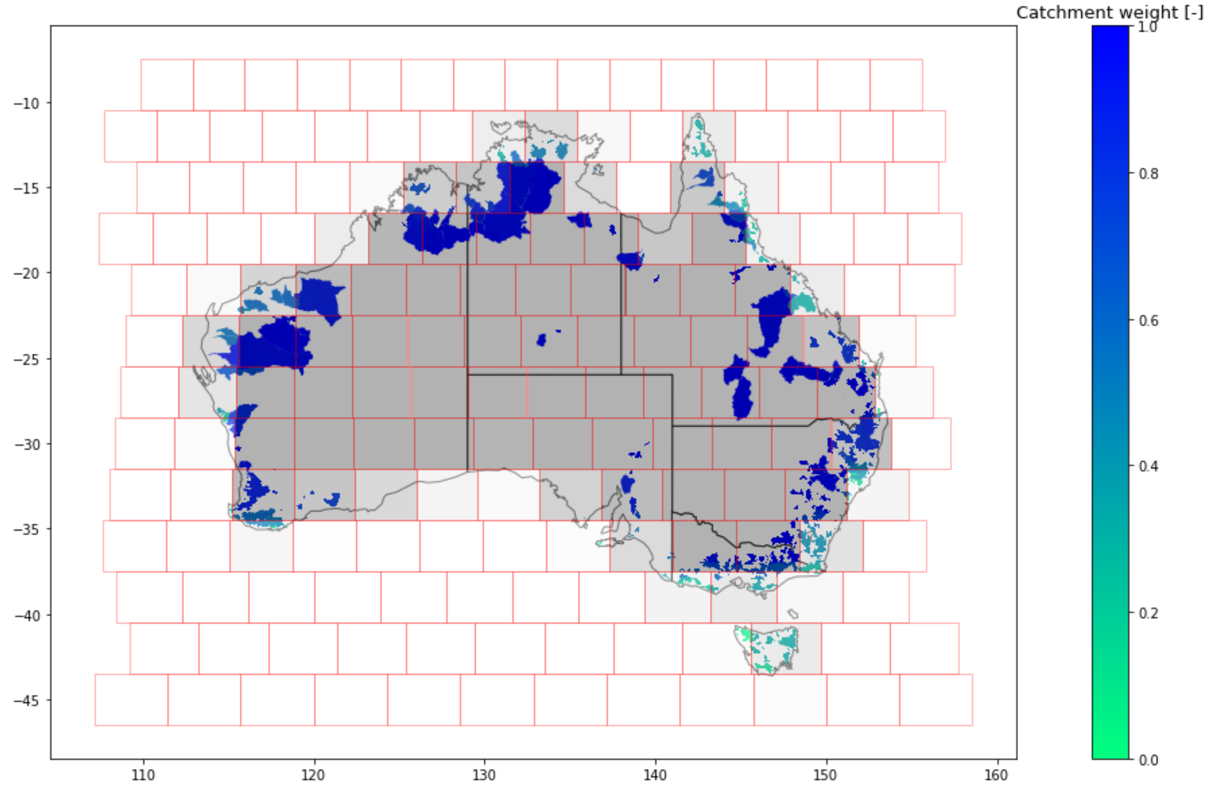



**Figure 6. The GRACE pixel location (red boxes) and the weight of catchments. The weight of each catchment is proportional to the land fraction of the underlying GRACE pixel.**

### 2.2.4 Actual Evapotranspiration

The following data sets were used for evaluation of the modelled outputs:

**OzFlux Network:** Daily evapotranspiration estimates were derived from flux stations from the OzFlux network (www.ozflux.org.au; Beringer et al., 2016; see Figure 7 for locations and Appendix B: ET monitoring site details) with average annual rainfall overlain to give an indication of the variety of climate areas sampled. Latent heat was obtained using the DINGO (Dynamic INtegrated Gap filling and partitioning for OzFlux) methodology for processing raw flux tower data (Beringer et al., 2017). Eddy covariance datasets were quality assured and quality controlled (QA/QC) using the OzFlux standard processing protocol OzFluxQCv2.8.5. The QA/QC processes and corrections involved in the OzFluxQC protocol are described in Eamus et al. (2013).

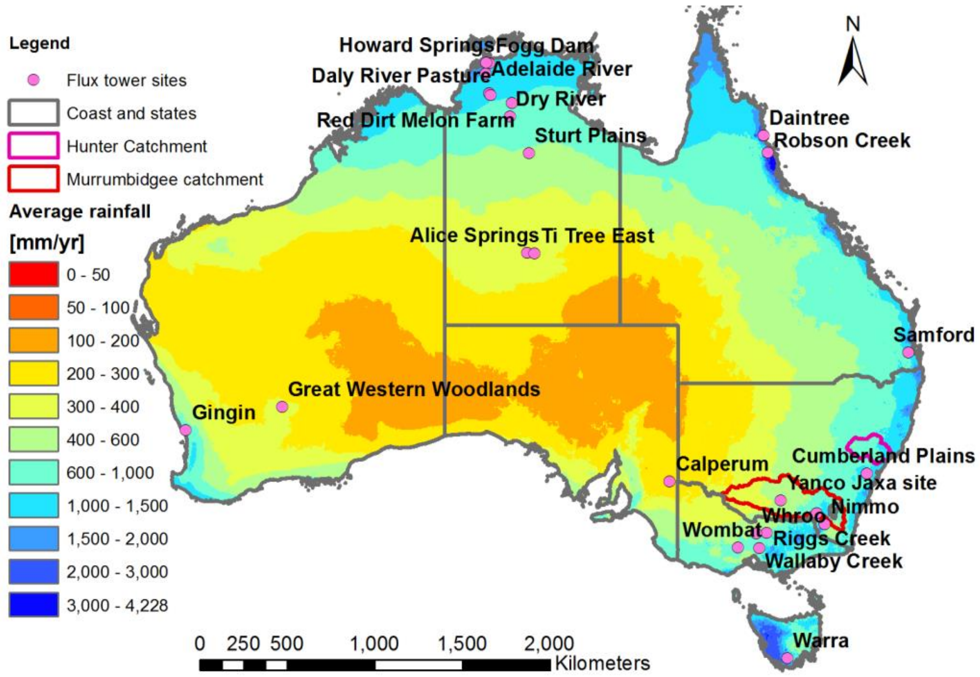



**Figure 7. Flux tower locations where ET and soil moisture are monitored and soil moisture monitoring catchment locations also shown. Average annual rainfall is shown to give an indication of the range of climate conditions sampled within Australia.**

The period 2001-2013 was used for scientific evaluation, being the intersection of years of available output for all models, with a median of 30% of months available for the 25 sites tested (after infilling using the DINGO). This data was not used in calibration of AWRA-L, but some flux tower data was used in calibration of CABLE (see section 3.3).

**Satellite retrieval based gridded estimates:** CSIRO MODIS reflectance-based Scaling ET (CMRSET; Guerschman et al., 2009) satellite ET covering 2001-2013 and the CSIRO developed Simplified Land Surface Temperature (SLST) algorithm (Guerschman et al., 2009; Van Niel et al., 2012), were compared to the observed point estimates of ET from flux towers. CMRSET produces ~250m gridded 8-day cycle national maps of Actual ET based on MODIS satellite data and AWAP climate data, see example AET map for Australia in Figure 8.

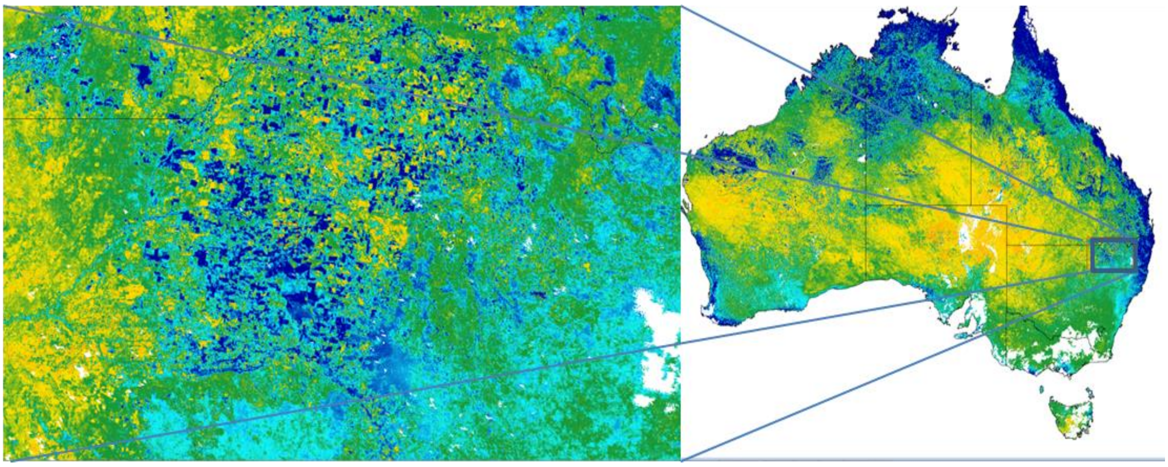


**Figure 8. CMRSET derived map of 8-day Actual Evapotranspiration for 04/07/2014 (noting white area shows no data, most likely affected by clouds). Courtesy Juan Pablo-Guerschman CSIRO.**

### 2.2.5 Vegetation Fraction ($F_{veg}$)

Vegetation fraction cover ($F_{veg}$) provides quantitative information about the vegetation dynamics within a given grid cell. Estimates used here are derived using Moderate-resolution Imaging Spectroradiometer (MODIS) Nadir BRDF-Adjusted Reflectance product (MCD43A4) collection 6 data following Guerschman et al. (2015).

### 2.2.6 Groundwater Deep Drainage

Shi et al. (2015) collated various datasets which could be used for evaluating AWRA-L modelled deep drainage across Australia:

**Long term average:** A long term average recharge dataset has been processed from 6343 individual field estimates of recharge collated by Crosbie et al (Crosbie et al., 2010a; Crosbie et al., 2010b) with some additional points added that were generated from the Bioregional Assessment Programme (www.bioregionalassessments.gov.au). The dataset was filtered to remove any data points that had recharge equal to zero or any points that had recharge more than two thirds of the mean annual rainfall. The remaining points were averaged for 2282 grid cells (0.05˚) that are coincident with the AWRA-L model by taking the geometric mean – see Figure 9. The majority of recharge estimates are based on chloride mass balance estimates, which represent long-term mean annual recharge at the point.

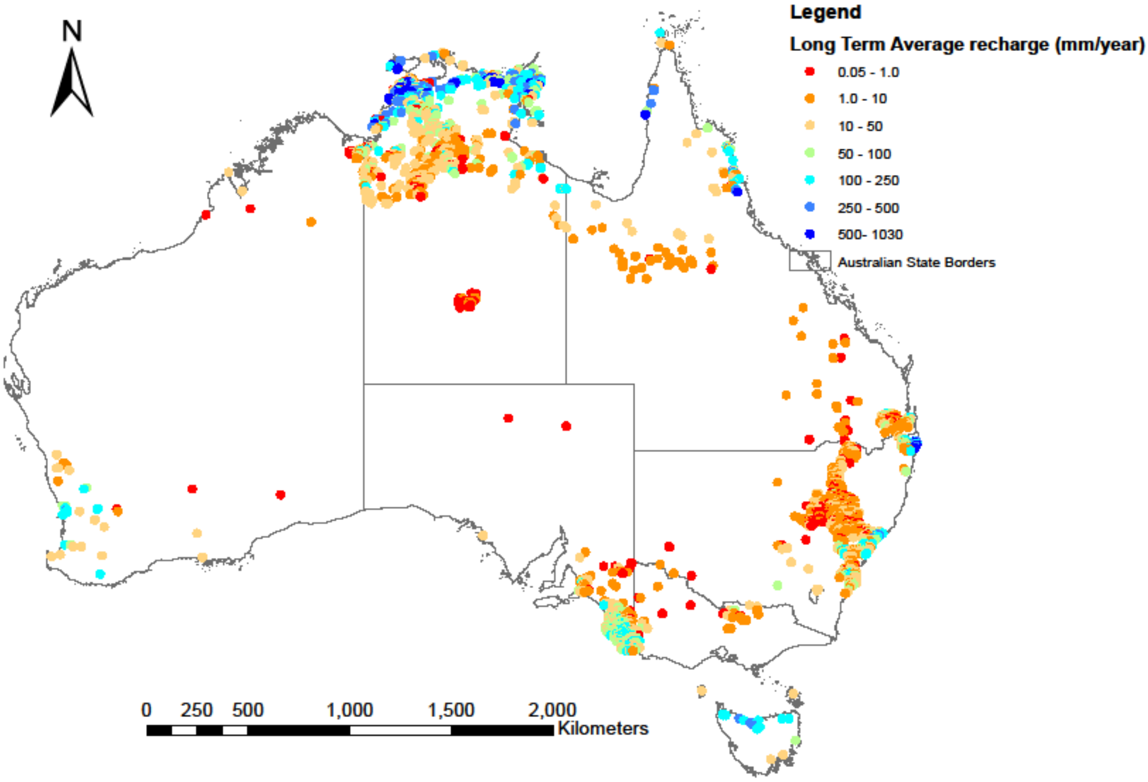



**Figure 9. Long term average recharge estimates**

**Annual recharge time series:** dataset was created using the water table fluctuation (WTF) method and data for the period 1970-2012 at 438 boreholes in the southeast of South Australia and southwest of Victoria – see Figure 10.

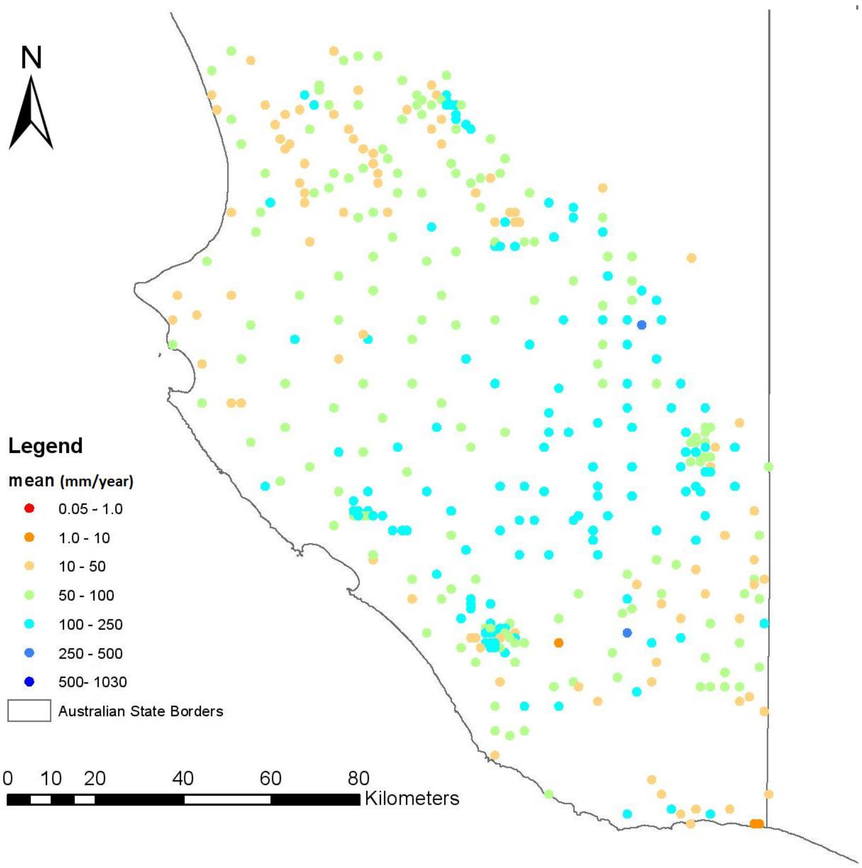



**Figure 10. Location of groundwater bores with annual recharge estimates created using the water table fluctuation method in the southeast corner of South Australia, near the Victorian state border.**

**Monthly time-series:** A further monthly time series dataset covering 6 sites over August 2000-December 2002 in the Tomago sand beds in NSW is available.

Considering the large variability of deep drainage at any point and uncertainties associated with derivation of evaluation datasets, validation of modelled 25 km$^2$ gridded deep drainage is difficult. Further, deep drainage does not necessarily end as groundwater recharge. Deep drainage is an estimate of the water that drains from the bottom of the deep soil layer (6 m) towards the groundwater stores and the recharge is the amount of water reaches to the groundwater stores. Nevertheless, it is of interest to be aware of how AWRA deep drainage estimates compare to other recharge estimates; although deep drainage and recharge are different variables by definition, they are expected to be correlated.

# 3 Models

## 3.1 AWRA-L

AWRA-L (Frost et al., 2018, 2016; Frost and Shokri, 2021; Van Dijk, 2010; Neil Viney et al., 2015; Viney et al., 2014) is a one dimensional, 0.05° grid based water balance model over the continent that has semi-distributed representation of the soil, groundwater and surface water stores.

**Soil layers:** AWRA-L includes three soil layers: Top 0-10cm, Shallow 10cm-100cm, and Deep 100cm-600cm soil. These layers are alternatively referred to as upper, lower and deep soil layers on our external website product. However, naming conventions consistent with the original model design and code are used here.

**Hydrological Response Units:** AWRA-L has three Hydrological Response Units (HRU; i.e. shallow rooted, deep rooted, and impervious landscapes. Shallow rooted vegetation is assumed to have roots to the extent of the shallow soil layer (to 1m), while deep rooted vegetation is assumed to have roots down to 6 m (i.e. the extent of the deep soil layer). The impervious HRU comprises urban landscapes and rocky outcrops and assumes no roots.

**Model processes:** AWRA-L models hydrological processes for:

- Saturation excess overland flow (depending on groundwater store saturation level)
- Infiltration and Hortonian (infiltration excess) overland flow
- Saturation, interflow, drainage and evapotranspiration from soil layers
- Baseflow, evaporation and capillary rise from the groundwater store

With the soil layers modelled separately for 3 (impervious area and shallow and deep rooted vegetation) HRUs – see Figure 1.

**Spatial datasets:** Various spatial datasets are also used to parameterise AWRA-L spatially including:

- Vegetation properties: Estimates of satellite observation derived forest height (1km lidar based estimated derived by Simard et al., 2011), maximum Leaf Area Index (LAI: from analysis of time series of MODIS LAI images), deep soil maximum root water uptake (Vaze et al., 2018), and importantly the proportion HRU (based on estimate of fraction persistent and recurrent vegetation as derived by Vaze et al (2018).
- Slope and hydraulic conductivity affecting infiltration capacity
- Soil drainage/storage parameters:

- Soil hydraulic conductivity and fractional water storage capacity from pedotransfer function applied to clay content from the Soil and Landscape Grid of Australia (Vaze et al., 2018)
- Topography and effective porosity affecting baseflow/saturation (Peeters et al., 2011)
- Hypsometric curves used for conversion from groundwater storage to head relative to the lowest point in the cell

For further details of the AWRA-L v7 algorithms and input data see Frost and Shokri (2021).

**Calibration and evaluation approach:** AWRA-L v7 contains 57 notionally optimisable parameters (4 parameters more than v6) – see Table 1 of Frost and Shokri (2021). Thirty-six parameters are chosen a priori based through previous experience or according to mapping data in order to reduce the number of parameters to be optimised and to better identify parameters that the model is sensitive to. The remaining 21 parameters are optimised across the continent to maximise a composite objective function combining the performance according to various water balance datasets. Automated calibration is undertaken on the National Computational Infrastructure (NCI; http://nci.org.au) supercomputer using distributed simulation of 295 gridded catchments (11320 grid cells), using pre-defined starting states, a full simulation period 01/07/1950 - 30/12/2011, and evaluation period 01/01/1981 - 30/12/2011. The period before the evaluation is used to ensure stable states for slow moving stores such as deep soil moisture and groundwater. The Shuffled Complex Evolution algorithm (Duan et al., 1993) is used for optimisation with 75,000 function evaluations set as the upper limit. The optimisation was undertaken on NCI High Performance Computers (HPC), with each calibration using 1400 CPUs and 1.13 TB memory for approximately 17 hours (more details about the CPU specifications: https://nci.org.au/our-systems/hpc-systems).The calibration was conducted 5 times with different random seed numbers to ensure the optimality of the resulting parameter set, with the final simulation seed chosen according to best objective function value. Input calibration data and defined objective functions were as detailed below.

For evaluation purposes, simulation was from 1950 - 2018, with evaluation for all data sets covering the intersection of the observed data and simulated data. 295 catchments not used in calibration are reserved for catchment based evaluation, and grid based outputs are compared to point based observations. A range of different statistics are used in calibration and evaluation depending on the data type, with the full extent of the data used for evaluation covering the AWRA-L simulation period. Final model performance compared to previous versions is judged according to summary evaluation statistics, based on data not used in calibration.

**Calibration and validation data:** streamflow, ET, soil moisture, vegetation fraction, and TWS at a set of 295 and 291 unimpaired catchments across Australia (see Figure 14) were used in calibration and validation respectively as follows:

- **Catchment streamflow**: covering the period of 1981-2011 was used in calibration and validation, to maintain consistency with previous model evaluations. Given there is no routing processed within AWRA-L currently, catchment streamflow is compared to aggregated runoff across AWRA-L grid cells according to the catchment boundary and area weighting.
- **Catchment evapotranspiration:** CSIRO MODIS reflectance-based Scaling ET (CMRSET; Guerschman et al., 2009) satellite retrieval based grid estimates of 8-day evapotranspiration covering 2001-2017, where 2001-2011 data are used for calibration catchments and 2001-2017 data are used for validation catchments.
- **Catchment soil moisture:** ASCAT product (https://manati.star.nesdis.noaa.gov/datasets/ASCATData.php) satellite retrieval based grid estimates of soil moisture, covering the period of 2002-2013 have been used where 2002-2011 data are used for calibration catchments and 2002-2013 data are used for validation catchments.
- **Terrestrial Water Storage (TWS) anomaly:** represents the water content change within the entire observed soil column at monthly timesteps as observed from GRACE mission satellite covering the period of 2002-2017, that is used for both calibration and validation catchments.
- **Vegetation fraction ($F_{veg}$):** estimates used here are derived using Moderate-resolution Imaging Spectroradiometer (MODIS) Nadir BRDF-Adjusted Reflectance product (MCD43A4) collection 6 data following Guerschman et al. (2015). The 8-day $F_{veg}$ product is used over the periods 2001-2011 for calibration catchments and 2001-2017 for validation catchments.

**Lumped versus spatial calibration:** a new spatial calibration approach was applied in AWRA-L v7, where model pixel output values are compared against spatially distributed satellite data for soil moisture, evapotranspiration, fraction vegetation and terrestrial water storage, rather than using lumped catchment average values of evapotranspiration and soil moisture as used previously. The correlation of each pixel of the model was calculated against the observation then the correlations were aggregated across catchments, with the median value weighted according to catchment size.

**Statistics used in calibration and evaluation:** various statistics are calculated for each catchment for streamflow, model grid cell for remotely sensed soil moisture, ET, Fveg and TWS, or sites for point based measurements including recharge, soil moisture, and ET to assess the model's performance depending on the variable type:

Relative bias (B)

$$B_i = \sum_{t=1}^{T} \frac{Q_{mit} - Q_{oit}}{\bar{Q}_{oi}} \quad (1)$$

Nash-Sutcliffe Efficiency (NSE)

$$NSE_i = 1 - \sum_{t=1}^{T} \frac{(Q_{mit} - Q_{oit})^2}{(Q_{oit} - \bar{Q}_{oit})^2} \quad (2)$$

Pearson's correlation coefficient ($r$)

$$r_i = \frac{\sum_{t=1}^{T}(Q_{oit} - \bar{Q}_{oi})(Q_{mit} - \bar{Q}_{mi})}{\sqrt{\sum_{t=1}^{T}(Q_{oit} - \bar{Q}_{oi})^2}\sqrt{\sum_{t=1}^{T}(Q_{mit} - \bar{Q}_{mi})^2}} \quad (3)$$

Where $Q_{mit}$ and $Q_{oit}$ represent the modelled simulations and observations respectively for site/pixel/catchment *i* and timestep *t* for *T* available observations. $\bar{Q}_{oi}$ and $\bar{Q}_{mi}$ are the mean of the observations and modelled outputs respectively over all timesteps.

The bias and monthly NSE statistics are seen as good metrics for judging AWRA-L model's performance for simulating streamflow and Terrestrial water storage. Pearson's correlation coefficient is a good indicator for variables where the bias (and absolute value) of the variable is not as important as matching the variability (e.g. soil moisture, actual ET or fraction vegetation). Finally, the Kling-Gupta Efficiency (KGE; Gupta et al., 2009) is also used for evaluation in decomposing streamflow performance across the country according to (a) correlation, (b) alpha (ratio of standard deviation of model over observed standard deviation, a measure of variability), and (c) beta (ratio of mean of model over observed mean, a measure of bias):

$$KGE = 1 - \sqrt{(r-1)^2 + \left(\frac{\bar{Q}_{mi}}{\bar{Q}_{oi}} - 1\right)^2 + \left(\frac{st.Q_{mit}}{st.Q_{oit}} - 1\right)^2} \quad (4)$$

where r is Pearson's correlation coefficient value and $st.Q_{oit}$ is standard deviation of observation values and $st.Q_{mit}$ is standard deviation of modelled values.

In addition, Root Square Nash-Sutcliffe Efficiency ($NSE_{sq}$),

$$NSEsq_i = 1 - \sum_{t=1}^{T} \frac{\left(\sqrt{Q_{mit}} - \sqrt{Q_{oit}}\right)^2}{\left(\sqrt{Q_{oit}} - \overline{\sqrt{Q_{oi}}}\right)^2} \quad (5)$$

is used as a low flow performance metric at the daily and monthly timescales.

Bias range (BR), is used as a metric to measure the amount of spread of bias across all the sites evaluated between the 5th and 95th percentiles, with lower values being better:

$$BR = |B_{95\%} - B_{5\%}| \quad (6)$$

where $B_{95\%}$ and $B_{5\%}$ are the 95th and 5th percentiles of bias.

De-seasonalised correlation (DR), is used to estimate correlation but to remove the effect (and potential false skill) of seasonality,

$$DR_i = \frac{\sum_{t=1}^{T}(Q_{oit}-\bar{Q}_{ois})(Q_{mit}-\bar{Q}_{mis})}{\sqrt{\sum_{t=1}^{T}(Q_{oit}-\bar{Q}_{ois})^2}\sqrt{\sum_{t=1}^{T}(Q_{mit}-\bar{Q}_{mis})^2}} \quad (7)$$

where $\bar{Q}_{ois}$ and $\bar{Q}_{mis}$ are the long-term monthly means for modelled and observed values averaged over the entire validation period and *s* represents the month correspond to $Q_{oit}$.

**Calibration objective:**

The following streamflow objective function is evaluated for each catchment simulation (as derived by Viney et al., 2009):

$$F_s = NSE_d - 5\,|\ln(1 + B)|^{2.5} \quad (8)$$

where $NSE_d$ is the daily Nash-Sutcliffe Efficiency and B is relative bias (B). Since the calibration and validation catchments are small enough, and there is no routing process in AWRA-L currently, the runoff aggregated to catchment boundaries is compared to streamflow.

In addition to $F_s$ for streamflow, daily soil moisture correlation ($r_{sm}$), 8-day evapotranspiration correlation ($r_{et}$), 8-day fraction vegetation correlation ($r_{fveg}$) and monthly NSE of de-seasonalised TWS ($NSE_{dsTWS}$) are calculated (by subtracting monthly means from TWS timeseries) for each catchment as different components of the objective function. For TWS the influence of leakage and highly uncertain data on coastal areas is dealt with using reduced weighting when a GRACE pixel has area covered by the Ocean. The weight of a pixel is reduced proportionally with reduction of the GRACE pixel land coverage (see Figure 6).

In the case of the spatially varying data within a catchment, the median value of the statistic is calculated across all cells within each catchment and then the median value is weighted according to the number of cells in each catchment (a proxy for catchment area).

Performance across the calibration catchments is then averaged for each variable type by using the following average:

$$OF_{cm} = \text{mean}\,(OF_{cm25\%}, OF_{cm50\%}, OF_{cm75\%}, OF_{cm100\%}) \quad (9)$$

where $OF_{cmX\%}$ is the $X_{th}$ ranked average percentile $OF_{cm}$ value for each catchment objective where $cm \in c = \{F_s, r_{sm}, r_{et}, r_{fveg}, NSE_{dsTWS}\}$.

This objective function aims to get an adequate fit over a wide range of sites, but also to exclude very poor fitting areas (i.e. those below the 25%), possibly influenced by poor data. Finally, the calibration of AWRA-L maximises the grand objective function across all variables as:

$$\text{grandOF} = 50\%\,NSE_{dsTWS} + 35\%F_s + 7.5\%r_{sm} + 2.5\%r_{et} + 5\%r_{fveg} \quad (10)$$

This weighting is a marked change from AWRA-L v6 with large weighting now applied on TWS (weighted 50%) along with the addition of vegetation fraction, where previously the focus was on streamflow (weighted 70%) with lower weights applied to satellite derived soil moisture and evapotranspiration (15% each). The weights were obtained as a result trial and error tests, starting from the previous model weighting scheme.

In addition to the optimised parameters, some parameters were manually tuned/specified during experimentation and differ from previous versions (see Table 1 of Frost and Shokri, 2020) For further details of calibration, evaluation of model performance and a-priori specification of model parameters see Viney *et al.* (2015), Frost and Wright (2018b, 2018a), and Van Dijk (2010c).

## 3.2 WaterDyn

The WaterDyn model, developed by CSIRO Marine and Atmospheric Research (Raupach et al., 2009), as part of the AWAP, is another daily national 0.05° grid-based biophysical model of the water balance between the atmosphere and soil which runs at a daily timestep, with monthly and weekly outputs published operationally by CSIRO.

Fluxes contributing to streamflow consist of two components: surface runoff and deep drainage. Surface runoff occurs only when the upper soil layer is completely saturated and is then equal to the rate of precipitation. Deep drainage is a function of the relative soil moisture and the saturated hydraulic conductivity of the soil layer.

WaterDyn, like AWRA-L, also uses daily input gridded data (0.05°) from AWAP although WaterDyn uses the recalibrated daily rainfall surfaces (monthly interpolated surfaces disaggregated daily according to the daily rainfall interpolations), as opposed to the standard daily rainfall surfaces as used by AWRA-L across Australia.

WaterDyn model has two soil layers (and no groundwater store) 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.

WaterDyn was parameterised using calibration, and investigation of parameter uncertainty, to streamflow from six unimpaired catchments within the Murrumbidgee (see

Raupach et al., 2009 for more details). Monthly simulation values were available for evaluation covering January 1900 to February 2014.

## 3.3 CABLE

The CSIRO Atmosphere Biosphere Land Exchange (CABLE) model is a community global land-surface model developed by CSIRO, the Bureau and partner universities (Kowalczyk et al., 2006; Wang et al., 2011). The CABLE model is being developed with the intention of use within the Australian Community Climate and Earth System Simulator climate model (ACCESS). CABLE is a land surface model, used to calculate the fluxes of momentum, energy, water and carbon between the land surface and the atmosphere and to model the major biogeochemical cycles of the land ecosystem.

Fluxes contributing to streamflow consist of two components: surface runoff and deep drainage. Drainage, compared to WaterDyn and AWRA-L, is modelled as gravitational drainage from the lowest soil layer. Drainage in the soil layers is modelled according to Richard's equation solution assuming a relationship between hydraulic conductivity and soil moisture content.

CABLE uses daily input climate gridded data (0.05°) from the Bureau operational AWAP service. It is noted that CABLE (like WaterDyn) uses the recalibrated daily rainfall surfaces (monthly interpolated surfaces disaggregated daily according to the daily rainfall interpolations), as opposed to the standard daily rainfall surfaces as used by AWRA-L model. Data are downscaled from daily to hourly time steps (on the half-hourly) using a weather generator (Haverd et al., 2013).

10 soil layers are included in this implementation of CABLE (0.022, 0.058, 0.07, 0.15, 0.30, 0.30, 0.30, 1.20, 3.0, and 4.5m thicknesses depth from topmost to bottommost layer). Secondly, the default CABLE v1.4 soil and carbon modules were replaced respectively by the Soil-Litter-Iso (SLI) soil model (Haverd and Cuntz, 2010) and the CASA-CNP biogeochemical model (Wang et al., 2010) – see Haverd et al (2013). 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.

CABLE parameters were calibrated/constrained according to:

- 50 unimpaired catchment streamflow records spread across Australia (10 from each bioclimatic region except desert) used to compare to long term streamflow (precipitation-ET) from the model. i.e. does not attempt to model short term temporal dynamics of streamflow.

- 6 OzFlux sites Evapotranspiration and gross primary production of Carbon (Howard Springs, Daly River Savanna, Daly River Pasture, Sturt Plains, Tumbarumba, Virginia Park)

Monthly simulation values were available for evaluation covering January 1900 to December 2013.

## 3.4 Summary of model characteristics

The salient features of AWRA-L and peer models (WaterDyn and CABLE) are summarised in Table 1.

**Table 1. Summary of AWRA-L, WaterDyn, and CABLE model characteristics**

| | WaterDyn | CABLE | AWRA-L (v5,v6 and v7) |
|---|---|---|---|
| Reference | Raupach et al (2009) | Wang et al (2010); Haverd et al (2013) | Viney et al (2015); Frost et al (2018); Frost et al (2021) |
| Developer | CSIRO/BoM/ABARES | CSIRO/BoM + universities | CSIRO/BoM |
| Purpose | Monitoring terrestrial water balance | Land surface scheme for the Australian Community Climate and Earth-System Simulator (ACCESS) | Water resources reporting, assessment, and monitoring |
| Soil layers<br><br>(spatially varying properties) | 2<br><br>(depth, saturated volumetric water content) | 10<br><br>(saturated hydraulic conductivity, field capacity, etc) | 3<br><br>(saturated hydraulic conductivity, % available water holding capacity) |
| Calibration | Parameter calibration and sensitivity analysis to 6 catchments in Murrumbidgee | Calibration to derived ET (50 catchments across 10 climate zones within Australia) and flux tower data | Streamflow over ~300 catchments and satellite soil moisture and ET |

## 3.5 Lumped-rainfall runoff models

Two lumped catchment conceptual rainfall-runoff models are used for streamflow comparison purposes against AWRA-L model:

- GR4J (Perrin et al., 2003): a 4-parameter model derived from empirical analysis over many catchments towards finding the most efficient/parsimonious model structure.
- Sacramento (Burnash, 1995): the Sacramento model is a conceptual catchment water balance model developed for the U.S. National Weather Service that models the rainfall–runoff process at daily time-steps. A 13-parameter implementation was used here.

