Australian Government
Bureau of Meteorology

Evaluation of the Bureau's Operational AWRA-L Model

Comparison of AWRA-L against Observed Hydrological Data and Peer Models

Citation: Frost, A. J., Ramchurn, A., Hafeez, M. (2016). Evaluation of the Bureau's Operational AWRA-L Model. Bureau of Meteorology Technical Report.

Version number/type	Date of issue
12	16/11/2016

Published by the Bureau of Meteorology

Contact details

Dr. Andrew Frost

Manager (acting) - Water Resources Modelling Unit

Bureau of Meteorology
PO Box 413, Darlinghurst NSW 1300

Tel: +61 2 9296 1517
Email: Andrew.Frost@bom.gov.au

Dr. Amgad Elmahdi

Section Head, Water Resources Assessment

Bureau of Meteorology
GPO Box 1289 Melbourne VIC 3001

Tel: +61 3 8638 8274
Email: Amgad.Elmahdi@bom.gov.au

Summary

This technical report details the scientific evaluation of the Bureau of Meteorology (BoM: here after called the Bureau) operational Australian Water Resources Assessment Landscape (AWRA-L version 5) modelling system using a range of the best available measurements/estimates of hydrological data including streamflow, soil moisture, actual evapotranspiration (ET) and groundwater recharge across the selected catchments/sets of point measurements at the continental scale. In addition, the performance of the operational AWRA-L version 5 model (here after called AWRA-L) is compared to two other national gridded land-surface peer models i.e. CABLE-SLI and WaterDyn. AWRA-L is also compared with individual conceptual rainfall runoff models using at-site calibration and nearest neighbour regionalisation for streamflow prediction purposes.

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 over which future model improvements can be compared against using the same comparison data. Aspirational targets for overall performance are also provided.

The results show that AWRA-L v5 performs relatively well according to streamflow nationally (with 295 unimpaired catchments used in calibration, and 291 separate catchments used in validation), rootzone (0-100cm) soil moisture, but relatively under perform in comparison for these two models for ET. Preliminary comparison of AWRA-L model deep drainage output against a long term average and annual time-series recharge dataset showed that observed drainage biases are driven predominantly by the saturated hydraulic conductivity rather than rainfall variability, noting high uncertainty in these recharge estimates. The better performance of AWRA-L model according to streamflow is due to better nationwide calibration and conceptual hydrological structure. CABLE is equivalent to AWRA-L in terms of soil moisture, and better according to ET as expected from its purpose as a model for land/atmosphere exchange model, along with calibration to flux tower and derived catchment ET. WaterDyn performs well for ET, but performs relatively worse for streamflow and root zone soil moisture.

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 is considered most fit for purpose water balance estimation purposes for water resources and agricultural applications.

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 for scientific understanding of hydrological processes. Overall, the spatial plots and the 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.

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

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

WIA: Water in Australia

WIRADA: Water Information Research and Development Alliance

1 Introduction

The Australian Water Resources Assessment Modelling System (AWRAMS) underpins the Bureau water information services that are mandated through the Water Act (2007). The science of AWRAMS (AWRAMS; see Elmahdi et al., 2015; Hafeez et al., 2015a; Hafeez et al., 2015b; Vaze et al., 2013) has been developed since July 2008 through the Water Information Research and Development Alliance (WIRADA) between CSIRO and the Australian Bureau of Meteorology (BoM). The AWRAMS has been operational at the Bureau since 2011-12 for regular use in the National Water Account (NWA) and water resources assessment reports. The AWRAMS has evolved from AWRA v 0.5 (2008) to AWRA v 5.0 (2015) with AWRA v 5.0 currently being used for reporting purposes by the Bureau. While a prototype AWRAMS was developed through WIRADA, the AWRAMS has been significantly refactored and enhanced through the Bureau AWRAMS Implementation (AWRAMSI) project over the last three years for superior performance and less simulation and calibration time. The operational AWRAMS has been used towards supplying retrospective water balance estimates published by the BoM within:

- Water in Australia (www.bom.gov.au/water/waterinaustralia): an annual national picture of water availability and use in a particular financial year
- Water resources assessments produced prior to Water In Australia (www.bom.gov.au/water/awra)
- Regional water information water resource assessments (www.bom.gov.au/water/rwi)
- National Water Account (NWA: www.bom.gov.au/water/nwa): that provides an annual set of water accounting reports for ten nationally significant water resource management regions. Adelaide, Burdekin, Canberra, Daly, Melbourne, Murray–Darling Basin, Ord, Perth, South East Queensland and Sydney.

The Bureau's AWRAMSI Project recoded the WIRADA prototype to make an operational AWRAMS that is more efficient, functional, and easily maintainable in a Linux platform. It is a Python based modelling system, with the core model algorithms implemented in high performing native languages (Fortran, C) and generic functionality provided by robust, open source libraries. The operational AWRAMS simulates Australian landscape and river water stores and fluxes for the past 100 years to now (Hafeez et al., 2015). These estimates are updated on a daily basis and provide the current and historical context of water availability in Australia. There are two main components to the AWRA modelling system:

- **AWRA-L**: a one dimensional, 0.05 degree grid based landscape water balance model over the continent that has semi-distributed representation of the soil, groundwater and surface water stores. The AWRA-L model, operational since November 2015, publishes daily updated outputs to the public, with daily gridded soil moisture, runoff, evapotranspiration, and deep drainage outputs (see Figure 1) available from yesterday back to 1911 online through www.bom.gov.au/water/landscape.
- **AWRA-R:** a node link network conceptual river model designed for both regulated and unregulated river system. Currently, implemented over a few regions (MDB, Melbourne and SEQ) for national water account purposes.

Since the operational AWRA-L modelled outputs have been made publicly available in November 2015, the modelled fluxes have been used internally and externally for various climatological, flood, water and agriculture applications across Australia. The Bureau's AWRA team has been regularly interacting with a wide range of stakeholders about their needs and how these can be met by a daily operational water balance model. These interactions have spanned Commonwealth agencies and State government water and agriculture agencies, catchment management authorities, water utilities, consultants, water industry professionals, research organisations, universities and farmers.

This technical report evaluates and compares the hydrologic performance of the AWRA-L model with a national water balance model (WaterDyn) and a global biogeochemical land surface scheme (CABLE), applied regionally. These models were compared against catchment streamflow, point estimates of flux tower derived evapotranspiration across Australia, and point estimates of 0-90cm profile soil moisture over the Murrumbidgee and Upper Hunter Catchments. Satellite derived estimates of evapotranspiration (CMRSET, SLST) and soil moisture (AMSR-E, ASCAT) are also compared to the point based observations.

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. Finally, locally calibrated nearest neighbour regionalised conceptual rainfall runoff models are also compared to AWRA-L, WaterDyn and CABLE to see how they perform relatively for streamflow (given a range of outputs are provided by the national models).

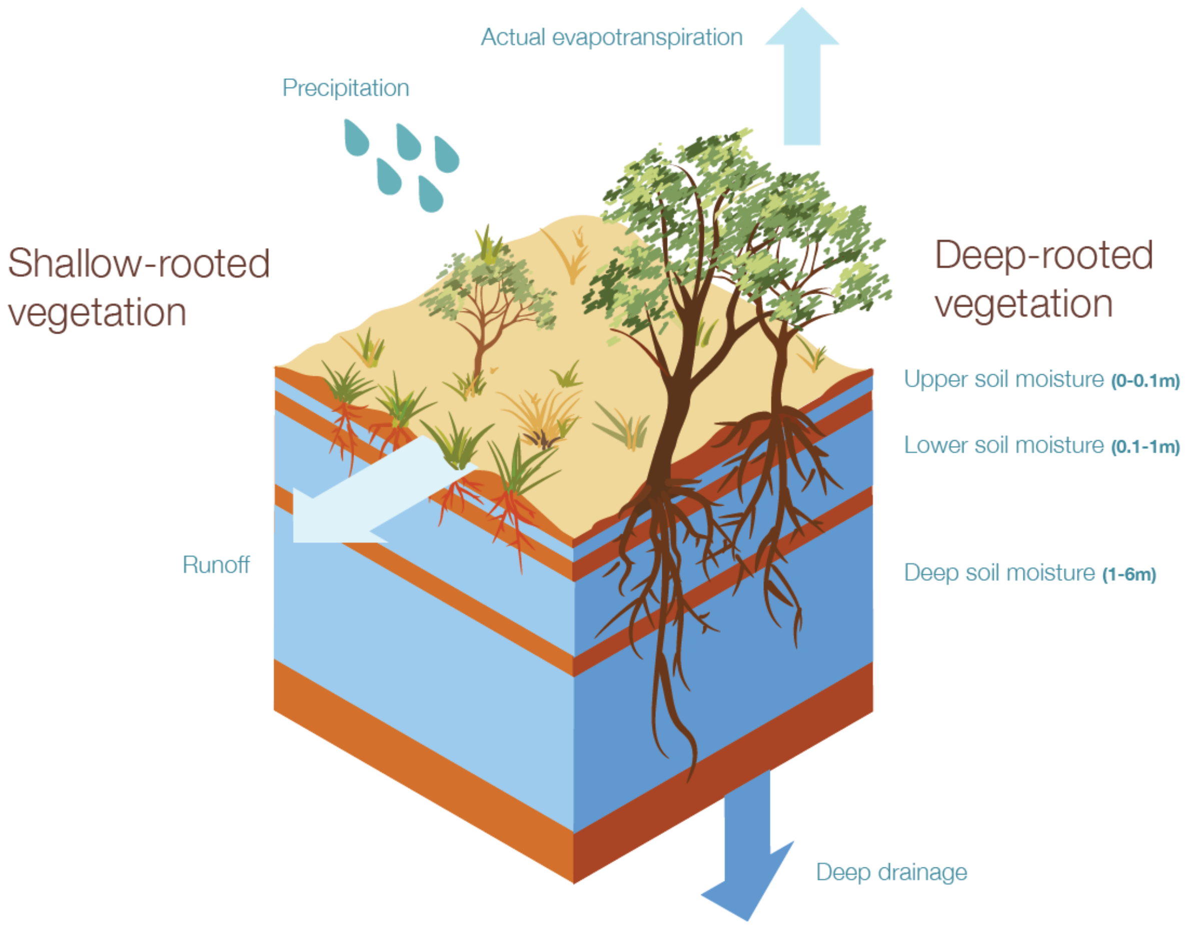


Figure 1. Conceptual AWRA-L grid cell with key water stores and fluxes shown

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 good 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, many universities and research groups have established an excellent ground based network since 2000 onwards for physical measurement of various 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 daily gridded Australian Water Availability Project (AWAP) climate data set that consists of air temperature (daily minimum and maximum) and daily precipitation from 1st January 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. 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 are covering at least the period of 1950 until 2013. For soil moisture evaluations, 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).

2.2 Evaluation data

2.2.1 Streamflow

A set of 782 unimpaired catchments with gauged flow records in unimpaired across Australia were collated Zhang et al. (2013), according to the following criteria: (a) catchment area is greater than 50 km^2, (b) the stream is unregulated (no dams or reservoirs), (c) no major impacts of irrigation and land use, (d) observed record has at least 10 years of data between 1975 and 2011. The catchments (delineated using the BoM's national catchment Geofabric product: www.bom.gov.au/water/geofabric) were collated towards being used in evaluation. The spatial distribution of catchments reserved for calibration and validation of AWRA-L is shown in Figure 2; with regional divisions showing areas of similar climate. 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 are used for validation.

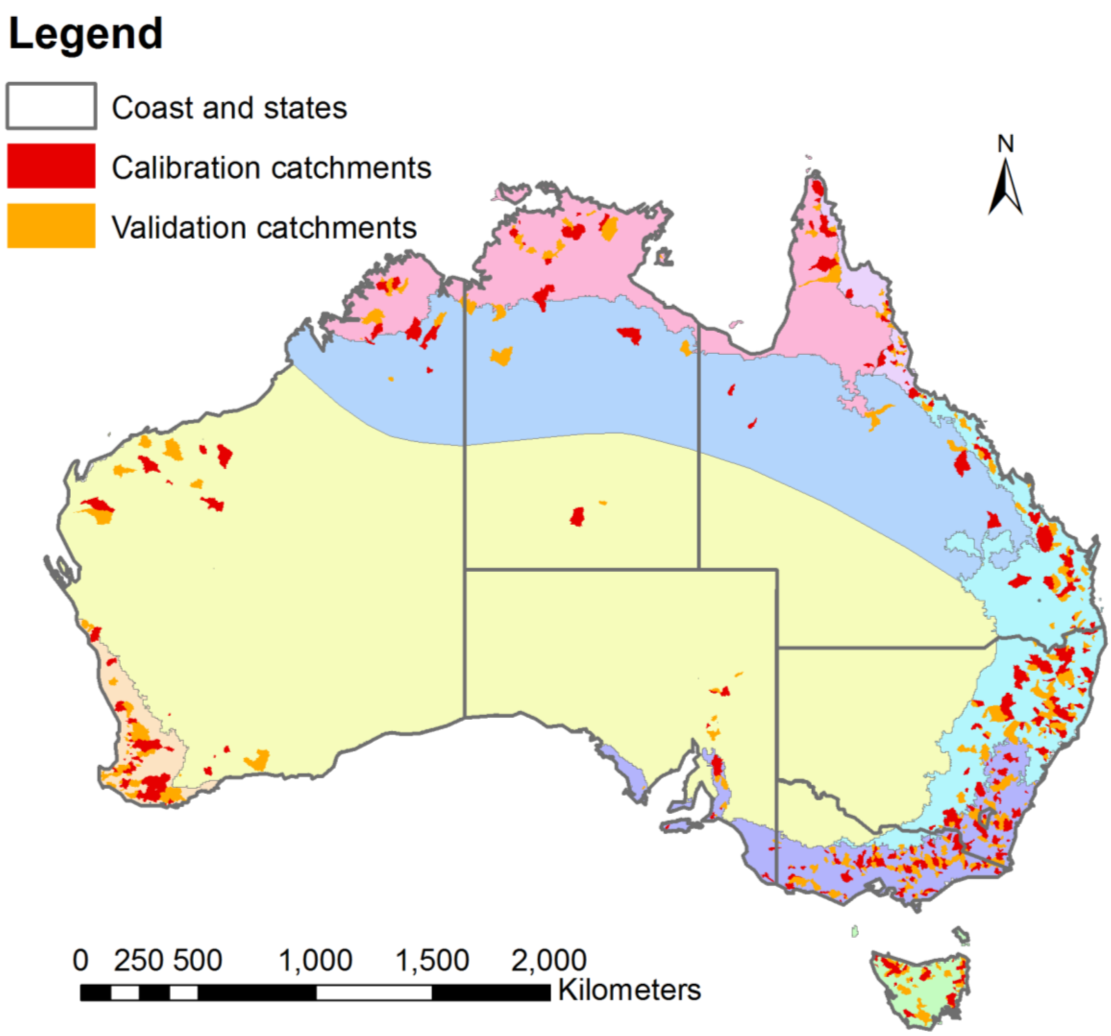


Figure 2. Location of unimpaired catchments used for model evaluation

2.2.2 Soil moisture

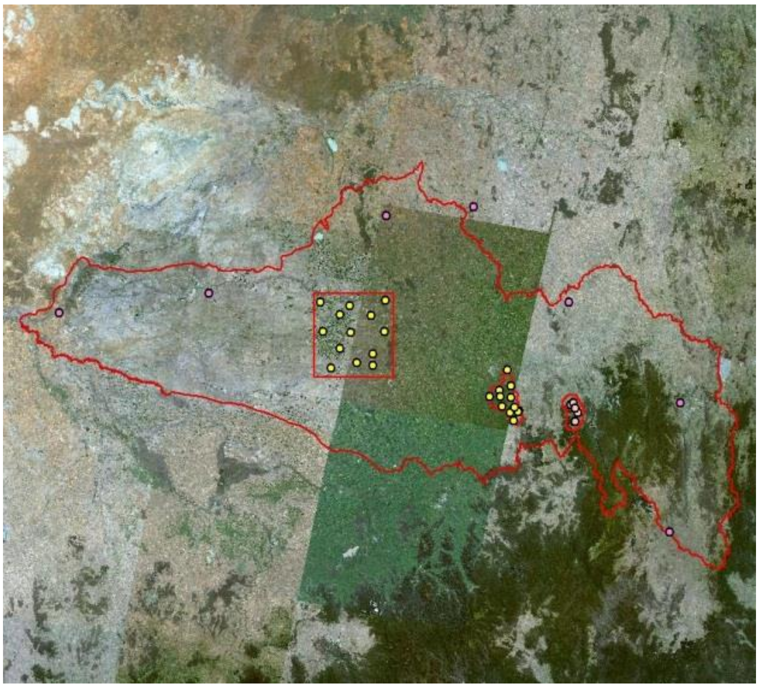
The following soil moisture data sets have been used for evaluation of the modelled outputs.

OzNet network: Time series 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, NSW (see Figure 3(a)) were used in evaluation (and not calibration) of the models. These time series were derived from reflectometer measurements from the OzNet network, setup and maintained by the University of Melbourne and Monash University (Smith et al., 2012). The reflectometers were calibrated according to independent measurements (Rüdiger et al., 2010), and had a median of 67% of monthly data available over the 2001 – 2013 period considered.

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, NSW (Rüdiger et al., 2007) - see Figure 3(b) - were used in evaluation (and not calibration) of the models. These time series were collated as part of the SASMAS project monitoring sites (managed by the University of Newcastle). There were 13 sites with profile (0-90cm) data available, 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 purposes. ASCAT is a Technische Universitat Wien (TUW) product (Bartalis et al., 2007), 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 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 has been used for calibration of the AWRA-L model covering the same time period – see section 3.1.

(a)



(b)

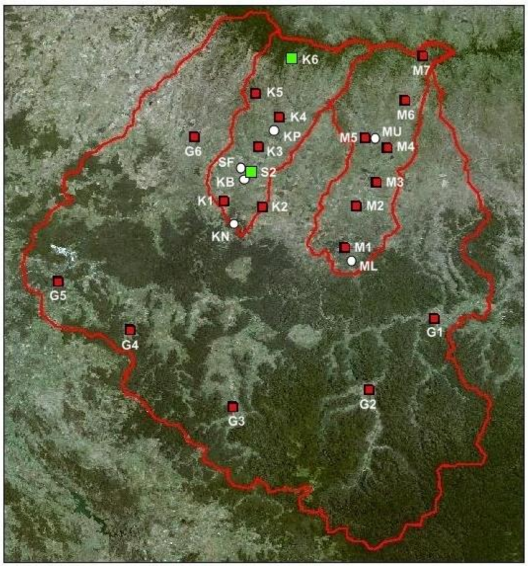


Figure 3. (a) OzNet Murrumbidgee soil moisture (from www.oznet.org.au/murrumbidgeesm.html), (b) SASMAS Goulburn soil moisture (from www.eng.newcastle.edu.au/sasmas/SASMAS/sasdata.html)

2.2.3 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 (Beringer, Hutley et al., 2016a; www.ozflux.org.au; see Figure 4 for locations (see Appendix A: 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, McHugh et al., 2016). 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).

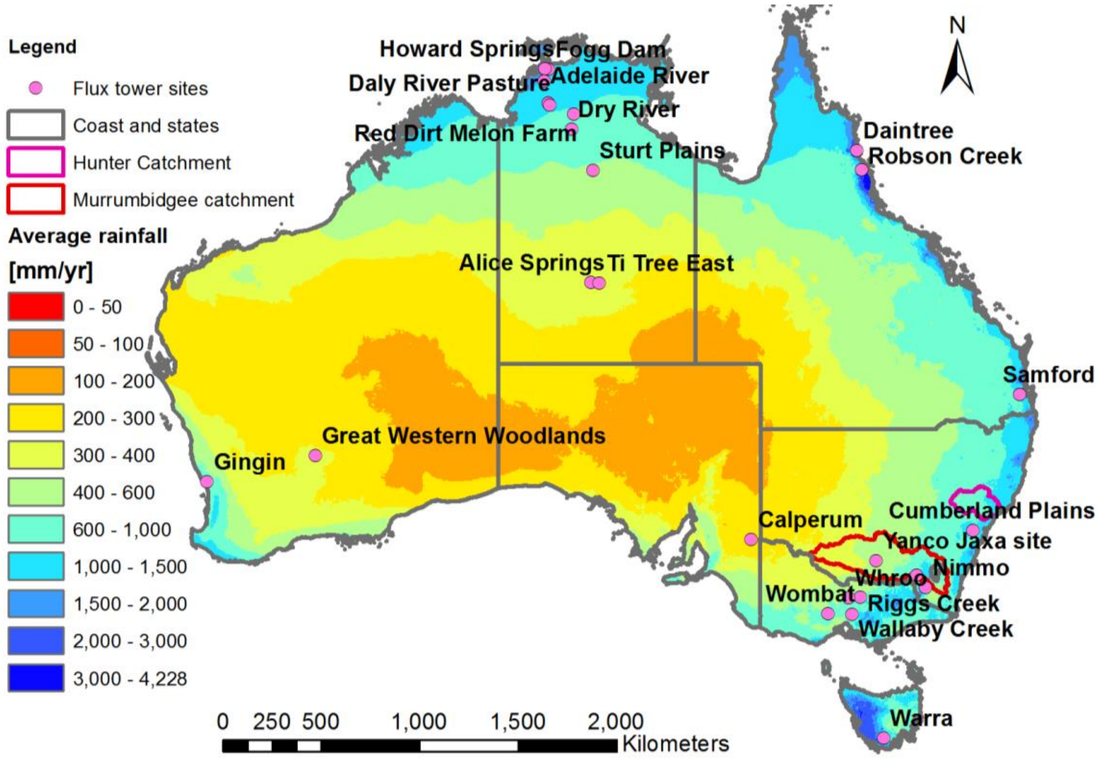


Figure 4. ET flux towers locations and soil moisture monitoring catchment locations also shown. Average annual rainfall is shown to give an indication of the range of climate conditions sampled within Australia.

The period 2001-2013 was used for scientific evaluation, being the intersection of years available for all models available, and had a median of 30% 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 2001-2010 and the CSIRO developed Simplified Land Surface Temperature (SLST) algorithm (Van Niel et al., 2012), were compared to the observed point estimates. CMRSET was used in AWRA-L calibration, and also for evaluation purposes. CMRSET is run operationally within the Bureau and 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 5.

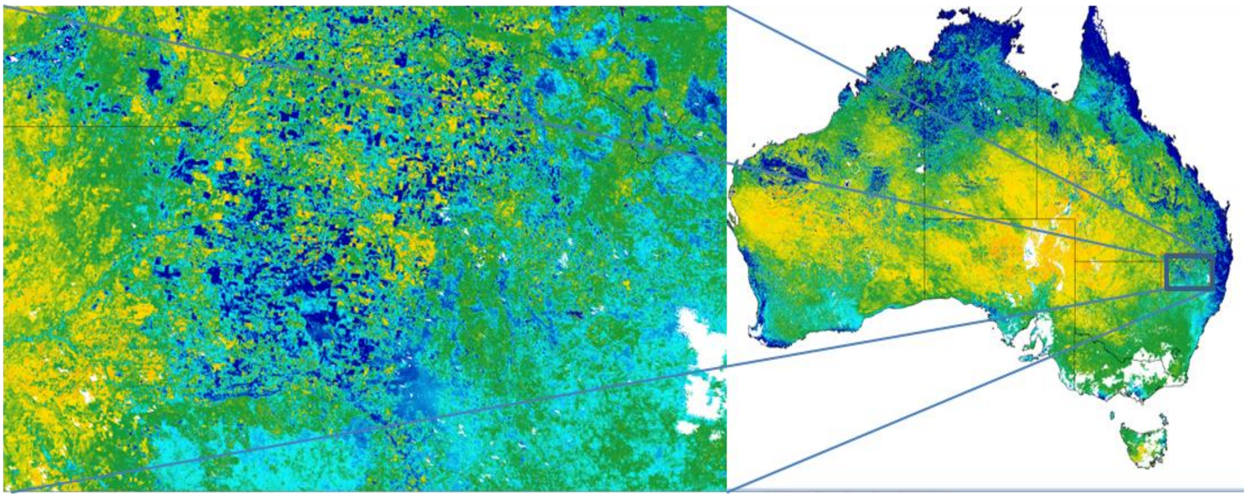


Figure 5. CMRSET derived map of 8 day Actual Evapotranspiration for 04/07/2014 (noting white area shows no data, most likely affected by clouds). Courtesy Juan Pablo-Guerschman CSIRO.

2.2.4 Groundwater Deep Drainage

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

1. **Long term average:** A long term average recharge dataset has been processed from 6343 individual field estimates of estimates collated by Crosbie et al (2010a;

2010b) with some additional points added that were generated from the Bioregional Assessment Programme (www.bioregionalassessments.gov.au). It 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 6. The majority of the recharge estimates are based on chloride mass balance estimates, which represent long-term mean annual recharge at the point.

2. **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 7.