These models are calibrated in a different way to AWRA-L, in that they are calibrated for individual catchments, rather than finding a single parameter set to cover the entire model domain. Once the parameters are found for the calibration catchments, they are transferred by nearest-neighbour regionalisation to the closest validation catchments nearby. Nearest-neighbour regionalisation is a method used as a practical approach to regionalisation/predictions in ungauged basins, to produce the best performance possible where calibration is possible, but to also allow prediction in areas where the model cannot be calibrated. It is noted that the following function was used as a calibration objective for the conceptual rainfall runoff models:

$$F_s = (NSE_d + NSE_m)/2 - 5 \, |\ln(1 + B)|^{2.5}$$

where the monthly NSE ($NSE_m$) is included along with the daily NSE ($NSE_d$), as used in AWRA-L v5 (see Viney et al., 2016), differing from the current AWRA-L streamflow objective shown in Eq (8). This is as the streamflow objective was changed in AWRA-L v6 to rely on the daily NSE only, however the results for the rainfall runoff modelling include the monthly component. For further details of the methods applied for the conceptual rainfall runoff modelling approach used here see Ramchurn and Frost (2013).

# 4 Evaluation approach

## 4.1 Evaluation criteria

The AWRA-L model was primarily developed for water resource applications across Australia. Therefore, evaluation criteria are primarily based on available observed hydrological data across Australia. Primary metrics are applied to streamflow for water resources assessment purposes, but secondary metrics are applied to ensure model works well across the water balance. Improvements in model performance should be judged on data reserved for validation (i.e. separate to calibration data) so that performance is tested for predictions in ungauged basins, following the principles outlined in Refsgaard and Henriksen (2004). It is to be noted that all observed datasets used for testing AWRA-L performance have uncertainty associated with them. Future improvements of AWRA-L (and other) models can be judged according to the performance of AWRA-L v7 using these metrics.

**Primary metric** – Assessment of AWRA-L against observed streamflow

Catchment Streamflow is assessed based on the following metrics:

- KGE and NSE for daily ($KGE_d$, $NSE_d$) and monthly ($KGE_m$, $NSE_m$) Streamflow
- Relative Bias (B) in long-term averages
- Bias range (BR)
- Root square NSE

**Secondary metric** – Assessment of AWRA-L against derived data

- Soil moisture: daily and monthly correlation of probe-based point soil moisture sampled for the profile (0-90cm) with AWRA-L soil moisture for 0-100cm depth across Australia (see Figure 4).
- Actual ET: correlation of daily and monthly flux tower ET to AWRA ET.
- Vegetation fraction: correlation of 8-day satellite based estimates of vegetation fraction to AWRA vegetation fraction.
- Deep drainage: correlation between long-term reliable point measurements of recharge with AWRA-L deep drainage
- Terrestrial water storage: correlations of monthly satellite based estimates of catchment terrestrial water storage with the summed water store components.

**Tertiary metric** – Behaviour

- Checking AWRA-L simulations of internal fluxes and checking sensibility of national AWRA simulations for reporting purposes (e.g., no major irregular spatial patterns due to regionalisation, time-series plots for select locations).

The assessment criteria according to the observed data (the Primary and Secondary metrics above) are provided in Table 2. Aspirational targets are provided based on how the Bureau would like the AWRA model to perform, based on what we consider to be reasonable baseline performance characteristics and past experience of performance with peer models. For example, for the majority of catchments this is to perform better than the average/climatology for streamflow therefore, we want to have less than 5% at zero NSE (equivalent to climatology) and to have at least half of the catchments above 0.5 NSE (considered good performance for how the model is intended to be used). AWRA-L is assessed against these criteria in Chapter 5.

Chapter 6 presents a brief examination of AWRA-L outputs according to the Tertiary metric based on behaviour of the model for reporting purposes.

**Table 2. AWRA-L assessment criteria**

| Variable | Assessed against | Assessment criteria | Comparison models | Aspirational target |
|---|---|---|---|---|
| **Streamflow** | Gauged streamflow (calibration and validation sites) | Daily NSE<br>Monthly NSE<br>Bias | CABLE<br>WaterDyn<br>Rainfall-runoff model (Sacramento and/or GR4J) – local calibration/ nearest neighbour regionalisation | Daily and monthly NSE:<br>(a) Less than 5% catchments with NSE<0<br>(b) greater than 50% catchments with NSE>0.5<br>Bias:<br>(a) 50 % of catchments with -30% <bias<30%,<br>(b) 90% of catchments with -50%<Bias<100%, and<br>(c) No systematic spatial pattern of under- or over-estimation (i.e. low Bias when aggregated, mean and median bias close to 0) |
| **Soil moisture** | Profile soil moisture from dedicated field observations | Daily and monthly correlation | CABLE<br>WaterDyn | 50% with daily correlation > 0.75<br>50% with monthly correlation >0.75 |
| **TWS** | Satellite estimates | Monthly correlation<br>De-seasonalised correlation | | 50% with monthly correlation > 0.75 |
| **Actual ET** | Flux ET | Monthly correlation | CABLE<br>WaterDyn | Monthly correlation<br>(a) 95% sites/cells with R>0.5<br>(b) >50% sites/cells with R>0.8 |
| **Vegetation fraction** | Satellite estimates | Monthly correlation<br>De-seasonalised correlation | | 50% with correlation > 0.75<br>50% with de-seasonalised correlation > 0.75 |
| **Deep drainage** | National Long-term average dataset<br>Annual time series dataset | Spatial Correlation / bias<br>Correlation | CABLE<br>WaterDyn | Long-term average dataset:<br>(a) 25% bias value below zero<br>(b) 75% bias value above zero<br>(c) Spatial correlation above 0.5<br>Annual times-series dataset:<br>Median annual correlation above 0.5 |

# 5 Evaluation according to observed data

Various statistics for calibration and validation catchments are now presented for each model to assess their performance against observed hydrological data sets including streamflow, soil moisture, terrestrial water storage, actual evapotranspiration, vegetation fraction and recharge. Before going into detail for each dataset, a summary diagram showing the performance of AWRA-L versions at a glance for the validation datasets is shown in Figure 11, with better performance according to statistics generally plotting further out on the radar plot diagrams. Median catchment/point statistics are presented, with values closer to 1 for correlation (R), de-seasonalised correlation and NSE being better, and values closer to zero for bias range being better.

For soil moisture (Fig 11. top left radar), the correlation of AWRA-L v7 improves compared to the AWRA-L v6, although not as well; as AWRA v5 (see Table 2). Particularly, the top layer soil moisture noticeably improves against all observations, but still does not satisfy the target performance of a correlation of at least 0.75.

AWRA-L v7 modelled evapotranspiration (Fig 11. top right radar) is considerably improved compared to the previous versions, where the correlations of AWRA-L v7 mostly meet or exceed the aspirational performance. This improvement is most significant for daily correlation with OzFlux ET and monthly correlation with satellite-based Fveg observations.

AWRA-L v7 groundwater (Fig 11. bottom left radar) improves in terms of yearly and monthly de-seasonalised correlation against satellite GRACE observations and improves somewhat for point based estimates of recharge (annual correlations). No significant changes are detected when it is compared in terms of long-term average recharge spatial correlations.

AWRA-L v7 modelled streamflow (Fig 11. bottom right radar) outperforms other versions in terms of daily/monthly NSE, daily sqNSE and daily KGE and bias range. Moreover, investigation of performance of AWRA-L for impervious area shows that AWRA-L v7 performs the best of all AWRA-L versions for the statistics presented (noted as NSE (daily imp) in the diagram).

In following sections, the evaluation performance is presented in more detail where statistics are presented using maps and boxplots, showing the cumulative distribution of the statistics across all sites, with the box indicating the 25% percentile, median and 75% percentile (e.g. 25% percentile for the 295 calibration sites means that 74 sites have lower values).

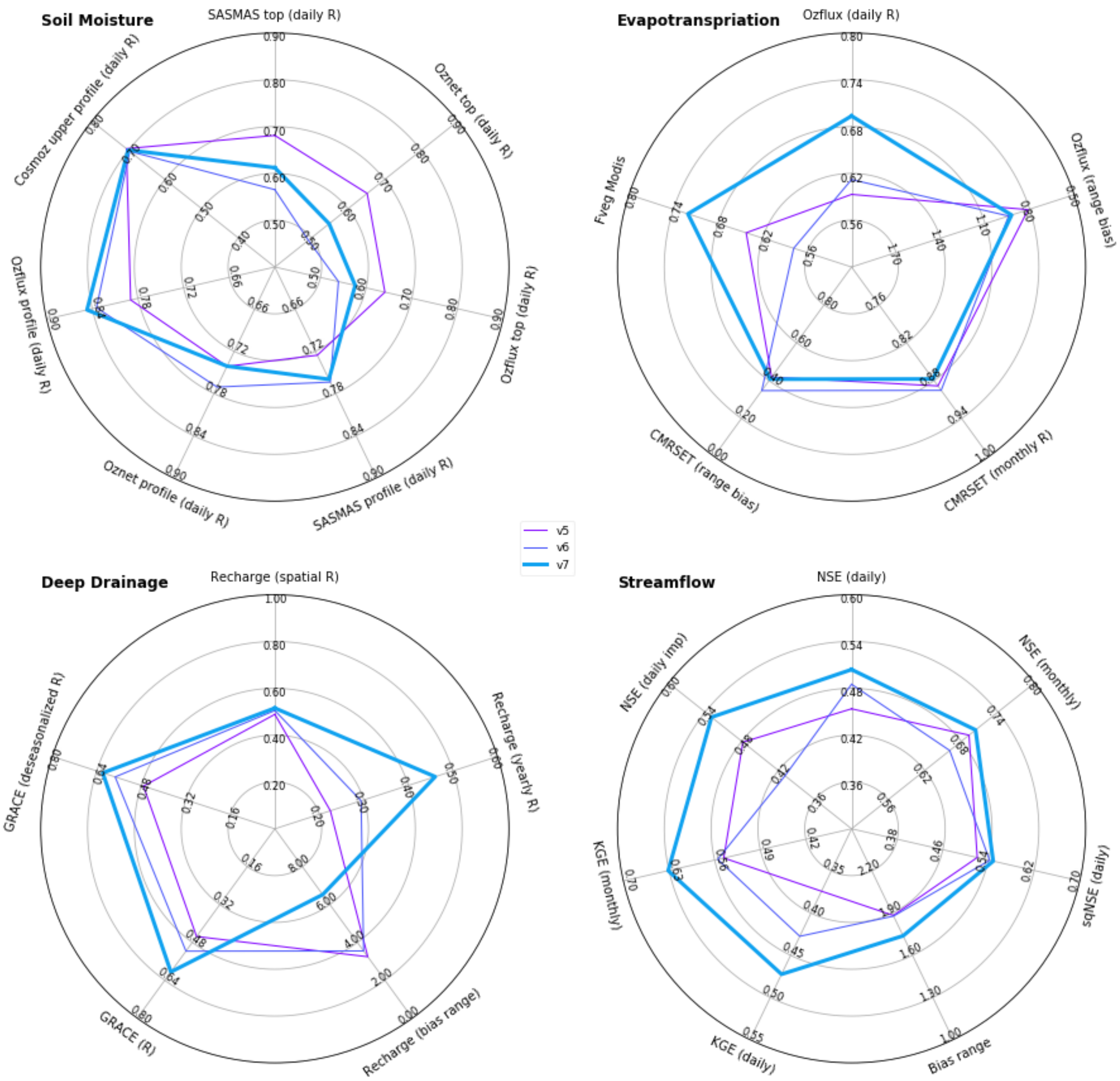



**Figure 11. Summary of the performances of the AWRA-L versions. The values are the median performance metrics among different observation sets based on validation sites. Note that R represents correlation and imp indicates impervious area. For all metrics, the outer is better.**

## 5.1 Streamflow

The AWRA-L model performance has been assessed against other national models (WaterDyn and CABLE) as well as typical rainfall-runoff catchment scale models (GR4J and Sacramento) across Australia. Calibration and validation daily NSE/KGE and monthly NSE, NSE of root squares, and relative bias are plotted in Figure 12. AWRA-L v7 improves over previous versions of AWRA-L for daily NSE and KGE (Figure 12 abcd), and also monthly NSE and bias (Figure 12 efgh). This improvement is attributed to the improved model parametrisation and structural changes related to baseflow and the updated calibration approach which puts 50% weight on terrestrial water storage in the objective function, along with updated input datasets.

For the national landscape/landsurface models, AWRA-L model performs better for streamflow than WaterDyn and CABLE according to monthly NSE and bias (Figure 12 efgh) over the AWRA-L calibration and validation catchments across Australia. This result is expected due to a) AWRA-L being designed to represent runoff characteristics more accurately; and b) AWRA-L is calibrated directly to streamflow characteristics nationally.

For the locally calibrated nearest neighbour regionalised rainfall runoff models, AWRA-L performs worse in the calibration catchments than the locally calibrated models, due to the differing calibration approach used (AWRA sacrifices local performance for as good performance as possible across multiple sites and variables). In particular, bias is near zero for the locally calibrated models due to each of the models having terms that can effectively match the average flow at a particular site where calibrated, while AWRA-L tries to minimise the bias over a set of sites. However, over the validation catchments AWRA-L bias has less spread about zero, providing confidence in the spatial predictive qualities of AWRA-L which has significance for predictions in ungauged basins. In general, AWRA-L v7's calibration, experienced degradation in terms of bias when it is compared with the previous version; however, the validation phase is associated with an enhancement. Significantly, AWRA-L v7 monthly performance for the validation catchments (Figure 12 f) is approaching the performance of nearest neighbour regionalised rainfall-runoff models (GR4J and Sacramento), even though AWRA-L is not calibrated purely to streamflow like the lumped rainfall-runoff models. AWRA-L performs approximately 0.04 worse for daily NSE (Figure 12 f) than the locally calibrated models.

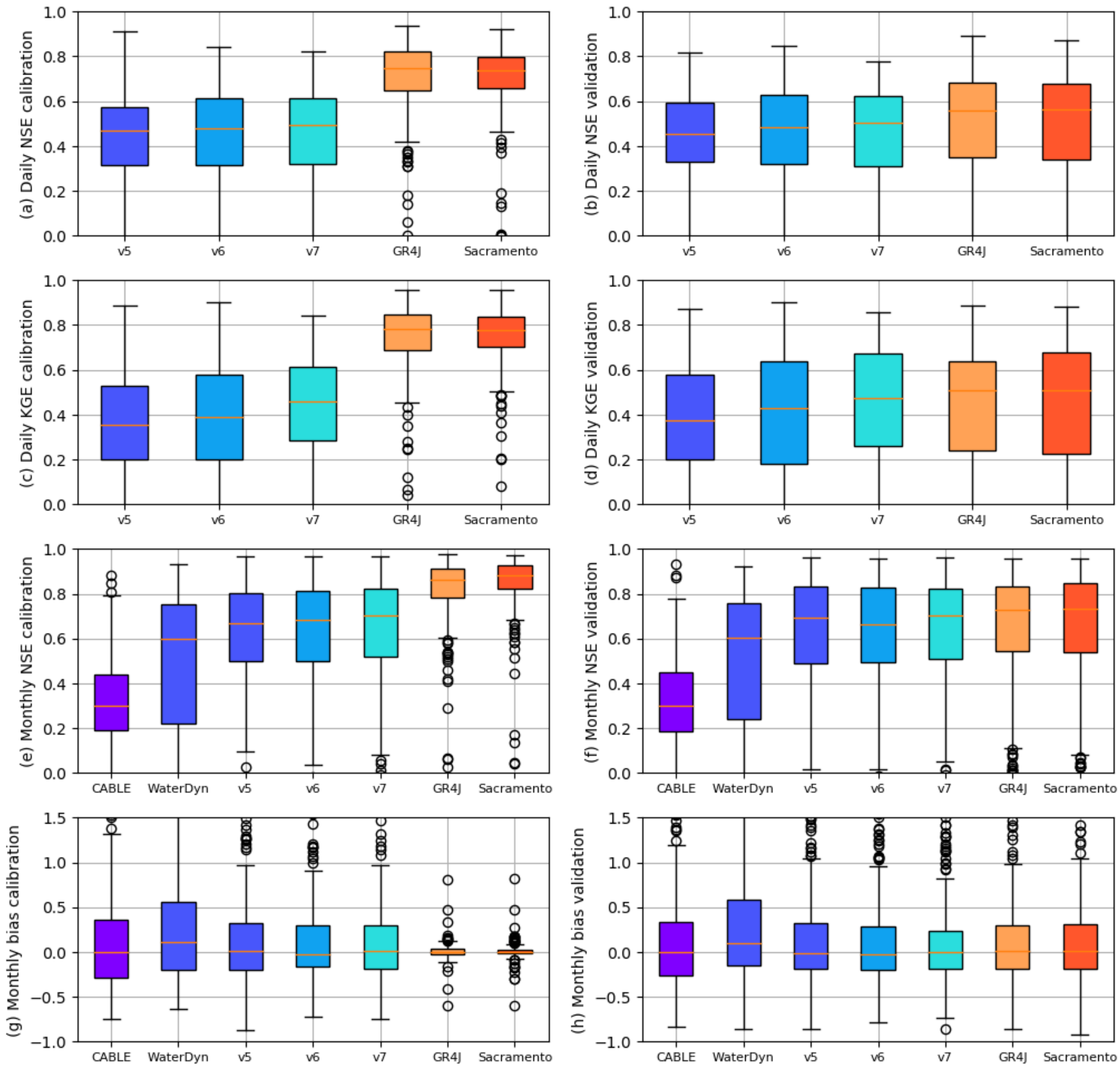



**Figure 12. Streamflow statistics for calibration (left) and validation (right) catchments; including Daily NSE (a,b) and KGE (c,d) and Monthly NSE (e,f) and Bias (g,h)**

The performance of AWRA-L according to daily NSE, monthly NSE and relative bias are presented in Table 3 to Table 5. Evaluation criteria listed in Table 2. AWRA-L assessment criteria are bold in the tables for model benchmarking purposes and comparison to the aspirational targets.

**Table 3. Daily NSE percentiles for each model**

| Calibration | 0% | 5% | 25% | 50% | 75% | 95% | 100 | Validation | 0% | 5% | 25% | 50% | 75% | 95% | 100% |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **AWRA-Lv5** | -16.50 | -0.95 | 0.32 | 0.47 | 0.57 | 0.72 | 0.91 | **AWRA-Lv5** | -44.57 | -0.27 | 0.33 | 0.45 | 0.59 | 0.73 | 0.82 |
| **AWRA-Lv6** | -31.24 | **-0.37** | 0.32 | 0.48 | 0.61 | 0.73 | 0.84 | **AWRA-Lv6** | -10.50 | **-0.06** | 0.32 | 0.49 | 0.63 | 0.73 | 0.85 |
| **AWRA-Lv7** | -33.69 | -1.50 | 0.32 | **0.49** | 0.61 | 0.73 | 0.82 | **AWRA-Lv7** | -26.06 | -1.07 | 0.31 | **0.50** | 0.62 | 0.72 | 0.78 |
| **GR4J** | 0.00 | 0.45 | 0.65 | 0.75 | 0.82 | 0.88 | 0.94 | **GR4J** | -12.44 | -0.60 | 0.35 | 0.56 | 0.68 | 0.79 | 0.89 |
| **Sacramento** | -1.94 | 0.47 | 0.66 | 0.74 | 0.80 | 0.86 | 0.92 | **Sacramento** | -7449.97 | -1.55 | 0.34 | 0.56 | 0.68 | 0.80 | 0.87 |
| **Benchmark** | | **-0.37** | | **0.49** | | | | **Benchmark** | | **-0.06** | | **0.50** | | | |

* Daily results for CABLE and WaterDyn model are not available for the comparison

**Table 4. Monthly NSE percentiles for each model**

| Calibration | 0% | 5% | 25% | 50% | 75% | 95% | 100% | Validation | 0% | 5% | 25% | 50% | 75% | 95% | 100% |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **CABLE** | -267.48 | -0.74 | 0.19 | 0.30 | 0.44 | 0.70 | 0.88 | **CABLE** | -23.62 | -0.43 | 0.18 | 0.30 | 0.45 | 0.72 | 0.93 |
| **WaterDyn** | -480.99 | -3.30 | 0.22 | 0.60 | 0.75 | 0.87 | 0.93 | **WaterDyn** | -46.98 | -2.26 | 0.24 | 0.60 | 0.76 | 0.89 | 0.92 |
| **AWRA-L v5** | -22.75 | -0.46 | 0.50 | 0.67 | 0.80 | 0.90 | 0.97 | **AWRA-L v5** | -46.73 | -0.33 | 0.49 | 0.69 | 0.83 | 0.92 | 0.96 |
| **AWRA-L v6** | -38.32 | **-0.35** | 0.50 | 0.68 | 0.81 | 0.92 | 0.97 | **AWRA-L v6** | -13.45 | **-0.19** | 0.50 | 0.66 | 0.83 | 0.92 | 0.96 |
| **AWRA-Lv7** | -36.81 | -0.88 | 0.52 | **0.70** | 0.82 | 0.91 | 0.97 | **AWRA-Lv7** | -20.20 | -0.86 | 0.51 | **0.70** | 0.82 | 0.91 | 0.96 |
| **GR4J** | -0.02 | 0.56 | 0.78 | 0.86 | 0.91 | 0.94 | 0.98 | **GR4J** | -16.67 | -0.24 | 0.54 | 0.73 | 0.83 | 0.91 | 0.95 |
| **Sacramento** | -5.01 | 0.64 | 0.82 | 0.88 | 0.93 | 0.96 | 0.97 | **Sacramento** | -1671.51 | -0.31 | 0.54 | 0.73 | 0.85 | 0.93 | 0.96 |
| **Benchmark** | | **-0.35** | | **0. 70** | | | | **Benchmark** | | **-0.19** | | **0.70** | | | |

**Table 5. Relative bias percentiles for each model (BR=Bias Range)**

| Calibration | 0% | 5% | 25% | 50% | 75% | 95% | 100% | BR% | Validation | 0% | 5% | 25% | 50% | 75% | 95% | 100% | BR |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **CABLE** | -0.75 | -0.54 | -0.29 | 0.00 | 0.36 | 1.43 | 126.63 | 1.97 | **CABLE** | -0.84 | -0.53 | -0.26 | 0.00 | 0.33 | 1.55 | 10.84 | 2.08 |
| **WaterDyn** | -0.64 | -0.48 | -0.20 | 0.12 | 0.56 | 1.86 | 109.90 | 2.34 | **WaterDyn** | -0.86 | -0.46 | -0.16 | 0.10 | 0.59 | 2.53 | 14.92 | 2.99 |
| **AWRA-L v5** | -0.87 | -0.48 | -0.20 | 0.02 | 0.32 | 1.32 | 20.71 | 1.80 | **AWRA-L v5** | -0.86 | -0.47 | **-0.18** | -0.01 | 0.32 | 1.41 | 8.98 | 1.88 |
| **AWRA-L v6** | -0.73 | **-0.41** | **-0.17** | -0.03 | 0.30 | 1.27 | 30.21 | 1.68 | **AWRA-L v6** | -0.78 | -0.44 | -0.21 | -0.02 | 0.29 | 1.44 | 5.64 | 1.88 |
| **AWRA-L v7** | -0.75 | -0.45 | -0.19 | 0.01 | **0.29** | **1.20** | 15.93 | 1.65 | **AWRA-L v7** | -0.86 | **-0.44** | -0.19 | 0.00 | **0.23** | **1.30** | 6.37 | 1.74 |
| **GR4J** | -0.60 | -0.07 | -0.03 | 0.00 | 0.03 | 0.09 | 0.81 | 0.16 | **GR4J** | -0.86 | -0.59 | -0.19 | 0.01 | 0.30 | 1.35 | 7.39 | 1.94 |
| **Sacramento** | -0.60 | -0.07 | -0.02 | 0.00 | 0.02 | 0.11 | 1.78 | 0.18 | **Sacramento** | -0.92 | -0.52 | -0.19 | 0.02 | 0.31 | 1.34 | 7.33 | 1.86 |
| **Benchmark** | | **-0.41** | **-0.17** | | **0.29** | **1.20** | | | | | **-0.44** | **-0.18** | | **0.23** | **1.30** | | |

AWRA-L v7 meets the aspirational daily NSE median target of 0.5, although slightly below 0 at 5%. Similarly, at the monthly scale, the AWRA-L v7 shows the improved performance with median NSE equal to 0.7, although the 5% value remains below 0. In terms of bias AWRA-L v7 does meet the criteria in validation for 50% of sites (25% to 75%) to be within -0.3 and 0.3, similar to the previous versions. It does not yet meet the criteria for the 90% of sites (5% to 95%) being between -0.5 and 1; with the 95% value showing a bias of 120% in validation. In this report, the Bias range is introduced to quantitatively measure the spread of bias across different models; as shown here the bias range is significantly reduced in AWRA-L v7

AWRA-L (a) daily NSE and (b) monthly relative bias are plotted in Figure 13 to evaluate spatial performance. AWRA-L v7 performs well (above 0.5 daily NSE) in Coastal NSW and Victoria, the majority of Queensland, the majority of Tasmania, South Western West Australia and coastal catchments in the Northern Territory. AWRA-L has lower performance for catchments along the Great Dividing Range (from Victoria to NSW/Queensland border) and also in Western Australia along the Darling Scarp. This appears to be partly due to positive bias in these areas. Possible reasons for this bias include (a) deep soil store rooting depth being insufficient (e.g. Jarrah forests of Darling Scarp having roots to 20 metres rather than 6m currently) causing underestimated ET, (b) losses to groundwater systems/transfer that are currently unaccounted for (ie. losses cannot be included in the system currently) and (c) losses due to inadequate routing procedure, amongst other possibilities.

Table 6 shows that performance has improved in general at the daily timestep according to NSE and KGE over the majority of catchments.

**Table 6. Percentage of 295 calibration and 291 validation catchments that show improvement in AWRA-L v7 compared to AWRA-L v6 according to daily NSE and KGE**

| Observed data | NSE | KGE |
|---|---|---|
| Calibration (295 sites) | 57% | 67% |
| Validation (291 sites) | 51% | 65% |

AWRA-L daily (a) KGE, (b) correlation, (c) KGE alpha (std. dev. modelled/std. dev. observed) and (d) KGE beta (mean modelled/mean observed) are plotted spatially in Figure 14 to further investigate spatial performance. In general, number of catchments with daily KGE greater than 0.5 has increased significantly from 99 to 128 in the calibration phase and from 117 to 137 in the validation. Underestimation of variability (alpha) appears to be an aspect of poor performance. Further investigation is required to determine the reasons for this underestimation.

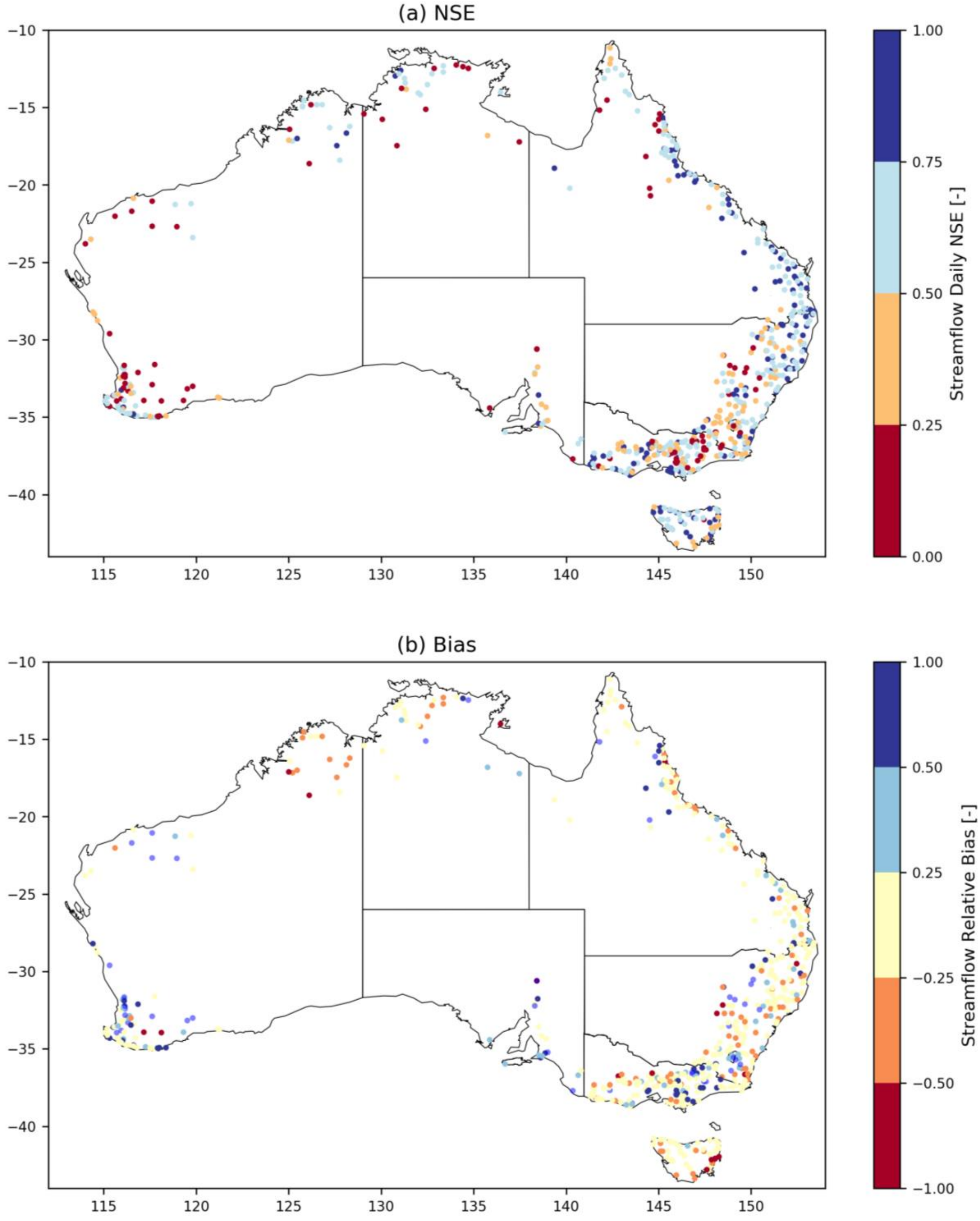



**Figure 13. Map of AWRA-L v7 runoff (a) daily NSE and (b) monthly relative bias compared to streamflow. Calibration and validation sites shown.**

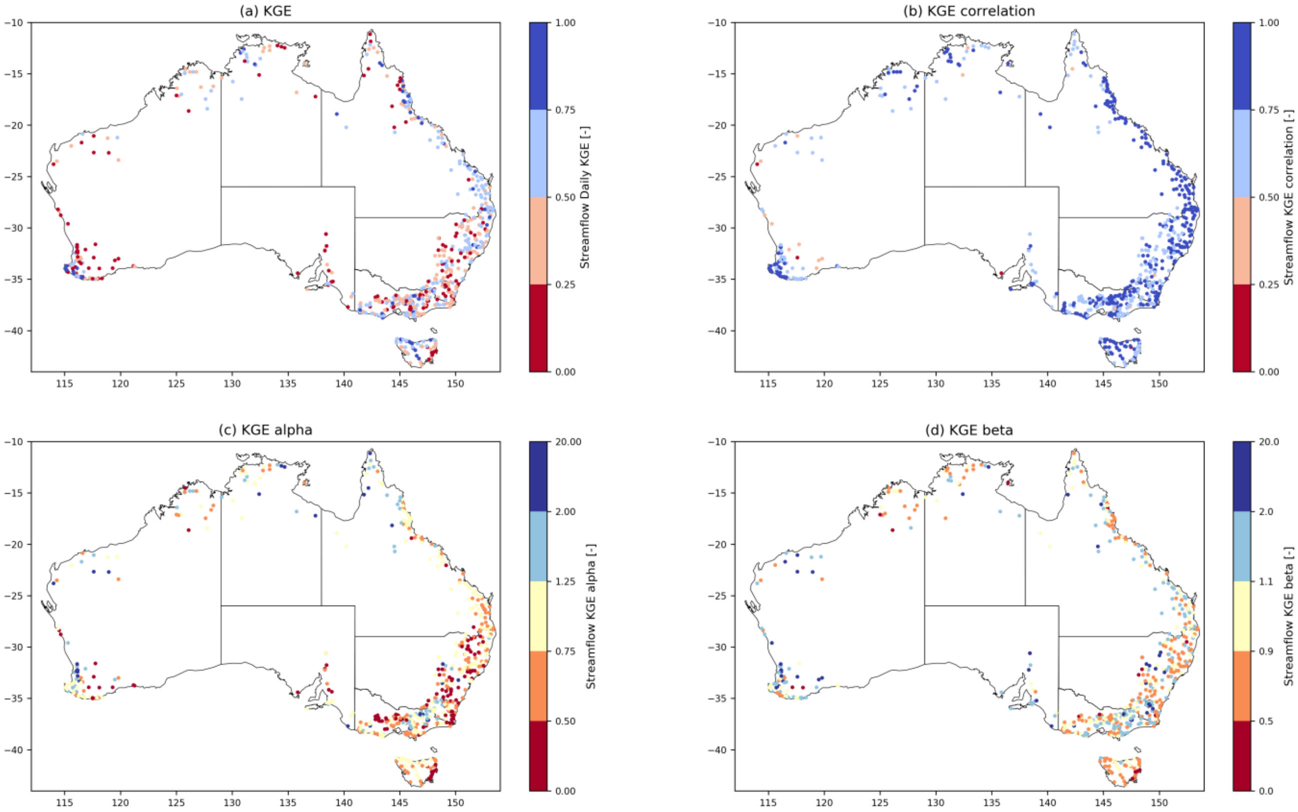



**Figure 14. Map of AWRA-L v7 runoff (a) KGE, (b) KGE correlation, (c) KGE alpha and (d) KGE beta compared to streamflow. Calibration and validation sites shown.**

Time series comparing outputs of AWRA-L v5 to v7 against monthly streamflow and other available water balance data are provided for a range of example catchments in Chapter 6, in Figure 40 to Figure 49. Various sites reproduce the monthly streamflow variability well (eg. Figure 40 Arthur Crk, Qld, Figure 41 Albert River, Qld, Figure 46 Dombakup Brook, WA, Figure 49 East Baines, NT). Regarding the positive biases observed in Figure 13b and Fig.14d, selected South and South Eastern catchments have a high level of baseflow compared to that observed (eg. Figure 43 Fifteen Mile Creek, Vic

Figure 44 Kyeamba Creek, NSW and Figure 45 Victor harbour, SA) during drier months coinciding with times of underestimated terrestrial water storage. Conversely Figure 47 (Yarragil Brook WA) shows that streamflow is overestimated due to too high a storage, and underestimated ET, likely due to underestimation of deep rooted vegetation ET. Thus, the soil storage capacity (and depth of soil layers) and associated root depth are likely areas for potential improvement of streamflow.

## Evaluation of AWRA for catchments containing impervious area

AWRA-L v5, v6 and v7 are tested for the 13 catchments that have significant impervious area, noting that these catchments were not in the calibration or validation set used. AWRA-L v7 improves over previous versions for most metrics presented in Figure 15 (daily and monthly NSE/KGE, daily sqNSE and bias). The exception is monthly NSE where the AWRA-L v5 performs the better for the median value and AWRA-L v7 mostly improves only for 95$^{th}$ percentile. AWRA-L v7 bias has significantly less spread around zero, suggesting improvement of AWRA-L v7 performance for streamflow modelling compared to the previous versions. The improvements here are notable as they occurred due to the introduction of the new HRU only and with no calibration over these types of catchments, or for related parameters. This means that the performance of the other part of the model were not affected by this improvement and that calibration may further improve this component.

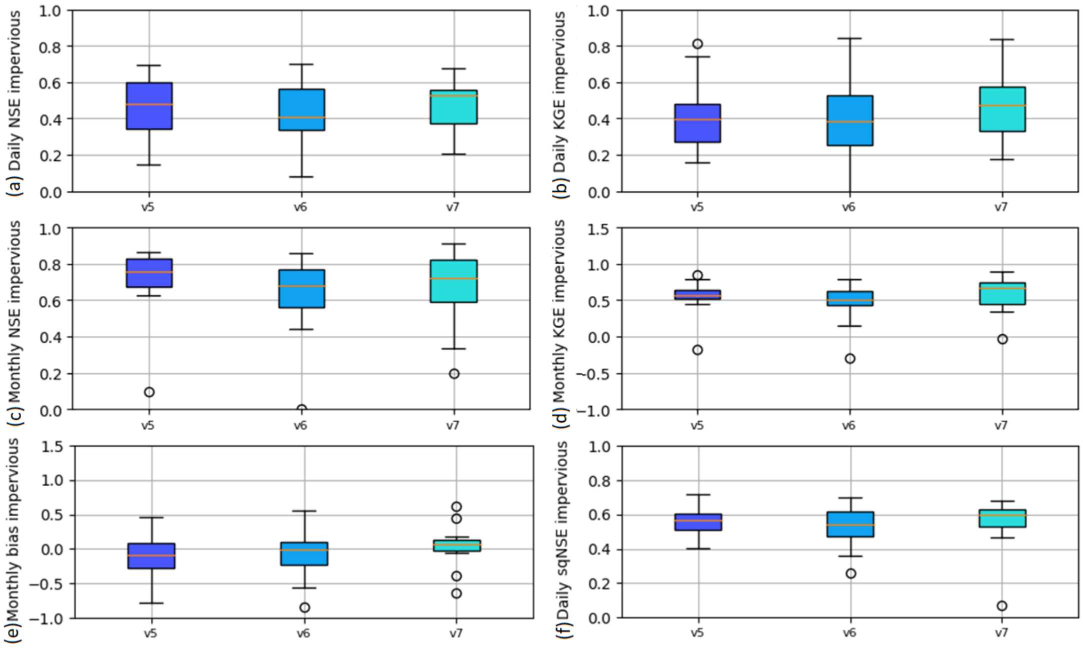



**Figure 15. Streamflow statistics for impervious validation catchments; including Daily and Monthly NSE (a,c) and KGE (b,d), monthly bias (e) and daily sqNSE (f).**

The location of selected impervious catchments and the corresponding monthly time series for all AWRA-L versions over 32 years are shown in Figure 3 and Figure 16, respectively.

**Table 7. Daily and monthly NSE percentiles for impervious catchments modelled by different version of AWRA-L**

| | 0% | 5% | 25% | 50% | 75% | 95% | 100% |
|---|---|---|---|---|---|---|---|
| **Models** | | | | **Daily NSE** | | | |
| **AWRA-L v5** | 0.15 | 0.25 | 0.34 | 0.48 | 0.60 | 0.68 | 0.69 |
| **AWRA-L v6** | 0.08 | 0.18 | 0.34 | 0.41 | 0.56 | 0.67 | 0.70 |
| **AWRA-L v7** | 0.20 | 0.22 | 0.37 | **0.53** | 0.56 | 0.65 | 0.68 |
| **Benchmark** | | | | **0.53** | | | |
| | **0%** | **5%** | **25%** | **50%** | **75%** | **95%** | **100%** |
| **Models** | | | | **Monthly NSE** | | | |
| **AWRA-L v5** | 0.10 | 0.41 | 0.67 | **0.76** | 0.83 | 0.86 | 0.86 |
| **AWRA-L v6** | 0.00 | 0.27 | 0.56 | 0.68 | 0.77 | 0.85 | 0.86 |
| **AWRA-L v7** | 0.20 | 0.28 | 0.59 | 0.72 | 0.82 | 0.87 | 0.92 |
| **Benchmark** | | | | **0.76** | | | |

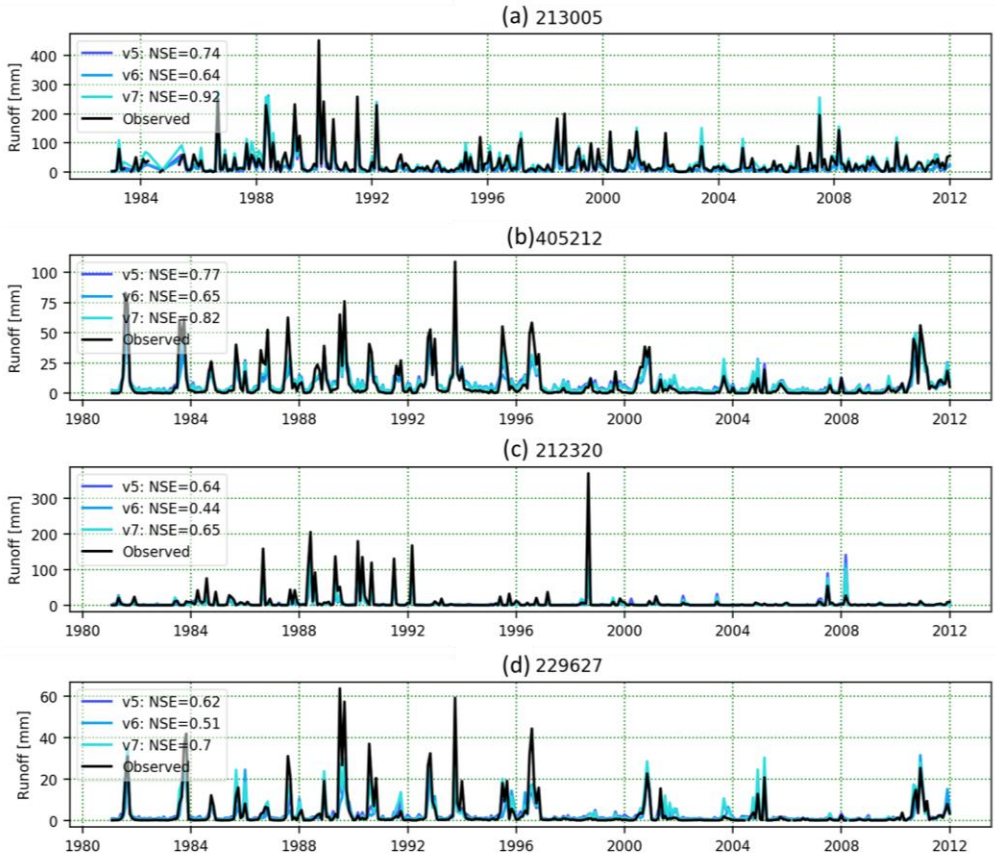



**Figure 16. Monthly runoff time series for four select catchments with impervious area comparing modelled and observed flow.**

## 5.2 Soil moisture

The point based profile (0-90cm) estimates of soil moisture are compared to the layers of each model (weighted according to degree of overlap) using monthly correlation for the OzNet Murrumbidgee data (Figure 17), Upper Hunter SASMAS data (Figure 18), OzFlux data (Figure 19) and CosmOz data (Figure 20) across the Australia. This comparison uses the entire record that is available covering the model simulations.

For the OzNet set, the performance of AWRA-L v7 is slightly better compared to AWRA-L v5 and v6 for daily and monthly for higher range values and similarly for the median and 25th percentile. For the SASMAS set, AWRA-L v7 outperforms previous versions. AWRA-L performs similarly to CABLE and better than WaterDyn at the monthly timescale over both sets of sites (noting WaterDyn was parameterised based on testing in 6 catchments in the Murrumbidgee). The overall result of this comparison is therefore that AWRA-L represents profile (0-90cm) soil moisture temporal dynamics as well as CABLE, and better than WaterDyn (particularly for the Upper Hunter SASMAS data).

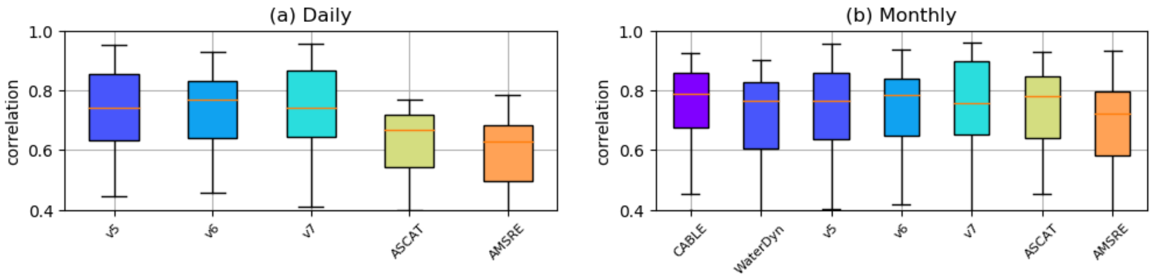



**Figure 17. (a) Daily and (b) Monthly correlation of models against Murrumbidgee OzNet data 2001-2013 profile (0-90cm) soil moisture. Satellite data is evaluated over a shorter period (AMSR-E: 2002-2011, ASCAT: 2007-2011) – and relates only to the top few cm.**

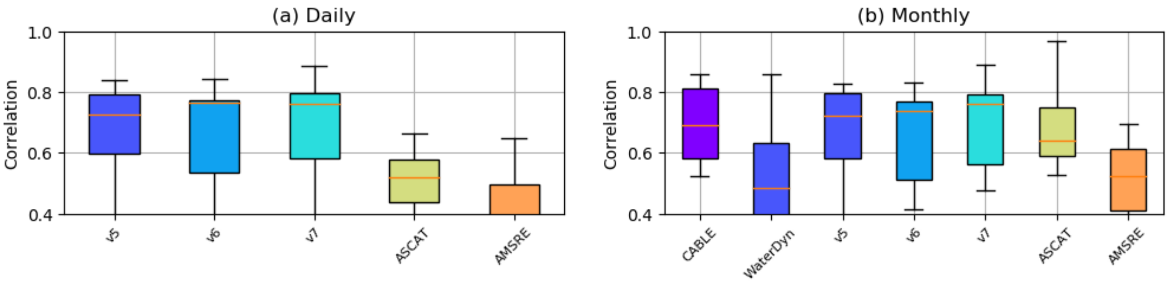



**Figure 18. (a) Daily and (b) Monthly correlation of models against Upper Hunter SASMAS 2007-2011 profile (0-90cm) soil moisture. Satellite data is evaluated over a shorter period (AMSR-E: 2002-2011, ASCAT: 2007-2011) and relates only to the top few cm.**

Figure 19 and Figure 20 shows the correlation of modelled soil moisture with OzFlux and CosmOz observations, respectively. AWRA-L v7 performs equally or better than previous versions for the OzFlux data, and AWRA-L v7 improves over v5 and v6 when evaluated over the CosmOz network.

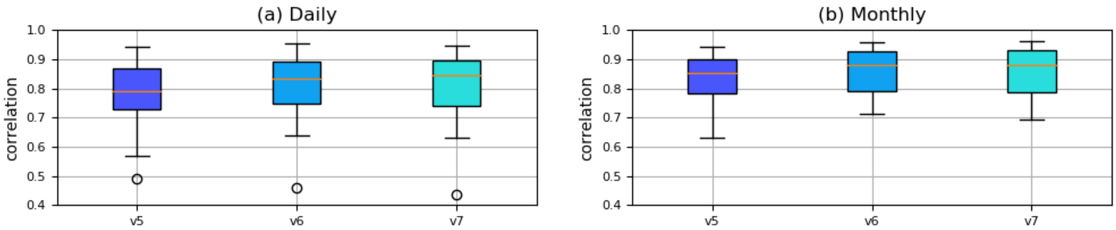



**Figure 19. (a) Daily and (b) Monthly correlation of models against OzFlux 2007-2017 soil profile (0-90cm) for 14 active sites.**

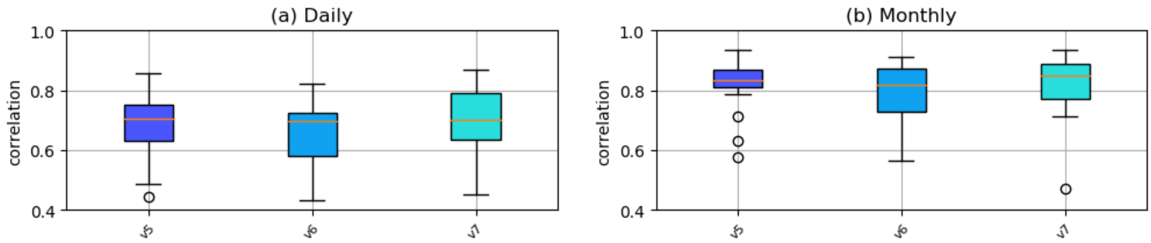



**Figure 20. (a) Daily and (b) Monthly correlation of models against CosmOz 2010-2017 16 tested sites.**

Table 8 and Table 9 present the daily and monthly profile (0-90cm) correlation statistics, for evaluation against the evaluation criteria listed in Table 2, which was 50% of sites with daily and monthly correlation greater than 0.75. AWRA-L is capable of providing estimates of profile soil moisture that successfully meet the aspirational targets defined in Table 2, with median daily and monthly correlation greater than 0.75 against all modelled soil moisture and satellite/land based observations except for CosmOz upper profile soil moisture with a median daily correlation value of 0.70.

**Table 8. Ranked correlation of profile (0-90cm) daily and monthly soil moisture AWRA-L and satellite estimates against OzNet (2001-2013) and SASMAS (2003-2011) data. Noting satellite data is evaluated over a shorter period (AMSR-E: 2002-2011, ASCAT: 2007-2011) – and relates only to the top few cm.**

| OzNet | 0% | 5% | 25% | 50% | 75% | 95% | 100% | SASMAS | 0% | 5% | 25% | 50% | 75% | 95% | 100% |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Daily Soil Moisture | | | | | | | | | | | | | | | |
| **AWRA-L v5** | 0.45 | 0.51 | 0.63 | 0.74 | 0.85 | 0.93 | 0.95 | **AWRA-L v5** | 0.27 | 0.34 | 0.60 | 0.73 | 0.79 | 0.84 | 0.84 |
| **AWRA-L v6** | 0.46 | 0.57 | 0.64 | **0.77** | 0.83 | 0.91 | 0.93 | **AWRA-L v6** | 0.23 | 0.37 | 0.53 | 0.76 | 0.77 | 0.83 | 0.84 |
| **AWRA-L v7** | 0.41 | 0.48 | 0.64 | 0.74 | 0.87 | 0.94 | 0.96 | **AWRA-L v7** | 0.28 | 0.29 | 0.58 | **0.76** | 0.79 | 0.86 | 0.89 |
| **ASCAT** | 0.40 | 0.43 | 0.54 | 0.67 | 0.72 | 0.75 | 0.77 | **ASCAT** | 0.06 | 0.18 | 0.44 | 0.52 | 0.58 | 0.65 | 0.66 |
| **AMSRE** | 0.30 | 0.39 | 0.50 | 0.63 | 0.68 | 0.77 | 0.78 | **AMSRE** | 0.08 | 0.13 | 0.37 | 0.39 | 0.49 | 0.58 | 0.65 |
| **Benchmark** | | | | **0.77** | | | | **Benchmark** | | | | **0.76** | | | |
| Monthly Soil Moisture | | | | | | | | | | | | | | | |
| **CABLE** | 0.38 | 0.52 | 0.67 | 0.79 | 0.86 | 0.90 | 0.93 | **CABLE** | 0.11 | 0.18 | 0.58 | 0.69 | 0.81 | 0.83 | 0.86 |
| **WaterDyn** | 0.29 | 0.39 | 0.61 | 0.77 | 0.83 | 0.85 | 0.90 | **WaterDyn** | 0.17 | 0.25 | 0.36 | 0.49 | 0.63 | 0.75 | 0.86 |
| **AWRA-L v5** | 0.40 | 0.52 | 0.63 | 0.76 | 0.86 | 0.93 | 0.96 | **AWRA-L v5** | 0.14 | 0.23 | 0.58 | 0.72 | 0.80 | 0.82 | 0.83 |
| **AWRA-L v6** | 0.42 | 0.56 | 0.65 | **0.79** | 0.84 | 0.92 | 0.93 | **AWRA-L v6** | 0.10 | 0.29 | 0.51 | 0.74 | 0.77 | 0.82 | 0.83 |
| **AWRA-L v7** | 0.37 | 0.51 | 0.65 | 0.76 | 0.90 | 0.94 | 0.96 | **AWRA-L v7** | 0.12 | 0.14 | 0.56 | **0.76** | 0.79 | 0.87 | 0.89 |
| **ASCAT** | 0.45 | 0.48 | 0.64 | 0.78 | 0.85 | 0.89 | 0.93 | **ASCAT** | 0.53 | 0.55 | 0.59 | 0.64 | 0.75 | 0.91 | 0.97 |
| **AMSRE** | 0.39 | 0.44 | 0.58 | 0.72 | 0.79 | 0.92 | 0.93 | **AMSRE** | 0.00 | 0.12 | 0.41 | 0.52 | 0.61 | 0.69 | 0.70 |
| **Benchmark** | | | | **0.79** | | | | **Benchmark** | | | | **0.76** | | | |

**Table 9. Ranked correlation of profile (0-90cm) daily and monthly soil moisture AWRA-L against OzFlux (2007-2017) and ranked correlation of upper profile daily and monthly soil moisture AWRA-L against CosmOz (2010-2017). Noting that CosmOz soil moisture generally relates to the top 10-30cm.**

| OzFlux | 0% | 5% | 25% | 50% | 75% | 95% | 100% | CosmOz | 0% | 5% | 25% | 50% | 75% | 95% | 100% |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Daily Soil Moisture | | | | | | | | | | | | | | | |
| **AWRA-L v5** | 0.49 | 0.55 | 0.73 | 0.79 | 0.87 | 0.91 | 0.94 | **AWRA-L v5** | 0.44 | 0.48 | 0.63 | 0.71 | 0.75 | 0.82 | 0.86 |
| **AWRA-L v6** | 0.46 | 0.60 | 0.75 | 0.83 | 0.89 | 0.93 | 0.96 | **AWRA-L v6** | 0.43 | 0.49 | 0.58 | 0.70 | 0.73 | 0.80 | 0.82 |
| **AWRA-L v7** | 0.44 | 0.58 | 0.74 | **0.85** | 0.89 | 0.93 | 0.95 | **AWRA-L v7** | 0.45 | 0.45 | 0.64 | **0.70** | 0.79 | 0.82 | 0.87 |
| **Benchmark** | | | | **0.85** | | | | **Benchmark** | | | | **0.70** | | | |
| Monthly Soil Moisture | | | | | | | | | | | | | | | |
| **AWRA-L v5** | 0.63 | 0.70 | 0.78 | 0.85 | 0.90 | 0.93 | 0.94 | **AWRA-L v5** | 0.58 | 0.62 | 0.81 | 0.83 | 0.87 | 0.93 | 0.93 |
| **AWRA-L v6** | 0.71 | 0.71 | 0.79 | 0.88 | 0.93 | 0.96 | 0.96 | **AWRA-L v6** | 0.56 | 0.67 | 0.73 | 0.82 | 0.87 | 0.91 | 0.91 |
| **AWRA-L v7** | 0.70 | 0.71 | 0.79 | **0.88** | 0.93 | 0.95 | 0.96 | **AWRA-L v7** | 0.47 | 0.65 | 0.77 | **0.85** | 0.89 | 0.91 | 0.93 |
| **Benchmark** | | | | **0.88** | | | | **Benchmark** | | | | **0.85** | | | |

The profile layer monthly correlation values are plotted for the OzNet and SASMAS sites (Figure 21) to give an indication of how AWRA-L performs spatially. No spatial patterns in performance are detected, except for a possible slightly lower correlation between 0.25-0.75 rather than above 0.75, for some sites where there is lower saturated conductivity.

Daily timeseries of select sites in each observational dataset are presented to demonstrate the performance of AWRA at reproducing the drying and wetting of the soil at different locations. The timeseries of select Murrumbidgee OzNet sites daily profile (0-90cm), Upper Hunter SASMAS daily profile (0-90cm) sites, Ozflux profile and CosmOz sites are shown in Figure 22 through to Figure 25. It is noted the volumetric soil moisture observations are scaled between 0-100 % based on the maximum and minimum observations in the series to match the AWRA-L percentage full scale. This approach assumes that at some point in the observed record the soil was fully saturated and dried out. The time series of the AWRA soil moisture profile against that observed for the OzNet sites provide an example of how well AWRA-L produces drying and wetting of the soil as experienced during the Millennium drought, particularly the drying years 2006 and 2007 (Potter et al., 2010) and the  wetting in 2010.

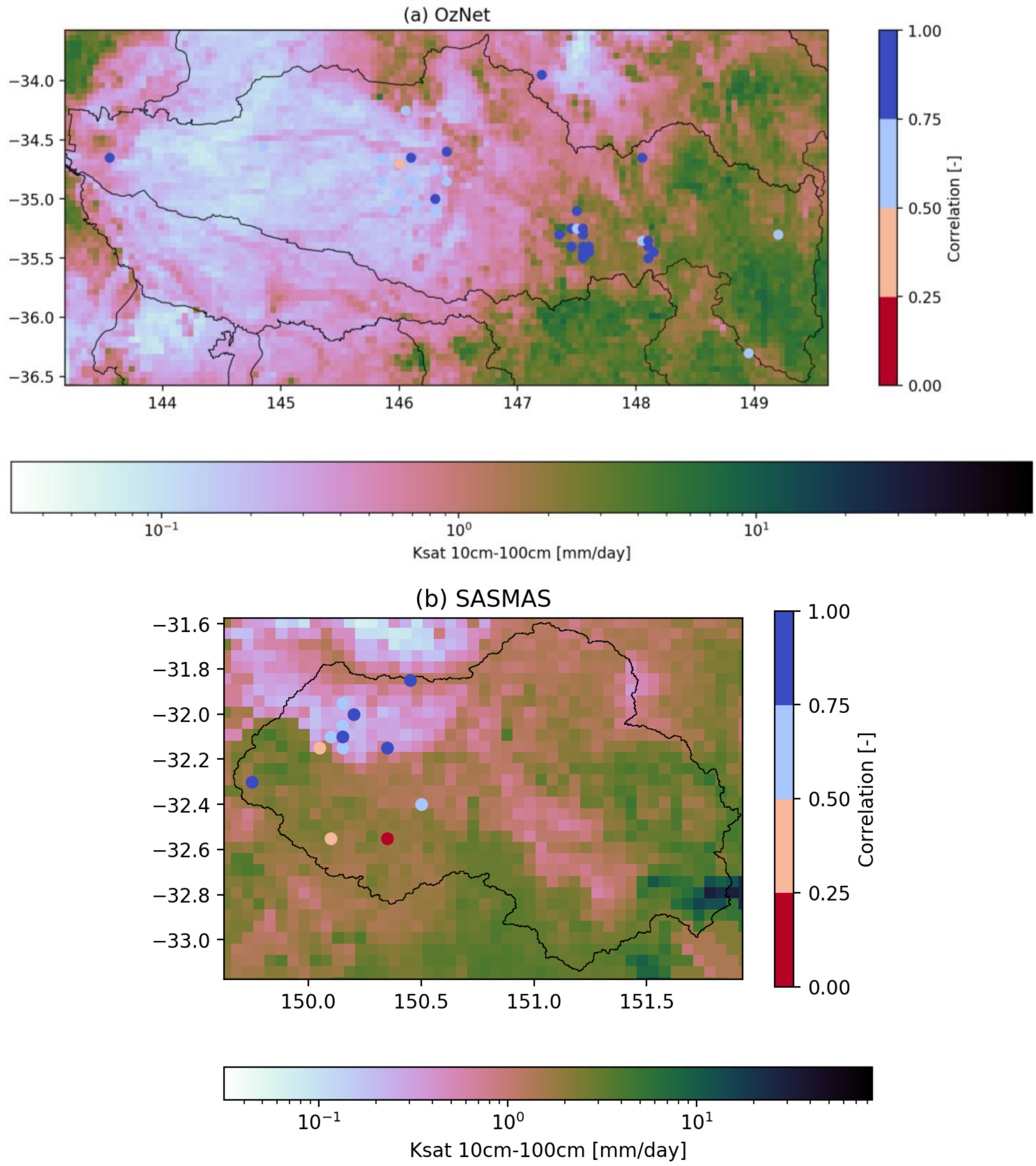



**Figure 21. AWRA-L Monthly correlation for profile soil moisture of a) Murrumbidgee (OzNet) and b) Upper Hunter (SASMAS) data. AWRA-L saturated hydraulic conductivity (Ksat) for shallow layer (10cm-100cm) underlain.**

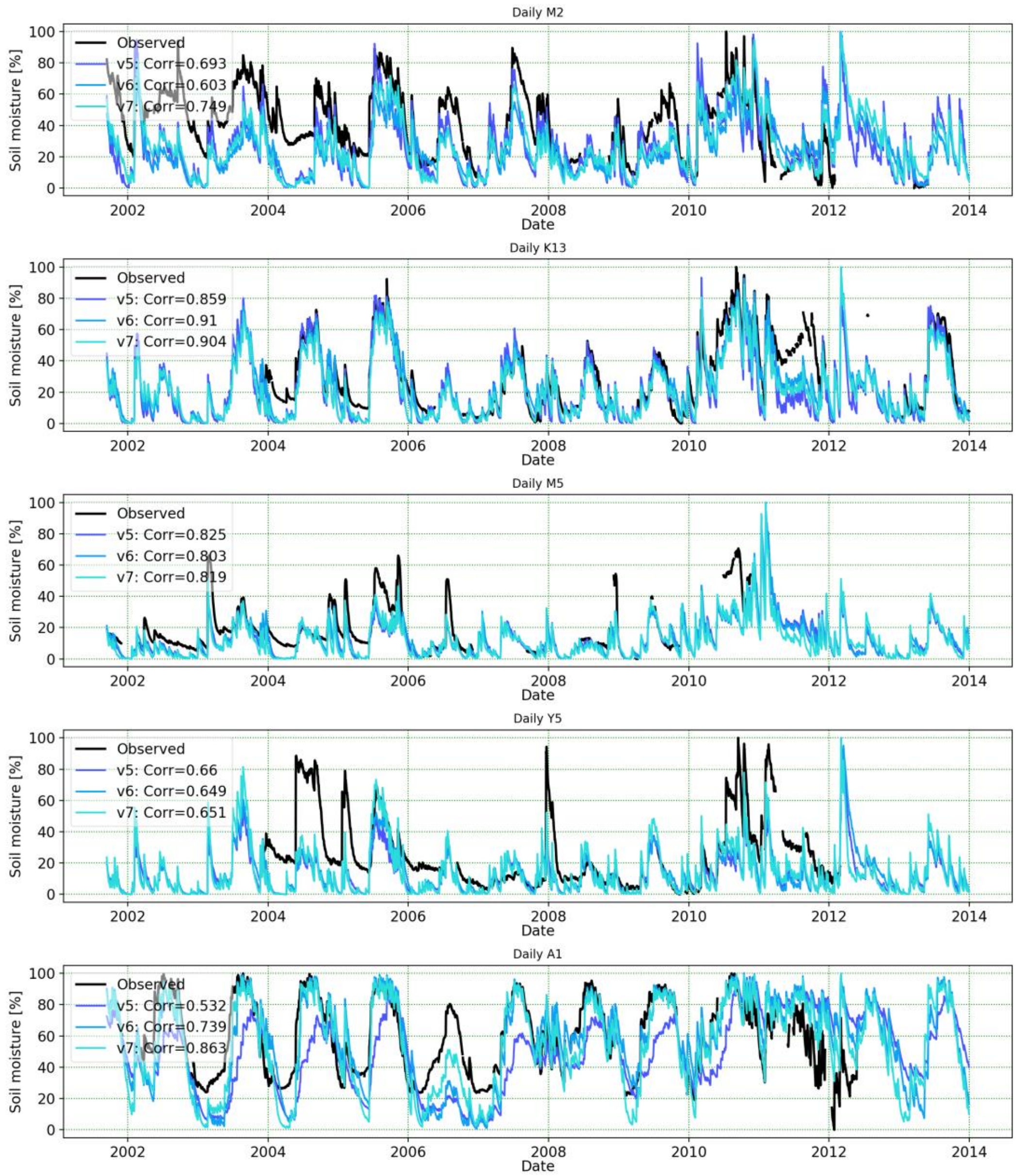



**Figure 22. Five Murrumbidgee OzNet sites daily profile (0-90cm) soil moisture (% full) and model estimates.**

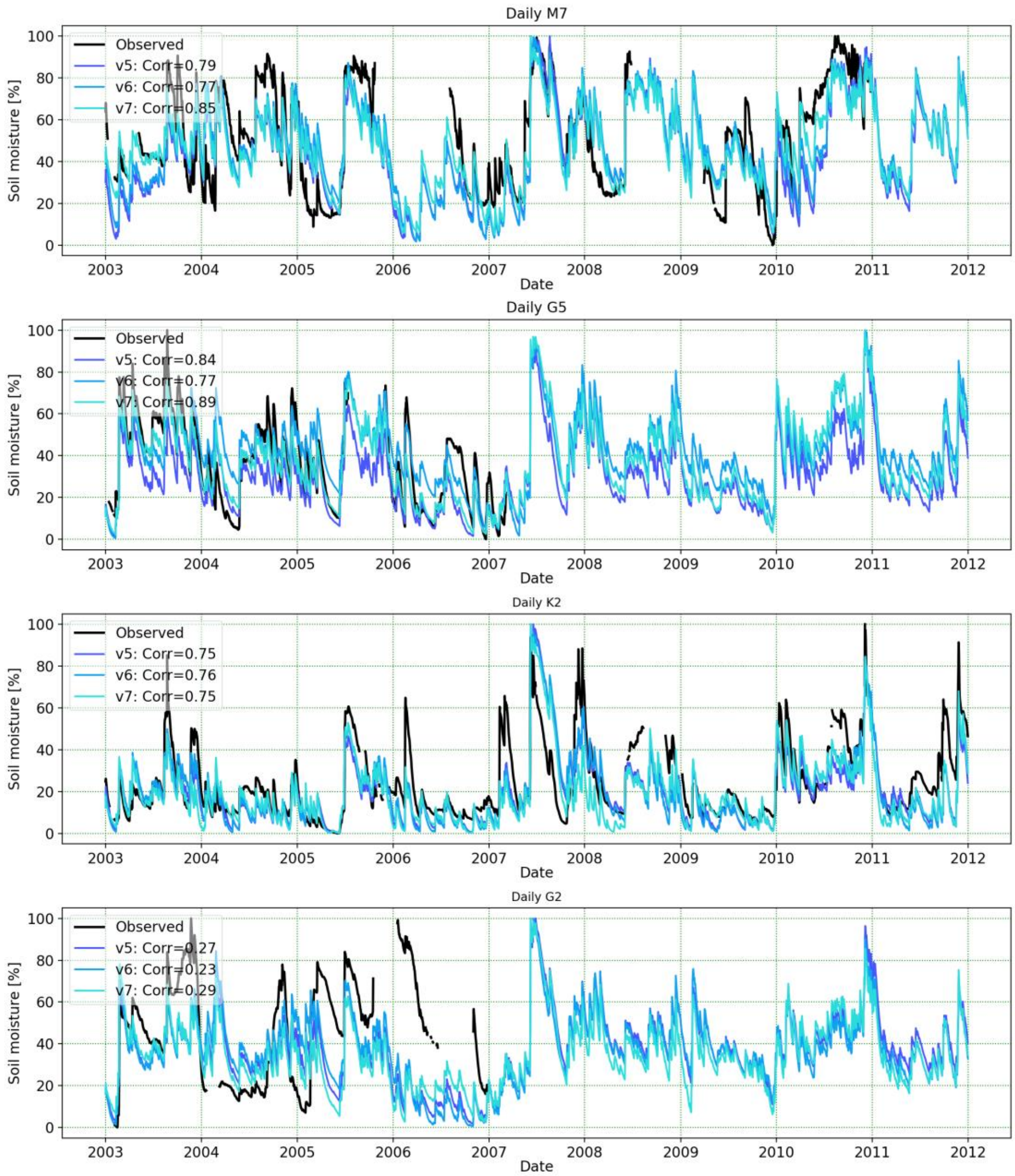



**Figure 23. Four Upper Hunter SASMAS sites daily profile (0-90cm) soil moisture (% full) and model/satellite estimates.**

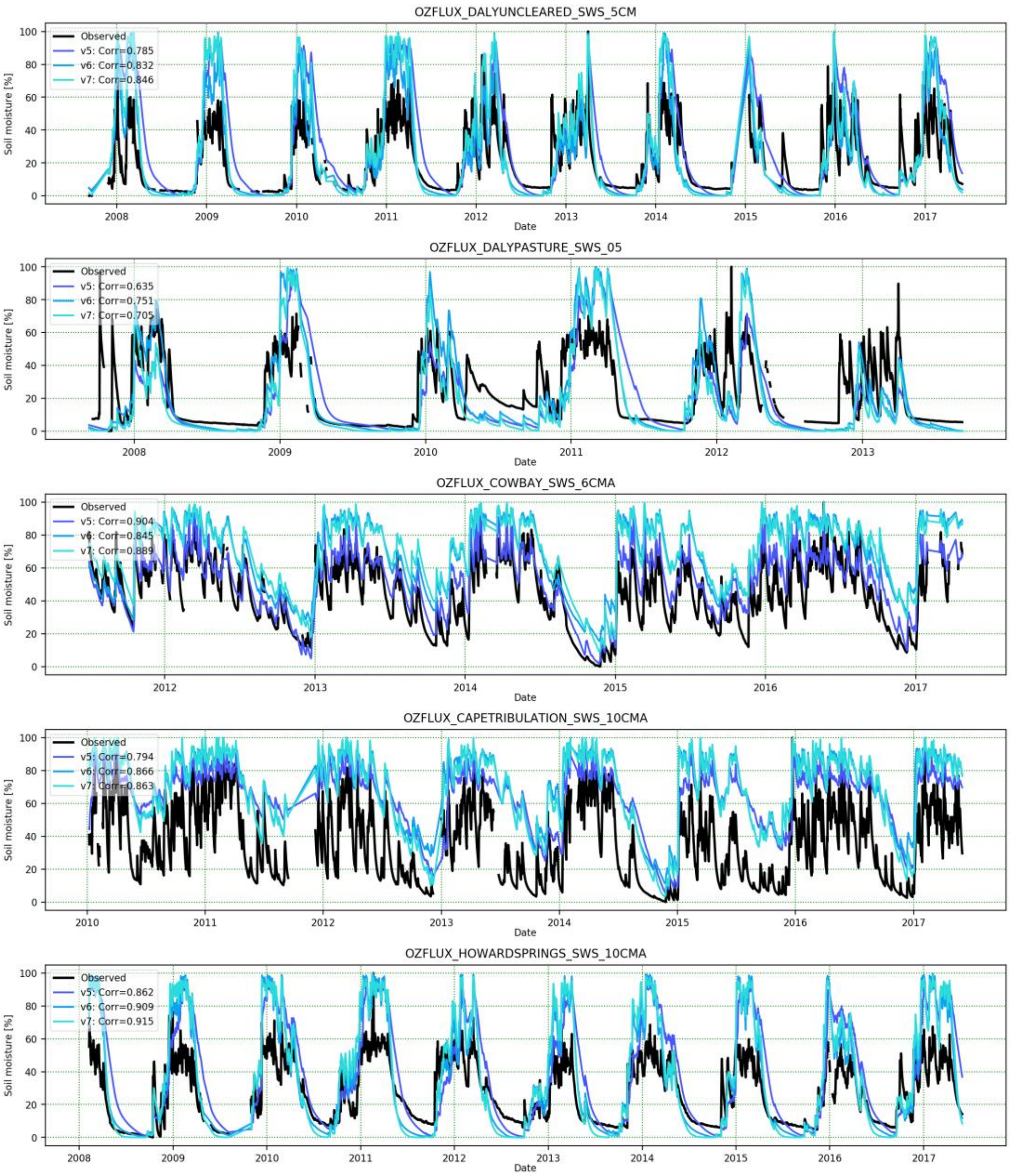



**Figure 24. Five OzFlux sites daily profile (0-90cm) soil moisture (% full) and model/satellite estimates**

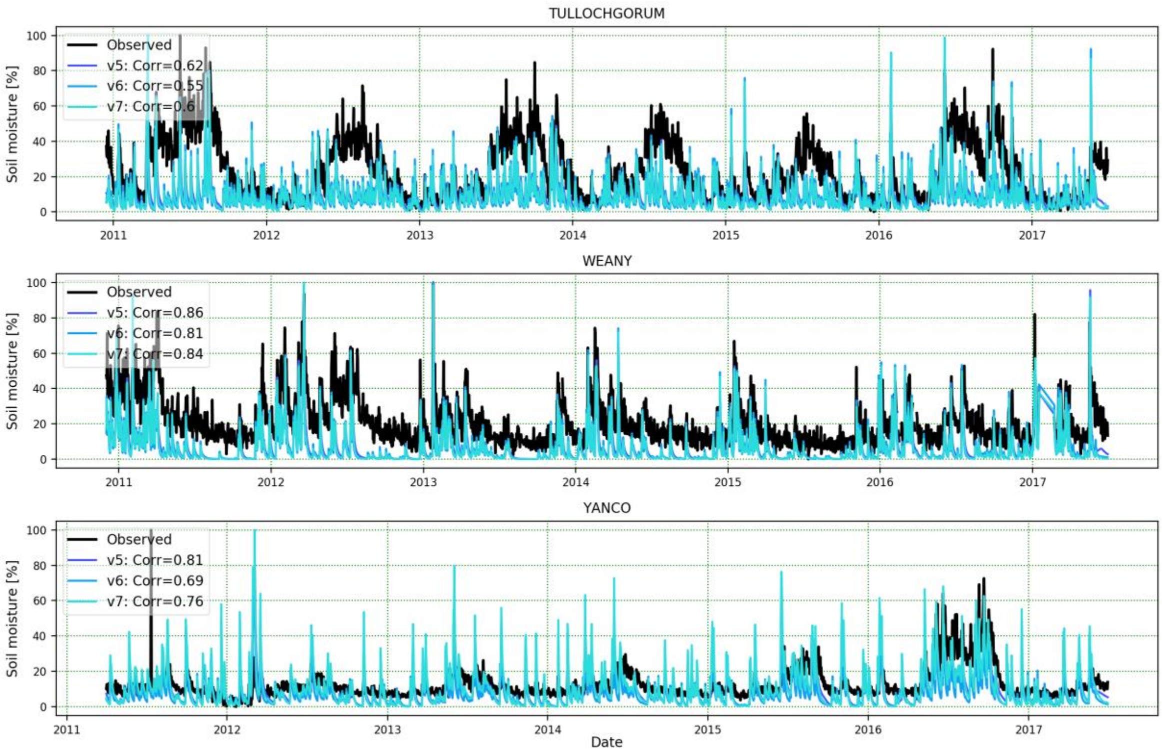



**Figure 25. Three CosmOz sites daily soil moisture and model/satellite estimates**

A brief evaluation of the performance of the models at reproducing the top layer soil moisture is presented in Appendix D. Following the results presented in Frost et al. (2015), AWRA-L performs relatively worse than CABLE and WaterDyn when evaluated against the Oznet Murrumbidgee, SASMAS Upper Hunter and nationwide OzFlux 0cm-5cm/8cm data and satellite based data (see Figs. 57-59), although there is an improvement in AWRA-L v7 compared to AWRA-L v6 due to changes in the model structure.