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

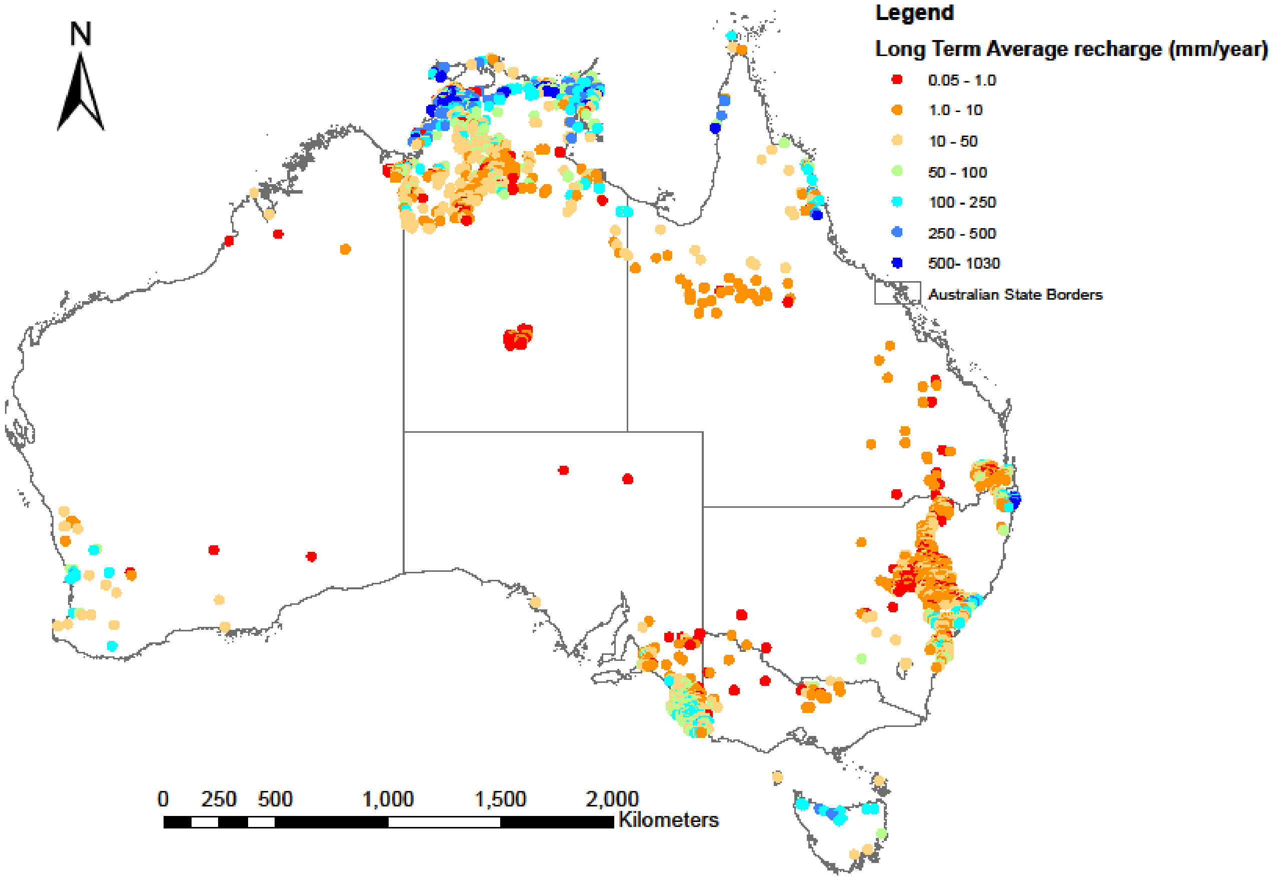


Figure 6. Shi et al (2015) long term average recharge estimates

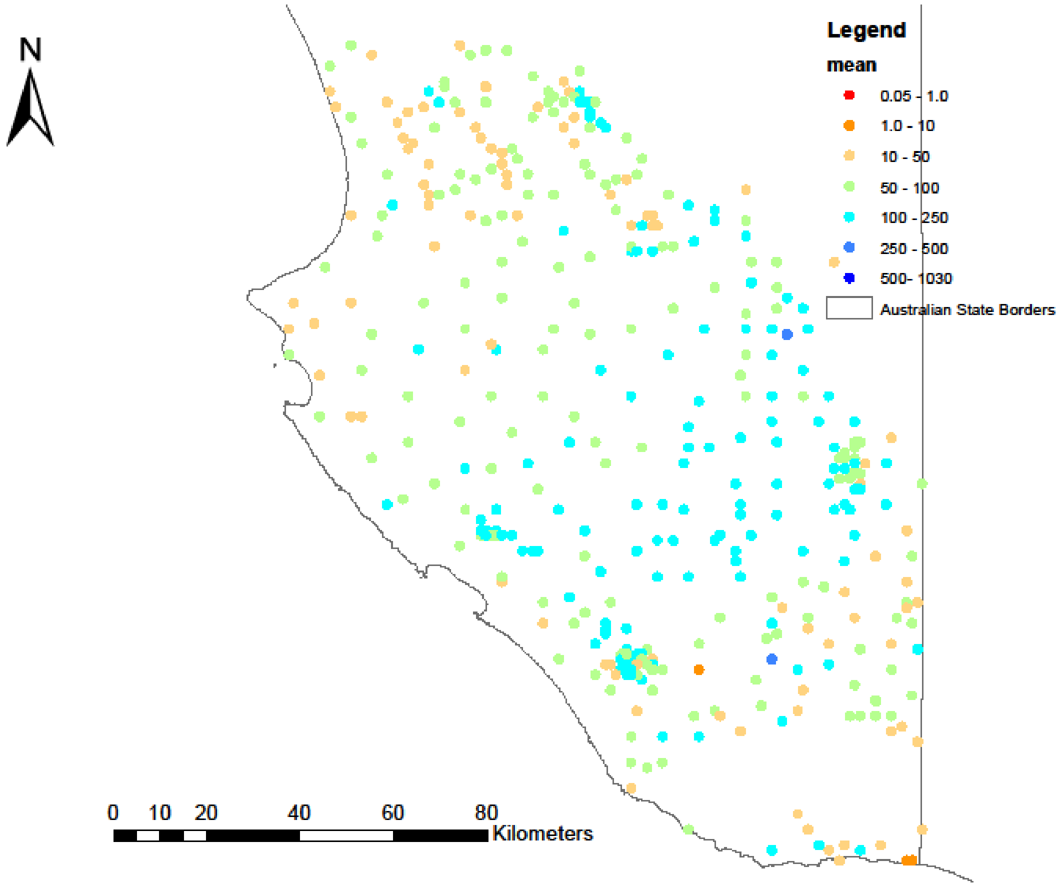


Figure 7. Shi et al (2015) location of annual estimates in South Australia/Victorian state border.

Considering huge variability of deep drainage at any point, and uncertainties associated with derivation of evaluation datasets, it is very hard to have absolute validation of modelled deep drainage of a 25 km^2 grid with the field data. The Bureau is working on a case study with Murray Darling Basin Authority (MDBA) on evaluating AWRA-L modelled deep drainage fluxes through comparison with field data and peer models over groundwater Sustainable Diversion Limit (SDL) areas in the Murray Darling Basin.

3 Models

3.1 AWRA-L

AWRA-L (Van Dijk, 2010; Viney et al., 2014; Viney et al., 2015) 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. AWRA-L is a three soil layer (top: 0-10cm, shallow: 10cm-100cm, deep: 100cm-600cm), two hydrological response unit (shallow rooted versus deep rooted) model (Figure 8).

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 2 (shallow and deep rooted) hydrological response units.

Various spatial datasets are also used to parameterise AWRA-L spatially (with key examples shown in Figure 9) 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) and importantly the proportion deep/shallow rooted (based on estimate of fraction persistent and recurrent vegetation as derived by Donohue et al., 2008).
- Slope and hydraulic conductivity affecting infiltration capacity
- Soil drainage/storage parameters:
- soil hydraulic conductivity using the Dane and Puckett (1994) pedotransfer function applied to clay content from the Soil and Landscape Grid of Australia (www.clw.csiro.au/aclep/soilandlandscapegrid)
- fractional water storage capacity from Australian Soil Resource Information System (ASRIS) level 4 (Johnston et al., 2003)
- Topology and effective porosity effecting baseflow/saturation (Peeters et al., 2011)

The top, middle and deep soil layer depths within AWRA-L are chosen to be 0.1m, 1m and 6m respectively. For further details of the AWRA-L v5.0 algorithms and input data see Viney et al (2015).

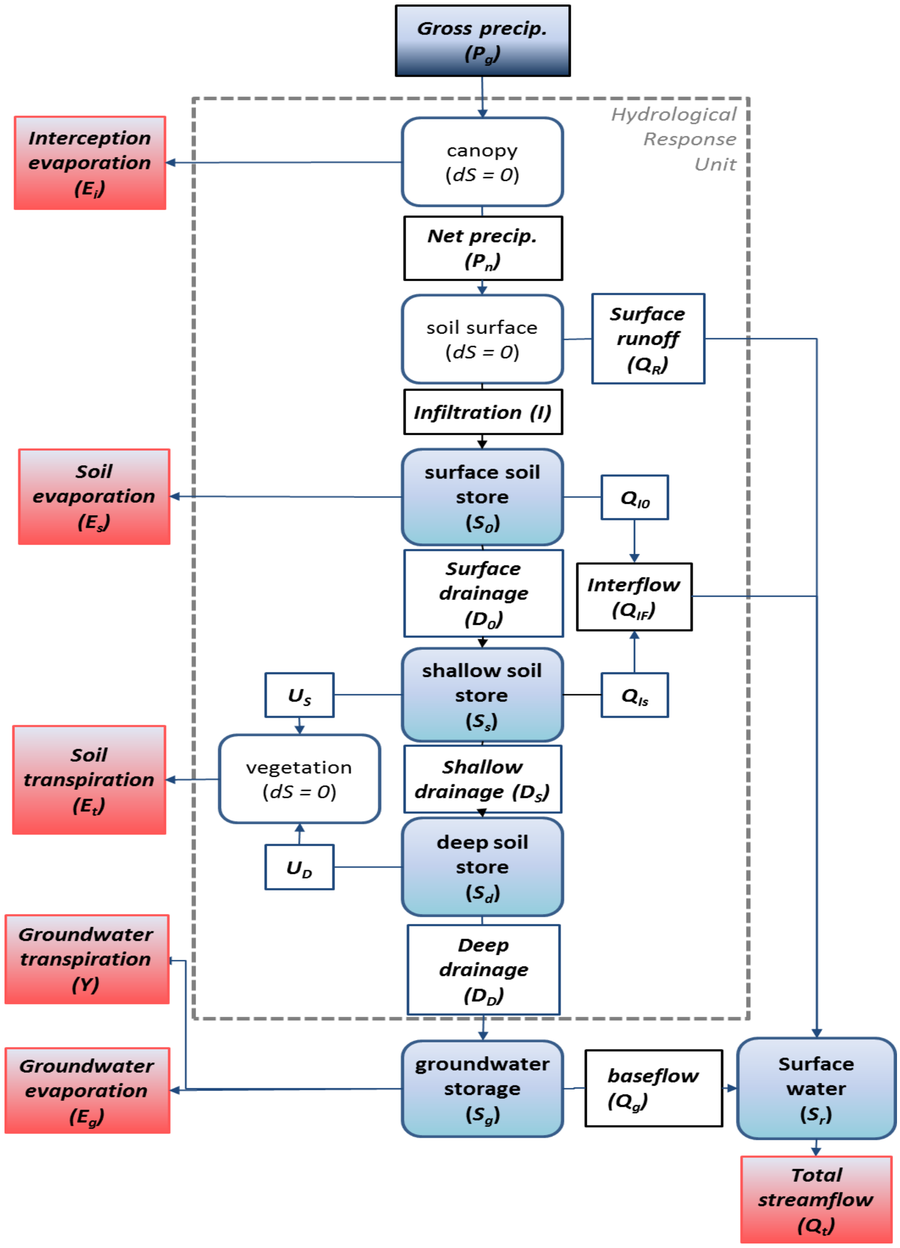


Figure 8. AWRA-L model conceptual diagram showing different hydrological processes

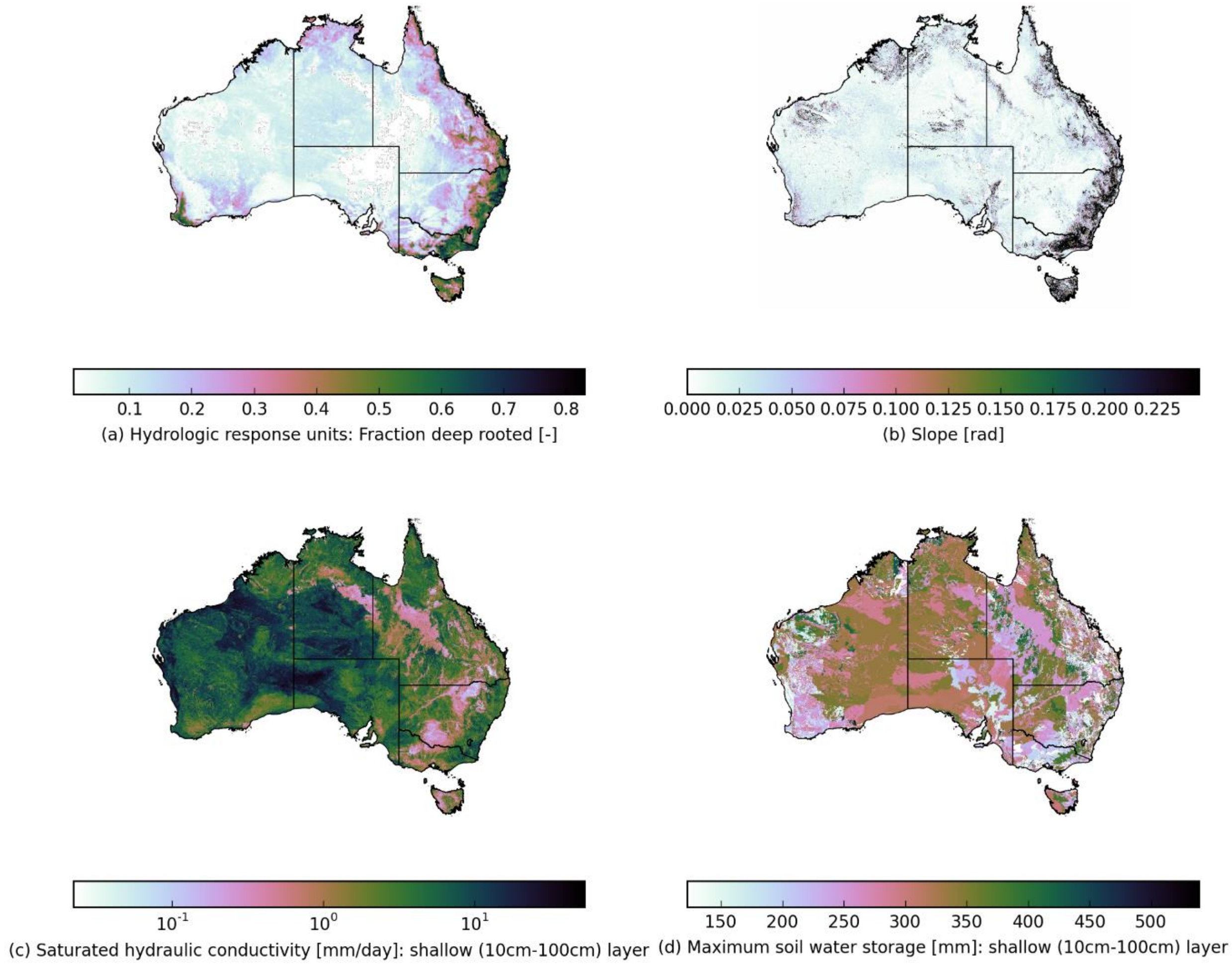


Figure 9. Example static spatial properties used in AWRA-L v5: (a) fraction of the cell a deep rooted hydrologic response unit, (b) average slope within a grid cell, and shallow (10-100cm) soil (c) saturated hydraulic conductivity and (d) maximum available water

The AWRA-L model has been calibrated to streamflow and catchment average soil moisture and ET across Australia. AWRA-L model parameters are currently calibrated nationally over 295 unimpaired catchments as identified within Zhang et al. (2013) – see Figure 2. Remaining 291 separate catchments are kept independently for scientific validation purposes. Three different datasets are used in calibration over these catchments across Australia including:

- **Catchment streamflow**: covering the period of 1981-2011
- **Catchment evapotranspiration:** CMRSET satellite ET - Satellite retrieval based grid estimates of evapotranspiration covering 2001-2010.
- **Catchment soil moisture:** AMSR-E product (Owe et al., 2008) - Satellite retrieval based grid estimates of soil moisture, covering the period of 2002-2011 have been used.

AWRA-L parameters (i.e. 21 parameters chosen to be free, rather than fixed) are optimised across the continent to maximise a composite function combing the performance according to streamflow, ET and soil moisture at all sites across Australia. The following streamflow objective function is evaluated for each catchment simulation (as derived by Viney et al., 2009 with the addition of a monthly NSE term):

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

where NSE_d and NSE_m are daily and monthly Nash-Sutcliffe Efficiency (Eq. 5) and B is relative bias (B) (Eq. 4 - see section 4.1). Daily soil moisture correlation (R_{SM}) and monthly evapotranspiration (R_{ET}) (defined in Eq. 6) are also used for each catchment according to the weighted function:

$$F = 0.7 * F_s + 0.15 * R_{SM} + 0.15 * R_{ET} \qquad (2)$$

Finally, the national calibration of AWRA-L maximises the grand objective function:

$$grandF = mean(F25\%, F50\%, F75\%, F100\%) \qquad (3)$$

where FX% being the Xth ranked site percentile F value. 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%).

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 run at a daily timestep, with monthly and weekly outputs published.

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 according to 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, according to run 26j (www.csiro.au/awap/doc/AWAP_readme_v9.txt).

3.3 CABLE

The CSIRO Atmosphere Biosphere Land Exchange (CABLE) model, is a community global land-surface model developed by CSIRO, BoM and other universities (Kowalczyk et al., 2006; Wang et al., 2011). 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 (with draining in the soil layers 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 (v5.0)***
Reference	Raupach et al (2009)	Wang et al (2011); Haverd et al (2013)	Viney et al (2015);Hafeez et al (2015)
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 (spatially varying properties)	2 (depth, saturated volumetric water content)	10 (saturated hydraulic conductivity, field capacity, etc)	3 (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 (v5.0)

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 different way to AWRA-L, in that they are calibrated for individual catchments, rather than finding a single parameter set to cover the entire Australia. 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 in practical approach to regionalisation, 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 calibration process only uses streamflow (rather than also using satellite derived soil moisture and evapotranspiration as now used in AWRA-L). For further details of the methods applied for the conceptual rainfall runoff modelling approach used here see Ramchurn and Frost (2014).

4 Evaluation approach

4.1 Statistics used in evaluation

The bias and monthly NSE statistics in particular are seen as good metrics for judging the models performance for AWRA-L purposes. 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 and actual ET).

Various statistics are calculated for each catchment/point to assess the models including:

Relative bias (B)

$$B_i = \sum_{t=1}^{T_i} \frac{Q_o^{ti} - Q_m^{ti}}{\bar{Q}_o^i} \tag{4}$$

Monthly and daily Nash-Sutcliffe Efficiency (NSE) and

$$NSE_i = 1 - \frac{\sum_{t=1}^{T_i} \left(Q_o^{ti} - Q_m^{ti}\right)^2}{\sum_{t=1}^{T_i} \left(Q_o^{ti} - \bar{Q}_o^i\right)^2} \tag{5}$$

Pearson's correlation coefficient (r)

$$r = \frac{\sum_{t=1}^{T_i} (Q_o^{ti} - \overline{Q_o^i})(Q_m^{ti} - \overline{Q_m^i})}{\sqrt{\sum_{t=1}^{T_i} (Q_o^{ti} - \overline{Q_o^i})^2} \sqrt{\sum_{t=1}^{T_i} (Q_m^{ti} - \overline{Q_m^i})^2}} \tag{6}$$

where Q_o^{ti} and Q_m^{ti} are the observed and modelled values for site i and time step t, for a total of T_i available observations, and $\overline{Q_o^i}$ is the mean observed and $\overline{Q_m^i}$ the mean modelled data for site *i*.

4.2 Evaluation criteria

The AWRA-L model was primarily developed for water resources application across Australia. Therefore, the evaluation criteria are primarily based on the available observed hydrological data across Australia. In general, improvements in model performance should be judged on data reserved for validation (i.e. separate to calibration data) – so that performance is more assured for predictions is ungauged basins – following the principles outlined in Refsgaard and Henriksen (2004). It is to be noted that all observed datasets have uncertainty associated with them, and are essentially a model also. Future improvement of the AWRA-L (and other) models can be judged according to the performance of AWRA-L v5 according to these metrics.

Primary metric – Assessment of AWRA-L against observed streamflow

- NSE for daily (NSE_d) and monthly (NSE_m) runoff
- Relative Bias (B) in long-term averages

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.
- Actual ET: Daily and monthly correlation of flux tower ET with AWRA-L ET.
- Deep drainage: Correlation between long-term reliable point measurements of recharge with AWRA-L deep drainage.

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 given what we consider would be a good performing model compared to similar peer models used for these purposes. For example: the majority of catchments to perform better than the average/climatology for streamflow – therefore want less than 5% at zero NSE (equivalent to climatology) – and 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 with simulations from alternative models	Aspirational target
Streamflow	Gauged streamflow (calibration and validation sites)	Daily NSE Monthly NSE Bias	CABLE WaterDyn Rainfall-runoff model (Sacramento and/or GR4J) – local calibration/ nearest neighbour regionalisation	**Daily NSE**: Less than 5% catchments with NSE<0 greater than 50% catchments with NSE>0.5 **Bias:** (a) 25%-75% catchments (50%) with -30%<bias<30%, (b) 5%-95% catchments (90%) with -50%<Bias<100%, and (c) No systematic spatial pattern of under- or over-estimation (i.e. low Bias when aggregated)
Soil moisture	Profile soil moisture from dedicated field observations	Daily and monthly correlation	CABLE WaterDyn	50% with daily correlation > 0.75 50% with monthly correlation >0.75
Actual ET	Flux ET	Monthly correlation	CABLE WaterDyn	Monthly correlation – 95% sites/cells with R>0.5, >50% sites/cells with R>0.8
Deep drainage	National Long term average dataset Annual time series dataset	Bias Annual correlation	CABLE WaterDyn	25% bias value below zero 75% bias value above zero 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, ET and recharge. Calibration/validation catchment statistics are presented using boxplots, showing the cumulative distribution of the statistics across all sites, with the box indicate the 25%, median and 75% (e.g. 25% for the 295 calibration sites means that 74 sites have lower values). Tables of statistics are presented in the case where there are insufficient sites for representation as a cumulative distribution (e.g. flux tower ET), or where alternative presentation of the statistics was meaningful (e.g. recharge).

5.1 Streamflow

AWRA-L model performance has been assessed for national models (WaterDyn and CABLE) as well as typical rainfall-runoff catchment scale models (GR4J and Sacramento) across Australia.

For the national landscape/landsurface models, the results show that AWRA-L model performs better than WaterDyn and CABLE according to monthly NSE and bias (Figure 10, Figure 11) 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. In particular, bias is near zero for the locally calibrated models (Figure 10(a)) 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 (Figure 11(b)) AWRA-L bias has less spread about zero, providing confidence in the spatial predictive qualities of AWRA-L. Significantly, AWRA-L v5 monthly performance on the validation catchments (Figure 10(b) and Figure 11(b)) 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 0.1 worse for daily NSE (Figure 12) than the locally calibrated models.