It is noted that interpretation of the results presented on soil moisture should consider:

- The difference in point scale observations compared to large grid scale (~5 km by 5 km) for the models, larger for satellite data) outputs, with the point not reflecting the sampling area represented by the models.
- Uncertainties in probe calibrations: with some sites being better calibrated than others.

- Inaccuracies of transfer and quality control, with some sites likely to have timing errors and/or the wrong data.
- Inaccuracies in satellite soil moisture product derivations.

It is expected these datasets will improve over time, with further calibration and quality control.

## 5.3 Terrestrial Water Storage (TWS)

The AWRA-L TWS estimates are compared to corresponding GRACE observations across all validation catchments. In addition to monthly correlation (used in calibration), de-seasonalised correlations (where the seasonal trends of timeseries are removed) are also analysed. From Figure 26 and Table 10, AWRA-L performance according to monthly correlation and de-seasonalised correlation at the 50% is 0.60. While AWRA-L v7 improves over its previous versions for correlation and de-seasonalised correlation, it is far off the median aspirational performance for TWS set nominally at 0.75. However, the improvement from version v5 to v7 was considerable. The main reason behind this improvement is the fact that now the TWS observations are a part of objective function in the calibration. It is worth noting that although TWS is now weighted at 50% in calibration, forming the main component of the objective function, other aspects of the water balance have further improved. Most notably streamflow performance has not been degraded, although now weighted relatively less.

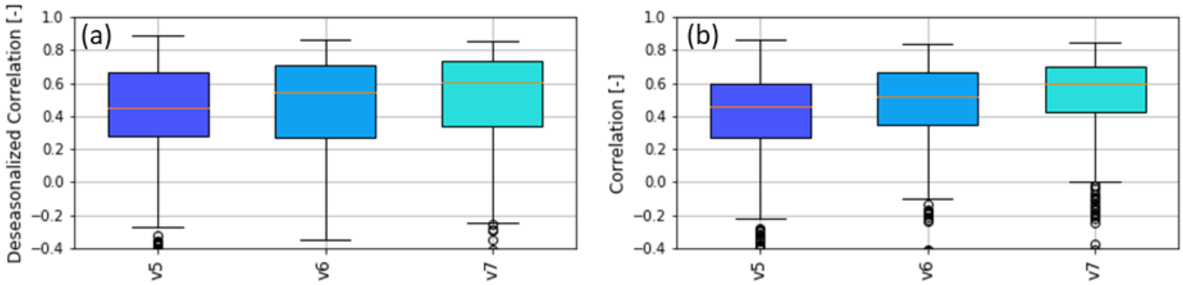



**Figure 26. (a) De-seasonalised Monthly correlation and (b) Monthly correlation of models against Terrestrial Water Storage (TWS) 2002-2017.**

**Table 10. Terrestrial Water Storage (TWS) evaluation criteria**

| | 0% | 5% | 25% | 50% | 75% | 95% | 100% |
|---|---|---|---|---|---|---|---|
| **Models** | **Monthly Correlation** | | | | | | |
| **AWRA-L v5** | -0.56 | -0.10 | 0.27 | 0.46 | 0.60 | 0.75 | 0.86 |
| **AWRA-L v6** | -0.66 | -0.04 | 0.35 | 0.52 | 0.66 | 0.76 | 0.83 |
| **AWRA-L v7** | -0.68 | 0.03 | 0.42 | **0.60** | 0.70 | 0.78 | 0.85 |
| **Benchmark** | | | | **0.60** | | | |
| | **0%** | **5%** | **25%** | **50%** | **75%** | **95%** | **100%** |
| **Models** | **Monthly De-seasonalised Correlation** | | | | | | |
| **AWRA-L v5** | -0.74 | -0.10 | 0.28 | 0.45 | 0.66 | 0.78 | 0.89 |
| **AWRA-L v6** | -0.77 | -0.05 | 0.27 | 0.54 | 0.71 | 0.80 | 0.87 |
| **AWRA-L v7** | -0.77 | 0.05 | 0.34 | **0.60** | 0.73 | 0.81 | 0.86 |
| **Benchmark** | | | | **0.60** | | | |

**Time series comparison of AWRA-L v5 to v7 outputs against GRACE is provided for a range of example catchments in Chapter 6, in Figure 40 to Figure 49. The AWRA-L values correlate reasonably well according to seasonality. The scale of anomaly is roughly equivalent between GRACE TWS and AWRA-L; with the exception of catchments located near Beaudesert, Qld (Figure 41) and Victor harbour, SA (Figure 45), both catchments contains predominantly coastal GRACE pixel (see Fig. 6). Several South Eastern catchments do not replicate the magnitude in shift observed from dry to wet for the large scale flooding that occurred in 2011 following the Millennium drought (Figure 42 Noojee, Vic and Figure 43 Fifteen Mile Creek, Vic). This is similar to findings by Fowler et al. (2020) for standard rainfall-runoff models. It is noted that the model does capture this change in some cases (eg.**

Figure 44 Kyeamba Creek, NSW) and associated streamflow patterns. Seasonal variability is underestimated by AWRA-L for Northern Australia (Figure 48 Dry River and Figure 49 East Baines). Therefore, although the calibration using TWS correlations has improved TWS variability, the calibration approach could be improved by a scale related metric rather than correlations.

## 5.4 Actual Evapotranspiration

The point based estimates of actual ET derived from infilled flux tower data (DINGO) at 25 sites was compared to the CABLE, WaterDyn, AWRA-L, CMRSET and SLST gridded outputs over the entire simulation period from 2001 to 2013,noting that the CMRSET and SLST do not cover this entire period, according to correlation (Figure 27) and relative bias (Figure 28). CABLE and WaterDyn are roughly equal in terms of monthly correlation. AWRA-L v7 improves over previous versions and performs similarly to CABLE and

WaterDyn in terms of monthly correlation of 50th percentile. AWRA-L v7 considerably outperforms previous versions and satellite estimates particularly for 50th percentile and above at the daily timescale, predominantly due to changes in alignment of climate inputs. However, CABLE and WaterDyn perform better according to monthly bias than AWRA-L. CABLE is expected to perform best here, as: (a) it is calibrated to the Tumbarumba, Howard Springs and Virginia Park ET (albeit over a different time period), while the other models are not, and (b) it contains a more complete formulation of land-surface energy and water related dynamics.

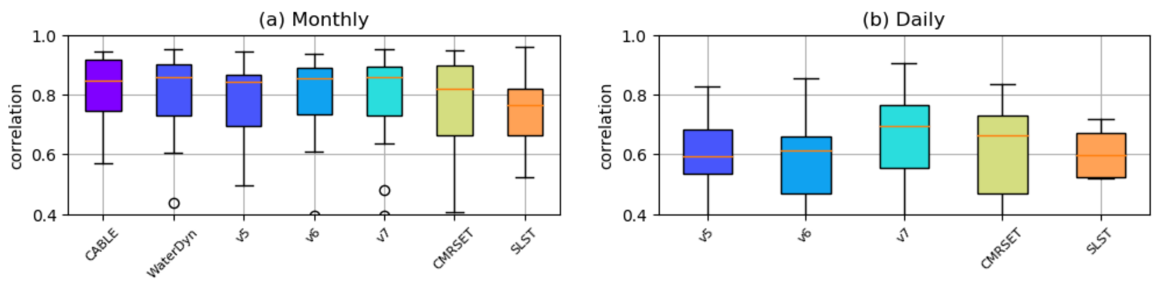



**Figure 27. Correlation over 2001-2013 of flux tower actual ET compared to modelled (a) Monthly and (b) Daily data**

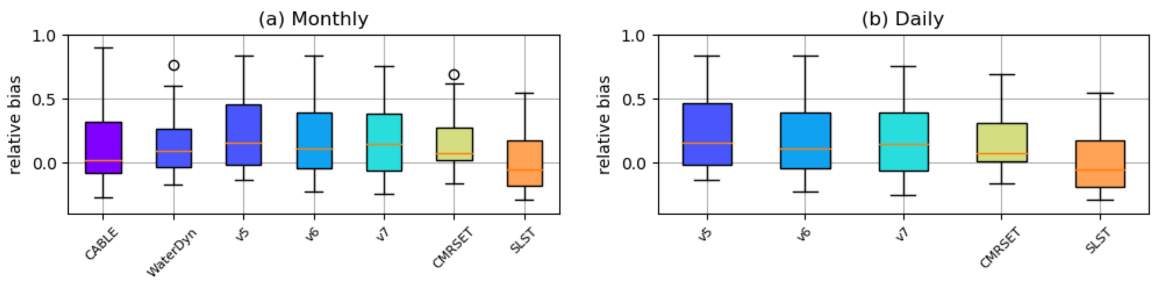



Figure 28. Relative bias over 2001-2013 of flux tower actual ET compared to modelled (a) Monthly and (b) Daily dataTable 11 and Table 12 present the monthly and daily correlation and bias statistics, for evaluation against the criteria listed in Table 2. AWRA-L v7 improves over previous versions and satellite observations and perform equivalently to WaterDyn in terms of 50th percentile meeting the target criteria (greater than 0.8), but experience degradation compared to WaterDyn particularly at the 5th percentile. Moreover, the performance of this version dropped compared to v6 regarding monthly

relative bias where lower bias is detected for Cable and WaterDyn. Overall, WaterDyn provides benchmark for future performance testing. In terms of daily correlation, AWRA-L v7 outperforms AWRA-L v5 and v6 particularly in 5th and 50th percentiles.

**Table 11. Monthly (a) correlation and (b) relative bias of modelled estimates compared to DINGO data 2001-2013. Noting satellite-based estimates CMRSET and SLST do not cover the same period as models.**

| Correlation | 0% | 5% | 25% | 50% | 75% | 95% | 100% | Relative bias | 0% | 5% | 25% | 50% | 75% | 95% | 100% |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **CABLE** | -0.03 | 0.32 | 0.75 | 0.85 | 0.92 | 0.94 | 0.95 | **CABLE** | -0.28 | -0.17 | -0.08 | 0.02 | 0.31 | 0.71 | 0.90 |
| **WaterDyn** | 0.44 | 0.61 | 0.73 | 0.86 | 0.90 | 0.95 | 0.95 | **WaterDyn** | -0.18 | -0.15 | -0.04 | 0.08 | 0.26 | 0.60 | 0.76 |
| **AWRA-L v5** | 0.26 | **0.50** | 0.69 | 0.84 | 0.87 | 0.93 | 0.94 | **AWRA-L v5** | -0.14 | -0.11 | -0.02 | 0.15 | 0.46 | 0.70 | 0.83 |
| **AWRA-L v6** | 0.26 | 0.44 | 0.73 | 0.85 | 0.89 | 0.92 | 0.94 | **AWRA-L v6** | -0.23 | -0.15 | -0.04 | **0.11** | 0.39 | 0.79 | 0.84 |
| **AWRA-L v7** | 0.28 | 0.38 | 0.73 | **0.86** | 0.89 | 0.92 | 0.95 | **AWRA-L v7** | -0.25 | -0.20 | -0.06 | 0.14 | 0.38 | 0.72 | 0.75 |
| **CMRSET** | 0.41 | 0.59 | 0.66 | 0.81 | 0.90 | 0.94 | 0.95 | **CMRSET** | -0.17 | -0.11 | -0.01 | 0.06 | 0.26 | 0.61 | 0.69 |
| **SLST** | 0.17 | 0.40 | 0.67 | 0.77 | 0.82 | 0.92 | 0.96 | **SLST** | -0.29 | -0.27 | -0.19 | -0.06 | 0.17 | 0.52 | 0.54 |
| **Benchmark** | | **0.50** | | **0.86** | | | | **Benchmark** | | | | **0.11** | | | |

**Table 12. Daily correlation of AWRA-L compared to DINGO data 2001-2013.**

| Correlation | 0% | 5% | 25% | 50% | 75% | 95% | 100% |
|---|---|---|---|---|---|---|---|
| **AWRAL v5** | 0.26 | 0.30 | 0.54 | 0.59 | 0.68 | 0.81 | 0.83 |
| **AWRAL v6** | 0.20 | 0.23 | 0.47 | 0.61 | 0.66 | 0.83 | 0.85 |
| **AWRAL v7** | 0.29 | 0.32 | 0.56 | **0.69** | 0.77 | 0.84 | 0.90 |
| **CMRSET** | 0.31 | 0.39 | 0.47 | 0.66 | 0.73 | 0.80 | 0.83 |
| **SLST** | 0.21 | 0.24 | 0.53 | 0.60 | 0.67 | 0.71 | 0.72 |
| **Benchmark** | | | | **0.69** | | | |

A second comparison (Figure 29) was undertaken using the time period that the satellite ET data was available (2000-2013). This gives an indication of how well the satellite data represents ET, compared to the three models. WaterDyn, CABLE and AWRA-L are superior in terms of median monthly correlation, although CMRSET produces some correlations at the high deciles that are higher than the models, while SLST performs relatively poorly. This suggests that this version of CMRSET provides some value in terms of a dataset that can be used for evaluation, calibration and assimilation into AWRA-L. However, given that some correlations in the lowest performing sites (25th percentile) are below that of AWRA-L, the use of the data in calibration may be diminishing ET performance in some cases.

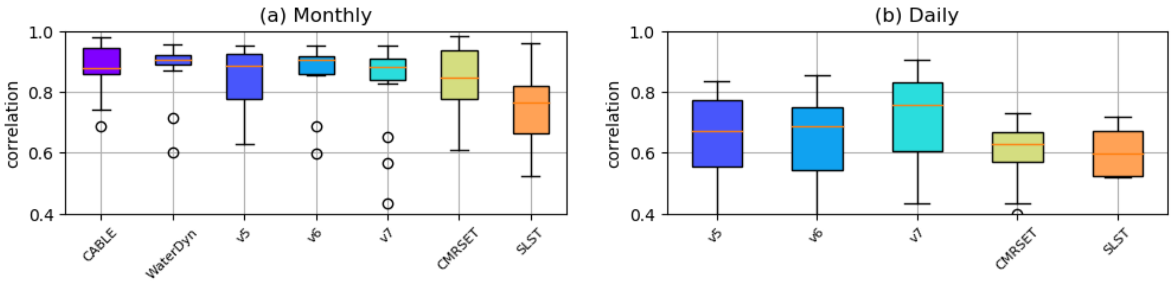



**Figure 29. Correlation over 2001-2010 of flux tower actual ET compared to modelled (a) Monthly and (b) Daily data**

Figure 30 shows the spatial plots of AWRA-L (a) correlation and (b) relative bias compared to the DINGO ET data. Spatially we see that AWRA-L overestimates in some areas for several sites in central Australia and eastern Australia when compared to DINGO ET, however the remaining sites compare favourably.

Finally, it is noted that there is significant uncertainty associated in closing energy balance from flux tower data. Wilson et al. (2002) carried out a comprehensive evaluation of energy balance closure across 22 sites using eddy covariance flux towers ranging from Mediterranean to temperate and arctic climate. Results indicated a general lack of energy balance closure at most sites, with a mean imbalance in the order of 20%. Further, the infilling procedure used here for infilling also has uncertainties. In particular, the method used to infill data up until the start of the calendar year,before the flux tower observations start, shows significant uncertainty (e.g. Cumberland in early 2011 before start in September 2011 – see Figure 52 in Appendix B: ET monitoring site details and time series). In general, evapotranspiration is difficult to definitively measure and all comparisons are therefore indicative only.

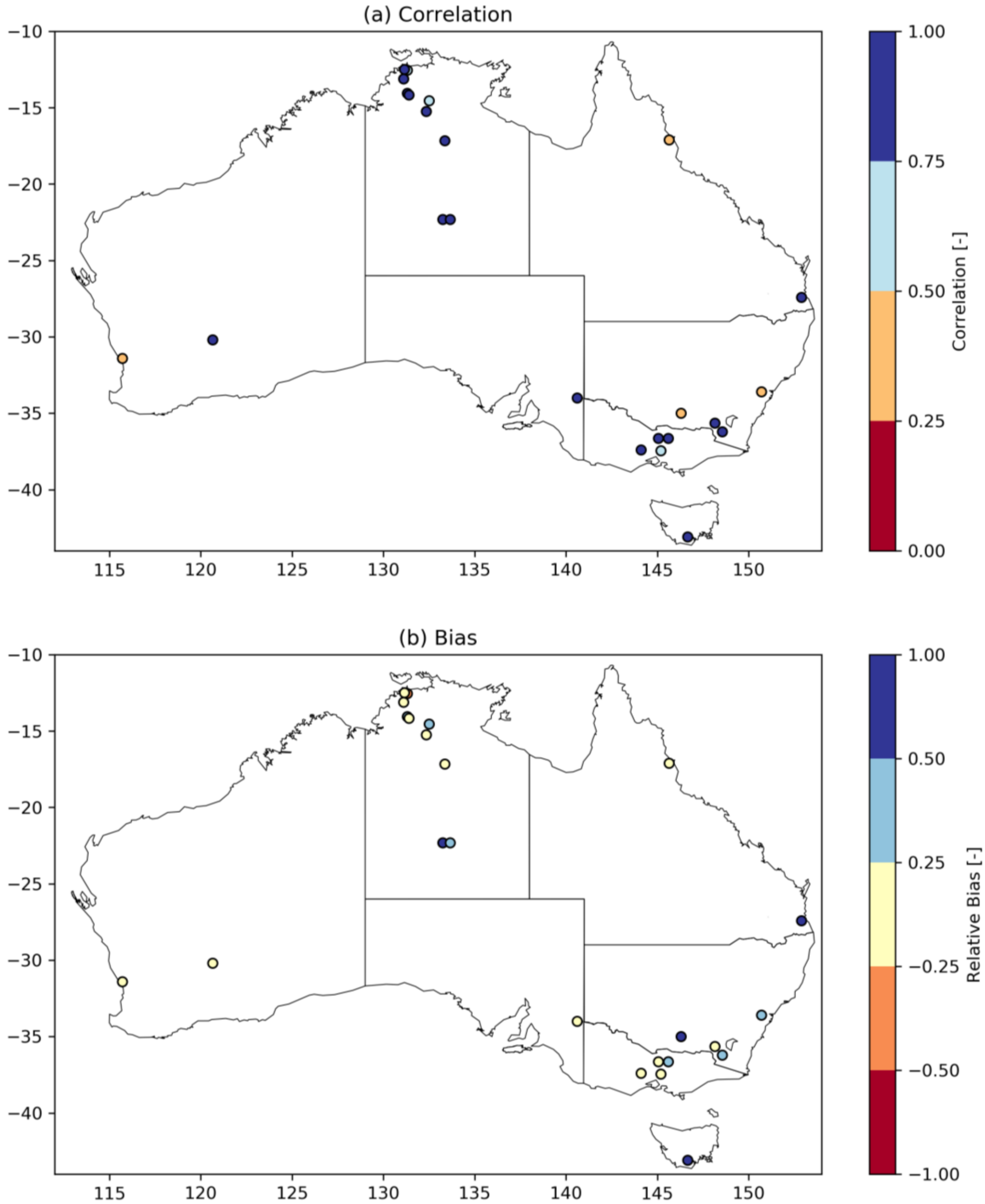



**Figure 30. AWRA-L ET monthly (a) correlation and (b) bias compared with flux tower measurements of ET**

## 5.5 Vegetation Fraction

The monthly estimates of AWRA-L vegetation fraction (Fveg) was assessed against satellite remote sensing data (derived from MODIS). With reference to Figure 31 and Table 13 the latest version of AWRA-L performs best (for all percentiles) monthly correlation and monthly de-seasonalised correlations compared to the satellite observations. The improvement in the vegetation estimation of the model is attributed to adding vegetation fraction to the calibration along with manual tuning of some vegetation growth parameters to reproduce observed dynamics. Using the new observation in the calibration process, the model performs better in this area.

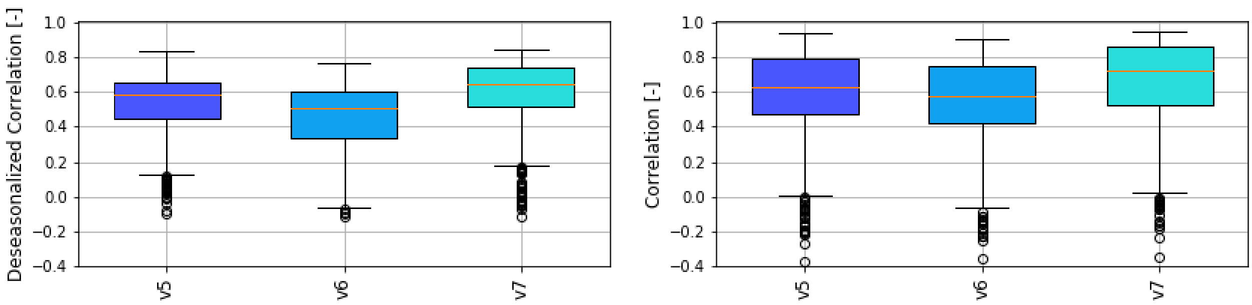



**Figure 31. (a) De-seasonalised Monthly correlation and (b) Monthly correlation of modelled Vegetation Fraction (Fveg) compared to satellite derived Fveg 2002-2017.**

**Table 13. Vegetation Fraction (Fveg) evaluation criteria**

| | 0% | 5% | 25% | 50% | 75% | 95% | 100% |
|---|---|---|---|---|---|---|---|
| **Models** | **Monthly Correlation** | | | | | | |
| **AWRA-L v5** | -0.46 | 0.09 | 0.48 | 0.63 | 0.79 | 0.90 | 0.94 |
| **AWRA-L v6** | -0.36 | 0.12 | 0.42 | 0.57 | 0.75 | 0.86 | 0.91 |
| **AWRA-L v7** | -0.35 | 0.19 | 0.53 | **0.72** | 0.86 | 0.92 | 0.95 |
| **Benchmark** | | | | **0.72** | | | |
| | **0%** | **5%** | **25%** | **50%** | **75%** | **95%** | **100%** |
| **Models** | **Monthly De-seasonalised Correlation** | | | | | | |
| **AWRA-L v5** | -0.10 | 0.16 | 0.45 | 0.58 | 0.66 | 0.73 | 0.83 |
| **AWRA-L v6** | -0.11 | 0.09 | 0.34 | 0.51 | 0.60 | 0.68 | 0.76 |
| **AWRA-L v7** | -0.11 | 0.18 | 0.51 | **0.65** | 0.74 | 0.80 | 0.84 |
| **Benchmark** | | | | **0.65** | | | |

Time series comparison of AWRA-L v5 to v7 outputs against monthly Fveg and other available data are provided for a range of example catchments in Chapter 6, in Figure 40 to Figure 49. Although Fveg correlation statistics are reasonable, the range of Fveg values in v7 is typically low compared to the MODIS based estimate. This is likely due to vegetation parameters being optimised and set to unrealistic values (as noted in Frost and Shokri, 2021), and is an area of potential future improvement.

## 5.6 Groundwater deep drainage

Modelled deep drainage is compared against the long term average national collated recharge dataset covering 2282 grid cells, with relative bias calculated (Figure 32) along with an overall correlation value. Modelled deep drainage was also compared to the annual time series recharge dataset spanning 1970-2012 covering 438 sites using the water table fluctuation method (Figure 33); with annual correlation and relative bias presented.

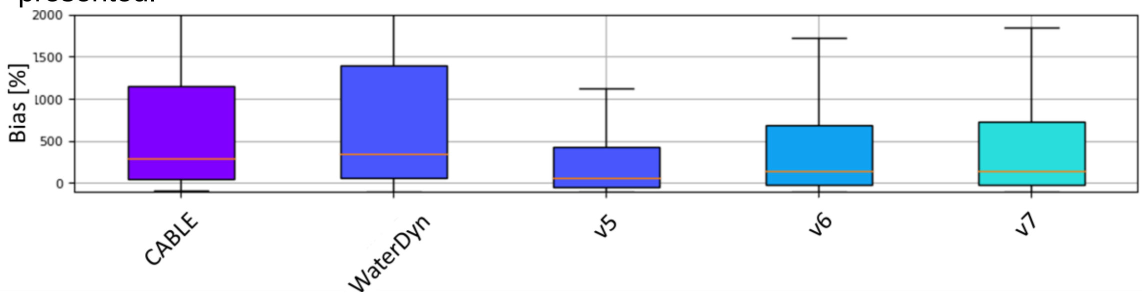



**Figure 32. Modelled outputs versus long term average recharge dataset (2282 grid cells across Australia) relative bias**

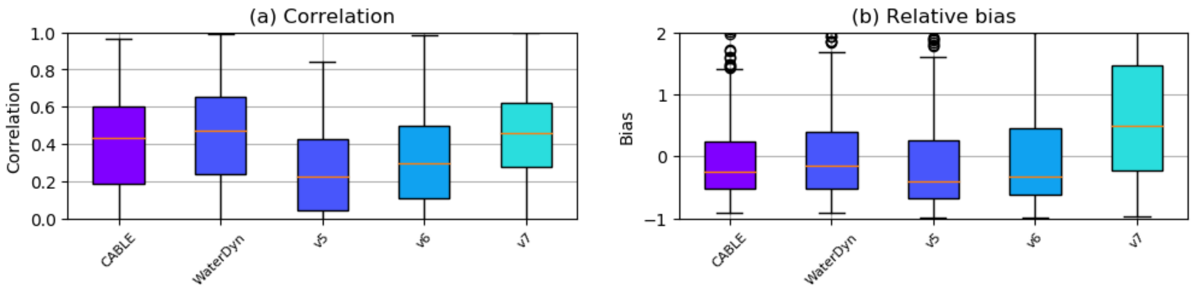



**Figure 33. Modelled outputs versus annual recharge dataset (438 sites in South Australia) (a) correlation and (b) relative bias**

With reference to Figure 32 and Table 14, AWRA-L performs well comparatively against the national long term average recharge dataset, with a low median bias and a reasonable spatial correlation of ~0.5 (noting WaterDyn and CABLE have values of ~0.65).

AWRA-L v7 improves significantly over previous version in term of correlation to annual timeseries (Figure 33a), and it performs equivalent to CABLE/WaterDyn. However, AWRA-L is positively biased (Figure 33b), where other models are negatively biased. The reasons for improved AWRA-L v7 groundwater performance is attributed to calibration to total water storage and vegetation, along with the changes to structure for improved soil drainage, baseflow and transpiration.

The three models are now compared against the specified evaluation criteria:

- Spatial correlation with long term average data above 0.5
- Bias for at least 25% of the long-term average sites to be below zero and bias for at least 25% to be greater than zero. Ideally it is expected that the number of positive and negative biased catchments to be equal.
- Annual correlation (for annual data) of at least 50% of sites to be greater than 0.5.

AWRA-L accords with the bias constraint, where the other models do not. All models do not achieve the aspirational target of 0.5 median annual correlation, while all are above the threshold of 0.5 spatial correlation with the long term average dataset.

**Table 14. Deep drainage evaluation criteria**

| | 0% | 5% | 25% | 50% | 75% | 95% | 100% | Long term average spatial correlation |
|---|---|---|---|---|---|---|---|---|
| **Model** | **National Long-Term Average Dataset relative bias** | | | | | | | |
| **CABLE** | -0.94 | -0.50 | 0.48 | 2.93 | 11.50 | 85.40 | 1827.24 | 0.66 |
| **WaterDyn** | -0.97 | -0.50 | 0.60 | 3.40 | 13.98 | 111.69 | 1514.61 | 0.64 |
| **AWRA-L v5** | -1.00 | -0.87 | -0.49 | 0.62 | **4.23** | 41.70 | 1196.93 | 0.49 |
| **AWRA-L v6** | -1.00 | -0.76 | **-0.20** | 1.42 | 6.81 | 58.17 | 2149.96 | 0.51 |
| **AWRA-L v7** | -1.00 | -0.82 | -0.21 | 1.38 | 7.26 | 56.33 | 1211.79 | **0.52** |
| **Benchmark** | | | **-0.20** | | **4.23** | | | **0.52** |
| **Model** | **Annual time series correlation** | | | | | | | |
| **CABLE** | -0.76 | -0.29 | 0.18 | 0.44 | 0.60 | 0.78 | 0.97 | |
| **WaterDyn** | -0.61 | -0.10 | 0.24 | 0.47 | 0.66 | 0.84 | 0.99 | |
| **AWRA-L v5** | -1.00 | -0.32 | 0.04 | 0.22 | 0.43 | 0.68 | 0.84 | |
| **AWRA-L v6** | -0.89 | -0.18 | 0.11 | 0.29 | 0.50 | 0.71 | 0.99 | |
| **AWRA-L v7** | -0.75 | -0.05 | 0.28 | **0.46** | 0.62 | 0.83 | 1.00 | |
| **Benchmark** | | | | **0.46** | | | | |

Example annual time series for two sites are presented for the 3 versions of AWRA-L model in Figure 34. This plot gives an indication of the variability between models and data.