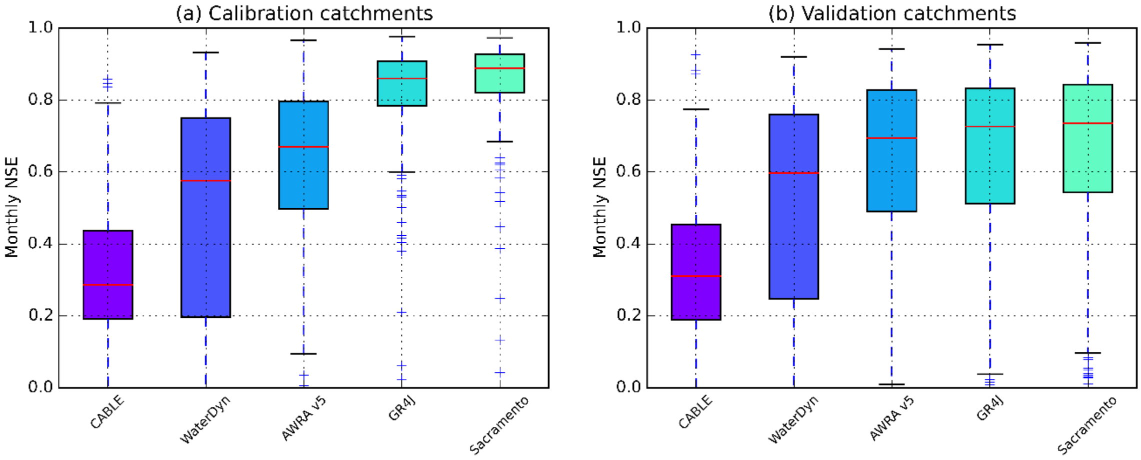


Figure 10. Monthly streamflow NSE for (a) calibration and (b) validation catchments

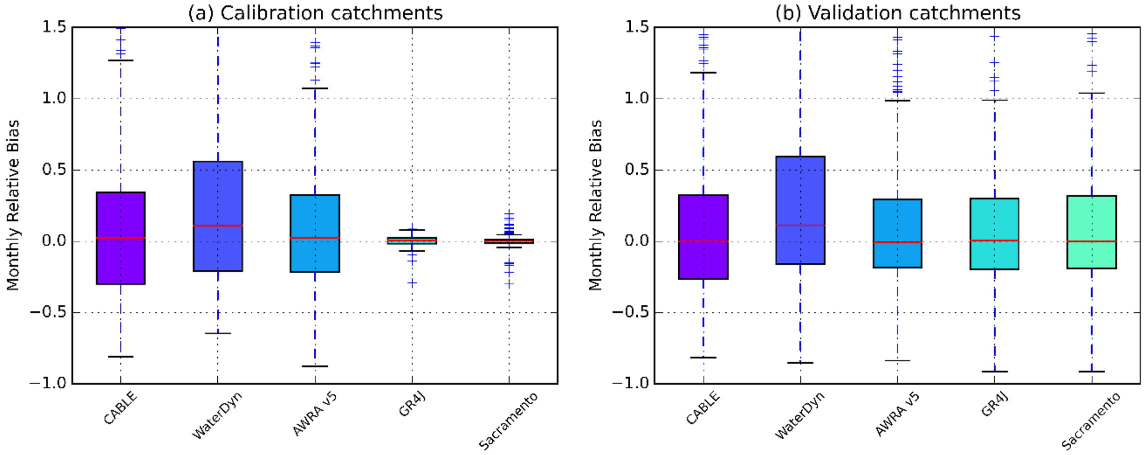


Figure 11. Monthly streamflow bias for (a) calibration and (b) validation catchments

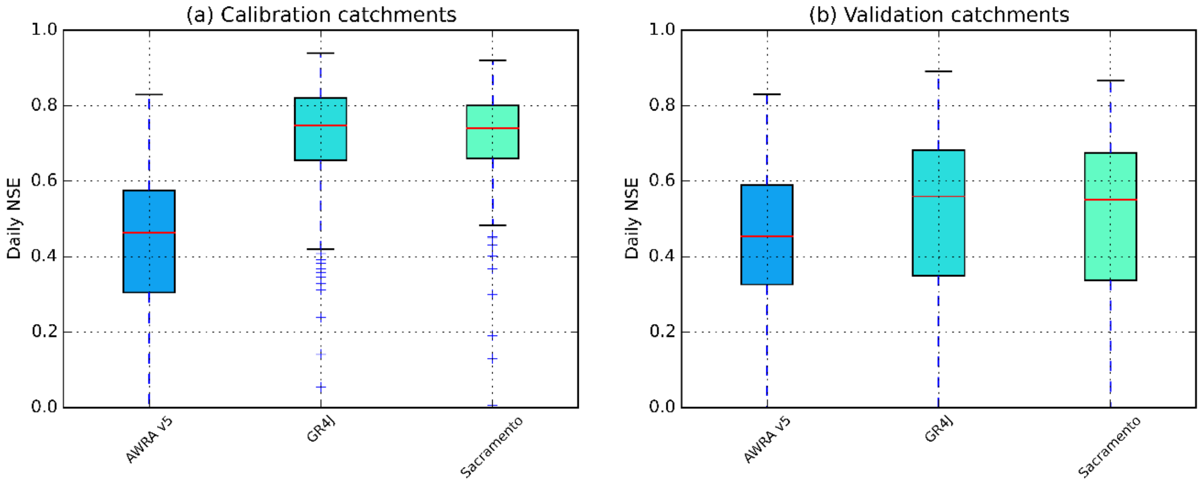


Figure 12. Daily streamflow NSE for (a) calibration and (b) validation catchments. Noting daily outputs not available for WaterDyn/CABLE.

Selected catchment monthly time series for AWRA-L (and WaterDyn and CABLE) are compared to those from the lumped rainfall models obtained when using a nearest neighbour method (which uses the parameters obtained from the closest locally calibrated catchment) in Figure 13. Catchment details and AWRA-L water balance time series are provided in Appendix D. This comparison indicates that AWRA-L provides more reliable estimates than a simple method of estimating flow in ungauged catchments using locally calibrated models in these locations. Furthermore, AWRA-L provides a range of water balance outputs (ET, soil moisture and deep drainage) – where the lumped conceptual model does not. *Overall, the result that AWRA-L performs well in validation provides confidence in its use for spatial prediction – across the country – for water resource assessment and scenario analysis purposes.*

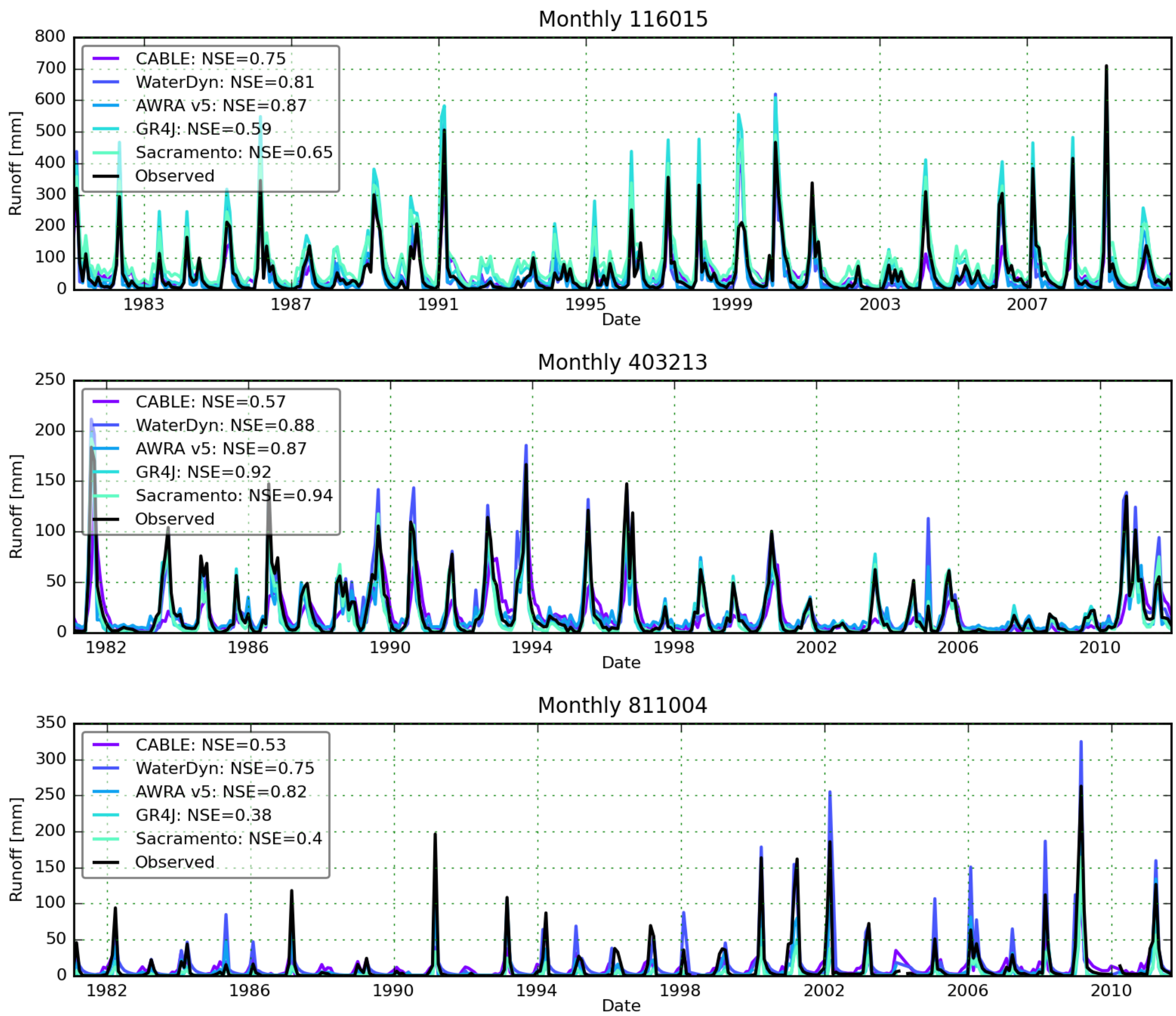


Figure 13. Select monthly catchment runoff time series comparing modelled and observed flow.

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 are bolded in the tables for model benchmarking purposes and comparison to the aspirational targets. Figure 14 shows a spatial plot of the AWRA-L (a) daily NSE and (b) monthly relative bias.

Table 3. Daily NSE percentiles for each model

Calibration	0%	5%	25%	50%	75%	95%	100%	Validation	0%	5%	25%	50%	75%	95%	100%
CABLE*								**CABLE***							
WaterDyn*								**WaterDyn***							
AWRA-L	-15.42	**-1.16**	0.30	**0.46**	0.58	0.71	0.83	**AWRA-L**	-41.79	**-0.30**	0.33	**0.45**	0.59	0.73	0.83
GR4J	0.00	0.47	0.66	0.75	0.82	0.88	0.94	**GR4J**	-11.12	-0.76	0.35	0.56	0.68	0.79	0.89
Sacramento	-1.95	0.50	0.66	0.74	0.80	0.86	0.92	**Sacramento**	-8184.58	-1.73	0.34	0.55	0.67	0.80	0.87
Benchmark		**-1.16**		**0.46**				**Benchmark**		**-0.30**		**0.45**			

* Daily results 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	-286.86	-0.72	0.19	0.29	0.44	0.71	0.86	**CABLE**	-23.80	-0.38	0.19	0.31	0.45	0.73	0.93
WaterDyn	-515.30	-3.25	0.20	0.58	0.75	0.86	0.93	**WaterDyn**	-43.96	-1.76	0.25	0.60	0.76	0.88	0.92
AWRA-L	-22.55	-0.51	0.50	0.67	0.80	0.89	0.97	**AWRA-L**	-43.77	-0.23	0.49	0.69	0.83	0.91	0.94
GR4J	-0.02	0.60	0.78	0.86	0.91	0.94	0.98	**GR4J**	-16.81	-0.24	0.51	0.73	0.83	0.91	0.95
Sacramento	-4.95	0.67	0.82	0.89	0.93	0.96	0.97	**Sacramento**	-1943.31	-0.25	0.54	0.74	0.84	0.93	0.96

Table 5. Relative bias percentiles for each model

Calibration	0%	5%	25%	50%	75%	95%	100%	Validation	0%	5%	25%	50%	75%	95%	100%
CABLE	-0.81	-0.54	-0.30	0.03	0.34	1.36	130.68	**CABLE**	-0.82	-0.56	-0.27	0.00	0.32	1.44	10.84
WaterDyn	-0.65	-0.48	-0.21	0.11	0.56	1.83	113.41	**WaterDyn**	-0.85	-0.46	-0.16	0.11	0.60	2.38	14.98
AWRA-L	-0.87	**-0.49**	**-0.21**	0.02	**0.32**	**1.36**	21.24	**AWRA-L**	-0.84	**-0.46**	**-0.18**	-0.01	**0.29**	**1.28**	8.69
GR4J	-0.29	-0.04	-0.02	0.01	0.02	0.05	0.09	**GR4J**	-0.91	-0.60	-0.20	0.00	0.30	1.34	7.41
Sacramento	-0.30	-0.02	-0.01	0.00	0.01	0.05	1.78	**Sacramento**	-0.92	-0.54	-0.19	0.00	0.32	1.41	7.36
Benchmark		**-0.49**	**-0.21**		**0.32**	**1.36**				**-0.46**	**-0.18**		**0.29**	**1.28**	

AWRA-L currently does not meet the aspirational daily NSE criteria (50% NSE of 0.45 rather than 0.5, 5% NSE at -0.3 rather than above zero); although it is not far off. In terms of bias AWRA-L does meet the criteria in validation for 50% of sites (25% to 75%) to be within -0.3 and 0.3. It does not 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 128% in validation. These results serve as a baseline benchmark for future improvements.

In terms of spatial performance, AWRA-L performs well (above 0.5 daily NSE) in Coastal NSW and Victoria, the majority of Queensland, the majority of Tasmania, South Wester 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 predominantly due to positive bias in these areas (Figure 13(b)). 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.

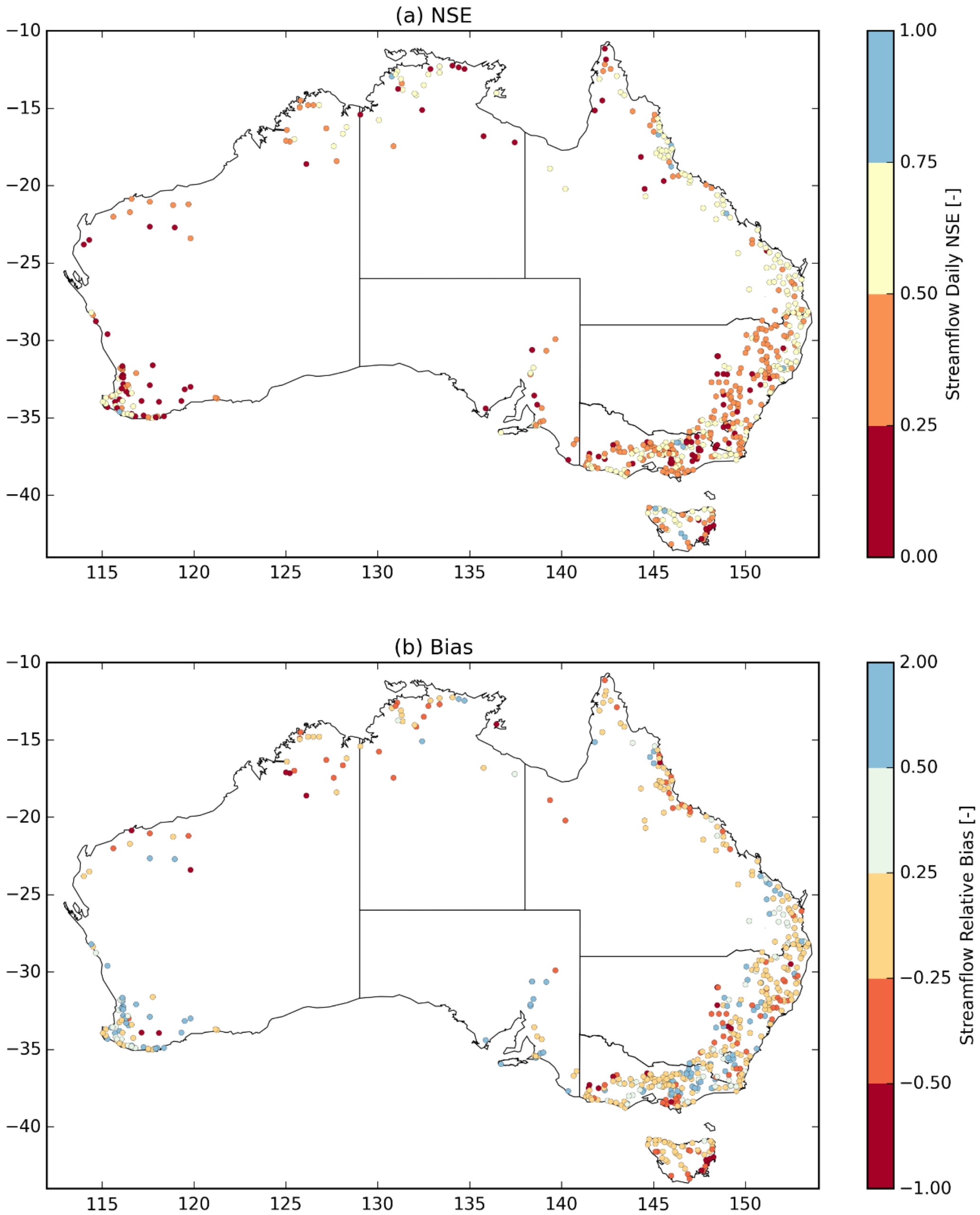


Figure 14. Map of AWRA-L runoff (a) daily NSE and (b) monthly relative bias compared to streamflow. Calibration and validation sites shown.

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) according to monthly correlation for the OzNet Murrumbidgee data and Upper Hunter SASMAS data (Figure 15).

This comparison uses the entire record that is available covering the model simulations (i.e. up until 2013). AWRA-L and CABLE perform similarly for profile soil moisture, with WaterDyn worse for the SASMAS Upper Hunter evaluation (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).

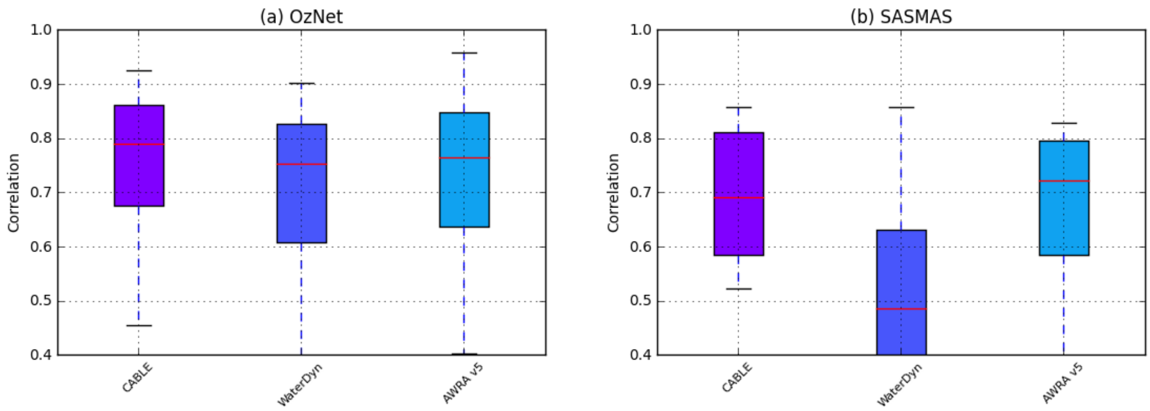


Figure 15. Correlation of models against monthly profile (0-90cm) soil moisture for (a) Murrumbidgee OzNet data 2001-2013 (b) Upper Hunter SASMAS 2003-2011 data

Table 6 presents the daily and monthly profile (0-90cm) correlation statistics, for evaluation against the Evaluation criteria listed in Table 2. AWRA-L is close to the 0.75 monthly and daily correlation performance for the SASMAS and OzNet datasets – exceeding it for the monthly OzNet dataset.

Table 6. 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	0.45	0.51	0.64	**0.74**	0.83	0.93	0.95	**AWRA-L**	0.27	0.34	0.60	**0.73**	0.79	0.84	0.84
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	0.51	0.61	0.68	0.77	0.78	**AMSRE**	0.08	0.13	0.37	0.39	0.49	0.58	0.65
Benchmark				**0.74**				**Benchmark**				**0.73**			
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	0.75	0.83	0.85	0.90	**WaterDyn**	0.17	0.25	0.36	0.49	0.63	0.75	0.86
AWRA-L	0.40	0.52	0.64	**0.76**	0.85	0.93	0.96	**AWRA-L**	0.14	0.23	0.58	**0.72**	0.80	0.82	0.83
Benchmark				**0.76**				**Benchmark**				**0.72**			

Following the analysis presented in Frost et al. (2015), the profile layer monthly correlation values are plotted for the OzNet and SASMAS (Figure 16) sites to give an indication of how AWRA-L performs spatially. Further, time series of the profile soil moisture for the models against the probe data is presented in Figure 17, to provide examples of how well AWRA-L produces drying and wetting of the soil as experienced in the Murrumbidgee during the Millennium drought (see Potter et al, 2010).

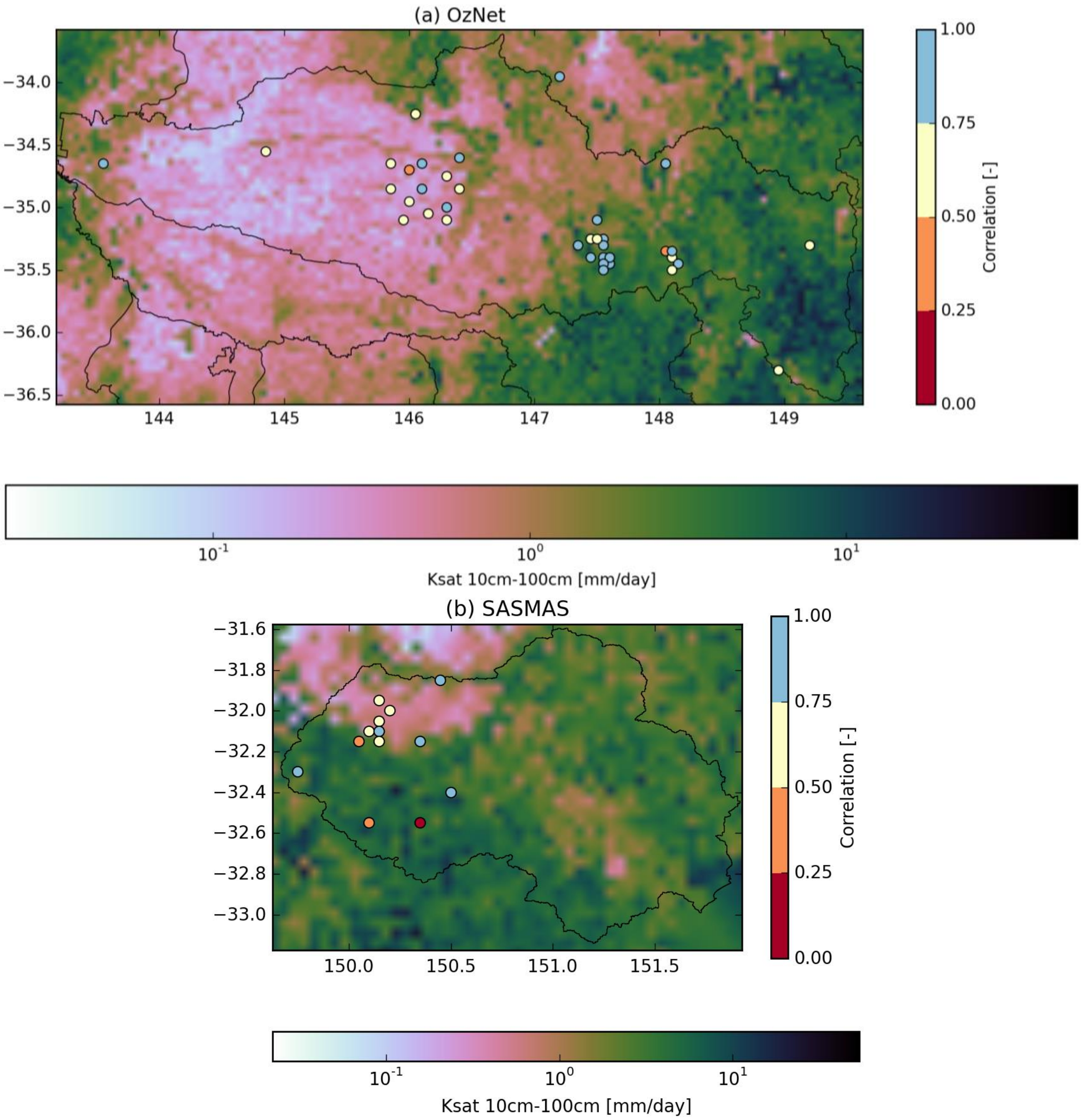


Figure 16. AWRA-L Monthly correlation for profile soil moisture of a) Murrumbidgee (OzNet) and b) SASMAS data. AWRA-L saturated conductivity (Ksat) for shallow layer (10cm-100cm) underlain.