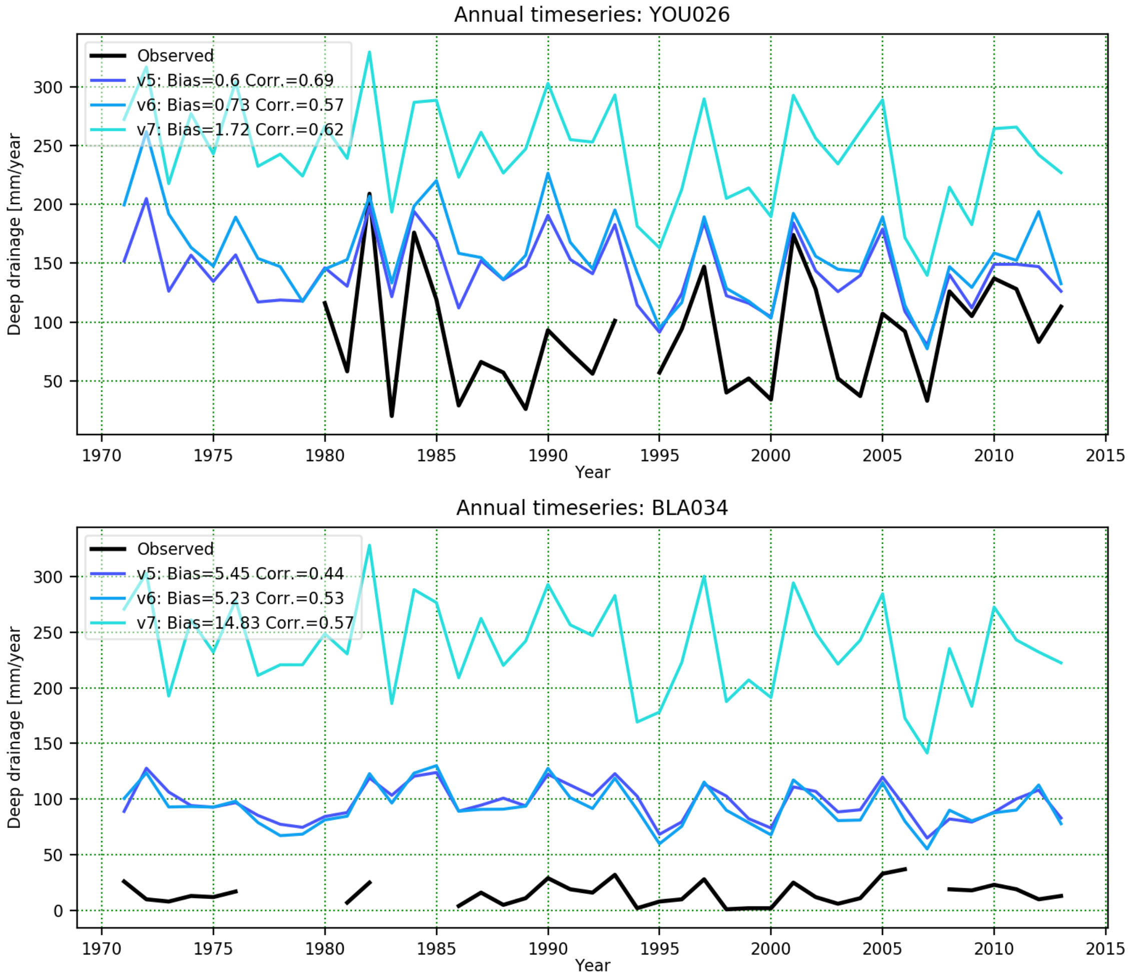



**Figure 34. Example annual deep drainage time-series for two sites.**

Figure 35 shows the relative bias value of the AWRA-L model compared to the long term average data Australia wide; with the AWRA-L v7 shallow layer saturated conductivity underlain. Figure 36 shows the AWRA-L performance according to the annual data spatially, overlaid on the AWRA-L shallow layer storage and saturated conductivity. The recharge values tend to be biased positively in areas where the saturated degree of conductivity is high, and negatively biased where saturated conductivity is low.

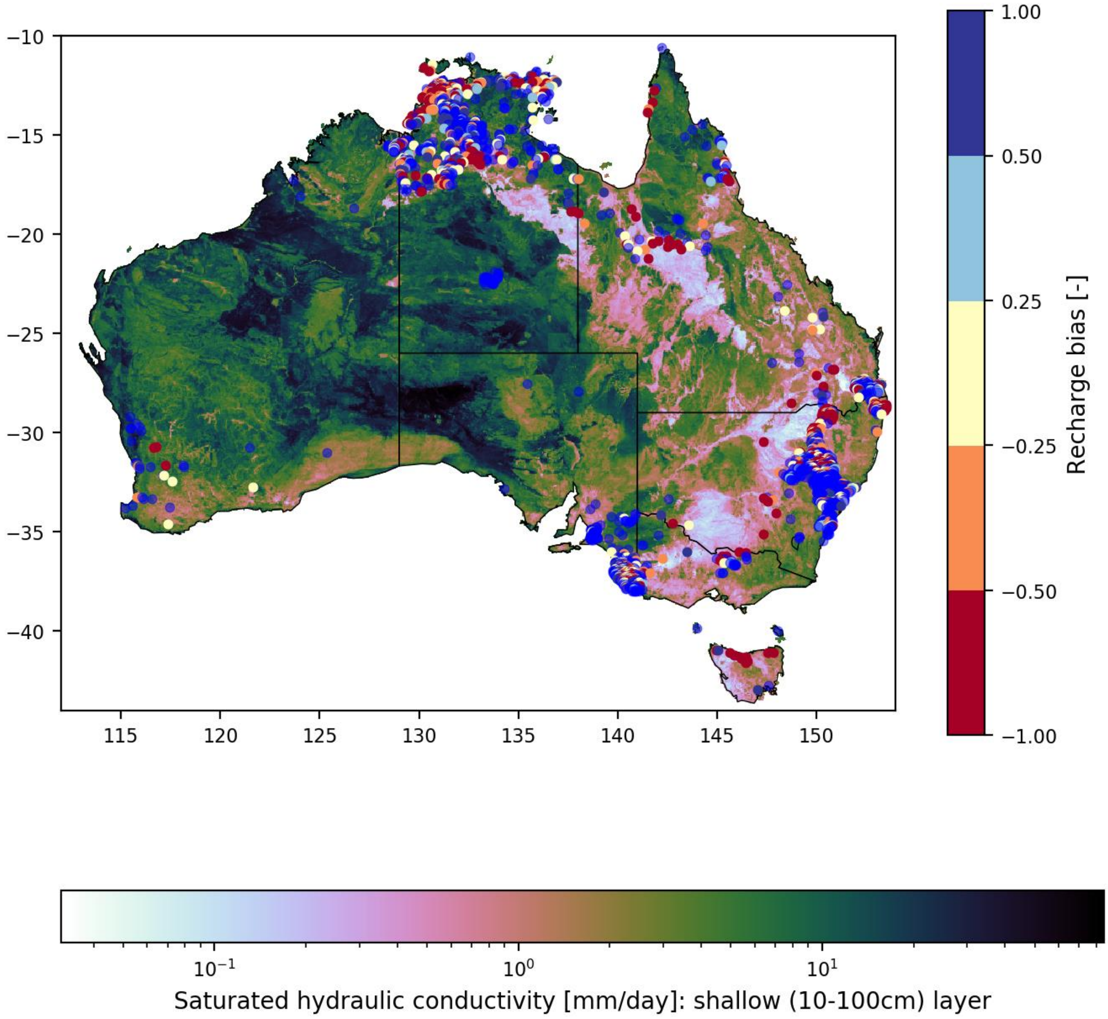



**Figure 35. AWRA-L relative bias of deep drainage compared to long term average estimates over Australia. AWRA-L v7 shallow layer soil saturated hydraulic conductivity (Ksat) is also mapped.**

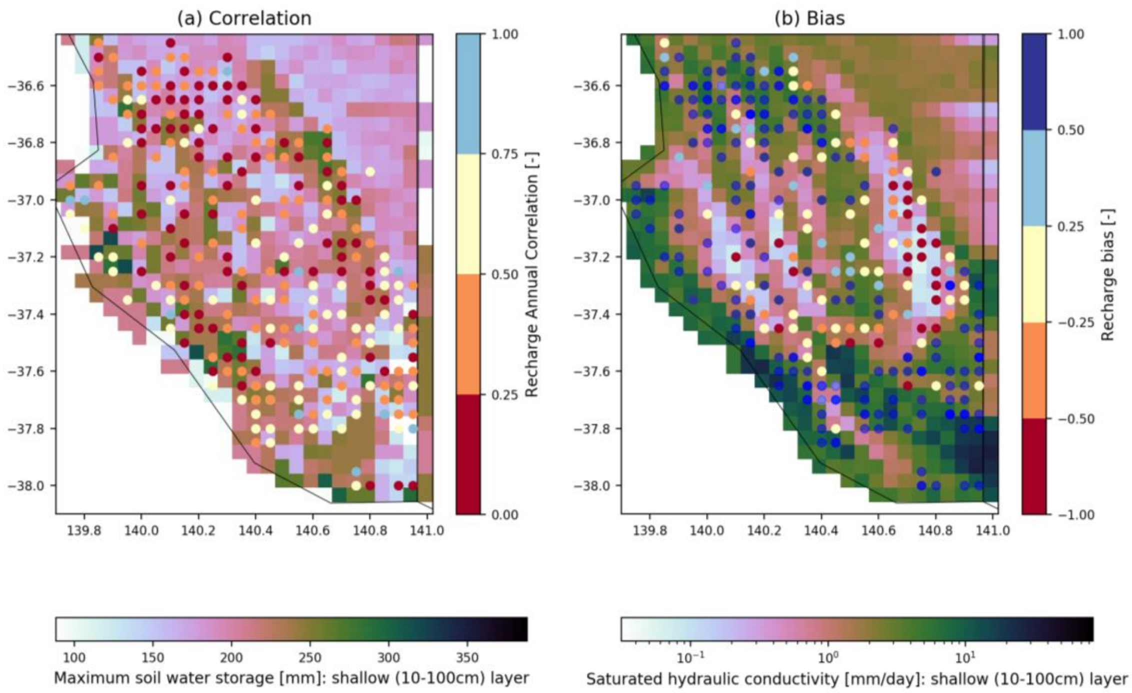



**Figure 36. AWRA-L correlation and relative bias of deep drainage compared to annual time series estimates over South Eastern South Australia. AWRA-L v7 shallow layer maximum soil storage (Ssmax) and saturated hydraulic conductivity (Ksat) is also mapped.**

These results appear to indicate that the saturated conductivity layer is controlling drainage too strongly; and other factors such as rainfall gradient and variability are not affecting drainage enough. This repeats similar results previously found for AWRA-L v5 and v6, although the degree of bias is lower.

It is noted that this timespan used here in evaluation is much smaller than that estimated according to the long term average through chloride mass balance, and also that land use changes may mean that the long term averages are not representative for the period compared. The simulation period considered was 1970-2012, consistent with the evaluation against recharge annual time series span. However, the pattern of bias against the long-term recharge dataset is consistent nationally.

## 5.7 Summary according to benchmark statistics

Table 15 shows a summary of the performance of all the models considered against all the available observation data sets across the water balance. This table provides a quick overview of the trade-offs in performance between different models and the opportunity to compare all models for overall performance across the water balance where the performances are compared to the aspirational targets for critical percentile. Based on these results, AWRA-L v6 was deemed to improve on AWRA-L v5 due to the improvements in soil moisture, recharge and runoff performance. Soil moisture and runoff were considered the focus variables since they are the most requested datasets and most used by the Bureau. The current version, AWRA-L v7, improves over previous versions due to the improvements in streamflow, soil moisture, evapotranspiration, vegetation and groundwater performance for most of the evaluation criteria.

**Table 15. Performance according to benchmark validation statistics. Percentile indicates the ranked site value for a given statistic. The red, white and blue colouring indicates the rank of the model according to the statistic. Dark blue indicates highest rank, white middle rank, dark red lowest rank. *Note: satellite-based estimates (ASCAT, AMSRE, CMRSET, SLST) do not cover full time range of observed data.**

| Data and statistic | Percentile | Target | Best | CABLE | WaterDyn | v5 | v6 | v7 | GR4J | Sacramento | ASCAT | AMSRE | CMRSET | SLST |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Streamflow (291 validation catchments) | | | | | | | | | | | | | | |
| Daily NSE | 0.05 | 0 | -0.06 | | | -0.27 | -0.06 | -1.07 | -0.60 | -1.55 | | | | |
| | 0.5 | 0.5 | 0.56 | | | 0.45 | 0.49 | 0.5 | 0.56 | 0.56 | | | | |
| Monthly NSE | 0.05 | 0 | -0.24 | -0.43 | -2.26 | -0.33 | -0.19 | -0.86 | -0.24 | -0.31 | | | | |
| | 0.5 | 0.5 | 0.74 | 0.3 | 0.6 | 0.67 | 0.68 | 0.7 | 0.73 | 0.74 | | | | |
| Monthly Relative Bias | 0.05 | >-0.5 | -0.44 | -0.53 | -0.46 | -0.47 | -0.44 | -0.44 | -0.59 | -0.52 | | | | |
| | 0.25 | >-0.3 | -0.16 | -0.26 | -0.16 | -0.18 | -0.21 | -0.19 | -0.19 | -0.19 | | | | |
| | 0.75 | <0.3 | 0.23 | 0.33 | 0.59 | 0.32 | 0.29 | 0.23 | 0.30 | 0.31 | | | | |
| | 0.95 | <1.5 | 1.3 | 1.55 | 2.53 | 1.41 | 1.44 | 1.3 | 1.35 | 1.34 | | | | |
| Streamflow (13 impervious catchments) | | | | | | | | | | | | | | |
| Daily NSE | 0.05 | 0 | 0.25 | | | 0.25 | 0.18 | 0.22 | | | | | | |
| | 0.5 | 0.5 | 0.53 | | | 0.48 | 0.41 | 0.53 | | | | | | |
| Monthly NSE | 0.05 | 0 | 0.41 | | | 0.41 | 0.27 | 0.28 | | | | | | |
| | 0.5 | 0.5 | 0.72 | | | 0.76 | 0.68 | 0.72 | | | | | | |
| Soil moisture | | | | | | | | | | | | | | |
| SASMAS 0-90cm Daily correlation | 0.5 | 0.75 | 0.76 | | | 0.73 | 0.76 | 0.76 | | | 0.52 | 0.39 | | |
| SASMAS 0-90cm Monthly correlation | 0.5 | 0.75 | 0.76 | 0.69 | 0.49 | 0.72 | 0.74 | 0.76 | | | 0.64 | 0.52 | | |
| OzNet 0-90cm Daily correlation | 0.5 | 0.75 | 0.77 | | | 0.74 | 0.77 | 0.74 | | | 0.67 | 0.63 | | |
| OzNet 0-90cm Monthly correlation | 0.5 | 0.75 | 0.79 | 0.79 | 0.77 | 0.76 | 0.79 | 0.76 | | | 0.78 | 0.72 | | |
| OzFlux 0-90 Daily correlation | 0.5 | 0.75 | 0.85 | | | 0.79 | 0.83 | 0.85 | | | | | | |
| OzFlux 0-90 Monthly correlation | 0.5 | 0.75 | 0.88 | | | 0.85 | 0.88 | 0.88 | | | | | | |
| CosmOz 0-90 Daily correlation | 0.5 | 0.75 | 0.55 | | | 0.70 | 0.69 | 0.70 | | | | | | |
| CosmOz 0-90 Monthly correlation | 0.5 | 0.75 | 0.72 | | | 0.83 | 0.82 | 0.85 | | | | | | |
| Terrestrial Water Storage | | | | | | | | | | | | | | |
| Monthly correlation | 0.5 | 0.75 | 0.60 | | | 0.46 | 0.52 | 0.60 | | | | | | |
| De-seasonalised Monthly correlation | 0.5 | 0.75 | 0.60 | | | 0.45 | 0.54 | 0.60 | | | | | | |
| Actual Evapotranspiration | | | | | | | | | | | | | | |
| Monthly correlation | 0.05 | 0.5 | 0.61 | 0.32 | 0.61 | 0.50 | 0.44 | 0.38 | | | | | 0.59 | 0.40 |
| | 0.5 | 0.8 | 0.86 | 0.85 | 0.86 | 0.84 | 0.85 | 0.86 | | | | | 0.81 | 0.77 |
| Daily correlation | 0.05 | 0.5 | 0.47 | | | 0.37 | 0.19 | 0.47 | | | | | 0.34 | 0.24 |
| | 0.5 | 0.75 | 0.69 | | | 0.59 | 0.61 | 0.69 | | | | | 0.66 | 0.60 |
| Recharge | | | | | | | | | | | | | | |
| Long Term Average Relative bias | 0.25 | <0 | -0.2 | 0.48 | 0.6 | -0.49 | -0.2 | -0.21 | | | | | | |
| | 0.75 | >0 | 4.23 | 11.5 | 13.98 | 4.23 | 6.81 | 7.26 | | | | | | |
| Long Term Average Spatial Correlation | 0.5 | 0.5 | 0.52 | 0.66 | 0.64 | 0.49 | 0.51 | 0.52 | | | | | | |
| Recharge annual correlation | 0.5 | 0.5 | 0.47 | 0.44 | 0.47 | 0.22 | 0.29 | 0.46 | | | | | | |
| Vegetation Fraction | | | | | | | | | | | | | | |
| Monthly correlation | 0.5 | 0.75 | 0.72 | | | 0.63 | 0.57 | 0.72 | | | | | | |
| De-seasonalised Monthly correlation | 0.5 | 0.75 | 0.65 | | | 0.58 | 0.51 | 0.65 | | | | | | |

# 6 Evaluation of AWRA-L for reporting purposes

Gridded outputs and select catchment based timeseries from AWRA-L v5, v6 and v7 are presented here to give an understanding of the AWRA-L model spatial and temporal dynamics and the extent of changes presented by the updated version.

**National annual spatial plots:** Annual totals and deciles for the years covering 2008-2017 are plotted in Figure 37 and Figure 38. These plots show year to year variability, in particular the ending of the Millennium drought and the spatial variability expected across Australia according to climate and catchment conditions. It is noted there is an area in arid Western Australia where rainfall is sparsely monitored, leaving a hole in the interpolated rainfall in the region. This further translates into the modelled water balance components having no flows in this area.

**Catchment Timeseries**: Ten catchments were selected for evaluation of states/flux time-series as shown in Figure 39; with site plots shown in Figure 40 to Figure 49. Key features of these sites are presented in Table 16.Time series of the following variables are plotted for each of these catchments: potential and actual evapotranspiration, runoff, deep drainage to the groundwater store, top 0-10cm soil moisture, shallow 10-100cm soil moisture, deep 100-600cm soil moisture, TWS, and vegetation fraction (Fveg). Observed streamflow, GRACE TWS anomaly, MODIS vegetation fraction, along with satellite-based ET and soil moisture is also plotted for comparison purposes. These plots give an indication of the seasonal and inter-annual variability present at each of these locations for the key water variables output by AWRA-L.

Key observations from these site plots include:

- Streamflow performance varies according to location, with some sites showing good performance (eg. Site 116013), and others overestimating (eg. Site 226222) or underestimating (e.g. site 410048) variance;
- Top layer 'upper' soil moisture (0-10cm) is low compared to satellite data for all AWRA versions, with v7 between the lower v6 and higher v5 values.
- Shallow 'lower' layer soil moisture (10-100cm) is similar for v5, v6 and v7;
- the deep storage and deep drainage shows drawdown over the Millennium drought period – as expected in this area;
- Deep drainage decreases for most sites for v7 compared to v5 and v6.
- AWRA-L v7 PET is higher than v5 and v6;
- In terms of Fveg, a considerable change has happened in version 7 which partially fixed a phase difference between observed and modelled maximum and minimum.

Overall, the spatial plots and time series give confidence in the use of AWRA-L for water resources assessment, as they broadly follow the expected catchment responses and spatial and temporal trends expected across Australia. The comparisons between the outputs of AWRA-L v5, AWRA-L v6 and AWRA-L v7 emphasise that users of the data will need to be aware of the model updates and in some cases review any relationships they have created between their models and AWRA data.

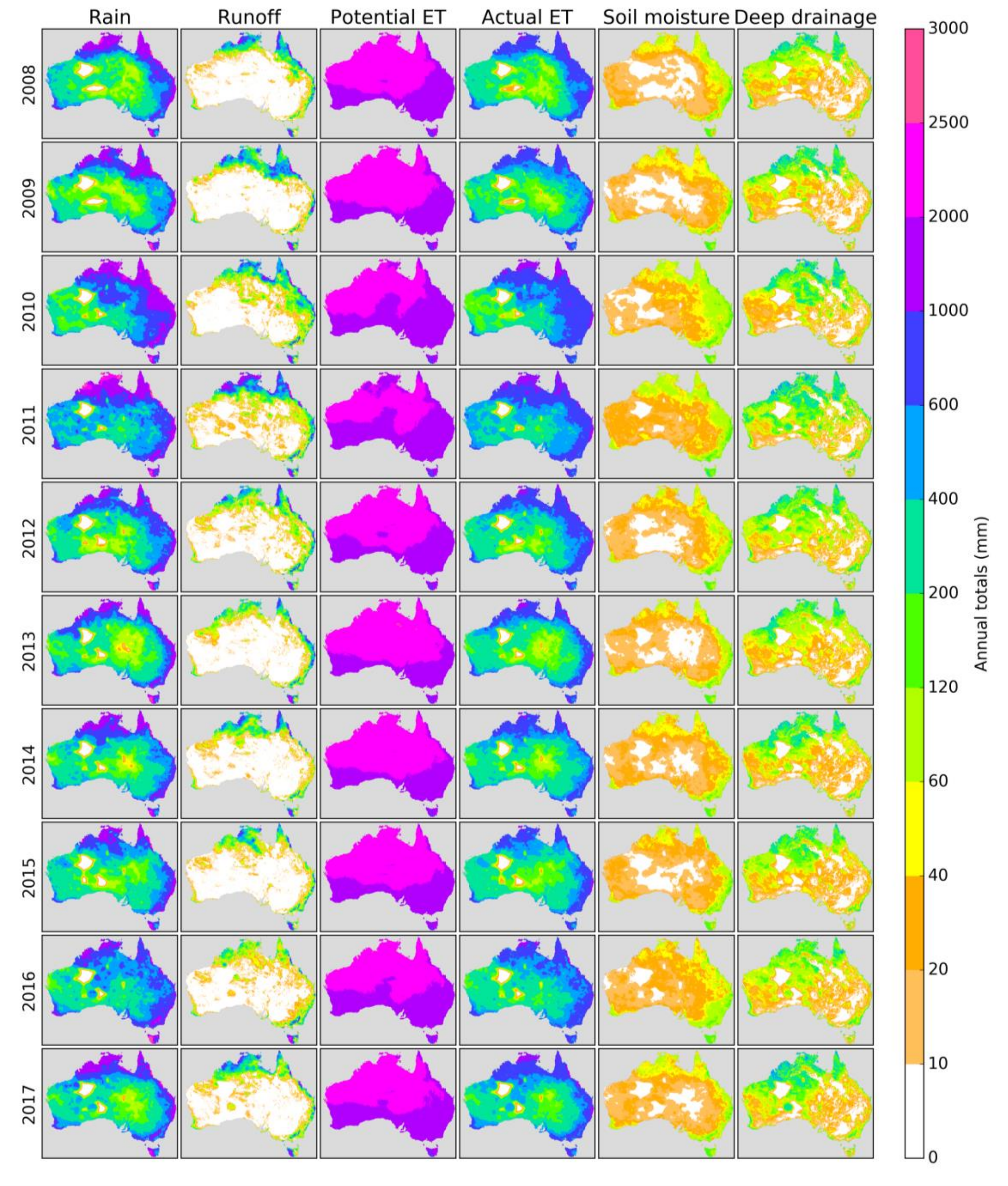



**Figure 37. AWRA-L v7 mean annual rainfall, runoff, PET, AET, soil moisture and deep drainage 2008-2017. Units=mm.**

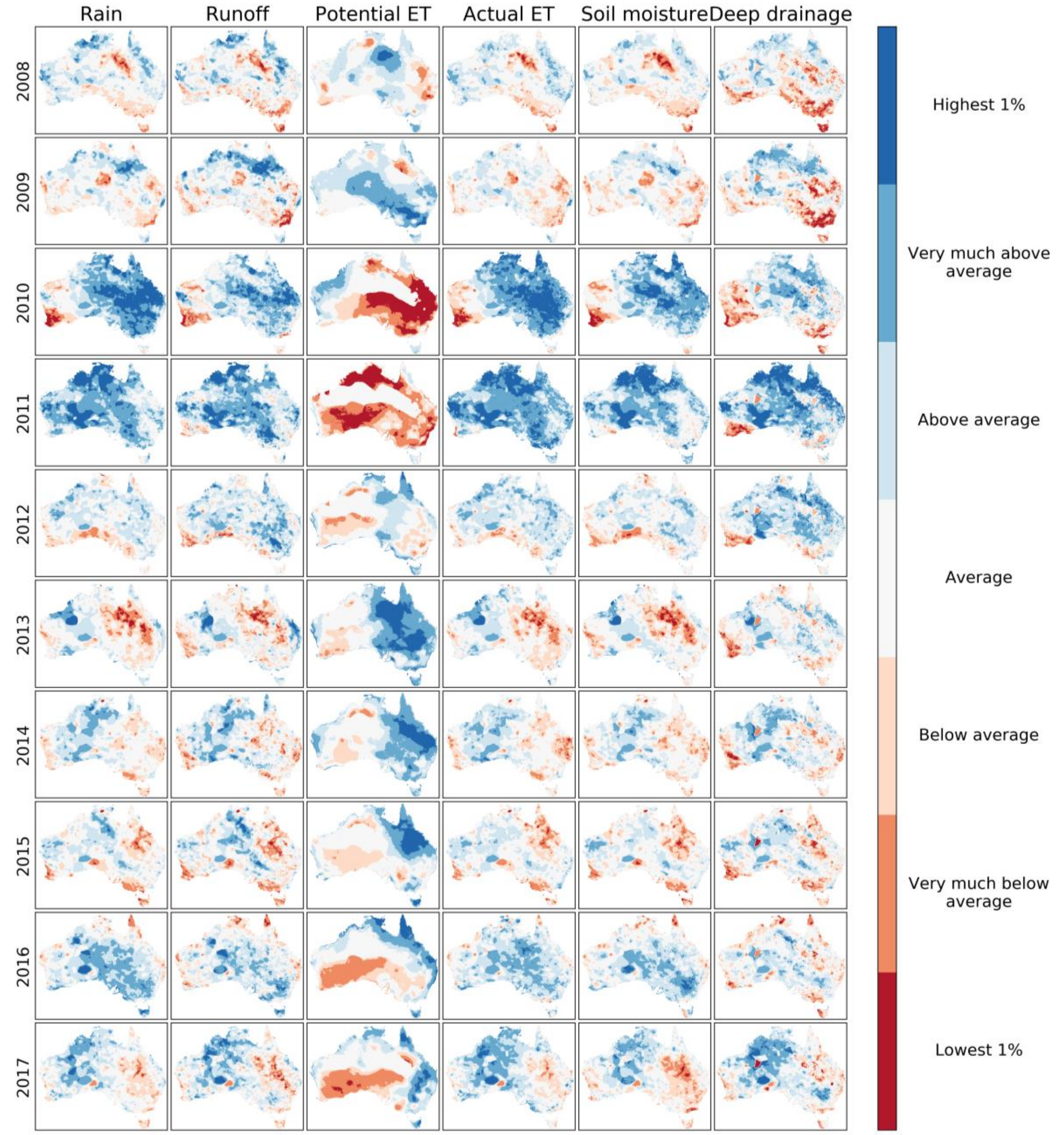



**Figure 38. AWRA-L v7 Annual rain, runoff, ET, soil moisture and deep drainage deciles for 2008-2017**

**Table 16. Selected catchments for detailed evaluation**

| ID | Name | River | State | Lat. (°) | Lon. (°) | Area (km²) | Elev. (m) | Slope (%) | Ave. Precip. (mm) | Ave.PET (mm) | Forest (-) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **145105** | Beaudesert Pump Stn | Albert | QLD | -28.02 | 153.06 | 266 | 326 | 8 | 1209 | 1443 | 0.6 |
| **116013** | Archer Ck | Millstream | QLD | -17.65 | 145.34 | 315 | 911 | 4 | 1589 | 1714 | 0.5 |
| **226222** | near Noojee (U/S Ada R) | Latrobe | VIC | -37.88 | 145.89 | 65 | 480 | 8 | 1352 | 1103 | 0.9 |
| **403213** | Greta South | Fifteen Mile Ck | VIC | -36.62 | 146.24 | 231 | 549 | 7 | 1032 | 1214 | 0.6 |
| **410048** | Ladysmith | Kyeamba Ck | NSW | -35.20 | 147.53 | 548 | 321 | 3 | 641 | 1217 | 0.3 |
| **501503** | U/S Victor Harbour Stw | Inman | SA | -35.54 | 138.58 | 165 | 168 | 4 | 701 | 1190 | 0.4 |
| **614044** | Yarragil Formation | Yarragil Brook | WA | -32.81 | 116.15 | 71 | 288 | 2 | 904 | 1489 | 0.7 |
| **607155** | Malimup Track | Dombakup Brook | WA | -34.58 | 115.97 | 116 | 87 | 1 | 1129 | 1288 | 0.7 |
| **814011** | Manbulloo Boundary | Dry | NT | -15.08 | 132.41 | 4786 | 204 | 0 | 896 | 2091 | 0.2 |
| **811004** | Victoria HWY | East Baines | NT | -15.77 | 130.03 | 2443 | 195 | 2 | 833 | 1988 | 0.2 |