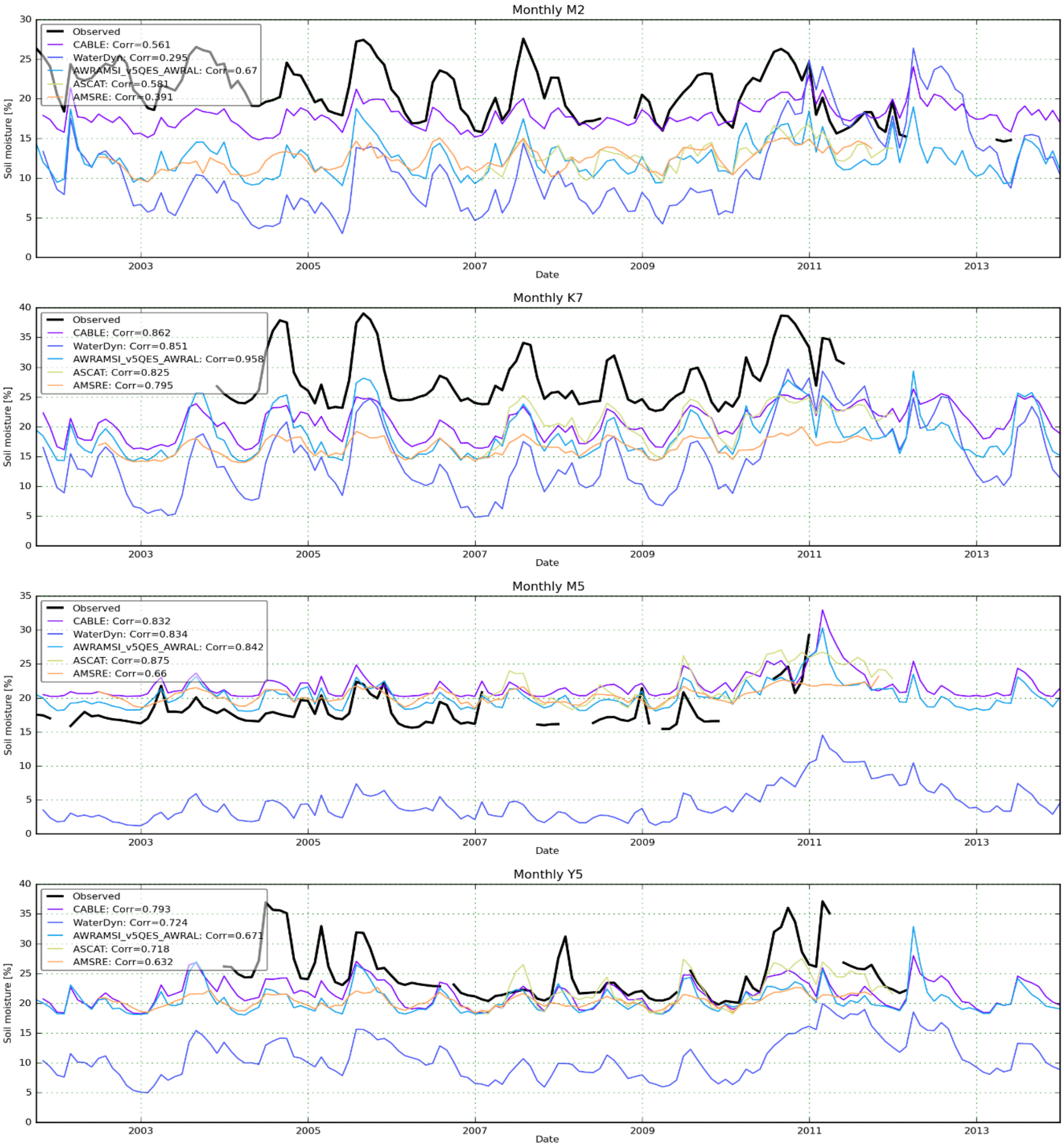


Figure 17. Four Murrumbidgee OzNet sites monthly profile (0-90cm) soil moisture and model/satellite estimates

A brief evaluation of the performance of the models at reproducing the top layer soil moisture is presented in Appendix C. Following the results presented in Frost et al. (2015), AWRA-L performs relatively worse compared to WaterDyn and CABLE when evaluated against the probe 0cm-5cm/8cm and satellite based data.

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

- The difference in point scale observations compared to large grid scale (~25 km by 25km for the models, larger for satellite data) outputs – with the point not reflecting the sampling area of the models evaluated.
- 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 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 and AWRA-L gridded outputs over the entire simulation period (2001-2013) available (Figure 18). CABLE and WaterDyn are roughly equal in terms of monthly correlation and better 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.

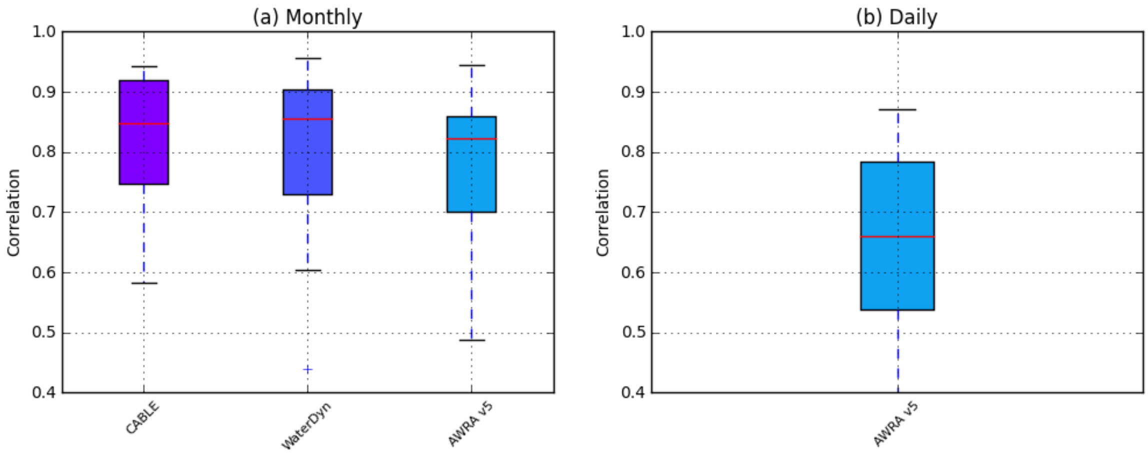


Figure 18. Correlation over 2001-2013 of DINGO derived actual ET compared to modelled (a) Monthly and (b) Daily data

Table 7 and Table 8 present the monthly and daily correlation and bias statistics, for evaluation against the criteria listed in Table 2. That is the monthly correlation for the 5% catchment is greater than 0.5 and 50% is greater than 0.9. AWRA-L is close to both of these – but below CABLE and WaterDyn particularly at the 5%. CABLE and WaterDyn provide benchmarks for future performance testing.

Table 7. 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 same period as models.

Correlation	0%	5%	25%	50%	75%	95%	100%	Relative bias	0%	5%	25%	50%	75%	95%	100%
CABLE	0.69	**0.72**	0.86	0.88	0.95	0.96	0.98	**CABLE**	-0.28	-0.18	-0.11	-0.05	0.27	0.51	0.59
WaterDyn	0.61	0.68	0.89	**0.90**	0.93	0.96	0.96	**WaterDyn**	-0.14	-0.14	-0.02	0.05	0.17	0.45	0.67
AWRAL	0.04	0.42	0.77	0.88	0.93	0.94	0.95	**AWRA-L**	-0.10	-0.08	-0.01	0.06	0.22	0.46	0.48
CMRSET	0.61	0.70	0.77	0.85	0.94	0.98	0.99	**CMRSET**	-0.41	-0.31	-0.13	-0.07	0.04	0.38	0.59
SLST	0.17	0.41	0.66	0.77	0.82	0.92	0.96	**SLST**	-0.29	-0.27	-0.19	-0.06	0.16	0.52	0.53
Benchmark		**0.42**		**0.88**											

Table 8. Daily correlation of AWRA-L compared to DINGO data 2001-2013.

Correlation	0%	5%	25%	50%	75%	95%	100%
AWRA-L	0.29	0.37	0.54	0.66	0.78	0.86	0.87

A second comparison (Figure 19) was undertaken using the time period that the satellite ET data was available (2001- 2010). 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 low deciles are below that of AWRA-L, the use of the data in calibration may be detracting ET performance in some cases. Further work is required to ensure that the actual ET dataset adds value to AWRA-L calibration.

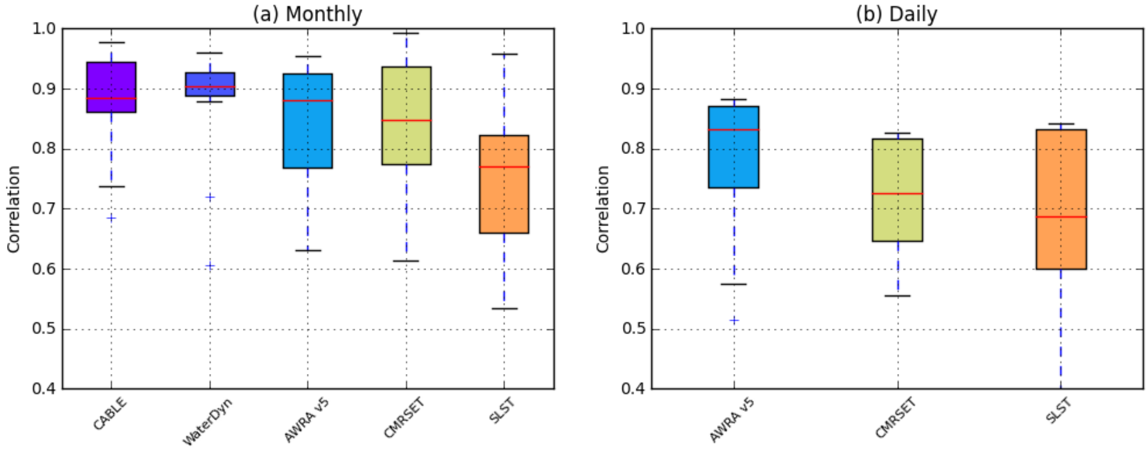


Figure 19. Correlation over 2001-2010 of DINGO derived actual ET compared to modelled (a) Monthly and (b) Daily data

Figure 20 shows the correlation of the three models against the CMRSET data. CABLE is most highly correlated to CMRSET, followed by AWRA-L and then WaterDyn. It appears that the calibration to CMRSET used in AWRA-L improves its performance relative to WaterDyn, according to CMRSET rather than flux tower data.

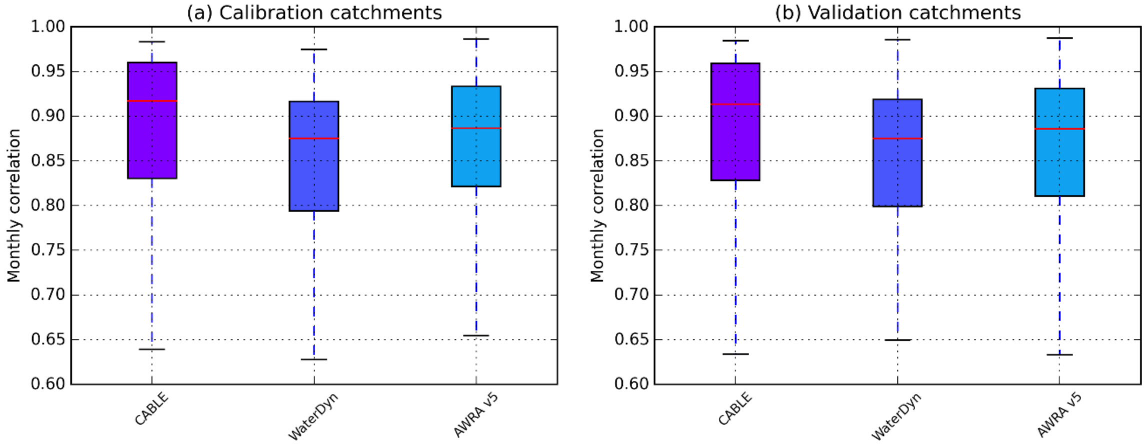


Figure 20. Monthly correlation of CMRSET against modelled ET over (a) calibration and (b) validation catchments

Figure 21 shows the spatial plots of the AWRA-L (a) correlation and (b) relative bias compared to the DINGO ET data.

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 indicate 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 29 in Appendix A: ET monitoring site details and time series) – and in future this data may be excluded from comparison.

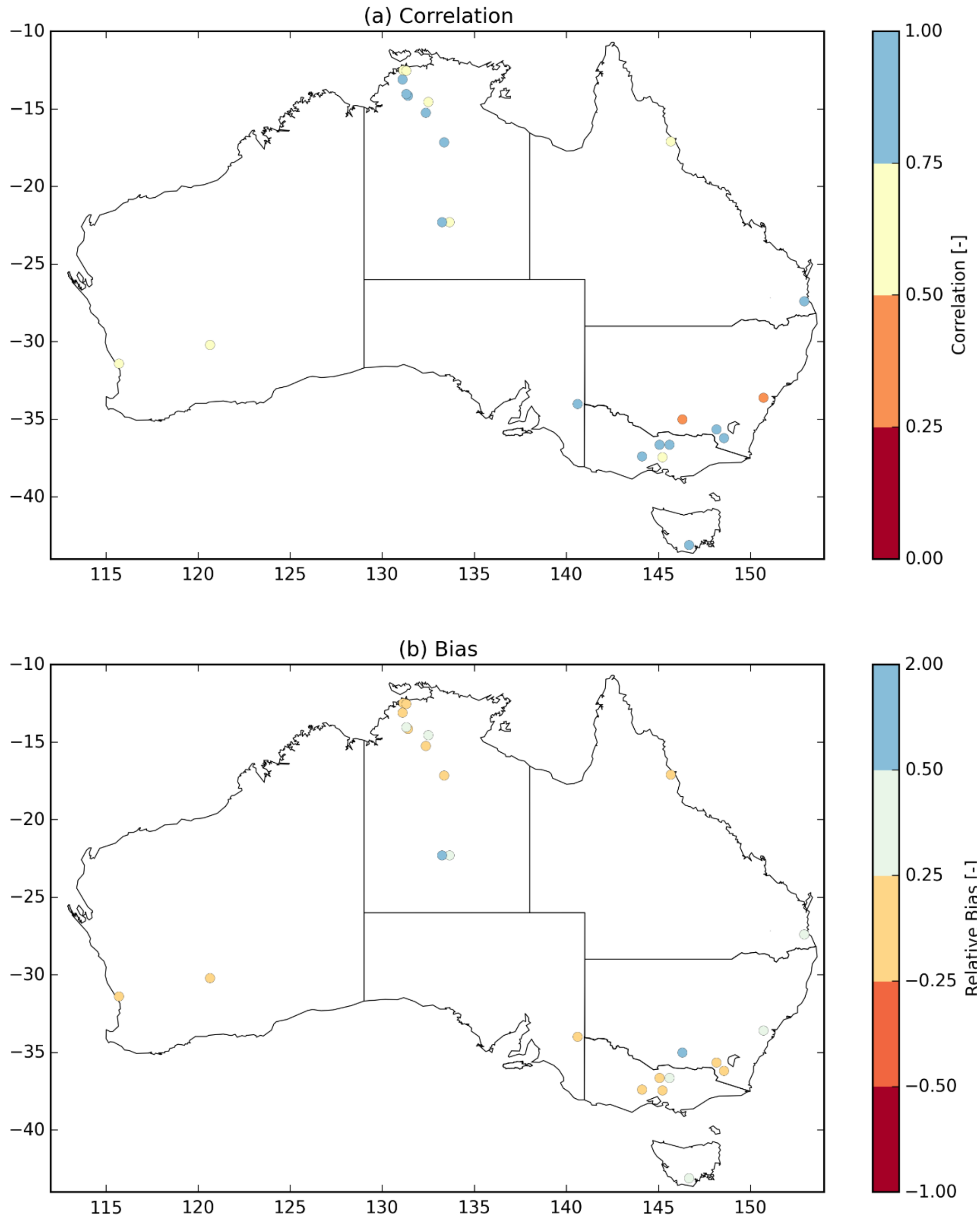


Figure 21. AWRA-L ET monthly (a) correlation and (b) bias compared with DINGO data

5.4 Groundwater deep drainage

Modelled deep drainage was compared against the Long Term Average national collated recharge dataset covering 2282 grid cells – with relative bias calculated (Figure 22). 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 23); with annual correlation and relative bias presented.

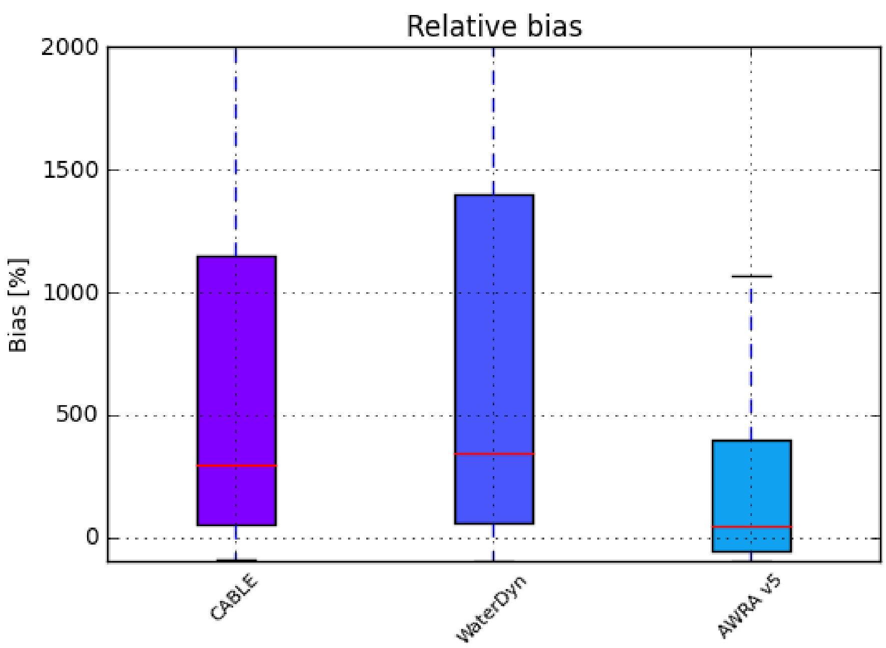


Figure 22. Modelled outputs versus Long Term Average recharge dataset (2282 grid cells across Australia) relative bias

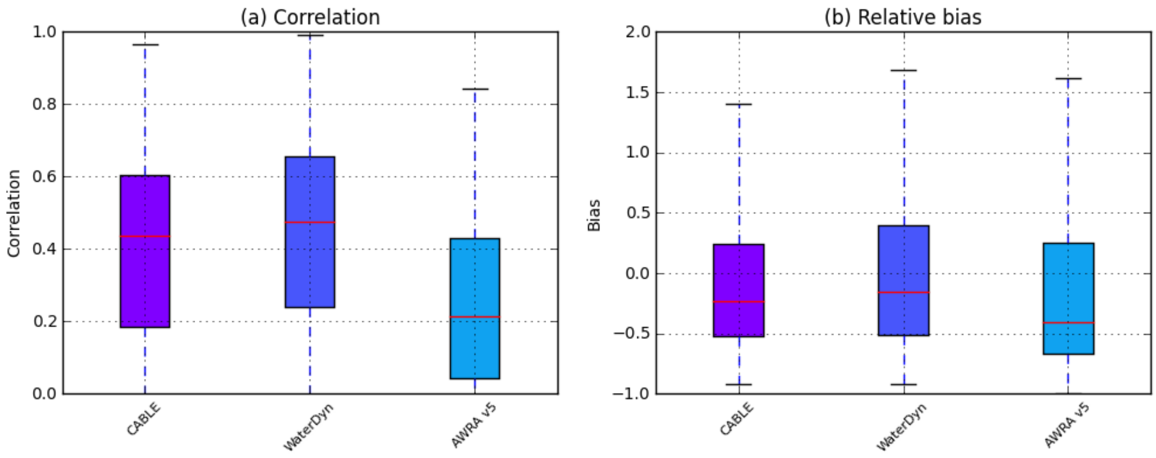


Figure 23. Modelled outputs versus annual recharge dataset (438 sites in South Australia) (a) correlation and (b) relative bias

AWRA-L performs well comparatively against the national long term average recharge dataset, with a median bias just above zero. WaterDyn and CABLE have median biases over 300%. However, for the annual time-series, AWRA-L is most biased (-40%), and has the lowest median correlation of the 3 models. This difference in results between the two datasets is attributed to variability in local performance in differing areas by AWRA-L.

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

- 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.
- 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 correlation.

Table 9. Deep drainage evaluation criteria

	0%	5%	25%	50%	75%	95%	100%
Model	**National Long Term Average Dataset relative bias**						
CABLE	-0.94	-0.50	0.48	2.93	11.50	85.40	1827.24
WaterDyn	-0.97	-0.50	0.60	3.40	13.98	111.69	1514.61
AWRA-L	-1.00	-0.92	**-0.57**	0.45	**3.95**	41.52	1198.05
Benchmark			**-0.57**		**3.95**		
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	-0.99	-0.32	0.04	**0.21**	0.43	0.68	0.84
Benchmark				**0.21**			

Example annual time series for two sites are presented for the 3 models in Figure 24. This plot gives an indication of the variability between models and data.

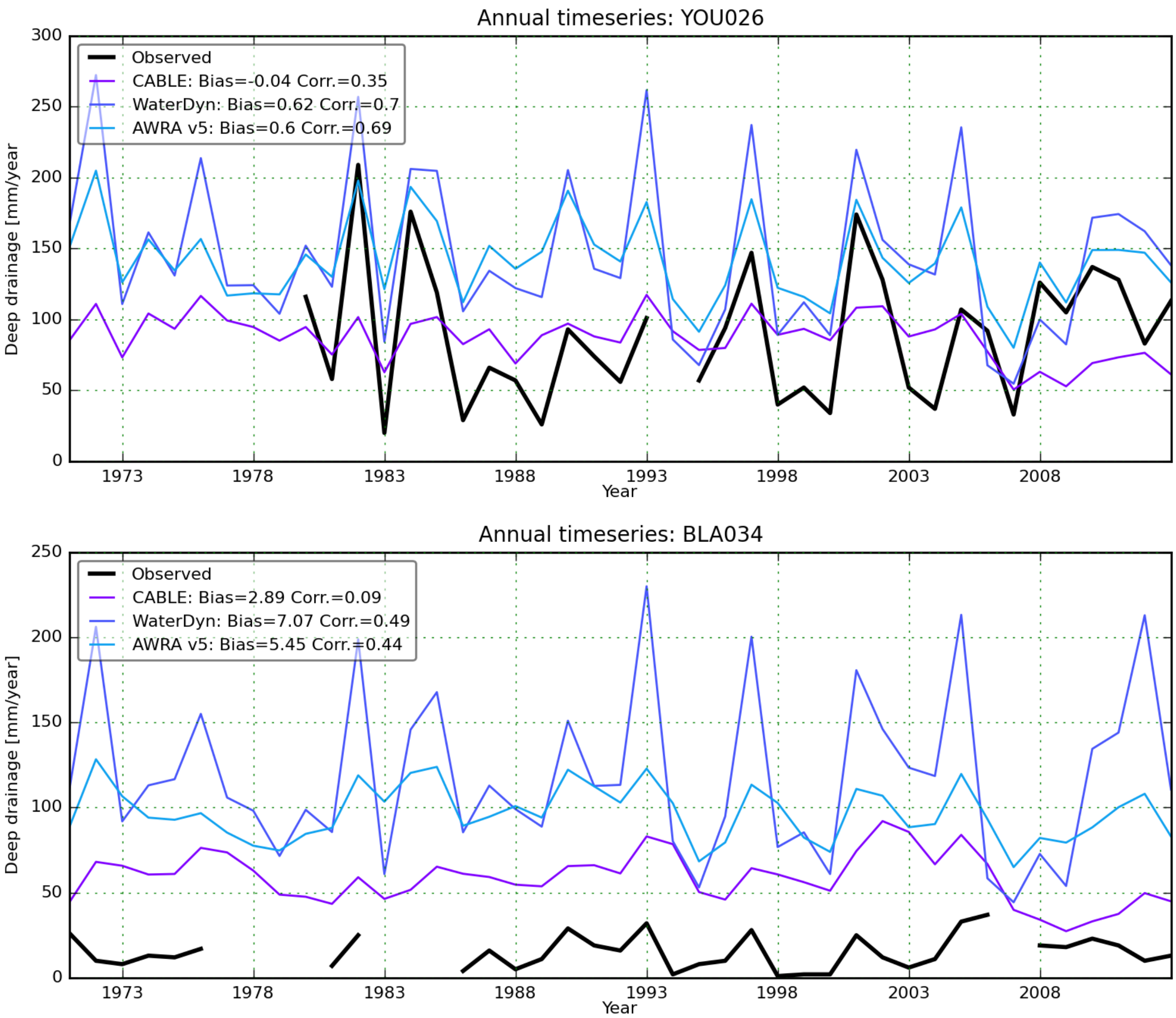


Figure 24. Example annual deep drainage time-series for two sites.

Figure 25 shows the relative bias value of the AWRA-L model compared to the Long Term Average data Australia wide; with the AWRA-L v5 shallow layer saturated soil conductivity underlain. Figure 26 shows the AWRA-L performance according to the annual data spatially – with the AWRA-L map of the maximum shallow layer soil storage and saturated conductivity is also underlain. 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.

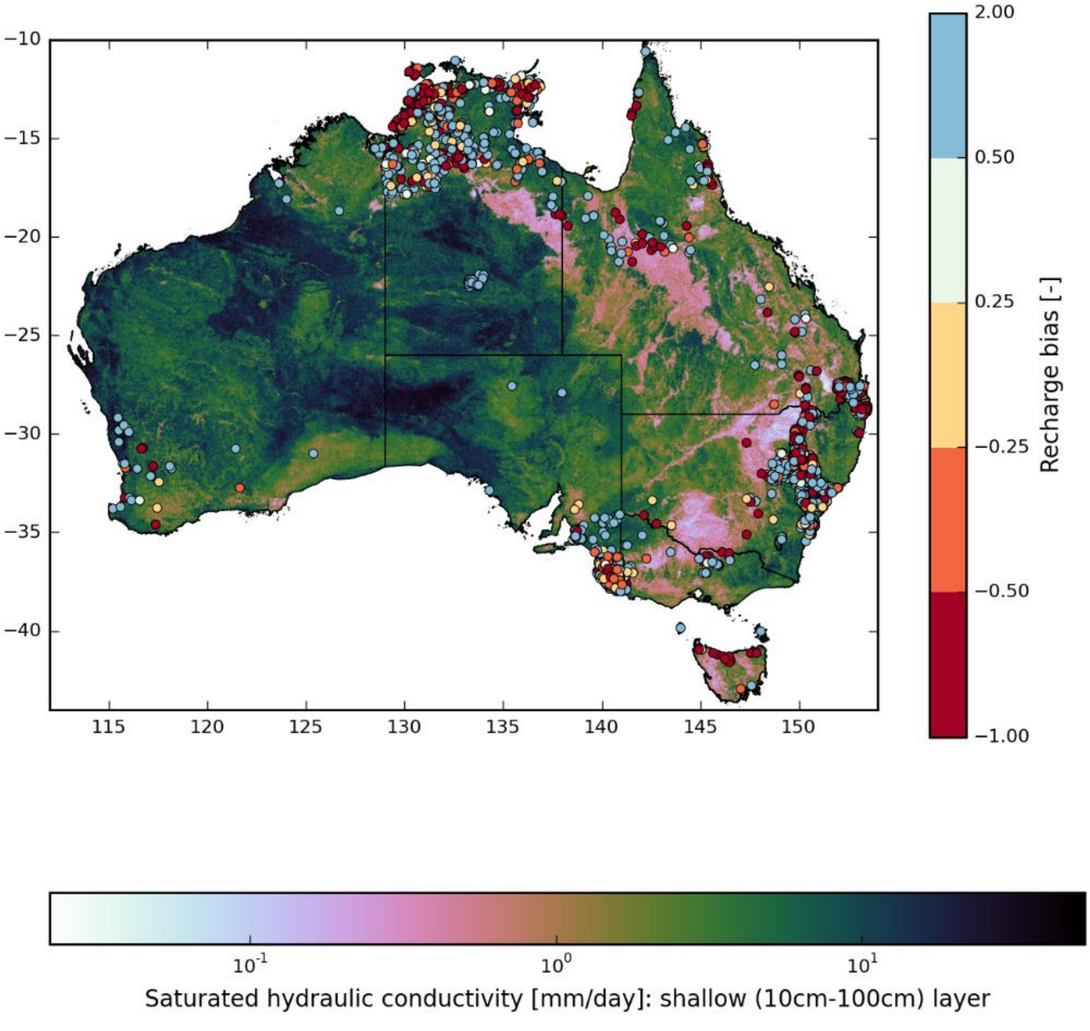


Figure 25. AWRA-L relative bias of deep drainage compared to Long Term Average estimates over Australia. AWRA-L v5 shallow layer soil saturated hydraulic conductivity (Ksat) is also mapped.

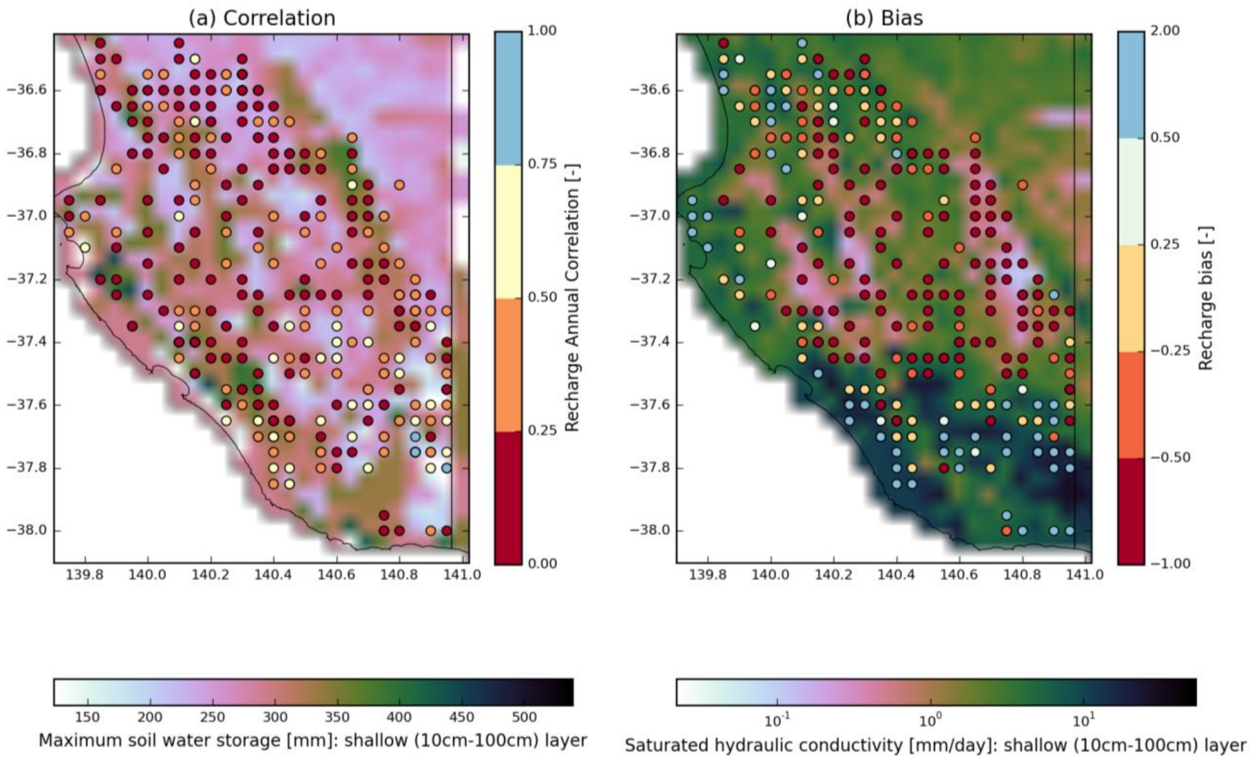


Figure 26. AWRA-L correlation and relative bias of deep drainage compared to annual time series estimates over South Eastern South Australia. AWRA-L v5 shallow layer maximum soil storage (Ssmax) and saturated hydraulic conductivity (Ksat) is also mapped.

These results appears 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.

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.

5.5 Summary according to benchmark statistics

Table 10 shows the summary performance according to the benchmark statistics.

Table 10. AWRA-L v5 Performance according to benchmark statistics

Streamflow	Percentile	Calibration catchments (295 Nationally)	Validation catchments (291 Nationally)	Aspirational target
Daily Nash-Sutcliffe Efficiency (NSE)	5%	-1.16	-0.30	Less than 5% catchments with NSE<0
	50%	0.46	0.45	Greater than 50% catchments with NSE>0.5
Relative Bias	25%	-21%	-18%	25%-75% catchments (50%) with -30%<bias<30%
	75%	32%	28%	
	5%	-49%	-46%	5%-95% catchments (90%) with -50%<Bias<100%
	95%	136%	128%	
Soil moisture	**Percentile**	**OzNet Murrumbidgee**	**SASMAS Hunter**	**Aspirational target**
0-90cm Daily correlation	50%	0.74	0.73	50% with daily correlation >0.75
0-90cm Monthly correlation	50%	0.76	0.72	50% with monthly correlation >0.75
Actual Evapotranspiration	**Percentile**	**National infilled flux tower data (DINGO)**		Aspirational target
Monthly correlation	5%	0.42		95% sites/cells with R>0.5
	50%	0.88		>50% sites/cells with R>0.8
Recharge	**Percentile**	**National Long Term average**	**Annual time-series in South Australia**	**Aspirational target**
Relative bias	25%	-57%	NA	25% bias value below zero
	75%	395%	NA	75% bias value above zero
Annual correlation	50%	NA	0.21	Median annual correlation above 0.5

Note: Red indicates where AWRA-L does not currently meet aspirational target.

6 Evaluation of AWRA-L v5 for reporting purposes

Gridded outputs and select catchment based time series from AWRA-L v5 are presented here to give an understanding of the AWRA-L model spatial and temporal dynamics.

Annual average maps for calendar years from 1911-2015 for ET, runoff, root zone soil moisture and Deep drainage are shown in Figure 27. These plots show 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 water balance components having no flows in this area.

Further annual total and decile map, and monthly average maps are provided in Appendix D: Maps and time series of water balance outputs from AWRA-L.

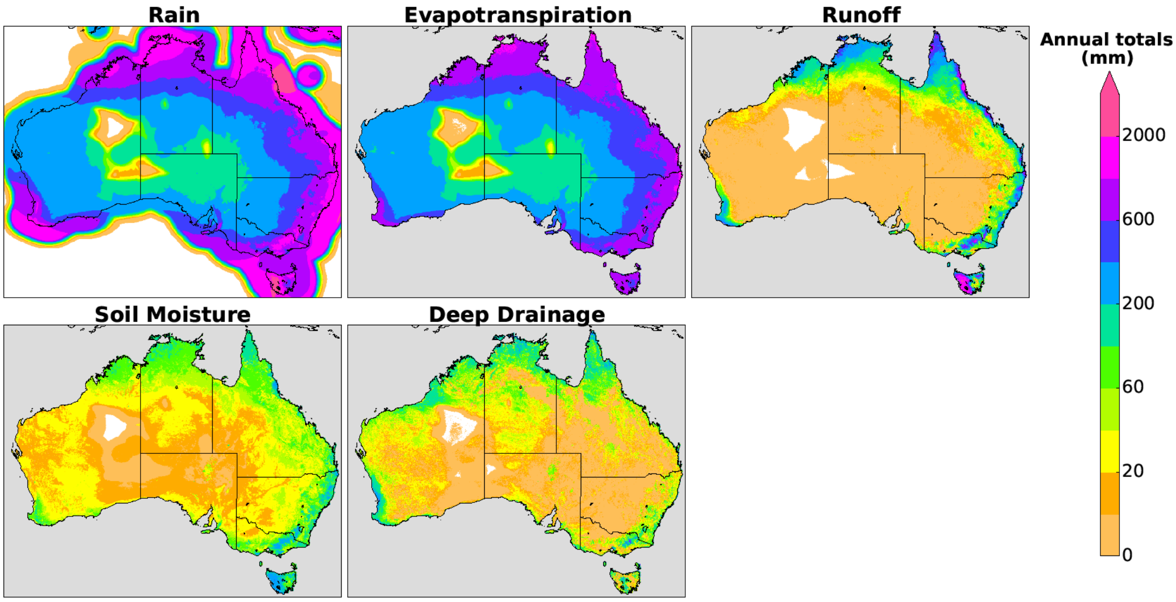


Figure 27. AWRA-L v5 mean annual rainfall, runoff, ET, soil moisture and deep drainage. Units=mm.

An example time series of monthly water balance outputs is presented to give an idea of the temporal dynamics of the AWRA-L in Figure 28. Variables plotted include rainfall, actual ET, runoff, soil moisture (top, shallow and deep layers), deep drainage and groundwater storage (Sg). The relevant water balance terms are compared to catchment streamflow, and satellite derived soil moisture and ET also. This site matches the observed data well. Further, the storage terms (in particular the deep soil and groundwater storage) shows drawdown over the Millennium drought period – as expected in this area. Further example sites spread across Australia are provided for user evaluation in Appendix D.

Overall from the spatial plots and the 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.

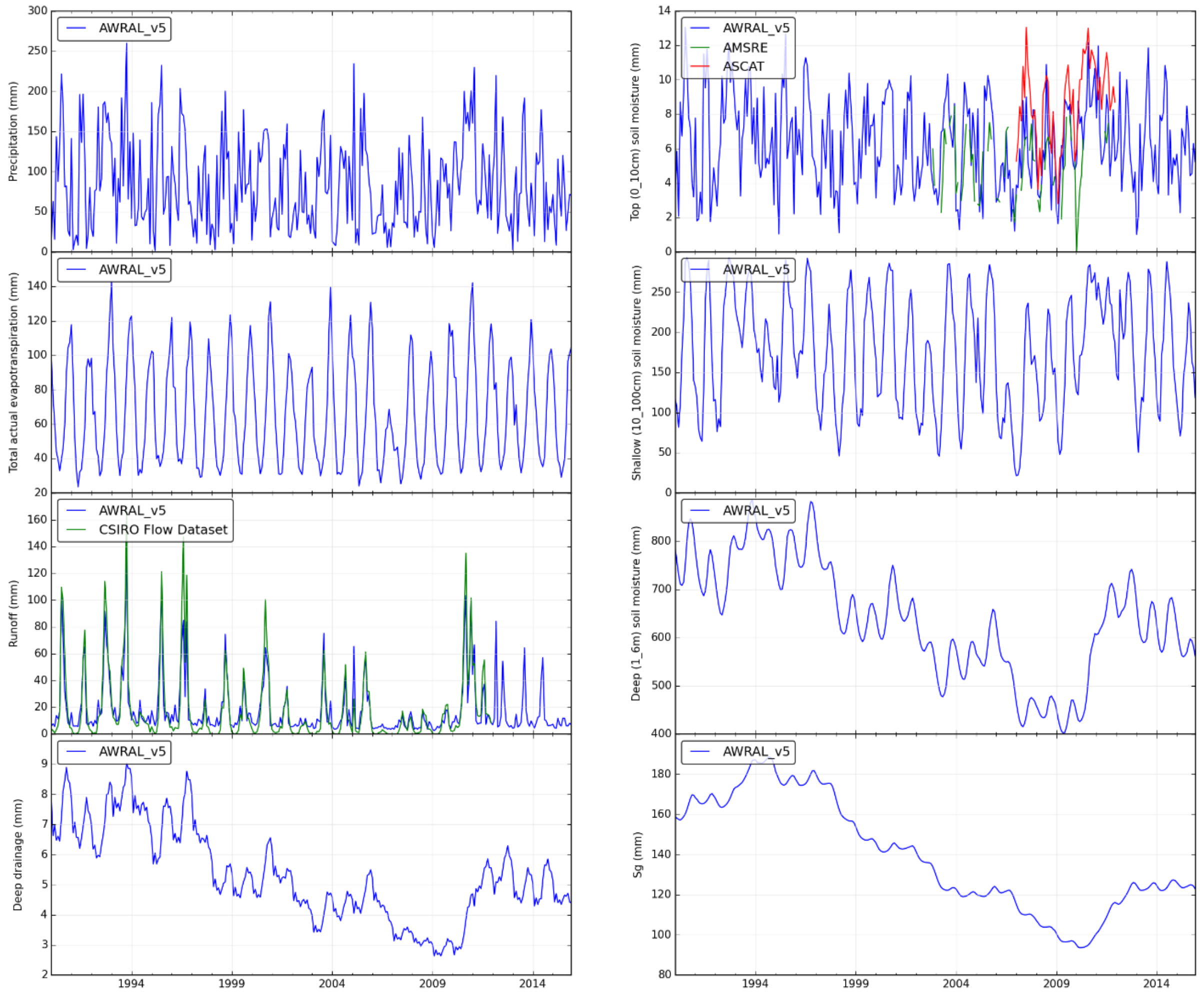


Figure 28. AWRA-L water balance outputs for FIFTEEN MILE CREEK @ GRETA SOUTH catchment (403213) in Victoria. Comparison time series of streamflow and satellite derived soil moisture and ET are also shown.

7 Conclusions

AWRA-L performance was evaluated using available Australian 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. Performance against key evaluation criteria was undertaken, and results presented in Table 10 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 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). Current AWRA-L performance according to daily Nash-Sutcliffe Efficiency at the 5% / 50% is -0.3 / 0.45 in validation, with aspirational performance set at 0.0 / 0.5 respectively.

AWRA-L and CABLE perform similarly for root-zone (profile 0-90mm) soil moisture, with WaterDyn worse. Current AWRA-L performance according to daily and monthly correlation at the 50% is 0.72-0.76 (for the Murrumbidgee and Hunter sites), with aspirational performance set at 0.75.

For actual ET, CABLE and WaterDyn are better than AWRA-L model. CABLE performs best according to ET, as expected from its purpose as a model for land/atmosphere exchange model, along with calibration to flux tower and streamflow derived ET (noting its poor performance according to runoff). Current AWRA-L performance according to monthly correlation to DINGO flux tower data at 5% / 50% is 0.42 / 0.88 with aspirational performance set at 0.5 / 0.8 respectively.

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 performance according to relative bias against the national long term average recharge dataset at 25% / 75% is -57% / 395% with aspirational performance set at being less than zero / greater than zero respectively. Secondly AWRA-L has a median annual correlation against the South Australian annual time-series dataset of 0.21, below the aspirational target of 0.5.

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 is considered most fit for purpose water balance estimation purposes.

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 current AWRA-L v5 operational model. Overall from the spatial plots and the 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.

It is noted that there are various improvements to AWRA-L underway including:

- Regional calibration of the model: towards better matching local variability
- Updating input spatial datasets based on newly released TERN soil properties
- Improving spatial resolution of AWRA-L model, from the current ~5km grid scale down to ~1km
- Increasing the number of Hydrological Response Units from two (shallow and deep rooted vegetation) to five to include:
    - impervious areas (e.g. Urban)
    - irrigated areas, and
    - permanently inundated areas/lakes.

Once these additions have been parameterised and shown to improve the model, they will be incorporated into the operational system. Finally, the AWRA-L model is being released as a community model to enable use and development of the system by a wide range of stakeholders.

References

Bartalis, Z., Wagner, W., Naeimi, V., Hasenauer, S., Scipal, K., Bonekamp, H., Figa, J. and Anderson, C., 2007. Initial soil moisture retrievals from the METOP-A Advanced Scatterometer (ASCAT). Geophysical Research Letters, 34(20): L20401.

Beringer, J., Hutley, L. B., McHugh, I., Arndt, S. K., Campbell, D., Cleugh, H. A., Cleverly, J., Resco 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., and Wardlaw, T, 2016. An introduction to the Australian and New Zealand flux tower network – OzFlux, Biogeosciences, 13, 5895-5916.

Beringer, J., McHugh, I., Hutley, L.B., Isaac, P. and Kljun, N., 2016. Dynamic INtegrated Gap-filling and partitioning for OzFlux (DINGO). Biogeosciences Discuss., 2016: 1-36.

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

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

Crosbie, R.S., Jolly, I.D., Leaney, F.W. and Petheram, C., 2010b. Can the dataset of field based recharge estimates in Australia be used to predict recharge in data-poor areas? Hydrol. Earth Syst. Sci., 14(10): 2023-2038.

Dane, J.H. and Puckett, W., 1994. Field soil hydraulic properties based on physical and mineralogical information. In: M.v.G.e. al. (Editor), International Workshop on Indirect Methods for Estimating the Hydraulic Properties of Unsaturated Soils. University of California, Riverside, CA, pp. 389–403.

Donohue, R.J., Roderick, M.L. and McVicar, T.R., 2008. Deriving consistent long-term vegetation information from AVHRR reflectance data using a cover-triangle-based framework. Remote Sensing of Environment, 112(6): 2938-2949.

Draper, C.S., Walker, J.P., Steinle, P.J., de Jeu, R.A.M. and Holmes, T.R.H., 2009. An evaluation of AMSR-E derived soil moisture over Australia. Remote Sensing of Environment, 113(4): 703-710.

Eamus, D., Cleverly, J., Boulain, N., Grant, N., Faux, R. and 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. Agricultural and Forest Meteorology, 182–183: 225-238.

Elmahdi, A., Hafeez, M., Smith, A., Frost, A., Vaze, J. and Dutta, D., 2015. Australian Water Resources Assessment Modelling System (AWRAMS)-informing water resources assessment and national water accounting, 36th Hydrology and Water Resources Symposium. Engineers Australia, Hobart, Australia.

Frost, A.J., Ramchurn, A. and Hafeez, M., 2015. Evaluation of AWRA-L for national drought and soil moisture monitoring, 36th Hydrology and Water Resources Symposium: The art and science of water. Barton, ACT: Engineers Australia, Hobart, pp. 1496-1504.

Grant, I., Jones, D., Wang, W., Fawcett, R., and Barratt, D., 2008. Meteorological and remotely sensed datasets for hydrological modelling: a contribution to the Australian Water Availability Project. 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., Van Dijk, A.I.J.M., Mattersdorf, G., Beringer, J., Hutley, L.B., Leuning, R., Pipunic, R.C. and Sherman, B.S., 2009. Scaling of potential evapotranspiration with MODIS data reproduces flux observations and catchment water balance observations across Australia. Journal of Hydrology, 369(1-2): 107-119.

Hafeez, M., Frost, A., Vaze, J., Dutta, D., Smith, A. and Elmahdi, A., 2015a. A New Integrated Continental Hydrological Simulation System: An overview of the Australian Water Resource Assessment Modelling System (AWRAMS). Water: Journal of the Australian Water Association, 42(3): 75-82.

Hafeez, M., Smith, A., Frost, A., Srikanthan, S., Elmahdi, A., Vaze, J. and Prosser, I., 2015b. The Bureau's Operational AWRA Modelling System in the context of Australian landscape and hydrological model products, 36th Hydrology and Water Resources Symposium. Engineers Australia, Hobart, Australia.

Haverd, V. and 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. Journal of Hydrology, 388(3–4): 438-455.

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. and Wang, Z., 2013. Multiple observation types reduce uncertainty in Australia's terrestrial carbon and water cycles. Biogeosciences, 10(3): 2011-2040.

Johnston, R.M., Barry, S.J., Bleys, E., Bui, E.N., Moran, C.J., Simon, D.A.P., Carlile, P., McKenzie, N.J., Henderson, B.L., Chapman, G., Imhoff, M., Maschmedt, D., Howe, D., Grose, C., Schoknecht, N., Powell, B. and Grundy, M., 2003. ASRIS: the database. Soil Research, 41(6): 1021-1036.

Kowalczyk, E.A., Wang, Y.P., Law, R.M., Davies, H.L., McGregor, J.L. and 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. and Tramutoli, V., 2012. A multi-sensor (SMOS, AMSR-E and ASCAT) satellite-based soil moisture products inter-comparison, 2012 IEEE International Geoscience and Remote Sensing Symposium, pp. 1135-1138.

Leroux, D., Kerr, Y.H., Albitar, A., Bindlish, R., Jackson, T.J., Berthelot, B. and Portet, G., 2013. Comparison between SMOS, VUA, ASCAT, and ECMWF soil moisture products over four watersheds in U.S. IEEE Transactions on Geoscience and Remote Sensing: 1-20.

Owe, M., de Jeu, R. and Holmes, T., 2008. Multisensor historical climatology of satellite-derived global land surface moisture. Journal of Geophysical Research: Earth Surface, 113(F1): F01002.

Peeters, L.J.M., Doble, R.C., Crosbie, R.S. and van Dijk, A.I.J.M., 2011. Incorporating topography-dependent groundwater storage in AWRA-L improves groundwater flux estimation, 19th International Congress on Modelling and Simulation, Perth, Australia, pp. 4064-4070.

Perrin, C., Michel, C. and AndrÃeassian, V., 2003. Improvement of a parsimonious model for streamflow simulation. Journal of Hydrology, 279(1-4): 275-289.

Potter, N. J., F. H. S. Chiew and A. J. Frost (2010). *An assessment of the severity of recent reductions in rainfall and runoff in the Murray-Darling Basin*, Journal of Hydrology 381(3-4): 52-64

Raupach, M.R., Briggs, P.R., Haverd, V., King, E.A., Paget, M. and 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.

Ramchurn, A. and Frost, A.J., 2013. Potential improvements to the Australian Water Resources Assessment system landscape (AWRA-L) model, 20th International Congress on Modelling and Simulation, Adelaide, Australia, pp. 3008-3014.

Refsgaard, J.C. and Henriksen, H.J., 2004. Modelling guidelines--terminology and guiding principles. Advances in Water Resources, 27(1): 71-82.

Renzullo, L.J., van Dijk, A.I.J.M., Perraud, J.M., Collins, D., Henderson, B., Jin, H., Smith, A.B. and McJannet, D.L., 2014. Continental satellite soil moisture data assimilation improves root-zone moisture analysis for water resources assessment. Journal of Hydrology, 519, Part D(0): 2747-2762.

Rüdiger, C., Hancock, G., Hemakumara, H.M., Jacobs, B., Kalma, J.D., Martinez, C., Thyer, M., Walker, J.P., Wells, T. and Willgoose, G.R., 2007. Goulburn River experimental catchment data set. Water Resources Research, 43(10): W10403.

Rüdiger, C., Western, A.W., Walker, J.P., Smith, A.B., Kalma, J.D. and Willgoose, G.R., 2010. Towards a general equation for frequency domain reflectometers. Journal of Hydrology, 383(3–4): 319-329.

Shi, X., R.S., C. and Vaze, J., 2015. Evaluation of AWRA recharge: Comparison to field estimates of recharge across Australia. , CSIRO.

Simard, M., Pinto, N., Fisher, J.B. and Baccini, A., 2011. Mapping forest canopy height globally with spaceborne lidar. Journal of Geophysical Research: Biogeosciences, 116(G4): G04021.

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. and Richter, H., 2012. The Murrumbidgee soil moisture monitoring network data set. Water Resources Research, 48(7): W07701.

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. and van Gorsel, E., 2012. Upscaling latent heat flux for thermal remote sensing studies: Comparison of alternative approaches and correction of bias. Journal of Hydrology, 468–469(0): 35-46.