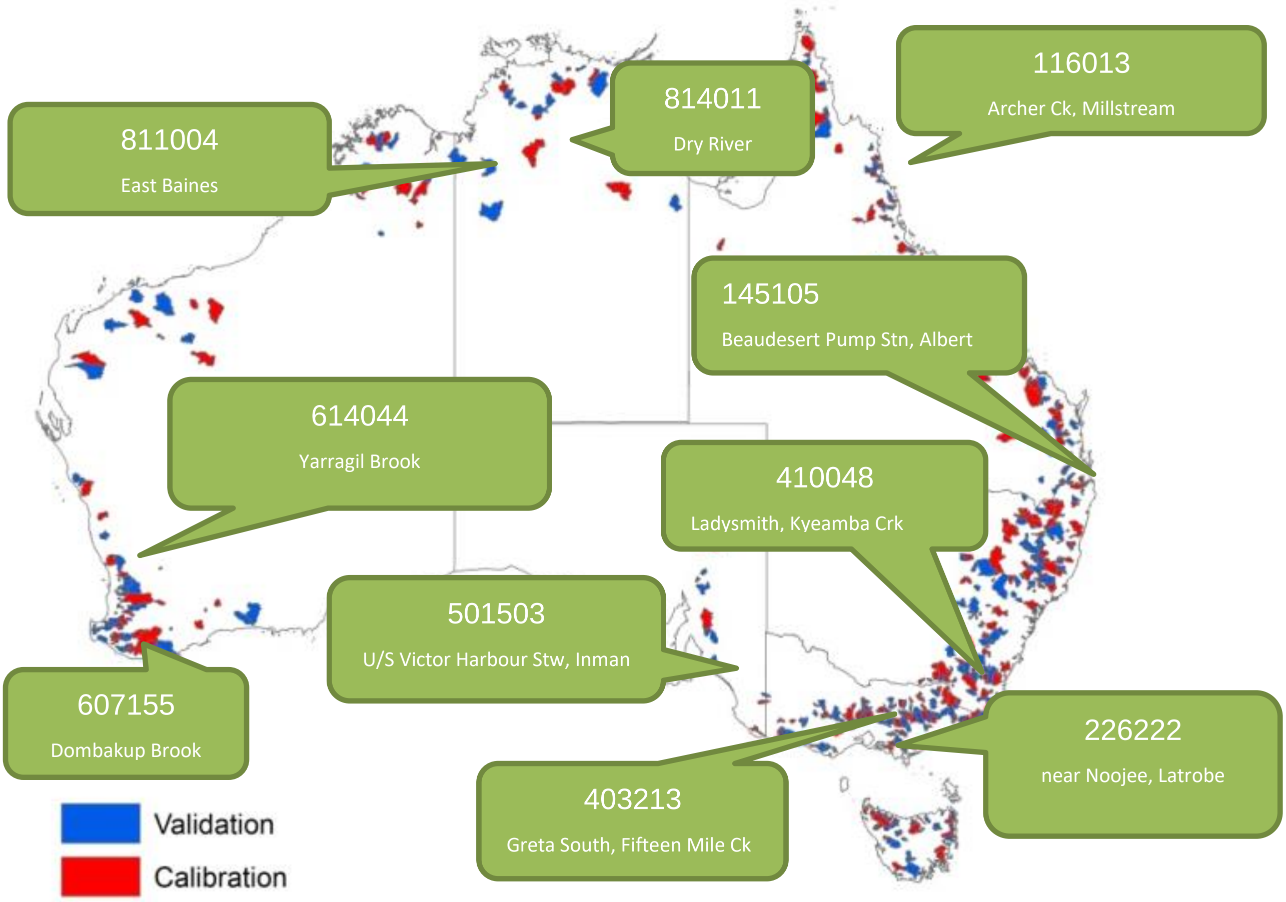



**Figure 39. Locations of selected catchments for detailed evaluation.**

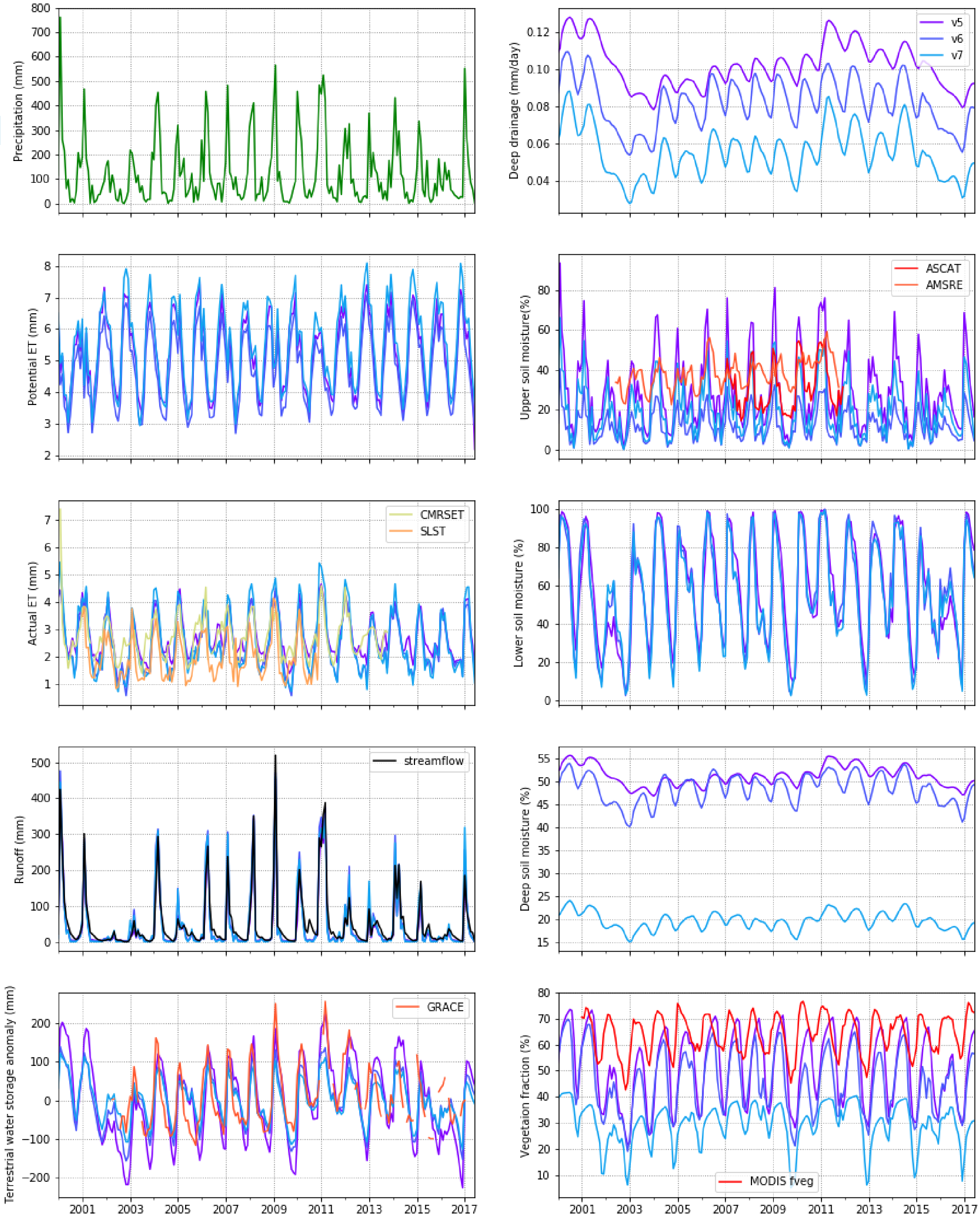



**Figure 40. 116013 Archer Creek @ Millstream QLD AWRA-L monthly simulations.**

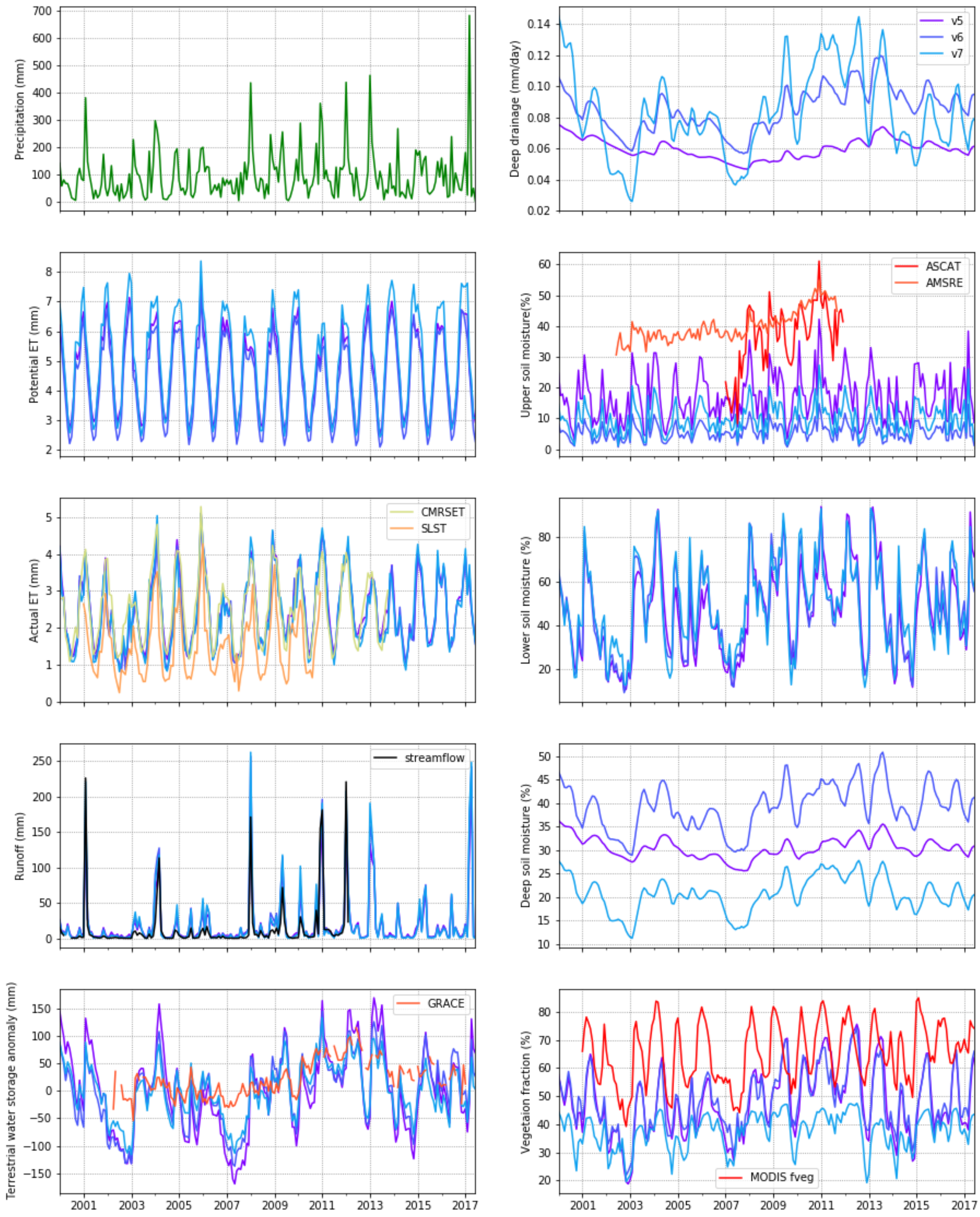



**Figure 41. 145105 Beaudesert, Albert River QLD AWRA-L monthly simulations.**

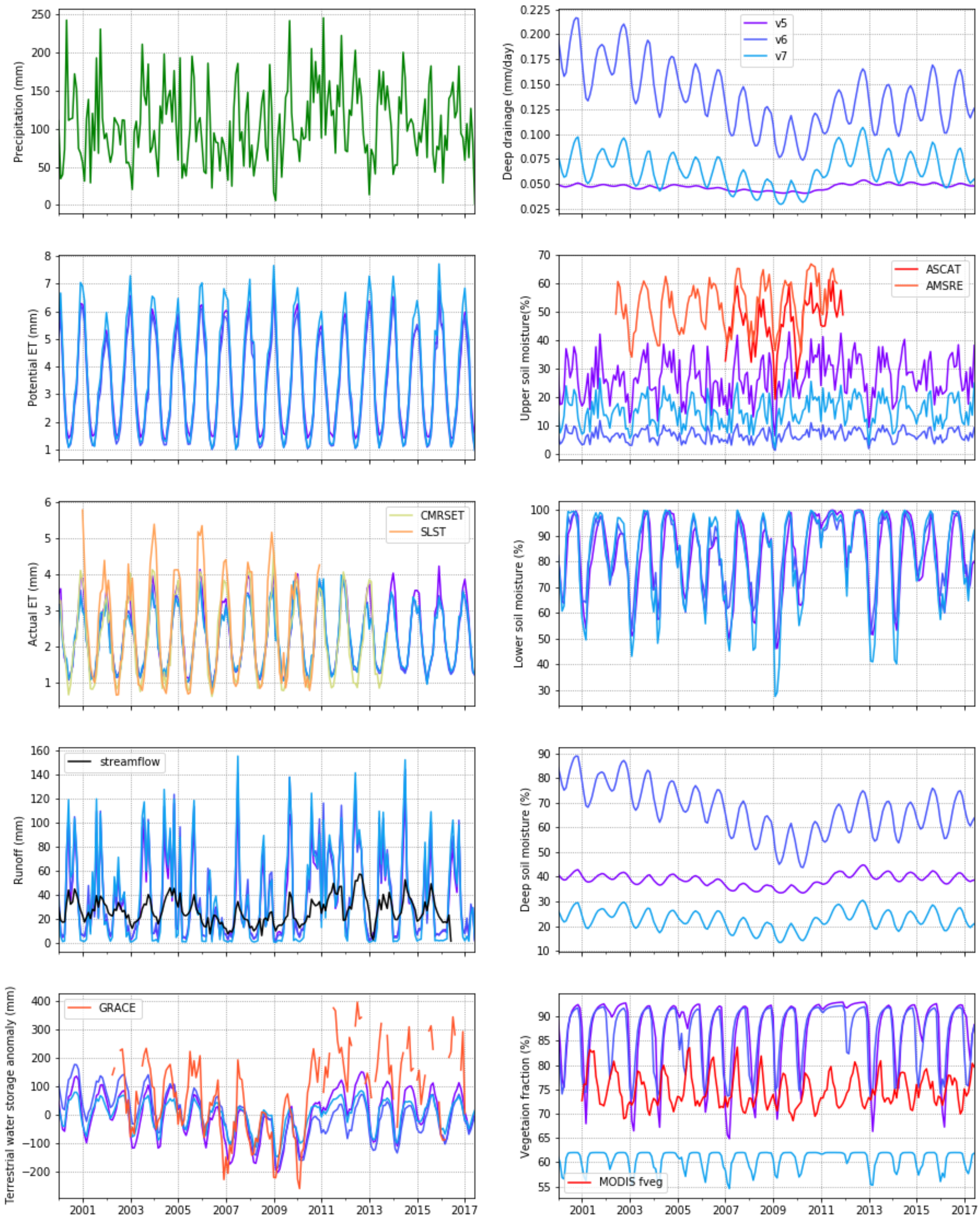



**Figure 42. 226222 near Noojee VIC AWRA-L monthly simulations.**

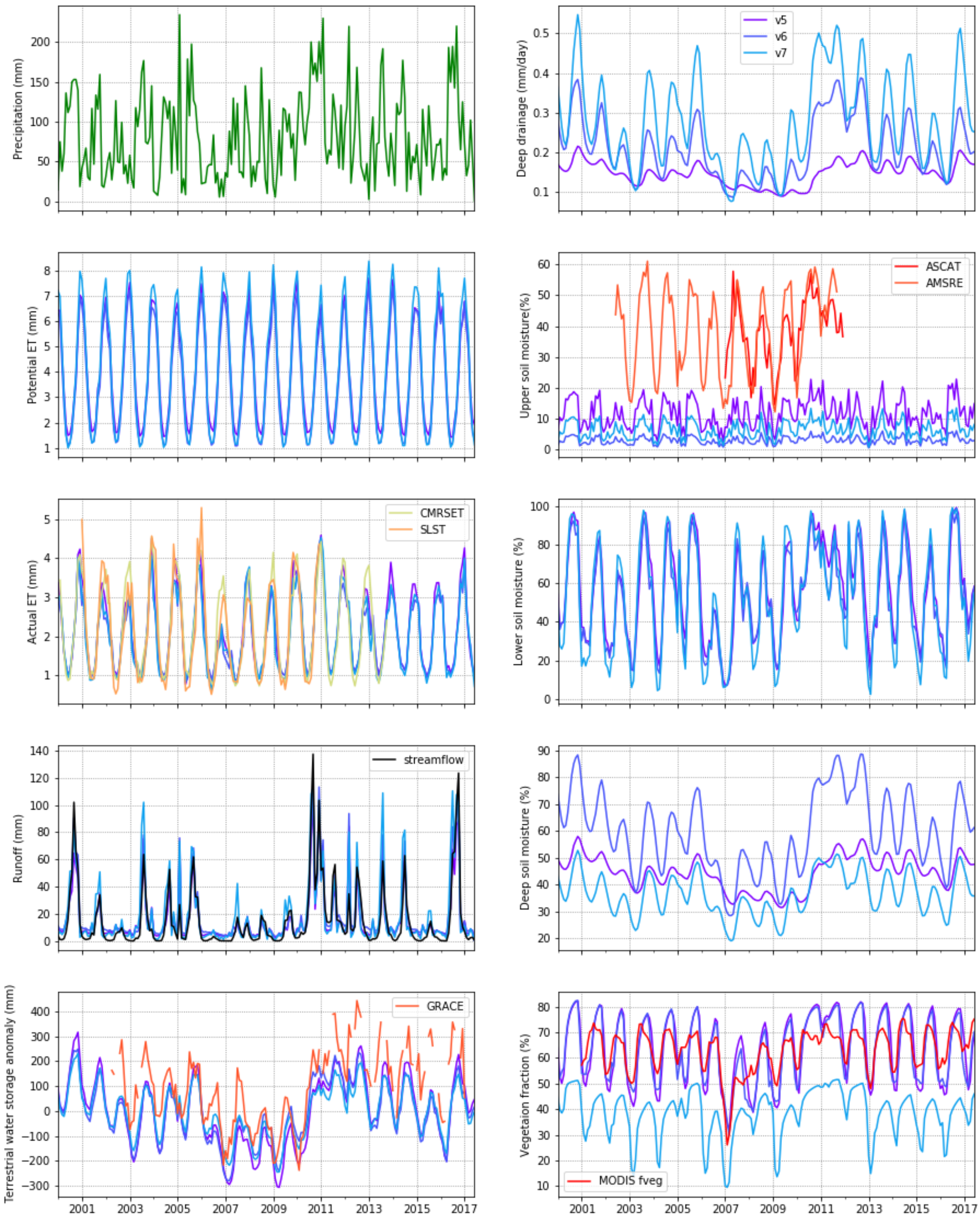



**Figure 43. 403213 FIFTEEN MILE CREEK, GRETA SOUTH AWRA-L monthly simulations.**

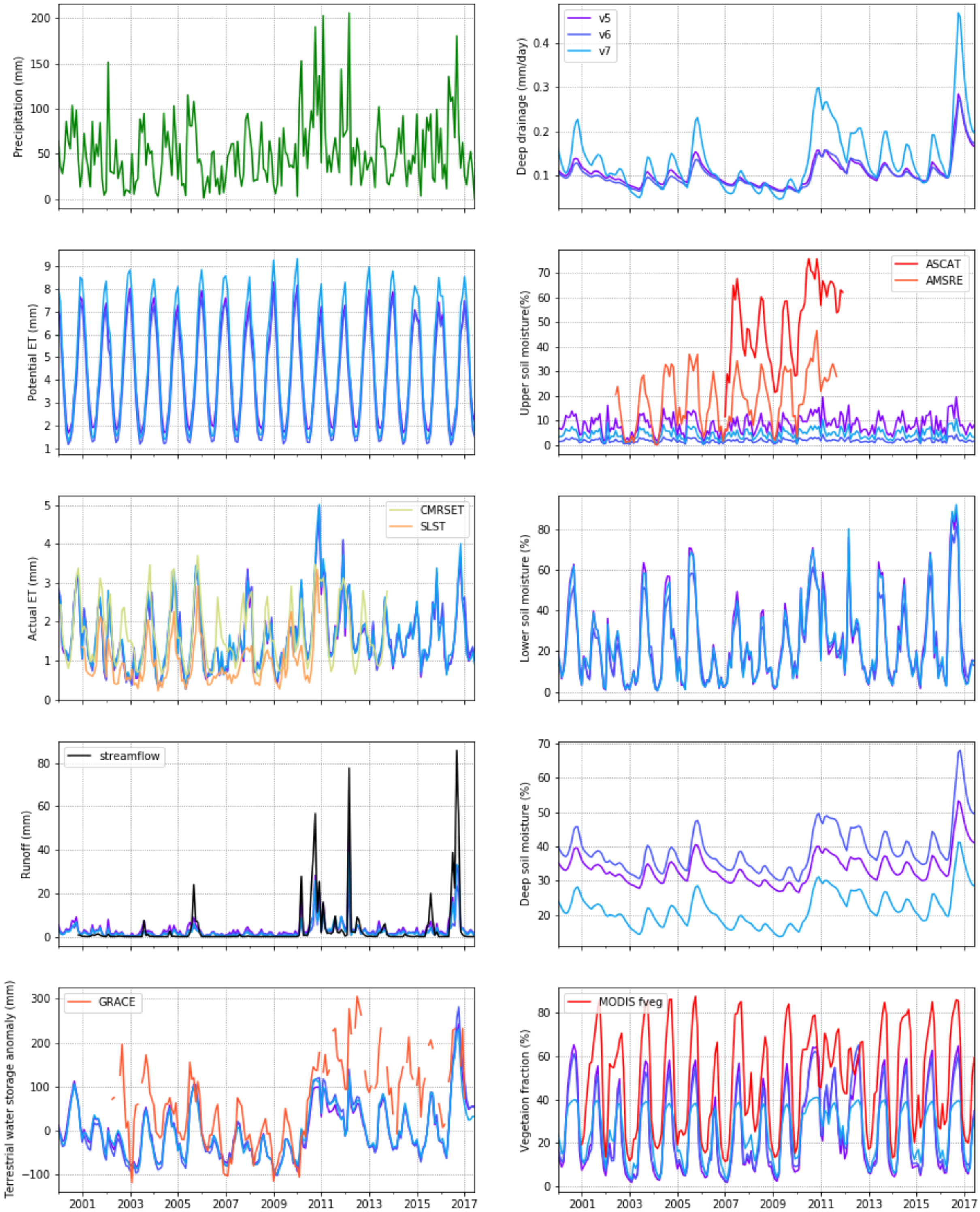



**Figure 44. 410048 Ladysmith, Kyeamba Creek NSW AWRA-L Monthly simulations.**

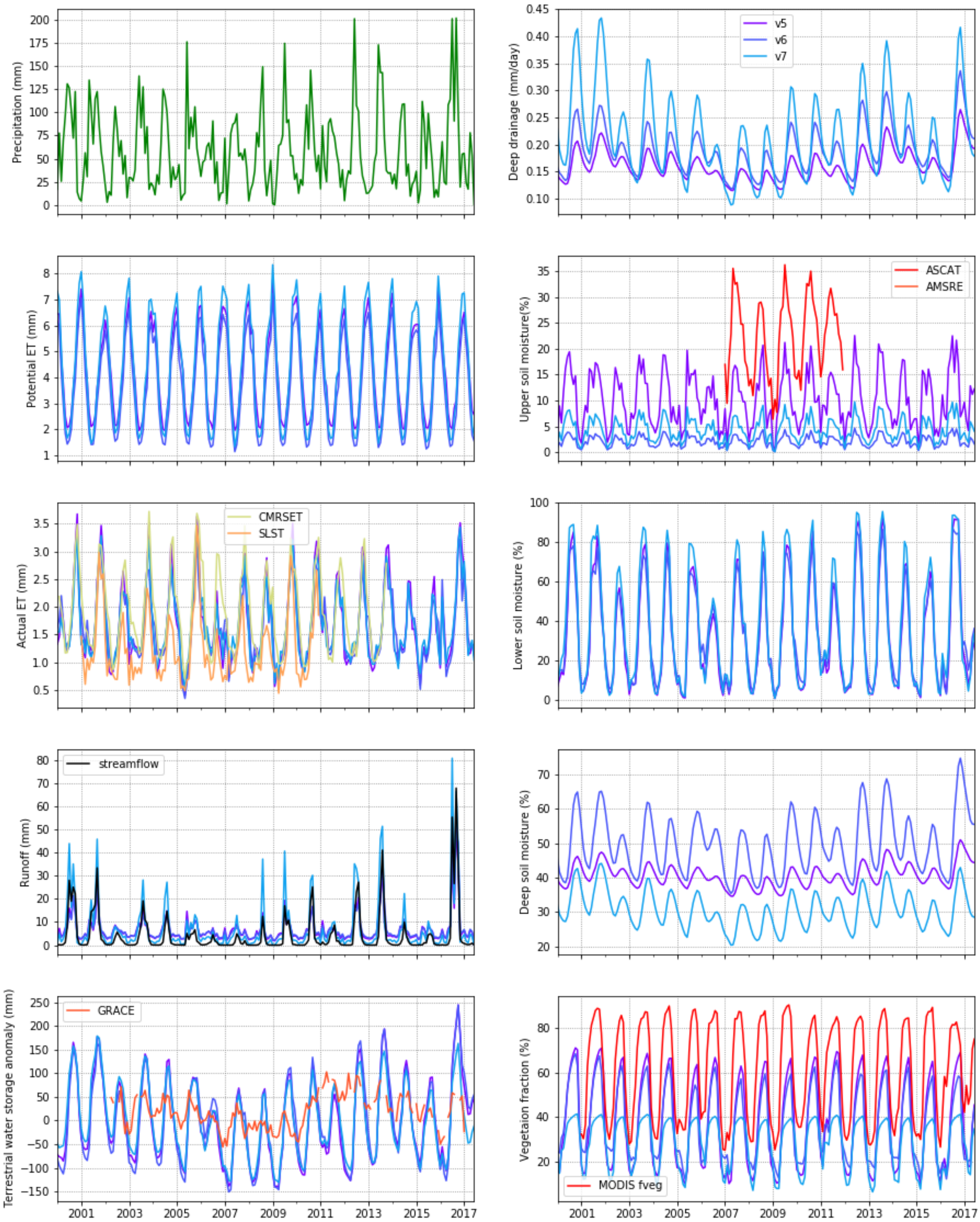



**Figure 45. 501503 US Victor harbour, Inman River SA AWRA-L monthly simulations.**

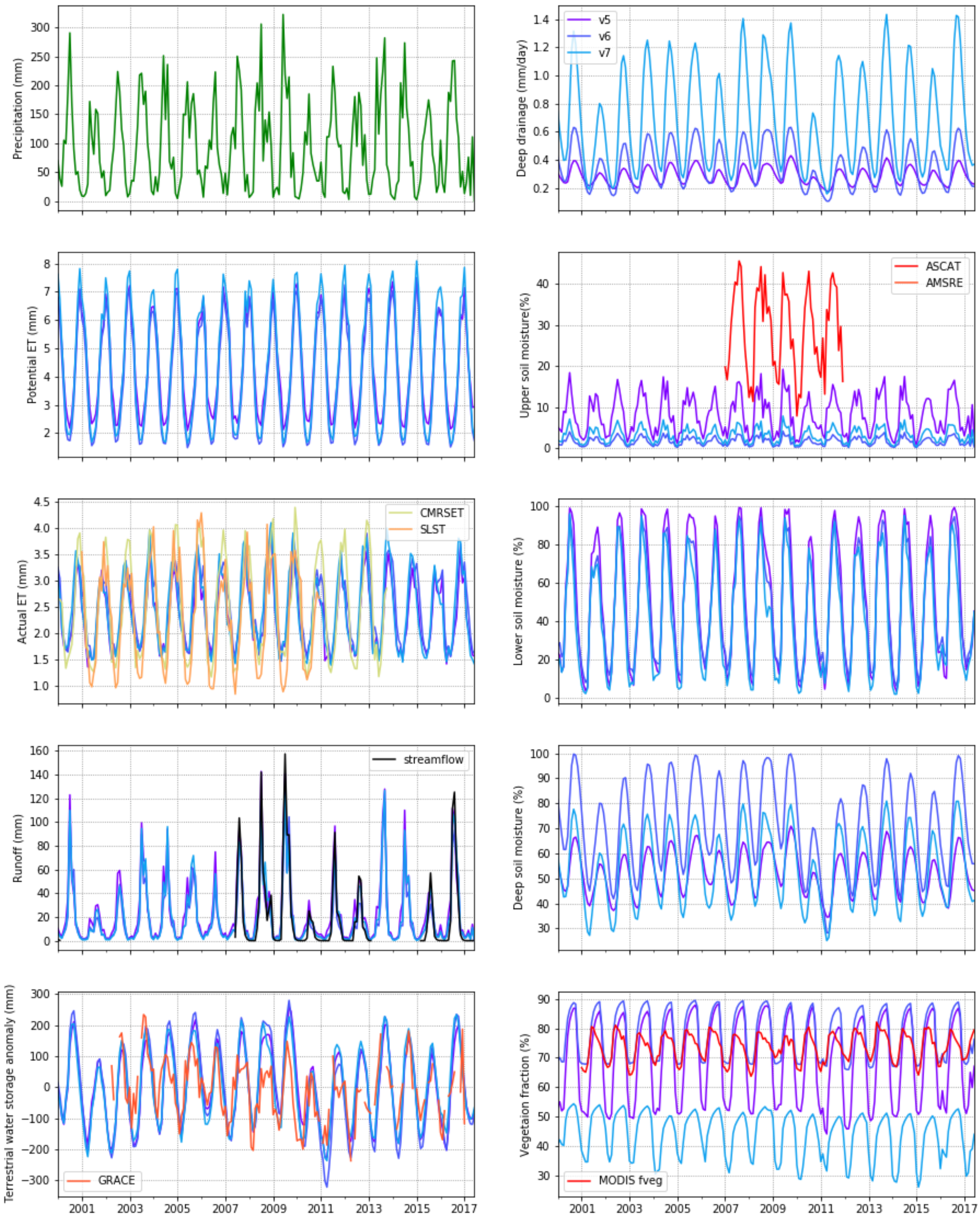



**Figure 46. 607155 Malimup Track, Dombakup Brook WA  AWRA-L Monthly simulations.**

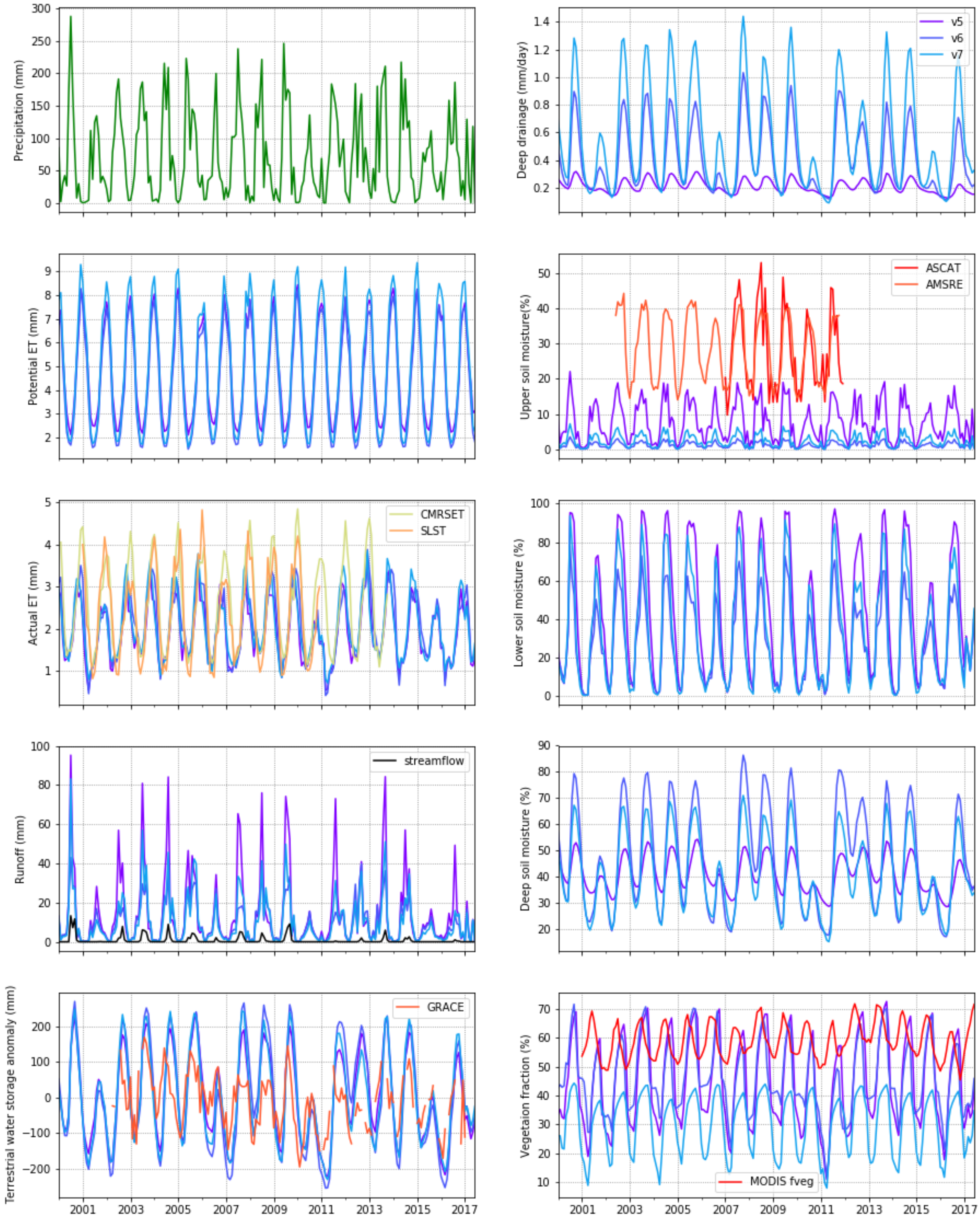



**Figure 47. 614044 Yarragil Formation, Yarragil Brook WA AWRA-L monthly simulations.**

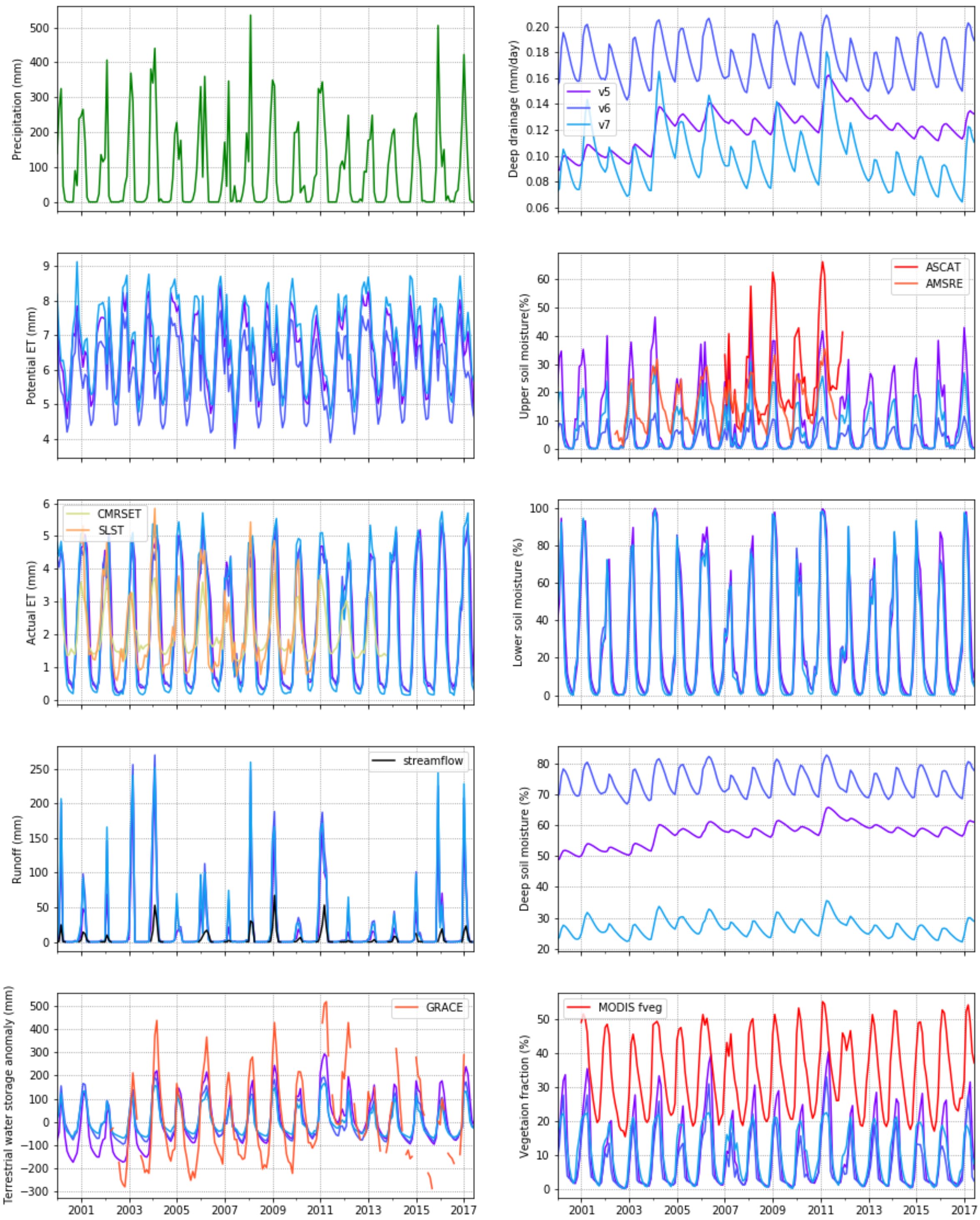



**Figure 48. 814011 Dry River NT AWRA-L monthly simulations.**

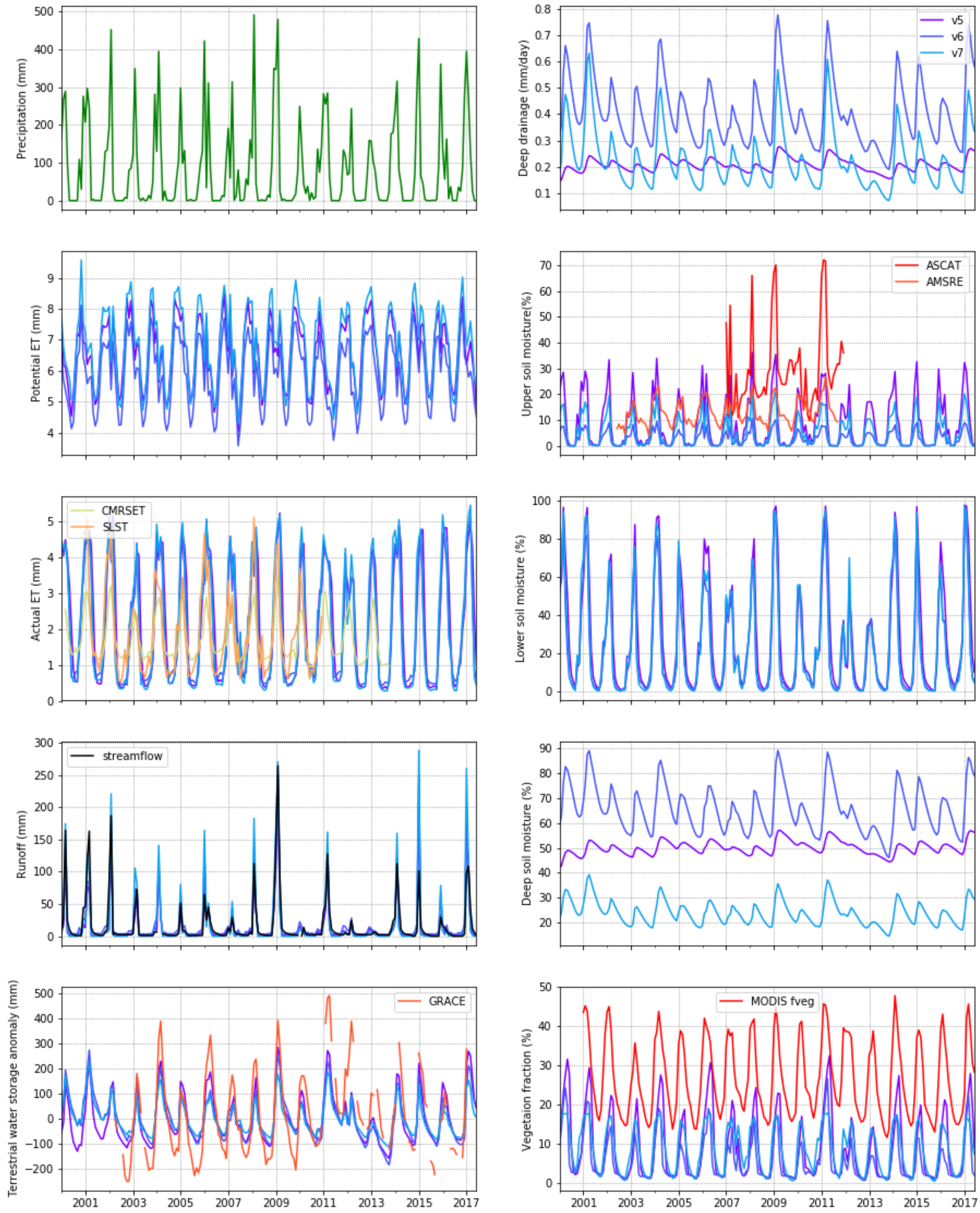



**Figure 49. 811004 Victoria HWY East Baines NT AWRA-L monthly simulations.**

# 7 Conclusions

AWRA-L performance was evaluated using available streamflow, soil moisture, evapotranspiration and groundwater recharge hydrological data sets and compared to two peer national models (WaterDyn and CABLE) as well as two locally calibrated nearest-neighbour regionalised rainfall-runoff models (GR4J and Sacramento). Performance against key evaluation criteria was undertaken, and results presented in Table 15 provide a benchmark from which future versions of AWRA-L and other models can be compared. Aspirational targets for model performance are set based on past experience. AWRA-L performs well across the range of variables tested. In certain cases, AWRA-L does not reach the aspirational targets set leaving room for future improvement.