Vaze, J., Viney, N., Stenson, M., Renzullo, L., Van Dijk, A., Dutta, D., Crosbie, R., Lerat, J., Penton, D., Vleeshouwer, J., Peeters, L., Teng, J., Kim, S., Hughes, J., Dawes, W., Zhang, Y., Leighton, B., Perraud, J., Joehnk, K., Yang, A., Wang, B., Frost, A., Elmahdi, A., Smith, A. and Daamen, C., 2013. The Australian Water Resource Assessment System (AWRA), Proceedings of the 20th International Congress on Modelling and Simulation, Adelaide, Australia.

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

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

Viney, N.R., Perraud, J., Vaze, J., Chiew, F.H.S., Post, D.A. and Yang, A., 2009. The usefulness of bias constraints in model calibration for regionalisation to ungauged catchments. In: R.S. Anderssen, R.D. Braddock and L.T.H. Newham (Editors), 18th

World IMACS / MODSIM Congress. Modelling and Simulation Society of Australia and New Zealand and International Association for Mathematics and Computers in Simulation, Cairns, Australia, pp. 3421-3427, http://www.mssanz.org.au/modsim09/I7/viney_I7a.pdf.

Wagner, W., Brocca, L., Naeimi, V., Reichle, R., Draper, C., Jeu, R.d., 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. and Paulik, C., 2014. Clarifications on the "Comparison Between SMOS, VUA, ASCAT, and ECMWF Soil Moisture Products Over Four Watersheds in U.S.". IEEE Transactions on Geoscience and Remote Sensing, 52(3): 1901-1906.

Wang, Y.P., Kowalczyk, E., Leuning, R., Abramowitz, G., Raupach, M.R., Pak, B., van Gorsel, E. and Luhar, A., 2011. Diagnosing errors in a land surface model (CABLE) in the time and frequency domains. Journal of Geophysical Research: Biogeosciences, 116(G1): G01034.

Wang, Y.P., Law, R.M. and Pak, B., 2010. A global model of carbon, nitrogen and phosphorus cycles for the terrestrial biosphere. Biogeosciences, 7(7): 2261-2282.

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. and Verma, S., 2002. Energy balance closure at FLUXNET sites. Agricultural and Forest Meteorology, 113(1–4): 223-243.

Zhang, Y.Q., Viney, N., Frost, A., Oke, A., Brooks, M., Chen, Y. and 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: ET monitoring site details and time-series

Table 11. 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:*	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

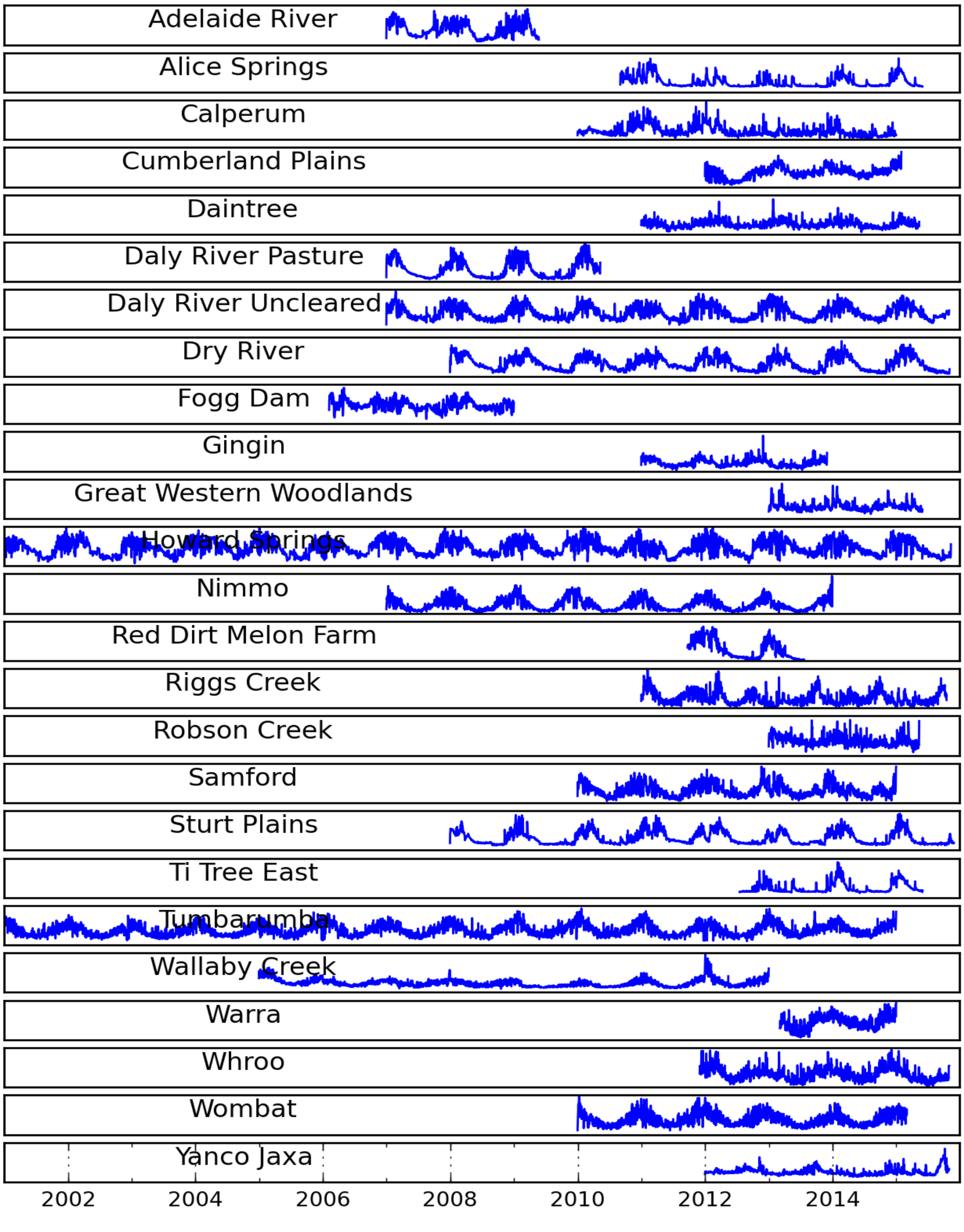


Figure 29. Indicative site time-series of DINGO Evapotranspiration (mm). Axis scale omitted for space purposes.

Appendix B: Soil moisture monitoring site details and time-series

Table 12. OzNet site details

OzNet Site	Start Date	End Date	Daily avail.	Monthly avail.
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%

Table 13. 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%

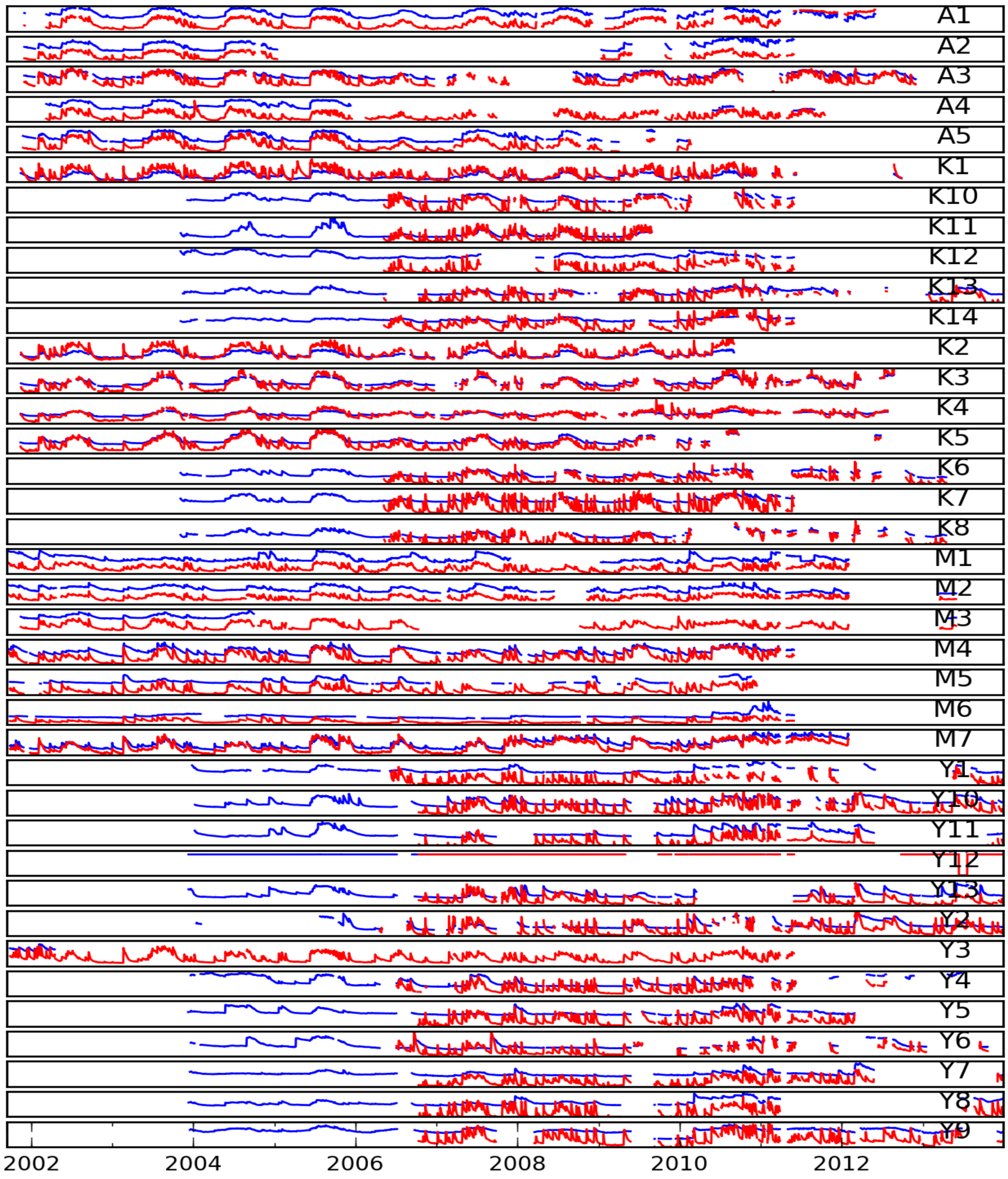


Figure 30. Indicative site daily time-series of OzNet top layer (red: 0-5/8cm) and profile (blue: 0-90cm) volumetric soil moisture. Site numbers are listed far right. Axis scale omitted for space purposes.

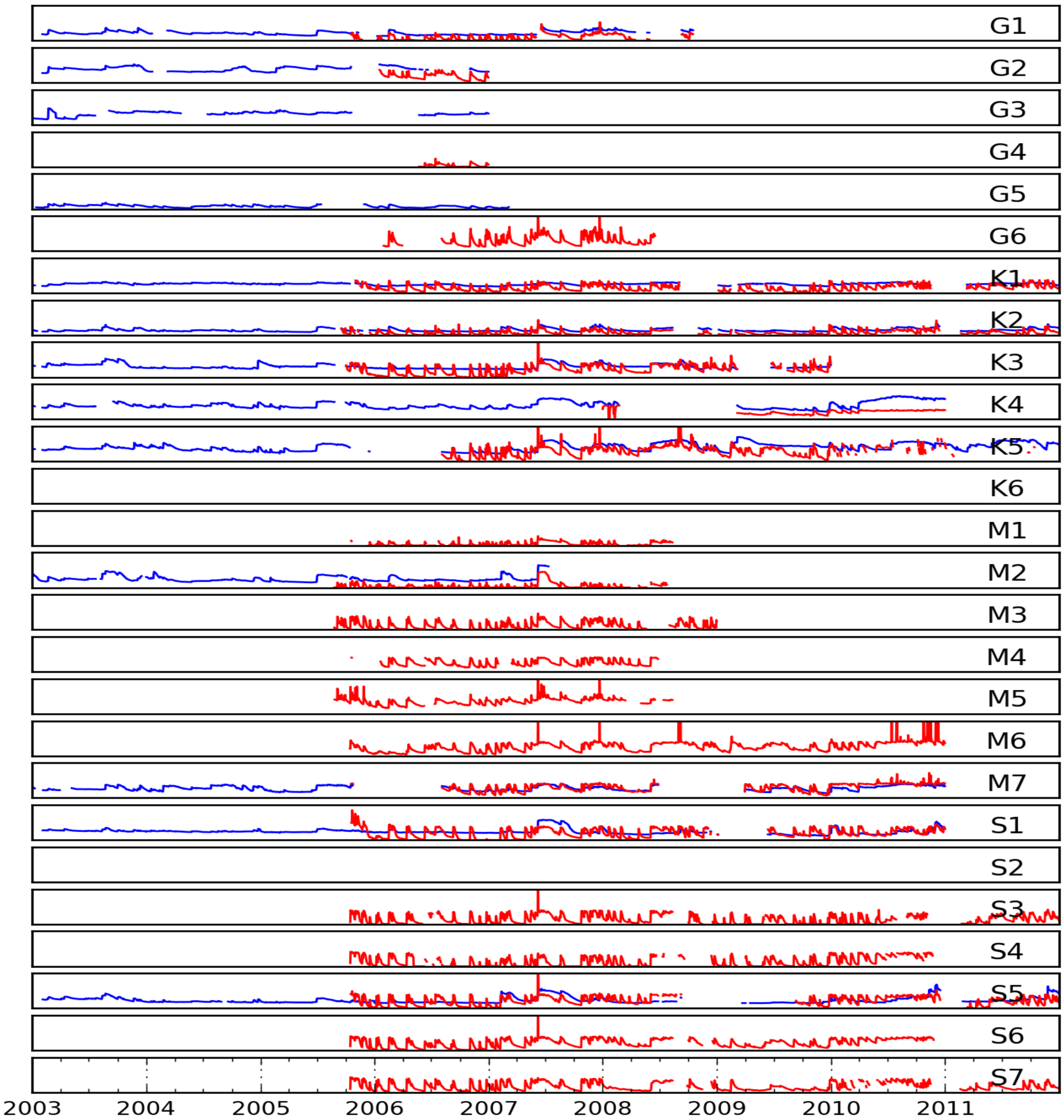


Figure 31. Indicative site daily time-series of SASMAS top layer (red: 0-5/8cm) and profile (blue: 0-90cm) volumetric soil moisture. Site numbers are listed far right. Axis scale omitted for space purposes.

Appendix C: Evaluation against top layer soil moisture

A comparison was undertaken using the time period that the satellite data was available (Jan 2007-Sept 2011) for the OzNet Murrumbidgee and SASMAS Hunter data (Figure 32. Correlation of monthly top layer (0-5/8cm) soil moisture of models against (a) Murrumbidgee OzNet and (b) Upper Hunter SASMAS data for Jan 2007-Sept 2011). This gives an indication of how well the satellite data represents surface and profile soil moisture, compared to AWRA-L. Only a monthly comparison is presented (noting a high proportion of missing daily data in satellite derived soil moisture – resulting in missing data for some sites for the ASCAT dataset in particular). CABLE and WaterDyn perform better than AWRA-L. ASCAT appears to perform slightly better than AMSRE in general – but this is purely due to the ASCAT missing data issue (with spurious perfect values and missing values for some sites). 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).

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. (2013) 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.

Finally, a comparison of the model versus AMSR-E satellite data is undertaken (Figure 33. Monthly correlation of AMSR-E against modelled soil moisture over (a) calibration and (b) validation catchments). Following the results presented for the top layer evaluation, CABLE and WaterDyn most closely match the AMSR-E data, with AWRA-L having lower correlation.

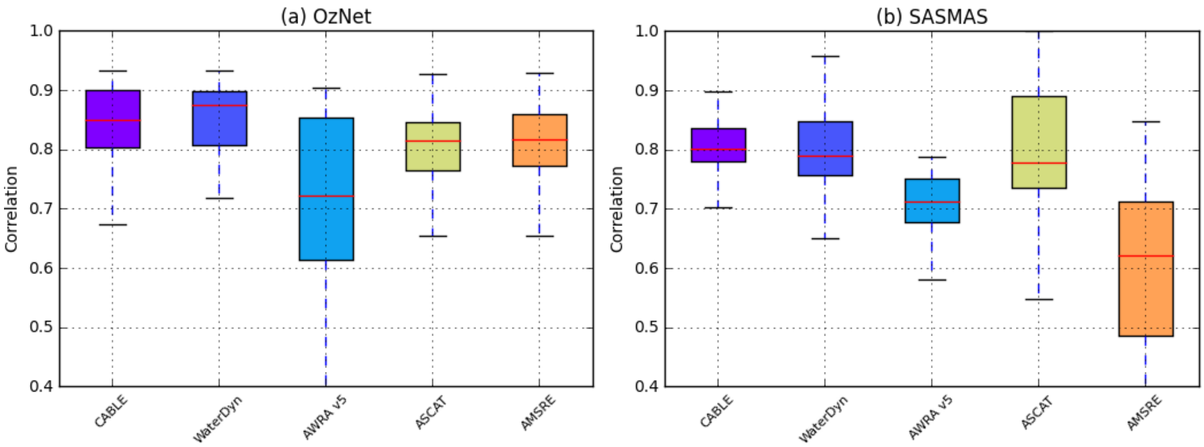


Figure 32. Correlation of monthly top layer (0-5/8cm) soil moisture of models against (a) Murrumbidgee OzNet and (b) Upper Hunter SASMAS data for Jan 2007-Sept 2011

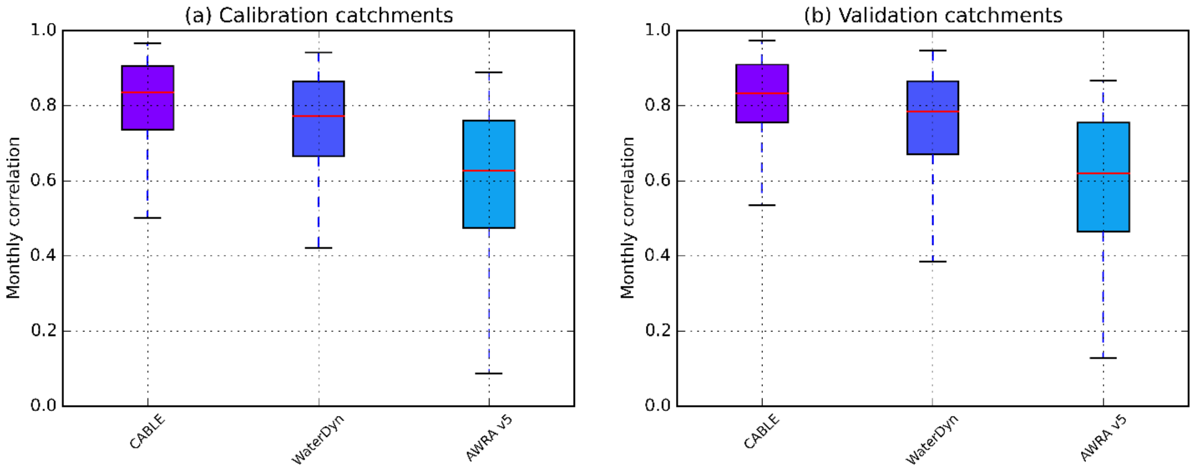


Figure 33. Monthly correlation of AMSR-E against modelled soil moisture over (a) calibration and (b) validation catchments

Appendix D: Maps and time series of water balance outputs from AWRA-L

Annual totals and deciles for the years covering 2007-2015 are plotted in Figure 34 and Figure 35.These plots show year to year variability, in particular the ending of the Millennium drought. Monthly average values are shown in Figure 36.

Ten catchments were selected for evaluation of states/flux time-series as shown in Figure 37. Key features of these sites are presented in Table 14.Time series of the following variables are plotted for each of these catchments: total evapotranspiration, runoff, deep drainage to the groundwater store, top 0-10cm soil moisture, shallow 10-100cm soil moisture, deep 100-600cm soil moisture and groundwater. Observed streamflow, along with satellite based ET and soil moisture is also plotted for comparison purposes. These plots give an indication of the seasonal and interannual variability present at each of these locations for the key water variables output by AWRA-L.

Table 14. Selected catchments for detailed evaluation

ID	Name	River	State	Lat. (°)	Lon. (°)	Area (km^2)	Elev. (m)	Slope (%)	Ave. Precip. (mm)	Ave.P ET (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

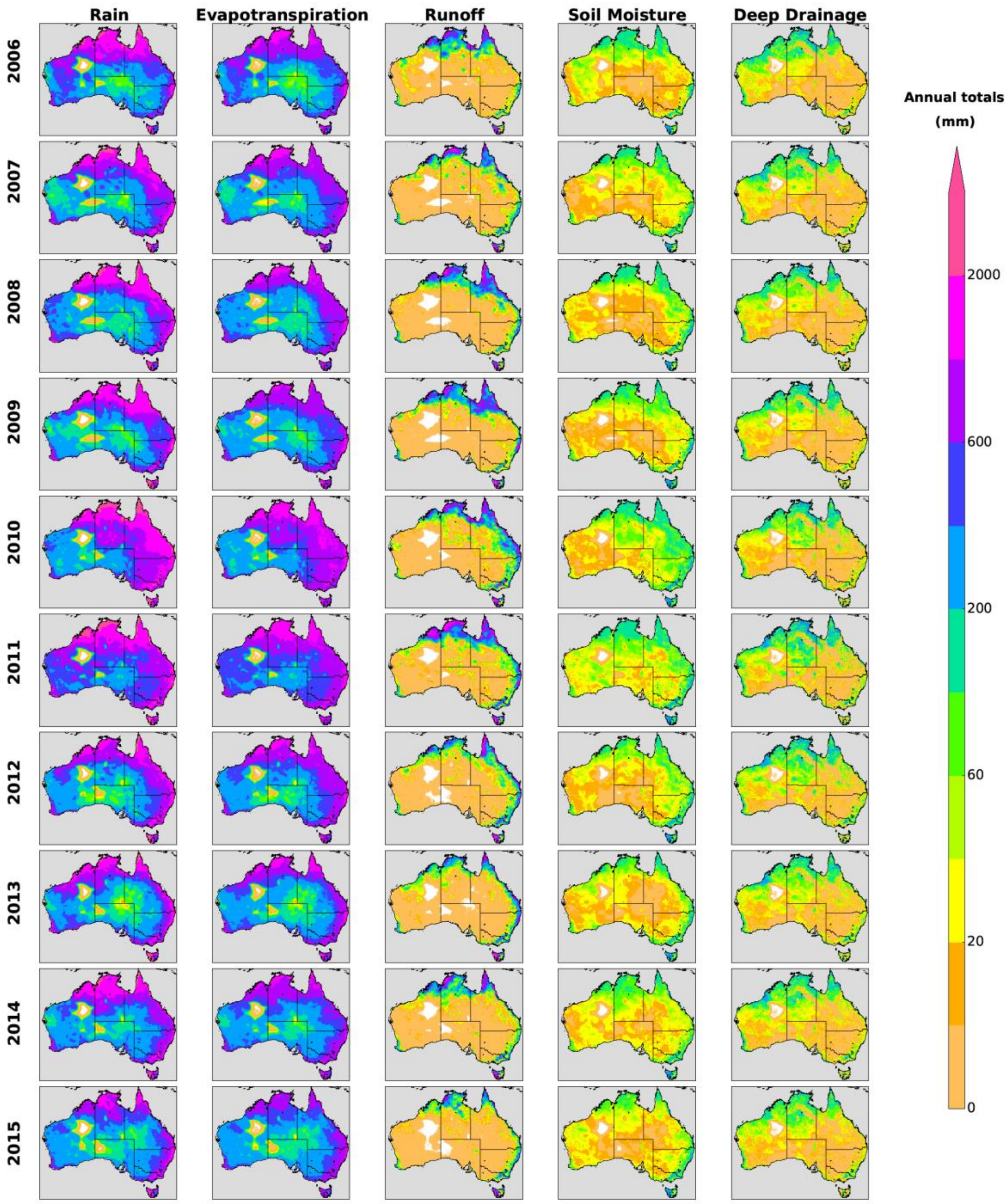


Figure 34. Annual rain, runoff, ET, soil moisture and deep drainage for 2007-2015