AWRA-L reproduces streamflow relatively well (compared to the other national models trialled) over the 291 catchments reserved for validation. It performs particularly well considering it is approaching the performance of locally calibrated-nearest neighbour regionalised rainfall-runoff models (and is superior in terms of bias). AWRA-L v7 outperforms v6 considerably in terms of KGE (both daily and monthly), slightly better performance in terms of daily NSE and greatly improved performance in terms of bias. These improvements are attributed to improved model structure, in particular baseflow representation, coupled with calibration to slower water balance dynamics represented by GRACE, and improved vegetation dynamics.

To test the benefit of adding a new impervious HRU to AWRA, streamflow performance was evaluated over 13 catchments with a high impervious fraction. AWRA v7 outperformed the previous versions with higher daily median NSE value and reduced bias. These improvements are significant as they occurred through the introduction of the new impervious HRU without any calibration and without compromising the performance of other parts of the model. Calibration of this HRU could be an area of improvement in future versions of the model.

AWRA-L and CABLE perform similarly for root-zone (profile 0-90cm) soil moisture, with WaterDyn performing worse. Current AWRA-L performance according to daily and monthly correlation at the 50% is 0.73-0.77 for the Murrumbidgee and Hunter sites, and 0.69-0.88 median values for the national OzFlux/CosmOz networks. In general, AWRA-L v7 improves over AWRA-L v6. Top layer soil moisture (0-10cm) is improved in AWRA-L v7, following changes to drainage parametrisation. It is noted that ASCAT replaced AMSR-E in calibration of AWRA-L v7 towards aligning with the assimilation of ASCAT in the operational forecasting modelling system.

AWRA-L TWS improves over the previous versions where its performance according to median monthly correlation and de-seasonalised correlation is 0.58 and 0.55, respectively. This improvement reflects the use of GRACE in catchment-based calibration, being weighted 50%. However, it is below the median aspirational performance for TWS set here at 0.75. Most notably, adding TWS into the calibration has improved other components of the water balance without diminishing streamflow performance.

For actual ET, CABLE and WaterDyn are better overall than AWRA-L model, although AWRA-L v7 median monthly correlation is equivalent to CABLE. AWRA-L v7 performance according to monthly correlation to DINGO flux tower data at 5% / 50% is 0.53 / 0.85 with aspirational performance set at 0.5 / 0.8 respectively. AWRA-L v7 is an improvement over v5 and v6 for 50% and 5% correlation, with a marked improvement in daily correlation due to shift in input data. Other changes influencing ET include incorporation of vapour pressure as an input field, calculation of maximum transpiration, use of vegetation in calibration, and updating of the satellite based CMRSET data used to calibrate AWRA-L to use 8-day rather than monthly estimates.

AWRA-L vegetation fraction performs well according to monthly correlation and de-seasonalised correlation at 50% is 0.65 and 0.72, respectively. While AWRA-L v7 improves over its previous versions, it is below the median aspirational performance for TWS set here as 0.75 leaving space for improvement in the future investigations.

The addition of two new observation groups to the calibration objective function, namely the vegetation fraction derived from MODIS and TWS from GRACE, has enabled calibration to previously unobserved parts of the model. This change, coupled with improvements to the model structure and model inputs, has improved the entire water balance and vegetation dynamics of the AWRA-L v7 model. Notably, streamflow performance in the validation catchments has improved as a result, even though it is weighted in calibration at a lower level than in previous versions of the model, reflecting the valuable contribution of these two new satellite derived observation groups.

AWRA-L appears to not match the spatial patterns of the national recharge dataset, due to drainage currently being overly dependent on saturated conductivity, and not enough on rainfall variability. However, it is noted that there is high uncertainty in this evaluation data. Current AWRA-L v7 performance according to relative bias against the national long term average recharge dataset at 25% / 75% is -21% / 423677% with aspirational performance set at being less than zero / greater than zero respectively, and a spatial correlation of 0.5 spatially. Secondly AWRA-L v7 has a median annual correlation against the South Australian annual time-series dataset of 0.46, slightly below the aspirational target of 0.5. AWRA-L v7 improves in terms of correlation performance compared to v6.

Each of the models have differing strengths and weaknesses. Overall, given runoff/streamflow is the dominant hydrological variable used in surface water resource assessment, and that AWRA-L performs well for root zone soil moisture (a key agricultural variable), AWRA-L v7 is considered most fit for purpose water balance estimation purposes of the models evaluated.

Various maps and time series are presented to give an understanding of the spatial and temporal dynamics of the model and key output variables according to the AWRA-L v5, v6 and v7. Overall, the water balance verification statistics, the spatial plots and time series give confidence in the use of AWRA-L for national water resource assessment, as they broadly follow the expected catchment responses and spatial and temporal trends expected across Australia.

The AWRA-L model is available as a community model to enable use by a wide range of stakeholders and further development and application by the wider research community. Please see: https://github.com/awracms/awra_cms for further details.

# References

Bartalis, Z., Wagner, W., Naeimi, V., Hasenauer, S., Scipal, K., Bonekamp, H., Figa, J., Anderson, C., 2007. Initial soil moisture retrievals from the METOP-A Advanced Scatterometer (ASCAT). Geophys. Res. Lett. 34, L20401. https://doi.org/10.1029/2007gl031088

Beringer, J., Hutley, L.B., McHugh, I., Arndt, S.K., Campbell, D., Cleugh, H.A., Cleverly, J., de Dios, V., Eamus, D., Evans, B., Ewenz, C., Grace, P., Griebel, A., Haverd, V., Hinko-Najera, N., Huete, A., Isaac, P., Kanniah, K., Leuning, R., Liddell, M.J., Macfarlane, C., Meyer, W., Moore, C., Pendall, E., Phillips, A., Phillips, R.L., Prober, S.M., Restrepo-Coupe, N., Rutledge, S., Schroder, I., Silberstein, R., Southall, P., Yee, M.S., Tapper, N.J., van Gorsel, E., Vote, C., Walker, J., Wardlaw, T., 2016. An introduction to the Australian and New Zealand flux tower network -- OzFlux. Biogeosciences 13, 5895–5916. https://doi.org/10.5194/bg-13-5895-2016

Beringer, J., McHugh, I., Hutley, L.B., Isaac, P., Kljun, N., 2017. Technical note: Dynamic~INtegrated Gap-filling and partitioning for OzFlux (DINGO). Biogeosciences 14, 1457–1460. https://doi.org/10.5194/bg-14-1457-2017

Burnash, R.J., 1995. The NWS river forecast system–catchment modelling, in: Singh, V.P. (Ed.), Computer Models of Watershed Hydrology. Water Resources Publications, Littleton, Colorado, pp. 311–366.

Crosbie, R., Jolly, I.D., Leaney, F.W., Petheram, C., Wohling, D., 2010. Review of Australian groundwater recharge studies. CSIRO: Water for a Healthy Country National Research Flagship.

Crosbie, R.S., Jolly, I.D., Leaney, F.W., Petheram, C., 2010. Can the dataset of field based recharge estimates in Australia be used to predict recharge in data-poor areas? Hydrol. Earth Syst. Sci. 14, 2023–2038. https://doi.org/10.5194/hess-14-2023-2010

Draper, C.S., Walker, J.P., Steinle, P.J., de Jeu, R.A.M., Holmes, T.R.H., 2009. An evaluation of AMSR-E derived soil moisture over Australia. Remote Sens. Environ. 113, 703–710.

Duan, Q.Y., Gupta, V.K., Sorooshian, S., 1993. Shuffled complex evolution approach for effective and efficient global minimization. J. Optim. Theory Appl. 76, 501–521. https://doi.org/10.1007/bf00939380

Eamus, D., Cleverly, J., Boulain, N., Grant, N., Faux, R., Villalobos-Vega, R., 2013. Carbon and water fluxes in an arid-zone Acacia savanna woodland: An analyses of seasonal patterns and responses to rainfall events. Agric. For. Meteorol. 182–183, 225–238. https://doi.org/http://dx.doi.org/10.1016/j.agrformet.2013.04.020

Fowler, K., Knoben, W., Peel, M., Peterson, T., Ryu, D., Saft, M., Seo, K.-W., Western, A., 2020. Many Commonly Used Rainfall-Runoff Models Lack Long, Slow Dynamics: Implications for Runoff Projections. Water Resour. Res. 56, e2019WR025286.

https://doi.org/10.1029/2019WR025286

Frost, A.J., Ramchurn, A., Smith, A., 2018. The Australian Landscape Water Balance model. AWRA-L v6). Tech. Descr. Aust. Water Resour. Assess. Landsc. Model. Version 6.

Frost, A.J., Ramchurn, A., Smith, A., 2016. The Bureau's Operational Landscape (AWRA-L) Model: Technical Description of AWRA-L version 5.

Frost, A.J., Shokri, A., 2021. The Australian Landscape Water Balance model (AWRA-L v7). Technical Description of the Australian Water Resources Assessment Landscape model version 7.

Frost, A.J., Wright, D.P., 2018a. The Australian Water Resource Assessment Landscape model – AWRA-L: improved performance and regional calibration, in: Proceedings from HWRS2018: Hydrology and Water Resources Symposium, Melbourne.

Frost, A.J., Wright, D.P., 2018b. Evaluation of the Australian Landscape Water Balance model (AWRA-L v6): Comparison of AWRA-L v6 against Observed Hydrological Data and Peer Models.

Grant, I., Jones, D., Wang, W., Fawcett, R., Barratt, D., 2008. Meteorological and remotely sensed datasets for hydrological modelling: a contribution to the Australian Water Availability Project, in: Proceedings of the Catchment-Scale Hydrological Modelling & Data Assimilation (CAHMDA-3) International Workshop on Hydrological Prediction: Modelling, Observation and Data Assimilation, January 9 2008–January 11 2008.

Guerschman, J.P., Scarth, P.F., McVicar, T.R., Renzullo, L.J., Malthus, T.J., Stewart, J.B., Rickards, J.E., Trevithick, R., 2015. Assessing the effects of site heterogeneity and soil properties when unmixing photosynthetic vegetation, non-photosynthetic vegetation and bare soil fractions from Landsat and MODIS data. Remote Sens. Environ. 161, 12–26. https://doi.org/https://doi.org/10.1016/j.rse.2015.01.021

Guerschman, J.P., Van Dijk, A.I.J.M., 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. J. Hydrol. 369, 107–119.

Gupta, H. V, Kling, H., Yilmaz, K.K., Martinez, G.F., 2009. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. J. Hydrol. 377, 80–91. https://doi.org/https://doi.org/10.1016/j.jhydrol.2009.08.003

Haverd, V., Cuntz, M., 2010. Soil–Litter–Iso: A one-dimensional model for coupled transport of heat, water and stable isotopes in soil with a litter layer and root extraction. J. Hydrol. 388, 438–455. https://doi.org/http://dx.doi.org/10.1016/j.jhydrol.2010.05.029

Haverd, V., Raupach, M.R., Briggs, P.R., Canadell, J.G., Isaac, P., Pickett-Heaps, C., Roxburgh, S.H., van Gorsel, E., Viscarra Rossel, R.A., Wang, Z., 2013. Multiple observation types reduce uncertainty in Australia's terrestrial carbon and water cycles. Biogeosciences 10, 2011–2040. https://doi.org/10.5194/bg-10-2011-2013

Jones, D.A., Wang, W., Fawcett, R., 2009. High-quality spatial climate data-sets for Australia. Aust. Meteorol. Oceanogr. J. 58, 233–248. https://doi.org/10.22499/2.5804.003

Kowalczyk, E.A., Wang, Y.P., Law, R.M., Davies, H.L., McGregor, J.L., Abramowitz, G., 2006. The CSIRO Atmosphere Biosphere Land Exchange model for use in climate models and as an offline model. CSIRO.

Lacava, T., Brocca, L., Faruolo, M., Matgen, P., Moramarco, T., Pergola, N., Tramutoli, V., 2012. A multi-sensor (SMOS, AMSR-E and ASCAT) satellite-based soil moisture products inter-comparison, in: 2012 IEEE International Geoscience and Remote Sensing Symposium. pp. 1135–1138. https://doi.org/10.1109/IGARSS.2012.6351348

Leroux, D.J., Kerr, Y.H., Bitar, A.A., Bindlish, R., Jackson, T.J., Berthelot, B., Portet, G., 2014. Comparison Between SMOS, VUA, ASCAT, and ECMWF Soil Moisture Products Over Four Watersheds in U.S. IEEE Trans. Geosci. Remote Sens. 52, 1562–1571. https://doi.org/10.1109/TGRS.2013.2252468

Owe, M., de Jeu, R., Holmes, T., 2008. Multisensor historical climatology of satellite-derived global land surface moisture. J. Geophys. Res. Earth Surf. 113, F01002. https://doi.org/10.1029/2007jf000769

Peeters, L., Crosbie, R., Van Dijk, A.I.J.M., 2011. Incorporating groundwater dynamics and clay content in AWRA-L. CSIRO: Water for a Healthy Country National Research Flagship.

Perrin, C., Michel, C., AndrÃeassian, V., 2003. Improvement of a parsimonious model for streamflow simulation. J. Hydrol. 279, 275–289. https://doi.org/http://dx.doi.org/10.1016/S0022-1694(03)00225-7

Potter, N.J., Chiew, F.H.S., Frost, A.J., 2010. An assessment of the severity of recent reductions in rainfall and runoff in the Murray-Darling Basin. J. Hydrol. 381, 52–64.

Ramchurn, A., Frost, A.J., 2013. Potential improvements to the Australian Water Resources Assessment system landscape ( AWRA-L ) model. Proc. 20th Int. Congr. Model. Simul.

Raupach, M.R., Briggs, P.R., Haverd, V., King, E.A., Paget, M., Trudinger, C.M., 2009. Australian Water Availability Project (AWAP) - CSIRO Marine and Atmospheric Research Component: Final Report for Phase 3. CSIRO Marine and Atmospheric Research, Canberra, Australia.

Refsgaard, J.C., Henriksen, H.J., 2004. Modelling guidelines--terminology and guiding principles. Adv. Water Resour. 27, 71–82.

Renzullo, L.J., van Dijk, A.I.J.M., Perraud, J.M., Collins, D., Henderson, B., Jin, H., Smith, A.B., McJannet, D.L., 2014. Continental satellite soil moisture data assimilation improves root-zone moisture analysis for water resources assessment. J. Hydrol. 519, Part, 2747–2762. https://doi.org/http://dx.doi.org/10.1016/j.jhydrol.2014.08.008

Rüdiger, C., Western, A.W., Walker, J.P., Smith, A.B., Kalma, J.D., Willgoose, G.R., 2010. Towards a general equation for frequency domain reflectometers. J. Hydrol. 383, 319–329. https://doi.org/http://dx.doi.org/10.1016/j.jhydrol.2009.12.046

Save, H., Bettadpur, S., Tapley, B.D., 2016. High-resolution CSR GRACE RL05 mascons. J. Geophys. Res. Solid Earth 121, 7547–7569. https://doi.org/https://doi.org/10.1002/2016JB013007

Shi, X., Crosbie, R.., Vaze, J., 2015. Evaluation of AWRA recharge: Comparison to field estimates of recharge across Australia. CSIRO Technical report.

Simard, M., Pinto, N., Fisher, J.B., Baccini, A., 2011. Mapping forest canopy height globally with spaceborne lidar. J. Geophys. Res. Biogeosciences 116, G04021. https://doi.org/10.1029/2011jg001708

Smith, A.B., Walker, J.P., Western, A.W., Young, R.I., Ellett, K.M., Pipunic, R.C., Grayson, R.B., Siriwardena, L., Chiew, F.H.S., Richter, H., 2012. The Murrumbidgee soil moisture monitoring network data set. Water Resour. Res. 48, W07701. https://doi.org/10.1029/2012wr011976

Van Dijk, A., 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., 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 Niel, T.G., McVicar, T.R., Roderick, M.L., van Dijk, A.I.J.M., Beringer, J., Hutley, L.B., van Gorsel, E., 2012. Upscaling latent heat flux for thermal remote sensing studies: Comparison of alternative approaches and correction of bias. J. Hydrol. 468–469, 35–46. https://doi.org/http://dx.doi.org/10.1016/j.jhydrol.2012.08.005

Vaze, J., Mateo, C.M., Wang, B., Teng, J., Marvanek, S., 2018. AWRA-L input spatial layers at 1 km and 5 km resolutions for the Australian continent - Source data and comparison between 1 km and 5 km resolutions.

Viney, N., Vaze, J., Crosbie, R., Wang, B., Dawes, W., Frost, A., 2015. AWRA-L v5.0: technical description of model algorithms and inputs. CSIRO, Australia.

Viney, N., Vaze, J., Crosbie, R., Wang, B., Dawes, W., Frost, A., 2014. AWRA-L v4.5: technical description of model algorithms and inputs. CSIRO, Australia.

Viney, N., Vaze, J., Crossbie, R., Wang, B., Dawes, W., Frost, A., 2015. AWRA-L v5.0: Technical description of model algorithms and inputs. CSIRO Land and Water.

Viney, N.R., Perraud, J., Vaze, J., Chiew, F.H.S., Post, D.A., Yang, A., 2009. The usefulness of bias constraints in model calibration for regionalisation to ungauged catchments. 18th World IMACS / MODSIM Congr.

Wagner, W., Brocca, L., Naeimi, V., Reichle, R., Draper, C., Jeu, R. de, Ryu, D., Su, C.-H., Western, A., Calvet, J.-C., Kerr, Y.H., Leroux, D.J., Drusch, M., Jackson, T.J., Hahn, S., Dorigo, W., Paulik, C., 2014. Clarifications on the "Comparison Between SMOS, VUA, ASCAT, and ECMWF Soil Moisture Products Over Four Watersheds in U.S." IEEE Trans. Geosci. Remote Sens. 52, 1901–1906. https://doi.org/10.1109/TGRS.2013.2282172

Wang, Y.P., Kowalczyk, E., Leuning, R., Abramowitz, G., Raupach, M.R., Pak, B., van Gorsel, E., Luhar, A., 2011. Diagnosing errors in a land surface model (CABLE) in the time and frequency domains. J. Geophys. Res. Biogeosciences 116, G01034. https://doi.org/10.1029/2010jg001385

Wang, Y.P., Law, R.M., Pak, B., 2010. A global model of carbon, nitrogen and phosphorus cycles for the terrestrial biosphere. Biogeosciences 7, 2261–2282. https://doi.org/10.5194/bg-7-2261-2010

Wilson, K., Goldstein, A., Falge, E., Aubinet, M., Baldocchi, D., Berbigier, P., Bernhofer, C., Ceulemans, R., Dolman, H., Field, C., Grelle, A., Ibrom, A., Law, B.E., Kowalski, A., Meyers, T., Moncrieff, J., Monson, R., Oechel, W., Tenhunen, J., Valentini, R., Verma, S., 2002. Energy balance closure at FLUXNET sites. Agric. For. Meteorol. 113, 223–243. https://doi.org/https://doi.org/10.1016/S0168-1923(02)00109-0

Zhang, Y.Q., Viney, N., Frost, A., Oke, A., Brooks, M., Chen, Y., Campbell, N., 2013. Collation of streamflow and catchment attribute dataset for 780 unregulated Australian catchments. CSIRO: Water for a Healthy Country National Research Flagship.

# Appendices

# Appendix A: Methods for linear weighting of soil moisture measurements

**OzFlux linear weighting:** Since the probe-based OzFlux measurement depths are variable across sites, the soil moisture is estimated for 0-30cm, 30-60cm and 60-90cm soil profile using linear weighting. The following eqs (12-15) are used for linear weighting of the measurements and a conceptual illustration is presented in Figure 50:

$$M_i = \frac{d_{i+1} - d_i}{2} + d_i \quad i = [1, n-1] \tag{12}$$

$$\begin{cases} if\ i = 1, & , ds - M_{n+1} \\ if\ 1 < i < n\ , & M_i = \frac{d_{i+1} - d_i}{2} + d_i \\ if\ i = n\ , & L_i = M_i \end{cases} \tag{13}$$

$$W_i = {}^{L_i}/_{ds} \tag{14}$$

$$P = \sum_{i=1}^{i=n} W_i s_i \tag{15}$$

where *di* represents sensor depths and $s_i$ represents sensor measurements correspond to $i^{th}$ sensor where $i \in [2, n]$. *ds* indicate shallow soil layer and $M_i$ indicates the midpoints between adjacent sensors, $L_i$ is total depth, $W_i$ is weighting coefficient (where $\sum W_i = 1$) and *P* indicates OzFlux profile soil moisture.

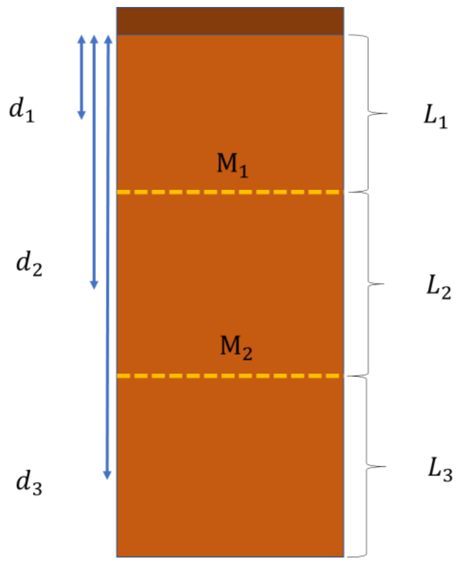



**Figure 50. Conceptual illustration of the calculation of OZFlux profile soil moisture**

**CosmOz linear weighing:** CosmOz soil moisture measurement depths (which vary through soil column) are correlated with percentage of soil moisture. This correlation is used to calculate a single estimate of soil moisture assumed for shallow layer as described in the following conceptual illustration and eqns (16-18):

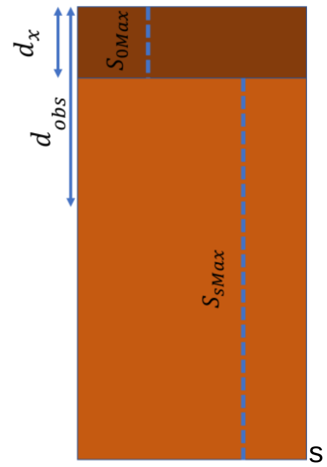



**Figure 51. Conceptual illustration of the calculation of CosmOz profile soil moisture**

$$d_x = \min(d_0, d_{obs}) \tag{16}$$

$$w_0 = \frac{S_0}{S_{0Max}}, w_s = \frac{S_s}{S_{sMax}} \tag{17}$$

$$S_{CosmOz} = \frac{w_0 \times d_x + w_s \times (d_{obs} - d_x)}{d_{obs}} \tag{18}$$

where $d_0$ is the thickness of upper soil layer. $S_{CosmOz}$ is weighted estimate of CosmOz referred to as upper profile to correspond with the CosmOz sensing depth, $d_{obs}$ is measurement depth, $S_0$ and $S_{0max}$ respectively represent soil moisture and the maximum capacity of soil moisture at top layer soil zone, $S_s$ and $S_{sMax}$ represent soil moisture and the maximum capacity of soil moisture at lower layer soil zone, respectively. $w_0$ and $w_s$ indicate relative wetness at soil upper and lower zone, respectively. Note that depending on the $d_x$ and $d_{obs}$ the eq (4) obtains smaller value for calculation of weighted estimate of CosmOz soil moisture.

# Appendix B: ET monitoring site details and time-series

**Table 17.Flux tower site details (data source noted)**

| Site Name | Citation | Temporal coverage |
|---|---|---|
| Adelaide Riv. | *Jason Beringer (2013 ) Adelaide River OzFlux tower site OzFlux: Australian and New Zealand Flux Research and Monitoring hdl: 102.100.100/14228* | 2007-01 - 2009-05 |
| Alice Springs | *James Cleverly (2011 ) Alice Springs Mulga OzFlux site OzFlux: Australian and New Zealand Flux Research and Monitoring hdl: 102.100.100/14217* | 2010-09 - 2013-12 |
| Calperum | *Calperum Tech (2013 ) Calperum Chowilla OzFlux tower site OzFlux: Australian and New Zealand Flux Research and Monitoring hdl: 102.100.100/14236* | 2010-01 - 2013-12 |
| Cumberland | *Elise Pendall (2015 ) Cumberland Plain OzFlux Tower Site OzFlux: Australian and New Zealand Flux Research and Monitoring hdl: 102.100.100/25164* | 2012-01 - 2013-12 |
| Daintree | *Mike Liddell (2013 ) Daintree Ozflux tower site OzFlux: Australian and New Zealand Flux Research and Monitoring hdl: 102.100.100/14244* | 2011-01- 2013-12 |
| Daly Pasture | *Jason Beringer (2013 ) Daly Pasture OzFlux tower site OzFlux: Australian and New Zealand Flux Research and Monitoring hdl: 102.100.100/14238* | 2007-01- 2010-05 |
| Daly Uncleared | *Jason Beringer (2013 ) Daly Uncleared OzFlux tower site OzFlux: Australian and New Zealand Flux Research and Monitoring hdl: 102.100.100/14239* | 2007-01- 2013-12 |
| Dry River | *Jason Beringer (2013 ) Dry River OzFlux tower site OzFlux: Australian and New Zealand Flux Research and Monitoring hdl: 102.100.100/14229* | 2008-01- 2013-12 |
| Fogg Dam | *Jason Beringer (2013 ) Fogg Dam OzFlux tower site OzFlux: Australian and New Zealand Flux Research and Monitoring hdl: 102.100.100/14233* | 2006-02- 2008-12 |
| Gingin | *Craig Macfarlane (2012 ) Gingin OzFlux: Australian and New Zealand Flux Research and Monitoring hdl: 102.100.100/14223* | 2011-01- 2013-11 |
| GWW | *Craig Macfarlane (2013 ) Great Western Woodlands OzFlux: Australian and New Zealand Flux Research and Monitoring hdl: 102.100.100/14226* | 2013-01- 2013-12 |
| Howard Spr | *Jason Beringer (2013 ) Howard Springs OzFlux tower site OzFlux: Australian and New Zealand Flux Research and Monitoring hdl: 102.100.100/14234* | 2001-01- 2013-12 |
| Nimmo | *Robert Simpson (2012 ) Nimmo High Plains OzFlux Tower Site OzFlux: Australian and New Zealand Flux Research and Monitoring hdl:* | 2007-01- 2013-12 |
| RDMF | *Jason Beringer (2014 ) Red Dirt Melon Farm OzFlux tower site OzFlux: Australian and New Zealand Flux Research and Monitoring hdl: 102.100.100/14245* | 2011-09- 2013-07 |
| Riggs Creek | *Jason Beringer (2014 ) Riggs Creek OzFlux tower site OzFlux: Australian and New Zealand Flux Research and Monitoring hdl: 102.100.100/14246* | 2011-01- 2013-12 |
| Robson Ck | *Mike Liddell (2013 ) Robson Creek Ozflux tower site OzFlux: Australian and New Zealand Flux Research and Monitoring hdl: 102.100.100/14243* | 2013-01- 2013-12 |
| Samford | *David Rowlings (2011 ) Samford Ecological Research Facility OzFlux tower site OzFlux: Australian and New Zealand Flux Research and Monitoring hdl:* | 2010-01- 2013-12 |
| Sturt Plains | *Jason Beringer (2013 ) Sturt Plains OzFlux tower site OzFlux: Australian and New Zealand Flux Research and Monitoring hdl: 102.100.100/14230* | 2008-01- 2013-12 |
| Ti Tree East | *James Cleverly (2013 ) Ti Tree East OzFlux Site OzFlux: Australian and New Zealand Flux Research and Monitoring hdl: 102.100.100/14225* | 2012-08- 2013-12 |
| Tumbarumba | *Eva vanGorsel (2013 ) Tumbarumba OzFlux tower site OzFlux: Australian and New Zealand Flux Research and Monitoring hdl: 102.100.100/14241* | 2001-01 - 2013-12 |
| Wallaby Ck | *Jason Beringer (2013 ) Wallaby Creek OzFlux tower site OzFlux: Australian and New Zealand Flux Research and Monitoring hdl: 102.100.100/14231* | 2005-01- 2012-12 |
| Warra | Emma White (2014 ) Warra OzFlux tower site OzFlux: Australian and New Zealand Flux Research and Monitoring hdl: 102.100.100/16188 | 2013-03 - 2013-12 |
| Whroo | *Jason Beringer (2013 ) Whroo OzFlux tower site OzFlux: Australian and New Zealand Flux Research and Monitoring hdl: 102.100.100/14232* | 2011-12- 2013-12 |
| Wombat | *Stefan Arndt (2013 ) Wombat State Forest OzFlux-tower site OzFlux: Australian and New Zealand Flux Research and Monitoring hdl: 102.100.100/14237* | 2010-01- 2013-12 |
| Yanco | *Jason Beringer (2013 ) Yanco JAXA OzFlux tower site OzFlux: Australian and New Zealand Flux Research and Monitoring hdl: 102.100.100/14235* | 2012-01- 2013-12 |

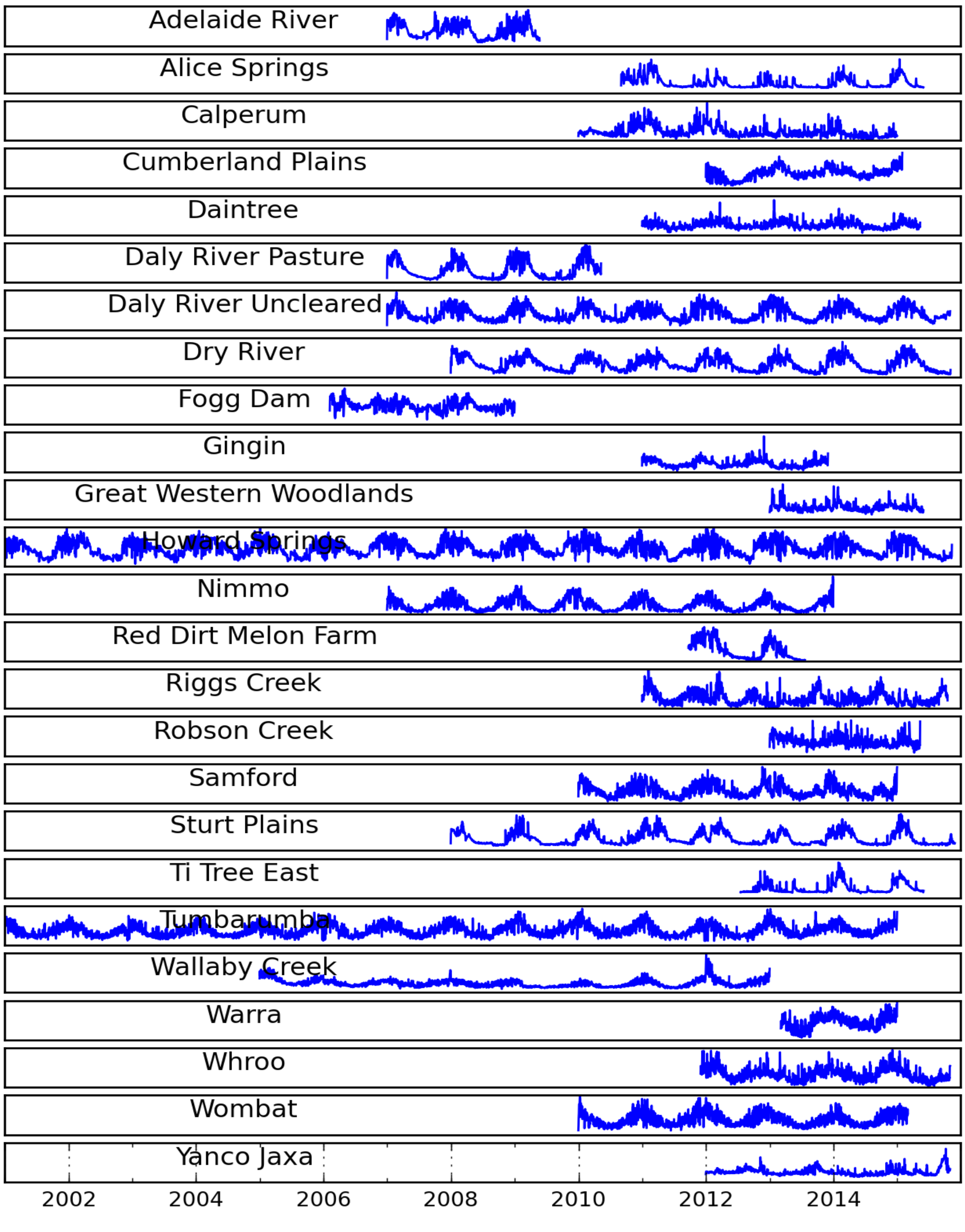



**Figure 52. Indicative site time-series of DINGO Evapotranspiration (mm). Axis scale omitted for space purposes.**

# Appendix C: Soil moisture monitoring site details and time-series

**Table 18. OzNet site details**

| OzNet | Start Date | End Date | Daily | Monthly |
|---|---|---|---|---|
| **A1** | 1/12/2001 | 31/05/2012 | 74% | 80% |
| **A2** | 1/12/2001 | 30/05/2011 | 37% | 41% |
| **A3** | 1/12/2001 | 30/11/2012 | 66% | 72% |
| **A4** | 1/12/2001 | 31/08/2011 | 34% | 36% |
| **A5** | 25/11/2001 | 5/02/2010 | 58% | 64% |
| **K1** | 15/11/2001 | 27/09/2012 | 73% | 76% |
| **K10** | 6/12/2003 | 31/05/2011 | 49% | 56% |
| **K11** | 6/11/2003 | 28/08/2009 | 46% | 47% |
| **K12** | 5/11/2003 | 31/05/2011 | 51% | 55% |
| **K13** | 16/11/2003 | 31/12/2013 | 65% | 71% |
| **K14** | 6/11/2003 | 31/05/2011 | 56% | 59% |
| **K2** | 16/11/2001 | 3/09/2010 | 69% | 72% |
| **K3** | 16/11/2001 | 24/08/2012 | 71% | 82% |
| **K4** | 15/11/2001 | 26/07/2012 | 80% | 84% |
| **K5** | 14/11/2001 | 25/06/2012 | 66% | 70% |
| **K6** | 5/11/2003 | 16/04/2013 | 62% | 70% |
| **K7** | 5/11/2003 | 31/05/2011 | 59% | 61% |
| **K8** | 5/11/2003 | 16/04/2013 | 52% | 60% |
| **M1** | 13/09/2001 | 1/02/2012 | 73% | 76% |
| **M2** | 13/09/2001 | 31/05/2013 | 79% | 84% |
| **M3** | 15/11/2001 | 31/05/2013 | 24% | 25% |
| **M4** | 15/09/2001 | 31/05/2011 | 75% | 79% |
| **M5** | 27/09/2001 | 15/12/2010 | 49% | 61% |
| **M6** | 27/09/2001 | 31/05/2011 | 71% | 77% |
| **M7** | 28/09/2001 | 1/02/2012 | 82% | 85% |
| **Y1** | 27/12/2003 | 31/12/2013 | 59% | 67% |
| **Y10** | 9/01/2004 | 31/12/2013 | 70% | 76% |
| **Y11** | 8/01/2004 | 31/12/2013 | 59% | 64% |
| **Y12** | 11/12/2003 | 31/12/2013 | 62% | 68% |
| **Y13** | 11/12/2003 | 31/12/2013 | 65% | 72% |
| **Y2** | 16/01/2004 | 31/12/2013 | 55% | 65% |
| **Y3** | 28/09/2001 | 17/04/2002 | 4% | 5% |
| **Y4** | 21/12/2003 | 23/06/2013 | 58% | 66% |
| **Y5** | 9/12/2003 | 28/02/2012 | 60% | 65% |
| **Y6** | 21/12/2003 | 20/10/2013 | 54% | 64% |
| **Y7** | 17/12/2003 | 31/12/2013 | 63% | 66% |
| **Y8** | 11/12/2003 | 31/12/2013 | 56% | 61% |
| **Y9** | 17/12/2003 | 25/12/2013 | 65% | 72% |