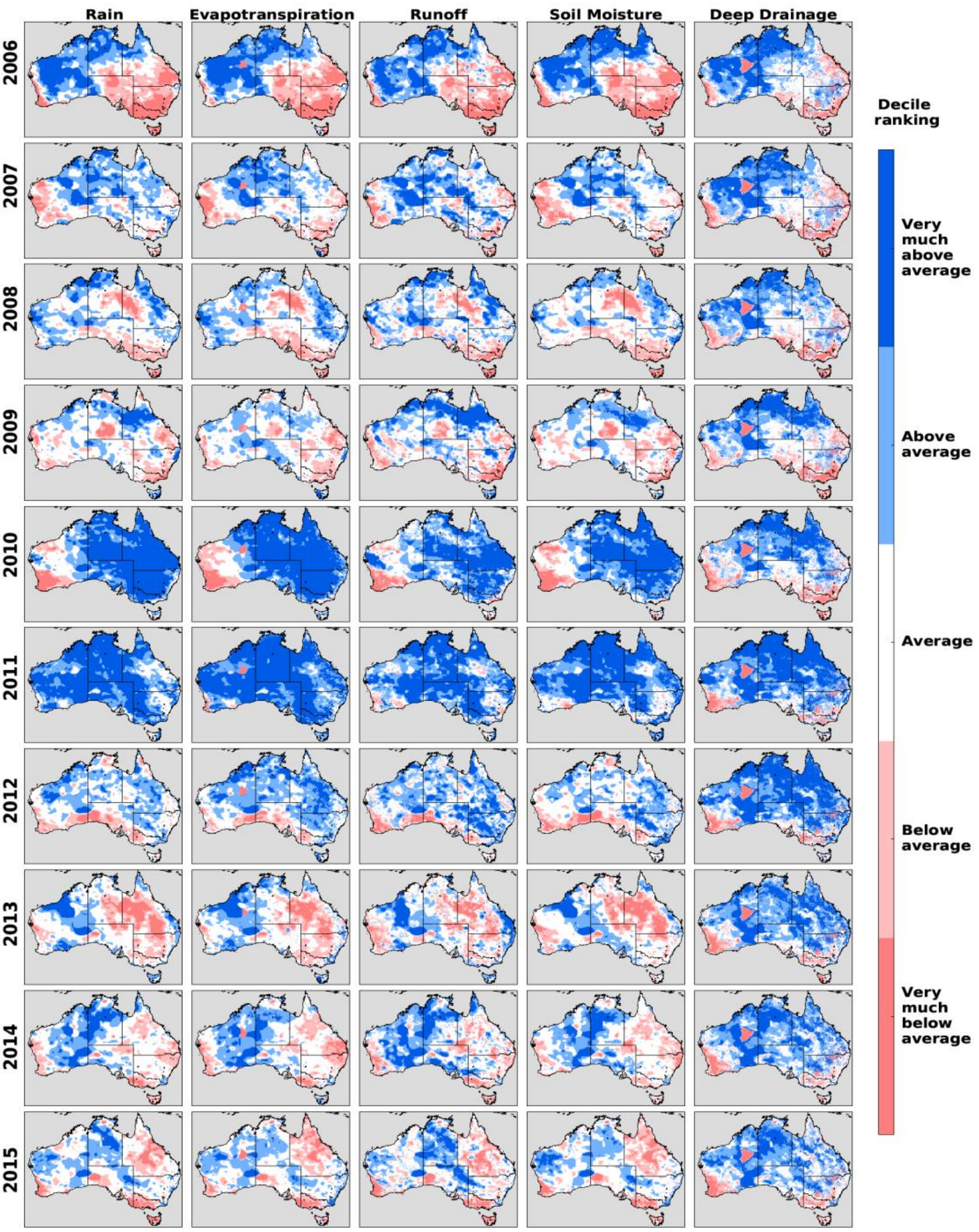


Figure 35. Annual rain, runoff, ET, soil moisture and deep drainage deciles for 2007-2015

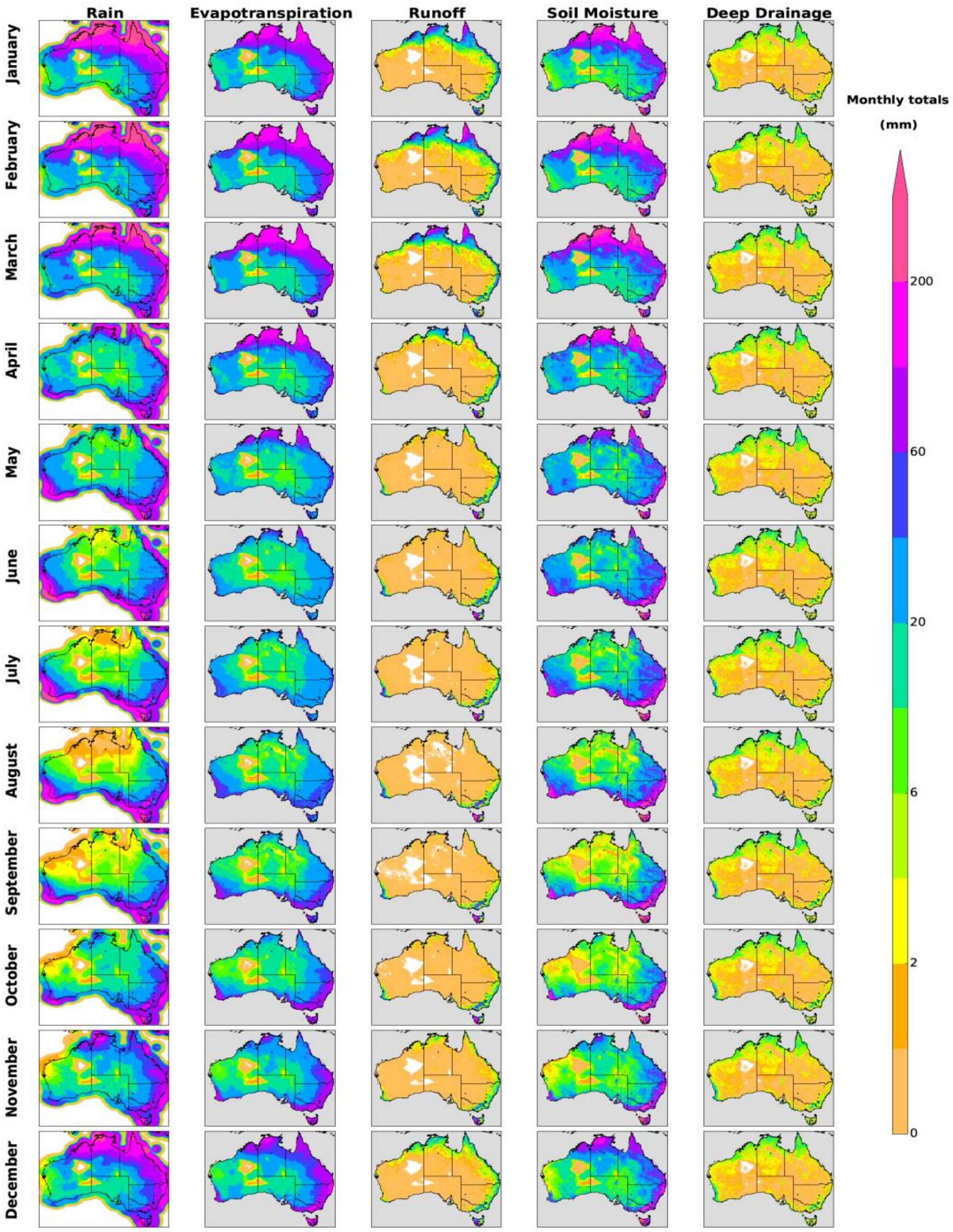


Figure 36. 1911-2015 Monthly mean rain, ET, runoff, soil moisture (1m) and deep drainage

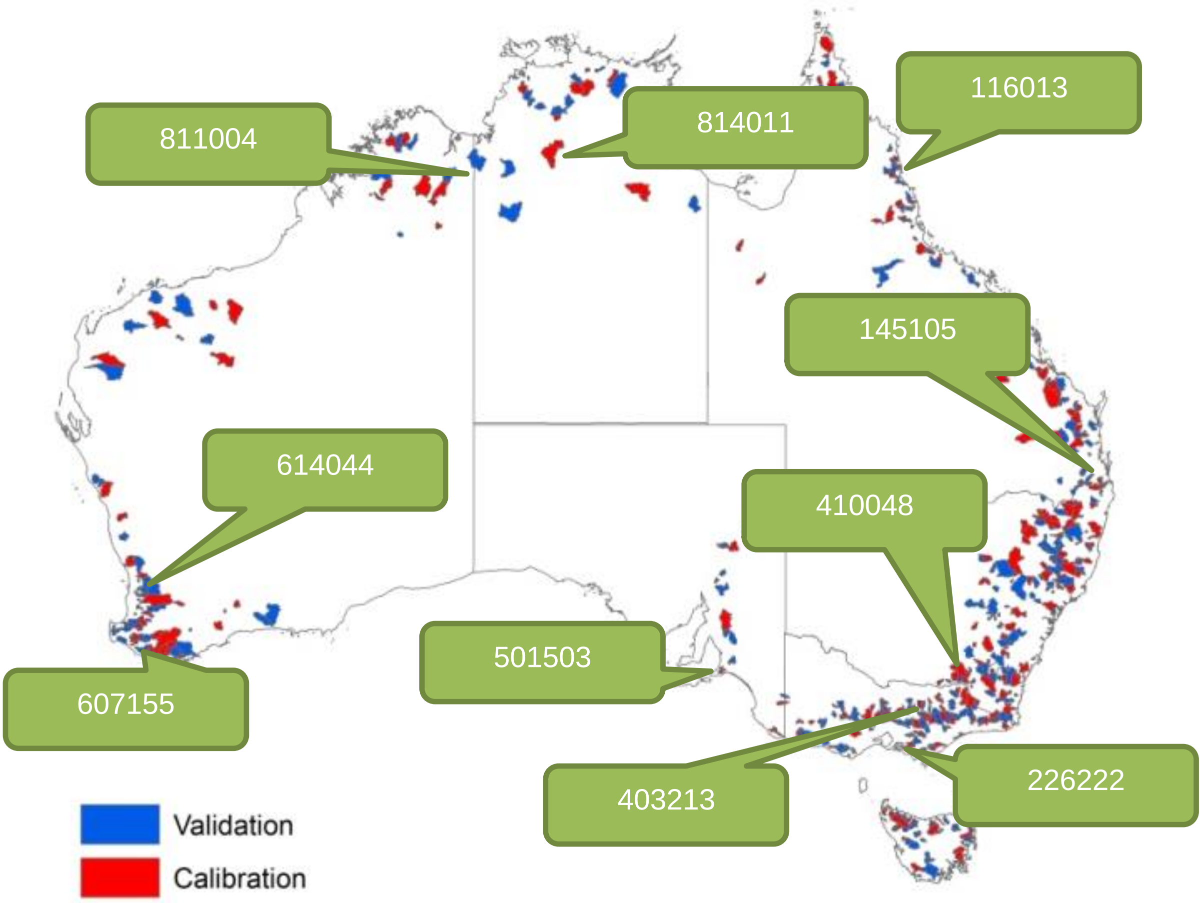


Figure 37. Locations of selected catchments for detailed evaluation

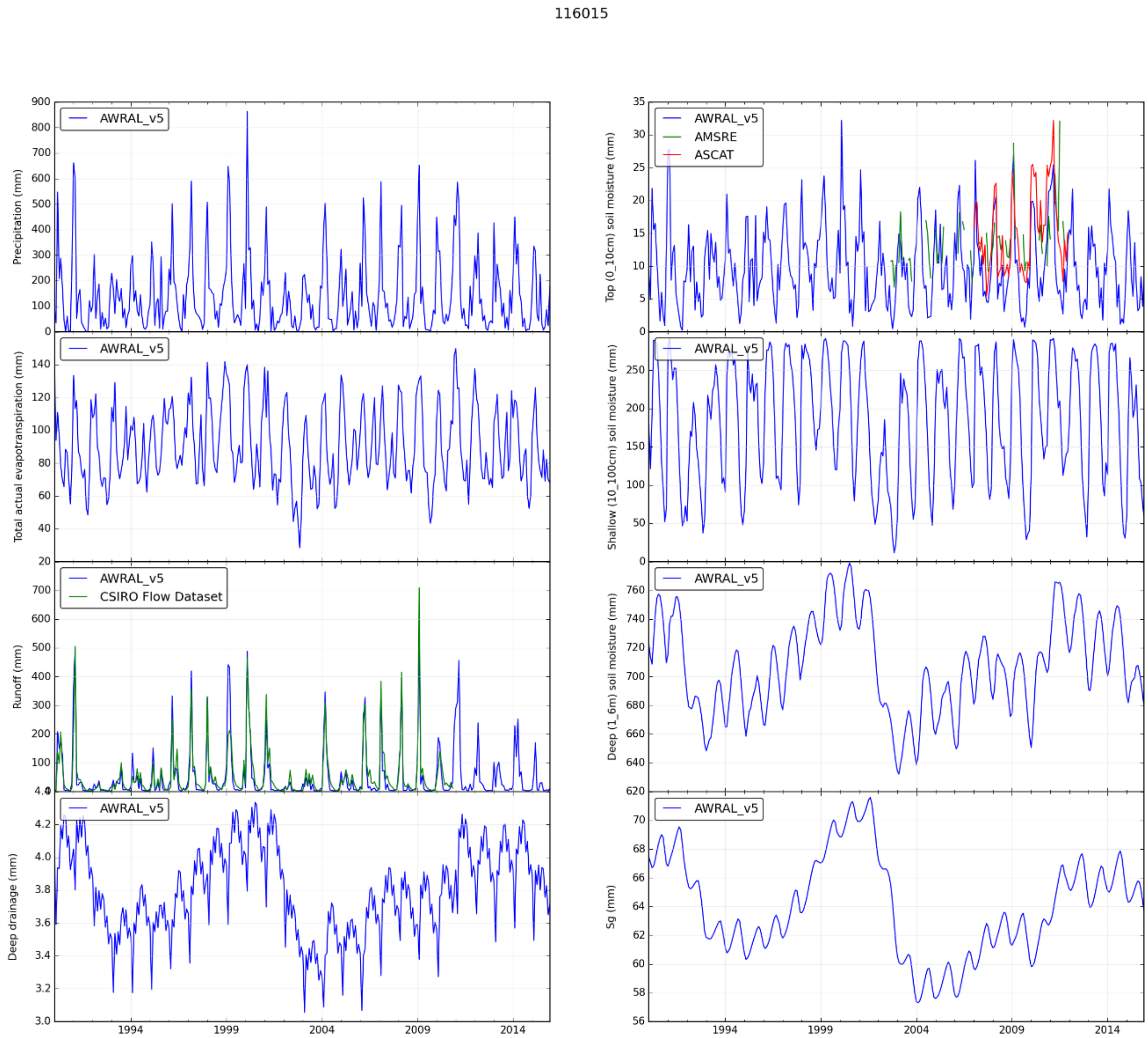


Figure 38. 16015 Archer Creek @ Millstream QLD AWRA-L v5 1990-2015 simulation monthly values. Catchment streamflow, CMRSET/SLST evapotranspiration and scaled AMSR-E/ASCAT soil moisture values also shown where relevant.

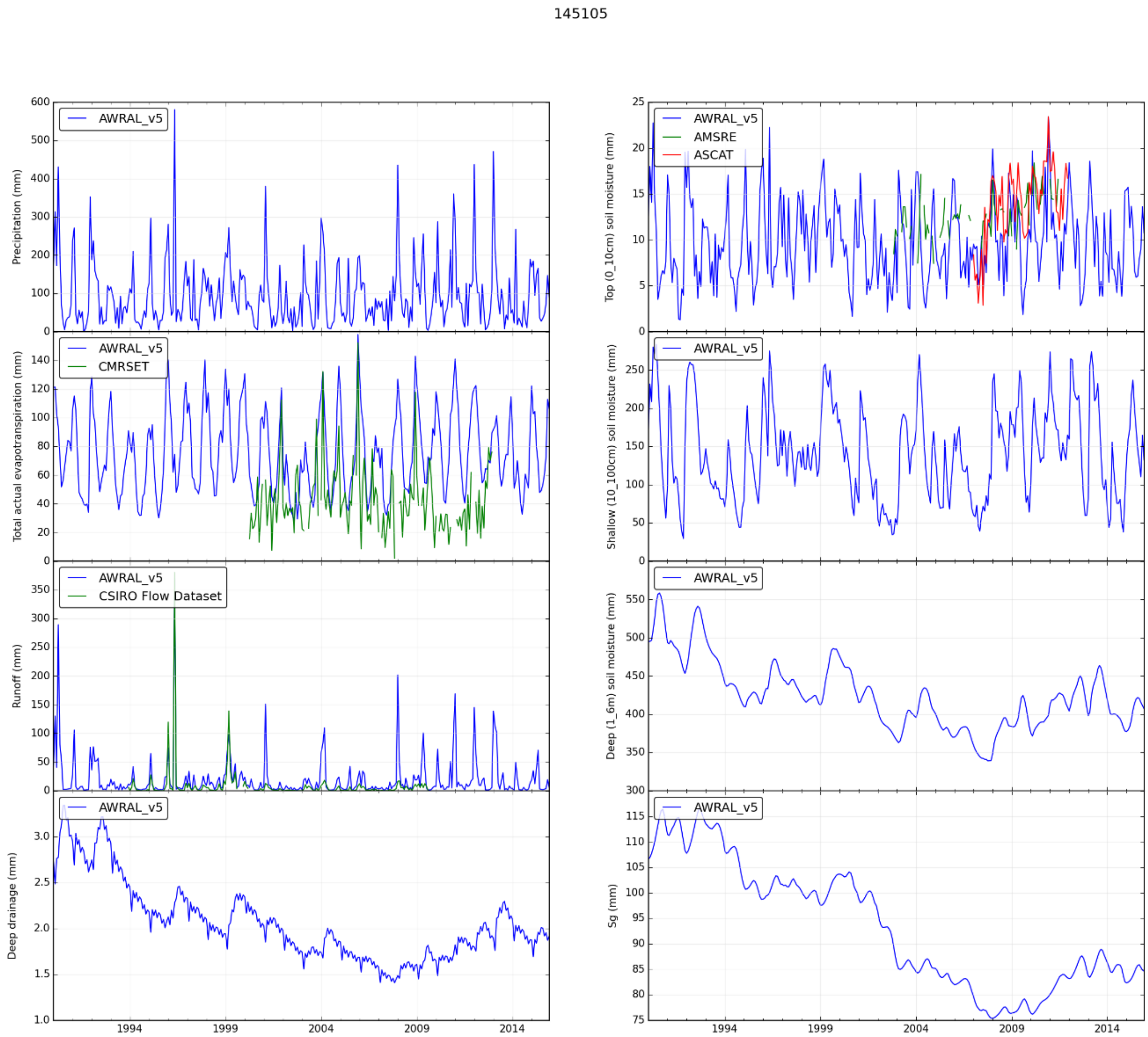


Figure 39. 145105 Beaudesert pump station Albert River QLD AWRA-L v5 1990-2015 simulation monthly values. Catchment streamflow, CMRSET/SLST evapotranspiration and scaled AMSR-E/ASCAT soil moisture values also shown where relevant.

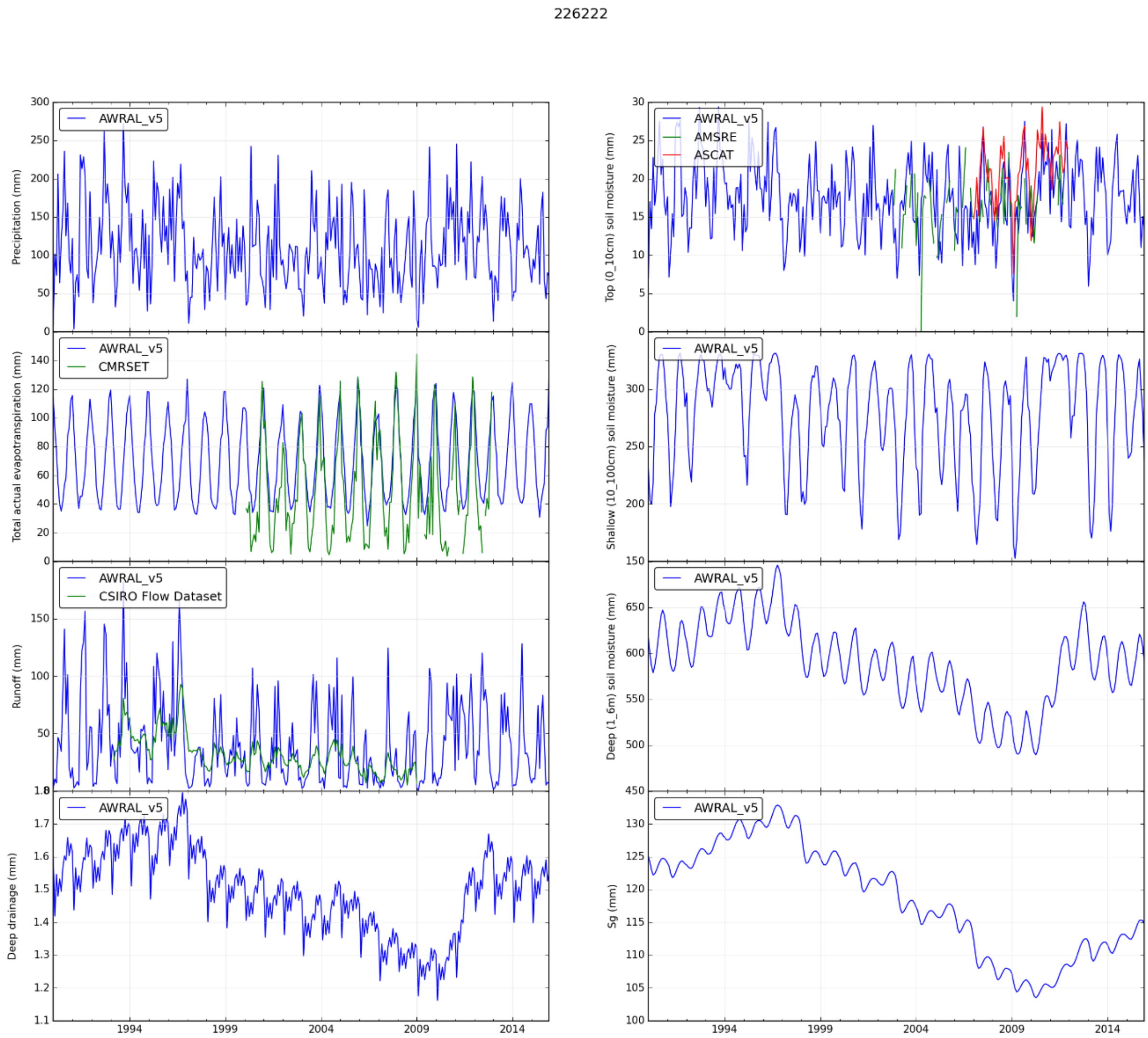


Figure 40. 226222 near Noojee VIC AWRA-L v5 1990-2015 simulation monthly values. Catchment streamflow, CMRSET/SLST evapotranspiration and scaled AMSR-E/ASCAT soil moisture values also shown where relevant

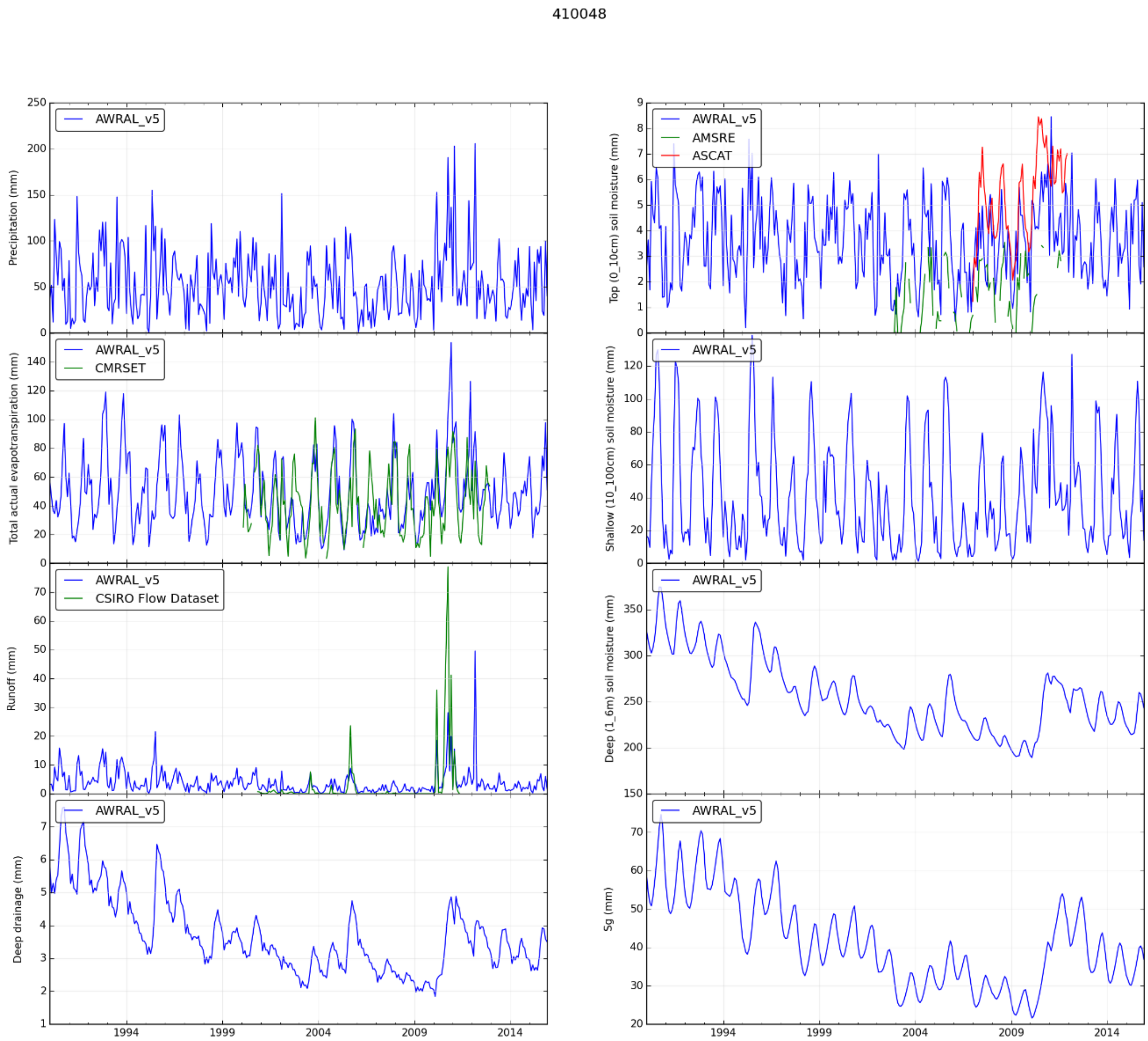


Figure 41. 410048 Ladysmith Kyeamba Creek NSW AWRA-L v5 1990-2015 simulation monthly values. Catchment streamflow, CMRSET/SLST evapotranspiration and scaled AMSR-E/ASCAT soil moisture values also shown where relevant.

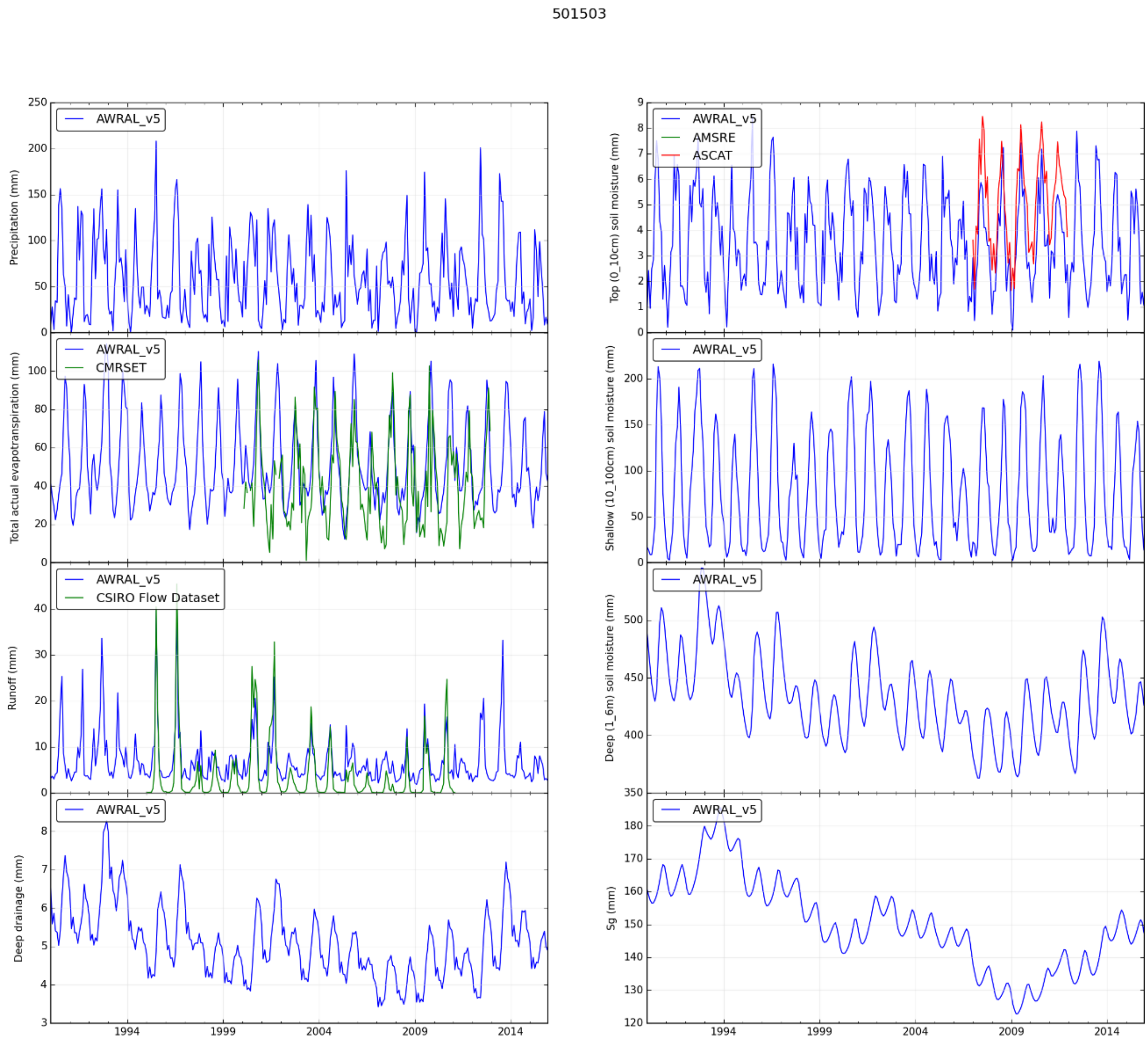


Figure 42. 501503 US Victor harbour, Inman River SA AWRA-L v5 1990-2015 simulation monthly values. Catchment streamflow, CMRSET/SLST evapotranspiration and scaled AMSR-E/ASCAT soil moisture values also shown where relevant.

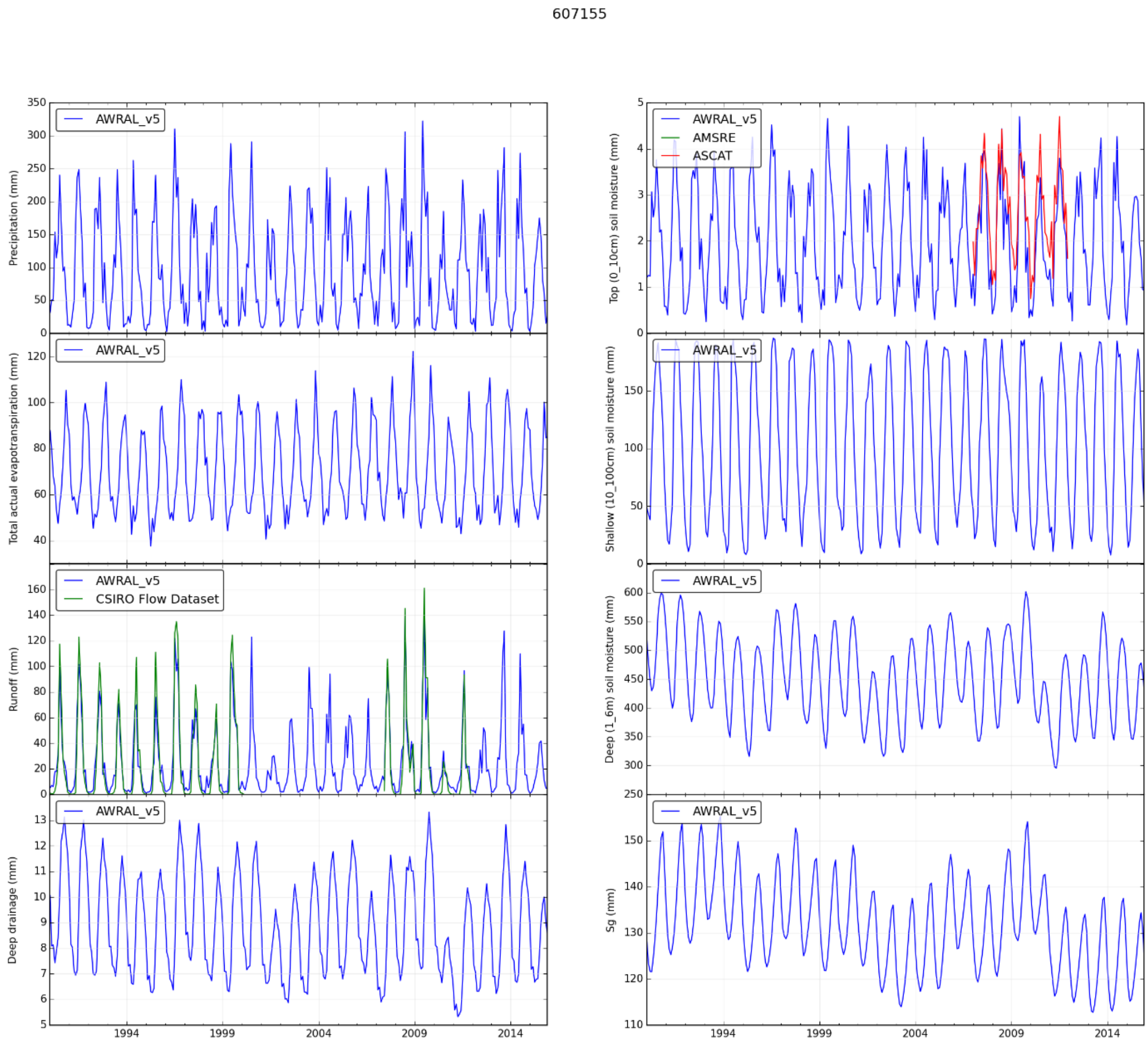


Figure 43. 607155 AWRA-L v5 1990-2015 simulation monthly values. Catchment streamflow, CMRSET/SLST evapotranspiration and scaled AMSR-E/ASCAT soil moisture values also shown where relevant

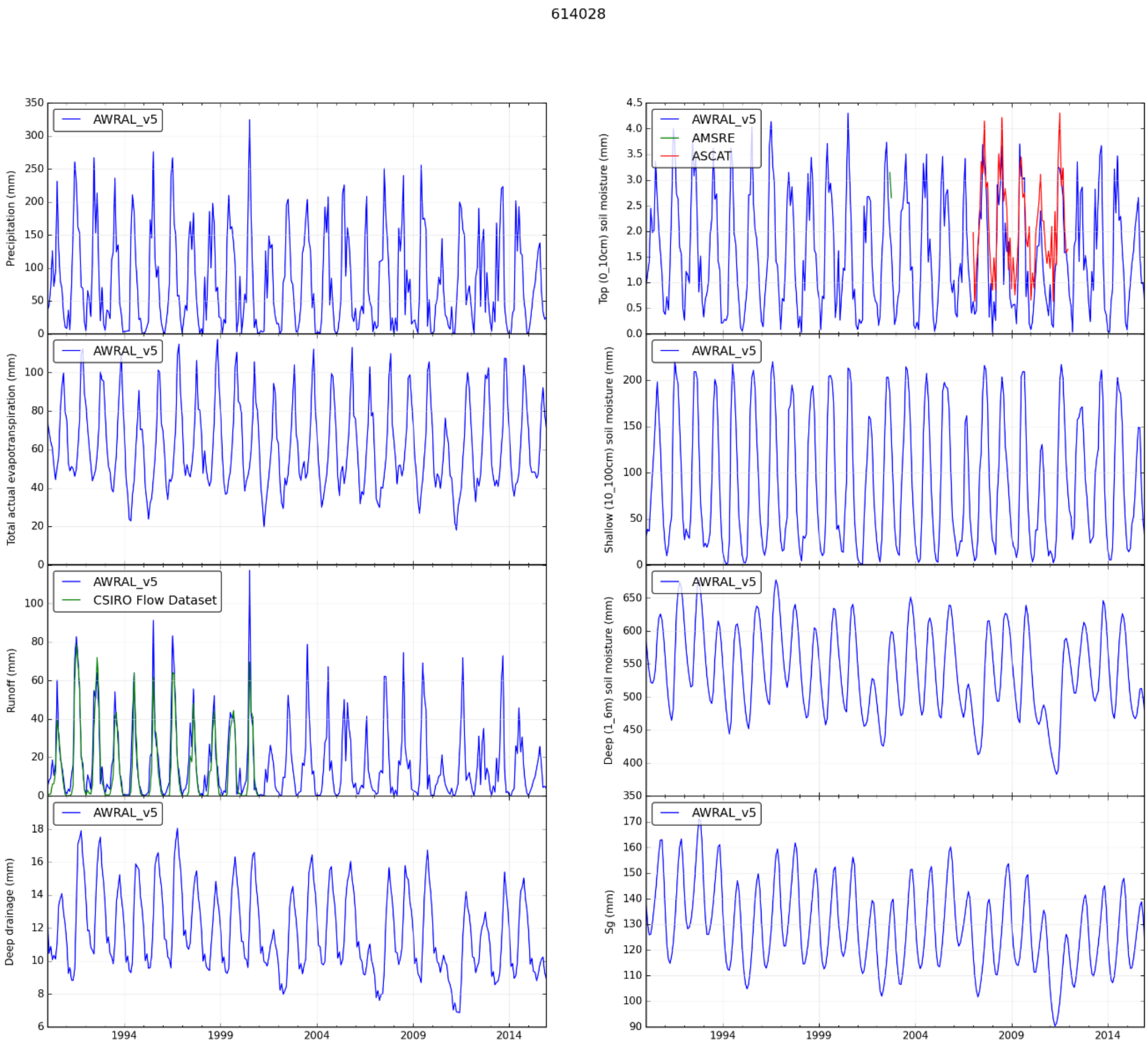


Figure 44. 614028 Malimup Track at Dombakup Brook WA AWRA-L v5 1990-2015 simulation monthly values. Catchment streamflow, CMRSET/SLST evapotranspiration and scaled AMSR-E/ASCAT soil moisture values also shown where relevant.

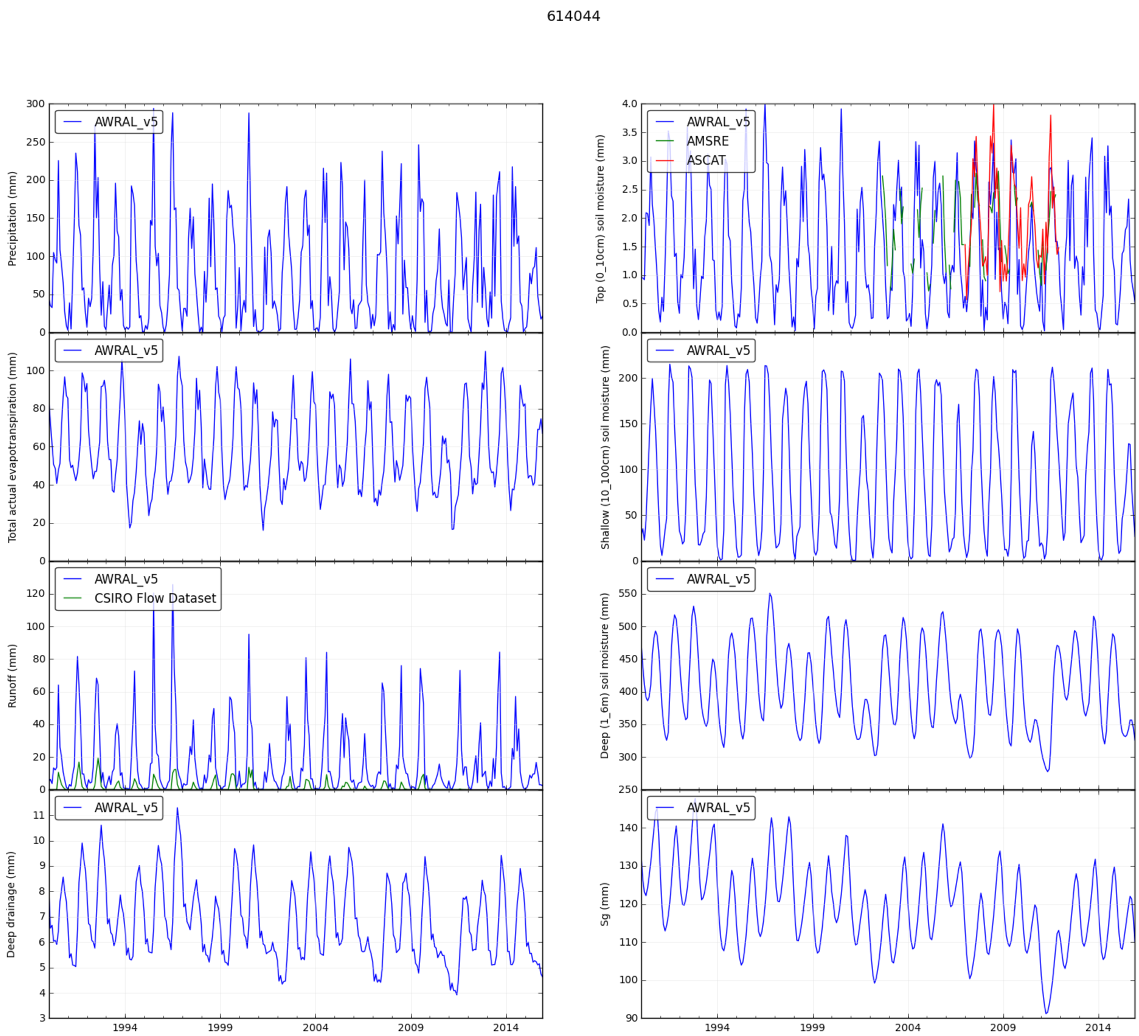


Figure 45. 614044 Yarragil Brook AWRA-L v5 1990-2015 simulation monthly values. Catchment streamflow, CMRSET/SLST evapotranspiration and scaled AMSR-E/ASCAT soil moisture values also shown where relevant.

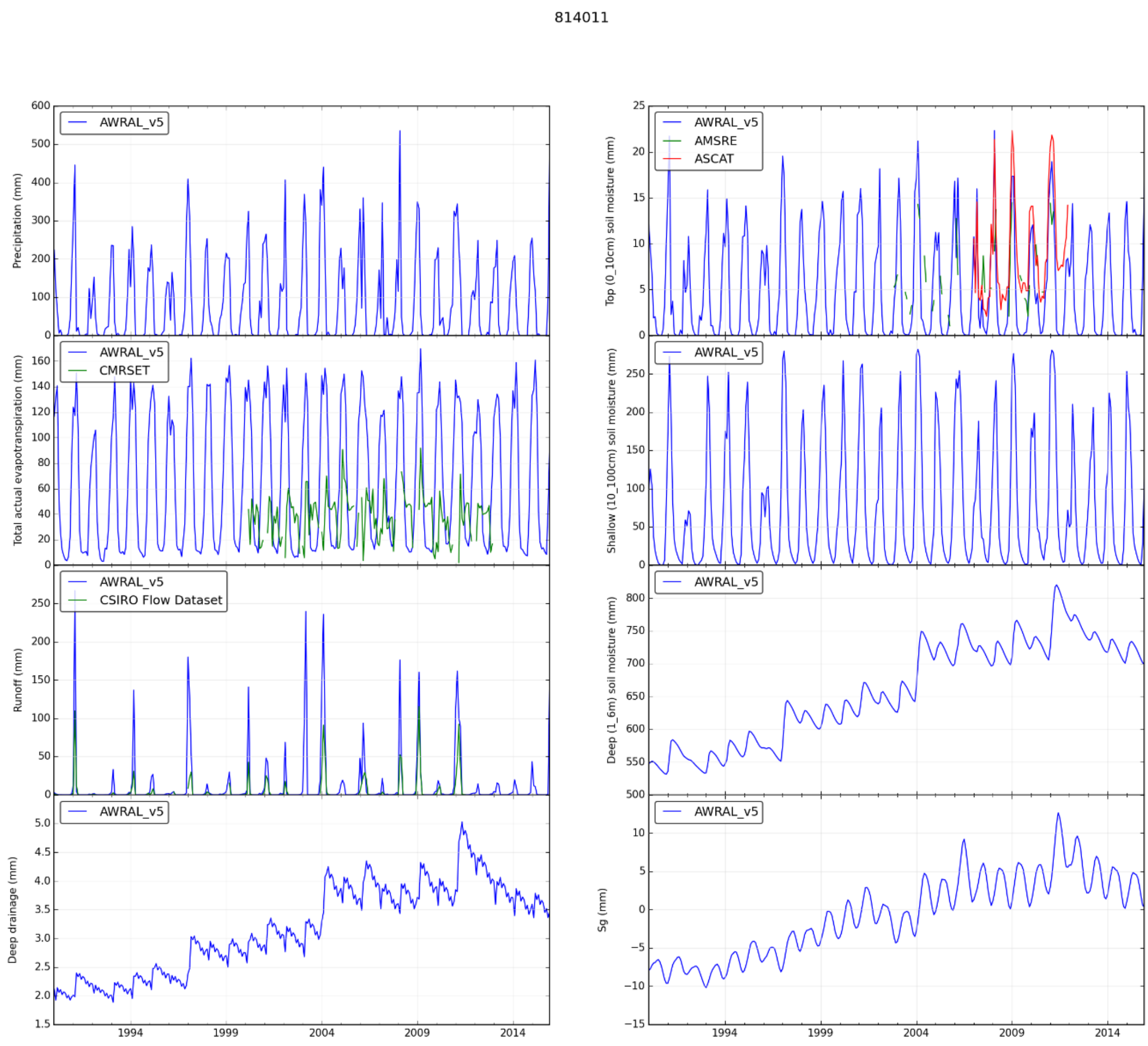


Figure 46. 814011 Dry River NT AWRA-L v5 1990-2015 simulation monthly values. Catchment streamflow, CMRSET/SLST evapotranspiration and scaled AMSR-E/ASCAT soil moisture values also shown where relevant.

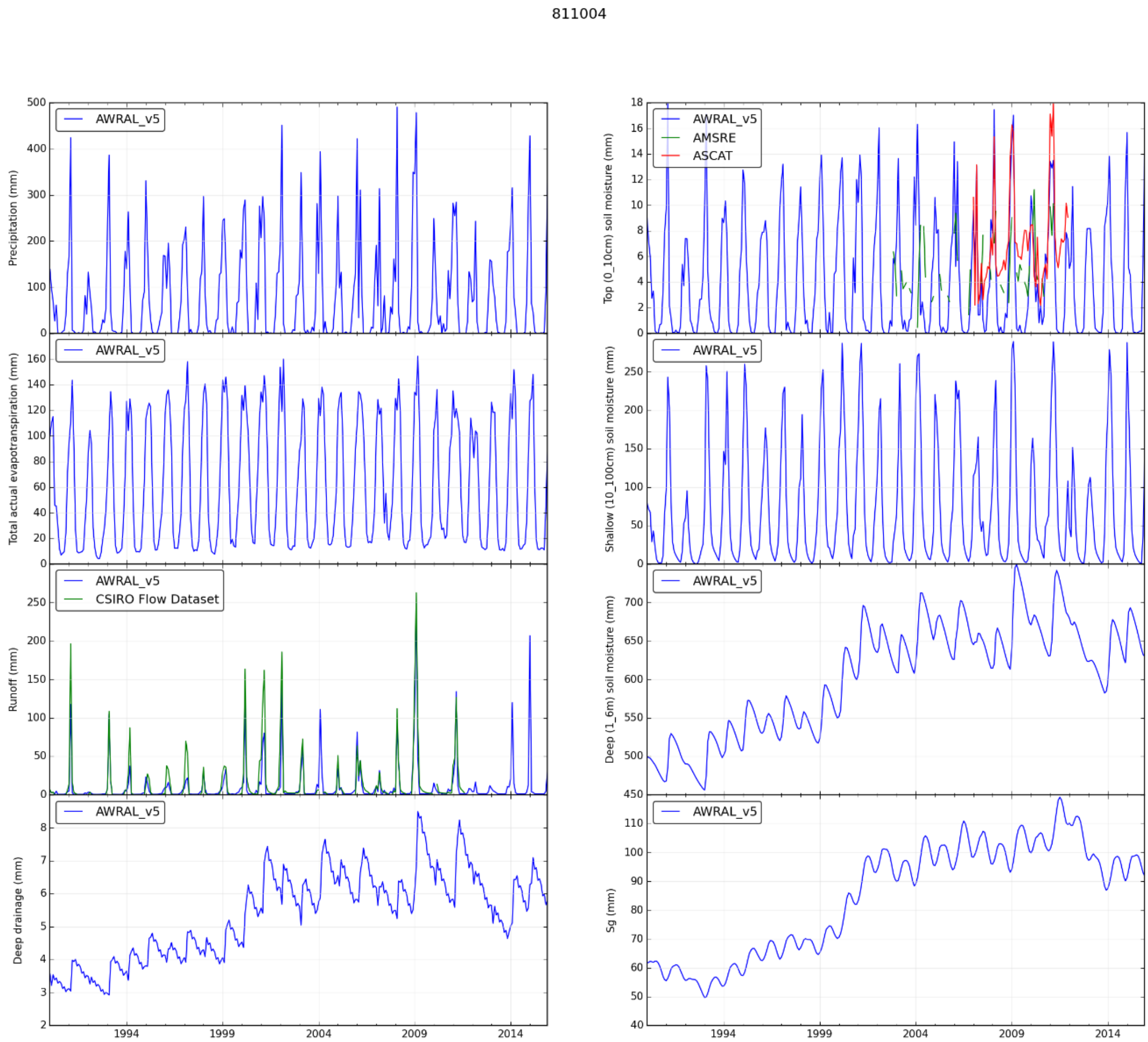


Figure 47. 811004 Victoria HWY East Baines NT AWRA-L v5 1990-2015 simulation monthly values. Catchment streamflow, CMRSET/SLST evapotranspiration and scaled AMSR-E/ASCAT soil moisture values also shown where relevant.