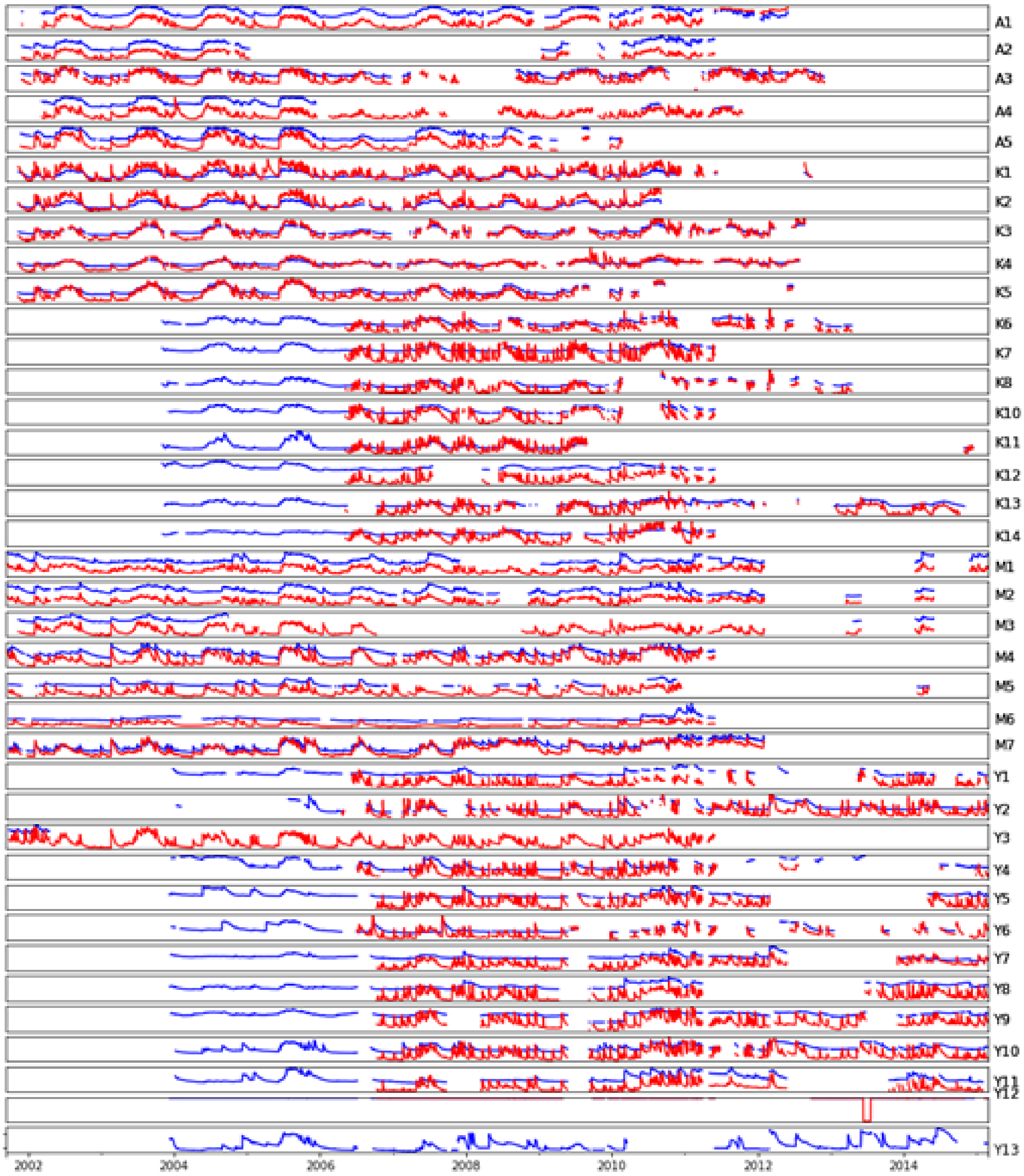



**Figure 53. Indicative site daily time-series of OzNet top layer (red: 0-5/8cm) and profile (blue: 0-90cm) volumetric soil moisture. Site numbers are listed far right. Axis scale omitted for space purpose**

**Table 19. SASMAS site details**

| Site | Start Date | End Date | Daily avail. | Monthly avail. |
|---|---|---|---|---|
| **G1** | 3/02/2003 | 16/10/2008 | 55% | 59% |
| **G2** | 3/02/2003 | 31/12/2006 | 34% | 38% |
| **G3** | 1/01/2003 | 31/12/2006 | 34% | 36% |
| **G4** | NA | NA | 0% | 0% |
| **G5** | 14/01/2003 | 6/03/2007 | 42% | 44% |
| **G6** | NA | NA | 0% | 0% |
| **K1** | 1/01/2003 | 31/12/2011 | 89% | 94% |
| **K2** | 1/01/2003 | 31/12/2011 | 90% | 97% |
| **K3** | 1/01/2003 | 31/12/2009 | 72% | 75% |
| **K4** | 1/01/2003 | 31/12/2010 | 74% | 76% |
| **K5** | 1/01/2003 | 31/12/2011 | 90% | 93% |
| **K6** | NA | NA | 0% | 0% |
| **M1** | NA | NA | 0% | 0% |
| **M2** | 1/01/2003 | 11/07/2007 | 49% | 51% |
| **M3** | NA | NA | 0% | 0% |
| **M4** | NA | NA | 0% | 0% |
| **M5** | NA | NA | 0% | 0% |
| **M6** | NA | NA | 0% | 0% |
| **M7** | 1/01/2003 | 31/12/2010 | 69% | 72% |
| **S1** | 4/02/2003 | 31/12/2010 | 82% | 84% |
| **S2** | NA | NA | 0% | 0% |
| **S3** | NA | NA | 0% | 0% |
| **S4** | NA | NA | 0% | 0% |
| **S5** | 4/02/2003 | 31/12/2011 | 88% | 94% |
| **S6** | NA | NA | 0% | 0% |
| **S7** | NA | NA | 0% | 0% |

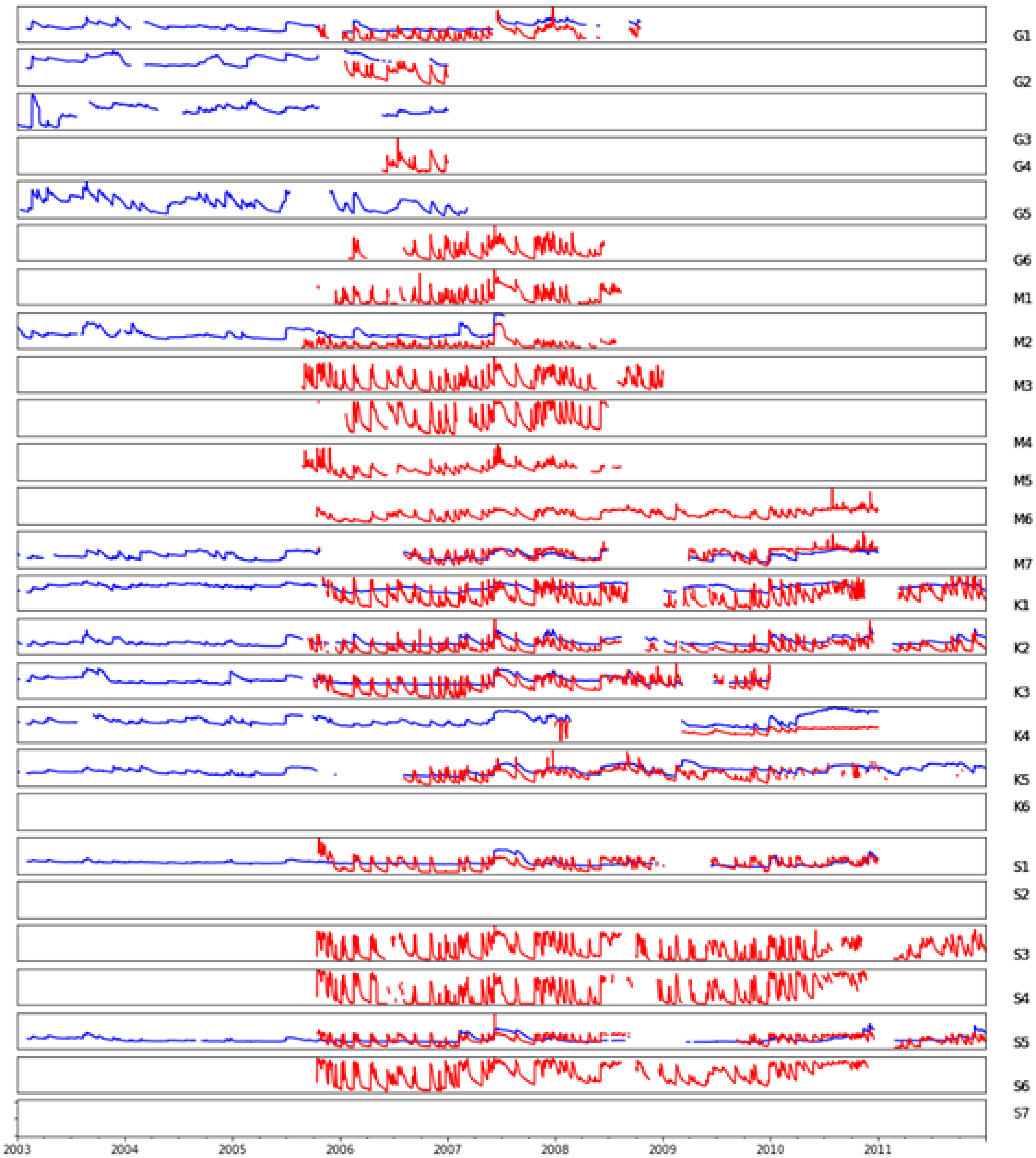



**Figure 54. Indicative site daily time-series of SASMAS top layer (red: 0-5/8cm) and profile (blue: 0-90cm) volumetric soil moisture. Site numbers are listed far right. Axis scale omitted for space purposes**

**Table 20. OzFlux (profile layer) site details**

| Site | Start Date | End Date | Daily avail. | Monthly avail. |
|---|---|---|---|---|
| OZFLUX_CAPETRIBULATION_SWS_10CMA | 7/01/2010 | 2/11/2018 | 67% | 76% |
| OZFLUX_COWBAY_SWS_6CMA | 1/07/2011 | 21/02/2019 | 58% | 66% |
| OZFLUX_DALYPASTURE_SWS_05 | 10/09/2007 | 8/09/2013 | 45% | 53% |
| OZFLUX_DALYUNCLEARED_SWS_5CM | 11/09/2007 | 16/02/2019 | 82% | 99% |
| OZFLUX_DRYRIVER_SWS_5CM | 1/09/2008 | 22/03/2019 | 62% | 75% |
| OZFLUX_HOWARDSPRINGS_SWS_10CMA | 9/02/2008 | 8/05/2019 | 87% | 99% |
| OZFLUX_HOWARDUNDERSTORY_SWS_10CMA | 6/09/2012 | 31/12/2016 | 33% | 37% |
| OZFLUX_LITCHFIELD_SWS_5CMA | 24/07/2015 | 8/05/2019 | 30% | 34% |
| OZFLUX_RIDGEFIELD_SWS_5CM | 23/03/2016 | 24/03/2018 | 16% | 18% |
| OZFLUX_ROBSON_SWS_6CMA | 26/01/2014 | 13/02/2019 | 41% | 45% |
| OZFLUX_STURTPLAINS_SWS_5CM | 29/08/2008 | 25/11/2014 | 40% | 47% |
| OZFLUX_WARRA_SWS_20A | 1/01/2016 | 31/12/2018 | 22% | 25% |
| OZFLUX_WOMBATSTATEFOREST_SWS_10CM_VN2 | 20/01/2010 | 1/03/2019 | 63% | 75% |
| OZFLUX_YANCO_SWS_3CM | 1/01/2014 | 8/05/2019 | 40% | 46% |

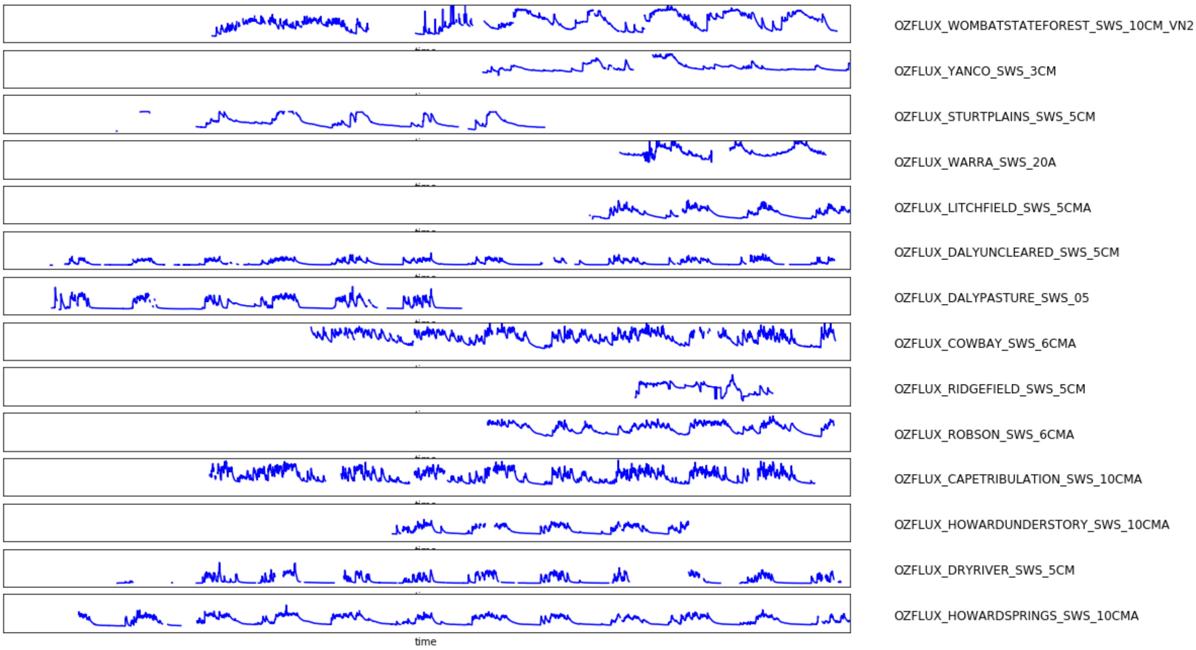



**Figure 55 Indicative site daily time-series of OzFlux profile volumetric soil moisture. Site numbers are listed far right. Axis scale omitted for space purposes.**

**Table 21. OzFlux (top layer) site details**

| Site | Start Date | End Date | Daily avail. | Monthly avail. |
|---|---|---|---|---|
| OZFLUX_COWBAY_SWS_6CMA | 4/09/2014 | 21/02/2019 | 34% | 38% |
| OZFLUX_CUMBERLANDPLAIN_SWS_20CMA | 16/11/2013 | 31/12/2017 | 33% | 36% |
| OZFLUX_CUMBERLANDPLAIN_SWS_20CMB | 16/11/2013 | 31/12/2017 | 33% | 36% |
| OZFLUX_CUMBERLANDPLAIN_SWS_8CMA | 16/11/2013 | 12/05/2019 | 43% | 49% |
| OZFLUX_CUMBERLANDPLAIN_SWS_8CMB | 16/11/2013 | 12/05/2019 | 43% | 49% |
| OZFLUX_CUMBERLANDPLAIN_SWS_CALC | 19/10/2012 | 31/12/2013 | 9% | 11% |
| OZFLUX_CUMBERLANDPLAIN_SWS_MERGE | 1/01/2013 | 31/12/2013 | 8% | 9% |
| OZFLUX_DALYPASTURE_SWS_05 | 10/09/2007 | 8/09/2013 | 45% | 53% |
| OZFLUX_DALYUNCLEARED_SWS_5CM | 11/09/2007 | 16/02/2019 | 38% | 47% |
| OZFLUX_DALYUNCLEARED_SWS_5CMA | 1/01/2008 | 31/12/2013 | 46% | 53% |
| OZFLUX_DRYRIVER_SWS_5CM | 1/09/2008 | 22/03/2019 | 66% | 77% |
| OZFLUX_LITCHFIELD_SWS_5CMA | 24/07/2015 | 8/05/2019 | 30% | 34% |
| OZFLUX_LITCHFIELD_SWS_5CMB | 8/08/2015 | 8/05/2019 | 28% | 33% |
| OZFLUX_OTWAY_SWS_7.5CM | 15/05/2009 | 31/12/2010 | 12% | 15% |
| OZFLUX_RIDGEFIELD_SWS_5CM | 18/03/2016 | 24/03/2018 | 12% | 18% |
| OZFLUX_RIGGS_SWS_5CMA | 1/01/2011 | 12/07/2017 | 43% | 53% |
| OZFLUX_RIGGS_SWS_5CMB | 1/01/2011 | 12/07/2017 | 32% | 41% |
| OZFLUX_ROBSON_SWS_6CMA | 1/09/2014 | 13/02/2019 | 36% | 39% |
| OZFLUX_STURTPLAINS_SWS_5CM | 1/01/2012 | 22/02/2018 | 48% | 53% |
| OZFLUX_YANCO_SWS_3CM | 1/01/2014 | 8/05/2019 | 40% | 46% |

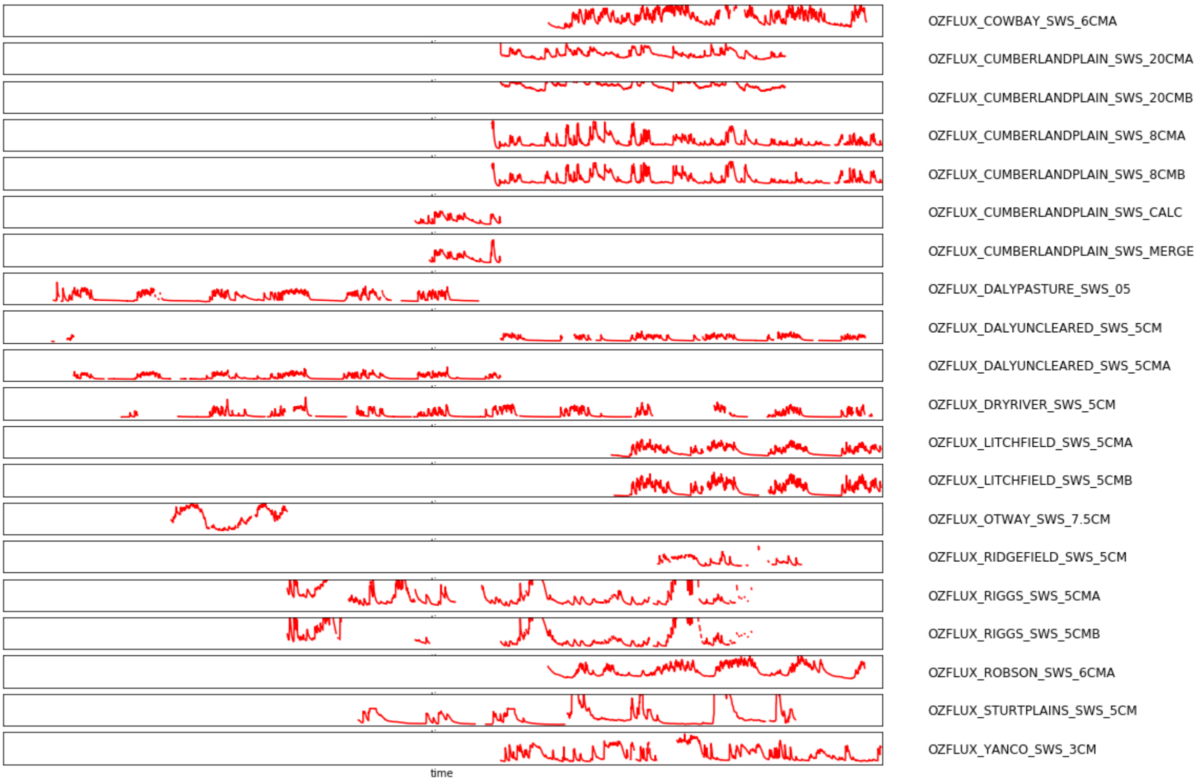



**Figure 56 Indicative site daily time-series of OzFlux top layer volumetric soil moisture. Site numbers are listed far right. Axis scale omitted for space purposes.**

**Table 22. CosmOz site details**

| Site | Start Date | End Date | Daily avail. | Monthly avail. |
|---|---|---|---|---|
| 01 BALDRY | 30/03/2011 | 12/03/2014 | 35% | 36% |
| 02 DALY | 7/06/2011 | 24/04/2019 | 75% | 90% |
| 03 GNANGARA | 17/05/2011 | 6/05/2018 | 47% | 68% |
| 04 GRIFFITH | 1/10/2011 | 8/05/2013 | 17% | 19% |
| 06 ROBSON | 28/10/2010 | 8/11/2018 | 90% | 95% |
| 07 TEMORA | 17/05/2013 | 20/11/2018 | 44% | 47% |
| 08 TULLOCHGORUM | 15/12/2010 | 17/01/2019 | 95% | 96% |
| 09 TUMBARUMBA | 3/04/2011 | 13/12/2018 | 77% | 87% |
| 10 WEANY | 2/12/2010 | 29/01/2019 | 94% | 96% |
| 11 YANCO | 1/04/2011 | 18/11/2018 | 89% | 90% |
| 12 NAMADGI | 23/08/2014 | 3/03/2019 | 50% | 55% |
| 13 MINERAL BANKS | 6/12/2013 | 27/09/2018 | 29% | 37% |
| 15 HAMILTON | 1/07/2015 | 4/02/2019 | 42% | 43% |
| 18 BISHES | 12/04/2016 | 22/09/2017 | 17% | 18% |
| 19 BENNETS | 12/04/2016 | 9/10/2018 | 25% | 30% |
| 21 BULLAWARRIE | 26/07/2016 | 11/12/2018 | 28% | 29% |

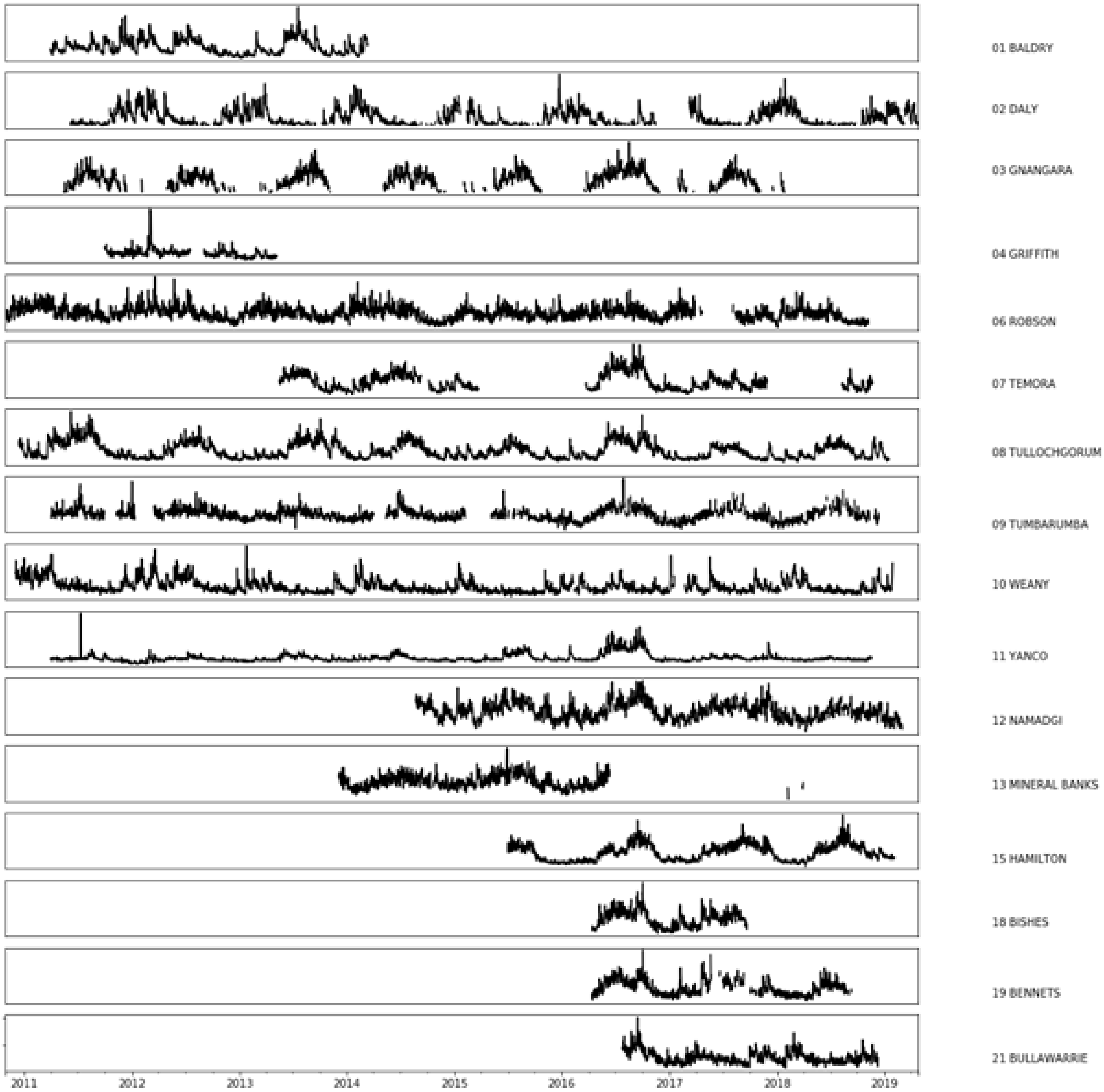



**Figure 57. Indicative site daily time-series of CosmOz volumetric soil moisture. Site numbers are listed far right. Axis scale omitted for space purposes.**

# Appendix D: Evaluation against top layer soil moisture

Top layer soil moisture was compared against point data sets using the time period that the satellite data was available for the OzNet Murrumbidgee (Figure 58), SASMAS Hunter data (Figure 59), OzFlux, (Figure 60). This gives an indication of how well the satellite data represents surface and profile soil moisture, compared to AWRA-L. CABLE and WaterDyn perform better than AWRA-L. AWRA-L v7 improves over AWRA v6 ASCAT appears to perform slightly better than AMSRE in general. AWRA-L performs better than AMSRE for surface soil moisture for the SASMAS Hunter sites, but not as well for the Murrumbidgee (for daily data also). For OzFlux data, AWRA-L v7 improves over AWRA-L v6 for monthly correlation while its performance is slightly degraded for daily correlation.

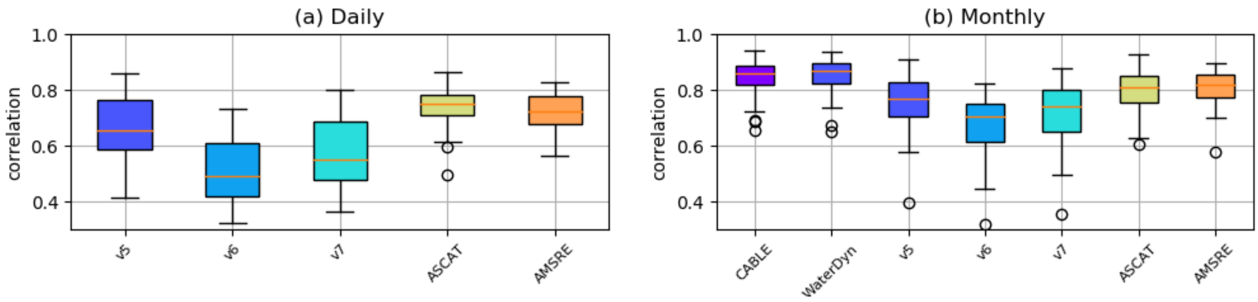



**Figure 58. Correlation of (a) daily and (b) monthly top layer (0-5/8cm) soil moisture of models against Murrumbidgee OzNet for Jan 2007-Sept 2011.**

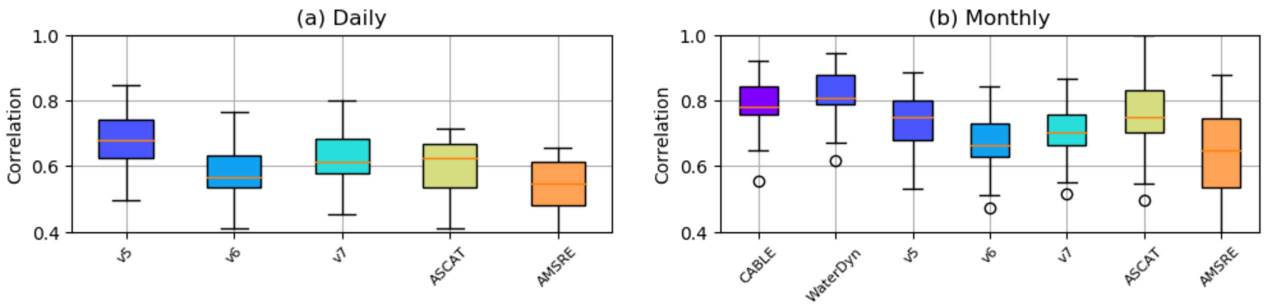



**Figure 59. Correlation of (a) daily and (b) monthly top layer (0-5/8cm) soil moisture of models against Upper Hunter SASMAS data for Jan 2007-Sept 2011.**

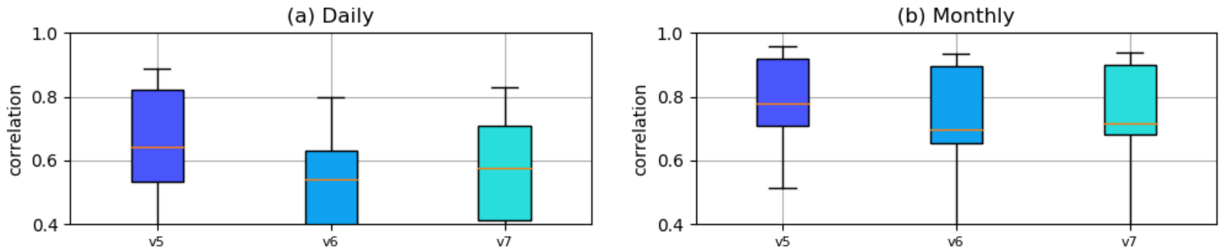



**Figure 60. Correlation of (a) daily and (b) monthly top layer (0-5/8cm) soil moisture of models against OzFlux data for 2007- 2017.**

Satellite data provides relatively accurate estimate of monthly and daily (not shown) surface soil moisture compared to AWRA-L in some areas (OzNet Murrumbidgee), while the model based estimates are superior in other areas (SASMAS). The difference in AMSR-E performance over the two areas follows that found (in comparing AMSR-E performance in these two areas) by Draper et al., ( 2009), noting that AMSR-E is no longer operational. There are multiple candidate satellite derived products available for evaluation/assimilation/calibration of water balance/land-surface models – and some debate over which satellite is best to use. See Lacava et al. (2012) for comparison of SMOS, AMSR-E and ASCAT, Leroux et al. (2014) for a comparison of SMOS, VUA (AMSR-E), ASCAT satellite based and ECMWF model forecast for surface soil moisture, and the subsequent clarification paper by Wagner et al. (2014) presenting differing results depending on the version of satellite data used and analysis method. However, all products do show use in terms of correlation to surface soil moisture. These datasets therefore serve as valuable tools for verification and calibration of AWRA-L, and more recent products will be evaluated for this purpose in future. In v7 top layer soil moisture equations are between v5 and v6, following the change in drainage equation in v6 (and resulting degradation of performance) and change back to a similar approach in v5 with v7. This approach can be tuned to be even closer to v5 if required in future.

# Appendix E: Details of catchments including impervious area

**Table 23. List of catchments added for testing AWRA-L including a significant proportion of impervious area**

| Catchment ID | River | Gauge name | % impervious | State |
|---|---|---|---|---|
| **5201.1** | North West Bay Rivulet | Margate Ws Int | 5 | TAS |
| **407221** | Jim Crow Creek | Yandoit | 10 | VIC |
| **229681B** | Wandin Yallock Creek | Seville East | 15 | VIC |
| **229627A** | Merri Creek | Summerhill Road Craigieburn | 17 | VIC |
| **212049** | Ropes Creek | Debrincat Avenue | 26 | NSW |
| **229618A** | Diamond Creek | Bridge Street Eltham | 14 | VIC |
| **16200.1** | Don River | Us Old Bass Hwy | 5 | TAS |
| **229215B** | Woori Yallock Creek | Woori Yallock | 13 | VIC |
| **213005** | Toongabbie Creek | Briens Road | 86 | NSW |
| **230204** | Riddells Creek | Riddells Creek | 14 | VIC |
| **405212** | Sunday Creek | Tallarook | 9 | VIC |
| **14209.1** | Claytons Rivulet | D/S Bass Hwy | 5 | TAS |
| **212320** | South Creek | Elisabeth Drive | 11 | NSW |